The Annals of Probability
2015, Vol. 43, No. 6, 3494–3513
DOI: 10.1214/14-AOP967

THE FREE ENERGY IN A MULTI-SPECIES SHERRINGTON–KIRKPATRICK MODEL

BY DMITRY PANCHENKO

University of Toronto

The authors of [*Ann. Henri Poincaré* **16** (2015) 691–708] introduced a multi-species version of the Sherrington–Kirkpatrick model and suggested the analogue of the Parisi formula for the free energy. Using a variant of Guerra's replica symmetry breaking interpolation, they showed that, under certain assumption on the interactions, the formula gives an upper bound on the limit of the free energy. In this paper we prove that the bound is sharp. This is achieved by developing a new multi-species form of the Ghirlanda–Guerra identities and showing that they force the overlaps within species to be completely determined by the overlaps of the whole system.

1. Introduction and main results. Recently, the following modification of the Sherrington–Kirkpatrick model [15] was introduced in [3]. Given $N \geq 1$, let us denote by

$$\sigma = (\sigma_1, \dots, \sigma_N) \in \Sigma_N = \{-1, +1\}^N \tag{1}$$

a configuration of N Ising spins. Consider a finite set $\mathscr{S}$ that will be fixed throughout the paper and, in particular, it does not change with N. We emphasize this because we will often omit the dependence of other objects on N. The elements of $\mathscr{S}$ will be called species and will be denoted by s or t. Let us divide all spin indices into disjoint groups indexed by the species

$$I = \{1, \dots, N\} = \bigcup_{s \in \mathscr{S}} I_s. \tag{2}$$

These sets will, obviously, vary with N, and we will assume that their cardinalities $N_s = |I_s|$ satisfy

$$\lim_{N \to \infty} \frac{N_s}{N} = \lambda_s \in (0, 1) \qquad \text{for all } s \in \mathscr{S}. \tag{3}$$

For simplicity of notation, we will omit the dependence of $\lambda_s^N := N_s/N$ on N and will simply write λ_s. The Hamiltonian proposed in [3] resembles the usual SK Hamiltonian,

$$H_N(\sigma) = \frac{1}{\sqrt{N}} \sum_{i,j=1}^{N} g_{ij} \sigma_i \sigma_j, \tag{4}$$

Received February 2014; revised August 2014.
MSC2010 subject classifications. 60K35, 60G09, 82B44.
Key words and phrases. Spin glasses, Sherrington–Kirkpatrick model.

where the interaction parameters (g_{ij}) are independent Gaussian random variables, only now they are not necessarily identically distributed but, instead, satisfy

$$\mathbb{E}g_{ij}^2 = \Delta_{st}^2 \qquad \text{if } i \in I_s,\, j \in I_t \text{ for } s, t \in \mathscr{S}. \tag{5}$$

In other words, the variance of the interaction between i and j depends only on the species they belong to. We will make the same assumptions on the matrix $\Delta^2 = (\Delta_{st}^2)_{s,t\in\mathscr{S}}$ as in [3], namely, that it is symmetric and nonnegative definite,

$$\Delta_{st}^2 = \Delta_{ts}^2 \qquad \text{for all } s, t \in \mathscr{S} \text{ and } \Delta^2 \geq 0. \tag{6}$$

Let us denote the overlap of the restrictions of two spin configurations to a given species $s \in \mathscr{S}$ by

$$R_s(\sigma^1, \sigma^2) = \frac{1}{N_s} \sum_{i \in I_s} \sigma_i^1 \sigma_i^2. \tag{7}$$

Then it is easy to see that the covariance of the Gaussian Hamiltonian (4) is given by

$$\frac{1}{N}\mathbb{E}H_N(\sigma^1)H_N(\sigma^2) = \sum_{s,t\in\mathscr{S}} \Delta_{st}^2 \lambda_s \lambda_t R_s(\sigma^1, \sigma^2) R_t(\sigma^1, \sigma^2). \tag{8}$$

This already gives some idea about the main new difficulty one encounters in this model compared to the classical Sherrington–Kirkpatrick model. Namely, now we will need to understand the joint distributions of the overlap arrays in the thermodynamic limit simultaneously for all species $s \in \mathscr{S}$. Our main goal will be to compute the limit of the free energy in this model,

$$F_N = \frac{1}{N}\mathbb{E}\log Z_N, \qquad \text{where } Z_N = \sum_{\sigma\in\Sigma_N} \exp H_N(\sigma). \tag{9}$$

Notice that we do not consider the inverse temperature parameter here, because it can be absorbed into the definition of the matrix Δ^2. One can also consider the externals fields that depend only on the species but, since it does not affect any arguments in the paper, for simplicity of notation we will omit them.

Under assumption (6), the authors in [3] proved, using the Guerra–Toninelli interpolation [7], that the free energy has a limit. They also proposed the following analogue of the Parisi formula [11, 12] for the free energy, which was proved for the original SK model by Talagrand in [16]; see also [17]. Given integer $r \geq 1$, consider a sequence

$$0 = \zeta_{-1} < \zeta_0 < \cdots < \zeta_{r-1} < \zeta_r = 1 \tag{10}$$

and, for each $s \in \mathscr{S}$, a sequence

$$0 = q_0^s \leq q_1^s \leq \cdots \leq q_{r-1}^s \leq q_r^s = 1. \tag{11}$$

We will also consider two types of nondecreasing combinations of these sequences as follows. For $0 \le \ell \le r$, we define

$$Q_\ell = \sum_{s,t\in\mathscr{S}} \Delta_{st}^2 \lambda_s \lambda_t q_\ell^s q_\ell^t \quad \text{and} \quad Q_\ell^s = 2\sum_{t\in\mathscr{S}} \Delta_{st}^2 \lambda_t q_\ell^t \text{ for } s \in \mathscr{S}. \tag{12}$$

The meaning of these definitions will become clear when we look at the covariance of the cavity fields in the Aizenman–Sims–Starr scheme in Section 5. Given these sequences, let us consider i.i.d. standard Gaussian random variables $(\eta_\ell)_{1\le\ell\le r}$ and, for $s \in \mathscr{S}$, define

$$X_r^s = \log \operatorname{ch} \sum_{1\le\ell\le r} \eta_\ell \bigl(Q_\ell^s - Q_{\ell-1}^s\bigr)^{1/2}. \tag{13}$$

Recursively over $0 \le \ell \le r-1$, we define

$$X_\ell^s = \frac{1}{\zeta_\ell} \log \mathbb{E}_\ell \exp \zeta_\ell X_{\ell+1}^s, \tag{14}$$

where $\mathbb{E}_\ell$ denotes the expectation with respect to $\eta_{\ell+1}$ only. Notice that X_0^s are nonrandom. Finally, we define the analogue of the Parisi functional by

$$\mathscr{P}(\zeta, q) = \log 2 + \sum_{s\in\mathscr{S}} \lambda_s X_0^s - \frac{1}{2} \sum_{0\le\ell\le r-1} \zeta_\ell (Q_{\ell+1} - Q_\ell). \tag{15}$$

The main result of the paper is the following.

THEOREM 1. *Under the assumption* (6), *the limit of the free energy is given by*

$$\lim_{N\to\infty} F_N = \inf \mathscr{P}(\zeta, q), \tag{16}$$

where the infimum is taken over $r \ge 1$ *and the sequences* (10) *and* (11).

In [3], the inequality $F_N \le \inf \mathscr{P}(\zeta, q)$ was proved under assumption (6) using the analogue of Guerra's replica symmetry breaking interpolation [6]. For convenience, we will reproduce this result in Section 2 in the formalism of the Ruelle probability cascades, which will also allow us to introduce several objects that will be used in the subsequent sections. In this paper we will prove the matching lower bound using the analogue of the Aizenman–Sims–Starr scheme [1] and, in this part, the assumption $\Delta^2 \ge 0$ will not be needed. The approach was applied previously in various situations in [10] and [4] and is based on the ultrametricity result in [8]. As we mentioned above, in the multi-species model we encounter a new nontrivial obstacle. Namely, we need to describe the joint distribution of the overlap arrays simultaneously for all species, and even though it is clear that the marginal distribution of each array will be generated by the Ruelle probability cascades as in the SK model, it is not at all clear what their joint distribution should be. We will develop an approach to overcome this obstacle in Sections 3 and 4. In Section 3 we will prove a multi-species version of the Ghirlanda–Guerra identities, which are similar to the original Ghirlanda–Guerra identities [5], but apply to generic overlaps that may depend on the overlaps of all species. Using these identities, we will show in Section 4 that the overlaps of different species are synchronized in the sense that they are deterministic functions of the overlaps of the whole system. This will describe the joint distribution of all overlaps and allow us to obtain the lower bound in Section 5 in a straightforward way using the Aizenman–Sims–Starr scheme. In the last section, we will mention several interesting open questions.

2. Guerra's replica symmetry breaking bound. Given $r \geq 1$, let $(v_\alpha)_{\alpha\in\mathbb{N}^r}$ be the weights of the Ruelle probability cascades [14] corresponding to the parameters (10); see, for example, Section 2.3 in [9] for the definition. For $\alpha, \beta \in \mathbb{N}^r$, we denote

$$\alpha \wedge \beta = \min\{0 \leq \ell \leq r \,|\, \alpha_1 = \beta_1, \ldots, \alpha_\ell = \beta_\ell, \alpha_{\ell+1} \neq \beta_{\ell+1}\}, \tag{17}$$

where $\alpha \wedge \beta = r$ if $\alpha = \beta$. Since the sequences defined in (12) are nondecreasing, we can consider Gaussian processes $C^s(\alpha)$ for $s \in \mathscr{S}$ and $D(\alpha)$ both indexed by $\alpha \in \mathbb{N}^r$ with the covariances

$$\mathbb{E}C^s(\alpha)C^s(\beta) = Q^s_{\alpha\wedge\beta} \quad \text{and} \quad \mathbb{E}D(\alpha)D(\beta) = Q_{\alpha\wedge\beta}. \tag{18}$$

These are the usual Gaussian fields that accompany the construction of the Ruelle probability cascades; see, for example, Section 2.3 in [9]. For each $s \in \mathscr{S}$ and each $i \in I_s$, let $C_i(\alpha)$ be a copy of the process $C^s(\alpha)$, and suppose that all these processes are independent of each other and of $D(\alpha)$. For $0 \leq x \leq 1$, consider an interpolating Hamiltonian defined on $\Sigma_N \times \mathbb{N}^r$ by

$$H_{N,x}(\sigma,\alpha) = \sqrt{x}H_N(\sigma) + \sqrt{1-x}\sum_{i=1}^N \sigma_i C_i(\alpha) + \sqrt{x}\sqrt{N}D(\alpha) \tag{19}$$

and the corresponding interpolating free energy

$$\varphi(x) = \frac{1}{N}\mathbb{E}\log\sum_{\sigma,\alpha} v_\alpha \exp H_{N,x}(\sigma,\alpha). \tag{20}$$

Then it is easy to check the following.

LEMMA 1. *Under assumption* (6), *the derivative of* $\varphi(x)$ *in* (20) *satisfies* $\varphi'(x) \leq 0$.

PROOF. Let us denote by $\langle\cdot\rangle_x$ the average with respect to the Gibbs measure $\Gamma_x(\sigma,\alpha)$ on $\Sigma_N \times \mathbb{N}^r$ defined by

$$\Gamma_x(\sigma,\alpha) \sim v_\alpha \exp H_{N,x}(\sigma,\alpha).$$

Then, obviously, for $0 < x < 1$,

$$\varphi'(x) = \frac{1}{N}\mathbb{E}\left\langle \frac{\partial H_{N,x}(\sigma,\alpha)}{\partial x}\right\rangle_x .$$

It is easy to check from the above definitions that

$$\begin{aligned}\frac{1}{N}\mathbb{E}&\frac{\partial H_{N,x}(\sigma^1,\alpha^1)}{\partial x} H_{N,x}(\sigma^2,\alpha^2)\\ &= \frac{1}{2}\sum_{s,t\in\mathscr{S}} \Delta^2_{st}\lambda_s\lambda_t\big(R_s(\sigma^1,\sigma^2)R_t(\sigma^1,\sigma^2)\\ &\qquad\qquad\qquad - 2R_s(\sigma^1,\sigma^2)q^t_{\alpha^1\wedge\alpha^2} + q^s_{\alpha^1\wedge\alpha^2}q^t_{\alpha^1\wedge\alpha^2}\big).\end{aligned}$$

In particular, this is zero when $(\sigma^1,\alpha^1) = (\sigma^2,\alpha^2)$ and, in general, can be rewritten as a quadratic form $(\Delta^2(R-q),(R-q))/2$, where

$$R = \big(\lambda_s R_s(\sigma^1,\sigma^2)\big)_{s\in\mathscr{S}}, \qquad q = \big(\lambda_s q^s_{\alpha^1\wedge\alpha^2}\big)_{s\in\mathscr{S}}.$$

Notice that here we used the symmetry of the matrix Δ^2. Finally, usual Gaussian integration by parts then gives (see, e.g., Lemma 1.1 in [9])

$$\varphi'(x) = -\frac{1}{2}\mathbb{E}\big\langle\big(\Delta^2(R-q),(R-q)\big)\big\rangle_x \le 0,$$

where the last inequality follows from the assumption $\Delta^2 \ge 0$ in (6). □

The lemma implies that $\varphi(1)\le\varphi(0)$. It is easy to see that

$$\varphi(0) = \log 2 + \frac{1}{N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha \prod_{i\le N}\operatorname{ch}C_i(\alpha)$$

and

$$\varphi(1) = F_N + \frac{1}{N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha \exp\sqrt{N}D(\alpha).$$

Now, standard properties of the Ruelle probability cascades imply that (see, e.g., the proof of Lemma 3.1 in [9]),

$$\begin{aligned}\frac{1}{N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha\prod_{i\le N}\operatorname{ch}C_i(\alpha) &= \frac{1}{N}\sum_{1\le i\le N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha\operatorname{ch}C_i(\alpha)\\ &= \sum_{s\in\mathscr{S}}\lambda_s X_0^s\end{aligned}\tag{21}$$

and

$$\frac{1}{N}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r} v_\alpha\exp\sqrt{N}D(\alpha) = \frac{1}{2}\sum_{0\le\ell\le r-1}\zeta_\ell(Q_{\ell+1}-Q_\ell).\tag{22}$$

Recalling (15), the inequality $\varphi(1)\le\varphi(0)$ can be written as $F_N \le \mathscr{P}(\zeta,q)$, which yields the upper bound in (16).

3. Multi-species Ghirlanda–Guerra identities. In order to prepare for the proof of the lower bound, we need to obtain some strong coupling properties for the overlaps in different species, which will be achieved in the next section using a multi-species version of the Ghirlanda–Guerra identities that we will now prove. Let us consider a countable dense subset $\mathscr{W}$ of $[0,1]^{|\mathscr{S}|}$. For a vector

$$w = (w_s)_{s\in\mathscr{S}} \in \mathscr{W}, \tag{23}$$

let $s_i(w) = \sqrt{w_s}$ for $i \in I_s$ and $s \in \mathscr{S}$, and consider the following p-spin Hamiltonian,

$$h_{N,w,p}(\sigma) = \frac{1}{N^{p/2}} \sum_{1\le i_1,\dots,i_p\le N} g^{w,p}_{i_1,\dots,i_p} \sigma_{i_1} s_{i_1}(w)\cdots\sigma_{i_p} s_{i_p}(w), \tag{24}$$

where $g^{w,p}_{i_1,\dots,i_p}$ are i.i.d. standard Gaussian random variables independent for all combinations of indices $p \ge 1$, $w \in \mathscr{W}$ and $i_1,\dots,i_p \in \{1,\dots,N\}$. If we define

$$R_w(\sigma^1,\sigma^2) = \sum_{s\in\mathscr{S}} \lambda_s w_s R_s(\sigma^1,\sigma^2), \tag{25}$$

where $R_s(\sigma^1,\sigma^2)$ was defined in (7), then it is easy to check that the covariance of (24) is

$$\mathbb{E} h_{N,w,p}(\sigma^1) h_{N,w,p}(\sigma^2) = R_w(\sigma^1,\sigma^2)^p. \tag{26}$$

Since the set $\mathscr{W}$ is countable, we can consider some one-to-one function $j:\mathscr{W}\to\mathbb{N}$. Then we let $x_{w,p}$ for $p\ge 1$, $w\in\mathscr{W}$ be i.i.d. random variables uniform on the interval $[1,2]$ and define a Hamiltonian

$$h_N(\sigma) = \sum_{w\in\mathscr{W}} \sum_{p\ge 1} 2^{-j(w)-p} x_{w,p} h_{N,w,p}(\sigma). \tag{27}$$

Note that, conditionally on $x = (x_{w,p})_{p\ge 1, w\in\mathscr{W}}$, this is a Gaussian process and its variance is bounded by 4. The Hamiltonian $h_N(\sigma)$ will play a role of a perturbation Hamiltonian, which means that, instead of $H_N(\sigma)$ in (4), from now on we will consider the perturbed Hamiltonian

$$H_N^{\text{pert}}(\sigma) = H_N(\sigma) + s_N h_N(\sigma), \tag{28}$$

where $s_N = N^{\gamma}$ for any $1/4 < \gamma < 1/2$. First of all, it is easy to see, using Jensen's inequality on each side, that

$$\begin{aligned} \frac{1}{N}\mathbb{E}\log\sum_{\sigma\in\Sigma_N}\exp H_N(\sigma) &\le \frac{1}{N}\mathbb{E}\log\sum_{\sigma\in\Sigma_N}\exp H_N^{\text{pert}}(\sigma) \\ &\le \frac{1}{N}\mathbb{E}\log\sum_{\sigma\in\Sigma_N}\exp H_N(\sigma) + \frac{2s_N^2}{N}, \end{aligned} \tag{29}$$

and, since $\lim_{N\to\infty} N^{-1}s_N^2 = 0$, the perturbation term does not affect the limit of the free energy. As in the Sherrington–Kirkpatrick and mixed p-spin models, the purpose of adding the perturbation term is to obtain the Ghirlanda–Guerra identities for the Gibbs measure

$$G_N(\sigma) = \frac{\exp H_N^{\mathrm{pert}}(\sigma)}{Z_N} \qquad \text{where } Z_N = \sum_{\sigma\in\Sigma_N} \exp H_N^{\mathrm{pert}}(\sigma), \tag{30}$$

corresponding to the perturbed Hamiltonian (28). We will denote the average with respect to $G_N^{\otimes\infty}$ by $\langle\cdot\rangle$. Now, given $n \geq 2$, let

$$R^n = \big(R_s(\sigma^\ell, \sigma^{\ell'})\big)_{s\in\mathscr{S},\ell,\ell'\leq n}$$

and consider an arbitrary bounded measurable function $f = f(R^n)$. For $p \geq 1$ and $w \in \mathscr{W}$, let

$$\begin{aligned} \Delta(f,n,w,p) = \Big|\mathbb{E}\big\langle f R_w(\sigma^1,\sigma^{n+1})^p\big\rangle - \frac{1}{n}\mathbb{E}\langle f\rangle \mathbb{E}\big\langle R_w(\sigma^1,\sigma^2)^p\big\rangle \\ - \frac{1}{n}\sum_{\ell=2}^{n} \mathbb{E}\big\langle f R_w(\sigma^1,\sigma^\ell)^p\big\rangle\Big|, \end{aligned} \tag{31}$$

where $\mathbb{E}$ denotes the expectation conditionally on the i.i.d. uniform sequence $x = (x_{w,p})_{p\geq 1, w\in\mathscr{W}}$. If we denote by $\mathbb{E}_x$ the expectation with respect to x then the following holds.

THEOREM 2. *For any $n \geq 2$ and any bounded measurable function $f = f(R^n)$,*

$$\lim_{N\to\infty} \mathbb{E}_x \Delta(f,n,w,p) = 0 \tag{32}$$

for all $p \geq 1$ and $w \in \mathscr{W}$.

PROOF. The proof is identical to the one of Theorem 3.2 in [9]. For a given $p \geq 1$ and $w \in \mathscr{W}$, equation (32) is obtained by utilizing the term $h_{N,w,p}(\sigma)$ in the perturbation (27). □

Theorem 2 implies that we can choose a nonrandom sequence $x^N = (x^N_{w,p})_{p\geq 1, w\in\mathscr{W}}$ changing with N such that

$$\lim_{N\to\infty} \Delta(f,n,w,p) = 0 \tag{33}$$

for the Gibbs measure G_N with the parameters x in the perturbation Hamiltonian (27) equal to x^N rather than random. In fact, the choice of x^N will be made below in a special way to coordinate with the Aizenman–Sim–Starr scheme. In this section, we will simply assume that we have any such sequence x^N. Moreover, let us now consider any subsequence $(N_k)_{k\geq 1}$ along which the array

$$\big(R_s(\sigma^\ell,\sigma^{\ell'})\big)_{s\in\mathscr{S},\ell,\ell'\geq 1}$$

of the overlaps within species for infinitely many replicas $(\sigma^\ell)_{\ell\geq 1}$ converges in distribution under the measure $\mathbb{E}G_N^{\otimes\infty}$. Again, later we will be interested in a special choice of such subsequence. Let

$$\big(R^s_{\ell,\ell'}\big)_{s\in\mathscr{S},\ell,\ell'\geq 1} \tag{34}$$

be the array with the limiting distribution, and similarly to (25), define

$$R^w_{\ell,\ell'}=\sum_{s\in\mathscr{S}}\lambda_s w_s R^s_{\ell,\ell'}. \tag{35}$$

Then equations (31) and (33) imply that the limiting array satisfies

$$\mathbb{E}f(R^n)(R^w_{1,n+1})^p=\frac{1}{n}\mathbb{E}f(R^n)\mathbb{E}(R^w_{1,2})^p+\frac{1}{n}\sum_{\ell=2}^n\mathbb{E}f(R^n)(R^w_{1,\ell})^p, \tag{36}$$

where, of course, now $R^n=(R^s_{\ell,\ell'})_{s\in\mathscr{S},\ell,\ell'\leq n}$. From this we will deduce the following multi-species form of the Ghirlanda–Guerra identities for such limiting arrays. Let us consider an array

$$Q_{\ell,\ell'}=\varphi\big((R^s_{\ell,\ell'})_{s\in\mathscr{S}}\big) \tag{37}$$

for any bounded measurable function φ of the overlaps in different species.

THEOREM 3. *For any $n\geq 2$ and any bounded measurable function $f=f(R^n)$,*

$$\mathbb{E}f(R^n)Q_{1,n+1}=\frac{1}{n}\mathbb{E}f(R^n)\mathbb{E}Q_{1,2}+\frac{1}{n}\sum_{\ell=2}^n\mathbb{E}f(R^n)Q_{1,\ell}. \tag{38}$$

PROOF. Since equation (36) holds for all $w\in\mathscr{W}$, both sides are continuous in w, and $\mathscr{W}$ is dense in $[0,1]^{|\mathscr{S}|}$, equation (36) holds for all $w\in[0,1]^{|\mathscr{S}|}$. Take any integers $p_s\geq 0$ for $s\in\mathscr{S}$, and let $p=\sum_{s\in\mathscr{S}}p_s$. If we recall the definition of $R^w_{\ell,\ell'}$ in (35),

$$\frac{\partial^p}{\prod_{s\in\mathscr{S}}\partial w_s^{p_s}}(R^w_{\ell,\ell'})^p=p!\prod_{s\in\mathscr{S}}(\lambda_s R^s_{\ell,\ell'})^{p_s}.$$

Computing this partial derivative on both sides of (36) implies

$$\begin{aligned}\mathbb{E}f(R^n)\prod_{s\in\mathscr{S}}(R^s_{1,n+1})^{p_s}&=\frac{1}{n}\mathbb{E}f(R^n)\mathbb{E}\prod_{s\in\mathscr{S}}(R^s_{1,2})^{p_s}\\&\quad+\frac{1}{n}\sum_{\ell=2}^n\mathbb{E}f(R^n)\prod_{s\in\mathscr{S}}(R^s_{1,\ell})^{p_s}.\end{aligned} \tag{39}$$

Approximating continuous functions by polynomials, this implies (38) for continuous functions φ in (37), and the general case follows. □

REMARK. In particular, Theorem 3 implies that the array $(Q_{\ell,\ell'})_{\ell,\ell'\geq 1}$ itself satisfies the usual Ghirlanda–Guerra identities,

$$\mathbb{E}f(Q^n)\psi(Q_{1,n+1}) = \frac{1}{n}\mathbb{E}f(Q^n)\mathbb{E}\psi(Q_{1,2}) + \frac{1}{n}\sum_{\ell=2}^{n}\mathbb{E}f(Q^n)\psi(Q_{1,\ell}), \tag{40}$$

for any bounded measurable function ψ and $f = f(Q^n)$, where $Q^n = (Q_{\ell,\ell'})_{\ell,\ell'\leq n}$. In the case when the array Q is also nonnegative definite, the main result in [8] will allow us to use the full force of the Ghirlanda–Guerra identities and, in particular, will imply that such arrays are ultrametric and can be generated by the Ruelle probability cascades; see Section 2.4 in [9].

4. Synchronizing the species. Now, let us consider any limiting distribution as in (34), and let us notice that the overlap

$$R(\sigma^\ell, \sigma^{\ell'}) = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^\ell\sigma_i^{\ell'} = \sum_{s\in\mathscr{S}}\lambda_s R_s(\sigma^\ell, \sigma^{\ell'})$$

of two configurations over the whole system in the limit will become

$$R_{\ell,\ell'} = \sum_{s\in\mathscr{S}}\lambda_s R^s_{\ell,\ell'}. \tag{41}$$

In this section, we will prove the main result that will allow us to characterize the limits that will arise in the Aizenman–Sims–Starr scheme.

THEOREM 4. *For any array* (34) *that satisfies* (38), *there exist nondecreasing* $(1/\lambda_s)$*-Lipschitz functions* $L_s : [0, 1] \to [0, 1]$ *such that* $R^s_{\ell,\ell'} = L_s(R_{\ell,\ell'})$ *almost surely for all* $s \in \mathscr{S}$ *and all* $\ell, \ell' \geq 1$.

The reason we can consider the domain and range of L_s to be $[0, 1]$ is because each array R^s is nonnegative definite and satisfies the Ghirlanda–Guerra identities (40), and therefore, its entries are nonnegative by Talagrand's positivity principle (Theorem 2.16 in [9]). Theorem 4 implies that the joint distribution of the overlap arrays for all species will be determined trivially by the overlap array $(R_{\ell,\ell'})_{\ell,\ell'\geq 1}$. On the other hand, the Ghirlanda–Guerra identities imply that this array can be generated using the Ruelle probability cascades, which will be used in Section 5. We begin with the following observation.

LEMMA 2. *If* $R^s_{\ell,\ell'} > R^s_{\ell,\ell''}$ *for some* $s \in \mathscr{S}$, *then* $R^t_{\ell,\ell'} \geq R^t_{\ell,\ell''}$ *for all* $t \in \mathscr{S}$.

PROOF. By Theorem 3, for any $s,t\in\mathscr{S}$, the arrays

$$(R^s_{\ell,\ell'})_{\ell,\ell'\ge 1},\qquad (R^t_{\ell,\ell'})_{\ell,\ell'\ge 1}\quad\text{and}\quad (R^s_{\ell,\ell'}+R^t_{\ell,\ell'})_{\ell,\ell'\ge 1}$$

satisfy the Ghirlanda–Guerra identities. Since all these arrays are nonnegative definite, the main result in [8] (or Theorem 2.14 in [9]) implies that these arrays are ultrametric, that is,

$$R^s_{\ell',\ell''}\ge \min(R^s_{\ell,\ell'},R^s_{\ell,\ell''}) \tag{42}$$

for any different $\ell,\ell',\ell''\ge 1$ and, similarly, for the other two arrays. In other words, given three replica indices, the smallest two overlaps are equal. Suppose now that $R^s_{\ell,\ell'}>R^s_{\ell,\ell''}$ but $R^t_{\ell,\ell'}<R^t_{\ell,\ell''}$. By ultrameticity of the first two arrays,

$$R^s_{\ell,\ell'}>R^s_{\ell,\ell''}=R^s_{\ell',\ell''}\quad\text{and}\quad R^t_{\ell',\ell''}=R^t_{\ell,\ell'}<R^t_{\ell,\ell''}.$$

However, this implies that

$$R^s_{\ell',\ell''}+R^t_{\ell',\ell''}<\min(R^s_{\ell,\ell'}+R^t_{\ell,\ell'},R^s_{\ell,\ell''}+R^t_{\ell,\ell''}),$$

violating ultrametricity of the third array. □

Let us state one obvious corollary of the above lemma.

COROLLARY 1. *The following statements hold*:

(a) *If* $R_{\ell,\ell'}>R_{\ell,\ell''}$, *then* $R^s_{\ell,\ell'}\ge R^s_{\ell,\ell''}$ *for all* $s\in\mathscr{S}$.
(b) *If* $R^s_{\ell,\ell'}>R^s_{\ell,\ell''}$ *for some* $s\in\mathscr{S}$, *then* $R_{\ell,\ell'}>R_{\ell,\ell''}$.

This already gives some indication that the overlaps in different species will be synchronized. However, keeping in mind the ultrametric tree structure of the Ruelle probability cascades that generate them, we need to show that the entire clusters are synchronized and the corresponding cascades are completely coupled. To prove this, for $q\in[0,1]$ and $s\in\mathscr{S}$, we will consider the array

$$R^{s,q}_{\ell,\ell'}=\mathrm{I}(R_{\ell,\ell'}\ge q)(R^s_{\ell,\ell'}+1). \tag{43}$$

First of all, we add $+1$ to the overlap $R^s_{\ell,\ell'}$ to ensure that the only way the right-hand side can be equal to zero is when $R_{\ell,\ell'}<q$ and not, for example, when $R^s_{\ell,\ell'}=0$. As in (42), by Theorem 3, the array $(R_{\ell,\ell'})_{\ell,\ell'\ge1}$ is ultrametric, which implies that the array $(\mathrm{I}(R_{\ell,\ell'}\ge q))_{\ell,\ell'\ge1}$ is nonnegative definite, as it consists of blocks on the diagonal with all entries equal to one. Therefore, the array

$$R^{s,q}=(R^{s,q}_{\ell,\ell'})_{\ell,\ell'\ge1}$$

is nonnegative definite as the Hadamard product of two such arrays. By Theorem 3, the array $R^{s,q}$ also satisfies the Ghirlanda–Guerra identities, so all the consequences of the Ghirlanda–Guerra identities for nonnegative definite arrays described, for example, in Section 2.4 in [9], hold in this case. One such consequence is the following. Let

$$\mu = \mathscr{L}(R_{1,2}) \quad \text{and} \quad \mu^{s,q} = \mathscr{L}\big(R^{s,q}_{1,2}\big) \tag{44}$$

be the distributions of one entry of the arrays R and $R^{s,q}$ correspondingly. Lemma 2.7 in [9] implies the following consequence of the Ghirlanda–Guerra identities, which was first observed in [13].

LEMMA 3. *For any $s \in \mathscr{S}$, $\ell \geq 1$ and $q \in [0,1]$, with probability one, the set*

$$A^s_\ell(q) = \{R^{s,q}_{\ell,\ell'} \,|\, \ell' \neq \ell\} = \{\mathrm{I}(R_{\ell,\ell'} \geq q)\big(R^s_{\ell,\ell'} + 1\big) \,|\, \ell' \neq \ell\} \tag{45}$$

is a dense subset of the support of $\mu^{s,q}$.

This will be the key to the proof of Theorem 4. Now, for any $q \in [0,1]$, let us define

$$\ell_s(q) = \inf\{x \geq 1 \,|\, x \in \operatorname{supp} \mu^{s,q}\} - 1. \tag{46}$$

Equivalently, one could take the infimum over $x > 0$, because $R^{s,q}_{\ell,\ell'} > 0$ if and only if $R^{s,q}_{\ell,\ell'} \geq 1$. To understand the meaning of this definition, let us notice that, whenever the set $A^s_\ell(q)$ in (45) is dense in the support of $\mu^{s,q}$ (which happens with probability one for a given q),

$$\ell_s(q) = \inf\{R^s_{\ell,\ell'} \,|\, \ell' \neq \ell, R_{\ell,\ell'} \geq q\}, \tag{47}$$

so $\ell_s(q)$ is just the smallest value that $R^s_{\ell,\ell'}$ can take whenever $R_{\ell,\ell'} \geq q$. This alternative definition, obviously, implies the following.

LEMMA 4. *For any $s \in \mathscr{S}$, the function $\ell_s(q)$ in (46) is nondecreasing in q.*

To obtain the functions L_s in Theorem 4, we will first need to regularize $\ell_s(q)$ as follows:

$$L_s(q) = \lim_{x \uparrow q} \ell_s(x) \tag{48}$$

for $q > 0$ and $L_s(0) = \ell_s(0)$. Theorem 4 will be now proved in two steps. First, we will show that $R^s_{\ell,\ell'} = L_s(R_{\ell,\ell'})$ almost surely. Second, we will show that L_s is $(1/\lambda_s)$-Lipschitz on the support of the distribution μ of $R_{1,2}$. Then we can redefine L_s outside of the support to be $(1/\lambda_s)$-Lipschitz extension which, obviously, does not change the first claim, $R^s_{\ell,\ell'} = L_s(R_{\ell,\ell'})$, since $R_{\ell,\ell'}$ belongs to the support of μ almost surely.

PROOF OF THEOREM 4. *Step* 1. We will use that the claim in Lemma 3 holds with probability one simultaneously for all $q \in \mathbb{Q} \cap [0,1]$. Let us fix some indices $\ell \neq \ell'$. If $\mu(\{0\}) = 0$, then all $R_{\ell,\ell'} > 0$ almost surely. If $\mu(\{0\}) > 0$ and $R_{\ell,\ell'} = 0$, then we must have $R^s_{\ell,\ell'} = 0$ for all $s \in \mathscr{S}$, and definition (46) implies that $\ell_s(0) = 0$. In this case,

$$R^s_{\ell,\ell'} = \ell_s(R_{\ell,\ell'}) = L_s(R_{\ell,\ell'}).$$

Let us now consider the case when $R_{\ell,\ell'} > 0$. First of all, for any $x < R_{\ell,\ell'}$ we must have that $\ell_s(x) \le R^s_{\ell,\ell'}$, because the function $\ell_s(x)$ is nondecreasing and, for any rational $q \le R_{\ell,\ell'}$, (47) implies that $\ell_s(q) \le R^s_{\ell,\ell'}$. Next, consider arbitrary $\varepsilon > 0$, and consider any rational q such that

$$q < R_{\ell,\ell'} \le q + \varepsilon. \tag{49}$$

Consider two possibilities. First, suppose that $R^s_{\ell,\ell'} = \ell_s(q)$. Since for $q \le x < R_{\ell,\ell'}$ we showed that

$$\ell_s(x) \le R^s_{\ell,\ell'} = \ell_s(q) \le \ell_s(x)$$

[so $\ell_s(x) = \ell_s(q)$ for such x], we get the desired claim,

$$R^s_{\ell,\ell'} = \ell_s(q) = \lim_{x \uparrow R_{\ell,\ell'}} \ell_s(x) = L_s(R_{\ell,\ell'}).$$

Second, suppose that $\ell_s(q) < R^s_{\ell,\ell'}$. By (47), we can find a sequence (ℓ_n) such that $R_{\ell,\ell_n} \ge q$ and $R^s_{\ell,\ell_n} \downarrow \ell_s(q)$. Since we assumed that $\ell_s(q) < R^s_{\ell,\ell'}$, for large enough n we must have $R^s_{\ell,\ell_n} < R^s_{\ell,\ell'}$ and, by Corollary 1, we get $R_{\ell,\ell_n} < R_{\ell,\ell'}$ and $R^t_{\ell,\ell_n} \le R^t_{\ell,\ell'}$ for all $t \in \mathscr{S}$. Therefore,

$$\begin{aligned} 0 &\le \lambda_s\big(R^s_{\ell,\ell'} - R^s_{\ell,\ell_n}\big) \le \sum_{t \in \mathscr{S}} \lambda_t\big(R^t_{\ell,\ell'} - R^t_{\ell,\ell_n}\big) \\ &= R_{\ell,\ell'} - R_{\ell,\ell_n} \le q + \varepsilon - q = \varepsilon. \end{aligned} \tag{50}$$

Using that $R^s_{\ell,\ell_n} \downarrow \ell_s(q)$ implies that $\ell_s(q) \le R^s_{\ell,\ell'} \le \ell_s(q) + \varepsilon\lambda_s^{-1}$. Finally, letting $q \uparrow R_{\ell,\ell'}$ and $\varepsilon \downarrow 0$ in such a way that (49) holds, again, implies the desired claim

$$R^s_{\ell,\ell'} = \lim_{q \uparrow R_{\ell,\ell'}} \ell_s(q) = L_s(R_{\ell,\ell'}).$$

Step 2. Let us now show that L_s is $(1/\lambda_s)$-Lipschitz on the support of the distribution μ of $R_{1,2}$. Take $q_1 < q_2$ in the support of μ. Let $q_2' = q_2 - \varepsilon_2$ for some small $\varepsilon_2 > 0$ such that $q_2' > 0$, and let $q_1' = \max(q_1 - \varepsilon_1, 0)$ for some small $\varepsilon_1 > 0$. Let us also make sure that q_1' and q_2' are rational. By (47), given $\varepsilon > 0$, we can find indices ℓ_j for $j = 1, 2$ such that

$$R_{\ell,\ell_j} \ge q_j' \quad \text{and} \quad \ell_s(q_j') \le R^s_{\ell,\ell_j} \le \ell_s(q_j') + \varepsilon. \tag{51}$$

Similarly to Lemma 3, Lemma 2.7 in [9] implies that the set $\{R_{\ell,\ell'} | \ell' \ne \ell\}$ is a dense subset of the support of $\mu = \mathscr{L}(R_{1,2})$ with probability one and, since we chose q_1 and q_2 in the support of μ, we can find other indices ℓ_j' for $j = 1, 2$ such that

$$q_j' \le R_{\ell,\ell_j'} \le q_j + \varepsilon.$$

If the index ℓ_j already satisfies this condition, we simply take $\ell'_j = \ell_j$. Otherwise, because of the first inequality in (51), we must have $R_{\ell,\ell'_j} < R_{\ell,\ell_j}$ and, by (47), Corollary 1 and the second inequality in (51),

$$\ell_s(q'_j) \le R^s_{\ell,\ell'_j} \le R^s_{\ell,\ell_j} \le \ell_s(q'_j) + \varepsilon.$$

In both cases, we have

$$q'_j \le R_{\ell,\ell'_j} \le q_j + \varepsilon \quad \text{and} \quad \ell_s(q'_j) \le R^s_{\ell,\ell'_j} \le \ell_s(q'_j) + \varepsilon.$$

Since $q_1 < q_2$, by taking $\varepsilon > 0$ small enough, we can assume that $R_{\ell,\ell'_1} < R_{\ell,\ell'_2}$. Then, as in (50),

$$\lambda_s(R^s_{\ell,\ell'_2} - R^s_{\ell,\ell'_1}) \le R_{\ell,\ell'_2} - R_{\ell,\ell'_1}.$$

Combining all the inequalities, we showed that

$$\lambda_s(\ell_s(q'_2) - \ell_s(q'_1) - \varepsilon) \le q_2 + \varepsilon - q'_1.$$

Letting $\varepsilon, \varepsilon_1, \varepsilon_2 \downarrow 0$ implies $\lambda_s(L_s(q_2) - L_s(q_1)) \le q_2 - q_1$, which proves that L_s is $(1/\lambda_s)$-Lipschitz on the support of μ. As we mentioned above, $(1/\lambda_s)$-Lipschitz extension of L_s outside of the support does not affect the fact that $R^s_{\ell,\ell'} = L_s(R_{\ell,\ell'})$ almost surely. $\square$

5. Lower bound via the Aizenman–Sims–Starr scheme. Given the main result in the previous section, the arguments of this section will be a standard exercise. To a reader familiar with the corresponding arguments in the setting of the classical SK model (e.g., Sections 3.5 and 3.6 in [9]) these arguments will be completely obvious. Otherwise, we recommend to study them first in the easier case of the SK model.

It is clear that small modifications of the vector $(\lambda_s)_{s\in\mathscr{S}}$ result in small changes both of the free energy for large N and the Parisi formula (16), so without loss of generality, we can assume that all λ_s are rational and can be written as

$$\lambda_s = \frac{k_s}{k}. \tag{52}$$

In the proof of the lower bound, we will use an obvious fact that

$$\liminf_{N\to\infty} F_N \ge \frac{1}{k} \liminf_{n\to\infty} (\mathbb{E}\log Z_{nk+k} - \mathbb{E}\log Z_{nk}). \tag{53}$$

Let us consider the right-hand side for a fixed $N = nk$, and in addition to partition (2), let us consider a partition of k new coordinates

$$I^+ = \{N+1, \dots, N+k\} = \bigcup_{s\in\mathscr{S}} I^+_s \tag{54}$$

into different species, so that $|I_s^+| = k_s$. Let us compare the partition functions Z_N and Z_{N+k}. If we denote $\rho = (\sigma, \varepsilon) \in \Sigma_{N+k}$ for $\sigma \in \Sigma_N$ and $\varepsilon \in \Sigma_k$, then we can write

$$H_{N+k}(\rho) = H'_N(\sigma) + \sum_{i \in I^+} \varepsilon_i z_{N,i}(\sigma) + r(\varepsilon), \tag{55}$$

where

$$H'_N(\sigma) = \frac{1}{\sqrt{N+k}} \sum_{i,j=1}^{N} g_{ij}\sigma_i\sigma_j, \tag{56}$$

$$z_{N,i}(\sigma) = \frac{1}{\sqrt{N+k}} \sum_{j=1}^{N} (g_{ij} + g_{ji})\sigma_j \tag{57}$$

and

$$r(\varepsilon) = \frac{1}{\sqrt{N+k}} \sum_{i,j \in I^+} g_{ij}\varepsilon_i\varepsilon_j. \tag{58}$$

On the other hand, the Gaussian process $H_N(\sigma)$ on Σ_N can be decomposed into a sum of two independent Gaussian processes

$$H_N(\sigma) \stackrel{d}{=} H'_N(\sigma) + y_N(\sigma), \tag{59}$$

where

$$y_N(\sigma) = \frac{\sqrt{k}}{\sqrt{N(N+k)}} \sum_{i,j=1}^{N} g'_{ij}\sigma_i\sigma_j \tag{60}$$

and (g'_{ij}) are independent copies of the Gaussian random variables (g_{ij}). Using that the term $r(\varepsilon)$ is of a small order, we can write

$$\mathbb{E}\log Z_{N+k} = \mathbb{E}\log \sum_{\sigma \in \Sigma_N} \prod_{i \in I^+} 2\mathrm{ch}\big(z_{N,i}(\sigma)\big) \exp H'_N(\sigma) + o(1) \tag{61}$$

and, using equation (59),

$$\mathbb{E}\log Z_N = \mathbb{E}\log \sum_{\sigma \in \Sigma_N} \exp\big(y_N(\sigma)\big) \exp H'_N(\sigma). \tag{62}$$

Finally, if we consider the Gibbs measure on Σ_N corresponding to the Hamiltonian $H'_N(\sigma)$ in (56),

$$G'_N(\sigma) = \frac{\exp H'_N(\sigma)}{Z'_N} \qquad \text{where } Z'_N = \sum_{\sigma \in \Sigma_N} \exp H'_N(\sigma), \tag{63}$$

then combining (61), (62) we can replace the right-hand side of (53) by

$$\frac{1}{k}\liminf_{n\to\infty}\Big(\mathbb{E}\log\sum_{\sigma\in\Sigma_N}\prod_{i\in I^+}2\mathrm{ch}\big(z_{N,i}(\sigma)\big)G'_N(\sigma) \tag{64}$$
$$-\mathbb{E}\log\sum_{\sigma\in\Sigma_N}\exp\big(y_N(\sigma)\big)G'_N(\sigma)\Big).$$

This is the analogue of the Aizenman–Sims–Starr representation in [1]; see Section 3.5 in [9]. From the construction it is clear that the Gaussian processes $z_{N,i}(\sigma)$ for $i\in I^+$ and $y_N(\sigma)$ are independent of each other and the randomness of the measure G'_N. For $s\in\mathscr{S}$ and $i\in I_s^+$,

$$\begin{aligned}\mathbb{E}z_{N,i}(\sigma^1)z_{N,i}(\sigma^2) &= \frac{1}{N+k}\sum_{t\in\mathscr{S}}\sum_{j\in I_t}2\Delta_{st}^2\sigma_j^1\sigma_j^2\\ &= \frac{N}{N+k}\sum_{t\in\mathscr{S}}2\Delta_{st}^2\lambda_t R_t(\sigma^1,\sigma^2)\\ &= 2\sum_{t\in\mathscr{S}}\Delta_{st}^2\lambda_t R_t(\sigma^1,\sigma^2)+O(N^{-1})\end{aligned} \tag{65}$$

and, similarly to the computation of the covariance in (8),

$$\begin{aligned}\mathbb{E}y_N(\sigma^1)y_N(\sigma^2) &= \frac{kN^2}{N(N+k)}\sum_{s,t\in\mathscr{S}}\Delta_{st}^2\lambda_s\lambda_t R_s(\sigma^1,\sigma^2)R_t(\sigma^1,\sigma^2)\\ &= k\sum_{s,t\in\mathscr{S}}\Delta_{st}^2\lambda_s\lambda_t R_s(\sigma^1,\sigma^2)R_t(\sigma^1,\sigma^2)+O(N^{-1}).\end{aligned} \tag{66}$$

Notice how these expressions resemble the definition in (12). Of course, one can ignore the lower error terms $O(N^{-1})$ from now on.

The same computation can be carried out just as easily in the case when the free energy F_N in (53) corresponds to the perturbed Hamiltonian $H_N^{\mathrm{pert}}(\sigma)$ in (28) instead of the original Hamiltonian $H_N(\sigma)$. Moreover, since the perturbation term $s_N h_N(\sigma)$ in (28) is of a smaller order, one can show that the perturbation term $s_{N+k}h_{N+k}(\rho)$ in the partition function Z_{N+k} can simply be replaced by the one in Z_N, $s_N h_N(\sigma)$. This is standard and is explained, for example, in Section 3.5 in [9]. In this case, we obtain the representation (64) with the Gibbs measure G'_N in (63) corresponding to the perturbed Hamiltonian

$$H'_N(\sigma)+s_N h_N(\sigma).$$

Also, in this case the expectation $\mathbb{E}$ in (64) includes the average $\mathbb{E}_x$ in the uniform random variables $x=(x_{w,p})$ in the definition of the perturbation Hamiltonian (27).

The proof of Theorem 2 applies verbatim to the measure G'_N, and right below Theorem 2 we mentioned that one can choose a nonrandom sequence $x^N=$

$(x^N_{w,p})_{p\geq 1, w\in\mathscr{W}}$ changing with N such that (33) holds for the Gibbs measure G'_N with the parameters x in the perturbation Hamiltonian (27) equal to x^N rather than random. By Lemma 3.3 in [9], one can choose this sequence x^N in such a way that the lower limit in (64) is not affected by fixing $x = x^N$ instead of averaging in x. To finish the proof, we will use Theorem 1.3 in [9] (a trivial modification of) which implies that

$$\mathbb{E}\log\sum_{\sigma\in\Sigma_N}\prod_{i\in I^+} 2\mathrm{ch}\bigl(z_{N,i}(\sigma)\bigr)G'_N(\sigma) - \mathbb{E}\log\sum_{\sigma\in\Sigma_N}\exp\bigl(y_N(\sigma)\bigr)G'_N(\sigma) \tag{67}$$

is a continuous functional of the distribution of the array

$$\bigl(R_s(\sigma^\ell,\sigma^{\ell'})\bigr)_{s\in\mathscr{S},\ell,\ell'\geq 1} \tag{68}$$

under the measure $\mathbb{E}G'^{\otimes\infty}_N$. Passing to a subsequence, if necessary, we can assume that this array converges in distribution to some array $(R^s_{\ell,\ell'})_{s\in\mathscr{S},\ell,\ell'\geq 1}$ that, by construction, satisfies Theorem 3. In particular, by Theorem 4,

$$R^s_{\ell,\ell'} = L_s(R_{\ell,\ell'}) \tag{69}$$

for some nondecreasing $(1/\lambda_s)$-Lipschitz functions L_s, where $R_{\ell,\ell'}$ is the overlap of the whole system in (41).

Let us consider sequence (10) and a sequence

$$0 = q_0 < q_1 < \cdots < q_{r-1} < q_r = 1 \tag{70}$$

such that the distribution ζ on [0, 1] defined by

$$\zeta\bigl(\{q_\ell\}\bigr) = \zeta_\ell - \zeta_{\ell-1} \qquad \text{for } \ell = 0,\ldots,r \tag{71}$$

is close to the distribution $\mathscr{L}(R_{1,2})$ of one element of the array $(R_{\ell,\ell'})_{\ell,\ell'\geq 1}$ in some metric that metrizes weak convergence of distributions on [0, 1]. As in Section 2, let $(v_\alpha)_{\alpha\in\mathbb{N}^r}$ be the weights of the Ruelle probability cascades corresponding to the parameters (10). Let $(\alpha^\ell)_{\ell\geq 1}$ be an i.i.d. sample from $\mathbb{N}^r$ according to these weights and, using sequence (70), define

$$Q_{\ell,\ell'} = q_{\alpha^\ell\wedge\alpha^{\ell'}}. \tag{72}$$

Since from Theorem 3 it is clear that the overlap array $(R_{\ell,\ell'})_{\ell,\ell'\geq 1}$ satisfies the Ghirlanda–Guerra identities, Theorems 2.13 and 2.17 in [9] imply that its distribution will be close to the distribution of the array $(Q_{\ell,\ell'})_{\ell,\ell'\geq 1}$. If for each $s\in\mathscr{S}$ we define the sequence in (11) by

$$q^s_\ell = L_s(q_\ell) \qquad \text{for } 0\leq l\leq r, \tag{73}$$

and let

$$Q^s_{\ell,\ell'} = L_s(Q_{\ell,\ell'}) = q^s_{\alpha^\ell\wedge\alpha^{\ell'}}, \tag{74}$$

equation (69) implies that the entire array $(Q^s_{\ell,\ell'})_{s,\in\mathscr{S},\ell,\ell'\geq 1}$ will be close in distribution to the array $(R^s_{\ell,\ell'})_{s\in\mathscr{S},\ell,\ell'\geq 1}$.

Let us now consider Gaussian processes $C^s(\alpha)$ for $s\in\mathscr{S}$ and $D(\alpha)$ indexed by $\alpha\in\mathbb{N}^r$ as in Section 2. For each $s\in\mathscr{S}$ and each $i\in I^+_s$, let $C_i(\alpha)$ be a copy of the process $C^s(\alpha)$, and suppose that all these processes are independent of each other and of $D(\alpha)$. Similarly to (67), consider

$$\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r}\prod_{i\in I^+}2\mathrm{ch}\big(C_i(\alpha)\big)v_\alpha-\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r}\exp\big(\sqrt{k}D(\alpha)\big)v_\alpha. \tag{75}$$

By (12), (18) and (74), the covariances of these Gaussian processes can be written as

$$\mathbb{E}C_i(\alpha^1)C_i(\alpha^2)=2\sum_{t\in\mathscr{S}}\Delta^2_{st}\lambda_t q^s_{\alpha^1\wedge\alpha^2}=2\sum_{t\in\mathscr{S}}\Delta^2_{st}\lambda_t Q^s_{1,2} \tag{76}$$

for $s\in\mathscr{S}$ and $i\in I^+_s$, and

$$\begin{aligned}\mathbb{E}\sqrt{k}D(\alpha^1)\sqrt{k}D(\alpha^2)&=k\sum_{s,t\in\mathscr{S}}\Delta^2_{st}\lambda_s\lambda_t q^s_{\alpha^1\wedge\alpha^2}q^t_{\alpha^1\wedge\alpha^2}\\&=k\sum_{s,t\in\mathscr{S}}\Delta^2_{st}\lambda_s\lambda_t Q^s_{1,2}Q^t_{1,2}.\end{aligned} \tag{77}$$

If we compare the covariances in (65) and (66) with (76) and (77), Theorem 1.3 in [9] implies that (75) is the same continuous functional of the distribution of the array

$$\big(Q^s_{\ell,\ell'}\big)_{s\in\mathscr{S},\ell,\ell'\geq 1}, \tag{78}$$

as (67) is of the array (68). Since both arrays, by construction, approximate in distribution the array $(R^s_{\ell,\ell'})_{s\in\mathscr{S},\ell,\ell'\geq 1}$, we proved that the quantities

$$\frac{1}{k}\bigg(\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r}\prod_{i\in I^+}2\mathrm{ch}\big(C_i(\alpha)\big)v_\alpha-\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r}\exp\big(\sqrt{k}D(\alpha)\big)v_\alpha\bigg) \tag{79}$$

can be used to approximate the lower limit of the free energy. It remains to observe that, similarly to (21) and (22), using standard properties of the Ruelle probability cascades (again, we refer to the proof of Lemma 3.1 in [9]),

$$\begin{aligned}\frac{1}{k}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r}\prod_{i\in I^+}2\mathrm{ch}\big(C_i(\alpha)\big)v_\alpha&=\frac{1}{k}\sum_{i\in I^+}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r}2\mathrm{ch}\big(C_i(\alpha)\big)v_\alpha\\&=\sum_{s\in\mathscr{S}}\lambda_s\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r}2\mathrm{ch}\big(C^s(\alpha)\big)v_\alpha\\&=\log 2+\sum_{s\in\mathscr{S}}\lambda_s X^s_0\end{aligned}$$

and

$$\frac{1}{k}\mathbb{E}\log\sum_{\alpha\in\mathbb{N}^r}\exp\big(\sqrt{k}D(\alpha)\big)v_\alpha=\frac{1}{2}\sum_{0\le\ell\le r-1}\zeta_\ell(Q_{\ell+1}-Q_\ell).$$

Therefore, (79) is precisely $\mathscr{P}(\zeta,q)$ defined in (15), and this finishes the proof of the lower bound.

6. Some open questions. An obvious question that arises is what happens when Δ^2 is not positive definite, for example, in the case of a bipartite model with two interacting species and no interactions within species, that is, $\Delta^2_{12}>0$, and $\Delta^2_{11}=\Delta^2_{22}=0$. Notice that our proof of the lower bound for the free energy still works in this case, but the Guerra-type upper bound in Section 2 utilized the condition $\Delta^2\ge 0$ in an essential way.

It was clear from the proof of the lower bound, in particular from the equations (71) and (73), that the parameters (ζ_ℓ) in (10) and (q_ℓ) in (11) can be interpreted as encoding the joint distribution of the overlaps within species and, therefore, the minimizer in formula (16) for the free energy has an important physical interpretation. As a result, as in the original Sherrington–Kirkpatrick model, there are many interesting questions about this formula that one can study. For example, can one extend the result in [2] to show the uniqueness of this minimizer? The main result in [2] implies that the functional in (15) is strictly convex in the vector $(\zeta_\ell)_{\ell\le r}$ for fixed parameters (11), which is sufficient to prove the uniqueness of the minimizer for one system, but not obviously for the multi-species case. Another important problem would be to understand the phase transition in this model and to describe the replica symmetric (RS) region when the minimizer corresponds to a distribution (71) concentrated on one point $q\in[0,1]$, that is,

$$\zeta_0=0,\qquad \zeta_1=1,\qquad \zeta_2=1,\qquad q_0=0,\qquad q_1=q,\qquad q_2=1.$$

For technical reasons (to define the Ruelle probability cascades) we assumed that the inequalities in (10) are strict, but the infimum in (16) may be achieved on the limiting case when some inequalities become equalities. If the infimum is replica symmetric, it is easy to write down the following critical point equations for the parameters $q^s=q^s_1$ for $s\in\mathscr{S}$:

$$\sum_{s\in\mathscr{S}}\lambda_s\Delta^2_{st}\big(q^s-\mathbb{E}\mathrm{th}^2\big(z\sqrt{Q^s}+h_s\big)\big)=0\qquad\text{for all } t\in\mathscr{S},$$

where z is a standard Gaussian random variable, $Q^s=2\sum_{t\in\mathscr{S}}\Delta^2_{st}\lambda_t q^t$ and $(h_s)_{s\in\mathscr{S}}$ is a vector of external fields corresponding to each species. (For simplicity of notation, we did not consider external fields above, but including them does not affect any arguments.) Assuming that Δ^2 is invertible, this system is equivalent to

$$q^s=\mathbb{E}\mathrm{th}^2\big(z\sqrt{Q^s}+h_s\big)\qquad\text{for all } t\in\mathscr{S}.$$

In the SK model, this reduces to one equation, and the uniqueness of its solution is known as the Latala–Guerra lemma; see Section A.14 in [17]. It would be interesting to see if the solution of the above system of equations is also unique. In that case, it should not be difficult to prove replica symmetry breaking above some analogue of the AT line (in this case, some surface) by the same method as in the SK model; see [18] or Theorem 13.3.1 in [17]. However, to characterize the replica symmetric region exactly, one would probably need to work much harder. Notice that the multi-species model allows for some interesting possibilities; for example, one can imagine that for some choice of parameters, the replica symmetry is broken in some species but not the others.

Acknowledgments. The author would like to thank Wei-Kuo Chen for several helpful discussions and the referee for several comments about the paper.

REFERENCES

[1] AIZENMAN, M., SIMS, R. and STARR, S. L. (2003). An extended variational principle for the SK spin-glass model. *Phys. Rev. B* **68** 214403.

[2] AUFFINGER, A. and CHEN, W.-K. (2015). The Parisi formula has a unique minimizer. *Comm. Math. Phys.* **335** 1429–1444. MR3320318

[3] 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

[4] CHEN, W.-K. (2013). The Aizenman–Sims–Starr scheme and Parisi formula for mixed p-spin spherical models. *Electron. J. Probab.* **18** 14. MR3126577

[5] GHIRLANDA, S. and GUERRA, F. (1998). General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity. *J. Phys. A* **31** 9149–9155. MR1662161

[6] GUERRA, F. (2003). Broken replica symmetry bounds in the mean field spin glass model. *Comm. Math. Phys.* **233** 1–12. MR1957729

[7] GUERRA, F. and TONINELLI, F. L. (2002). The thermodynamic limit in mean field spin glass models. *Comm. Math. Phys.* **230** 71–79. MR1930572

[8] PANCHENKO, D. (2013). The Parisi ultrametricity conjecture. *Ann. of Math.* (2) **177** 383–393. MR2999044

[9] PANCHENKO, D. (2013). *The Sherrington–Kirkpatrick Model*. Springer, New York. MR3052333

[10] PANCHENKO, D. (2014). The Parisi formula for mixed p-spin models. *Ann. Probab.* **42** 946–958. MR3189062

[11] PARISI, G. (1979). Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett.* **43** 1754–1756.

[12] PARISI, G. (1980). A sequence of approximate solutions to the S–K model for spin glasses. *J. Phys. A* **13** L–115.

[13] PARISI, G. and TALAGRAND, M. (2004). On the distribution of the overlaps at given disorder. *C. R. Math. Acad. Sci. Paris* **339** 303–306. MR2092018

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

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

[16] TALAGRAND, M. (2006). The Parisi formula. *Ann. of Math.* (2) **163** 221–263. MR2195134

[17] TALAGRAND, M. (2011). *Mean-Field Models for Spin Glasses*. *Ergebnisse der Mathematik und Ihrer Grenzgebiete*. 3. *Folge a Series of Modern Surveys in Mathematics* **54**. Springer, Berlin.

[18] TONINELLI, F. L. (2002). About the Almeida–Thouless transition line in the Sherrington–Kirkpatrick mean-field spin glass model. *Europhysics Letters* **60** 764.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF TORONTO
BAHEN CENTRE, 40 ST. GEORGE ST.
TORONTO, ONTARIO M5S 2E4
CANADA
E-MAIL: panchenk@math.toronto.edu