The Annals of Applied Probability
2018, Vol. 28, No. 3, 1536–1572
https://doi.org/10.1214/17-AAP1337

MAX κ-CUT AND THE INHOMOGENEOUS POTTS SPIN GLASS

BY AUKOSH JAGANNATH[*,1], JUSTIN KO[*,2] AND SUBHABRATA SEN[†]

University of Toronto[*] and Stanford University[†]

We study the asymptotic behavior of the Max κ-cut on a family of sparse, inhomogeneous random graphs. In the large degree limit, the leading term is a variational problem, involving the ground state of a constrained inhomogeneous Potts spin glass. We derive a Parisi-type formula for the free energy of this model, with possible constraints on the proportions, and derive the limiting ground state energy by a suitable zero temperature limit.

1. Introduction. Networks arise in various applications in economics, engineering and social sciences. In a typical social science application, the vertices of the network represent individuals, while their relationships are represented by the edges. The study of structural properties of these networks, and algorithms to find these structures are extremely important in this context. Various random graph models have been introduced to study such real-life networks (see, e.g., [25, 26])—and questions about networks translate directly into questions about random graphs under this approach. Graph partition problems are natural class of algorithmic questions which arise in this context. In these problems, the goal is to find a partition of the vertex set that maximizes some objective function, typically given by a function of the edges. Graph partition problems are of interest in applications as diverse as community detection [16, 46] and VLSI design [28]. In this paper, we focus on the Max κ-cut, an important example in this class.

For any graph $G = (V, E)$, the Max κ-cut problem (henceforth denoted as MaxCut_κ) seeks to divide the vertices, V, into κ (not necessarily equal) parts such that the number of edges between distinct parts is maximized. For $\kappa = 2$, this reduces to the well-known MAXCUT problem (see [45] for a survey of the MAXCUT problem). From the point of view of complexity theory, these questions are usually NP hard in the worst case. This motivates a study of average case complexity, often formalized by studying this problem on random graph instances. As a first attempt, one seeks to determine the typical behavior of these quantities on a random graph—this provides a valuable benchmark for comparing the performance of specific algorithms on random instances.

Such questions have been studied in classical settings, such as the Erdős–Rényi random graph and random regular graph ensembles. The key insight in this setting

Received March 2017; revised July 2017.
[1]Supported in part by NSF Grant OISE-1604232.
[2]Supported in part by NSERC Grant RGPIN-2015-04637.
MSC2010 subject classifications. 90C27, 82D30, 82B44, 60K35, 60G15, 90C06.
Key words and phrases. MAX CUT, spin glasses, Potts model, inhomogeneous random graphs.

is a connection between statistical physics and the MaxCut_κ problem, enunciated as follows. Any κ-cut can be represented by an assignment of spins $\sigma \in [\kappa]^N$ to the vertices of the graph. Further, setting $A = (A_{ij})$ to be the adjacency matrix of the random graph G, we have

$$\frac{\mathsf{MaxCut}_\kappa(G)}{N} = \frac{1}{2N} \max_{\sigma \in [\kappa]^N} \sum_{i,j=1}^{N} A_{i,j}\, \mathbb{1}(\sigma_i \neq \sigma_j). \tag{1.1}$$

In statistical physics parlance, (1.1) establishes a direct relation between the MaxCut_κ and the ground state of the antiferromagnetic Potts model on the graph. Connections between graph partition problems and statistical physics are, by now, classical [22]. For a textbook introduction to the physical perspective on these questions, we refer the reader to [33, 34].

Physicists predict that the antiferromagnetic nature of the Max κ-cut should force the quantity to behave as the ground state of a disordered spin glass, and its behavior in graphs with large degrees should be well approximated by properties of ground states in mean field spin glasses. For the MAXCUT problem on sparse Erdős–Rényi and random regular graphs, this idea was partially formalized in [17] and [48]. The authors of [17] deduced that for $G \sim G(N, \frac{c}{N})$, as $N \to \infty$, we have

$$\frac{\mathsf{MAXCUT}(G)}{N} = \frac{c}{4} + \mathsf{P}_* \sqrt{\frac{c}{4}} + o_c(\sqrt{c}). \tag{1.2}$$

Here, P_* is the limiting ground state energy of the Sherrington–Kirkpatrick model [40, 50]. Here and henceforth in the paper, we say that a sequence of random variables, (X_N), satisfies $X_N = o_c(\sqrt{c})$ if and only if there is a deterministic function $g(c) = o(\sqrt{c})$ such that $\mathbb{P}[|X_N| \leq g(c)] \to 1$ as $N \to \infty$. The first term in the right-hand side comes from the standard observation that a typical partition of the vertices will contribute $Nc/4$ edges to the cut in expectation. The second term is the leading order correction, and specifies the difference in size between a typical cut and the MAXCUT. An analogous formula for the MaxCut_κ on sparse Erdős–Rényi and random regular graphs was derived in [48].

For practical applications, it is thus of natural interest to determine the typical value of these quantities on random graph ensembles that capture natural properties of realistic networks. In practice, networks are typically observed to be sparse and "inhomogeneous" [1, 3, 18, 31]. The simplest random graph models, such as Erdős–Rényi and random regular graphs, lead to instances where the degree distributions are relatively concentrated—a feature seldom observed in real networks. To address this issue, a plethora of models have been introduced, which faithfully capture some of the observed characteristics of real networks. In this paper, we seek to establish formulae similar to (1.2) for a general family of graph models using the framework of [17] and [48]. Our approach leads naturally to the study of an inhomogeneous Potts spin glass model, which has yet to be studied rigorously in the mathematical literature.

Let us first explain the class of random graph models that we study. Our framework will be similar to the one introduced by Söderberg [51] and adopted by Bollobás, Janson and Riordan [8]. Furthermore, this model has natural connections to the theory of graphons for dense sequences of random graphs [9, 10]. Consider a symmetric kernel $K : [0,1]^2 \to [0,\infty)$. We will assume throughout that $K \in L^1([0,1]^2, \mathrm{d}x)$. Given such a kernel, consider the following model for a sequence of inhomogeneous random graphs $\mathcal{G}_N = (V_N, E_N)$. For all $N \geq 1$, we let the vertex set be $V_N = [N]$. The edges will then be added independently with probability

$$\mathbb{P}\big[\{i,j\} \in E_N\big] = \min\left\{c\frac{\tilde{K}(i,j)}{N}, 1\right\}, \tag{1.3}$$

where $\tilde{K}$ is the average of K within blocks,

$$\tilde{K}_N(i,j) = N^2 \int_{[\frac{i-1}{N},\frac{i}{N}]\times[\frac{j-1}{N},\frac{j}{N}]} K(x,y)\,\mathrm{d}x\,\mathrm{d}y. \tag{1.4}$$

This specifies the random graph model. The parameter c controls the degree of the vertices. We note that this model is more restricted compared to that of Bollobás, Janson and Riordan [8]. In the notation of [8], we restrict ourselves to the case where the ground space $\mathcal{S} = [0,1]$ and the measure μ is the Lebesgue measure. Further, the model introduced in [8] is governed by the value of the kernel K on a set of measure zero. Here, we average over small partitions of the kernel, so that we may avoid technical subtleties on sets of measure zero. For reasonable kernels, such as continuous ones, this distinction will be negligible.

To state our main result in a concrete setting, let us first work in the case when K is block constant. That is, we assume that there are numbers

$$0 = t_0 < t_1 < t_2 < \cdots < t_{M-1} < 1 = t_M$$

such that K is constant on each square of the form $[t_{j-1}, t_j] \times [t_{k-1}, t_k]$ for $0 \leq j,k \leq M$. Further, we set $\rho^s = t_s - t_{s-1}$, $s = 1,\ldots,M$. For any such block kernel K, let $\mathbf{K}$ denote the $M \times M$ matrix of the values of the kernel on the blocks. For technical reasons, we will work with block kernels such that the matrix $\mathbf{K}$ is positive definite. Finally, we note that any block constant kernel with finitely many blocks is almost surely bounded. By a standard application of the Efron–Stein inequality [11], it suffices to study the asymptotic behavior of $\mathbb{E}[\mathsf{MaxCut}_\kappa(G_N)]/N$.

To analyze this quantity, we introduce the following notation. For any finite set $\mathscr{S}$, let $\mathscr{D}$ be the space of proportions, given by

$$\mathscr{D} = \left\{(d_1^s,\ldots,d_\kappa^s)_{s\in\mathscr{S}} \Big| d_k^s \geq 0, \sum_{k=1}^{\kappa} d_k^s = 1 \ \forall s \in \mathscr{S}\right\}. \tag{1.5}$$

In our setting, $\mathscr{S} = [M]$. Any $d \in \mathscr{D}$ can be expressed as $d = (d^s)_{s\in\mathscr{S}}$, where $(d^s)_{s\in\mathscr{S}}$ is a collection of probability measures on $[\kappa]$. The distribution d^s governs

the proportion of vertices in block s which belong to the partition i, $1 \le i \le \kappa$. We will refer to the elements $d \in \mathscr{D}$ as proportions. The following theorem characterizes the value of the MaxCut_κ problem for inhomogeneous graphs G_N with block constant kernels and large degrees, up to corrections which are $o(\sqrt{c})$.

THEOREM 1.1. *We have, as* $N \to \infty$,

$$\lim_{N\to\infty} \mathbb{E}\left[\frac{\mathsf{MaxCut}_\kappa(G_N)}{N}\right] = \sup_{d\in\mathscr{D}}\left[\frac{c}{2}\sum_{s,t=1}^{M} \mathbf{K}(s,t)\rho^s\rho^t\left(1-\langle d^s, d^t\rangle\right) + \frac{\sqrt{c}}{2}\mathcal{P}(d)\right] + o(\sqrt{c}).$$

REMARK. We take this opportunity to comment on the positive definite assumption on the matrix $\mathbf{K}$. One prominent example where $\mathbf{K}$ is not positive definite is the random bipartite graph, where the kernel consists of two off-diagonal blocks. However, note that for $\kappa = 2$, the behavior of the MAXCUT on this graph is trivial, and very different from that established in Theorem 1.1.

Note that the leading term in Theorem 1.1 is a variational problem involving the empirical distribution of spins within each block. This variational problem has two terms: the first term governs the expected cut-size, while the second term, of order $\sqrt{c}$, governs the extra contribution which is attained by optimization. It remains to introduce $\mathcal{P}(d)$. It turns out that $\mathcal{P}(d)$ is the limiting ground state energy of the inhomogeneous Potts spin glass model, subject to constraints on the composition of spins within each block. We introduce this model in the rest of the section, and define the constant $\mathcal{P}(d)$ rigorously using a Parisi-type formula for the limiting free energy.

1.1. *The inhomogeneous Potts model.* We consider a natural generalization of the Potts spin glass model that allows for inhomogeneous coupling interactions between species. The configuration space for this model is $\Sigma_N = [\kappa]^N$ for some $\kappa \ge 2$. Let $\mathscr{S}$ be the finite set in (1.5), each element of which is called a species. For each N, we are given a partition of $[N]$ indexed by the species as

$$[N] = \bigcup_{s\in\mathscr{S}} I_s.$$

We say that i belongs to species s if $i \in I_s$. Conversely, we denote by $s(i)$ the species to which i belongs. Let $N_s = |I_s|$. Naturally, this quantity varies in N. To obtain a reasonable limiting structure, we assume that the proportions converge:

$$\rho_N^s = \frac{N_s}{N} \to \rho^s \in (0,1). \tag{1.6}$$

The Hamiltonian for this model, H_N, is the centered Gaussian process

$$H_N(\sigma) = \frac{1}{\sqrt{N}} \sum_{i,j=1}^{N} g_{i,j}\, \mathbb{1}(\sigma_i = \sigma_j), \tag{1.7}$$

where $g_{i,j}$ are independent, centered Gaussian random variables with covariance

$$\mathbb{E} g_{i,j}^2 = \Delta_{s,t}^2, \qquad s,t \in \mathscr{S}, i \in I_s, j \in I_t. \tag{1.8}$$

We assume, following [7, 44], that the matrix $\Delta := \Delta_{s,t}^2$ is symmetric and positive definite in s and t. Observe that if we define, for $\sigma^1, \sigma^2 \in \Sigma_N$,

$$R_{1,2}^s(k,k') = \frac{1}{N_s} \sum_{i \in I_s} \mathbb{1}(\sigma_i^1 = k)\, \mathbb{1}(\sigma_i^2 = k') \tag{1.9}$$

and define the $\kappa \times \kappa$ species overlap matrix

$$R_{1,2}^s = \big(R_{1,2}^s(k,k')\big)_{k,k' \le \kappa} \tag{1.10}$$

then H_N has covariance

$$\mathrm{Cov}(\sigma^1, \sigma^2) = N \sum_{s,t \in \mathscr{S}} \Delta_{s,t}^2 \rho_N^s \rho_N^t \big(R_{1,2}^s, R_{1,2}^t\big), \tag{1.11}$$

where $(\cdot,\cdot)$ denotes the Frobenius (or Hilbert–Schmidt) inner product.

This model is an inhomogeneous extension of the Potts spin glass model, which has been studied extensively in the physics literature [13, 20, 21, 36], and analyzed rigorously in [42, 43]. In particular, we break the symmetry between sites. When $\kappa = 2$, this type of inhomogeneity was introduced in an equivalent form by Barra, Contucci, Mignone and Tantari in [7] where a Guerra-type [24] upper bound for the free energy was obtained. The matching lower bound was obtained by Panchenko in [44].

Our goal is to compute constrained free energies of the type

$$F_N(A) = \frac{1}{N} \mathbb{E} \log \sum_{\sigma \in A} e^{H_N(\sigma)} \tag{1.12}$$

for a specific choices of A. Recall the space of proportions $\mathscr{D}$ defined in (1.5). Given a $d \in \mathscr{D}$, we have the associated *constrained state space*

$$\Sigma_N^\varepsilon(d) = \Big\{ \sigma \in \Sigma_N \Big| \sum_{i \in I_s} \frac{\mathbb{1}(\sigma_i = k)}{N_s} \in \big[d_k^s - \varepsilon, d_k^s + \varepsilon\big] \Big\}. \tag{1.13}$$

We will use the notation $\Sigma_N(d) := \Sigma_N^0(d)$ to denote the constrained state space where the proportions of spins within species are exactly equal to the proportion d. Let $\mathscr{D}_N \subset \mathscr{D}$ be the space of feasible constraints for configuration spaces of N coordinates

$$\mathscr{D}_N = \big\{ d \in \mathscr{D} \big| \Sigma_N(d) \neq \varnothing \big\}. \tag{1.14}$$

We are particularly interested in computing constrained free energies as in (1.12) with $A = \Sigma(d)$. Note that since the space of feasible configurations is at most polynomial growth, classical concentration arguments show that the free energy of the whole system, $F_N(\Sigma_N)$, is asymptotically given by the maximum of $F_N(\Sigma_N(d))$ over $\mathscr{D}$.

As in the Potts spin glass model, among others, the overlap (1.10) will play a key role. In particular, we will find that the array of overlaps from i.i.d. draws of configurations from (a perturbation of) the Gibbs measure, will be determined by a path $\pi : [0,1] \to \Gamma_\kappa^{|\mathscr{S}|}$. Here, Γ_κ is the space of $\kappa \times \kappa$ positive definite matrices and π is effectively the family of quantile transform of the limiting law of the overlap of two independent copies σ^1, σ^2 from the Gibbs measure.

We now turn to the main result. To this end, we denote the space of *left-continuous monotone functions* on Γ_κ as

$$\Pi = \{\pi : [0,1] \to \Gamma_\kappa : \pi \text{ is left-continuous, } \pi(x) \le \pi(x') \text{ for } x \le x'\}, \tag{1.15}$$

where $\pi(x) \le \pi(x')$ means that $\pi(x') - \pi(x) \in \Gamma_\kappa$. Similarly, we let

$$\mathbf{\Pi} = \{(\pi^s)_{s\in\mathscr{S}} : [0,1] \to \Gamma_\kappa^{|\mathscr{S}|} | \pi^s \in \Pi\}. \tag{1.16}$$

We also have the following metric on $\mathbf{\Pi}$

$$\Delta(\boldsymbol{\pi}, \tilde{\boldsymbol{\pi}}) = \int_0^1 \max_{s\in\mathscr{S}} \|\pi^s(x) - \tilde{\pi}^s(x)\|_1 \, \mathrm{d}x. \tag{1.17}$$

For $d \in \mathscr{D}$ and $r > 0$, we define the following sequences of parameters. Let $(x_i)_{i=1}^r$ be a strictly increasing sequence of numbers

$$0 = x_{-1} < x_0 < \cdots < x_r = 1. \tag{1.18}$$

For each species, let $(Q_i^s)_{i=1}^r$ be an increasing sequence of $\kappa \times \kappa$ positive semidefinite matrices

$$0 = Q_0^s \le Q_1^s \le \cdots \le Q_r^s = \operatorname{diag}(d_1^s, \ldots, d_\kappa^s) = D^s. \tag{1.19}$$

Given these sequences, for each species we can define the Gaussian vector $(z_p^s) \in \mathbb{R}^\kappa$ such that

$$\mathbb{E} z_p^s (z_q^s)^T = 2\delta_{p,q} \Big(\sum_{t\in\mathscr{S}} \Delta_{s,t}^2 \rho^t Q_p^t - \sum_{t\in\mathscr{S}} \Delta_{s,t}^2 \rho^t Q_{p-1}^t \Big). \tag{1.20}$$

The nonrandom value X_0^s is defined recursively as

$$\begin{aligned} X_r^s &= \log \sum_{k\le\kappa} \exp\Big(\sum_{1\le p\le r} z_p^s(k) + \lambda_k^s \Big), \\ X_k^s &= \frac{1}{x_k} \log \mathbb{E}_k \exp(x_k X_{k+1}^s) \qquad \text{for } 0 \le k < r, \end{aligned} \tag{1.21}$$

where $\mathbb{E}_k$ denotes the expectation with respect to only z_{k+1}. Finally, define the functional

$$\begin{aligned}
&\mathscr{P}\big(r, x, d, (\lambda^s, Q^s)_{s\in\mathscr{S}}\big) \\
&\quad = \sum_{s\in\mathscr{S}} \rho^s X_0^s - \sum_{s\in\mathscr{S}} \sum_{k\le\kappa} \rho^s \lambda_k^s d_k^s \\
&\qquad - \frac{1}{2}\sum_{\ell=0}^{r-1} x_\ell \sum_{s,t\in\mathscr{S}} \Delta_{st}^2 \rho^s \rho^t \big((Q_{\ell+1}^s, Q_{\ell+1}^t) - (Q_\ell^s, Q_\ell^t)\big).
\end{aligned} \tag{1.22}$$

The next result characterizes the limiting free energy in these models.

THEOREM 1.2. *For any $\kappa \ge 2$, set of species $\mathscr{S}$, and sequences $\rho_N^s \to \rho^s$ we have the following*:

1. *For any $d \in \mathscr{D}$ and $\varepsilon_N \to 0$ sufficiently slowly, the limit of the constrained free energy is given by*

$$\lim_{N\to\infty} F_N\big(\Sigma_N^{\varepsilon_N}(d)\big) = \inf_{x,r,(\lambda^s,Q^s)_{s\in\mathscr{S}}} \mathscr{P}\big(r, x, d, (\lambda^s, Q^s)_{s\in\mathscr{S}}\big). \tag{1.23}$$

2. *The limit of the unconstrained free energy is given by*

$$\lim_{N\to\infty} F_N(\Sigma_N) = \sup_{d\in\mathscr{D}} \inf_{x,r,(\lambda^s,Q^s)_{s\in\mathscr{S}}} \mathscr{P}\big(r, x, d, (\lambda^s, Q^s)_{s\in\mathscr{S}}\big). \tag{1.24}$$

Before moving forward, we explain the nontrivial obstacles encountered in the inhomogeneous Potts model. First, the symmetry between sites is broken. As a result, interactions both within the species and between the species must be considered. Second, the natural overlap structure are matrices (1.10) and are a priori not necessarily positive definite in the limit.

Each of these issues has been studied in the past in [44] and [43]. The synchronization property in [44] connected the species overlaps with the average of the overlaps over the entire system. Similarly, the synchronization property in [43] proved the overlaps concentrated on the space of Gram matrices in the limit. Another consequence of this result implied the overlap matrices could also be recovered from the trace of the matrix. At the heart of both of these synchronization arguments were generalized Ghirlanda–Guerra-type identities that implied an ultrametric underlying structure of the overlaps [39]. The synchronization combines local and global ultrametric properties forcing a rigid distribution structure.

We prove an analogue of the Ghirlanda–Guerra identities, which combines those in the inhomogeneous SK and Potts models. This results in a simultaneous synchronization mechanism of the overlap matrices both within and between species. In our setting, we will be able recover the structure of the overlap matrices $R_{\ell,\ell'}^s$ deterministically from the average of the traces of overlaps

$$\sum_{s\in\mathscr{S}} \rho^s \operatorname{tr}(R_{\ell,\ell'}^s).$$

These techniques will reduce the problem to a familiar setting, allowing us to derive a formula for the free energy using the Guerra upper bound [24] and the Aizenman–Sims–Starr scheme [2] using the characterization method introduced in [4] and formalized further in [38, 41]. The resulting functional order parameter is a vector of monotone paths of $\kappa \times \kappa$ matrices.

In applications, we will be interested, not only in the free energy, but also the maximum of (1.7) subject to the same constraints. To understand the connection between the two, we recall the classical fact from statistical mechanics that the ground state energy of a system can be obtained as the "zero temperature limit" of the free energy of this system—a limit usually referred to as annealing [30]. In this case, we take the beaten path, and define free energies of the form

$$F_N^\beta(A) = \frac{1}{N}\mathbb{E}\log \sum_{\sigma\in A} e^{\beta H_N(\sigma)}, \tag{1.25}$$

where H_N is as in (1.7) for some fixed Δ. Here, β is the inverse temperature, and the zero temperature limit corresponds to sending $\beta \to \infty$. We note that the free energy (1.25) corresponds to the Hamiltonian $H_N(\sigma)$ defined as in (1.7), with respect to $\Delta^\beta = \beta^2\Delta$ instead of Δ. Therefore, a straightforward modification of Theorem 1.2 implies that for any $d \in \mathscr{D}$,

$$\lim_{N\to\infty} F_N^\beta\big(\Sigma_N^{\varepsilon_N}(d)\big) = \inf_{x,r,(\lambda^s,Q^s)_{s\in\mathscr{S}}} \mathscr{P}_\beta\big(r,x,d,(\lambda^s,Q^s)_{s\in\mathscr{S}}\big). \tag{1.26}$$

The functional $\mathscr{P}_\beta(r,x,d,(\lambda^s,Q^s)_{s\in\mathscr{S}})$ is identical to (1.22) with Δ^β in place of Δ. As a corollary to the theorem above, we obtain an expression for the limiting constrained ground state energies.

COROLLARY 1.3. *For any configuration $d \in \mathscr{D}$, we have, as $N \to \infty$, for some sequence $\varepsilon_N \to 0$ sufficiently slowly,*

$$\begin{aligned}\lim_{N\to\infty}\frac{1}{N}\mathbb{E}\Big[\max_{\sigma\in\Sigma_N^{\varepsilon_N}(d)} H(\sigma)\Big] &= \lim_{\beta\to\infty}\frac{1}{\beta}\inf_{x,r,(\lambda^s,Q^s)_{s\in\mathscr{S}}} \mathscr{P}_\beta\big(r,x,d,(\lambda^s,Q^s)_{s\in\mathscr{S}}\big)\\ &:= \mathcal{P}(d).\end{aligned}$$

Note that $\mathcal{P}(d)$ is the relevant constant for the MaxCut_κ in Theorem 1.1. Exact variational formulas for ground state energies like $\mathcal{P}(d)$ have been obtained recently in several models [5, 6, 14, 27]. However, in this setting this remains an interesting question.

1.2. *Applications.* We return to study of graph partitioning problems in this section and study some examples. We first note that given a general kernel $K \in L^1$, to determine the MaxCut_κ up to $o(\sqrt{c})$ corrections, it is enough to restrict ourselves

to block-constant kernels. To this end, given a kernel K and $M \geq 1$, we construct the kernel K_1 by "coarsening" the kernel K,

$$K_1(x,y) = M^2 \sum_{i,j=1}^{M} \mathbb{1}\left(x \in \left[\frac{i-1}{M}, \frac{i}{M}\right], y \in \left[\frac{j-1}{M}, \frac{j}{M}\right]\right) \times \int_{[\frac{i-1}{M}, \frac{i}{M}]\times[\frac{j-1}{M}, \frac{j}{M}]} K(s,t)\,\mathrm{d}s\,\mathrm{d}t. \tag{1.27}$$

Let $\tilde{G}_N$ denote the sequence of graphs formed from the kernel K_1 using (1.3).

Then we have the following.

LEMMA 1.4. *For any kernel K and c, $0 < \delta < 1/2$, we can choose $M := M(c)$ such that for all N sufficiently large,*

$$\left|\mathbb{E}\left[\frac{\mathsf{MaxCut}_\kappa(G_N)}{N}\right] - \mathbb{E}\left[\frac{\mathsf{MaxCut}_\kappa(\tilde{G}_N)}{N}\right]\right| \leq c^{1/2-\delta}.$$

Any kernel K is naturally associated with the integral operator $T_K(f)(x) := \int K(x,y)f(y)\,\mathrm{d}y$. We assume that for $f \in L^\infty([0,1])$, $\iint f(x)f(y)K(x,y)\,\mathrm{d}x\,\mathrm{d}y \geq 0$. We note that in this case, the corresponding discretized kernel K_1, described in Lemma 1.4, inherits the positive definite character. Therefore, $\mathsf{MaxCut}_\kappa(\tilde{G}_N)$ can be determined by an application of Theorem 1.2. We now turn to some examples to which our results apply.

EXAMPLE 1 (Finite species block model). The first example concerns the simple case when the kernel K has an explicit block structure. This model has been proposed and studied intensely by Söderberg [51] and Bollobás, Janson and Riordan [8]. These models have also been studied as "stochastic block models" in statistics, machine learning and theoretical computer science in connection to the community detection problem [16, 32, 35]. Our results apply directly to this model in case the kernel is positive semidefinite.

EXAMPLE 2 (Rank 1 model). The next example concerns the Rank 1 model for random graphs. In this model, we have a function $\psi : [0,1] \to \mathbb{R}^+$ such that $K(x,y) = \psi(x)\psi(y)$. $\psi(x)$ governs the "activity" of the vertex and the probability of a connection is determined by the product of the activities of the two vertices. This model has been extensively studied; see, for example, [12, 15, 37]. Prominent features of interest include the existence and size of a giant component, the degree distribution, the typical distances between the vertices, etc. We refer to [8], Section 16, for an extensive survey of the related family of models and connections to earlier results. The kernel is positive semidefinite in this case. Further, $\int K < \infty$ whenever $\psi \in L^1$. In this case, the kernel is in our framework and our result applies.

We note that for this example, if $\psi > 0$ is constant on blocks, then we reduce to the example of block kernels discussed earlier. The approximation scheme for this example basically approximates the function ψ by a piecewise constant function on $[0, 1]$. From an algorithmic viewpoint, our result says that for evaluating the first order correction term, we can coarsen the model to a setup where there are finitely many species with the same activity.

EXAMPLE 3 (Dubins's model). Consider the Dubins kernel $K(x, y) = \frac{1}{\max\{x,y\}}$. Observe that $K \in L^1$, and is symmetric and positive definite as a function so that our results apply. (The associated integral operator is a bounded operator from $L^2([0, 1])$ to $L^2([0, 1])$ as shown in [8].) In this case, $\mathbb{P}[\{i, j\} \in E_N] = c/j \wedge 1$ for $j \geq i, c$. This corresponds to the situation where the graph is formed by a sequential addition of vertices, and the jth vertex joins to the existing vertices independently with probability c/j. This model is naturally inhomogeneous, in that the older vertices usually have higher degrees and play a crucial role in determining the structure of the graph. An infinite version of this model was introduced by Dubins in 1984, who wished to determine the critical c such that the graph would have an infinite path (see [29, 49]). The critical constant $c = 1/4$ was partially determined by Kalikow and Weiss [29] and finally determined by Shepp [49]. Durett [19] determined that $c = 1/4$ is also the critical threshold for the emergence of a giant component in the finite graph. We refer the reader to [8] for a detailed survey of the model and related results.

1.3. *Outline*. Before turning to the body of the paper, let us briefly outline the structure of the paper and the proof of the main results. Theorem 1.1 follows using the interpolation idea introduced in [17] and generalized in [48]. Its proof, included in Section 8, compares the MaxCut_κ value on inhomogeneous random graphs to the ground state of the inhomogeneous Potts spin glass. The proof of Theorem 1.2 follows the method outlined above. The starting point of the proof is the characterization of a family of arrays that follow a natural generalization of the Ghirlanda–Guerra identities [23] for this setting. This combines the synchronization mechanisms of [43, 44] and is included in Section 2. In Section 3, we construct a perturbation of the system that does not affect the limiting free energy but allows us to use the derived invariance properties. With these results in hand, we prove the upper bound in Section 4 using a Guerra-type interpolation and the matching lower bound using an Aizenman–Sims–Starr scheme in Section 6. Before proving the lower bound, we briefly study the continuity of certain functionals used in the lower bound in Section 5.

2. Invariant arrays and their characterization. In this section, we study an invariance property that combines the multispecies and vector spin Ghirlanda–Guerra type identities [43, 44] for the limit points of doubly infinite array of overlaps, (1.10), of independent draws from the Gibbs measure. This will allow us to characterize these limit points.

Let $\mathcal{R}_\kappa$ be the space of arrays of the form $R^s_{\ell,\ell'}$ such that there is a collection of vectors $(v_i(\ell,s))_{i,\ell\geq 1, s\in\mathscr{S}}$ in $\mathbb{R}^\kappa$ such that

$$R^s_{\ell,\ell'} = \sum_i v_i(\ell,s)\otimes v_i(\ell',s), \tag{2.1}$$

and such that

$$\big(R^s_{\ell,\ell'}e_k, e_k\big)\in[0,1] \qquad \forall k\in[\kappa]. \tag{2.2}$$

We equip $\mathcal{R}_\kappa$ with the induced topology from the product topology on a countable product of $[0,1]$ with itself so that, in particular, it is compact Polish. Observe that the overlap array (1.10) is in $\mathcal{R}_\kappa$ for each N.

We now introduce the aforementioned invariance property. For any $m, p\geq 1$, $(\nu^s_k)_{s\in\mathscr{S},k\in[m]}\in\mathbb{R}^\kappa$, and $\varphi:(\mathbb{R}^m)^{|\mathscr{S}|}\to\mathbb{R}$, let

$$\mathcal{Q}_{\ell,\ell'} = \varphi\big(\big[\big((R^s_{\ell,\ell'})^{\circ p}\nu^s_1,\nu^s_1\big),\dots,\big((R^s_{\ell,\ell'})^{\circ p}\nu^s_m,\nu^s_m\big)\big]_{s\in\mathscr{S}}\big). \tag{2.3}$$

The $(R^s_{\ell,\ell'})^{\circ p}$ term appearing above is the Hadamard pth power of $R^s_{\ell,\ell'}$. We say that a random variable with values in $\mathcal{R}_\kappa$ is weakly exchangeable if

$$R^s_{\pi_s(\ell)\pi_s(\ell')} \overset{(d)}{=} R^s_{\ell,\ell'} \tag{2.4}$$

for all collections (π_s) of permutations of $\mathbb{N}$ of finitely many coordinates. We say that a random variable in $\mathcal{R}$ is IP-invariant if for all $n\geq 2$, bounded f, and choice of $\mathcal{Q}$ as above, we have

$$\mathbb{E}f(R^n)\mathcal{Q}_{1,n+1} = \frac{1}{n}\mathbb{E}f(R^n)\mathbb{E}\mathcal{Q}_{1,2} + \frac{1}{n}\sum_{\ell=2}^n \mathbb{E}f(R^n)\mathcal{Q}_{1,\ell}, \tag{2.5}$$

where R^n is the array $(R^s_{\ell,\ell'})_{\ell,\ell'\in[n],s\in\mathscr{S}}$.

We now turn to the main result of this section. Recall that ρ^s is a probability measure on $\mathscr{S}$, which we denote equivalently by $\mathrm{d}\rho$. Let

$$\bar{R}_{\ell,\ell'} = \int R^s_{\ell,\ell'}\,\mathrm{d}\rho. \tag{2.6}$$

Our goal is to prove that random arrays in $\mathcal{R}$ that satisfy (2.5) synchronize, in the sense that $R^s_{\ell,\ell'}$ is actually a Lipschitz function of the trace.

THEOREM 2.1. *Suppose that R is a $\mathcal{R}_\kappa$-valued random variable that is* IP*-invariant and weakly exchangeable. Then there are deterministic, Lipschitz functions Ψ_s, depending on the law of R, such that*

$$R^s_{\ell,\ell'} = \Psi_s\big(\mathrm{tr}(\bar{R}_{\ell,\ell'})\big) \tag{2.7}$$

almost surely.

The proof of this result is essentially by composition of the synchronization theorems from [43, 44].

LEMMA 2.2. *Suppose that R in $\mathcal{R}$, is weakly exchangeable and* IP*-invariant. Then there exists deterministic, Lipschitz functions, depending on the law of* $\operatorname{tr}(\bar{R}_{\ell,\ell'})$, *such that*

$$\operatorname{tr}\bigl(R^s_{\ell,\ell'}\bigr) = L_s\bigl(\operatorname{tr}(\bar{R}_{\ell,\ell'})\bigr). \tag{2.8}$$

PROOF. We begin by observing that for any element of $\mathcal{R}$, the array of traces $T = (T^s_{\ell,\ell'})_{\ell,\ell'\geq 1} = (\operatorname{tr}(R^s_{\ell,\ell'}))_{\ell,\ell'\geq 1}$ is a Gram matrix for each s. To see this, simply observe that if we let

$$V_\ell(s) = \sum_i v_i(\ell, s) \otimes e_i$$

then

$$\operatorname{tr}\bigl(R^s_{\ell,\ell'}\bigr) = \bigl(V_\ell(s), V_{\ell'}(s)\bigr)_{\mathrm{HS}}.$$

Thus T is a Gram–de Finetti array for each $s \in \mathscr{S}$. Applying (2.5) with f and φ being functions of T, we see that T satisfies the multispecies Ghirlanda–Guerra identities ([44], equation (36)),

$$\mathbb{E} f\bigl(T^n\bigr)\mathcal{Q}_{1,n+1} = \frac{1}{n}\mathbb{E} f\bigl(T^n\bigr)\mathbb{E}\mathcal{Q}_{1,2} + \frac{1}{n}\sum_{\ell=2}^{n}\mathbb{E} f\bigl(T^n\bigr)\mathcal{Q}_{1,\ell},$$

where $T^n = (T_{\ell,\ell'})_{\ell,\ell'\in[n]}$. Recall from [44], Section 4, that for such arrays there exist Lipschitz functions L_s, that depend on the law of T^n, such that

$$T^s_{\ell,\ell'} = L_s\biggl(\int T^s_{\ell,\ell'}\,\mathrm{d}\rho\biggr)$$

almost surely. Applying this result to our setting yields a family of Lipschitz functions depending on the law of $\operatorname{tr}(\bar{R}_{\ell,\ell'})$ such that

$$\operatorname{tr}\bigl(R^s_{\ell,\ell'}\bigr) = L_s\bigl(\operatorname{tr}(\bar{R}_{\ell,\ell'})\bigr)$$

almost surely, as desired. $\square$

LEMMA 2.3. *Suppose that R in $\mathcal{R}_\kappa$, is weakly exchangeable and* IP*-invariant. Then there exist deterministic, Lipschitz, monotone functions $\Theta^s : \mathbb{R}_+ \to \Gamma_\kappa$ which depend on the law of R such that*

$$R^s_{\ell,\ell'} = \Theta^s\bigl(\operatorname{tr}\bigl(R^s_{\ell,\ell'}\bigr)\bigr)$$

almost surely.

PROOF. For fixed $s \in \mathscr{S}$, we may apply (2.5), with $\mathcal{Q}$ of the form

$$\mathcal{Q}_{\ell,\ell'} = \varphi\bigl(\bigl(\bigl(R^s_{\ell,\ell'}\bigr)^{\circ p}\nu^s_1, \nu^s_1\bigr), \ldots, \bigl(\bigl(R^s_{\ell,\ell'}\bigr)^{\circ p}\nu^s_m, \nu^s_m\bigr)\bigr).$$

As a result, taking f to be a function of this species as well, we see that the array satisfies

$$\mathbb{E} f\big(R^n(s)\big)\mathcal{Q}_{1,n+1} = \frac{1}{n}\mathbb{E} f\big(R^n(s)\big)\mathbb{E}\mathcal{Q}_{1,2} + \frac{1}{n}\sum_{\ell=2}^{n}\mathbb{E} f\big(R^n(s)\big)\mathcal{Q}_{1,\ell},$$

where $R^n(s) = (R^s_{\ell,\ell'})_{\ell,\ell'\in[n]}$. It was shown in [43], Theorem 3, that for such arrays, there is a Lipschitz, monotone function Θ^s depending on the law of $R^n(s)$ such that

$$R^s_{\ell,\ell'} = \Theta^s\big(\mathrm{tr}\big(R^s_{\ell,\ell'}\big)\big)$$

almost surely, as desired. □

PROOF OF THEOREM 2.1. Applying the previous two lemmas, we obtain families $(\Theta^s)_{s\in\mathscr{S}}$ and $(L_s)_{s\in\mathscr{S}}$. The result then follows by taking

$$\Psi_s = \Theta^s \circ L_s. \qquad \square$$

3. Perturbation for invariance. In this section, we show that after a small perturbation, the limiting overlap array will satisfy a generalized form of the Ghirlanda–Guerra identities appearing in [44] and [43]. This argument is standard and can be safely skipped by the expert reader. For completeness, we include it here. The key observation is that, as with the Potts model, it is crucial that we restrict ourselves to configurations with fixed proportions of states.

Let $h_\theta(\sigma)$ be a Gaussian process with covariance

$$C^\theta_{\ell,\ell'} = \mathrm{Cov}\big(h_\theta(\sigma^\ell)h_\theta(\sigma^{\ell'})\big) = \prod_{s\in\mathscr{S}}\prod_{j\le m}\big(\big((R^s_{\ell,\ell'})^{\circ p}\nu^s_j, \nu^s_j\big)\big)^{n^s_j}, \tag{3.1}$$

where $\theta = (m, p, n^s_1, \dots, n^s_m, \nu^s_1, \dots, \nu^s_m)_{s\in\mathscr{S}}$ are the parameters in the covariance. The Gaussian process $h_\theta(\sigma)$ can be constructed explicitly using a similar construction as in [43], Section 5. We will provide a brief nonconstructive existence proof here.

LEMMA 3.1. *The covariance structure $C^\theta_{\ell,\ell'}$ is positive semidefinite for all $R^s_{\ell,\ell'} \in \mathcal{R}_\kappa$.*

PROOF. Clearly, if $R^s_{\ell,\ell'} \in \mathcal{R}_\kappa$, then $(R^s_{\ell,\ell'})^{\circ p} \in \mathcal{R}_\kappa$ for all $p \ge 1$. By the definition on (2.1), we can find some collection of vectors $(v_i(\ell, s))_{i,\ell\ge1,s\in\mathscr{S}}$ such that

$$\big(R^s_{\ell,\ell'}\big)^{\circ p} = \sum_i v_i(\ell, s)\otimes v_i\big(\ell', s\big) = \sum_i v_i(\ell, s)v_i^{\mathrm{T}}\big(\ell', s\big).$$

Given $v_j^s \in \mathbb{R}^\kappa$, we have

$$\begin{aligned}\big((R^s_{\ell,\ell'})^{\circ p} v_j^s, v_j^s\big)_{\ell,\ell'\geq 1} &= \Big(\sum_i v_i(\ell,s) v_i^{\mathrm{T}}(\ell',s) v_j^s, v_j^s\Big)_{\ell,\ell'\geq 1} \\ &= \sum_i \big(v_i(\ell',s)^{\mathrm{T}} v_j^s, v_i(\ell,s)^{\mathrm{T}} v_j^s\big)_{\ell,\ell'\geq 1}\end{aligned} \tag{3.2}$$

is a Gram array, and hence positive semidefinite. Since Hadamard products preserves positivity,

$$\prod_{s\in\mathscr{S}} \prod_{j\leq m} \big(\big((R^s_{\ell,\ell'})^{\circ p} v_j^s, v_j^s\big)\big)^{n_j^s}$$

is positive semidefinite because it is the Hadamard product of finitely many arrays of the form (3.2). Hence there exists a Gaussian process indexed with σ^ℓ with covariance given by (3.1). $\square$

Let v_j^s take rational values in $[-1,1]^\kappa$ and define the space of parameters

$$\Theta := \big\{\theta : m, p, n_1^s, \ldots, n_m^s \in \mathbb{N}, v_1^s, \ldots, v_m^s \in \mathbb{Q}\cap[-1,1] \text{ for all } s\in\mathscr{S}\big\}. \tag{3.3}$$

Since Θ is countable, we can find a enumeration map $j(\theta):\Theta\to\mathbb{N}$. Let $(u_\theta)_{\theta\in\Theta}$ be a sequence in $[1,2]$ and let $(h_\theta)_{\theta\in\Theta}$ of be pairwise independent copies of h_θ. Finally, define

$$h_N(\sigma) = \sum_{\theta\in\Theta} \frac{1}{2^{j(\theta)}} u_\theta h_\theta(\sigma). \tag{3.4}$$

Let $d_N \in \mathscr{D}_N$ be such that $d_N \to d \in \mathscr{D}$, and consider the perturbed Gibbs measure on $\Sigma_N(d_N)$ given by

$$G^{\mathrm{pert}}_{d_N} = \frac{\exp H_N^{\mathrm{pert}}(\sigma)}{Z_N(d_N)}, \qquad H_N^{\mathrm{pert}} = H_N(\sigma) + s_N h_N(\sigma), \tag{3.5}$$

where $\sigma\in\Sigma_N(d_N)$ and $s_N = N^\alpha$ for $1/4<\alpha<1/2$. We then have the following.

THEOREM 3.2. *There is a choice of (u_θ) such that the following hold*:

- *The perturbation is small in the sense that*

$$\lim_{N\to\infty}\left|\frac{1}{N}\mathbb{E}\log\sum_{\sigma\in\Sigma_N(d)}\exp\big(H_N(\sigma)\big) - \frac{1}{N}\mathbb{E}\log\sum_{\sigma\in\Sigma_N(d)}\exp\big(H_N^{\mathrm{pert}}(\sigma)\big)\right| = 0.$$

- *If $R^N = (R^s_{\ell,\ell'})_{s\in\mathscr{S},\ell,\ell'\geq 1}$ is the overlap array drawn from $\mathbb{E}(G^{\mathrm{pert}}_{d_N})^\infty$, then any weak limit point, R^∞, satisfies (2.5).*

PROOF. The proof of this fact is almost identical to Chapter 3.2 in [40], so we omit most details. The only essential difference is the same as that in [43], namely to point out why restricting the configuration space is important. This is because the main integration by parts step in the proof of [40], Theorem 3.2, uses in an essential way that the self overlap, $R_{\ell,\ell}$, and thus the variance of the field h_N, is constant. In our setting, the relevant term, namely $C^{\theta}_{\ell,\ell}$ is plainly constant on $\Sigma_N(d_N)$ by inspection of (3.1). □

4. Upper bound—Guerra interpolation. We now turn to proving the upper bound for the restricted free energy in (1.23) by a Guerra interpolation argument [24]. Recall the definition of $\Sigma_N^{\varepsilon}(d)$ (1.13) and let the corresponding partition function be denoted by $Z_N^{\varepsilon}(d) = \sum_{\sigma \in \Sigma_N^{\varepsilon}(d)} \exp(H_N(\sigma))$. We will prove

$$\limsup_{N\to\infty} \frac{1}{N} \mathbb{E} \log Z_N^{\varepsilon}(d) \le \mathscr{P}\big(r, x, d, \big(\lambda^s, Q^s\big)_{s\in\mathscr{S}}\big) + O(\varepsilon). \tag{4.1}$$

Given a sequence of strictly increasing $(x_i)_{i=1}^r$ as in (1.18), let $(v_\alpha)_{\alpha\in\mathbb{N}^r}$ be the weights of the Ruelle probability cascades [47] associated with that sequence. For $\alpha, \beta \in \mathbb{N}^r$, define

$$|\alpha \wedge \beta| = \min\{0 \le p \le r-1 | \alpha_1 = \beta_1, \dots, \alpha_p = \beta_p, \alpha_{p+1} \ne \beta_{p+1}\} \tag{4.2}$$

and $|\alpha \wedge \beta| = r$ if $\alpha = \beta$. For each species, let $(Z_s^{\alpha}(k))_{k\le\kappa}$ be the centered Gaussian vector with covariance

$$\mathbb{E} Z_s^{\alpha} Z_s^{\beta} = 2 \sum_{t\in\mathscr{S}} \Delta_{st}^2 \rho_N^t Q^t_{|\alpha\wedge\beta|}. \tag{4.3}$$

Similarly, let

$$\mathbb{E} Y^{\alpha} Y^{\beta} = \sum_{s,t\in\mathscr{S}} \Delta_{st}^2 \rho_N^s \rho_N^t \big(Q^s_{|\alpha\wedge\beta|}, Q^t_{|\alpha\wedge\beta|}\big). \tag{4.4}$$

For each $s \in \mathscr{S}$ and each $i \in I_s$, let Z_i^{α} be an independent copy of Z_s^{α}. The processes Z_i^{α}, Y^{α} and $H_N(\sigma)$ are all independent. Finally, define the interpolating Hamiltonian,

$$H_t(\sigma, \alpha) = \sqrt{t} H_N(\sigma) + \sqrt{t}\sqrt{N} Y^{\alpha} + \sqrt{1-t} \sum_{i\le N} Z_i^{\alpha}(\sigma_i), \tag{4.5}$$

and the corresponding interpolating free energy function

$$\varphi_N^{\varepsilon}(t) = \frac{1}{N} \mathbb{E} \log \sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\sigma\in\Sigma_N^{\varepsilon}(d)} e^{H_t(\sigma,\alpha)}. \tag{4.6}$$

We then have the following result.

LEMMA 4.1. *For any $\varepsilon > 0$ and $N \geq \frac{1}{\varepsilon}$,*

$$\partial_t \varphi_N^{\varepsilon}(t) \leq C\varepsilon^2,$$

for some constant $C(\kappa, \Delta, |\mathscr{S}|)$, uniformly in N. Furthermore, if $d \in \mathscr{D}_N$ and $\varepsilon = 0$, then

$$\partial_t \varphi_N^{0}(t) \leq 0.$$

PROOF. For any $d \in \mathscr{D}$, the set $\Sigma_N^{\varepsilon}(d)$ is nonempty for $N \geq \frac{1}{\varepsilon}$. Recall that by Gaussian integration by parts,

$$\begin{aligned}\partial_t \varphi_N^{\varepsilon}(t) &= \frac{1}{N}\mathbb{E}\left\langle \frac{\partial H_t(\sigma,\alpha)}{\partial t}\right\rangle \\ &= \frac{1}{N}\mathbb{E}\left\langle \mathbb{E}\frac{\partial H_t(\sigma^1,\alpha^1)}{\partial t} H_t(\sigma^1,\alpha^1) - \mathbb{E}\frac{\partial H_t(\sigma^1,\alpha^1)}{\partial t} H_t(\sigma^2,\alpha^2)\right\rangle,\end{aligned} \tag{4.7}$$

where $\langle\cdot\rangle$ is with respect to the Gibbs measure $G(\sigma,\alpha) \propto v_\alpha \exp(H_t(\sigma,\alpha))$ on $\Sigma_N^{\varepsilon}(d) \times \mathbb{N}^r$. If we write $Z_i^{\alpha}(\sigma_i) = \sum_{k\leq\kappa} Z_i^{\alpha}(k)\mathbb{1}(\sigma_i = k)$, then

$$\begin{aligned}&\mathbb{E}\frac{\partial H_t(\sigma^1,\alpha^1)}{\partial t} H_t(\sigma^2,\alpha^2) \\ &\quad= \frac{1}{2}\left(\mathbb{E}H_N(\sigma^1)H_N(\sigma^2) + N\mathbb{E}Y^{\alpha^1}Y^{\alpha^2} - \sum_{s\in\mathscr{S}}\sum_{i\in I_s}\mathbb{E}Z_s^{\alpha^1}(\sigma_i^1)Z_s^{\alpha^2}(\sigma_i^2)\right) \\ &\quad= \frac{N}{2}\sum_{s,t\in\mathscr{S}}\rho_N^s\rho_N^t\Delta_{st}^2\big[(R_{1,2}^s, R_{1,2}^t) + (Q_{|\alpha^1\wedge\alpha^2|}^t, Q_{|\alpha^1\wedge\alpha^2|}^s) \\ &\qquad - 2(Q_{|\alpha^1\wedge\alpha^2|}^t, R_{1,2}^s)\big] \\ &\quad= \frac{N}{2}\sum_{s,t\in\mathscr{S}}\Delta_{st}^2(R_{1,2}^s - Q_{|\alpha^1\wedge\alpha^2|}^s, R_{1,2}^t - Q_{|\alpha^1\wedge\alpha^2|}^t)\rho_N^s\rho_N^t \geq 0.\end{aligned} \tag{4.8}$$

The last quantity is nonnegative comes from the fact that if we define the matrix

$$A_{1,2,\alpha^1,\alpha^2} = \big((R_{1,2}^s - Q_{|\alpha^1\wedge\alpha^2|}^s, R_{1,2}^t - Q_{|\alpha^1\wedge\alpha^2|}^t)\big)_{s,t}$$

then it is positive definite as it is a Gram-matrix. The Hadamard product of this matrix with the positive definite matrix $\Delta = (\Delta_{s,t}^2)_{s,t}$, is still positive definite by the Schur product theorem. Thus (4.8) can be written as

$$\frac{N}{2}\big(\Delta^2 \circ (A_{1,2,\alpha^1,\alpha^2})\rho, \rho\big) \geq 0, \tag{4.9}$$

where $\circ$ denotes the Hadamard product. For $\sigma \in \Sigma_N^{\varepsilon}(d)$, the self overlap matrix $R_{1,1}^s$ is a diagonal matrix with entries $R_{1,1}^s(k,k) \in [d_k^s - \varepsilon, d_k^s + \varepsilon]$. Therefore, the

diagonal terms of $R^s_{1,1} - Q^s_{|\alpha^1 \wedge \alpha^1|}$ satisfy

$$\left|R^s_{1,1}(k,k) - Q^s_r(k,k)\right| \le \varepsilon \qquad \forall s \in \mathscr{S}, k \le \kappa.$$

Since there are κ nonzero terms in the inner product $(R^s_{1,1} - Q^s_{|\alpha^1 \wedge \alpha^1|}, R^t_{1,1} - Q^t_{|\alpha^1 \wedge \alpha^1|})$, our upper bound of the diagonals (4.8) imply

$$\begin{aligned} \mathbb{E}\frac{\partial H_t(\sigma^1, \alpha^1)}{\partial t} H_t(\sigma^2, \alpha^2) &= \frac{N}{2}\big(\Delta^2 \circ (A_{1,1,\alpha^1\alpha^1})\rho, \rho\big) \\ &\le \frac{N\|\Delta\|\mathscr{S}|^2\kappa}{2}\varepsilon^2. \end{aligned} \tag{4.10}$$

Plugging (4.9) and (4.10) in (4.7), implies

$$\begin{aligned} \partial_t \varphi^\varepsilon_N(t) &= \mathbb{E}\left\langle \frac{1}{2}\big(\Delta^2 \circ (A_{1,1,\alpha^1,\alpha^1})\rho, \rho\big)\right\rangle - \mathbb{E}\left\langle \frac{1}{2}\big(\Delta^2 \circ (A_{1,2,\alpha^1,\alpha^2})\rho, \rho\big)\right\rangle \\ &\le \frac{\|\Delta\|\mathscr{S}|^2\kappa}{2}\varepsilon^2. \end{aligned}$$

In the case $d \in \mathscr{D}_N(d)$, the set $\Sigma_N(d) = \Sigma^0_N(d)$ is nonempty. For $\sigma \in \Sigma_N(d)$, the overlap array $R^s_{\ell,\ell'} = Q^s_r$ for all $s \in \mathscr{S}$. In particular, $A_{1,1,\alpha^1,\alpha^1} = 0$, which implies

$$\partial_t \varphi^0_N(t) = -\mathbb{E}\left\langle \frac{1}{2}\big(\Delta^2 \circ (A_{1,2,\alpha^1,\alpha^2})\rho, \rho\big)\right\rangle \le 0. \qquad \square$$

As a consequence $\frac{1}{N}\mathbb{E}\log Z^\varepsilon_N(d)$ is bounded above by

$$\begin{aligned} &\frac{1}{N}\mathbb{E}\log \sum_{\alpha \in \mathbb{N}^r} v_\alpha \sum_{\sigma \in \Sigma^\varepsilon_N(d)} \exp \sum_{i \le N} Z^\alpha_i(\sigma_i) \\ &\qquad - \frac{1}{N}\mathbb{E}\log \sum_{\alpha \in \mathbb{N}^r} v_\alpha \exp \sqrt{N} Y^\alpha + O\big(\varepsilon^2\big). \end{aligned} \tag{4.11}$$

We now introduce the Lagrange multipliers $(\lambda^s_k)_{k \le \kappa} \in \mathbb{R}^k$, which are dual to the proportions $\sum_{i \in I_s} \mathbb{1}(\sigma_i = k)$. If we add and subtract $\sum_{i \le N}\sum_{k \le \kappa} \lambda^{s(i)}_k \mathbb{1}(\sigma_i = k)$ in the exponent of the first term in (4.11), then the first term is bounded by

$$\begin{aligned} &-\sum_{s \in \mathscr{S}}\sum_{k \le \kappa} \rho^s_N d^s_k \lambda^s_k \\ &\quad + \frac{1}{N}\mathbb{E}\log \sum_{\alpha \in \mathbb{N}^r} v_\alpha \sum_{\sigma \in \Sigma^\varepsilon_N(d)} \exp \sum_{i \le N}\left(Z^\alpha_i(\sigma_i) + \sum_{k \le \kappa} \mathbb{1}(\sigma_i = k)\lambda^{s(i)}_k\right) + O(\varepsilon). \end{aligned}$$

Since $\Sigma^\varepsilon_N(d) \subset \Sigma_N$, summing over $\sigma \in \Sigma_N$ in the larger set only increases our upper bound. We can now factor into species using the basic properties of the

Ruelle probability cascades (see the discussion after Theorem 2.9 in [40]) to see

$$\begin{aligned}
&\frac{1}{N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\sigma\in\Sigma_N}\exp\sum_{i\le N}\Big(Z_i^\alpha(\sigma_i)+\sum_{k\le\kappa}\mathbb{1}(\sigma_i=k)\lambda_k^{s(i)}\Big)\\
&\quad=\frac{1}{N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha \prod_{i\le N}\sum_{\sigma_i\le\kappa}\exp\Big(Z_i^\alpha(\sigma_i)+\sum_{k\le\kappa}\mathbb{1}(\sigma_i=k)\lambda_k^{s(i)}\Big)\\
&\quad=\sum_{s\in\mathscr{S}}\rho_N^s\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha\sum_{\sigma\le\kappa}\exp\big(Z_s^\alpha(\sigma)+\lambda_\sigma^s\big)\\
&\quad=\sum_{s\in\mathscr{S}}\rho_N^s X_0^s,
\end{aligned}\tag{4.12}$$

where X_0^s was defined in (1.21). Similarly, we see

$$\begin{aligned}
&\frac{1}{N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha\exp\sqrt{N}Y^\alpha\\
&\quad=\frac{1}{2}\sum_{\ell=0}^{r-1}x_\ell\sum_{s,t\in\mathscr{S}}\Delta_{st}^2\rho_N^s\rho_N^t\big((Q_{\ell+1}^s,Q_{\ell+1}^t)-(Q_\ell^s,Q_\ell^t)\big).
\end{aligned}\tag{4.13}$$

Referring back to (4.11), equations (4.12) and (4.13) imply

$$\limsup_{N\to\infty}\frac{1}{N}\mathbb{E}\log Z_N^\varepsilon(d)\le\mathscr{P}\big(r,x,d,(\lambda^s,Q^s)_{s\in\mathscr{S}}\big)+O(\varepsilon).\tag{4.14}$$

In the case $d_N\in\mathscr{D}_N$, $\sum_{i\in I_s}\mathbb{1}(\sigma_i=k)=N_sd_k^s$ for all $s\in\mathscr{S}$ and $k\le\kappa$. The above computation implies

$$\frac{1}{N}\mathbb{E}\log Z_N(d_N)\le\mathscr{P}\big(r,x,d,(\lambda^s,Q^s)_{s\in\mathscr{S}}\big).\tag{4.15}$$

5. Continuity and decoupling theorems. Before we turn to the proof of the matching lower bounds, we briefly pause to study the analytical properties of some of the functionals used in the upper bound, as well as relevant functionals for the lower bound. These functionals will be in terms of the Gaussian processes (Z_s^α) and (Y^α) from (4.3) and (4.4), respectively.

Many of the proofs in this section are essentially identical to arguments either from [43], Section 3, or are standard arguments and can be seen, for example, in [40, 52]. Thus to make the presentation concise, we explain only the parts where these arguments deviate from standard arguments and outline the rest.

5.1. *Decoupling size of the constraints.* When we compute the lower bound for the free energy, we will find that the cavity method will naturally impose an additional constraint on free energy, namely, that the cavity coordinates satisfy

the additional constraint that they lie in some $\Sigma(d)$. As this constraint does not appear in the upper bound from Section 4, we will need to remove this to obtain the matching lower bound. To this end, define the functional on $\mathscr{D}_N$

$$f_N^s(d) := \frac{1}{N}\mathbb{E}\log \sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\sigma\in\Sigma_N(d)} \exp \sum_{i\le N} Z_i^\alpha(\sigma_i), \tag{5.1}$$

where Z_i^α are i.i.d. copies of Z_s^α. Furthermore, we fix the covariance structure of Z_i^α and make the dependence of X_0^s on the parameter λ explicit

$$X_0^s(\lambda) = \frac{1}{N}\mathbb{E}\log \sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\sigma\in\Sigma_N} \exp \sum_{i\le N} Z_s^\alpha(\sigma_i) + \sum_{k\le\kappa} \lambda_k^s \mathbb{1}(\sigma_i = k). \tag{5.2}$$

We have the following result.

THEOREM 5.1. *If $d_N^s \in \mathscr{D}_N$ and $\lim_{N\to\infty} d_N^s \to d^s$, then*

$$\lim_{N\to\infty} f_N^s(d_N) = \inf_{\lambda^s}\Big(-\sum_{k\le\kappa}\lambda_k^s d_k^s + X_0^s(\lambda)\Big). \tag{5.3}$$

PROOF. We begin by observing that there is a constant L such that

$$\sup_{d\in\mathscr{D}_N} \big|f_{N,\varepsilon}^s(d) - f_N^s(d)\big| \le L\sqrt{\varepsilon}, \tag{5.4}$$

where $f_{N,\varepsilon}^s(d)$ is the same functional as in (5.1) but summed over $\sigma \in \Sigma_N^\varepsilon(d)$. This is the analogue of [43], Lemma 3, adapted to the covariance structure of the Gaussian processes Z_i defined in (4.3). We also fix the covariance structure of Z_i to remove its dependence on d. Let $\sigma \in \Sigma_N^\varepsilon(d)$ and $\tilde\sigma$ be a vector in $\Sigma_N(d)$ with the minimal number of different coordinates from σ. Let

$$\tilde f_{N,\varepsilon}^s = \frac{1}{N}\mathbb{E}\log \sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\sigma\in\Sigma_N^\varepsilon(d)} \exp \sum_{i\in S_N} Z_i^\alpha(\tilde\sigma_i).$$

Finally, let $\tilde Z_i^\alpha$ be independent copies of Z_i^α, and consider the "smart path"

$$Z_t(\alpha,\sigma) = \sum_{i\le N}\big(\sqrt{t}\,Z_i^\alpha(\sigma_i) + \sqrt{1-t}\,\tilde Z_i^\alpha(\tilde\sigma_i)\big).$$

We will then show that

$$\varphi(t) = \frac{1}{N}\mathbb{E}\log \sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\sigma\in\Sigma_N^\varepsilon(d)} \exp Z_t(\alpha,\sigma),$$

has a derivative that is bounded by $C\varepsilon$. To see this, let

$$\begin{aligned} C\big((\sigma^1,\alpha^1),(\sigma^2,\alpha^2)\big) &= \frac{1}{N}\mathbb{E}\frac{\partial Z_t(\sigma^1,\alpha^1)}{\partial_t} Z_t\big(\sigma^2,\alpha^2\big) \\ &= \frac{1}{N}\sum_{i\le N}\Big(\sum_{t\in\mathscr{S}}\Delta_{st}^2\rho^t\big(Q^t_{|\alpha\wedge\beta|}(\sigma_i^1,\sigma_i^2) - Q^t_{|\alpha\wedge\beta|}(\tilde\sigma_i^1,\tilde\sigma_i^2)\big)\Big) \\ &\le \|\Delta\|\cdot|\mathscr{S}|\cdot\kappa\cdot\varepsilon. \end{aligned}$$

In the last line, we use that $Q^t \le D^t$ and that $\|D^t\| \le 1$ along with the observation that

$$\sum \mathbb{1}(\sigma_i \ne \tilde\sigma_i) \le \lceil \kappa N\varepsilon\rceil.$$

Differentiating and integrating by parts, we obtain

$$|\varphi'(t)| = \big|\mathbb{E}\langle C\big((\sigma^1,\alpha^1),(\sigma^1,\alpha^1)\big) - C\big((\sigma^1,\alpha^1),(\sigma^2,\alpha^2)\big)\rangle\big| \le 2L\varepsilon,$$

for some constant $L = L(\kappa,\mathscr{S},\Delta)$. Integrating this inequality, we can conclude

$$|f^s_{N,\varepsilon} - \tilde f^s_{N,\varepsilon}| \le 2L\varepsilon. \tag{5.5}$$

For $\sigma\in\Sigma(d)$, let us denote by $\mathcal{N}(\sigma)$ the number of configurations $\rho\in\Sigma_\varepsilon(d)$ such that $\tilde\rho = \sigma$. Then we can rewrite and bound $\tilde f^s_{N,\varepsilon}(d)$ as follows:

$$\begin{aligned} \tilde f_{N,\varepsilon}(d) &= \frac{1}{N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\sigma\in\Sigma(d)}\mathcal{N}(\sigma)\exp\beta\sum_{i\le N} Z_i^\alpha(\sigma_i) \\ &\le f^s_N(d) + \frac{1}{N}\max_{\sigma\in\Sigma(d)}\log\mathcal{N}(\sigma). \end{aligned}$$

Using a combinatorial argument, the term containing $\mathcal{N}(\sigma)$ can be made arbitrarily small by choosing ε small enough. For any $\sigma\in\Sigma(d)$, the number $\mathcal{N}(\sigma)$ is bounded by the number of configurations ρ such that $\sum_{i\le N} I(\rho_i\ne\sigma_i) \le LN\varepsilon$. By the classical large deviation estimate for Bernoulli random variables, a number of different ways to choose $LN\varepsilon$ coordinates is bounded by $2^N\exp(-NI(1-L\varepsilon))$, where

$$I(x) = \frac{1}{2}\big((1+x)\log(1+x) + (1-x)\log(1-x)\big),$$

and there are $\kappa^{LN\varepsilon}$ ways to choose ρ_i different from σ_i on these coordinates. Therefore,

$$\begin{aligned} \frac{1}{N}\max_{\sigma\in\Sigma(d)}\log\mathcal{N}(\sigma) &\le L\varepsilon\log\kappa + \log 2 - I(1-L\varepsilon) \\ &= L\varepsilon\log\kappa + \log\Big(1+\frac{L\varepsilon}{2-L\varepsilon}\Big) + \frac{L\varepsilon}{2}\log\frac{2-L\varepsilon}{\varepsilon} \le L\sqrt{\varepsilon}, \end{aligned}$$

for small enough ε. We showed that $f_{N,\varepsilon}(d) \le f_N(d) + L\sqrt{\varepsilon}$. Combining this with (5.5) yields the estimate

$$|f_N^s - f_{N,\varepsilon}^s| \le 2L\sqrt{\varepsilon}.$$

Note that (5.4), implies that the map $d \mapsto f_N^s(d)$ is Hölder $1/2$.

The remaining steps to complete the proof are identical to [43], Section 3. An additivity argument and (5.4) will imply the following.

LEMMA 5.2 ([43], Lemma 4, 5). *If $d_N \in \mathscr{D}_N$ and $\lim_{N\to\infty} d_N = d \in \mathscr{D}$, then the limit*

$$f(d) := \lim_{N\to\infty} f_N(d_N) = \lim_{\varepsilon\to 0} f_\varepsilon(d)$$

exists and is concave. In addition, for all $d^1, d^2 \in \mathscr{D}$

$$|f(d^1) - f(d^2)| \le L\|d^1 - d^2\|_\infty^{1/2}$$

for some constant L that depends $\kappa, \mathscr{S}, \Delta$.

Next, a direct computation using the recursive property of the Ruelle probability cascades will imply the following result.

LEMMA 5.3 ([43], Lemma 6). *For any $\lambda = (\lambda_k)_{k\le\kappa} \in \mathbb{R}^\kappa$*

$$X_0^s(\lambda) = \max_{d\in\mathscr{D}}\Big(f^s(d) + \sum_{k\le\kappa} \lambda_k^s d_k^s\Big). \tag{5.6}$$

Notice $f_N^s(d)$ is continuous and bounded on $\mathscr{D}$. Since it is also concave by Lemma 5.2, we can take the Legendre transform of (5.6) to obtain

$$\lim_{N\to\infty} f_M^s(d) = \inf_\lambda\Big(-\sum_{k\le\kappa} \lambda_k^s d_k^s + X_0^s(\lambda)\Big). \qquad \square$$

5.2. *Continuity theorems.* Along with the decoupling theorem, we will also need to prove continuity of the functionals appearing in the Aizenman–Sims–Starr scheme for this system.

Define the two functionals on a subset $S \subset [\kappa]^N$,

$$f_N^Z(S, s; \boldsymbol{\pi}) := \frac{1}{N_s}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\sigma\in S}\exp\sum_{i\le N_s} Z_i^\alpha(\sigma_i), \tag{5.7}$$

$$f_N^Y(\boldsymbol{\pi}) := \frac{1}{N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha \exp\sqrt{N}Y^\alpha, \tag{5.8}$$

where $(Z_i^\alpha)_{i\le N}$ are i.i.d. copies of Z_s^α defined in (4.3). Observe that this functional depends on $\boldsymbol{\pi} = (\pi^s)_{s\in\mathscr{S}}$ through the covariance structures (4.3) and (4.4).

LEMMA 5.4. *For every $S \subset [\kappa]^N$, the functional $f_N^Z(S, s; \boldsymbol{\pi})$ is Lipschitz in $\boldsymbol{\pi}$*

$$\bigl|f_N^Z(S, s; \boldsymbol{\pi}) - f_N^Z(S, s; \tilde{\boldsymbol{\pi}})\bigr| \le L \int_0^1 \max_{s \in \mathscr{S}} \bigl\|\pi^s(x) - \tilde{\pi}^s(x)\bigr\|_1 \,\mathrm{d}x.$$

PROOF. Observe that any two monotone paths $\boldsymbol{\pi}$ and $\tilde{\boldsymbol{\pi}}$, can be associated with a single sequence

$$x_{-1} = 0 \le x_0 \le \cdots \le x_{r-1} \le x_r = 1$$

and, for every $s \in \mathscr{S}$, two sequences

$$0 < Q_0^s \le \cdots < Q_{r-1}^s < Q_r^s = \operatorname{diag}(d_1, \ldots, d_\kappa),$$
$$0 < \tilde{Q}_0^s \le \cdots < \tilde{Q}_{r-1}^s < \tilde{Q}_r^s = \operatorname{diag}(d_1, \ldots, d_\kappa).$$

Consider the Gaussian processes Z_s^α and $\tilde{Z}_s^\alpha$ with the covariance (4.3) with Q and $\tilde{Q}$, respectively, and consider the smart path between these two processes:

$$Z_{t,i}^\alpha = \sqrt{t} Z_i^\alpha + \sqrt{1-t}\tilde{Z}_i^\alpha.$$

If we take N copies of this process, and let (v_α) be the Ruelle probability cascade [47] associated to (x_k), then if we define

$$\varphi(t) := \frac{1}{N_s}\mathbb{E}\log \sum_{\alpha \in \mathbb{N}^r} v_\alpha \sum_{\sigma \in S} \exp \sum_{i \le N_s} Z_{t,i}^\alpha(\sigma_i),$$

we have $\varphi(1) = f_Z(S, s; \boldsymbol{\pi})$ and $\varphi(0) = f_Z(S, s; \tilde{\boldsymbol{\pi}})$. If $H_{N,t}(\sigma, \alpha) = \sum_{i \le N} Z_{t,i}^\alpha(\sigma_i)$, let $\langle\cdot\rangle_t$ denote the average with respect to the Gibbs measure

$$G_t(\sigma, \alpha) \propto v_\alpha \exp H_{N,t}(\sigma, \alpha). \tag{5.9}$$

We can now compute $\varphi'(t)$ using integration by parts. We first note that the covariance

$$\frac{1}{N_s}\mathbb{E}\frac{\partial H_{N,t}(\sigma^1, \alpha^1)}{\partial t} H_{N,t}(\sigma^2, \alpha^2)$$
$$= \frac{1}{N_s}\sum_{i \le N}\sum_{s \in \mathscr{S}} \Delta_{s,s(i)}^2 \rho^s\bigl(Q_{\alpha^1 \wedge \alpha^2}^s(\sigma_i^1, \sigma_i^2) - \tilde{Q}_{\alpha^1 \wedge \alpha^2}^s(\sigma_i^1, \sigma_i^2)\bigr).$$

The term on the right is bounded in absolute value by

$$\|\Delta\| \cdot |\mathscr{S}| \max_{s \in \mathscr{S}} \bigl\| Q_{\alpha^1 \wedge \alpha^2}^s - \tilde{Q}_{\alpha^1 \wedge \alpha^2}^s \bigr\|_1.$$

Recalling that the marginal of $G_t(\sigma, \alpha)$ (5.9) on $\mathbb{N}^r$ has the same distribution as the weights of v_α ([40], Theorem 4.4), a standard Gaussian integration by parts

argument will show

$$\begin{aligned}|\varphi'(t)| &\le \mathbb{E}\Big\langle \|\Delta\| \cdot |\mathscr{S}| \max_{s\in\mathscr{S}} \big\| Q^s_{\alpha^1\wedge\alpha^2} - \tilde{Q}^s_{\alpha^1\wedge\alpha^2}\big\|_1 \Big\rangle_t \\ &\le \|\Delta\||\mathscr{S}| \sum_{0\le p\le r} \max_{s\in\mathscr{S}} \big\| Q^s_{\alpha^1\wedge\alpha^2} - \tilde{Q}^s_{\alpha^1\wedge\alpha^2}\big\|_1 \mathbb{E} \sum_{\alpha^1\wedge\alpha^2=p} v_{\alpha^1} v_{\alpha^2} \\ &= \|\Delta\||\mathscr{S}| \int_0^1 \max_{s\in\mathscr{S}} \big\|\pi^s(x) - \tilde{\pi}^s(x)\big\|_1 \,\mathrm{d}x.\end{aligned}$$

Integrating this inequality yields the result. □

The following argument is well known and follows by a direct computation using properties of Ruelle probability cascades; see [40, 43].

LEMMA 5.5. *The functional $f_N^Y(\boldsymbol{\pi})$ is Lipschitz in $\boldsymbol{\pi}$,*

$$\big|f_N^Y(\boldsymbol{\pi}) - f_N^Y(\tilde{\boldsymbol{\pi}})\big| \le L \int_0^1 \max_{s\in\mathscr{S}} \big\|\pi^s(x) - \tilde{\pi}^s(x)\big\|_1 \,\mathrm{d}x.$$

We finally observe here that discrete paths are dense in π.

LEMMA 5.6. *For any path $\boldsymbol{\pi} \in \boldsymbol{\Pi}$ and $\varepsilon > 0$, there exists a finite sequence of points $(x_p)_{p=1}^r$ such that the discrete path*

$$\boldsymbol{\pi}^*(x) := \boldsymbol{\pi}(x_p) \qquad \text{for } x_{p-1} < x \le x_p$$

satisfies

$$\Delta\big(\boldsymbol{\pi}, \boldsymbol{\pi}^*\big) < \varepsilon.$$

We now state a general continuity theorem regarding functionals of this form. Such results are completely standard; see, for example, [40, 43]. Let $S \subset [\kappa]^N$, $(w_\alpha)_{\alpha\in\mathscr{A}}$ be the weights (possibly random) of a probability density distribution on a countable set $\mathscr{A}$, and

$$R_{\mathscr{A}} = \big(R^s_{\alpha^1,\alpha^2}\big)_{s\in\mathscr{S},\alpha^1,\alpha^2\in\mathscr{A}},$$

where each $R^s_{\alpha^1,\alpha^2}$ is a $\kappa\times\kappa$ matrix. We think of this array as fixed and nonrandom. It will be the values that some abstract overlap structure can take. Finally, let (I_s) be partition of $[N]$ into species.

We define the functionals

$$f_{1,N} = \frac{1}{N}\mathbb{E}\log \sum_{\alpha\in\mathscr{A}} w_\alpha \sum_{\varepsilon\in S} \exp \sum_{i\le N} Z_i^\alpha(\varepsilon_i) \tag{5.10}$$

and

$$f_{2,N} = \frac{1}{N}\mathbb{E}\log\sum_{\alpha\in\mathscr{A}} w_\alpha \exp\sqrt{N}Y^\alpha. \tag{5.11}$$

Here, $Z_i^\alpha(\sigma)$ is a centered Gaussian process is such that, if $i \in I_s \subset [N]$, the covariance structure of the Gaussian vector $Z_i^\alpha = (Z_i^\alpha(k))_{k\le\kappa}$ is given by

$$\mathrm{Cov}\big(Z_i^{\alpha^1}, Z_i^{\alpha^2}\big) = C_Z^s\big(\big(R^s_{\alpha^1,\alpha^2}\big)_{s\in\mathscr{S}}\big), \tag{5.12}$$

for some C_z^s, a continuous function of the overlaps $R^s_{\alpha^1,\alpha^2}$. Similarly, Y is a centered Gaussian process with covariance given by

$$\mathrm{Cov}\big(Y^{\alpha^1}, Y^{\alpha^2}\big) = C_Y\big(\big(R^s_{\alpha^1,\alpha^2}\big)_{s\in\mathscr{S}}\big). \tag{5.13}$$

The following result is standard and follows from basic properties of Gaussian processes, and the fact that log-sums of exponentials have at most linear growth at infinity. Let $(\alpha(\ell))_{\ell\ge1}$ be i.i.d. drawn from $\mathscr{A}$ with law w_α, and denote

$$R^n = \big(R^s_{\alpha(\ell),\alpha(\ell')}\big)_{\ell,\ell'\in[n],s\in\mathscr{S}}.$$

We have the following continuity property.

LEMMA 5.7. *For any $\varepsilon > 0$, there are continuous bounded functions g_ε^Z and g_ε^Y such that*

$$\big|f_{1,N} - \mathbb{E}g_\varepsilon^Z(R^n)\big| \le \varepsilon, \qquad \big|f_{2,N} - \mathbb{E}g_\varepsilon^Y(R^n)\big| \le \varepsilon. \tag{5.14}$$

These functions depend at most on N, S, C_z, C_y and ε.

6. Lower bound via an Aizenman–Sims–Starr scheme. For fixed $d \in \mathscr{D}$ and a sequence of realizable proportions $d_N \in \mathscr{D}_N$ (1.14) converging to d, we prove the matching constrained lower bound:

$$\liminf_{N\to\infty}\frac{1}{N}\mathbb{E}\log Z_N(d_N) \ge \inf_{x,r,(\lambda^s,Q^s)_{s\in\mathscr{S}}} \mathscr{P}\big(r,x,d,(\lambda^s,Q^s)_{s\in\mathscr{S}}\big). \tag{6.1}$$

By part (1) of Theorem 3.2, computing the lower bound of the free energy with respect to

$$H_N^{\mathrm{pert}}(\sigma) = H_N(\sigma) + s_N h_N(\sigma) \tag{6.2}$$

introduced in Section 3 and the corresponding constrained partition function

$$Z_N(d_N) = \sum_{\sigma\in\Sigma_N(d_N)} \exp\big(H_N^{\mathrm{pert}}(\sigma)\big), \tag{6.3}$$

is equivalent to computing the lower bound in (6.1). We will continue to work with the perturbed Hamiltonian (6.2) throughout the remainder of this section.

For $M \geq 1$, our starting point is the following inequality:

$$\begin{aligned} &\liminf_{N\to\infty} \frac{1}{N}\mathbb{E}\log Z_N(d_N) \\ &\qquad \geq \liminf_{N\to\infty} \frac{1}{M}\big(\mathbb{E}\log Z_{N+M}(d_{N+M}) - \mathbb{E}\log Z_N(d_N)\big). \end{aligned} \tag{6.4}$$

We will write the free energy in terms of the Gibbs measure $G_{d_N}^{\mathrm{pert}}$ defined in (3.5) using the cavity method via an Aizenman–Sims–Starr scheme [2]. To this end, let us denote a configuration $\tilde{\rho} \in [\kappa]^{N+M}$ by $\tilde{\rho} = (\varepsilon, \sigma)$, where $\varepsilon = (\varepsilon_1, \dots, \varepsilon_M) \in [\kappa]^M$ are called the cavity coordinates and $\sigma = (\sigma_{M+1}, \dots, \sigma_{N+M}) \in [\kappa]^N$ are called the bulk coordinates. We define $\mathcal{M}_s, \mathcal{N}_s$ to be the subset of the respective cavity and bulk coordinates that belong to species s. Let M_s, N_s be the cardinality of $\mathcal{M}_s$ and $\mathcal{N}_s$.

We control the rate of convergence of d_N so that for some constant L_κ,

$$|d_{N,k}^s - d_k^s| \leq \frac{L_\kappa}{N_s} \qquad \text{for all } k \leq \kappa,\ s \in \mathscr{S} \quad \text{and} \quad d_{N,k}^s = 0 \qquad \text{if } d_k^s = 0. \tag{6.5}$$

A generalized version of [43], Lemma 11, will allow us to split constrained configuration space into a product set of the bulk coordinates and the species-wise cavity coordinates along a subsequence.

We first introduce some more notation. Let $\mathcal{A} \subset [N+M]$ and $\mathcal{A}_s$ is the subset of $\mathcal{A}$ in species s. If $A := |\mathcal{A}|$ is the cardinality of the set $\mathcal{A}$, we define

$$\Sigma_A(d) := \Big\{ (\sigma_i)_{i\in A} \Big| \sum_{i\in\mathcal{A}_s} \mathbb{1}(\sigma_i = k) = A_s d_k^s,\ \forall k \leq \kappa,\ s \in \mathscr{S} \Big\} \tag{6.6}$$

to be the configurations of spins in $\mathcal{A}$ that satisfy the constraint d. If $\mathcal{A} = [N+M]$, then (6.6) coincides with (1.13). The definition in (6.6) naturally implies

$$\Sigma_A(d) = \prod_{s\in\mathscr{S}} \Sigma_{A_s}(d^s). \tag{6.7}$$

If we use the sets $\mathcal{N}$ and $\mathcal{M}$ in place of $\mathcal{A}$ in (6.7), the observations means we can break the cavity and bulk coordinates into a product set over species. We will now show that the entire constrained system $\Sigma_{N+M}(d)$ contains a product set over the bulk and cavity coordinates.

LEMMA 6.1. *For every $M > 0$, there exists a constraint $\delta_M \in \mathscr{D}_M$ such that*

$$|\delta_{M,k}^s - d_k^s| \leq \frac{2L_\kappa}{M_s} \qquad \textit{for all } k \leq \kappa,\ s \in \mathscr{S}, \tag{6.8}$$

and we can find a subsequence of N such that

$$\Sigma_{N+M}(d_{N+M}) \supseteq \Sigma_N(d_N) \times \Sigma_M(\delta_M). \tag{6.9}$$

PROOF. We first fix $s \in \mathscr{S}$ and apply [43], Lemma 11, to the subset of spins in Σ_{N+M} belonging to species s. There exists a sequence $\delta_M^s \in \mathscr{D}_M$ such that $|\delta_{M,k}^s - d_k^s| \le \frac{2L_\kappa}{M_s}$ for all $k \le \kappa$, and

$$N_s d_N^s + M_s \delta_M^s = (N_s + M_s) d_{N+M}^s \tag{6.10}$$

for infinitely many N_s. Therefore, we can find a subsequence of N such that that associated N_s satisfies (6.10) and

$$\Sigma_{N_s+M_s}(d_{N+M}) \supseteq \Sigma_{N_s}(d_N) \times \Sigma_{M_s}(\delta_M).$$

Repeating the argument over each species and extracting a further subsequence each iteration, we can conclude

$$\prod_{s\in\mathscr{S}} \Sigma_{N_s+M_s}(d_{N+M}) \supseteq \prod_{s\in\mathscr{S}} \Sigma_{N_s}(d_N^s) \times \prod_{s\in\mathscr{S}} \Sigma_{M_s}(\delta_M^s).$$

Writing the product sets using the observation in (6.7) completes the proof. □

With this observation, we can restrict the first sum in (6.4) to the product set (6.9), so that (6.4) is bounded below by

$$\begin{aligned}\frac{1}{M}\Bigg(&\mathbb{E}\log \sum_{\sigma\in\Sigma_N(d_N)} \sum_{\varepsilon\in\Sigma_M(\delta_M)} \exp\bigl(H_{N+M}^{\mathrm{pert}}(\sigma,\varepsilon)\bigr)\\ &- \mathbb{E}\log \sum_{\sigma\in\Sigma_N(d_N)} \exp\bigl(H_N^{\mathrm{pert}}(\sigma)\bigr)\Bigg).\end{aligned} \tag{6.11}$$

We will now separate the unperturbed portion of the Hamiltonians into its cavity fields,

$$H_{N+M}(\varepsilon,\sigma) = H'_N(\sigma) + \sum_{i\le M} Z_{N,i}^{\sigma}(\varepsilon_i) + r(\varepsilon), \tag{6.12}$$

$$H_N(\sigma) = H'_N(\sigma) + \sqrt{M} Y_N^{\sigma}. \tag{6.13}$$

These cavity fields are the same as those appearing in [43], equation (107) and (109), except with different covariance structure because of the inhomogeneity. In our case, these fields are independent Gaussian processes with covariance

$$\mathbb{E}H'_N(\sigma^\ell)H'_N(\sigma^{\ell'}) = \frac{N^2}{N+M} \sum_{s,t\in\mathscr{S}} \Delta_{s,t}^2 \rho_N^s \rho_N^t \bigl(R_{\ell,\ell'}^s, R_{\ell,\ell'}^t\bigr), \tag{6.14}$$

$$\mathbb{E}Z_{N,i}^{\sigma^\ell}(\varepsilon)Z_{N,j}^{\sigma^{\ell'}}(\varepsilon') = 2\delta_{\{i=j\}} \sum_{s\in\mathscr{S}} \Delta_{s(i),s}^2 \rho_N^s R_{\ell,\ell'}^s(\varepsilon,\varepsilon') + O\bigl(N^{-1}\bigr), \tag{6.15}$$

$$\mathbb{E}Y_N^{\sigma^\ell} Y_N^{\sigma^{\ell'}} = \sum_{s,t\in\mathscr{S}} \Delta_{s,t}^2 \rho_N^s \rho_N^t \bigl(R_{\ell,\ell'}^s, R_{\ell,\ell'}^t\bigr) + O\bigl(N^{-1}\bigr). \tag{6.16}$$

Here, $\rho_N^s := \frac{\sum_{M+1=1}^{N+M} \mathbb{1}(i \in I_s)}{N}$ is the proportion of the bulk coordinates that belong to species s. This distinction is not critical because this proportion converges to the same ρ^s defined in (1.6). The $r(\varepsilon)$ term is $O(N^{-1})$ and can be omitted without affecting (6.1).

By a standard interpolation argument (see, e.g., [40], Theorem 3.6), if $(z_{N,i}^{\sigma}(\varepsilon))_{i \le M}$ and y_N^{σ} are centered Gaussian processes with covariances

$$\mathbb{E} z_{N,i}^{\sigma^\ell}(\varepsilon) z_{N,j}^{\sigma^{\ell'}}(\varepsilon') = 2\delta_{\{i=j\}} \sum_{s \in \mathscr{S}} \Delta_{s(i),s}^2 \rho^s R_{\ell,\ell'}^s(\varepsilon, \varepsilon'), \tag{6.17}$$

$$\mathbb{E} y_N^{\sigma^\ell} y_N^{\sigma^{\ell'}} = \sum_{s,t \in \mathscr{S}} \Delta_{s,t}^2 \rho^s \rho^t \big(R_{\ell,\ell'}^s, R_{\ell,\ell'}^t\big), \tag{6.18}$$

then (6.11) is bounded below by

$$\frac{1}{M} \mathbb{E} \log \Big\langle \sum_{\varepsilon \in \Sigma_M(\delta)} \exp \sum_{i \le M} z_{N,i}^{\sigma}(\varepsilon) \Big\rangle_{G_N'} - \frac{1}{M} \mathbb{E} \log \langle \exp \sqrt{M} y_N^{\sigma} \rangle_{G_N'} + o(1). \tag{6.19}$$

Here, $\langle \cdot \rangle_{G_N'}$ is the average in σ with respect to the perturbed Gibbs measure (3.5). By Lemma 5.7, the functionals appearing in (6.19) are continuous functionals of the distribution of the overlap matrix array

$$R^N = \big(R_{\ell,\ell'}^s\big)_{\ell,\ell' \ge 1, s \in \mathscr{S}} \tag{6.20}$$

of i.i.d. draws from the perturbed Gibbs measure $G_{d_N}^{\mathrm{pert}}$. We will now relate (6.19) to the Ruelle probability cascades allowing us to compute its value explicitly.

In the following, we take a subsequence along which the limit inferior of (6.19) is achieved. Since $\mathcal{R}$ is compact, we may take a subsequential weak limit of R^N along our minimizing subsequence, which we denote by R^∞. For ease of notation, we will continue to denote this subsequence with N. By the choice of the perturbation Hamiltonian, Theorem 3.2, we have the limiting array R^∞ satisfies equation (2.5). By the characterization theorem, Theorem 2.1, we see the order parameter for this system, that is the quantity which determines the law of the system, will be the law of $\mathrm{tr}(\bar{R}_{12}^\infty)$.

With this in mind, we make the following approximation. The array, $(\mathrm{tr}(\bar{R}_{\ell,\ell'}^\infty))_{\ell,\ell' \ge 1}$, by definition of $\mathcal{R}$, is a Gram–de Finetti array. Furthermore, taking $\nu_n^s = \sqrt{\rho^s} e_n$, $p_s = 1$, and $\varphi := \sum_{s \in \mathscr{S}} \sum_{i \le \kappa} (\nu_i^s, R_{\ell,\ell'}^{s,\infty} \nu_i^s)$ in (2.5) the array $(\mathrm{tr}(\bar{R}_{\ell,\ell'}^\infty))_{\ell,\ell' \ge 1}$ also satisfies the classical Ghirlanda–Guerra identities. By [40], Thereom 2.13, the law of $(\mathrm{tr}(\bar{R}_{\ell,\ell'}^\infty))_{\ell,\ell' \ge 1}$ is uniquely determined by ζ, the law of $\mathrm{tr}(\bar{R}_{1,2}^\infty)$. Let $\zeta^n \to \zeta$ weakly in $\Pr[0,1]$, such that ζ^n consists of a finite number of atoms. This yields sequences

$$\begin{aligned} x_{-1}^n = 0 < x_0^n < \cdots < x_{r-1}^n < x_r^n = 1, \\ 0 = q_0^n < \cdots < q_{r-1}^n < q_r^n = 1, \end{aligned} \tag{6.21}$$

such that $\zeta^n([0, q_p^n]) = x_p^n$. By [40], Theorem 2.17, if $\mathrm{tr}(\bar{R}_{1,2}^n)$ has distribution ζ^n, then the approximating array of traces $(\mathrm{tr}(\bar{R}_{1,2}^n))_{\ell,\ell'\geq 1}$ converge weakly to $(\mathrm{tr}(\bar{R}_{\ell,\ell'}^\infty))_{\ell,\ell'\geq 1}$.

Let v_α be the weights of the Ruelle probability cascades associated with the sequence (x^n) in (6.21). If $(a_\ell)_{\ell\geq 1}$ are i.i.d. samples from $\mathbb{N}^r$ according to v_α, then it is well known [40], Section 3.6, that $T_{\ell,\ell'}^n = q^n_{\alpha^\ell\wedge\alpha^{\ell'}}$ will be close in distribution to $(\mathrm{tr}(\bar{R}_{\ell,\ell'}^\infty))_{\ell,\ell'\geq 1}$. More precisely,

$$(T_{\ell,\ell'}^n)_{\ell,\ell'\geq 1} \xrightarrow{d} (\mathrm{tr}(\bar{R}_{\ell,\ell'}^\infty))_{\ell,\ell'\geq 1}. \tag{6.22}$$

Since the limiting array R^∞ is IP-Invariant (2.5), by Theorem 2.1 we can find a family of Lipschitz functions $(\Phi_s)_{s\in\mathscr{S}}$ on $[0,1]$ such that

$$R^\infty = (R_{\ell,\ell'}^{s,\infty})_{\ell,\ell'\geq 1,s\in\mathscr{S}} = (\Phi_s(\mathrm{tr}(\bar{R}_{\ell,\ell'}^\infty)))_{\ell,\ell'\geq 1,s\in\mathscr{S}}.$$

Since Φ_s is Lipschitz, (6.22) implies the $\kappa\times\kappa$ matrices $Q_{\ell,\ell'}^{s,n} := \Phi_s(T_{\ell,\ell'}^n)$ will be close in distribution to $R_{\ell,\ell'}^{s,\infty}$. If we let $Q^n := (Q_{\ell,\ell'}^{s,n})_{\ell,\ell'\geq 1,s\in\mathscr{S}}$ denote the approximating array generated from i.i.d. samples under Ruelle probability cascades corresponding to order parameter ζ^n then (6.22) also implies

$$Q^n \xrightarrow{d} R^\infty. \tag{6.23}$$

For n sufficiently large, we will bound (6.19) arbitrarily closely with functionals of the infinite array Q^n, which we will now show. Let $((Z_{i,n}^\alpha(k))_{k\leq\kappa})_{i\leq M}$ and Y_n^α be centered Gaussian processes with covariances

$$\mathbb{E}Z_{i,n}^\alpha Z_{j,n}^\beta = 2\delta_{\{i=j\}}\sum_{t\in\mathscr{S}}\Delta_{s(i),t}^2\rho^t Q_{|\alpha\wedge\beta|}^{t,n}, \tag{6.24}$$

$$\mathbb{E}Y_n^\alpha Y_n^\beta = \sum_{s,t\in\mathscr{S}}\Delta_{s,t}^2\rho^s\rho^t(Q_{|\alpha\wedge\beta|}^{s,n}, Q_{|\alpha\wedge\beta|}^{t,n}). \tag{6.25}$$

Notice that the covariance structure in (6.24) and (6.25) depend on the overlap array in exactly the same way as (6.17) and (6.18). A direct application of the continuity in Lemma 5.7 and the convergence in distribution of Q^n in (6.23) will imply the following result.

LEMMA 6.2. *For every $\tilde{\varepsilon} > 0$, there is an $n > 0$ such that*

$$\begin{aligned}\liminf_{N\to\infty}\frac{1}{M}\mathbb{E}\log\Big\langle\sum_{\varepsilon\in\Sigma_M(\delta)}\exp\sum_{i\leq M}z_{N,i}^\sigma(\varepsilon)\Big\rangle_{G_N'} - \frac{1}{M}\mathbb{E}\log\langle\exp\sqrt{M}y_N^\sigma\rangle_{G_N'}\\ \geq \frac{1}{M}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r}v_\alpha\sum_{\varepsilon\in\Sigma_M(\delta)}\exp\sum_{i\leq M}Z_{i,n}^\alpha(\varepsilon_i)\\ - \frac{1}{M}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r}v_\alpha\exp\sqrt{M}Y_n^\alpha - \tilde{\varepsilon}.\end{aligned} \tag{6.26}$$

PROOF. The covariance structure of (6.24) and (6.25) are identical to (6.17) and (6.18). By Lemma 5.7, the functionals appearing in (6.26) can be approximated by the same bounded continuous functions $\mathbb{E}g_\varepsilon^Z(R)$ and $\mathbb{E}g_\varepsilon^Y(R)$ of the overlap arrays. Since the distribution of the arrays R^N and Q^n both converge weakly to R^∞, by first taking a subsequence N along which R^N converges to R^∞ and then approximating with Q^n we have for any $\tilde{\varepsilon} > 0$,

$$\left|\mathbb{E}g_\varepsilon^Z(R^\infty) - \mathbb{E}g_\varepsilon^Y(R^\infty) - \mathbb{E}g_\varepsilon^Z(Q^n) + \mathbb{E}g_\varepsilon^Y(Q^n)\right| \le \tilde{\varepsilon},$$

by choosing n sufficiently large. Applying the triangle inequality will complete the proof. □

Since $\Sigma_M(\delta_M) = \prod_{s\in\mathscr{S}} \Sigma_{M_s}(\delta_{M_s})$, we express the first term in (6.26) as a weighted average. By the properties of the Ruelle probability cascades (see [40], Theorem 2.9),

$$\begin{aligned}
&\frac{1}{M}\mathbb{E}\log \sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\varepsilon\in\Sigma_M(\delta)} \exp \sum_{i\le M} Z_{i,n}^\alpha(\varepsilon_i) \\
&\quad = \frac{1}{M}\mathbb{E}\log \sum_{\alpha\in\mathbb{N}^r} v_\alpha \prod_{s\in\mathscr{S}} \sum_{\varepsilon\in\Sigma_{M_s}(\delta_{M_s})} \exp \sum_{i\in M_s} Z_{i,n}^\alpha(\varepsilon_i) \\
&\quad = \sum_{s\in\mathscr{S}} \rho_M^s \frac{1}{M_s}\mathbb{E}\log \sum_{\alpha\in\mathbb{N}^r} v_\alpha \sum_{\varepsilon\in\Sigma_{M_s}(\delta_{M_s})} \exp \sum_{i\in M_s} Z_{i,n}^\alpha(\varepsilon_i).
\end{aligned}$$

Every sequence $(x_p^n)_{p=0}^r$ and $(Q_p^{n,s})_{p=0}^r$ defines a discrete path $(\pi_n^s(x))_{s\in\mathscr{S}} \in \Pi$. Using the notation of the functionals $f_{M_s}^Z(\Sigma_{M_s}(\delta_{M_s}), s; \boldsymbol{\pi}_n)$ and $f_M^Y(\boldsymbol{\pi}_n)$ defined on (5.7) and (5.8), we have shown that

$$\liminf_{N\to\infty} \frac{1}{N}\mathbb{E}\log Z_N(d_N) \ge \left(\sum_{s\in\mathscr{S}} \rho_M^s f_{M_s}^Z(\Sigma_{M_s}(\delta_{M_s}), s; \boldsymbol{\pi}_n) - f_M^Y(\boldsymbol{\pi}_n)\right) - \varepsilon.$$

Lemma 5.4 and Lemma 5.5 imply the functionals $f_{M_s}^Z(\Sigma_{M_s}(\delta_{M_s}), s; \boldsymbol{\pi}_n)$ and $f_M^Y(\boldsymbol{\pi}_n)$ are Lipschitz. Sending $\varepsilon \to 0$ and noticing that the paths $\boldsymbol{\pi}_n \to \boldsymbol{\pi}_\infty$ in Δ, we have shown

$$\begin{aligned}
&\liminf_{N\to\infty} \frac{1}{N}\mathbb{E}\log Z_N(d_N) \\
&\quad \ge \left(\sum_{s\in\mathscr{S}} \rho_M^s f_{M_s}^Z(\Sigma_{M_s}(\delta_{M_s}), s; \boldsymbol{\pi}_\infty) - f_M^Y(\boldsymbol{\pi}_\infty)\right) \qquad \text{for any } M > 0.
\end{aligned} \tag{6.27}$$

All that remains is to remove the dependence on $\Sigma_M(\delta_M)$. This will be a direct application of the decoupling proved in Lemma 5.6.

LEMMA 6.3. *There exists a path $\boldsymbol{\pi}^* \in \Pi$ such that*

$$\lim_{M\to\infty}\Big(\sum_{s\in\mathscr{S}} \rho_M^s f_{M_s}^Z(s,\boldsymbol{\pi}^*) - f_M^Y(\boldsymbol{\pi}^*)\Big) \geq \mathscr{P}\big(r,x,d,(\lambda^s,Q^s)_{s\in\mathscr{S}}\big) - \varepsilon.$$

PROOF. We now make the paths $\boldsymbol{\pi}_\infty^M$ dependence on M explicit. Since $(\boldsymbol{\pi}_\infty^M)_{M\geq 1}$ is a countable collection of bounded monotone paths of $\kappa\times\kappa$ matrices, there exists a subsequence in M such that $\boldsymbol{\pi}_\infty^M \to \boldsymbol{\pi}^*$. Given $\varepsilon > 0$, by Lemma 5.6 we can find a discrete path $\boldsymbol{\pi}_\varepsilon$ such that $\Delta(\boldsymbol{\pi}_\varepsilon, \boldsymbol{\pi}^*) < C\varepsilon$, where C is the maximum Lipschitz constant over all $f_{M_s}^Z(s)$. Since $\boldsymbol{\pi}_\varepsilon$ is discrete, applying Section 5.1 to the summation appearing in (6.27) shows

$$\begin{aligned}\lim_{M\to\infty}\sum_{s\in\mathscr{S}} \rho_M^s f_{M_s}^Z(s,\boldsymbol{\pi}^*) &\geq \lim_{M\to\infty}\sum_{s\in\mathscr{S}} \rho_M^s f_M^Z(\boldsymbol{\pi}_\varepsilon) - C\varepsilon\\ &\geq \inf_{\lambda}\sum_{s\in\mathscr{S}} \rho^s\Big(-\sum_{k\leq\kappa}\lambda_k^s d_k^s + X_0^s(\boldsymbol{\pi}_\varepsilon)\Big) - C\varepsilon\end{aligned}$$

since $\rho_M^s \to \rho^s$, $\delta_{M,k}^s \to d_k^s$ and f_M^z is Lipschitz.

The term $f_M^Y(\boldsymbol{\pi}_\infty)$ is actually independent of M, and a similar computation using the properties of Ruelle probability cascades like in (4.13) shows

$$\lim_{M\to\infty} f_M^Y(\boldsymbol{\pi}^*) = \frac{1}{2}\sum_{\ell=0}^{r-1} x_\ell \sum_{s,t\in\mathscr{S}} \Delta_{st}^2\rho^s\rho^t\big((Q_{\ell+1}^s, Q_{\ell+1}^t) - (Q_\ell^s, Q_\ell^t)\big).$$

Combining the computations of the two terms above, we now take $\varepsilon\to 0$ and get

$$\liminf_{N\to\infty}\frac{1}{N}\mathbb{E}\log Z_N(d_N) \geq \mathscr{P}\big(d,\lambda,\boldsymbol{\pi}^*\big) \tag{6.28}$$

since $\boldsymbol{\pi}_\varepsilon \to \boldsymbol{\pi}^*$ as $\varepsilon\to 0$. $\square$

Since the path $\boldsymbol{\pi}^*$ can be described as the limit of the discrete approximating sequences (1.18) and (1.19), we have shown

$$\liminf_{N\to\infty}\frac{1}{N}\mathbb{E}\log Z_N(d_N) \geq \inf_{x,r,(\lambda^s,Q^s)_{s\in\mathscr{S}}} \mathscr{P}\big(r,x,d,(\lambda^s,Q^s)_{s\in\mathscr{S}}\big).$$

7. Proof of Theorem 1.2.

PROOF OF THEOREM 1.2. We begin by proving part (1) of the theorem.

We start with the lower bound of $F_N(\Sigma_N^{\varepsilon_N}(d))$. For any $d\in\mathscr{D}$, let us choose $\varepsilon_N = \frac{L_\kappa}{N}$. We choose a sequence of $d_N \in \mathscr{D}_N$ such that $\|d_N - d\|_\infty = \varepsilon_N$ to satisfy

the condition in (6.5). Since $\Sigma_N^{\varepsilon}(d) \supseteq \Sigma_N(d_N)$, the lower bound (6.1) computed in Section 6 implies

$$\begin{aligned} \liminf_{N\to\infty} \frac{1}{N}\mathbb{E}\log Z_N^{\varepsilon_N}(d) &\geq \liminf_{N\to\infty} \frac{1}{N}\mathbb{E}\log Z_N(d_N) \\ &\geq \inf_{x,r,(\lambda^s,Q^s)_{s\in\mathscr{S}}} \mathscr{P}\big(r,x,d,(\lambda^s,Q^s)_{s\in\mathscr{S}}\big). \end{aligned} \tag{7.1}$$

We now obtain the matching upper bound of $F_N(\Sigma_N^{\varepsilon_N}(d))$. Since $\varepsilon_N \to 0$, the $O(\varepsilon)$ term in the upper bound (4.14) computed in Section 4 vanishes yielding

$$\limsup_{N\to\infty} \frac{1}{N}\mathbb{E}\log Z_N^{\varepsilon_N}(d) \leq \inf_{x,r,(\lambda^s,Q^s)_{s\in\mathscr{S}}} \mathscr{P}\big(r,x,d,(\lambda^s,Q^s)_{s\in\mathscr{S}}\big). \tag{7.2}$$

Combining the inequalities in (7.1) and (7.2), we arrive at the formula for the constrained free energy

$$\lim_{N\to\infty} \frac{1}{N}\mathbb{E}\log Z_N(d) = \inf_{x,r,(\lambda^s,Q^s)_{s\in\mathscr{S}}} \mathscr{P}\big(r,x,d,(\lambda^s,Q^s)_{s\in\mathscr{S}}\big). \tag{7.3}$$

Part (2) of the Theorem is a direct consequence of part (1). By classical Gaussian concentration inequalities, observe that the limit of the free energy $F_N(\Sigma_N)$ is asymptotically given by the supremum of $F_N(\Sigma_N(d))$ over $d \in \mathscr{D}$. By taking the supremum over $d \in \mathscr{D}$ in (7.3), we arrive at the formula for the free energy

$$\lim_{N\to\infty} \frac{1}{N}\mathbb{E}\log Z_N = \sup_{d\in\mathscr{D}} \inf_{x,r,(\lambda^s,Q^s)_{s\in\mathscr{S}}} \mathscr{P}\big(r,x,d,(\lambda^s,Q^s)_{s\in\mathscr{S}}\big). \tag{7.4}$$

□

PROOF OF COROLLARY 1.3. We include a proof of the corollary for the sake of completeness. By Hölder's inequality, $\frac{F_N^{\beta}(\Sigma_N^{\varepsilon_N}(d))}{\beta}$ is increasing in β and, therefore,

$$\lim_{\beta\to\infty} \frac{1}{\beta} \inf_{x,r,(\lambda^s,Q^s)_{s\in\mathscr{S}}} \mathscr{P}_{\beta}\big(r,x,d,(\lambda^s,Q^s)_{s\in\mathscr{S}}\big) \tag{7.5}$$

exists. Moreover, we have

$$\frac{1}{N}\mathbb{E}\Big[\max_{\Sigma_N^{\varepsilon_N}(d)} H_N(\sigma)\Big] \leq \frac{F_N^{\beta}(\Sigma_N^{\varepsilon_N}(d))}{\beta} \leq \frac{\log q}{\beta} + \frac{1}{N}\mathbb{E}\Big[\max_{\Sigma_N^{\varepsilon_N}(d)} H_N(\sigma)\Big].$$

Taking the limit as $N \to \infty$ and then as $\beta \to \infty$ completes the proof. □

8. Proofs regarding cuts problems. In this section, we prove Theorem 1.1. We start with a proof of Lemma 1.4.

PROOF OF LEMMA 1.4. The proof of this lemma proceeds by a direct coupling argument. We will denote the adjacency matrix corresponding to $\tilde{G}_N$ as

$\tilde{A} = (\tilde{A}_{i,j})$. We will realize the two graphs on the same probability space and couple the adjacency matrices such that

$$\begin{aligned}
&\mathbb{P}[A_{i,j} \neq \tilde{A}_{i,j}] \\
&\quad = \left|\mathbb{P}[A_{i,j} = 1] - \mathbb{P}[\tilde{A}_{i,j} = 1]\right| \\
&\quad \leq cN\left|\int_{[\frac{i-1}{N}, \frac{i}{N}]\times[\frac{j-1}{N}, \frac{j}{N}]} K(x, y)\,\mathrm{d}x\,\mathrm{d}y - \int_{[\frac{i-1}{N}, \frac{i}{N}]\times[\frac{j-1}{N}, \frac{j}{N}]} K_1(x, y)\,\mathrm{d}x\,\mathrm{d}y\right|.
\end{aligned}$$

We note that for any two graphs G_N and $\tilde{G}_N$, we have

$$\left|\frac{\mathsf{MaxCut}_\kappa(G_N)}{N} - \frac{\mathsf{MaxCut}_\kappa(\tilde{G}_N)}{N}\right| \leq \frac{1}{2N}\sum_{i,j=1}^{N} |A_{i,j} - \tilde{A}_{i,j}|.$$

This implies that

$$\begin{aligned}
&\left|\mathbb{E}\left[\frac{\mathsf{MaxCut}_\kappa(G_N)}{N}\right] - \mathbb{E}\left[\frac{\mathsf{MaxCut}_\kappa(\tilde{G}_N)}{N}\right]\right| \\
&\quad \leq \frac{c}{2}\sum_{i,j=1}^{N}\left|\int_{[\frac{i-1}{N}, \frac{i}{N}]\times[\frac{j-1}{N}, \frac{j}{N}]} K(x, y)\,\mathrm{d}x\,\mathrm{d}y - \int_{[\frac{i-1}{N}, \frac{i}{N}]\times[\frac{j-1}{N}, \frac{j}{N}]} K_1(x, y)\,\mathrm{d}x\,\mathrm{d}y\right| \\
&\quad = \frac{c}{2}\big(\|K - K_1\|_1 + o(1)\big).
\end{aligned}$$

Thus the proof is complete once we choose M sufficiently large such that $\|K - K_1\|_1 \leq 1/c^{1/2+\delta}$. $\square$

To complete the proof of Theorem 1.1, we first introduce the following Gaussian optimization problem. Set $J = (J_{i,j})$ a symmetric matrix such that $\{J_{i,j} : i \leq j\}$ are independent $N(0, \tilde{K}_N(i, j))$ random variables. We define

$$\begin{aligned}
\tilde{Z}_N = \frac{1}{2N}\max_{\sigma\in[\kappa]^N}\Bigg[&\frac{c}{N}\sum_{i,j=1}^{N}\tilde{K}_N(i, j)\,\mathbb{1}(\sigma_i \neq \sigma_j) \\
&+ \sqrt{c}\sum_{i,j=1}^{N}\frac{J_{i,j}}{\sqrt{N}}\,\mathbb{1}(\sigma_i \neq \sigma_j)\Bigg].
\end{aligned} \tag{8.1}$$

The following lemma establishes that in the "large degree" limit, we can study the asymptotic behavior of the MaxCut_κ problem via that of the Gaussian optimization problem $\tilde{Z}_N$.

LEMMA 8.1. *As $N \to \infty$, we have*

$$\mathbb{E}\left[\frac{\mathsf{MaxCut}_\kappa(G_N)}{N}\right] = \mathbb{E}[\tilde{Z}_N] + o(\sqrt{c}).$$

PROOF. The lemma follows from a direct application of [48], Theorem 1.1. Specifically, we note that

$$\frac{\mathsf{MaxCut}_\kappa(G_N)}{N} = \frac{1}{2N}\max_{\sigma\in[\kappa]^N}\sum_{i,j=1}^N A_{i,j}\,\mathbb{1}(\sigma_i\neq\sigma_j).$$

The proof follows directly upon an application of the result in [48]. □

Thus finally, it comes down to the study of $\mathbb{E}[\tilde{Z}_n]$ as $N\to\infty$. Fix a probability distribution on $[\kappa]$, $d^s=(d_1^s,\dots,d_\kappa^s)$, $s=1,\dots,M$. Consider the Hamiltonian

$$H(\sigma) = \sum_{i,j=1}^N \frac{J_{i,j}}{\sqrt{N}}\,\mathbb{1}(\sigma_i=\sigma_j), \tag{8.2}$$

where $J=(J_{i,j})$ is as described above. For $d=(d^1,\dots,d^M)$, probability measures on $[\kappa]$, and a sequence ε_N decaying to zero sufficiently slowly, recall the restricted configuration space (1.13)

$$\Sigma_N^{\varepsilon_N}(d) = \Big\{\sigma\in[\kappa]^N \Big| \sum_{i\in I_s}\frac{\mathbb{1}(\sigma_i=k)}{N_s}\in\big[d_k^s-\varepsilon_N, d_k^s+\varepsilon_N\big], s\in[M]\Big\}. \tag{8.3}$$

Recall the restricted ground state energy $\mathcal{P}(d)$, introduced in Corollary 1.3. This will allow us to deduce the following lemma.

LEMMA 8.2.

$$\lim_{N\to\infty}\mathbb{E}[\tilde{Z}_N] = \sup_d\left[\frac{c}{2}\sum_{s,t=1}^M \mathbf{K}(s,t)\rho^s\rho^t\big(1-\langle d^s,d^t\rangle\big)+\frac{\sqrt{c}}{2}\mathcal{P}(d)\right]+o(\sqrt{c}).$$

PROOF. We start with the lower bound. We define

$$\begin{aligned}\tilde{Z}_N(d) &= \frac{1}{2N}\max_{\sigma\in\Sigma_N^{\varepsilon_N}(d)}\left[\frac{c}{N}\sum_{i,j=1}^N \tilde{K}_N(i,j)\,\mathbb{1}(\sigma_i\neq\sigma_j)+\sqrt{c}\sum_{i,j=1}^N\frac{J_{i,j}}{\sqrt{N}}\,\mathbb{1}(\sigma_i\neq\sigma_j)\right]\\ &= \frac{c}{2}\sum_{s,t=1}^M \mathbf{K}(s,t)\rho^s\rho^t\big(1-\langle d^s,d^t\rangle\big)+\frac{\sqrt{c}}{2N}\max_{\sigma\in\Sigma_N^{\varepsilon_N}(d)}\sum_{i,j=1}^N\frac{J_{i,j}}{\sqrt{N}}\,\mathbb{1}(\sigma_i\neq\sigma_j)\\ &\quad+o(1).\end{aligned}$$

For fixed probability vectors $d^1,\dots,d^M$, we have

$$\begin{aligned}\liminf_{N\to\infty}\mathbb{E}[\tilde{Z}_N] &\geq \liminf_{N\to\infty}\mathbb{E}\big[\tilde{Z}_N(d)\big]\\ &= \frac{c}{2}\sum_{s,t=1}^M \mathbf{K}(s,t)\rho^s\rho^t\big(1-\langle d^s,d^t\rangle\big)+\frac{\sqrt{c}}{2}\mathcal{P}(d).\end{aligned}$$

We take the supremum over all possible probability vectors $d^1, \dots, d^M$ to get the requisite lower bound. To establish the upper bound, we define

$$\tilde{M} = \sup_{d} \left[\frac{c}{2} \sum_{s,t=1}^{M} \mathbf{K}(s,t) \rho^s \rho^t \big(1 - \langle d^s, d^t \rangle\big) + \frac{\sqrt{c}}{2} \mathcal{P}(d) \right]. \tag{8.4}$$

Therefore, we have

$$\begin{aligned} \mathbb{E}\big[\tilde{Z}_N(d)\big] &\le \frac{c}{2} \sum_{s,t=1}^{M} \mathbf{K}(s,t) \rho^s \rho^t \big(1 - \langle d^s, d^t \rangle\big) + \frac{\sqrt{c}}{2} \mathcal{P}(d) + o(1) \\ &\le \tilde{M} + o(1), \end{aligned}$$

uniformly over all choices of $d^1, \dots, d^M$. We note that empirical distributions within each block may assume only finitely many values and hence, summing over these values, we have

$$\mathbb{P}[\tilde{Z}_N > \tilde{M} + t] = \sum_{d^1, \dots, d^M} \mathbb{P}\big[\tilde{Z}_N(d) > \tilde{M} + t\big].$$

Now,

$$\mathbb{P}\big[\tilde{Z}_N(d) > \tilde{M} + t\big] \le \mathbb{P}\big[\tilde{Z}_N(d) - \mathbb{E}\big[\tilde{Z}_N(d)\big] > t\big] \le \exp\big[-CNt^2\big],$$

where the last inequality follows by Gaussian concentration. Finally, plugging this tail bound, we have

$$\mathbb{P}[\tilde{Z}_N > \tilde{M} + t] \le A N^{\kappa \tilde{M}} \exp\big[-CNt^2\big].$$

We note that

$$\begin{aligned} \mathbb{E}[\tilde{Z}_N] &\le \int_0^\infty \mathbb{P}[\tilde{Z}_N > x] \, \mathrm{d}x \le \tilde{M} + \delta_N + \int_{\delta_N}^\infty \mathbb{P}[\tilde{Z}_N > \tilde{M} + t] \, \mathrm{d}t \\ &\le \tilde{M} + \delta_N + A N^{\kappa \tilde{M}} \frac{\exp[-CN\delta_N^2]}{\sqrt{N}\delta_N}. \end{aligned}$$

Finally, we choose $\delta_N = C_0 \sqrt{\frac{\log N}{N}}$ for some constant C_0 sufficiently large. This establishes that

$$\mathbb{E}[\tilde{Z}_N] \le \tilde{M} + o(1),$$

completing the proof of the upper bound. $\square$

PROOF OF THEOREM 1.1. The theorem follows directly upon combining Lemma 8.1 and Lemma 8.2. $\square$

REFERENCES

[1] AIELLO, W., CHUNG, F. and LU, L. (2001). A random graph model for power law graphs. *Exp. Math.* **10** 53–66. MR1822851

[2] AIZENMAN, M., SIMS, R. and STARR, S. L. (2003). Extended variational principle for the Sherrington–Kirkpatrick spin-glass model. *Phys. Rev. B* **68** 214403.

[3] ALBERT, R. and BARABÁSI, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Modern Phys.* **74** 47–97. MR1895096

[4] ARGUIN, L.-P. and AIZENMAN, M. (2009). On the structure of quasi-stationary competing particle systems. *Ann. Probab.* **37** 1080–1113. MR2537550

[5] AUFFINGER, A. and CHEN, W.-K. (2017). Parisi formula for the ground state energy in the mixed p-spin model. *Ann. Probab.* **45** 4617–4631. MR3737919

[6] AUFFINGER, A. and CHEN, W.-K. (2014). Free energy and complexity of spherical bipartite models. *J. Stat. Phys.* **157** 40–59. MR3249903

[7] BARRA, A., CONTUCCI, P., MINGIONE, E. and TANTARI, D. (2015). Multi-species mean field spin glasses. Rigorous results. *Ann. Henri Poincaré* **16** 691–708. MR3311887

[8] BOLLOBÁS, B., JANSON, S. and RIORDAN, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms* **31** 3–122. MR2337396

[9] BORGS, C., CHAYES, J. T., LOVÁSZ, L., SÓS, V. T. and VESZTERGOMBI, K. (2008). Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.* **219** 1801–1851. MR2455626

[10] BORGS, C., CHAYES, J. T., LOVÁSZ, L., SÓS, V. T. and VESZTERGOMBI, K. (2012). Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. of Math.* (2) **176** 151–219. MR2925382

[11] BOUCHERON, S., LUGOSI, G. and MASSART, P. (2013). *Concentration Inequalities*: *A Nonasymptotic Theory of Independence*, *with a Foreword by Michel Ledoux*. Oxford Univ. Press, Oxford. MR3185193

[12] BRITTON, T., DEIJFEN, M. and MARTIN-LÖF, A. (2006). Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.* **124** 1377–1397. MR2266448

[13] CALTAGIRONE, F., PARISI, G. and RIZZO, T. (2012). Dynamical critical exponents for the mean-field Potts glass. *Phys. Rev. E* **85** 051504.

[14] CHEN, W.-K. and SEN, A. (2017). Parisi formula, disorder chaos and fluctuation for the ground state energy in the spherical mixed p-spin models. *Comm. Math. Phys.* **350** 129–173. MR3606472

[15] CHUNG, F. and LU, L. (2002). Connected components in random graphs with given expected degree sequences. *Ann. Comb.* **6** 125–145. MR1955514

[16] DECELLE, A., KRZAKALA, F., MOORE, C. and ZDEBOROVÁ, L. (2011). Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. *Phys. Rev. E* **84** 066106.

[17] DEMBO, A., MONTANARI, A. and SEN, S. (2017). Extremal cuts of sparse random graphs. *Ann. Probab.* **45** 1190–1217. MR3630296

[18] DOROGOVTSEV, S. N. and MENDES, J. F. F. (2002). Evolution of networks. *Advances in Physics* **51** 1079–1187.

[19] DURRETT, R. (2003). Rigorous result for the CHKNS random graph model. In *Discrete Random Walks* (*Paris*, 2003). *Discrete Math. Theor. Comput. Sci. Proc., AC* 95–104. Assoc. Discrete Math. Theor. Comput. Sci., Nancy. MR2042377

[20] ELDERFIELD, D. and SHERRINGTON, D. (1983). The curious case of the Potts spin glass. *J. Phys. C, Solid State Phys.* **16** L497.

[21] ELDERFIELD, D. and SHERRINGTON, D. (1983). Novel non-ergodicity in the Potts spin glass. *J. Phys. C, Solid State Phys.* **16** L1169.

[22] FU, Y. and ANDERSON, P. W. (1986). Application of statistical mechanics to NP-complete problems in combinatorial optimisation. *J. Phys. A: Math. Gen.* **19** 1605.
[23] GHIRLANDA, S. and GUERRA, F. (1998). General properties of overlap probability distributions in disordered spin systems. towards parisi ultrametricity. *J. Phys. A: Math. Gen.* **31** 9149.
[24] GUERRA, F. (2003). Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.* **233** 1–12. MR1957729
[25] HOLLAND, P. W., LASKEY, K. B. and LEINHARDT, S. (1983). Stochastic blockmodels: First steps. *Soc. Netw.* **5** 109–137.
[26] JACKSON, M. O. (2008). *Social and Economic Networks*. Princeton Univ. Press, Princeton, NJ. MR2435744
[27] JAGANNATH, A. and TOBASCO, I. (2017). Low temperature asymptotics of spherical mean field spin glasses. *Comm. Math. Phys.* **352** 979–1017. MR3631397
[28] KAHNG, A. B., LIENIG, J., MARKOV, I. L. and HU, J. (2011). *VLSI Physical Design: From Graph Partitioning to Timing Closure*. Springer, New York.
[29] KALIKOW, S. and WEISS, B. (1988). When are random graphs connected. *Israel J. Math.* **62** 257–268. MR0955131
[30] KIRKPATRICK, S., GELATT, C. D. JR. and VECCHI, M. P. (1983). Optimization by simulated annealing. *Science* **220** 671–680. MR0702485
[31] ŁUCZAK, T. (1992). Sparse random graphs with a given degree sequence. In *Random Graphs, Vol.* 2 (*Poznań*, 1989). 165–182. Wiley, New York. MR1166614
[32] MASSOULIÉ, L. (2014). Community detection thresholds and the weak Ramanujan property. In *STOC'*14*—Proceedings of the* 2014 *ACM Symposium on Theory of Computing* 694–703. ACM, New York. MR3238997
[33] MÉZARD, M. and MONTANARI, A. (2009). *Information, Physics, and Computation*. Oxford Univ. Press, Oxford. MR2518205
[34] MÉZARD, M., PARISI, G. and VIRASORO, M. (1987). *Spin Glass Theory and Beyond: An Introduction to the Replica Method and Its Applications* **9**. World Scientific, Singapore.
[35] MOSSEL, E., NEEMAN, J. and SLY, A. (2013). A proof of the block model threshold conjecture. Preprint. Available at arXiv:1311.4115.
[36] NISHIMORI, H. and STEPHEN, M. J. (1983). Gauge-invariant frustrated Potts spin-glass. *Phys. Rev. B* **27** 5644–5652.
[37] NORROS, I. and REITTU, H. (2006). On a conditionally Poissonian graph process. *Adv. in Appl. Probab.* **38** 59–75. MR2213964
[38] PANCHENKO, D. (2013). Spin glass models from the point of view of spin distributions. *Ann. Probab.* **41** 1315–1361. MR3098679
[39] PANCHENKO, D. (2013). The Parisi ultrametricity conjecture. *Ann. of Math.* (2) **177** 383–393. MR2999044
[40] PANCHENKO, D. (2013). *The Sherrington–Kirkpatrick Model*. Springer, New York. MR3052333
[41] PANCHENKO, D. (2014). The Parisi formula for mixed p-spin models. *Ann. Probab.* **42** 946–958. MR3189062
[42] PANCHENKO, D. (2018). Free energy in the mixed p-spin models with vector spins. *Ann. Probab.* **46** 865–896. MR3773376
[43] PANCHENKO, D. (2018). Free energy in the Potts spin glass. *Ann. Probab.* **46** 829–864. MR3773375
[44] PANCHENKO, D. (2015). The free energy in a multi-species Sherrington-Kirkpatrick model. *Ann. Probab.* **43** 3494–3513. MR3433586
[45] POLJAK, S. and TUZA, Z. (1993). *The Max-Cut Problem—a Survey*. Acad. Sinica, Taipei.
[46] PORTER, M. A., ONNELA, J.-P. and MUCHA, P. J. (2009). Communities in networks. *Notices Amer. Math. Soc.* **56** 1082–1097. MR2568495

[47] RUELLE, D. (1987). A mathematical reformulation of Derrida's REM and GREM. *Comm. Math. Phys.* **108** 225–239. MR0875300

[48] SEN, S. (2016). Optimization on sparse random hypergraphs and spin glasses. Preprint. Available at arXiv:1606.02365.

[49] SHEPP, L. A. (1989). Connectedness of certain random graphs. *Israel J. Math.* **67** 23–33. MR1021358

[50] SHERRINGTON, D. and KIRKPATRICK, S. (1975). Solvable model of a spin-glass. *Phys. Rev. Lett.* **35** 1792–1796.

[51] SÖDERBERG, B. (2002). General formalism for inhomogeneous random graphs. *Phys. Rev. E* (3) **66** 066121, 6. MR1953933

[52] TALAGRAND, M. (2010). *Mean Field Models for Spin Glasses*: *Volume I*: *Basic Examples* **54**. Springer, New York.

A. JAGANNATH
J. KO
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TORONTO
TORONTO, ONTARIO M5S 2E4
CANADA
E-MAIL: aukosh@math.toronto.edu
jko@math.toronto.edu

S. SEN
DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305-4020
USA
E-MAIL: ssen90@stanford.edu