



Statistical physics of principal minors: Cavity approachA. Ramezanpour **Department of Physics, College of Sciences, Shiraz University, Shiraz 71454, Iran
and Medical Systems Biophysics and Bioengineering, Leiden Academic Centre for Drug Research,
Faculty of Science, Leiden University, 2333 CC Leiden, The Netherlands*M. A. Rajabpour †*Instituto de Física, Universidade Federal Fluminense, Campus da Praia Vermelha Sao Domingos, 24210-346 Niteroi-RJ, Brasil*

(Received 18 March 2024; accepted 29 May 2024; published 18 June 2024)

Determinants are useful to represent the state of an interacting system of (effectively) repulsive and independent elements, like fermions in a quantum system and training samples in a learning problem. A computationally challenging problem is to compute the sum of powers of principal minors of a matrix which is relevant to the study of critical behaviors in quantum fermionic systems and finding a subset of maximally informative training data for a learning algorithm. Specifically, principal minors of positive square matrices can be considered as statistical weights of a random point process on the set of the matrix indices. The probability of each subset of the indices is in general proportional to a positive power of the determinant of the associated submatrix. We use Gaussian representation of the determinants for symmetric and positive matrices to estimate the partition function (or free energy) and the entropy of principal minors within the Bethe approximation. The results are expected to be asymptotically exact for diagonally dominant matrices with locally treelike structures. We consider the Laplacian matrix of random regular graphs of degree $K = 2, 3, 4$ and exactly characterize the structure of the relevant minors in a mean-field model of such matrices. No (finite-temperature) phase transition is observed in this class of diagonally dominant matrices by increasing the positive power of the principal minors, which here plays the role of an inverse temperature.

DOI: [10.1103/PhysRevE.109.064141](https://doi.org/10.1103/PhysRevE.109.064141)**I. INTRODUCTION**

A principal minor of a matrix \mathbf{A} is determinant of a square submatrix which is formed by a number of rows and columns with the same indices from the matrix. This provides a measure of independence for the selected subset of rows (or columns) which is useful for instance in sampling problems where diversity matters or in computing the entropy of physical systems with fermionic statistics. The sum of powers of principal minors (SPPM) of a matrix appears in various areas of physics and mathematics [1]. SPPM finds applications in the study of determinantal point processes [2,3], Rényi entropy of free fermions and quantum spin chains [4], and partition function of the Hubbard model [1]. When the power is one, the sum is a simple determinant that can be computed in polynomial time. However, for other powers, the problem is generally considered to be NP-hard, indicating that there may not be a polynomial time algorithm to calculate them [5,6]. Nonetheless, recent advancements have been made in the approximate analytical computation of these quantities [1]. In this paper we study statistical properties of the principal minors of positive, symmetric, and diagonally dominant matrices which can be represented by Gaussian integrals and

estimated by the cavity method of statistical physics. We employ this method to characterize the spectral entropy of principal minors in this subclass of matrices and utilize it as an approximate optimization algorithm to find the maximal minor configurations.

The principal minors of a matrix has a very natural meaning in the context of the graph theory. Consider, the adjacency matrix of a graph G with n nodes, which is a matrix that describes the connections between the nodes of the graph, where the entry in row i and column j of the matrix is 1 if there is an edge from node i to node j and 0 otherwise. Then the Laplacian matrix of a graph is defined as the difference between the degree matrix and the adjacency matrix of the graph. The degree matrix is a diagonal matrix where the entry in the i th row and i th column is equal to the degree of the i th node, i.e., the number of edges incident to it, and the nondiagonal entries are 0. The Laplacian matrix \mathbf{L} is a symmetric, positive semidefinite matrix, which means all the principal minors are non-negative. It has numerous applications in graph theory notably in the enumeration of the spanning trees. A spanning tree of an undirected graph is a tree that includes all of the graph's vertices, while using only a subset of its edges, such that no cycles (closed paths) are formed. The matrix-tree theorem, see for example Ref. [7], then states that the number of spanning trees of the graph G is equal to any cofactor of the Laplacian matrix \mathbf{L} . Specifically, if we remove any row and column (say, i) from \mathbf{L} to obtain a $(n - 1) \times (n - 1)$ matrix

*Contact author: aramezanpour@gmail.com†Contact author: mohammadali.rajabpour@gmail.com

$\mathbf{L}[i]$, then the number of spanning trees of G is given by the determinant of $\mathbf{L}[i]$.

There is a generalization of matrix-tree theorem which is called principal minor matrix tree theorem which is used to enumerate the number of rooted spanning forests [8]. A forest is formed by a set of disjoint tree subgraphs of the graph G . A rooted spanning forest is a subgraph of G that is a forest and contains all vertices of the graph. And each tree in the forest is rooted at a designated root vertex in that tree. Let $\mathbf{L}_{\mathbf{c}}$ be the submatrix of \mathbf{L} formed by selecting the rows and columns corresponding to the vertices in \mathbf{c} then the number of spanning forests rooted at \mathbf{c} is $\det \mathbf{L}_{\mathbf{c}}$. This interpretation of the principal minors makes the definition of the weighted partition function of the rooted spanning forests as the sum of powers of principal minors natural. In this paper we interpret the power as the inverse of a temperature-like parameter. In the limit of large powers, i.e., small temperatures, the SPPM gives the principal minor with the largest contribution which is a desired quantity in determinantal point processes [2] and the study of random spanning trees [9].

Here we introduce a statistical model in which the principal minors of a positive semidefinite matrix serve as the statistical weights. Although the matrix need not necessarily be a graph Laplacian, our focus is primarily on this type of matrix. To address this problem, we utilize the cavity method (Bethe approximation or belief propagation) [10–13], which is commonly used in spin glasses but has not been previously applied to studying the sum of powers of principal minors. Examples of applications of the method in similar problems, mainly in the study of spectral properties of sparse matrices, can be found in Refs. [14–20]. The cavity method should be asymptotically exact in interacting systems which have a locally treelike interaction graph [13]. The sum of principal minors to power β here is considered as the partition function of principal minors of the matrix at inverse temperature β . We employ the Gaussian representation of determinants to find estimations of the free energy and entropy of the relevant principal minors for Laplacian of random regular graphs of degree $K = 2, 3, 4$. These results along with an exact study of diagonally dominant matrices defined on fully connected graphs indicate on the absence of a finite-temperature phase transition in this class of matrices. In a diagonally dominant matrix, the magnitude of a diagonal element $|A_{ii}|$ in each row is larger than or equal to the sum of magnitudes of the off-diagonal elements $\sum_{j \neq i} |A_{ij}|$.

The nontrivial structure of the ground states in random regular graphs of degree $K = 2$ (or chains), however, results in numerical instabilities for large β and degrades the efficiency of the simulated annealing algorithm in finding a maximal principal minor in these systems; the entropy density of relevant minor configurations is discontinuous at zero temperature, jumping to a nonzero value for any finite temperature. The situation is much better for (random) Laplacian of graphs with larger degrees $K > 2$ as one approaches the mean-field limit, where the ground states have a trivial structure in the configuration space of the principal minors. Nevertheless, it is possible to observe discontinuities in the entropy density at zero temperature by introducing a chemical potential which controls the number of present indices in the minors. These zero-temperature transitions are similar in nature to those of

the smaller degree $K = 2$. We also use the zero-temperature limit of the Bethe equations, the so-called MaxSum (MS) equations [13,21], as an approximate optimization algorithm to find a maximal principal minor of diagonally dominant matrices. The standard algorithms are in general based on the spectral decomposition of the matrix with a time complexity of order N^3 for an arbitrary matrix of dimension N [2]. The computational complexity can of course be reduced to N for sampling of principal minors of dimensions $n \ll N$ [9,22,23]. The time complexity of the MaxSum algorithm here is proportional to N for sparse and diagonally dominant matrices where the (Gaussian) Bethe equations are expected to work.

The paper is structured as follows. In Sec. II we define the main quantities of the problem. In Sec. III we present the results obtained by the Gaussian representation of the determinants. This section includes subsections that deal with the finite- and zero-temperature limits of the Bethe equations. The summary of results and concluding remarks are given in Sec. IV. In the Appendices, we give a brief introduction to the Bethe approximation (A), describe the details of the population dynamics which is used to solve the Bethe equations (B), write the simplified Bethe equations for random regular graphs (C), and present an exact treatment of principal minors in matrices defined by homogeneous fully connected graphs (D).

II. THE PROBLEM STATEMENT AND DEFINITIONS

Consider a positive square matrix \mathbf{A} of size N with elements A_{ij} indexed by $i, j = 1, \dots, N$. The nonzero elements of \mathbf{A} define the structure or interaction graph of the matrix. The set of neighbors of node i in this graph is denoted by ∂i . The 2^N principal minors $\det \mathbf{A}(\mathbf{c})$ are labeled with configurations $\mathbf{c} = \{c_i = 0, 1 : i = 1, \dots, N\}$. That is $c_i = 1$ shows that row (column) i is included in the square submatrix $\mathbf{A}(\mathbf{c})$. Each configuration is assigned a Boltzmann weight,

$$P_{\beta}(\mathbf{c}) = \frac{[\det \mathbf{A}(\mathbf{c})]^{\beta}}{Z(\beta)} = \frac{e^{-\beta E(\mathbf{c})}}{Z(\beta)}, \quad (1)$$

where the associated energy function reads

$$E(\mathbf{c}) = -\ln[\det \mathbf{A}(\mathbf{c})], \quad (2)$$

and the positive power β plays the role of an inverse temperature. The energy of all-zero minor configuration is zero that is $\det \mathbf{A}(\mathbf{0}) = 1$. Here the partition function (normalization factor) is

$$Z(\beta) = \sum_{\mathbf{c}} [\det \mathbf{A}(\mathbf{c})]^{\beta} = \int de e^{-\beta N[e - \frac{s(e)}{\beta}]} = e^{-\beta N f(\beta)}, \quad (3)$$

with the energy density $e = E/N$ and the free energy density $f(\beta) = -[\ln Z(\beta)]/(\beta N)$. The entropy density is defined by $s(e) = \ln[\Omega(e)]/N$, where $\Omega(e)$ is the number of configurations with energy density e . In the thermodynamic limit $N \rightarrow \infty$, the integrand is concentrated on the mean energy, that is, $f = e - \frac{1}{\beta} s$.

In the following, we are going to change the parameter β to study the energy and entropy landscape of the SPPM problem given in Eq. (3). The range of energy values (principal minors of \mathbf{A}) represent the relevant minors for different values of β .

Specially, the minimum energy value and configuration, i.e., the ground state(s) are obtained at zero temperature ($\beta \rightarrow \infty$). The entropy spectrum $s(e)$ shows the number of such relevant submatrix configurations and the free energy f is a measure of the sum over the weights of these configurations at inverse temperature β .

From the above system we can compute the Rényi entropy of the probability distribution $P_\beta(\mathbf{c})$,

$$\begin{aligned} R_\beta(n) &= \frac{1}{1-n} \ln \left[\sum_{\mathbf{c}} P_\beta^n(\mathbf{c}) \right] \\ &= \frac{1}{1-n} [\ln Z(n\beta) - n \ln Z(\beta)]. \end{aligned} \quad (4)$$

Specifically,

$$R_\beta(1) = - \sum_{\mathbf{c}} P_\beta(\mathbf{c}) \ln P_\beta(\mathbf{c}), \quad (5)$$

which can be obtained by a Legendre transformation of the free energy

$$\frac{R_\beta(1)}{N} = s(\beta) = \beta[e(\beta) - f(\beta)]. \quad (6)$$

Note that here $e(\beta)$ is the average energy density at inverse temperature β . A parametric plot of the above quantity gives the entropy $s(e)$ as a function of energy. Here one is interested in the values of the free energy, entropy, and energy of the relevant minors at different inverse temperatures β , the structure of these minors in the configuration space and the nature of possible phase transitions in this system.

III. GAUSSIAN REPRESENTATION

Let us assume that \mathbf{A} is a positive and symmetric matrix. Thus all the square submatrices are positive and symmetric. Given a configuration \mathbf{c} of the indices we define the submatrix $\mathbf{A}(\mathbf{c})$ with elements

$$A_{ij}(\mathbf{c}) = c_i A_{ij} c_j + (1 - c_i) \delta_{i,j}. \quad (7)$$

A diagonal regularization takes care of the case $i = j$ and $c_i = 0$. Then, we employ the Gaussian integrals to write

$$\det \mathbf{A}(\mathbf{c}) = \frac{(2\pi)^N}{\left\{ \int_{-\infty}^{+\infty} \prod_{i=1}^N dx_i e^{-\frac{1}{2} \sum_{i,j} x_i [c_i A_{ij} c_j + (1-c_i) \delta_{i,j}] x_j} \right\}^2}. \quad (8)$$

In this way

$$\begin{aligned} \ln \det \mathbf{A}(\mathbf{c}) &= N \ln(2\pi) - 2 \ln \left\{ \int \prod_i dx_i e^{-\frac{1}{2} \sum_{i,j} x_i [c_i A_{ij} c_j + (1-c_i) \delta_{i,j}] x_j} \right\}, \\ &= N \ln(2\pi) - 2N g(\mathbf{c}), \end{aligned} \quad (9)$$

or

$$\ln \det \mathbf{A}(\mathbf{c}) = N \ln(2\pi) - 2N g(\mathbf{c}), \quad (10)$$

where for brevity we defined

$$\int \prod_i dx_i e^{-\frac{1}{2} \sum_{i,j} x_i [c_i A_{ij} c_j + (1-c_i) \delta_{i,j}] x_j} = e^{N g(\mathbf{c})}. \quad (11)$$

Recall that $E(\mathbf{c}) = -\ln \det \mathbf{A}(\mathbf{c})$, therefore,

$$Z(\beta) = \sum_{\mathbf{c}} e^{-\beta E(\mathbf{c})} = e^{\beta N \ln(2\pi)} \sum_{\mathbf{c}} e^{-2\beta N g(\mathbf{c})}. \quad (12)$$

Note that $g(\mathbf{c})$ is a global function of the c_i . In the following, we write this quantity as a sum of local energy contributions by introducing other auxiliary variables to the problem. This allows us to apply the standard methods of estimating the free energy of systems with a local energy function.

A. Bethe approximation of the Gaussian integral

In this section we use Bethe approximation to find an estimation of the Gaussian integral in Eq. (11). In Appendix A, we briefly explain the method for a simple statistical model, see also Refs. [10–13,21]. The Bethe or belief propagation (BP) equations for the Gaussian integrals are recursive equations for cavity probability distributions $m_{i \rightarrow j}(x_i)$. This is probability density of x_i in the absence of interaction with variable x_j assuming that the interaction graph defined by $\mathbf{A}(\mathbf{c})$ has a tree structure. To write these equations, we consider the local interactions of x_i with its neighboring variables which are assumed to be statistically independent of each other,

$$\begin{aligned} m_{i \rightarrow j}(x_i) &\propto e^{-\frac{1}{2}((1-c_i)+c_i A_{ii})x_i^2} \\ &\times \prod_{k \neq i,j} \left[\int dx_k e^{-x_i c_i A_{ik} c_k x_k} m_{k \rightarrow i}(x_k) \right]. \end{aligned} \quad (13)$$

It is known that even in interaction graphs which are not tree (loopy graphs) these equations converge to a unique solution as long as $\mathbf{A}(\mathbf{c})$ is a diagonally dominant matrix [24,25], that is $|A_{ii}| \geq \sum_{j \neq i} |A_{ij}|$.

Now consider the following ansatz for the BP messages:

$$m_{i \rightarrow j}(x_i) \propto e^{-\frac{x_i^2}{2v_{i \rightarrow j}}}. \quad (14)$$

Note that because of the symmetry of the problem we are assuming that $\langle x_i \rangle = 0$. This results in a set of BP equations for the variances,

$$\frac{1}{v_{i \rightarrow j}} = (1 - c_i) + c_i A_{ii} - \sum_{k \neq i,j} c_i A_{ik}^2 c_k v_{k \rightarrow i}. \quad (15)$$

That is, the Gaussian ansatz is consistent with the structure of the Gaussian BP equations.

Given the solution to the BP equations, we write $g(\mathbf{c})$ in terms of the local contributions of the variables and interactions to the Gaussian integral [21],

$$N g(\mathbf{c}) = \sum_i \Delta g_i(\mathbf{c}) - \sum_{i < j} \Delta g_{ij}(\mathbf{c}), \quad (16)$$

where

$$\begin{aligned} e^{\Delta g_i} &= \int dx_i e^{-\frac{1}{2}[(1-c_i)+c_i A_{ii}]x_i^2} \\ &\times \prod_{j \neq i} \left[\int dx_j e^{-x_i c_i A_{ij} c_j x_j} m_{j \rightarrow i}(x_j) \right], \end{aligned} \quad (17)$$

$$e^{\Delta g_{ij}} = \int dx_i dx_j e^{-x_i c_i A_{ij} c_j x_j} m_{i \rightarrow j}(x_i) m_{j \rightarrow i}(x_j). \quad (18)$$

Later, we will also need a cavity contribution of the variables $\Delta g_{i \rightarrow j}$, which is like Δg_i but excluding one of the neighboring interactions,

$$e^{\Delta g_{i \rightarrow j}} = \int dx_i e^{-\frac{1}{2}[(1-c_i)+c_i A_{ii}]x_i^2} \times \prod_{k \neq i, j} \left[\int dx_k e^{-x_i c_i A_{ik} c_k x_k} m_{k \rightarrow i}(x_k) \right]. \quad (19)$$

Within the Gaussian ansatz for the BP messages, the above expressions for the local variable and interaction contributions are simplified to

$$2\Delta g_i = \ln(2\pi v_i), \quad (20)$$

$$2\Delta g_{ij} = -\ln\{v_{i \rightarrow j} v_{j \rightarrow i} \det[B(ij)]\}, \quad (21)$$

where

$$\frac{1}{v_i} = (1 - c_i) + c_i A_{ii} - \sum_{k \neq i} c_i A_{ik}^2 c_k v_{k \rightarrow i}, \quad (22)$$

$$B(ij) = \begin{pmatrix} \frac{1}{v_{i \rightarrow j}} & c_i A_{ij} c_j \\ c_i A_{ij} c_j & \frac{1}{v_{j \rightarrow i}} \end{pmatrix}. \quad (23)$$

Note that the sign of off-diagonal elements A_{ij} is irrelevant as long as the Gaussian Bethe equations are valid. That is, we can change the sign of any such element of the matrix without changing the determinant or the energy of the system within the above approximation.

In the next subsection, we use the above expressions for the determinants of $\mathbf{A}(\mathbf{c})$ to compute the partition function of the original problem.

B. A higher-level Bethe approximation of the partition function

Now we are ready to do the sum over the configurations

$$\sum_{\mathbf{c}} e^{-2\beta N g(\mathbf{c})} = e^{-\beta N f_g}, \quad (24)$$

which is needed in the partition function,

$$Z(\beta) = e^{-\beta N [f_g - \ln(2\pi)]}. \quad (25)$$

Here we defined f_g as the nontrivial contribution to the free energy $f = f_g - \ln(2\pi)$.

In the previous subsection we wrote $g(\mathbf{c})$ in terms of the Δg_i and Δg_{ij} which in turn depend on the BP messages $v_{i \rightarrow j}$. These messages satisfy the BP equations with a unique solution for diagonally dominant matrices. Therefore, we can write

$$\sum_{\mathbf{c}} e^{-2\beta N g(\mathbf{c})} = \sum_{\mathbf{c}} \sum_{\mathbf{v}_{\rightarrow}} e^{-2\beta(\sum_i \Delta g_i - \sum_{i < j} \Delta g_{ij})} \mathbb{I}(\mathbf{v}_{\rightarrow}), \quad (26)$$

that is, considering the $v_{i \rightarrow j}$ as auxiliary variables which are constrained by the indicator function $\mathbb{I}(\mathbf{v}_{\rightarrow}) = \prod_{i \rightarrow j} \delta(v_{i \rightarrow j} - v_{i \rightarrow j}^{\text{BP}})$ to satisfy the BP equations.

Again we can resort to the Bethe approximation to estimate the sum over the extended set of variables $\mathbf{c}, \mathbf{v}_{\rightarrow}$. We can do this because both the $g(\mathbf{c})$ and the BP constraints on the $v_{i \rightarrow j}$ are local functions of these variables. If the graph associated to matrix \mathbf{A} has a tree structure, then the interaction graph of the c_i and the auxiliary variables $v_{i \rightarrow j}$ is also a tree. Consider the cavity probability distribution $M_{i \rightarrow j}(c_i, v_{i \rightarrow j})$ of the variables

in the absence of interactions with node j . The cavity variables $\{(c_k, v_{k \rightarrow i}) : k \in \partial i \setminus j\}$ are independent of each other in a tree interaction graph. Thus $M_{i \rightarrow j}(c_i, v_{i \rightarrow j})$ is proportional to the product of the cavity probabilities $\{M_{k \rightarrow i}(c_k, v_{k \rightarrow i}) : k \in \partial i \setminus j\}$ for variables that are consistent with the hard constraint $\mathbb{I}(v_{i \rightarrow j})$. In addition, $M_{i \rightarrow j}(c_i, v_{i \rightarrow j})$ is also proportional to the Boltzmann factor $e^{\Delta g_i - \Delta g_{ij}} = e^{\Delta g_{i \rightarrow j}}$ which gives the statistical weight of the cavity variables [see Eq. (26)]. Therefore, the higher-level BP equations for the cavity probability distributions read

$$M_{i \rightarrow j}(c_i, v_{i \rightarrow j}) \propto \prod_{k \in \partial i \setminus j} \left[\sum_{c_k} \int dv_{k \rightarrow i} M_{k \rightarrow i}(c_k, v_{k \rightarrow i}) \right] \times \mathbb{I}(v_{i \rightarrow j}) e^{-2\beta \Delta g_{i \rightarrow j}}. \quad (27)$$

We solve these equations by a population dynamics algorithm which is explained in Appendix B. The probability distributions $M_{i \rightarrow j}(c_i, v_{i \rightarrow j})$ are represented by populations of the variables $P_{i \rightarrow j} = \{(c_i^a, v_{i \rightarrow j}^a) : a = 1, \dots, \mathcal{N}_p\}$ on each directed link ($i \rightarrow j$). Then members of the populations from the right-hand side of the equation are selected to update the members of the population in left-hand side according to the Boltzmann weights of the variables and the hard constraints of the BP equations [26]. All members of the populations are updated in a single iteration of the algorithm. In the stationary state of the population dynamics, say, after t_{eq} iterations, we obtain an estimation of the free energy,

$$N f_g = \sum_i \Delta f_i - \sum_{i < j} \Delta f_{ij}, \quad (28)$$

with local free energies that are given by

$$e^{-\beta \Delta f_i} = \sum_{c_i} \prod_{j \in \partial i} \left[\sum_{c_j} \int dv_{j \rightarrow i} M_{j \rightarrow i}(c_j, v_{j \rightarrow i}) \right] e^{-2\beta \Delta g_i}, \quad (29)$$

$$e^{-\beta \Delta f_{ij}} = \sum_{c_i, c_j} \int dv_{ij} e^{-2\beta \Delta g_{ij}} M_{i \rightarrow j}(c_i, v_{i \rightarrow j}) M_{j \rightarrow i}(c_j, v_{j \rightarrow i}). \quad (30)$$

Here $v_{ij} = \{v_{i \rightarrow j}, v_{j \rightarrow i}\}$.

TABLE I. The matrices studied in this paper with a brief description of the defining parameters. These are $N \times N$ matrices with indices $i, j = 1, \dots, N$. The RRGs have degree K and ∂i denotes the set of neighbors of i . In random matrices the off-diagonal elements are independent random numbers uniformly distributed in $(-2, 0)$.

Matrix	Elements
Laplacian of RRGs (Sec. III B)	$A_{ii} = K, A_{ij} = -1 (j \in \partial i), \partial i = K$
Quasi-Laplacian of RRGs (Sec. III B)	$A_{ii} = \Lambda, A_{ij} = -\gamma/K (j \in \partial i), \partial i = K, 0 < \gamma \leq \Lambda$
Random Laplacian of RRGs (Sec. III C)	$A_{ii} = -\sum_{j \neq i} A_{ij}, A_{ij} \in (-2, 0) (j \in \partial i), \partial i = K$
Scaling limit I of complete graphs (App. D)	$A_{ii} = N - 1, A_{ij} = -1 (j \neq i)$
Scaling limit II of complete graphs (App. D)	$A_{ii} = \Lambda, A_{ij} = -\gamma/N (j \neq i, 0 < \gamma \leq \Lambda)$

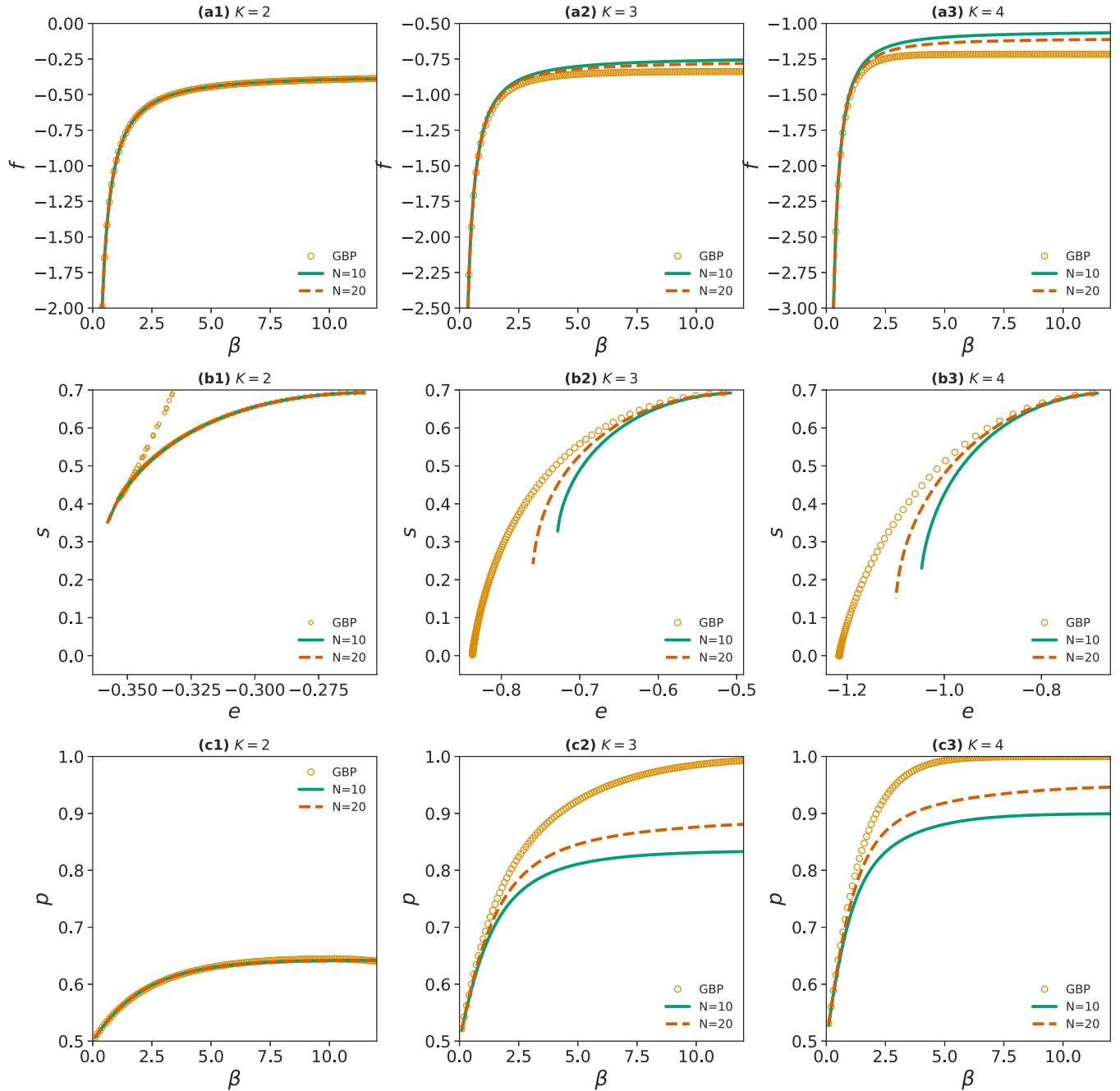


FIG. 1. The asymptotic behavior of minors for Laplacian of random regular graphs of degree $K = 2, 3, 4$. The free energy f , entropy vs energy $s(e)$, and probability of presence p are reported (Gaussian BP) and compared with exact numerical results for small sizes ($N = 10, 20$). β is the inverse of temperature in the partition function of the principal minors. The parameters of the population dynamics algorithm are as follows: population size $\mathcal{N}_p = 10^4$, equilibration time $t_{\text{eq}} = 10^6$, and averaging time $\Delta t_{\text{avg}} = 10^4$.

The above computations are simplified for homogeneous interaction graphs where by symmetry all equations for directed links ($i \rightarrow j$) reduce to a single equation. Here we consider the Laplacian of random regular graphs (RRGs) with the same degree (number of neighbors) K for all nodes. See Table I for definitions of the matrices we shall study in this work. In Appendix C we present the resulted equations which again are solved by a population dynamics algorithm. In this case a single population of the messages ($c_i, v_{i \rightarrow j}$) is enough to solve the higher-level BP equations. Figure 1 shows the

main quantities computed in this way for random regular graphs in the thermodynamic limit ($N \rightarrow \infty$). The results are obtained by a population of size $\mathcal{N}_p = 10^4$ after $t_{\text{eq}} = 10^6$ iterations of the population dynamics to reach a steady state where the average quantities are stationary. For comparison we also present the exact numerical results for small problem instances ($N = 10, 20$). As the figure shows finite-size effects are more pronounced for larger degrees $K = 3, 4$, where the entropy $s(e)$ decreases continuously to zero by increasing β . And the presence probability $p = \langle c_i \rangle$ approaches

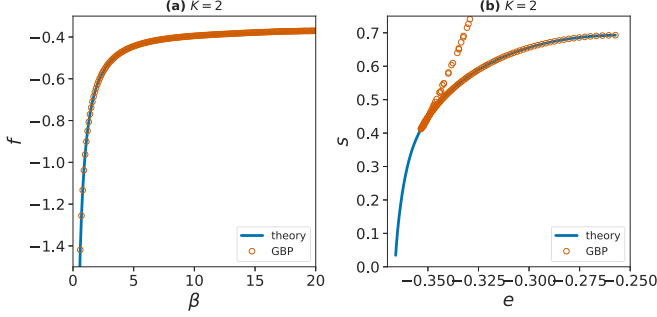


FIG. 2. The asymptotic behavior of minors for Laplacian of a closed chain. The free energy f and entropy vs energy $s(e)$ are reported (Gaussian BP) and compared with the exact solution (theory) of Ref. [1]. The parameters of the population dynamics algorithm are as follows: population size $\mathcal{N}_p = 10^4$, equilibration time $t_{\text{eq}} = 10^6$, and averaging time $\Delta t_{\text{avg}} = 10^4$.

monotonically to 1 for large β , where nearly all the indices are present in the relevant minors. No sign of a finite-temperature phase transition is observed here for $K = 3, 4$. By increasing β , the Boltzmann weight is smoothly concentrated more and more on minor configurations with a larger number of present indices.

The case $K = 2$ displays more interesting behaviors with a discontinuity in the entropy for large β and a probability p which approaches to a nontrivial value as β goes to infinity. We know that there are a subexponential number of ground states with extensive Hamming distances in the configuration space [1]. The ground states are dimer coverings with pairs of adjacent nodes which are separated with one unmatched node. However, no finite-temperature phase transition is expected to happen also for $K = 2$ because the entropic contributions to the free energy can easily destroy the ordered states of this one-dimensional system. We also observe an instability in the

BP equations for large β close to the discontinuity. Figure 2 compares the exact theoretical solution of Ref. [1] with the result we obtain by the above Gaussian BP equations. This clearly shows the region of instability where the entropy $s(e)$ is not anymore a concave function of the energy density. In fact, the number of iterations t_{eq} which is necessary to get close to the exact theoretical solution increases rapidly as one approaches the point of discontinuity.

To have a better picture of the phase space of these problems we consider two coupled replicas of the system at equilibrium. By controlling the strength of coupling between the replicas we can investigate the structure of the relevant minors around a given point of the configuration space. We define the partition function of two replicas as

$$Z(\beta, h) = \sum_{\mathbf{c}, \mathbf{c}'} e^{-\beta[E(\mathbf{c})+E(\mathbf{c}')-h\sum_i \delta_{c_i, c'_i}]}. \quad (31)$$

This gives the free energy f of the replicas as a function of β and the coupling h . In terms of the energy densities and the overlap $q = \frac{1}{N} \sum_i \delta_{c_i, c'_i}$, we have

$$Z(\beta, h) = e^{-\beta N f} = \int d e d e' d q e^{-\beta N [e + e' - h q - \frac{1}{\beta} s(e, e'; q)]}. \quad (32)$$

Here the entropy density is $s(e, e'; q) = \frac{1}{N} \ln \Omega(e, e'; q)$ where $\Omega(e, e'; q)$ is the number of two-replica configurations of energy densities e and e' with overlap q . In the thermodynamic limit, by the saddle-point approximation one obtains

$$f = 2e - hq - \frac{1}{\beta} s(e; q), \quad (33)$$

$$s(e; q) = \beta(2e - hq - f), \quad (34)$$

where the last equation for $s(e; q)$ is obtained after a Legendre transformation. Note that by symmetry the equilibrium energy densities of the two replicas are the same.

The higher-level BP equations here are marginal probability distributions for the replica variables

$$M_{i \rightarrow j}(c_i, c'_i, v_{i \rightarrow j}, v'_{i \rightarrow j}) \propto e^{\beta h \delta_{c_i, c'_i}} \prod_{k \in \partial i \setminus j} \left[\sum_{c_k, c'_k} \int d v_{k \rightarrow i} d v'_{k \rightarrow i} M_{k \rightarrow i}(c_k, c'_k, v_{k \rightarrow i}, v'_{k \rightarrow i}) \right] \mathbb{I}(v_{i \rightarrow j}) \mathbb{I}(v'_{i \rightarrow j}) e^{-2\beta(\Delta g_{i \rightarrow j} + \Delta g'_{i \rightarrow j})}. \quad (35)$$

As before we consider random regular graphs and solve these equations by population dynamics. The algorithm is very similar to the one presented in Appendix B except that we have the Boltzmann weight $e^{\beta h \delta_{c_i, c'_i}}$ that couples the two replicas. Figure 3 shows how the entropy, energy, and overlap of the two systems are related to each other in random regular graphs of degree $K = 2, 3$. For $K = 4$ (not shown) the qualitative behaviors are very similar to that of $K = 3$. Again we see the instability of the BP equations for large β and q in case $K = 2$. The discontinuity of the entropy for $K = 2$ is again observed in the way that the energy density e behaves with the overlap q . Compare it with the case $K = 3$ where e decreases monotonically by increasing q . Moreover, we see that when $K = 2$ the overlap changes abruptly at $h = 0$ for large β which is a signal of clustering of the ground states in this case. In contrast, for $K = 3$ the overlap smoothly approaches 1

with increasing β even at $h = 0$ which is consistent with the absence of phase transitions in these systems.

We end this section with an exact study of principal minors in matrices associated to fully connected graphs. In Appendix D we consider $N \times N$ matrices with elements $A_{ij} = \Lambda \delta_{i,j} - \Gamma(1 - \delta_{i,j})$ in two scaling limits: (I) $\Lambda = N - 1$, $\Gamma = 1$ and (II) $\Lambda = \text{finite}$, $\Gamma = \gamma/N$ as $N \rightarrow \infty$. In the latter case we assume that $0 \leq \gamma \leq \Lambda$ to represent positive and diagonally dominant matrices. We show that the behavior of the Laplacian of a complete graph, i.e., in the scaling limit (I), is very similar to that of Laplacian of random regular graphs of degree $K = 4$. There are $N - 1$ ground states with Hamming distances 2 where only one of the indices is not present in the minor configuration. The scaling limit (II), however, displays more interesting behaviors depending on the value of Λ . For $\Lambda > 1$ there are $N - 1$ ground states with only one index

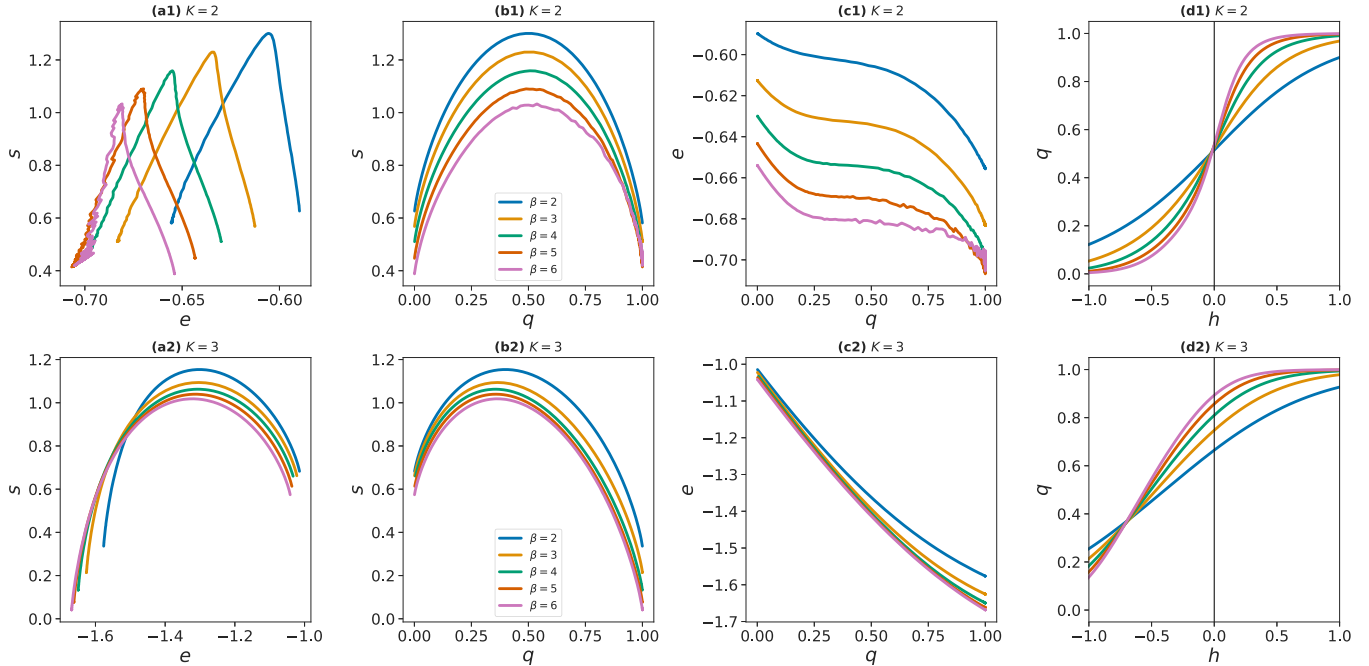


FIG. 3. The asymptotic behavior of two coupled minor configurations (replicas) for Laplacian of random regular graphs of degree $K = 2, 3$. h is the strength of coupling and q is the overlap of the two replicas. The parameters of the population dynamics algorithm are as follows: population size $\mathcal{N}_p = 10^4$, equilibration time $t_{\text{eq}} = 10^6$, and averaging time $\Delta t_{\text{avg}} = 10^4$.

absent in the maximal minor configurations. On the other side, for $\Lambda < 1$ there are $N - 1$ ground states with only one index present in the configurations. For $\Lambda = 1$ all configurations have the same energy and the ground-state entropy is maximal. Nevertheless, in both scaling limits we do not observe any finite-temperature phase transition; there is always a unique minimum of the free-energy function which changes smoothly by increasing the inverse temperature β .

Figures 4 and 5 show how finite connectivity of random regular graphs changes the above mean-field picture. We deviate from the Laplacian of random regular graphs in two different directions as depicted in Fig. 6. Let us start with the results displayed in Fig. 4. Here the diagonal matrix elements Λ can be smaller or larger than the connectivity degree $K = 2, 3, 4$ and the off-diagonal elements are $\Gamma = \Lambda/K$. This means that we are multiplying the Laplacian by Λ/K which in turn modifies the value of minors by a factor $(\Lambda/K)^l$ for a minor configuration of size l . That is $-\ln(\Lambda/K)$ is like a chemical potential which now controls the number of present indices in the minors. We see that for $\Lambda \leq 1$ the behavior of different degrees K is very similar to that of the complete graphs except the smaller ground-state entropy at $\Lambda = 1$. Exact enumerations in small systems show that like the complete graphs, in this regime the limit $\beta \rightarrow \infty$ of the entropy density coincides with the zero-temperature entropy.

For $\Lambda > 1$, we observe numerical instabilities which can be attributed to the discontinuous behavior of the entropy density at zero temperature. Consider for instance the case $K = 2$ (or a chain of length $N \gg 1$) where the energy of a connected cluster of l present indices is given by $E(l) = -\ln(l+1) - l \ln(\Lambda/K)$. A sequence of clusters of present indices (represented by 1s) which are separated by a single

absent index (shown by 0) make an ordered minor configuration for the chain graph; for instance, $\dots - 0 - 111 - 0 - 111 - 0 - 111 - 0 - \dots$ with $l = 3$ present indices in each cluster. The energy density of such a minor configuration is given by $e(l) = -[\ln(l+1) + l \ln(\Lambda/K)]/(l+1)$. This energy is minimized for a non-negative integer that is closer to $l^* = \frac{\Lambda}{K}e - 1$. The optimal size of clusters increases linearly with Λ starting from the all-zero minor configuration. A maximal minor configuration here can be considered as a close packing of clusters of effective size l^* centered around the zeros (absent indices). As mentioned before, for the Laplacian ($K = 2, \Lambda = 2$) an optimal configuration is a dimer covering, or a close packing of (nonoverlapping) rods of length 3. Moreover, it is easy to see that the number of such optimal configurations is of order N (by translation) so the entropy density at zero temperature is zero. For any finite cluster size l the Hamming distances between the ground states is extensive but such states are not stable for any finite temperature, because the extensive entropy of excitations dominates the finite energies of the domain walls in this one-dimensional system.

We see in Fig. 4 that for degrees $K = 3, 4$ the main quantities change smoothly as long as $\Lambda \geq K$, where the number density of present indices in the optimal configurations is 1. On the other hand, for $1 < \Lambda < K$ we observe a discontinuous entropy density at zero temperature as Λ approaches 1, very similar to what happens for the smaller degree $K = 2$. The ground states are random arrangements of an extensive number of zeros which are well separated on the interaction graph to maximize the number of spanning forests rooted at the zeros. The interval of Λ values in which a numerical instability is displayed is of course reduced by increasing the degree K approaching the mean-field limit. Figure 7

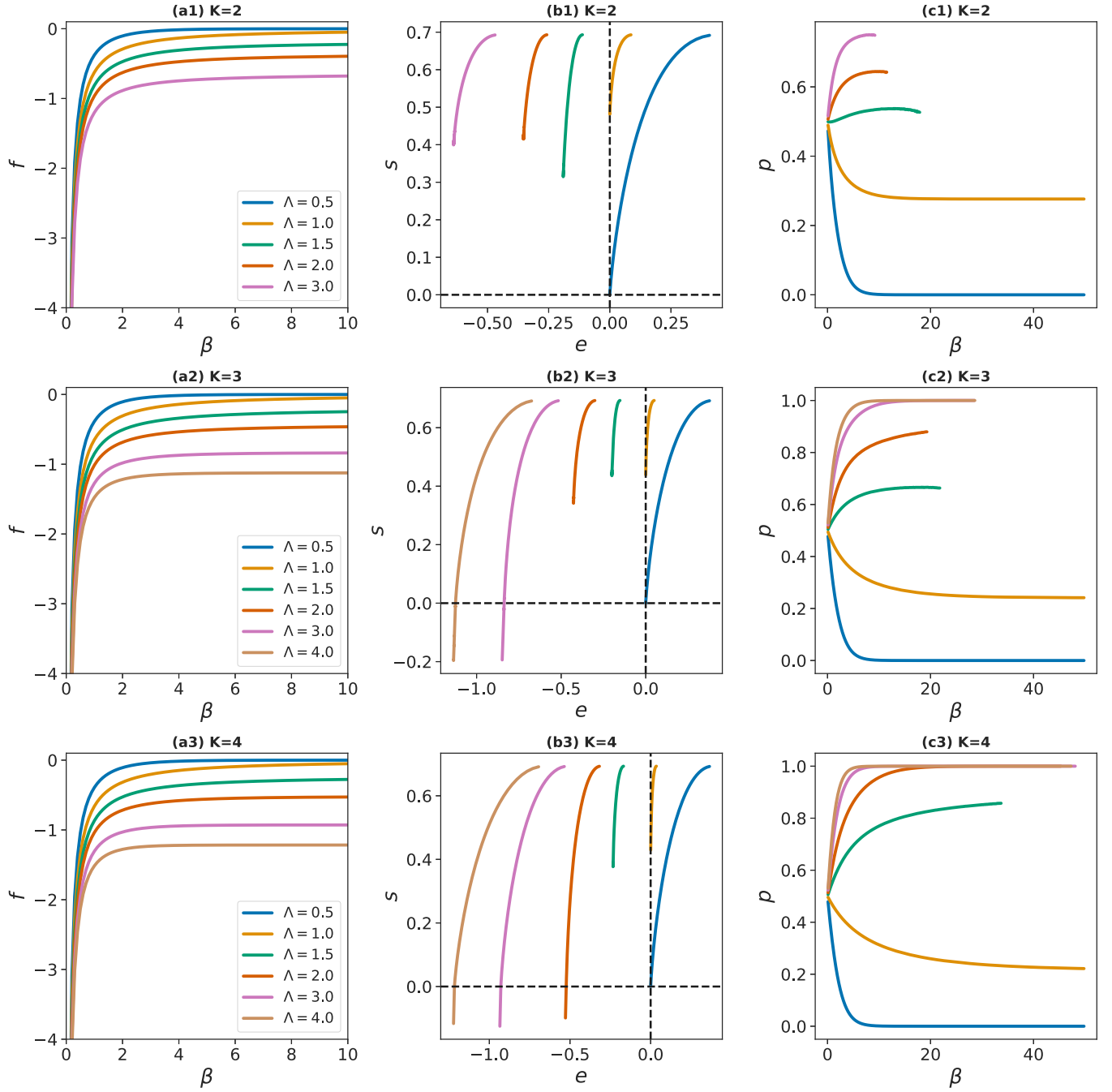


FIG. 4. The asymptotic behavior of minors vs Λ for quasi-Laplacian of random regular graphs of degree $K = 2, 3, 4$. The free energy f , entropy vs energy $s(e)$, and probability of presence p are reported for $\Lambda = 0.5$ to $\Lambda = K + 1$ in steps of size $\Delta\Lambda = 0.5$. The quasi-Laplacian matrix is defined by diagonal elements Λ and off-diagonal elements $-\Lambda/K$. The results are displayed for any $0 < \beta < 50$ as long as the entropy is concave and greater than -0.2 . The parameters of the population dynamics algorithm are as follows: population size $\mathcal{N}_p = 10^4$, equilibration time $t_{\text{eq}} = 10^6$, and averaging time $\Delta t_{\text{avg}} = 10^4$.

displays the exact numerical results we obtain for the sum of two-variable correlations $\chi = \frac{1}{N} \sum_{i < j} [\langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle]$ and $\chi_{\text{SG}} = \frac{1}{N} \sum_{i < j} [\langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle]^2$ with $\sigma_i = 2c_i - 1 = \pm 1$. The two susceptibilities χ and χ_{SG} are expected to diverge with N near a magnetic or spin-glass phase transition, respectively [13]. We observe that correlations remain short-ranged as β increases in the two cases ($K = 2, \Lambda = 2$) and ($K = 3, \Lambda = 1.5$) which display a discontinuity in the entropy density.

Figure 5 shows the results for matrices with diagonal elements K and off-diagonal elements $-\gamma/K$ when $\gamma \in (0, K)$. By decreasing the magnitude of the off-diagonals compared to the diagonals we approach the mean-field behavior. Here the picture is simpler, with curves that smoothly approach (up to very large β) the presence probability $p = 1$, where nearly all the indices are present in the optimal minor configurations. In Fig. 8 we report exact numerical results for variations in the number of minor configurations just above the optimal

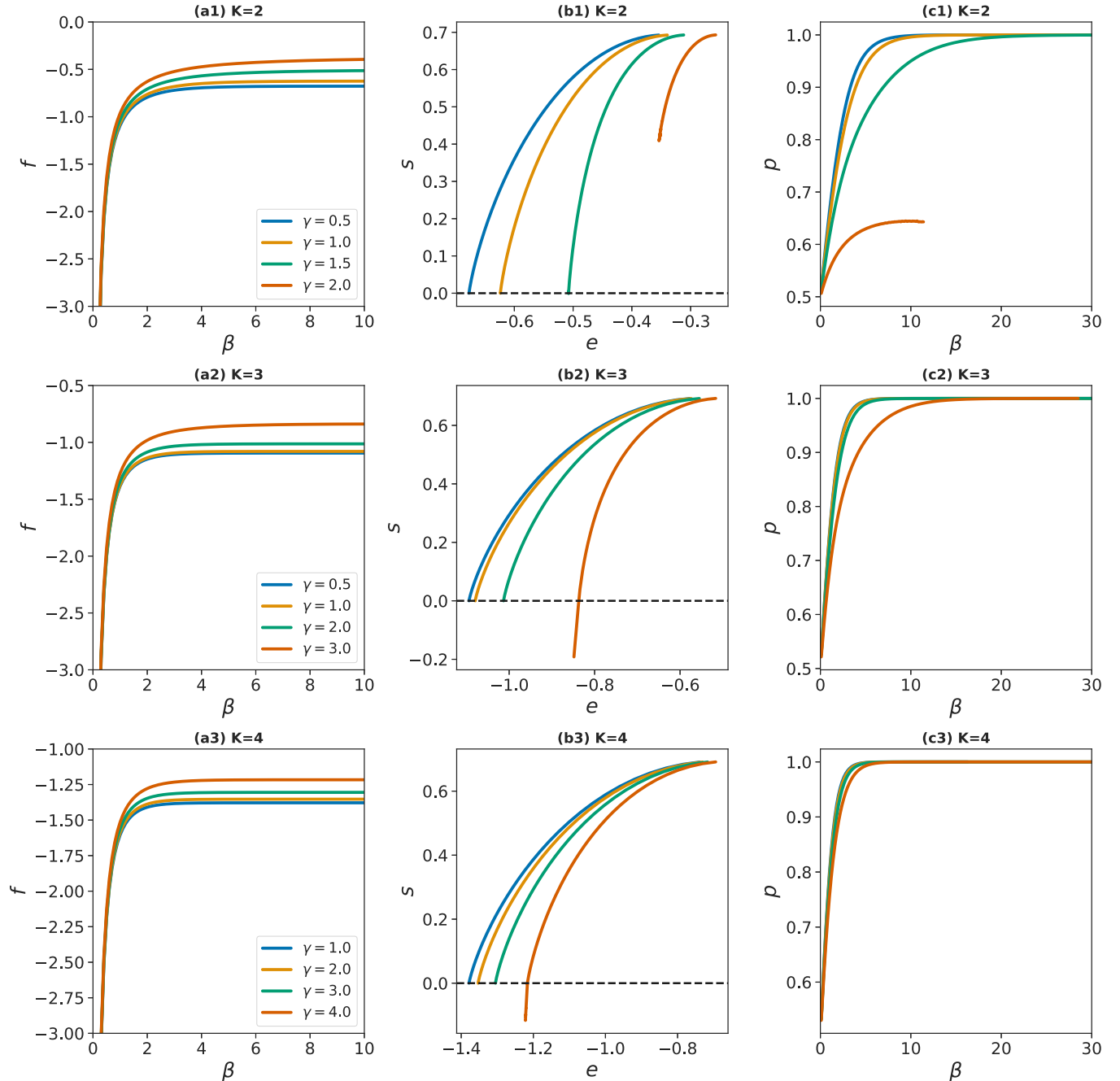


FIG. 5. The asymptotic behavior of minors vs γ for quasi-Laplacian of random regular graphs of degree $K = 2, 3, 4$. The free energy f , entropy vs energy $s(e)$, and probability of presence p are reported for $\gamma \in (0, K)$. The quasi-Laplacian matrix is defined by diagonal elements K and off-diagonal elements $-\gamma/K$. The results are displayed for any $0 < \beta < 50$ as long as the entropy is concave and greater than -0.2 . The parameters of the population dynamics algorithm are as follows: population size $\mathcal{N}_p = 10^4$, equilibration time $t_{\text{eq}} = 10^6$, and averaging time $\Delta t_{\text{avg}} = 10^4$.

ones in small systems of sizes $N = 18, 20, 22$. We see that the entropy gap remains nonzero (for $K = 2$) or approaches zero (for $K = 3, 4$) when the number of present indices changes by deviating from the Laplacian in the diagonal direction of Fig. 6. Again we observe that the entropy density displays a discontinuity for $K = 2$, and for $K = 3, 4$ when Λ is close to 1. On the other hand, the entropy gap tends to zero when we approach the mean-field limit, except for $K = 2$ and γ close to K .

C. Zero-temperature limit

At zero temperature the Gibbs probability measure is concentrated on the ground states of the system with minimum energy $E = -\ln \det \mathbf{A}(\mathbf{c})$, i.e., the maximal minors of the matrix \mathbf{A} . To study the ground states we take the limit $\beta \rightarrow \infty$ of the higher-level BP equations. Let us assume that $M_{i \rightarrow j}(c_i, v_{i \rightarrow j}) = e^{\beta w_{i \rightarrow j}(c_i, v_{i \rightarrow j})}$ as $\beta \rightarrow \infty$. Then, from Eq. (27) we obtain the so-called MS equations [21] for the

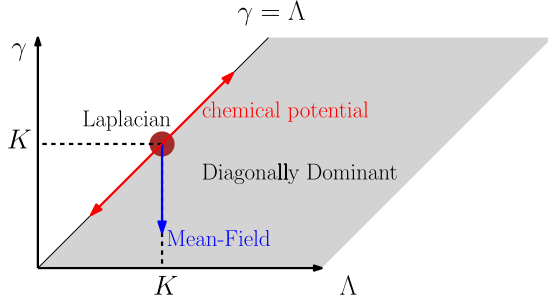


FIG. 6. Deviation from the Laplacian in the region of diagonally dominant matrices. Moving in the diagonal direction $\gamma = \Lambda$ is like introducing a chemical potential for the size of minor configurations. The mean-field picture is recovered in the other direction $\gamma \rightarrow 0$.

cavity messages $w_{i \rightarrow j}(c_i, v_{i \rightarrow j})$, see also Appendix A,

$$\begin{aligned} w_{i \rightarrow j}(c_i, v_{i \rightarrow j}) &= -2\Delta g_{i \rightarrow j} + \max_{\{c_k, v_{k \rightarrow i} : k \in \partial i \setminus j\}; \mathbb{I}(v_{i \rightarrow j})} \sum_{k \neq i, j} w_{k \rightarrow i}(c_k, v_{k \rightarrow i}) \\ &\quad - C_{i \rightarrow j}. \end{aligned} \quad (36)$$

The constant $C_{i \rightarrow j}$ is chosen such that $\max_{c_i, v_{i \rightarrow j}} w_{i \rightarrow j}(c_i, v_{i \rightarrow j}) = 0$.

To solve the above equations by iteration we use a discrete representation of the variances $v_{i \rightarrow j} = n_{i \rightarrow j} \delta v$ with integers $n_{i \rightarrow j}$ for a small δv . The maximum over the variables

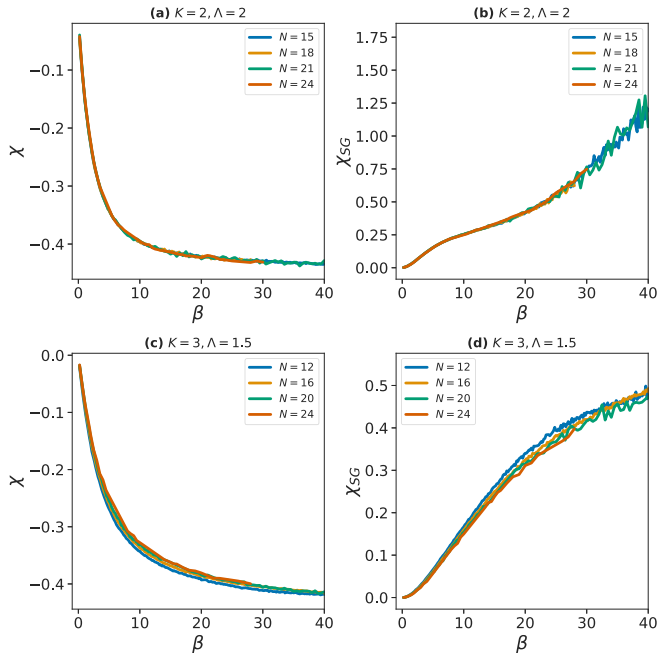


FIG. 7. The sum of two-variable correlations when the entropy density displays a discontinuity. The ferromagnetic and spin-glass susceptibilities, χ and χ_{SG} are exactly computed for random regular graphs of small sizes N . The matrices have diagonal elements Λ and off-diagonal elements $-\Lambda/K$. Panels (a) and (b) show the results for random regular graphs of degree $K = 2$ when $\Lambda = 2$. Panels (c) and (d) show the results for random regular graphs of degree $K = 3$ when $\Lambda = 1.5$.

$\{c_k, v_{k \rightarrow i} : k \in \partial i \setminus j\}$ can efficiently be computed by using repeated convolutions of the messages $w_{k \rightarrow i}$. In this way, the time complexity of each iteration of the algorithm in a graph of degree K is of order $N(KN_v)^2$, where N_v is the number of the possible values of discrete variances. In each iteration all the cavity messages are updated in a random and sequential way. The number of iterations needed to solve the equations is of order 100, independent of the problem size N .

In the same way, we can take the limit $\beta \rightarrow \infty$ of the free energies in Eqs. (29),

$$\Delta f_i = - \max_{c_i} \left\{ \max_{\{c_j, v_{j \rightarrow i} : j \in \partial i\}} \left[\sum_{j \in \partial i} w_{j \rightarrow i}(c_j, v_{j \rightarrow i}) - 2\Delta g_i \right] \right\}, \quad (37)$$

and

$$\Delta f_{ij} = - \max_{\{s_j, v_{ij}\}} [w_{i \rightarrow j}(c_i, v_{i \rightarrow j}) + w_{j \rightarrow i}(c_j, v_{j \rightarrow i}) - 2\Delta g_{ij}]. \quad (38)$$

These equations are useful to obtain an estimation of the ground-state energy within the Bethe approximation.

In practice, we use a reinforcement algorithm to find a minimum energy configuration of the system [27]. The idea is to reinforce the MS equations by introducing a bias feedback to the equations. The bias is provided by the local MS messages to gradually converge the algorithm towards a ground state of the system. More precisely, the reinforced MS equations at iteration $t + 1$ read as follows:

$$\begin{aligned} w_{i \rightarrow j}^{t+1}(c_i, v_{i \rightarrow j}) &= r(t)w_i^t(c_i) - 2\Delta g_{i \rightarrow j} + \max_{\{c_k, v_{k \rightarrow i} : k \in \partial i \setminus j\}; \mathbb{I}(v_{i \rightarrow j})} \\ &\quad \times \sum_{k \neq i, j} w_{k \rightarrow i}^t(c_k, v_{k \rightarrow i}) - C_{i \rightarrow j}, \end{aligned} \quad (39)$$

with the initial MS messages $w_{i \rightarrow j}^0 = w_i^0 = 0$. The reinforcement parameter $r(t)$ increases linearly with time as $r(t + 1) = r(t) + \delta r$, starting from $r(0) = 0$ with $\delta r \ll 1$. Similarly, the local MS messages are given by

$$\begin{aligned} w_i^{t+1}(c_i) &= r(t)w_i^t(c_i) \\ &\quad + \max_{\{c_j, v_{j \rightarrow i} : j \in \partial i\}} \left[\sum_{j \in \partial i} w_{j \rightarrow i}^t(c_j, v_{j \rightarrow i}) - 2\Delta g_i \right]. \end{aligned} \quad (40)$$

These messages then provide a candidate for the minimum energy configuration,

$$c_i^* = \arg \max w_i^t(c_i). \quad (41)$$

Note that the MaxSum algorithm is not expected to work well in a loopy graph but the algorithm can always be used to provide a candidate for the optimal configuration and therefore an upper bound for the minimum energy even in a loopy interaction graph.

We start by applying the above algorithm to Laplacian of a closed chain where we know the ground-state energy and configurations. Recall that the maximal minors of a chain are dimer coverings which are separated by extensive Hamming distances in the configuration space. We will see that this nontrivial structure of the ground states can make it difficult

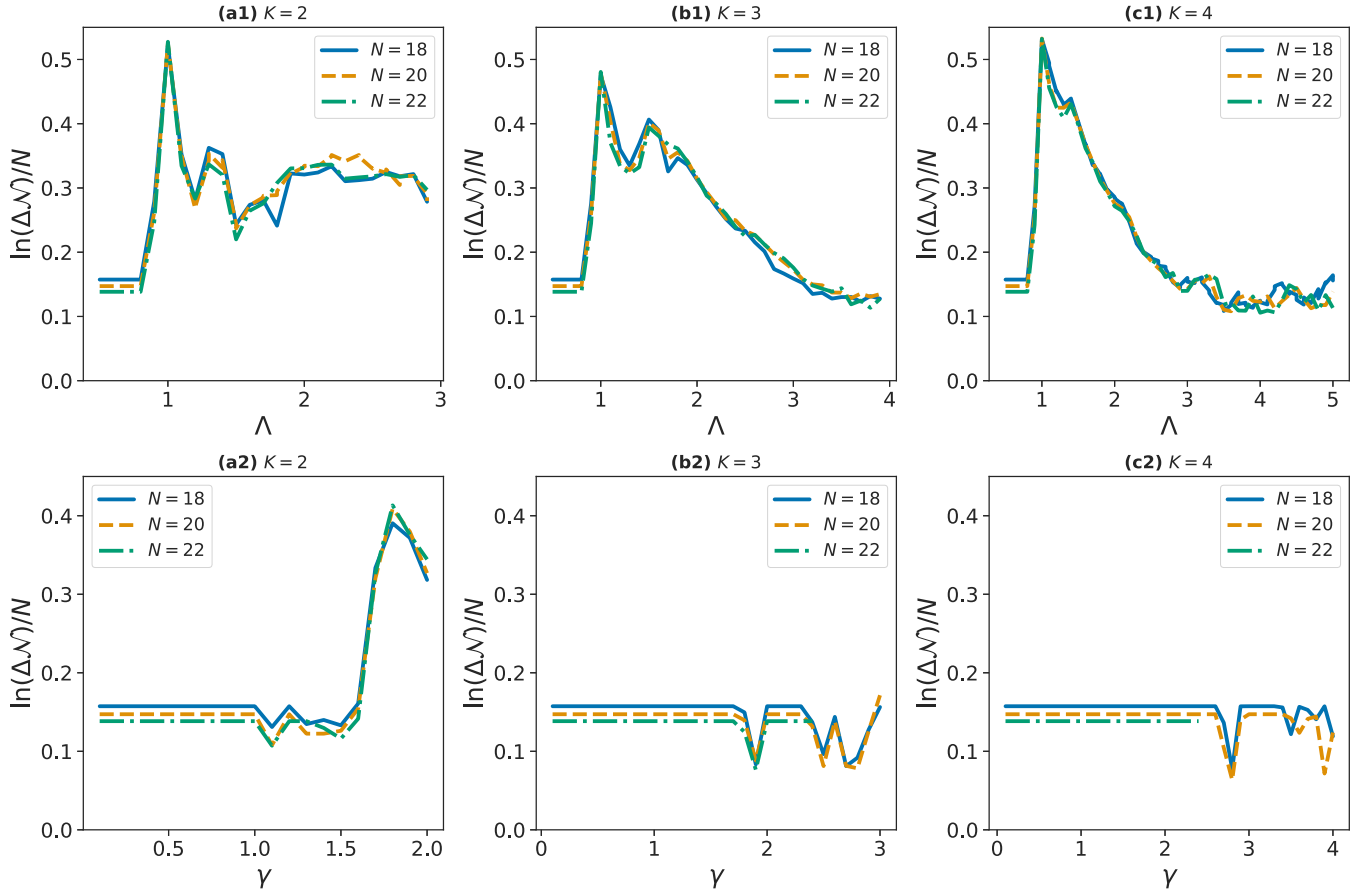


FIG. 8. The gap in the number of minors $\Delta\mathcal{N}$ above the optimal minors in random regular graphs of degree $K = 2, 3, 4$. The results are obtained by exhaustive enumeration for small sizes $N = 18, 20, 22$. Panels (a1), (b1), and (c1) show the changes with the chemical potential when the diagonal elements are Λ and off-diagonal elements are $-\Lambda/K$. Panels (a2), (b2), and (c2) show the variations in the mean-field direction when the diagonal elements are K and off-diagonal elements are $-\gamma/K$. The curves start with horizontal lines which correspond to $\frac{1}{N} \ln(N)$.

for a local optimization algorithm to find an optimal solution of the problem. In fact, an algorithm may try to minimize energy of the system by constructing a configuration that is locally like one of the ground states of the system. If these states are very different, then it could be very difficult to overcome these energy and entropy barriers and to end up with one of the optimal states of the system. Figure 9 shows how the minimum energy density suggested by the reinforced MS algorithm approaches the theoretical value $e_0 = -\ln(3)/3$ expected for a very large chain. For comparison, we also report the results obtained by a simulated annealing (SA) algorithm; in each step, one variable c_i is selected and its value is changed with probability $\min\{1, e^{-\beta\Delta E}\}$, where ΔE is the resulted change in the energy function. In one iteration of the algorithm all variables are selected in a random and sequential way, as in the MS algorithm. We start from a high temperature (small β) and in each iteration increase the inverse temperature by $\delta\beta$ to approach a low-energy configuration. We see in Fig. 9 that it is more difficult for the SA algorithm to find a ground state than the MS algorithm. A similar behavior is observed in Fig. 10 for random Laplacian of a closed chain. Here the off-diagonal elements are random and independent numbers uniformly distributed in $(-2, 0)$ to have the average -1 as in the ordered case. The diagonal elements in each row are

minus the sum of the off-diagonal elements in that row, like a Laplacian. Note that the MS algorithm finds a good estimation of the minimum energy in a smaller number of iterations compared to the SA algorithm. Moreover, a single iteration of SA algorithm is more time-consuming than the MS algorithm because we need to compute the matrix determinants for each update of the configuration. This does not allow us in practice to study larger systems by the simulated annealing algorithm.

Figure 11 displays the results we obtain for random Laplacian of random regular graphs of degree $K = 3, 4$. The random Laplacians are constructed as described in the previous paragraph. We observe that the SA algorithm is more effective for large connectivity degrees compared to the case of chain ($K = 2$) because the energy landscape is simpler for $K > 2$. For the same reason, the MS algorithm very quickly converges to the ground state of these systems. The optimal states of random Laplacians are indeed very close to the ground states of the ordered systems with only a small fraction of indices not present in the optimal states. The ground states in the ordered versions of these systems are minor configurations with nearly all indices, except one or two of them, present in the configuration. For instance, only one of the indices is not present in the ground states of the Laplacian of random regular graphs of degree $K = 4$. There are

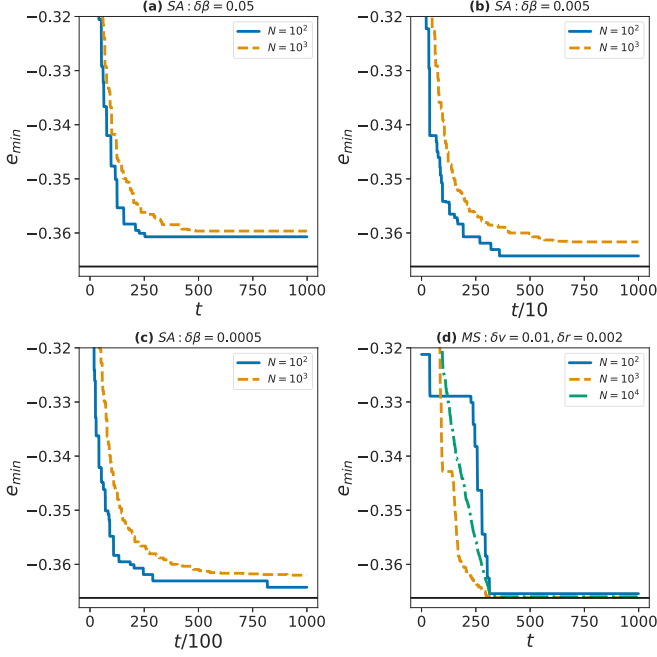


FIG. 9. Finding a maximal minor of Laplacian of a closed chain of size N by an optimization algorithm. The results are obtained by the SA and MS algorithms with the indicated parameters. The horizontal lines show the exact value expected in the thermodynamic limit. The ground states are dimer coverings of energy density $e_0 = -\ln(3)/3$. Here t is the number of iterations in the algorithms. In each iteration all the variables are updated once.

therefore N ground states in such a homogeneous system and the Hamming distance between two optimal configurations is 2. This trivial structure of the ground states makes it also easy for a local optimization algorithm like SA to find a ground state.

IV. CONCLUSION AND DISCUSSION

In summary, a Gaussian representation of minors for symmetric, positive, and diagonally dominant (DD) matrices was employed to estimate the free energy, entropy, and energy of relevant minors at different temperatures by the Bethe approximation. The estimation is expected to be asymptotically exact for matrices which have a locally treelike graph representation. Specifically, we studied the energy and entropy landscape of the Laplacian of random regular graphs of degree $K = 2, 3, 4$. For large degrees $K = 3, 4$ the interesting quantities change smoothly by increasing the inverse temperature β . The case $K = 2$, however, displays a discontinuous entropy and numerical instability close to this discontinuity which separates the zero- and finite-temperature behaviors. Nevertheless, we do not observe a finite-temperature phase transition in these systems for $K = 2, 3, 4$. The same result is obtained by an exact treatment of principal minors in DD matrices which are defined by homogeneous fully connected graphs.

The zero-temperature limit of the Bethe equations is used as an optimization (MaxSum) algorithm to find an estimation of the ground-state energy and optimal configurations (maximal minors) of random DD matrices. Here the

MaxSum algorithm is more efficient than the standard simulated annealing algorithm regarding the number of iterations that are needed to find a good estimation of the ground state(s). The time complexity of the MaxSum algorithm is proportional to N in a finite-connectivity graph which is represented by a sparse $N \times N$ matrix.

The first part of the study enables us to estimate the sum of powers of principal minors which is relevant to computation of the partition function and the Shannon-Rényi entropy of quantum systems such as the Hubbard model and the transverse field Ising model [1]. In the second part, we obtain an approximate message-passing optimization algorithm which can be applied to sampling problems where a subset of maximally independent configurations are needed as in the determinantal point processes.

It seems that restriction to diagonally dominant matrices is the main reason behind the absence of finite-temperature phase transitions in such systems. This is the case also for positive but quasi-one-dimensional matrices where entropy is dominated for any finite temperature. It would be interesting to investigate the nature of the possible phase transitions for an arbitrary positive matrix. If the problem of finding the optimal minors is in general a computationally hard problem, then one expects to observe spin-glass phases in these systems. In this study, we were mainly focused on regular quasi-Laplacian matrices with no randomness in matrix elements. A more detailed investigation of random and diagonally dominant matrices is needed to characterize the phase diagram of this class of matrices.

Recall that the principal minors of a Laplacian give the number of possible spanning forests that are generated by a set of trees rooted at the zeros of the minor configuration (the absent indices). By maximizing the number of such forests we are indeed increasing the strength of an effective repulsive interaction between the zeros of the minor configuration. In this study we used a chemical potential to control the expected number of roots (zeros) in the relevant minors. We showed that a maximal minor in a chain ($K = 2$) is a close packing of the roots with an effective distance l^* which is determined by the chemical potential. It would be interesting to investigate the statistical properties of this system of repulsive zeros, for example in a Hamming space from the perspective of the coding theory.

ACKNOWLEDGMENTS

M.A.R. thanks CNPq and FAPERJ (Grant No. E-26/210.062/2023) for partial support. This work was performed using the Academic Leiden Interdisciplinary Cluster Environment (ALICE), the compute resources provided by Leiden University.

APPENDIX A: BETHE APPROXIMATION: BELIEF PROPAGATION ALGORITHM

In this section we briefly describe the Bethe approximation and the resulting BP algorithm for an Ising model defined on an interaction graph G . Consider N Ising variables of configurations $\mathbf{s} = \{s_i = \pm 1 : 1, \dots, N\}$ and energy function $E[\mathbf{s}] = -\sum_{(ij)} J_{ij} s_i s_j - \sum_i h_i s_i$. Note that the

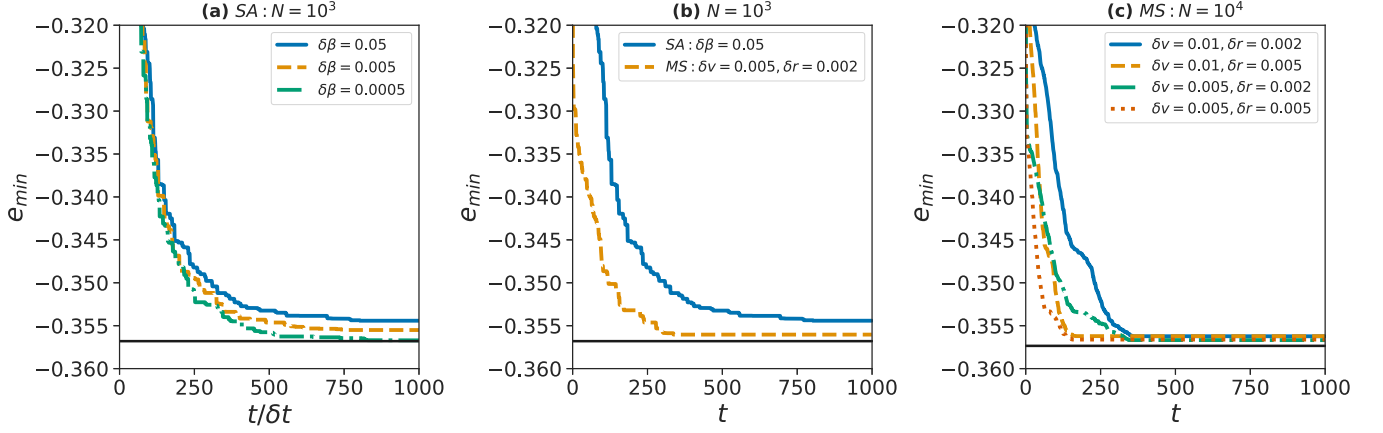


FIG. 10. Finding a maximal minor of a random Laplacian of a closed chain of size N by an optimization algorithm. The results are obtained by the SA and MS algorithms with the indicated parameters. The off-diagonal elements of the matrices are uniformly distributed in $(-2, 0)$ and the diagonal elements in each row are minus the sum of other elements in that row. The horizontal lines show the numerical value predicted by the MaxSum equations. The ground states are random configurations close to the dimer coverings of the ordered system. The number of iterations in the SA algorithm is scaled with $\delta t = 1, 10, 100$ for $\delta\beta = 0.05, 0.005, 0.0005$, respectively.

interaction term is a sum over edges (ij) of the interaction graph G . In thermal equilibrium at inverse temperature β , the Boltzmann factor $e^{-\beta E[s]}/Z$ gives the statistical weight of configurations \mathbf{s} , where the partition function $Z = \sum_{\mathbf{s}} e^{-\beta E[s]}$ is a normalization constant. Here we are interested in the marginal probabilities of local variables $\mu_i(s_i)$ at equilibrium. In limit $\beta \rightarrow \infty$ these marginals can be used to find the ground state(s) of the system.

Bethe equations are indeed recursive equations for the cavity marginals $\mu_{i \rightarrow j}(s_i)$, see Fig. 12. This is the probability of state s_i for variable i in the absence of interaction with variable j , i.e., when the interaction term $-J_{ij}s_i s_j$ is removed from the energy function. It is also assumed that in the absence of this interaction s_i is independent of the state of the other neighbors of variable j . This assumption is valid when the interaction

graph G is a tree, or has very large loops such that locally it is like a tree. Then, the Bethe equations for the cavity marginals read

$$\mu_{i \rightarrow j}(s_i) \propto e^{\beta h_i s_i} \prod_{k \in \partial i \setminus j} \left[\sum_{s_k} e^{\beta J_{ik} s_i s_k} \mu_{k \rightarrow i}(s_k) \right], \quad (\text{A1})$$

were ∂i is the set of neighbors of variable i in the interaction graph. The normalization constant is obtained from $\sum_{s_i} \mu_{i \rightarrow j}(s_i) = 1$. These equations can be solved by iteration starting from random initial messages $\mu_{i \rightarrow j}(s_i)$ and updating the messages according to the above equations. Having the cavity marginals one finds the local marginals $\mu_i(s_i)$ by considering the effects of all neighboring variables,

$$\mu_i(s_i) \propto e^{\beta h_i s_i} \prod_{k \in \partial i} \left[\sum_{s_k} e^{\beta J_{ik} s_i s_k} \mu_{k \rightarrow i}(s_k) \right]. \quad (\text{A2})$$

Moreover, an estimation of the free energy $F = -\frac{1}{\beta} \ln Z$ can be obtained from the cavity marginals. To this end, we need to compute the contributions of all variables and interactions in the partition function. Assuming that the interaction graph

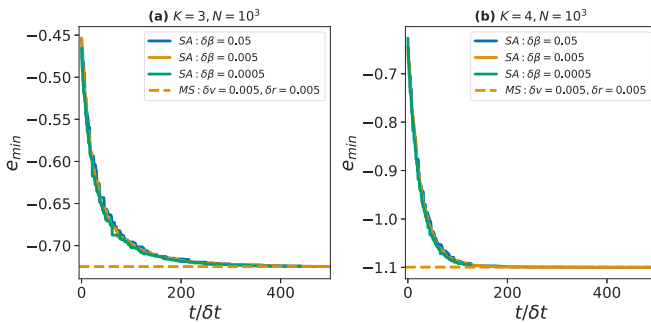


FIG. 11. Finding a maximal minor of a random Laplacian of random regular graphs of degree $K = 3, 4$ by an optimization algorithm. The results are obtained by the SA and MS algorithms with the indicated parameters. The off-diagonal elements of the matrices are uniformly distributed in $(-2, 0)$ and the diagonal elements in each row are minus the sum of other elements in that row. The horizontal lines show the minimum energies obtained by the MaxSum algorithm in a few hundred iterations. The fraction of present indices are around $p = 0.88 (K = 3)$ and $p = 0.99 (K = 4)$. The number of iterations in the SA algorithm is scaled with $\delta t = 1, 10, 100$ for $\delta\beta = 0.05, 0.005, 0.0005$, respectively.

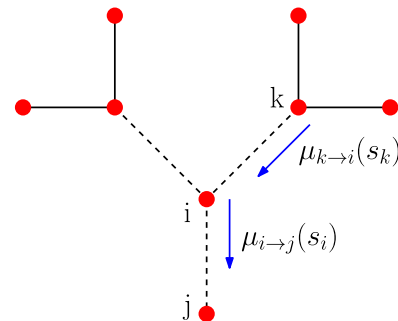


FIG. 12. An illustration of the BP algorithm in a tree interaction graph. The node variable i sends the cavity message $\mu_{i \rightarrow j}(s_i)$ to node j depending on the messages $\mu_{k \rightarrow i}(s_k)$ that it receives from the other neighbors.

is a tree, the free energy is given by

$$F = \sum_i \Delta F_i - \sum_{(ij)} \Delta F_{ij}. \quad (\text{A3})$$

This is obtained by writing a recursive equation for the partition function starting from an arbitrary variable [21]. The local free-energy changes are related to the cavity marginals as follows:

$$e^{-\beta \Delta F_i} = \sum_{s_i} e^{\beta h_i s_i} \prod_{k \in \partial i} \left[\sum_{s_k} e^{\beta J_{ik} s_i s_k} \mu_{k \rightarrow i}(s_k) \right], \quad (\text{A4})$$

$$e^{-\beta \Delta F_{ij}} = \sum_{s_i, s_j} e^{\beta J_{ij} s_i s_j} \mu_{i \rightarrow j}(s_i) \mu_{j \rightarrow i}(s_j). \quad (\text{A5})$$

To obtain some information about the ground states, we take the limit $\beta \rightarrow \infty$ and assume that $\mu_{i \rightarrow j}(s_i) = e^{\beta m_{i \rightarrow j}(s_i)}$. The BP equations in this limit (called MaxSum equations) are

$$m_{i \rightarrow j}(s_i) = h_i s_i + \sum_{k \in \partial i \setminus j} \max_{s_k} [J_{ik} s_i s_k + m_{k \rightarrow i}(s_k)] - C_{i \rightarrow j}, \quad (\text{A6})$$

where $C_{i \rightarrow j}$ is obtained by normalization condition $\max_{s_i} m_{i \rightarrow j}(s_i) = 0$. As before, the equations are solved by iteration. It is useful for a better convergence to introduce reinforcement in the algorithm; at each step of the iteration algorithm a bias field (feedback) is added to the system to favor the more probable states of the variables. More precisely, the reinforced MaxSum equations at time step $t + 1$ read as follows:

$$m_{i \rightarrow j}^{t+1}(s_i) = h_i s_i + r(t) m_i^t(s_i) + \sum_{k \in \partial i \setminus j} \max_{s_k} [J_{ik} s_i s_k + m_{k \rightarrow i}^t(s_k)] - C_{i \rightarrow j}. \quad (\text{A7})$$

Similarly, one obtains the local messages,

$$m_i^{t+1}(s_i) = h_i s_i + r(t) m_i^t(s_i) + \sum_{k \in \partial i} \max_{s_k} [J_{ik} s_i s_k + m_{k \rightarrow i}^t(s_k)] - C_i. \quad (\text{A8})$$

The reinforcement parameter $r(t)$ is zero at the beginning of the algorithm and increases slowly by the number of iterations. At each time step, one can find an approximate ground state of the system by computing the maximal states of the local messages, that is,

$$s_i^* = \arg \max_{s_i} m_i^t(s_i). \quad (\text{A9})$$

For more details and applications of the BP and MaxSum algorithms in other problems see [21].

APPENDIX B: THE POPULATION DYNAMICS ALGORITHM

In this section we describe the population dynamics that is used to solve the higher-level BP equations in the Gaussian

representation of the minors,

$$M_{i \rightarrow j}(c_i, v_{i \rightarrow j}) \propto \prod_{k \in \partial i \setminus j} \left[\sum_{c_k} \int dv_{k \rightarrow i} M_{k \rightarrow i}(c_k, v_{k \rightarrow i}) \right] \times \mathbb{I}(v_{i \rightarrow j}) e^{-2\beta \Delta g_{i \rightarrow j}}. \quad (\text{B1})$$

Recall that the variances $v_{i \rightarrow j}$ are solutions to the BP equations for the Gaussian variables,

$$\frac{1}{v_{i \rightarrow j}} = (1 - c_i) + c_i A_{ii} - \sum_{k \neq i, j} c_i A_{ik}^2 c_k v_{k \rightarrow i}, \quad (\text{B2})$$

and

$$2\Delta g_{i \rightarrow j} = \ln(2\pi v_{i \rightarrow j}). \quad (\text{B3})$$

The probability distributions $M_{i \rightarrow j}(c_i, v_{i \rightarrow j})$ are represented by populations of the variables $P_{i \rightarrow j} = \{(c_i^a, v_{i \rightarrow j}^a) : a = 1, \dots, \mathcal{N}_p\}$ on each directed link ($i \rightarrow j$). In addition, we introduce populations of the Boltzmann weights $W_{i \rightarrow j} = \{e^{-2\beta \Delta g_{i \rightarrow j}^a} : a = 1, \dots, \mathcal{N}_p\}$ which are associated to members $(c_i^a, v_{i \rightarrow j}^a)$ of the $P_{i \rightarrow j}$. More about the population dynamics in general can be found in Ref. [26].

At the beginning of the algorithm we set the initial values $v_{i \rightarrow j}^a = 0$, $c_i^a = 0, 1$ (with equal probability) and $\Delta g_{i \rightarrow j}^a = \infty$. In each iteration of the population dynamics we go through all the directed links ($i \rightarrow j$) in a random and sequential way and do the following steps:

- (i) select a member $(c_k^{a_k}, v_{k \rightarrow i}^{a_k})$ from $P_{k \rightarrow i}$ for $k \in \partial i \setminus j$;
- (ii) compute $v_{i \rightarrow j}(c_i)$ and $\Delta g_{i \rightarrow j}(c_i)$ [Eqs. (B2) and (B3)] for $c_i = 0, 1$ given the messages $(c_k^{a_k}, v_{k \rightarrow i}^{a_k})$;
- (iii) select a member $(c_i^{a_0}, v_{i \rightarrow j}^{a_0})$ of $P_{i \rightarrow j}$ and replace it with $(0, v_{i \rightarrow j}(0))$ with probability $e^{-2\beta[\Delta g_{i \rightarrow j}^{a_0} - \Delta g_{i \rightarrow j}(0)]}$. If accepted, then replace $\Delta g_{i \rightarrow j}^{a_0}$ with $\Delta g_{i \rightarrow j}(0)$;
- (iv) select a member $(c_i^{a_1}, v_{i \rightarrow j}^{a_1})$ of $P_{i \rightarrow j}$ and replace it with $(1, v_{i \rightarrow j}(1))$ with probability $e^{-2\beta[\Delta g_{i \rightarrow j}^{a_1} - \Delta g_{i \rightarrow j}(1)]}$. If accepted, then replace $\Delta g_{i \rightarrow j}^{a_1}$ with $\Delta g_{i \rightarrow j}(1)$;

Note that the members of the populations are selected randomly and uniformly. The updates are repeated for t_{eq} iterations to reach a steady state where the average quantities are stationary.

In the stationary state of the population dynamics, we obtain an estimation of the free energy,

$$N f_g = \sum_i \Delta f_i - \sum_{i < j} \Delta f_{ij}, \quad (\text{B4})$$

with local free energies that are given by

$$e^{-\beta \Delta f_i} = \sum_{c_i} \prod_{j \in \partial i} \left[\sum_{c_j} \int dv_{j \rightarrow i} M_{j \rightarrow i}(c_j, v_{j \rightarrow i}) \right] e^{-2\beta \Delta g_i}, \quad (\text{B5})$$

$$e^{-\beta \Delta f_{ij}} = \sum_{c_i, c_j} \int dv_{ij} e^{-2\beta \Delta g_{ij}} M_{i \rightarrow j}(c_i, v_{i \rightarrow j}) M_{j \rightarrow i}(c_j, v_{j \rightarrow i}). \quad (\text{B6})$$

Here $v_{ij} = \{v_{i \rightarrow j}, v_{j \rightarrow i}\}$. Recall that

$$2\Delta g_i = \ln(2\pi v_i), \quad (\text{B7})$$

$$2\Delta g_{ij} = -\ln \{v_{i \rightarrow j} v_{j \rightarrow i} \det[B(ij)]\}, \quad (\text{B8})$$

where

$$\frac{1}{v_i} = (1 - c_i) + c_i A_{ii} - \sum_{k \neq i} c_i A_{ik}^2 c_k v_{k \rightarrow i}, \quad (\text{B9})$$

$$B(ij) = \begin{pmatrix} \frac{1}{v_{i \rightarrow j}} & c_i A_{ij} c_j \\ c_i A_{ij} c_j & \frac{1}{v_{j \rightarrow i}} \end{pmatrix}. \quad (\text{B10})$$

To compute the Δf_i we repeat the following steps for Δt_{avg} times:

- (i) select a member $(c_k^{a_k}, v_{k \rightarrow i}^{a_k})$ from $P_{k \rightarrow i}$ for $k \in \partial i$;
- (ii) compute $\Delta g_i(c_i)$ [Eq. (B7)] and $w_i(c_i) \equiv e^{-2\beta \Delta g_i(c_i)}$ for $c_i = 0, 1$ given the messages $(c_k^{a_k}, v_{k \rightarrow i}^{a_k})$.

From the above computation we obtain the averages $Z_i(c_i) \equiv \langle w_i(c_i) \rangle$, and $\Delta E_i(c_i) \equiv 2 \langle w_i(c_i) \Delta g_i(c_i) \rangle$. Then $\Delta f_i = -\ln[Z_i(0) + Z_i(1)]/\beta$, the probability of $c_i = 1$ is $p_i = Z_i(1)/[Z_i(0) + Z_i(1)]$, and we define $\Delta e_i = [\Delta E_i(0) + \Delta E_i(1)]/[Z_i(0) + Z_i(1)]$.

In the same way, to compute the Δf_{ij} we repeat the following steps for Δt_{avg} times:

- (i) select $(c_i^a, v_{i \rightarrow j}^a), (c_j^b, v_{j \rightarrow i}^b)$ from $P_{i \rightarrow j}$ and $P_{j \rightarrow i}$;
- (ii) compute Δg_{ij} [Eq. (B7)] and $w_{ij} \equiv e^{-2\beta \Delta g_{ij}}$ given the above messages;

From the above computation we obtain the averages $Z_{ij} \equiv \langle w_{ij} \rangle$ and $\Delta E_{ij} \equiv 2 \langle w_{ij} \Delta g_{ij} \rangle$. Then $\Delta f_{ij} = -\ln(Z_{ij})/\beta$ and we define $\Delta e_{ij} = \Delta E_{ij}/Z_{ij}$.

Finally, the free energy density f and energy density e are obtained

$$Nf = \sum_i \Delta f_i - \sum_{i < j} \Delta f_{ij} - N \ln(2\pi), \quad (\text{B11})$$

$$Ne = \sum_i \Delta e_i - \sum_{i < j} \Delta e_{ij} - N \ln(2\pi). \quad (\text{B12})$$

APPENDIX C: THE BETHE EQUATIONS FOR RANDOM REGULAR GRAPHS

Consider the Laplacian of a random regular graph of degree K . Here by symmetry all nodes are equivalent. Thus the BP equations for the cavity marginals $M_{i \rightarrow j}(c_i, v_{i \rightarrow j})$ of the Gaussian representation

$$M_{i \rightarrow j}(c_i, v_{i \rightarrow j}) \propto \prod_{k \in \partial i \setminus j} \left[\sum_{c_k} \int dv_{k \rightarrow i} M_{k \rightarrow i}(c_k, v_{k \rightarrow i}) \right] \times \mathbb{I}(v_{i \rightarrow j}) e^{-2\beta \Delta g_{i \rightarrow j}}, \quad (\text{C1})$$

reduce to a single equation for $M_{\rightarrow}(c, v)$ which does not depend on the indices i and $i \rightarrow j$. We take $M_{\rightarrow}(c, v) = (1 - p)\delta(v - 1)\delta_{c,0} + p\rho(v)\delta_{c,1}$ with p for the probability of presence of index i in the minor configuration. We recall that for $c = 0$ the variance is $v = 1$ independent of the states of the neighbors. The probability distribution of v for $c = 1$ is shown by $\rho(v)$. Then the BP equations for p and $\rho(v)$ read

$$p = \frac{z_1(K - 1)}{z_0(K - 1) + z_1(K - 1)}, \quad (\text{C2})$$

$$\rho(v) = \frac{z(v : K - 1)}{z_1(K - 1)}, \quad (\text{C3})$$

where $z_0(K - 1) = e^{-\beta \ln(2\pi)}$,

$$z_1(K - 1) = \sum_{l=0}^{K-1} C(l, K - 1) p^l (1 - p)^{K-1-l} \times \int \prod_{k=1}^l dv_k \rho(v_k) e^{-\beta \ln\left(\frac{2\pi}{K - \sum_k v_k}\right)}, \quad (\text{C4})$$

and

$$z(v : K - 1) = \sum_{l=0}^{K-1} C(l, K - 1) p^l (1 - p)^{K-1-l} \times \int \prod_{k=1}^l dv_k \rho(v_k) e^{-\beta \ln\left(\frac{2\pi}{K - \sum_k v_k}\right)} \times \delta\left(v - \frac{1}{K - \sum_k v_k}\right). \quad (\text{C5})$$

with $C(l, K) = K!/(l!(K - l)!)$. Again we solve these equations with a population dynamics algorithm as described in Appendix B. The only difference is that here we need only one population of variables (c^a, v^a) and the Boltzmann weights $e^{-2\beta \Delta g_{\rightarrow}}$.

After solving the above equations for p and $\rho(v)$, the free energy is given by

$$-\beta f_g = \ln[z_0(K) + z_1(K)] - \frac{K}{2} \ln(z_{00} + 2z_{01} + z_{11}), \quad (\text{C6})$$

where $z_0(K), z_1(K)$ are as before and

$$z_{00} = (1 - p)^2, \quad (\text{C7})$$

$$z_{01} = (1 - p)p, \quad (\text{C8})$$

$$z_{11} = p^2 \int dv dv' \rho(v) \rho(v') e^{\beta \ln(1 - vv')}. \quad (\text{C9})$$

Recall that

$$-\beta f = \frac{1}{N} \ln Z = -\beta [f_g - \ln(2\pi)]. \quad (\text{C10})$$

On the other hand, the average energy density is

$$e = -\frac{1}{N} \langle \ln \det A(\mathbf{c}) \rangle = 2 \langle g \rangle - \ln(2\pi). \quad (\text{C11})$$

Here

$$2 \langle g \rangle = \frac{g_0(K) + g_1(K)}{z_0(K) + z_1(K)} - \frac{K}{2} \frac{g_{11}}{z_{00} + 2z_{01} + z_{11}}, \quad (\text{C12})$$

with

$$g_0(K) = z_0(K) \ln(2\pi), \quad (\text{C13})$$

$$g_1(K) = \sum_{l=0}^K C(l, K) p^l (1 - p)^{K-l} \times \int \prod_{k=1}^l dv_k \rho(v_k) \ln\left(\frac{2\pi}{K - \sum_k v_k}\right) e^{-\beta \ln\left(\frac{2\pi}{K - \sum_k v_k}\right)}, \quad (\text{C14})$$

$$g_{11} = -p^2 \int dv dv' \rho(v) \rho(v') \ln(1 - vv') e^{\beta \ln(1 - vv')}. \quad (\text{C15})$$

Finally, by a Legendre transformation, the entropy is obtained

$$s = \beta(e - f). \quad (\text{C16})$$

A parametric plot of this quantity vs e gives the entropy spectrum $s(e)$.

APPENDIX D: MINORS OF HOMOGENEOUS COMPLETE GRAPHS: A MEAN-FIELD MODEL

Consider a symmetric matrix \mathbf{A} associated to a complete graph of size N with elements $A_{ij} = \Lambda \delta_{ij} - \Gamma(1 - \delta_{ij})$. We assume that $\Lambda > 0$ and $\Lambda \geq N\Gamma > 0$ to have a positive and diagonally dominant matrix. A minor configuration of size l here is represented by a $l \times l$ matrix with diagonal elements Λ and off-diagonal elements $-\Gamma$. Such a matrix has one eigenvalue $\Lambda - (l - 1)\Gamma$ and $l - 1$ eigenvalues $\Lambda + \Gamma$. Thus the determinant (for $0 < l \leq N$) is given by

$$\det[\mathbf{A}(\mathbf{c})] = [\Lambda - (l - 1)\Gamma][\Lambda + \Gamma]^{l-1}, \quad (\text{D1})$$

and the energy is

$$\begin{aligned} E_l &= -\ln\{\det[\mathbf{A}(\mathbf{c})]\} \\ &= -(l - 1)\ln(\Lambda + \Gamma) - \ln[\Lambda - (l - 1)\Gamma]. \end{aligned} \quad (\text{D2})$$

The partition function is

$$Z = \sum_{\mathbf{c}} e^{\beta \ln \det[\mathbf{A}(\mathbf{c})]} = 1 + \sum_{l=1}^{N-1} \Omega(l, N) e^{-\beta E_l}, \quad (\text{D3})$$

where $\Omega(l, N) = N!/[l!(N - l)!]$. Or, in terms of the energy and entropy densities

$$Z \simeq N \int_0^1 dx e^{N[s(x) - \beta e(x)]} = N \int_0^1 dx e^{-N\beta f(x)}, \quad (\text{D4})$$

where $x = l/N$ is the number density of the present indices, and

$$s(x) = -x \ln(x) - (1 - x) \ln(1 - x), \quad (\text{D5})$$

$$e(x) = -x \ln(\Lambda + \Gamma) - \frac{1}{N} \ln(\Lambda - N\Gamma x), \quad (\text{D6})$$

$$f(x) = e(x) - \frac{1}{\beta} s(x). \quad (\text{D7})$$

The thermodynamic behavior depends on the scaling of the matrix elements Λ , Γ , with N .

1. Scaling limit I: $\Lambda = N - 1$, $\Gamma = 1$

This is Laplacian of a complete graph. Here we use the scaling $\beta \rightarrow \beta / \ln(N)$ and replace $e(x)/\ln(N)$ with $e(x)$ as $N \rightarrow \infty$. Ignoring the subextensive terms in the energy density, we get

$$e(x) = -x. \quad (\text{D8})$$

The minimum of the free-energy function

$$f(x) = e(x) - \frac{1}{\beta} s(x), \quad (\text{D9})$$

at x^* determines the equilibrium values of the main quantities $e(x^*)$, $s(x^*)$, $f(x^*)$. Recall that x^* is the presence probability p we used in the main text. Here the free energy is always minimized for a number density $\frac{1}{2} \leq x^* < 1$, which is the sole solution to

$$\beta = \ln\left(\frac{x^*}{1 - x^*}\right). \quad (\text{D10})$$

This system does not display a phase transition. By increasing β from zero to infinity, the density of relevant minors x^* increases smoothly from $1/2$ to 1 . And the minimum of free-energy function remains an extremum [$\frac{df}{dx}(x^*) = 0$] as $\beta \rightarrow \infty$. There are N ground-state configurations of energy $E_{N-1} = -(N - 2)\ln N$ where only one of the nodes is not present in the minor configuration. So the Hamming distance between two ground states is 2.

2. Scaling limit II: $\Lambda = \text{finite}$, $\Gamma = \gamma/N$

This can be considered as a mean-field model of random regular graphs of degree $K = \Lambda$. Here the energy density as a function of x in the thermodynamic limit reads

$$e(x) = -x \ln(\Lambda). \quad (\text{D11})$$

Note that value of Γ is not relevant in both the above scaling limits. Again, we do not observe a finite-temperature phase transition; the minimum of free energy at x^* changes smoothly as β approaches infinity. That is, equation

$$\frac{df}{dx}(x^*) = -\ln \Lambda + \frac{1}{\beta} \ln\left(\frac{x^*}{1 - x^*}\right) = 0 \quad (\text{D12})$$

always has a unique solution $0 < x^* < 1$ due to the monotonic behaviors of the two parts of the equation

$$\beta \ln \Lambda = \ln\left(\frac{x^*}{1 - x^*}\right). \quad (\text{D13})$$

However, depending on the value of Λ the solution x^* approaches 0 or 1 as $\beta \rightarrow \infty$ for $\Lambda < 1$ and $\Lambda > 1$, respectively. For $\Lambda = 1$ the solution is always $x^* = \frac{1}{2}$. The ground states have a trivial structure: When $\Lambda < 1$ the minimum energy is obtained by N minor configurations with a single present index, putting aside the all-zero configuration. When $\Lambda > 1$ the minimum energy is obtained by N minor configurations with a single absent index. The case $\Lambda = 1$ has the maximal ground-state degeneracy where all minor configurations are equivalent in energy. Note that in all the above scenarios the value of entropy at zero-temperature coincides with the limit $\beta \rightarrow \infty$ of the entropy function. It means that the entropy changes continuously as the temperature increases.

- [1] M. N. Najafi, A. Ramezanpour, and M. A. Rajabpour, A field theory representation of sum of powers of principal minors and physical applications, [arXiv:2403.09874](https://arxiv.org/abs/2403.09874).
- [2] A. Kulesza and B. Taskar, Determinantal point processes for machine learning, *Found. Trends Mach. Learn.* **5**, 123 (2012).
- [3] J. Gillenwater, Approximate inference for determinantal point-processes, Ph.D. thesis, University of Pennsylvania, 2014.
- [4] K. Najafi and M. A. Rajabpour, Formation probabilities and Shannon information and their time evolution after quantum quench in the transverse-field XY chain, *Phys. Rev. B* **93**, 125139 (2016).
- [5] L. Gurvits, On the complexity of mixed discriminants and related problems, in *Mathematical Foundations of Computer Science*, Vol. 5 (Springer, Berlin, 2005), pp. 447–458.
- [6] N. Ohsaka and T. Matsuoka, On the (in) tractability of computing normalizing constants for the product of determinantal point processes, in *Proceedings of the International Conference on Machine Learning* (ACM, New York, 2020).
- [7] J. A. Bondy and U. S. Murty, *Graph Theory with Applications* (Macmillan, London, 1976).
- [8] S. Chaiken, A combinatorial proof of the all minors matrix tree theorem, *SIAM. J. Algebr. Discr. Methods* **3**, 319 (1982).
- [9] N. Anari, Y. U. Liu, and T. D. Vuong, Optimal sublinear sampling of spanning trees and determinantal point processes via average-case entropic independence, in *Proceedings of the IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS'22)* (IEEE, Los Alamitos, CA, 2022), pp. 123–134.
- [10] M. Mézard and G. Parisi, The Bethe lattice spin glass revisited, *Eur. Phys. J. B* **20**, 217 (2001).
- [11] M. Mézard and G. Parisi, The cavity method at zero temperature, *J. Stat. Phys.* **111**, 1 (2003).
- [12] J. S. Yedidia, W. T. Freeman, and Y. Weiss, Understanding belief propagation and its generalizations, in *Exploring Artificial Intelligence in the New Millennium* (Morgan Kaufmann, Burlington, MA, 2003).
- [13] M. Mezard and A. Montanari, *Information, Physics, and Computation* (Oxford University Press, Oxford, 2009).
- [14] G. Semerjian and L. F. Cugliandolo, Sparse random matrices: The eigenvalue spectrum revisited, *J. Phys. A: Math. Gen.* **35**, 4837 (2002).
- [15] T. Rogers, I. P. Castillo, R. Kühn, and K. Takeda, Cavity approach to the spectral density of sparse symmetric random matrices, *Phys. Rev. E* **78**, 031116 (2008).
- [16] B. Huang and T. Jebara, Approximating the permanent with belief propagation, [arXiv:0908.1769](https://arxiv.org/abs/0908.1769).
- [17] T. Rogers and I. P. Castillo, Cavity approach to the spectral density of non-Hermitian sparse matrices, *Phys. Rev. E* **79**, 012101 (2009).
- [18] F. L. Metz, I. Neri, and D. Bollé, Localization transition in symmetric random matrices, *Phys. Rev. E* **82**, 031135 (2010).
- [19] Y. Kabashima, H. Takahashi, and O. Watanabe, Cavity approach to the first eigenvalue problem in a family of symmetric random sparse matrices, *J. Phys.: Conf. Ser.* **233**, 012001 (2010).
- [20] G. Biroli, G. Semerjian, and M. Tarzia, Anderson model on Bethe lattices: Density of states, localization properties and isolated eigenvalue, *Prog. Theor. Phys. Suppl.* **184**, 187 (2010).
- [21] C. Baldassi, A. Braunstein, A. Ramezanpour, and R. Zecchina, Statistical physics and network optimization problems, in *Mathematical Foundations of Complex Networked Information Systems*, Lecture Notes in Mathematics, Vol. 2141, edited by F. Fagnani, S. Fosson, and C. Ravazzi (Springer, Cham, 2015).
- [22] M. Dereziński, D. Calandriello, and M. Valko, Exact sampling of determinantal point processes with sublinear time preprocessing, *Advances in Neural Information Processing Systems*, Vol. 32, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., 2019).
- [23] J. Poulson, High-performance sampling of generic determinantal point processes, *Phil. Trans. R. Soc. A.* **378**, 20190059 (2020).
- [24] O. Shental, P. H. Siegel, J. K. Wolf, D. Bickson, and D. Dolev, Gaussian belief propagation solver for systems of linear equations, in *Proceedings of the IEEE International Symposium on Information Theory* (IEEE, Los Alamitos, CA, 2008), pp. 1863–1867.
- [25] Q. Su and Y. C. Wu, Convergence analysis of the variance in Gaussian belief propagation, *IEEE Trans. Sign. Process.* **62**, 5119 (2014).
- [26] L. Zdeborova, Statistical physics of hard optimization problems, Ph.D. thesis, Université Paris Sud-Paris XI, 2008.
- [27] A. Braunstein and R. Zecchina, Learning by message passing in networks of discrete synapses, *Phys. Rev. Lett.* **96**, 030201 (2006).