Variational Bounds for the Generalized Random Energy Model

Cristian Giardinà[1] and Shannon Starr[2]

We compute the pressure of the random energy model (REM) and generalized random energy model (GREM) by establishing variational upper and lower bounds. For the upper bound, we generalize Guerra's "broken replica symmetry bounds," and identify the random probability cascade as the appropriate random overlap structure for the model. For the REM the lower bound is obtained, in the high temperature regime using Talagrand's concentration of measure inequality, and in the low temperature regime using convexity and the high temperature formula. The lower bound for the GREM follows from the lower bound for the REM by induction. While the argument for the lower bound is fairly standard, our proof of the upper bound is new.

1. INTRODUCTION

While the analysis and properties of finite range spin glass systems (like the Edwards-Anderson model) is still a very debated issue, even in the physics community, in recent years there has been a large progress in the mathematical understanding of mean-field models.[22] This advance was triggered by the introduction of a "quadratic interpolation" technique, pioneered in Ref. 19 to establish the existence of thermodynamic limit for the Sherrington-Kirkpatrick (SK) model[27] and further developed by Guerra[18] to prove an upper bound for the pressure which coincides with the Parisi replica symmetry breaking solution of the model. Motivated by the cavity picture, this bound was generalized to a variational bound by the introduction of Random Overlap Structures (ROSt) and associated random

---

[1] EURANDOM, P.O. Box 513-5600, MB Eindhoven, The Netherlands; e-mail: giardina@eurandom. tue.nl

[2] University of Rochester, Mathematics Department, Hylan Building, Rochester, NY 14627; e-mail: sstarr@math.rochester.edu

weights in Ref. 1. Guerra's bound is recovered when the weights are chosen from the inhomogeneous Poisson point process studied by Ruelle[26] and the ROSt is a hierarchical one. Recently Talagrand[30] was able to prove that Guerra's bound is optimal, by showing that the correction term in the bound goes to zero in the thermodynamic limit, thus establishing the rigorous validity of the Parisi solution.

In this paper we solve the Generalized Random Energy Model (GREM)[13,14] by a simple analysis that partially follows the one developed for the SK model. We first obtain a variational upper bound for the pressure by the definiton of the appropriate auxiliary system (i.e., ROSt) for the model. As a guideline in this step we use the basic covariance inequality that was identified in Ref. 9. To show that the upper bound is optimal when the ROSt random weights are chosen according to the Poisson-Dirichlet point process we use a different strategy than the one developed for the SK model. Indeed the corresponding lower bound is easily obtained from a complete control of the high temperature region and convexity of the pressure. To study the high temperature region we propose a new induction argument, which starting from infinite temperature covers all the temperature values up to the critical one.

Despite the recent progress, there is not yet a direct proof of the most prominent property of the Parisi solution, namely *ultrametricity*. A distance d in a metric space is said to be ultrametric if the standard triangular inequality is replaced by the much stronger inequality $d(x, z) \leq \max\{d(x, y), d(y, z)\}$. Equivalently, in an ultrametric space all triangles are equilateral or isosceles with longer equal sides. The Generalized Random Energy Model (GREM), introduced by Derrida[13] as a model that possesses replica symmetry breaking, is ultrametric by construction, since its Hamiltonian is defined as the sum of independent Gaussian random variables positioned on the branches of a hierarchical tree. As a first step in the direction of achieving a proof of ultrametricity for mean-field models, generalized non-hierarchical models have been considered in Ref. 3, where it has been shown that they exhibit GREM-like behaviour. The analysis of the present paper, which sets the GREM model into the general variational scheme developed for the SK model, could be helpful in further studying non-hierarchical models.

The paper is organized as follow. In Sec. 2 we treat the basic case of the Random Energy Model (REM), where the energy levels are independent Gaussian random variables. This is a necessary warm-up since the GREM model will be treated as a nested succession of REM-like systems. The full analysis is presented in Sec. 3. In the Appendix we recall some useful concentration of measure estimates.

2. REM

The random energy model (REM) is a statistical mechanical model, where the energy levels are independent and identically distributed Gaussian random

variables. More precisely, for a system of size N, the Hamiltonian is a Gaussian centered family with covariance matrix

$$C_H(\sigma, \sigma') := \mathbb{E}[H_N(\sigma)H_N(\sigma')] = \frac{N}{2}\delta(\sigma, \sigma') \tag{1}$$

where $\sigma, \sigma' \in \{+1, -1\}^N$ are vectors, whose components are Ising spin variables. The dependence on N in (1) is such that thermodynamic observables (energy, free energy, etc.) are extensive in the volume, while the factor $1/2$ is included as a matter of convention. In this paper we denote by X a standard Gaussian, $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = 1$. Thus an explicit representation of the Hamiltonian is

$$H_N(\sigma) = \sqrt{\frac{N}{2}}\,X(\sigma) \tag{2}$$

where $\{X(\sigma)\}_{\sigma \in \Sigma_N}$ are 2^N i.i.d. copies of the random variable X. An equivalent representation, more in the spirit of statistical mechanics, is to consider a "lattice" Λ with $N = |\Lambda|$ sites and a random Hamiltonian

$$H_\Lambda(\sigma) = \sqrt{\frac{N}{2^{N+1}}}\sum_{X \subseteq \Lambda} J_X \sigma_X \tag{3}$$

defined on the spin configurations $\sigma : \Lambda \to \{+1, -1\}$. In (3) the J_X's are a family of i.i.d. Gaussian random variables, with $\mathbb{E}[J_X^2] = 1$, and $\sigma_X := \prod_{i \in X} \sigma_i$ for each $X \subseteq \Lambda$. Henceforth, all Gaussian random variables will be understood to have expectation equal to 0.

The main quantity we are going to study is the quenched *pressure*. We denote by $\Sigma_N = \{-1, +1\}^N$ the space of all possible spin configurations. For a finite system we define the *random* partition function

$$Z_N(\beta) = \sum_{\sigma \in \Sigma_N} e^{-\beta H_N(\sigma)}, \tag{4}$$

and the *quenched* pressure

$$P_N(\beta) = \mathbb{E}\left[\frac{1}{N}\ln Z_N(\beta)\right]. \tag{5}$$

We will be interested in the thermodynamic limit

$$P(\beta) = \lim_{N \to \infty} P_N(\beta) \tag{6}$$

In the following we will sometimes drop the $N-$dependence in the Hamiltonian $H(\sigma)$ in order to alleviate notation. We will also introduce additional randomness by considering an auxiliary system which is coupled to the Hamiltonian. We will denote by $\mathbb{E}[\cdot]$ the expectation with respect to all random variables involved.

2.1. Upper Bound

We start by recalling the quadratic interpolation technique.

Lemma 2.1. *Let $H(\sigma)$ be a Gaussian family, indexed by $\sigma \in \Sigma_N$, with covariance $C_H(\sigma; \sigma')$. Let $\alpha \in \mathcal{A}$ be an index ranging over the set $\mathcal{A}$, and let $K(\alpha)$ and $V(\sigma, \alpha)$ be Gaussian random variables, independent of $H(\sigma)$ and of each other, with covariances $C_K(\alpha; \alpha')$ and $C_V(\sigma, \alpha; \sigma', \alpha')$, respectively. Suppose that*

$$C_H(\sigma; \sigma') + C_K(\alpha; \alpha') \geq C_V(\sigma, \alpha; \sigma', \alpha') \tag{7}$$

for all $\sigma, \sigma' \in \Sigma_N$ and $\alpha, \alpha' \in \mathcal{A}$, and suppose that

$$C_H(\sigma; \sigma) + C_K(\alpha; \alpha) = C_V(\sigma, \alpha; \sigma, \alpha) \tag{8}$$

for every $\sigma \in \Sigma_N$ and $\alpha \in \mathcal{A}$. Moreover, suppose that there is a random weight $w : \mathcal{A} \to [0, \infty)$ such that, almost surely, $\sum_{\alpha \in \mathcal{A}} w(\alpha)$ is strictly positive and finite. Then,

$$P_N(\beta) \leq \mathbb{E}\left[\frac{1}{N} \ln \frac{\sum_{\sigma, \alpha} w(\alpha) e^{-\beta V(\sigma, \alpha)}}{\sum_{\alpha} w(\alpha) e^{-\beta K(\alpha)}}\right]$$

$$= \mathbb{E}\left[\frac{1}{N} \ln \sum_{\sigma, \alpha} w(\alpha) e^{-\beta V(\sigma, \alpha)}\right] - \mathbb{E}\left[\frac{1}{N} \ln \sum_{\alpha} w(\alpha) e^{-\beta K(\alpha)}\right], \tag{9}$$

as long as the right-hand-side is well-defined (i.e., not $\infty - \infty$).

Proof: We refer to Refs. 1,19 for full details. Here we only recall the basic idea. For $t \in [0, 1]$ define an interpolating Hamiltonian

$$\tilde{H}(\sigma, \alpha; t) = \sqrt{1 - t}\, [H(\sigma) + K(\alpha)] + \sqrt{t}\, V(\sigma, \alpha). \tag{10}$$

and an associated random partition function

$$Z_{N, t}(\beta) = \sum_{\sigma, \alpha} w(\alpha) e^{-\beta \tilde{H}(\sigma, \alpha; t)}. \tag{11}$$

Let $\Omega_{N, \beta, t}$ denote expectation with respect to the multiple-replica product measure, where the weight for a configuration (σ, α) of a generic copy is given by Gibbs measure associated to $\tilde{H}(\sigma, \alpha)$ times a generic weights $w(\alpha)$. In particular, for a function $f(\sigma, \alpha, \sigma', \alpha')$ of two replicas, one has

$$\Omega_{N, \beta, t}\{f(\sigma, \alpha; \sigma', \alpha')\}$$

$$= \sum_{\sigma, \alpha} \sum_{\sigma', \alpha'} w(\alpha) w(\alpha') \frac{e^{-\beta[\tilde{H}(\sigma, \alpha; t) + \tilde{H}(\sigma', \alpha'; t)]}}{Z_{N, t}^2} f(\sigma, \alpha; \sigma', \alpha').$$

Then one actually has (because of the equality along the diagonal Eq. (8))

$$\frac{1}{N} \, \mathbb{E}\Big[\ln \sum_{\sigma,\alpha} w(\alpha) e^{-\beta V(\sigma,\alpha)}\Big] - \frac{1}{N} \, \mathbb{E}\Big[\ln \sum_{\alpha} w(\alpha) e^{-\beta K(\alpha)}\Big] - P_N(\beta)$$

$$= \frac{\beta^2}{2} \int_0^1 \mathbb{E}[\Omega_{N,\beta,t}\{C_H(\sigma;\sigma') + C_K(\alpha;\alpha') - C_V(\sigma,\alpha;\sigma',\alpha')\}] \, dt. \quad (12)$$

This is proved by differentiating the quantity

$$\mathbb{E}\Big[N^{-1} \ln \sum_{\alpha,\sigma} w(\alpha) e^{-\beta \tilde{H}_N(\sigma,\alpha;t)}\Big],$$

with respect to t and using the generalized Wick's rule. Because of (7), the right hand side of (12) is obviously positive and Eq. (9) follows. □

Remark 1. The same basic argument works to bound $\mathbb{E}[F(Z_N(\beta))]$ for other functions such as $F(z) = z^a$.

Remark 2. An identity such as (12) is usually called a sum-rule. The process $K(\alpha)$ has to be thought of as a large reservoir which acts on the original system $H(\sigma)$ through the interaction $V(\sigma,\alpha)$.

We are going to use the previous lemma to establish an optimal upper bound for the REM model. A key element is to choose the correct formula for the random weight $w(\alpha)$. The correct formula for mean field spin glasses seems to generally be given by Ruelle's random probability cascade. For the REM, it is given by a single level of that, which is sometimes called the Poisson-Dirichlet process.[3] Let us give a brief description of this (the unnormalized version) in order to facilitate the following proposition.

Given $0 < m < 1$ consider the Poisson point process on $(0, \infty)$ with intensity measure equal to $m w^{-m-1} \, dw$. Almost surely, the points can be labelled $\{w_1, w_2, \ldots\}$ with $w_1 > w_2 > \cdots > 0$. Moreover, $\sum_{\alpha=1}^{\infty} w_\alpha$ is strictly positive and finite, almost surely. The distribution of $\{w_\alpha\}_\alpha$ has a remarkable invariance property: If $(f_1, f_2, \ldots)$ are i.i.d. copies of the random variable, f, which are assumed to be independent of $\{w_\alpha\}_\alpha$, then (modulo permutations) the distribution of $\{e^{f_\alpha} w_\alpha\}_\alpha$ is the same as $\{c w_\alpha\}_\alpha$ where c is the nonrandom number $c = (\mathbb{E}[e^{mf}])^{1/m}$. This is easily proved using the generalized Laplace transform. (For a proof see, for example, Ref. 29, page 481.)

One may note that instead of considering the Poisson point process $\{w_\alpha\}_\alpha$ with intensity equal to $m w^{-m-1} \, dw$, one could instead consider $w_\alpha = e^{y_\alpha/m}$ for some $\{y_\alpha\}_\alpha$. Then $-\infty < \cdots < y_2 < y_1 < \infty$ is a Poisson point process on $\mathbb{R}$ with intensity measure $e^{-y} \, dy$ (independent of m). One thinks of $-y_\alpha$'s as the free

---

[3] See Ref. 25 but also see Ref. 26. For a rather more abstract version, see Ref. 4.

energies. In this notation m is explicit. Henceforth $\{y_\alpha\}_\alpha$ will refer to the point process just described.

A consequence of the invariance property mentioned above is

$$\sum_{\alpha=1}^{\infty} e^{y_\alpha/m} \exp(f_\alpha) \stackrel{\mathcal{D}}{=} \mathbb{E}\big[\exp(mf)\big]^{1/m} \sum_{\alpha=1}^{\infty} e^{y_\alpha/m}, \tag{13}$$

which will be useful.

Proposition 2.2. *Let* $C_H(\sigma, \sigma')$ *be given by Eq. (1). Choose* $w(\alpha) = \exp[y(\alpha)/m]$ *for* $0 < m < 1$. *For each* $b \geq 1$ *let*

$$C_K(\alpha, \alpha') = (b-1)\frac{N}{2}\,\delta(\alpha, \alpha') \tag{14}$$

$$C_V(\sigma, \alpha; \sigma', \alpha') = b\,\frac{N}{2}\,\delta(\sigma, \sigma')\delta(\alpha, \alpha'). \tag{15}$$

Then one obtains the optimal upper bound for the REM,

$$P_N(\beta) \leq \inf_{0<m<1}\left[\frac{1}{4}m\beta^2 + \frac{1}{m}\ln 2\right] \tag{16}$$

Proof: We note that Lemma 2.1 is applicable because

$$b\,\delta(\sigma, \sigma')\delta(\alpha, \alpha') \leq \delta(\sigma, \sigma') + (b-1)\delta(\alpha, \alpha') \tag{17}$$

We compute separately the two terms in Eq. (9). For the first one, due to Eq. (15), we have

$$\frac{1}{N}\,\mathbb{E}\left[\ln\sum_{\sigma,\,\alpha} w(\alpha)\exp[-\beta V(\sigma, \alpha)]\right]$$

$$= \frac{1}{N}\,\mathbb{E}\left[\ln\sum_{\alpha}\exp\left[\frac{y(\alpha)}{m}\right]\sum_{\sigma}\exp\left[-\beta\sqrt{\frac{bN}{2}}X(\sigma, \alpha)\right]\right]$$

$$= \frac{1}{N}\,\mathbb{E}\left[\ln\sum_{\alpha}\exp\left[\frac{y(\alpha)}{m}\right]Z_N(\beta\sqrt{b}\,;\alpha)\right] \tag{18}$$

where $Z_N(\beta\sqrt{b}\,;\alpha)$ are independent copies (labeled by the α's) of the random variable $Z_N(\beta\sqrt{b}) = \sum_{\sigma}\exp[-\beta\sqrt{\frac{bN}{2}}X(\sigma)]$. By applying the invariance property of Eqs. (13) to (18) with $\exp[f_\alpha] = Z_N(\beta\sqrt{b};\alpha)$ we obtain

$$\frac{1}{N}\,\mathbb{E}\left[\ln\sum_{\sigma,\,\alpha} w(\alpha)\exp[-\beta V(\sigma, \alpha)]\right] =$$

$$= \frac{1}{N} \mathbb{E} \left[\ln \sum_\alpha \exp \left[\frac{y(\alpha)}{m} \right] \right] + \frac{1}{mN} \ln \mathbb{E} \left[Z_N^m(\beta \sqrt{b}) \right]. \qquad (19)$$

Then we consider the second term in Eq. (9). Taking into account the choice (14) we have

$$\frac{1}{N} \mathbb{E} \left[\ln \sum_\alpha \exp \left[\frac{y(\alpha)}{m} \right] \exp[-\beta K(\alpha)] \right]$$

$$= \frac{1}{N} \mathbb{E} \left[\ln \sum_\alpha \exp \left[\frac{y(\alpha)}{m} \right] \exp \left[-\beta \sqrt{\frac{(b-1)N}{2}} X_\alpha \right] \right]. \qquad (20)$$

Using again the invariance property Eq. (13) with $\exp[f_\alpha] = \exp[-\beta \sqrt{\frac{(b-1)N}{2}} X_\alpha]$ and computing the average we obtain

$$\frac{1}{N} \mathbb{E} \left[\ln \sum_\alpha \exp \left[\frac{y(\alpha)}{m} \right] \exp[-\beta K(\alpha)] \right]$$

$$= \frac{1}{N} \mathbb{E} \left[\ln \sum_\alpha \exp \left[\frac{y(\alpha)}{m} \right] \right] + \beta^2 \frac{(b-1)m}{4}. \qquad (21)$$

Putting together Eqs. (19) and (21) we obtain:

$$P_N(\beta) \leq \frac{1}{mN} \ln \mathbb{E} \left[Z_N^m(\beta \sqrt{b}) \right] - \beta^2 \frac{(b-1)m}{4}. \qquad (22)$$

Now we use the simple fact that

$$Z_N^m(\beta \sqrt{b}) \leq Z_N(m\beta \sqrt{b}). \qquad (23)$$

This is a general fact in statistical mechanics: since the entropy is positive by definition, the free energy is increasing in β. Indeed, considering $f_N(\beta) = -\frac{1}{N\beta} \ln(Z_N(\beta))$, the *random* free energy, one immediately checks that $f_N'(\beta) = \frac{1}{\beta} (u_N(\beta) - f_N(\beta)) = \frac{1}{\beta^2} s_N(\beta) \geq 0$, where $u_N(\beta)$ is the *random* internal energy, and $s_N(\beta)$ is the *random* entropy. Therefore, for any $0 < m \leq 1$, we have:

$$-\frac{1}{Nm\beta} \ln(Z_N(m\beta)) \leq -\frac{1}{N\beta} \ln(Z_N(\beta)), \qquad (24)$$

which is equivalent to (23) when we replace β by $\beta \sqrt{b}$. By inserting Eqs. (23) into (22) it is now easy to compute the expectation $\mathbb{E}[Z_N(m\beta \sqrt{b})]$ and we arrive at the upper bound

$$P_N(\beta) \leq \frac{1}{4} m\beta^2 + \frac{1}{m} \ln 2. \qquad (25)$$

Note that there is not anymore dependence on b in the bound. Finally the optimal bound is obtained by minimization in m which yields the (16) and completes the proof. □

Remark 3. It is interesting to note that for the REM (and for the REM only) the main result of quadratic interpolation can be viewed as an implementation of Jensen's inequality:

$$\mathbb{E}[\ln(Z_N(\beta))] = \frac{1}{m} \mathbb{E}\big[\ln Z_N^m(\beta)\big] \leq \frac{1}{m} \ln \mathbb{E}\big[Z_N^m(\beta)\big]. \tag{26}$$

This holds for general $m > 0$, but one needs $m \leq 1$ to apply (23). If one uses the "sum-rule" then one can get an explicit form for the *error* coming from Jensen's inequality in this case.

Remark 4. The inequality of (16) takes two different forms depending on whether β is greater than or less than $\beta_c := 2\sqrt{\ln(2)}$. For $\beta > \beta_c$, the right-hand-side of (16) is optimized at $m = \beta_c/\beta$. For $\beta < \beta_c$ the infimum over $0 < m < 1$ is attained by a limit $m \to 1$. Therefore, one has

$$P_N(\beta) \leq \mathcal{Q}(\beta), \tag{27}$$

where

$$\mathcal{Q}(\beta) = \begin{cases} \dfrac{1}{4}\beta^2 + \ln(2) & \text{for } \beta < \beta_c \\ \beta\sqrt{\ln(2)} & \text{for } \beta \geq \beta_c \end{cases} = \begin{cases} \dfrac{1}{4}(\beta^2 + \beta_c^2) & \text{for } \beta < \beta_c \\ \dfrac{1}{2}\beta\beta_c & \text{for } \beta \geq \beta_c \end{cases} \tag{28}$$

Note that the two pieces match at $\beta = \beta_c$.

2.2. Lower Bound

The main new result in this paper is the adaptation of the quadratic interpolation method to obtain an asymptotically sharp upper bound on $P_N(\beta)$, as we just considered above. One should view the result of the previous section as an analogue, for the REM, of Guerra's bounds for the SK model in Ref. 18. On the other hand, for the REM, unlike for the SK model, there are easy proofs of the same lower bound in the $N \to \infty$ limit. The exact formula for the pressure of the random energy model is well-known. Derrida calculated it when he introduced the model in Ref. 13 while a mathematically rigorous version of his argument is included in Refs. 5, 23. There are also proofs which rely more on

large-deviation *theory*, such as Refs. 15, 16.[4] We also encourage the reader to see the more recent, deep analysis of Refs. 6, 7. (There is much interest in the REM as far as the statistics of energy levels is concerned because in the bulk there is a kind of universality. See, for instance, Ref. 8.)

We will also present a proof of the lower bound. This is included primarily for completeness, for *nonexperts*. However, let us digress briefly to justify this for the experts: there is a great desire to obtain a purely variational proof of the lower bound for the the SK model. In light of that, it seemed worthwhile to explore how 'variational' the proof of the lower bound really is, for the REM. (Of course, it is not as variational as one would like).

Since $P_N(\beta) \leq \mathcal{Q}(\beta)$ for all N, it follows that $P(\beta) \leq \mathcal{Q}(\beta)$ in the limit $N \to \infty$. We want to show the opposite is also true, to establish that $P(\beta) = \mathcal{Q}(\beta)$ for all $\beta \geq 0$. The key to obtaining the lower bound is to understand the high temperature region, $\beta < \beta_c$.

Proposition 2.3.

$$P(\beta) = \mathcal{Q}(\beta) \quad \text{for} \quad \beta \leq \beta_c. \tag{29}$$

The proof of this result is provided in the next subsection, while here we stress its consequences. Stated otherwise, the upper bound of Proposition 2.2 saturates in the $N \to \infty$ limit, at least when $\beta \leq \beta_c$. It is a remarkable fact that this high-temperature result gives the sharpness of the upper bound also in the low-temperature region as follows.

Corollary 2.4. *For $\beta \in [0, +\infty)$*

$$P(\beta) = \mathcal{Q}(\beta). \tag{30}$$

Proof: It is a basic fact, easily seen from the definition (5), that $P_N(\beta)$ is convex in β for each N. Therefore, the limiting function $P(\beta)$ is also convex. Hence, for any β_0 and any $\beta \geq \beta_0$, we have

$$P(\beta) \geq P(\beta_0) + (\beta - \beta_0)DP(\beta_0), \tag{31}$$

where D is any convex combination of the left-handed and right-handed derivatives, which we denote D_- and D_+, respectively. We now take $\beta_0 \uparrow \beta_c$. Since we know from Proposition 2 that $P(\beta) = \mathcal{Q}(\beta)$ for $\beta < \beta_c$, we easily calculate $\lim_{\beta_0 \uparrow \beta_c} P(\beta_0) = 2\ln(2)$ while $\lim_{\beta_0 \uparrow \beta_c} D_- P(\beta_0) = \lim_{\beta_0 \uparrow \beta_c} \mathcal{Q}'(\beta_0) = \sqrt{\ln(2)}$. Putting this together completes the proof. □

---

[4] In Ref. 16, there is also a conjecture for the more general quantity $N^{-1} \ln E[Z_N(\beta)^a]$ for $a \in \mathbb{R}$. This is substantiated in the preliminary section of Ref. 31.

2.3. High Temperature Region

The proof of Proposition 2.3 will be obtained through a sequence of lemmata. The crux of the argument is standard. For example, see Ref. 29, Proposition 1.1.5. (there is another approach in Ref. 5, Theorem 9.1.2, called the "truncated second moment method"). We start with the following result, which is another variational calculation.

Lemma 2.5. *Let $\Omega_{\beta,N}$ refer to the (expectation associated to the) random probability measure on $\Sigma_N \times \Sigma_N$ specified by*

$$\Omega_{\beta,N}\{f(\sigma,\sigma')\} := Z_N^{-2}(\beta) \sum_{\sigma,\sigma' \in \Sigma_N} e^{-\beta H(\sigma)} e^{-\beta H(\sigma')} f(\sigma,\sigma').$$

For $0 \le \beta \le \beta_c$ we have

$$\frac{1}{N} \ln \mathbb{E}\big[Z_N(\beta)\Omega_{\beta,N}\{\delta(\sigma,\sigma')\}\big] \le \frac{\beta\beta_c}{2}. \tag{32}$$

Proof: Note that

$$\Omega_{\beta,N}\{\delta(\sigma,\sigma')\} = \frac{\sum_{\sigma,\sigma' \in \Sigma_N} e^{-\beta H(\sigma)} e^{-\beta H(\sigma')} \delta(\sigma,\sigma')}{Z_N^2(\beta)} = \frac{Z_N(2\beta)}{Z_N^2(\beta)}. \tag{33}$$

By (23), we know that for $0 < m < 1$

$$Z_N(2\beta) \le Z_N^{1/m}(2m\beta). \tag{34}$$

Using Hölder's inequality with $p = 1/m$, $q = 1/(1-m)$ and $\frac{1}{p} + \frac{1}{q} = 1$

$$Z_N(2m\beta) \le Z_N^m(\beta) Z_N^{1-m}\left(\frac{m}{1-m}\beta\right). \tag{35}$$

Therefore,

$$Z_N(\beta)\Omega_{\beta,N}\{\delta(\sigma,\sigma')\} = \frac{Z_N(2\beta)}{Z_N(\beta)} \le \frac{Z_N^{1/m}(2m\beta)}{Z_N(\beta)} \le Z_N^{(1-m)/m}\left(\frac{m}{1-m}\beta\right).$$

Since $m/(1-m)$ can take any positive value as m ranges over $(0,1)$ this means

$$Z_N(\beta)\Omega_{\beta,N}\{\delta(\sigma,\sigma')\} \le Z_N^{1/r}(r\beta) \tag{36}$$

for every $r \in [0, +\infty)$. Moreover, for $r \ge 1$ we can use Jensen's inequality to obtain

$$\mathbb{E}\big[Z_N^{1/r}(r\beta)\big] \le \big(\mathbb{E}[Z_N(r\beta)]\big)^{1/r} = \exp\left(N\left[\frac{\beta_c^2}{4r} + \frac{r\beta^2}{4}\right]\right). \tag{37}$$

It is easy to see that the optimal value is $r = \beta_c/\beta$, which does satisfy the constraint $r \ge 1$ because of the hypothesis $\beta \le \beta_c$. Choosing this r and putting (36) and (37) together yields (32). □

A second useful estimate is the following concentration of measure property.

Lemma 2.6. *For any β, and any $t > 0$,*

$$\mathbb{P}\{|N^{-1}\ln Z_N(\beta) - P_N(\beta)| \geq \beta t\} \leq 2e^{-Nt^2/2}.$$

The analogous result for the SK model is Corollary 2.2.5 of Ref. 29. For completeness we will give a proof of Lemma 2.6 in Appendix A and, particularly, show that a straightforward generalization of Talagrand's proof applies equally well to all models that satisfy thermodynamic stability. [10,12]

The proof of Proposition 2.3 will essentially follow from the next result, which we prove first.

Lemma 2.7. *For any $0 \leq \beta < \beta_c$,*

$$\limsup_{N\to\infty} \sup_{0\leq\beta'\leq\beta} \frac{1}{N} \ln \mathbb{E}\left[\Omega_{\beta',N}\{\delta(\sigma,\sigma')\}\right] < 0. \tag{38}$$

Proof: The proof will obtained by induction. Let us define the succession of temperatures given by $\beta_0 = 0$, $\beta_{n+1} = g(\beta_n)$ for $n \in \mathbb{N}$, where g is a definite function. As we will see, we can choose

$$g(\beta) = \beta + a\beta_c[1 - (\beta/\beta_c)]^2 \quad \text{for any} \quad 0 < a < 1/2. \tag{39}$$

and it will follow that $\beta_n \uparrow \beta_c$ as $n \to \infty$.

We first note that (38) is true for $\beta = \beta_0 = 0$, because one has $\mathbb{E}\left[\Omega_{0,N}\{\delta(\sigma,\sigma')\}\right] = 2^{-N}$. For the induction step, we will prove that if (38) is true for $\beta \in [0, \beta_n]$, then it is also true for every $\beta \in [0, \beta_{n+1}]$. Then, since $\beta_n \uparrow \beta_c$ the statement of Lemma 2.7 follows.

To prove the induction step, we first observe that

$$\frac{d}{d\beta} P_N(\beta) = \frac{\beta}{2}\left(1 - \mathbb{E}\left[\Omega_{N,\beta}\{\delta(\sigma,\sigma')\}\right]\right) \geq 0. \tag{40}$$

Indeed, this is a simple calculation using the generalized Wick's rule. Since $Z_N(0) = 2^N$ (deterministically) we have $P_N(0) = \ln(2)$. Then the sum rule

$$P_N(\beta) = \ln(2) + \frac{\beta^2}{4} - \int_0^\beta \frac{\beta'}{2} \mathbb{E}\left[\Omega_{N,\beta'}\{\delta(\sigma,\sigma')\}\right] d\beta' \tag{41}$$

follows.

Suppose now that (38) is true for $\beta \in [0, \beta_n]$. Consider a generic $\beta > \beta_n$, not necessarily smaller than β_{n+1}, and for $t > 0$ let $A_N(\beta, t)$ be the event in Lemma 2.6:

$$A_N(\beta, t) = \{|N^{-1}\ln Z_N(\beta) - P_N(\beta)| \geq \beta t\}.$$

On $A_N(\beta, t)^c$ we have $Z_N(\beta) \geq e^{N[P_N(\beta) - \beta t]}$. Let us employ the following short-hand: given a set A, let I_A denote the indicator of A and let us denote $\mathbb{E}[I_A X]$ by $\mathbb{E}[X, A]$ for every random variable X. Then we conclude that

$$\mathbb{E}\big[\Omega_{\beta, N}\{\delta(\sigma, \sigma')\}, \ A_N(\beta, t)^c\big] \ \leq \ \mathbb{E}\left[\frac{Z_N(\beta)}{e^{N[P_N(\beta) - \beta t]}}\Omega_{\beta, N}\{\delta(\sigma, \sigma')\}, \ A_N(\beta, t)^c\right]$$

$$\leq \ e^{-N[P_N(\beta) - \beta t]}\mathbb{E}\big[Z_N(\beta)\Omega_{N, \beta}\{\delta(\sigma, \sigma')\}\big].$$

Therefore

$$\frac{1}{N}\ln\mathbb{E}\big[\Omega_{N, \beta}\{\delta(\sigma, \sigma')\}, \ A_N(\beta, t)^c\big] \ \leq \ -P_N(\beta) + \frac{1}{2}\beta\beta_c + \beta t$$

follows from Lemma 2.5.

Since we assumed $\beta > \beta_n$, a lower bound for $P_N(\beta)$ is given by $P_N(\beta_n)$ and a lower bound for $P_N(\beta_n)$ is given by $\frac{1}{4}[\beta_c^2 + \beta_n^2] - o(1)$, using Eq. (41) and the induction hypothesis, where $o(1)$ represents a quantity whose limit is 0 when $N \to \infty$. One thing which is important is that while $o(1)$ does depend on β_n, it is independent of $\beta > \beta_n$. Therefore,

$$\frac{1}{N}\ln\mathbb{E}\big[\Omega_{\beta, N}\{\delta(\sigma, \sigma')\}, \ A_N(\beta, t)^c\big] \ \leq \ -\frac{1}{4}(\beta_c - \beta_n)^2$$

$$+ \frac{1}{2}\beta_c[\beta - \beta_n] + \beta t + o(1). \quad (42)$$

On the other hand, one always has $0 \leq \Omega_{\beta, N}\{\delta(\sigma, \sigma')\} \leq 1$. Hence,

$$\mathbb{E}\big[\Omega_{\beta, N}\{\delta(\sigma, \sigma')\}, \ A_N(\beta, t)\big] \ \leq \ \mathbb{E}[1, \ A_N(\beta, t)] \ = \ \mathbb{P}(A_N(\beta, t)).$$

So, by Lemma 2.6,

$$\frac{1}{N}\ln\mathbb{E}\big[\Omega_{\beta, N}\{\delta(\sigma, \sigma')\}, \ A_N(\beta, t)\big] \ \leq \ -\frac{1}{2}t^2 + N^{-1}\ln(2). \quad (43)$$

Putting Eqs. (42) and (43) together we obtain

$$\frac{1}{N}\ln\mathbb{E}\big[\Omega_{\beta, N}\{\delta(\sigma, \sigma')\}\big] \leq N^{-1}\ln(2) + \max\left\{-\frac{1}{2}t^2 + N^{-1}\ln(2),\right.$$

$$\left.-\frac{1}{4}(\beta_c - \beta_n)^2 + \frac{1}{2}\beta_c[\beta - \beta_n] + \beta t - o(1)\right\}. \quad (44)$$

If we now take

$$\beta < \beta_n + \frac{1}{2}\beta_c[1 - (\beta_n/\beta_c)]^2$$

then it is clear that by choosing t positive, but small enough, we will have a strictly negative limsup of the left hand side of (44) as $N \to \infty$. Choosing any

$0 < a < 1/2$, let us take $g(\beta) = \beta + a\beta_c[1 - (\beta/\beta_c)]^2$. Then for $\beta_{n+1} = g(\beta_n)$, we have proved the induction step: for β in the range $[0, \beta_{n+1}]$, inequality (38) also holds. □

We complete this section with the proof of Proposition 2.3.

Proof: By Lemma 2.7, the integrand of the third term in the right hand side of Eq. (41) approaches 0 uniformly as $N \to \infty$, as long as $0 \leq \beta < \beta_c$. This gives the desired result. □

3. THE GENERALIZED RANDOM ENERGY MODEL

In this section we extend the method developed in the previous section to treat the GREM. This is essentially a "correlated random energy model" on a hierarchical graph—that is, a tree.

3.1. Set-Up and Basics

The GREM is a family of models, taking various parameters for the definition. Let $n \in \mathbb{N}_+$ be an integer, equal to the number of levels in the hierarchical tree. Let $K_1, \ldots, K_n$ be positive integers such that $K_1 + K_2 + \cdots + K_n = N$, where N is the system size. Also, let $a_1, \ldots, a_n$ be real numbers such that $0 < a_i$ for $i = 1, \ldots, n$ and $a_1 + a_2 + \cdots + a_n = 1$.

Definition 3.1. Given $N \in \mathbf{N}$ and $\sigma \in \Sigma_N$, for $i = 1, \ldots, n$, let $\pi_i(\sigma)$ be the canonical projection over the subset Σ_{K_i} generated by the lexicographical partition $\mathcal{P}$ of the coordinates $(\sigma_1, \ldots, \sigma_N)$ into the first K_1 coordinates, the successive K_2 coordinates and so on up to the last K_n coordinates. Namely, $\Sigma_N = \Sigma_{K_1} \times \cdots \times \Sigma_{K_n}$, $\otimes_{i=1}^n \pi_i = 1_{\Sigma_N}$ and $\pi_i(\sigma) = (\sigma_{K_1+\ldots+K_{i-1}+1}, \ldots, \sigma_{K_1+\ldots+K_i})$.

Then the GREM Hamiltonian is a family of Gaussian random variables having the covariance

$$\mathbb{E}\left[H_N(\sigma)H_N(\sigma')\right] = \frac{N}{2} \sum_{i=1}^n a_i \prod_{j=1}^i \delta(\pi_j(\sigma), \pi_j(\sigma')) \tag{45}$$

An explicit form of GREM Hamiltonian is

$$H_N(\sigma) = \sqrt{\frac{N}{2}} \sum_{i=1}^n \sqrt{a_i} \, X(\pi_1(\sigma), \ldots, \pi_i(\sigma)). \tag{46}$$

where, for each $i = 1, \ldots, n$, the family of random variables $\{X(\pi_1(\sigma), \ldots, \pi_i(\sigma))\}_{\sigma \in \Sigma_N}$ are $2^{K_1+K_2+\ldots+K_i}$ i.i.d. Gaussians, and each family is independent of the others.

Remark 5. The Hamiltonian (46) corresponds to a tree with branching number that at each level is a power of two. We will stick to this case to simplify the notation, while the more general case of arbitrary branching number (with the constraint of having approximately 2^N leaves in the last layer) is completely equivalent in the thermodynamic limit. In order to make statements that apply in the limit, we will consider sequences of N's and $K_1, \ldots, K_n$ such that there are rational numbers $\kappa_1, \ldots, \kappa_n$, all nonnegative, and summing to 1, with $K_i = N\kappa_i$ for each $i = 1, \ldots, n$.

We now prove the variational expression for the pressure of the GREM. The strategy is to apply the results obtained for the REM model in the previous section at each level in the hierarchy. In order to denote the dependence on the parameters $a = (a_1, \ldots, a_n)$ and $\kappa = (\kappa_1, \ldots, \kappa_n)$, let us write the GREM pressure as

$$P_N^{(n)}(\beta; a, \kappa) = \frac{1}{N} \mathbb{E}\left[\ln Z_N(\beta; a, \kappa)\right]. \tag{47}$$

and its thermodynamic limit as $P^{(n)}(\beta; a, \kappa) := \lim_{N \to \infty} P_N^{(n)}(\beta; a, \kappa)$.

3.2. Upper Bound

Proposition 3.2. *Consider the GREM model, for which $C_H(\sigma, \sigma')$ is given by (45). For an index $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathcal{A}^n$ let the random weights $w(\alpha)$ be given by*

$$w(\alpha_1, \ldots, \alpha_n) = \exp\left[\frac{y(\alpha_1)}{m_1}\right] \exp\left[\frac{y(\alpha_1, \alpha_2)}{m_2}\right] \cdots \exp\left[\frac{y(\alpha_1, \ldots, \alpha_n)}{m_n}\right] \tag{48}$$

where the Poisson point processes is now a cascade with intensity measure $e^{-y}dy$. Namely, $y(\alpha_1)$ is the usual PPP, then for each given α_1, $y(\alpha_1, \alpha_2)$ is an independent copy (labelled by α_1) of the PPP, ... and so on up to $y(\alpha_1, \ldots, \alpha_n)$ which, for each given $\alpha_1, \ldots, \alpha_{n-1}$, is an independent copy of the PPP (labelled by $\alpha_1, \ldots, \alpha_{n-1}$). We also choose a sequence $0 < m_1 \leq m_2 \leq \cdots \leq m_n < 1$ and

$$C_K(\alpha, \alpha') = (b - 1)\frac{N}{2} \sum_{i=1}^{n} a_i \prod_{j=1}^{i} \delta(\alpha_j, \alpha'_j) \tag{49}$$

$$C_V(\sigma, \alpha, \sigma', \alpha') = b\frac{N}{2} \sum_{i=1}^{n} a_i \prod_{j=1}^{i} \delta(\pi_j(\sigma), \pi_j(\sigma')) \, \delta(\alpha_j, \alpha'_j), \tag{50}$$

where b is a real number such that b > 1. Then we obtain the optimal upper bound:

$$P_N^{(n)}(\beta; \boldsymbol{a}, \boldsymbol{\kappa}) \leq \inf_{0 < m_1 \leq \cdots \leq m_n < 1} \sum_{i=1}^{n} \left[\frac{\kappa_i}{m_i} \ln(2) + \frac{\beta^2}{4} m_i a_i \right]. \tag{51}$$

Proof: It will be along the lines of the proof of Proposition 2.2. Lemma 2.1 is applicable because for each $i = 1, \ldots, n$ one has

$$b \prod_{j=1}^{i} \delta(\pi_j(\sigma), \pi_j(\sigma')) \delta(\alpha_j, \alpha_j') \leq \prod_{j=1}^{i} \delta(\pi_j(\sigma), \pi_j(\sigma')) + (b-1) \prod_{j=1}^{i} \delta(\alpha_j, \alpha_j') \tag{52}$$

For the first term of Eq. (9), using Eq. (50), we have

$$\frac{1}{N} \mathbb{E} \left[\ln \sum_{\sigma, \alpha} w(\alpha) \exp[-\beta V(\sigma, \alpha)] \right]$$

$$= \frac{1}{N} \mathbb{E} \left(\ln \sum_{\sigma, \alpha} \exp \left[\frac{y(\alpha_1)}{m_1} \right] \exp \left[\frac{y(\alpha_1, \alpha_2)}{m_2} \right] \cdots \exp \left[\frac{y(\alpha_1, \cdots, \alpha_n)}{m_n} \right] \right.$$

$$\left. \times \exp \left[-\beta \sqrt{\frac{bN}{2}} \sum_{i=1}^{n} \sqrt{a_i} \, X(\pi_1(\sigma), \pi_2(\sigma), \ldots, \pi_i(\sigma), \alpha_1, \alpha_2, \ldots, \alpha_i) \right] \right) \tag{53}$$

Since the sum over configurations $\sigma \in \Sigma_N, \alpha \in \mathcal{A}^n$ can be decomposed into n sums over each subset $\pi_i(\sigma) \in \Sigma_{K_i}, \alpha_i \in \mathcal{A}$ for $i = 1, \ldots, n$, the invariance property (13) can now be applied telescopically, starting at the nth level and tracing back up to the first level. After this simplification we obtain

$$\frac{1}{N} \mathbb{E} \left[\ln \sum_{\sigma, \alpha} w(\alpha) \exp[-\beta V(\sigma, \alpha)] \right]$$

$$= \frac{1}{N} \mathbb{E} \left[\ln \sum_{\alpha} w(\alpha) \right] + \sum_{i=1}^{n} \frac{1}{m_i N} \ln \mathbb{E} \left[Z_{K_i}^{m_i}(\beta \sqrt{ba_i}) \right]$$

$$\leq \frac{1}{N} \mathbb{E} \left[\ln \sum_{\alpha} w(\alpha) \right] + \sum_{i=1}^{n} \frac{1}{m_i N} \ln \mathbb{E} \left[Z_{K_i}(m_i \beta \sqrt{ba_i}) \right]$$

$$= \frac{1}{N} \mathbb{E} \left[\ln \sum_{\alpha} w(\alpha) \right] + \sum_{i=1}^{n} \left[\frac{\kappa_i}{m_i} \ln(2) \right] + \frac{\beta^2}{4} b \sum_{i=1}^{n} a_i m_i \tag{54}$$

where in the third line we have made use again of Eq. (23). For the second term of Eq. (9), using Eq. (49) and the invariance property (13) we have

$$\frac{1}{N} \mathbb{E} \left[\ln \sum_\alpha w(\alpha) \exp[-\beta K(\alpha)] \right]$$

$$= \frac{1}{N} \mathbb{E} \left[\ln \sum_\alpha w(\alpha) \right] + \beta^2 \frac{(b-1)}{4} \sum_{i=1}^n a_i m_i \qquad (55)$$

Putting together Eqs. (54) and (55) we arrive at the upper bound stated in the Proposition. □

Remark 6. In the following we make the assumption

$$\frac{\kappa_1}{a_1} < \frac{\kappa_2}{a_2} < \cdots < \frac{\kappa_n}{a_n} \qquad (56)$$

in order to have a totally nondegenerate sequence of transition temperatures. To express the inequality of (51) in a more transparent form it is convenient to introduce a succession of critical temperatures: for $i = 1, \ldots, n$ let $\beta_i^* = \beta_c \sqrt{\frac{\kappa_i}{a_i}}$ (where $\beta_c = 2\sqrt{\ln(2)}$ as in the REM). Under the condition (56), this implies $\beta_1^* < \cdots < \beta_n^*$. Because of the constraint $0 < m_1 \leq \cdots \leq m_n < 1$ the optimal m_i is

$$m_i = \min\{1, \beta_i^*/\beta\} \qquad (57)$$

for $i = 1, \ldots, n$, the value 1 being attained by taking $m_i \uparrow 1$ in the infimum of Eq. (51). Therefore, one has

$$P_N^{(n)}(\beta; \boldsymbol{a}, \boldsymbol{\kappa}) \leq \mathcal{Q}^{(n)}(\beta; \boldsymbol{a}, \boldsymbol{\kappa}), \qquad (58)$$

where

$$\mathcal{Q}^{(n)}(\beta; \boldsymbol{a}, \boldsymbol{\kappa})$$

$$= \begin{cases} \displaystyle\sum_{k=1}^n \frac{1}{4} a_k(\beta^2 + (\beta_k^*)^2) & \text{for } \beta < \beta_1^* \\[2em] \displaystyle\sum_{k=1}^i \frac{1}{2} a_k \beta \beta_k^* + \sum_{k=i+1}^n \frac{1}{4} a_k(\beta^2 + (\beta_k^*)^2) & \text{for } \beta_i^* \leq \beta \leq \beta_{i+1}^* \\[2em] \displaystyle\sum_{k=1}^n \frac{1}{2} a_k \beta \beta_k^* & \text{for } \beta \geq \beta_n^* \end{cases} \qquad (59)$$

3.3. Lower Bound

Let us denote the REM pressure (Eq. (5)) as $P_N^{(1)}(\beta)$, and its thermodynamic limit (Eq. (6)) as $P^{(1)}(\beta)$. This is not really an abuse of notation because if

$n = 1$ then the GREM is the REM, and $a_1 = \kappa_1 = 1$. In the same way, we write $\mathcal{Q}^{(1)}(\beta) = \mathcal{Q}(\beta)$, where $\mathcal{Q}(\beta)$ is defined in Eq. (28).

Then the lower bound is the following:

Proposition 3.3. *For all $\beta \geq 0$,*

$$P^{(n)}(\beta; \boldsymbol{a}, \boldsymbol{\kappa}) \geq \mathcal{Q}^{(n)}(\beta; \boldsymbol{a}, \boldsymbol{\kappa}). \tag{60}$$

Proof: The proof will follow if we show that

$$P_N^{(n)}(\beta; \boldsymbol{a}, \boldsymbol{\kappa}) \geq \sum_{i=1}^n \kappa_i P_{K_i}^{(1)}\left(\sqrt{a_i/\kappa_i}\,\beta\right). \tag{61}$$

Indeed, taking the thermodynamic limit $N \to \infty$ on both sides and using

$$\mathcal{Q}^{(n)}(\beta; \boldsymbol{a}, \boldsymbol{\kappa}) = \sum_{i=1}^n \kappa_i \mathcal{Q}^{(1)}(\sqrt{a_i/\kappa_i}\,\beta) \quad \text{for} \quad \text{for all } \beta \geq 0. \tag{62}$$

we obtain the Lemma statement. To prove (61) we introduce the interpolating pressure

$$\frac{1}{N} \mathbb{E} \ln \sum_{\sigma \in \Sigma_N} e^{-\beta \tilde{H}(\sigma, t)} \tag{63}$$

with

$$\tilde{H}(\sigma, t) = \sqrt{t} \left[\sqrt{\frac{N}{2}} \sum_{i=1}^n \sqrt{a_i}\, X(\pi_1(\sigma), \ldots, \pi_i(\sigma)) \right]$$

$$+ \sqrt{1-t} \left[\sqrt{\frac{N}{2}} \sum_{i=1}^n \sqrt{a_i}\, Y(\pi_i(\sigma)) \right] \tag{64}$$

where the X's and Y's are families of i.i.d. Gaussian random variables, each independent from the other. A straightforward differentiation of Eq. (63) combined with integration by parts yields Eq. (61). □

APPENDIX: PROOF OF LEMMA 2.6: CONCENTRATION OF MEASURE

In this Appendix we prove the "concentration of measure" inequality (Lemma 2.6). The proof is a generalization[5] of the proof for the SK model (Corollary 2.2.5 of Ref. 29).

---

[5] In Ref. 29 the generalization of Corollary 2.2.5 to p-spin models is implicit: for example in the proof of Theorem 6.1.2. But our generalization to all models satisfying "thermodynamic stability" is new.

We show that the standard deviation inequality apply to a large class of Gaussian spin-glass models, which includes both mean-field (SK, p-spin, REM, GREM) and finite dimensional models (Edwards-Anderson, Random Field). We recall the basic result for a function of Gaussian variables:

Theorem A.1. (Talagrand) *Consider a Lipschitz function F on $\mathbb{R}^M$, of Lipschitz constant A. If $J_1, \ldots, J_M$ are independent, standard normal random variables, then for each $t > 0$,*

$$\mathbb{P}\{|F(\boldsymbol{J}) - \mathbb{E}[F(\boldsymbol{J})]| \geq At\} \leq 2e^{-t^2/4}. \tag{65}$$

This is Theorem 2.2.4 of Ref. 29. Talagrand proves this using the "smart path" method, which is his adaptation of the quadratic interpolation argument. Thus, his proof in Ref. 29 differs from his earlier proofs[28] and from the proofs of others.[24] This is good for those studying spin glasses. Particularly, one technique is unifying and simplifying various tools. Another application of quadratic interpolation is continuity of the pressure with respect to the covariance of a spin glass Hamiltonian: Corollary 3.3 of Ref. 2. Also, a result which uses the same idea, and surprisingly predates the applications in spin glasses, is Slepian's lemma.[21]

Let us consider the general Hamiltonian given by

$$H_\Lambda(\sigma; \boldsymbol{J}) = -\frac{1}{\sqrt{2}} \sum_{X \in \Lambda} \Delta_X J_X \sigma_X, \tag{66}$$

where $\Lambda \subset \mathbb{Z}^d$, $\sigma_X = \prod_{i \in X} \sigma_i$, $\{\Delta_X\}_{X \in \Lambda} \geq 0$ and the $\{J_X\}_{X \in \Lambda}$ are a family of i.i.d. standard Gaussian random variables, $\mathbb{E}[J_X] = 0$, $\mathbb{E}[J_X J_Y] = \delta_{X,Y}$. In order to have a bounded quenched pressure we assume the following thermodynamic stability condition holds: there exists a constant $c < \infty$ such that

$$\sup_{\Lambda \subset \mathbb{Z}^d} \frac{1}{|\Lambda|} \sum_{X \subset \Lambda} \Delta_X^2 \leq c \tag{67}$$

We remark that this condition immediately entails the existence of thermodynamic limit for short-range models[12] and it also implies the validity of the Ghirlanda-Guerra identities both for short-range and mean-field models.[10,11,17] To prove Lemma 2.6, we need to show that the random pressure function

$$P_\Lambda(\boldsymbol{J}) = \frac{1}{|\Lambda|} \ln \sum_{\sigma \in \Sigma_N} e^{-\beta H_\Lambda(\sigma; \boldsymbol{J})} \tag{68}$$

is Lipschitz. For this, we note that

$$P_\Lambda(\boldsymbol{J}) - P_\Lambda(\boldsymbol{J}') = \int_0^1 \frac{dP_\Lambda(t\boldsymbol{J} + (1-t)\boldsymbol{J}')}{dt} \, dt. \tag{69}$$

On the other hand

$$\frac{d P_\Lambda(t\,\boldsymbol{J} + (1-t)\boldsymbol{J}')}{dt} = \frac{\beta}{\sqrt{2}\,|\Lambda|} \sum_{X \subset \Lambda} \omega_t(\sigma_X)\Delta_X(J_X - J'_X) \tag{70}$$

with

$$\omega_t(\sigma_X) = \frac{\sum_\sigma \sigma_X e^{-\beta H_\Lambda(\sigma; t\,\boldsymbol{J} + (1-t)\boldsymbol{J}')}}{\sum_\sigma e^{-\beta H_\Lambda(\sigma; t\,\boldsymbol{J} + (1-t)\boldsymbol{J}')}} \tag{71}$$

From Eq. (69), by using $|\sigma_X| \leq 1$, Cauchy–Schwarz inequality and the thermodynamic stability condition (67), it then follows

$$|P_\Lambda(\boldsymbol{J}) - P_\Lambda(\boldsymbol{J}')| \leq \beta \sqrt{\frac{c}{2|\Lambda|}} \, \|\boldsymbol{J} - \boldsymbol{J}'\| . \tag{72}$$

where $\| \cdot \|$ denotes the L_2-norm. Therefore, $P_\Lambda(\boldsymbol{J})$ is Lipschitz, with Lipschitz constant $A = \beta\sqrt{\frac{c}{2|\Lambda|}}$. Applying Theorem A.1 it gives

$$\mathbb{P}\{|P_\Lambda(\boldsymbol{J}) - \mathbb{E}[P_\Lambda(\boldsymbol{J})]| \geq t\} \leq 2 \exp\left(-\frac{t^2|\Lambda|}{2c\beta^2}\right) . \tag{73}$$

This result apply equally well to all general Hamiltonian of the form (66). The REM model is obtained from Eq. (66) with the choice $|\Lambda| = N$, $\Delta_X = \sqrt{N2^{-N}}$ (this is indeed Hamiltonian (3)). In this case the condition (67) gives $c = 1$ and the statement of Lemma 2.6 is proved.

ACKNOWLEDGMENTS

REFERENCES

1. M. Aizenman, R. Sims and S. Starr, Extended variational principle for the Sherrington-Kirkpatrick spin-glass model. *Phys. Rev. B* **68**:214403 (2003).
2. M. Aizenman, R. Sims and S. Starr, Mean-Field Spin Glass models from the Cavity–ROSt perspective. *Preprint* (2006) math-ph/0607060.
3. E. Bolthausen and N. Kistler, On a non-hierarchical version of the Generalized Random Energy Model. *Ann. Appl. Probab.* **16**:1–16 (2006).
4. E. Bolthausen and A.-S. Sznitman, On Ruelle's probability cascades and an abstract cavity method. *Comm. Math. Phys.* **197**:247–276 (1998).
5. A. Bovier, *Statistical Mechanics of Disordered Systems: A Mathematical Perspective* (Cambridge University Press, Cambridge, UK, 2006).

6. A. Bovier and I. Kurkova, Derrida's generalised random energy models. I. Models with finitely many hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.* **40**:439–480 (2004).

7. A. Bovier and I. Kurkova, Derrida's generalised random energy models. II. Models with continuous hierarchies. *Ann. Inst. H. Poincaré Probab. Statist.* **40**:481–495 (2004).

8. A. Bovier and I Kurkova, Local energy statistics in disordered systems: a proof of the local REM conjecture. *Commun. Math. Phys.* **263**:513–533 (2006).

9. P. Contucci, M. Degli Esposti, C. Giardinà and S. Graffi, Thermodynamical limit for correlated Gaussian random energy models. *Commun. Math. Phys.* **236**:55–63 (2003).

10. P. Contucci and C. Giardinà, The Ghirlanda-Guerra identities. *Preprint* (2005) math-ph/0505055, to appear in *J. Statist. Phys.*

11. P. Contucci and C. Giardinà, Spin-glass stochastic stability: A rigorous proof. *Ann. Henri Poincaré* **6**:915–923 (2005).

12. P. Contucci and S. Graffi, Monotonicity and thermodynamic limit for short range disordered models. *J. Statist. Phys.* **115**:581–589 (2004).

13. B. Derrida, Random-energy model: limit of a family of disordered models. *Phys. Rev. Lett.* **45**:79–82 (1980).

14. B. Derrida and E. Gardner, Magnetic properties and the function $q(x)$ of the generalised random-energy model. *J. Phys. C: Solid State* **19**:5783–5798 (1986).

15. T. C. Dorlas and J. R. Wedagedera, Large deviations and the random energy model. *Internat. J. Modern Phys. B* **15**:1–15 (2001).

16. T. Eisele, On a third-order phase transition. *Comm. Math. Phys.* **90**:125–159 (1983).

17. S. Ghirlanda and F. Guerra, General properties of overlap probability distributions in disordered spin systems. Towards Parisi ultrametricity. *J. Phys. A: Math. Gen.* **31**:9149–9155 (1998).

18. F. Guerra, Broken replica symmetry bounds in the mean field spin glass model. *Commun. Math. Phys.* **233**:1–12 (2003).

19. F. Guerra and F. L. Toninelli, The thermodynamic limit in mean field spin glass models. *Commun. Math. Phys.* **230**:71–79 (2002).

20. F. Guerra and F. L. Toninelli, Quadratic replica coupling in the Sherrington-Kirkpatrick mean field spin glass model. *J. Math. Phys.* **43**:3704–3717 (2002).

21. K. Joag-Dev, M. D. Perlman and L. D. Pitt, Association of normal random variables and Slepian's Inequality. *Ann. Probab.* **11**:451–455 (1983).

22. M. Mezard, G. Parisi and M. A. Virasoro, *Spin Glass Theory and Beyond.* (World Scientific, Singapore, 1987).

23. E. Olivieri and P. Picco, On the existence of thermodynamics for the random energy model. *Commun. Math. Phys.* **96**:125–144 (1984).

24. G. Pisier, Probabilistic methods in the geometry of Banach spaces. in *Probability and analysis (Varenna, 1985)* Lecture Notes in Math., v. 1206, pp. 167–241, Springer, Berlin (1986).

25. J. Pitman and M. Yor, The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25**:855–900 (1997).

26. D. Ruelle, A mathematical reformulation of Derrida's REM and GREM. *Commun. Math. Phys.* **108**:225–239 (1987).

27. D. Sherrington and S. Kirkpatrick, Solvable model of a spin-glass. *Phys. Rev. Lett.* **35**:1792–1796 (1975).

28. M. Talagrand, Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.* **81**:73–205 (1995).

29. M Talagrand, *Spin Glasses: A Challenge for Mathematicians* (Springer, Berlin, 2003).

30. M. Talagrand, The Parisi formula. *C. R. Math. Acad. Sci. Paris* **337**:111–114 (2003); *Ann. Math.* **163**:221–263 (2006).

31. M. Talagrand, Large deviations, Guerra's and A.S.S. Schemes, and the Parisi hypothesis. *Preprint* (2005). To appear in the proceedings of the Cortona conference.