GENERALIZED RANDOM ENERGY MODELS IN A TRANSVERSAL MAGNETIC FIELD: FREE ENERGY AND PHASE DIAGRAMS

Chokri Manai and Simone Warzel

3:2 2022
GENERALIZED RANDOM ENERGY MODELS IN A TRANSVERSAL MAGNETIC FIELD: FREE ENERGY AND PHASE DIAGRAMS

CHOKRI MANAI AND SIMONE WARZEL

We determine explicit variational expressions for the quantum free energy of mean-field spin glasses in a transversal magnetic field, whose glass interaction is given by a hierarchical Gaussian potential as in Derrida’s generalized random energy model (GREM), its continuous version (CREM), or the nonhierarchical GREM. The corresponding phase diagrams, which generally include glass transitions as well as transitions to quantum paramagnetic and mixed quantum-classical phases, are discussed. In the glass phase, the free energy is generally determined by both the parameters of the classical model and the transversal field.

1. Introduction

Studying the fate of spin glass physics with respect to quantum effects induced by a transversal field has been a topic of continuing interest in the physics community. In the past 10 years this subject received an additional boost due to its relevance as a testing ground for quantum adiabatic algorithms and for many-body localized systems [1; 6; 13; 23]. Ever since exact solutions of the free energy of mean-field spin glasses became available, Parisi’s famous replica calculations [29; 33] have been extended to approximations of the quantum free energy. Notwithstanding numerous works (see, e.g., [12; 20; 30; 21; 37; 41]), an ultimate consensus on various aspects of quantum spin glasses such as the quantum Sherrington–Kirkpatrick (SK) model seems to be lacking even from the physics point of view. From the point of view of disordered many-body systems the field offers toy models for exotic behavior of many-body wave functions on Fock space. A hint to the existence of partially nonergodic states is provided by the results of this work.

Although the theory classical mean-field spin glasses became an established branch of probability [8; 22; 31; 39; 40], efforts of mathematicians in the area of quantum glasses are so far rather limited. Crawford [15] laid the foundations for the quantum SK model. For this model, by generalizing classical arguments from [3], the absence of replica symmetry breaking in the high-temperature regime was proved in [25], and, more recently, the existence of spin-glass order at low temperatures and transversal field could also be established [24]. Adhikari and Brennecke [2] used a path-integral approach and Parisi’s formula for vector-spin models to rewrite the free energy of the quantum SK model as a rather involved variational problem in terms of infinite-dimensional path overlaps.

MSC2020: 82B44, 82D30.

Keywords: quantum spin glass, phase transition, free energy.
The main aim of this work is to derive reasonably explicit variational expressions, which allow us to determine the structure of the phase diagram, for the quantum versions of three classic hierarchical and related mean-field spin glasses: (i) Derrida’s generalized random energy model (GREM) [18], (ii) its continuous version (CREM), and (iii) the nonhierarchical GREM by Bolthausen and Kistler [7]. These models were invented so as to incorporate the effects of correlations of energy levels into the oversimplified random energy model (REM) [16; 17]. The GREM’s and CREM’s built-in ultra-metric structure constitutes the backbone of Parisi’s solution of the SK mean-field spin glass — a fact which received its mathematical blessing through Talagrand’s proof [38] and further analysis of the Gibbs measure [31].

Although the built-in ultra-metric structure and the prearrangement of types in the GREM or CREM are somewhat artificial as long as one does not clarify the role of ultrametricity in quantum glasses, it is nevertheless surprising that no physics prediction exists for the quantum version of these classic hierarchical mean-field spin glasses. All the more so since in 1990, Goldschmidt presented his formula for the free energy of the quantum REM [21], which was recently confirmed through a mathematical analysis [26]. This gap is closed with the present paper. We find formulae which express the principle that the types decide within the groups whether to collectively follow the transversal field or stay in their classical order. The free energy is then computed as a minimum over all group decompositions. We call this principle hierarchical peeling. It is based purely on a combination of a probabilistic-geometric decomposition of the spin-configuration space and operator-theoretic techniques, which are further developments of ideas in [26; 27]. In passing, we also generalize basic interpolation techniques to the quantum set-up. These main new technical tools are presented in Section 2.

We start the paper with a short introduction to classical hierarchical models. The quantum free energy is then presented in Section 1B. The introduction closes with a discussion of the nonhierarchical GREM and its quantum Parisi-type formula. The proofs of the novel quantum formulae are postponed to Section 3.

1A. Classical GREM and CREM. The GREM and CREM potential $U$ is a centered Gaussian process on the Hamming cube $Q_N := \{-1, 1\}^N$, whose covariance matrix is given by

$$
\mathbb{E} [U(\sigma)U(\sigma')] = NA(q(\sigma, \sigma')),
$$

(1-1)

where $A : [0, 1] \rightarrow [0, 1]$ is a nondecreasing, right-continuous, and normalized function, i.e., $A(1) = 1$, which does not depend on $N$. Moreover, $q$ denotes the normalized lexicographic overlap of spin configurations $\sigma, \sigma' \in Q_N$, i.e.,

$$
q(\sigma, \sigma') := \begin{cases} 1 & \text{if } \sigma = \sigma', \\
\frac{1}{N}(\min\{1 \leq i \leq N; \sigma_i \neq \sigma_i'\} - 1) & \text{else}.
\end{cases}
$$

(1-2)

The induced metric $d_A(\sigma, \sigma') = \mathbb{E} [\|U(\sigma) - U(\sigma')\|^2]^{1/2}$ on the Hamming cube is an ultrametric.

In the GREM one further assumes that the distribution function $A$ is a step function with $n \in \mathbb{N}$ jumps of height $a_k$ at the values $0 = x_0 < x_1 < x_2 < \cdots < x_n = 1$. The Gaussian potential $U$ can then be expressed in terms of independent standard Gaussian variables. To this end, one decomposes $\sigma \in Q_N$
into \( n \) blocks (“types”), \( \sigma = (\sigma_1, \ldots, \sigma_n) \), each of which is represented by a spin vector on a reduced Hamming cube,

\[
\sigma_k \in \mathcal{Q}_N^{(k)} := \{-1, 1\}^{[x_kN] - [x_{k-1}N]}, \quad k \in \{1, \ldots, n\}.
\] (1-3)

Introducing independent standard Gaussian variables \( X_{\sigma_1}, X_{\sigma_1\sigma_2}, \ldots, X_{\sigma_1\sigma_2\cdots\sigma_n} \) one then rewrites

\[
U(\sigma) = \sqrt{N}(\sqrt{a_1}X_{\sigma_1} + \sqrt{a_2}X_{\sigma_1\sigma_2} + \cdots + \sqrt{a_n}X_{\sigma_1\sigma_2\cdots\sigma_n})
\] (1-4)
in the sense of distributional equality. The pressure or negative free energy

\[
\Phi_N(\beta) := \frac{1}{N} \ln Z_N(\beta)
\]
is given in terms of the partition function \( Z_N(\beta) := \sum_{\sigma \in \mathcal{Q}_N} e^{-\beta U(\sigma)} \), and converges for any distribution function \( A \) almost surely [14]. The limit depends on the concave hull \( \bar{A} \) of \( A \), i.e., the smallest concave function which is greater than or equal to \( A \). In the GREM, the concave hull \( \bar{A} \) is a piecewise linear function determined by the values \( \{y_0, y_1, \ldots, y_m\} \subset \{x_0, x_1, \ldots, x_n\} \) where \( A \) and \( \bar{A} \) agree. The increments of the concave hull \( \bar{a}_l := A(y_l) - A(y_{l-1}) \), the interval lengths \( L_l := y_l - y_{l-1} \), and the slopes \( \gamma_l := \bar{a}_l/L_l \) determine the limit of the pressure, which is given by [19; 14]

\[
\lim_{N \to \infty} \Phi_N(\beta) = \Phi(\beta) = \sum_{l=1}^{m} \varphi^{(l)}(\beta)
\] (1-5)

with the partial pressures

\[
\varphi^{(l)}(\beta) := \begin{cases} 
\frac{1}{2}\beta^2 \bar{a}_l + L_l \ln 2 & \text{if } \beta \leq \sqrt{(2\ln 2)\gamma_l^{-1}} =: \beta_l, \\
\beta \sqrt{(2\ln 2)} \bar{a}_l L_l & \text{if } \beta > \sqrt{(2\ln 2)\gamma_l^{-1}}.
\end{cases}
\] (1-6)

(For future reference, we note that this formula still holds if the weights \( (a_k) \) do not add up to 1.) The glass transition in the GREM occurs in steps with the components of the systems’ spins corresponding to \( l \) freezing at \( \beta_l \). Since \( \beta_m > \cdots > \beta_2 > \beta_1 \), the highest freezing temperature is found at \( \beta_c = \beta_1 \).

The CREM includes distribution functions \( A \) which are not step functions. Since they can be represented as a (uniform) limit of step functions, it is not surprising that the corresponding limit of the pressure \( \Phi(\beta) \) turns into an integral. The increments \( \bar{a}_l \) are replaced by the right derivative \( \bar{a}(x) \) of \( \bar{A}(x) \) which exists everywhere as a consequence of the convexity of \( \bar{A} \). This allows one to give an explicit expression for the limiting pressure [11; 8]

\[
\Phi(\beta) = \sqrt{2\ln 2} \beta \int_{0}^{x(\beta)} \sqrt{\bar{a}(x)} \, dx + \frac{\beta^2}{2} \left( 1 - \bar{A}(x(\beta)) \right) + (1 - x(\beta)) \ln 2
\] (1-7)

with the function

\[
x(\beta) := \sup \{x \mid \bar{a}(x) > (2\ln 2)/\beta^2 \}.
\] (1-8)

The glass transition in the CREM occurs at \( \beta_c = \sqrt{(2\ln 2)/\lim_{x \to 0} \bar{a}(x)} \).
1B. Quantum GREM and CREM and a Parisi formula. If a transversal magnetic field in the \( x \)-direction is turned on, the total Hamiltonian acting on the Hilbert space \( \ell^2(\mathbb{Q}_N) \), which is unitarily equivalent to the tensor product \( \otimes_{j=1}^N \mathbb{C}^2 \), is

\[
H_N = U - B, \tag{1-9}
\]

where \( B \) is the sum of the \( x \)-Pauli matrices \( s_j \) with (possibly random) weights \( b_j \in \mathbb{R} \),

\[
(B \psi)(\sigma) := \sum_{j=1}^N b_j (s_j \psi)(\sigma), \quad (s_j \psi)(\sigma) := \psi(\sigma_1, \ldots, -\sigma_j, \ldots, \sigma_N), \tag{1-10}
\]

and \( U \) acts as a random potential. Before further specifying \( U \) and \( B \), we record a simple observation: the quantum partition function \( \text{Tr} e^{-\beta (U-B)} \) and the diagonal matrix-elements of \( e^{-\beta (U-B)} \) in terms of the standard orthonormal \( z \)-basis, \( \{|\sigma\rangle \mid \sigma \in \mathbb{Q}_N\} \) for which \( \langle \sigma | \psi \rangle = \psi(\sigma) \), only depend on the absolute values \( |b_j| \). Here and in the following we use Dirac’s bra-ket notation for matrix elements and scalar products.

**Lemma 1.1.** Let \( U \) be an arbitrary potential on \( \mathbb{Q}_N \) and \( B, B' \) two transversal field with weights \( b_j \) and \( b'_j \) which only differ by a sign, i.e., \( |b_j| = |b'_j| \) for all \( j \). Then, for all \( \sigma \in \mathbb{Q}_N \):

\[
\langle \sigma | e^{-\beta (U-B)} | \sigma \rangle = \langle \sigma | e^{-\beta (U-B')} | \sigma \rangle. \tag{1-11}
\]

**Proof.** Expanding the exponential, we write \( \langle \sigma | e^{-\beta (U-B)} | \sigma \rangle \) as a convergent series of terms of the form

\[
\langle \sigma | A_1 \cdots A_k | \sigma \rangle \tag{1-12}
\]

where each \( A_j \) is either \(-U\) or some \( b_j s_j \). As \( s_j \) flips the sign of the \( j \)-th coordinate \( \sigma_j \), the term \( (1-12) \) vanishes unless each operator \( s_j \) occurs \( n_j \) times, where \( n_j \) is an even number. We conclude that \( \langle \sigma | e^{-\beta (U-B)} | \sigma \rangle \) only depends on the squares \( b_j^2 \) which proves (1-11).

If all the weights \( b_i \geq 0 \) are nonnegative, the Trotter product formula shows that \( H_N \) generates a positive semigroup, i.e., for any \( t \geq 0 \) and \( \sigma, \sigma' \in \mathbb{Q}_N \),

\[
\langle \sigma | e^{-t H_N} | \sigma' \rangle \geq 0.
\]

This is in general not true for an arbitrary transversal magnetic field \( B \), but due to Lemma 1.1 we can assume without loss of generality that the weights \( (b_j) \) are indeed nonnegative if we are only interested in properties which can be derived from diagonal matrix elements such as the quantum partition function.

In the remaining part of this section, we restrict ourselves to the case where the weights \( (b_j) \) are independent copies of an absolutely integrable random variable \( b \) and they shall be independent of the Gaussian potentials \( U \). We are mainly interested in the thermodynamic properties of the hierarchical quantum spin glasses which are encoded in the quantum partition function

\[
Z_N(\beta, b) := \text{Tr} [e^{-\beta H_N}]
\]
or, equivalently, in the pressure (or negative free energy)

\[ \Phi_N(\beta, b) := \frac{1}{N} \ln Z_N(\beta, b). \]

In the special case that the weights \( b = \Gamma \) are (almost surely) constant, we will sometimes write \( B = \Gamma T \) and denote the pressure by \( \Phi_N(\beta, \Gamma) \).

Our first main result concerns the free energy of the QGREM. We show that the free energy converges almost surely to a nonrandom limit, for which we derive an explicit expression in terms of the classical partial free energies (1-6) and the paramagnetic free energy. With the notation of Section 1A, we have the following:

**Theorem 1.2.** For the GREM specified by \( U \) as in (1-4) in terms of its distribution function \( A \), any \( \beta \geq 0 \), and an absolutely integrable random variable \( b \), the quantum free energy converges almost surely:

\[
\lim_{N \to \infty} \Phi_N(\beta, b) = \Phi(\beta, b) := \max_{0 \leq k \leq m} \left[ \sum_{l=1}^{k} \varphi^{(l)}(\beta) + (1 - y_k)E[\ln(2 \cosh(\beta b))] \right]. \tag{1-13}
\]

The maximum is taken over all points \( \{y_0, y_1, \ldots, y_m\} \) supporting the convex hull \( \tilde{A} \) of \( A \).

The proof of this theorem is found in Section 3A. We stress that as in the classical case the concave hull \( \tilde{A} \), rather than \( A \), remains the determining function for the limit. The second term in (1-13) is the pressure of the random quantum paramagnet given by

\[
p(\beta b) := \frac{1}{N} E[\ln \text{Tr}[e^{\beta B}]] = \frac{1}{N} E\left[ \ln \prod_{j=1}^{N} \text{Tr}[e^{\beta b_j} s_j] \right]
= \frac{1}{N} \sum_{j=1}^{N} E[\ln(2 \cosh(\beta b_j))] = E[\ln(2 \cosh(\beta b))]. \tag{1-14}
\]

If \( b = \Gamma > 0 \) is constant, the structure of the limit in (1-13) becomes more transparent if we introduce the critical field strengths

\[
\Gamma^{(l)}_c := \frac{1}{\beta} \text{arccosh}\left(\frac{1}{2} \exp\left(\frac{\varphi^{(l)}(\beta)}{L_l}\right)\right), \quad l \in \{1, \ldots, m\}.
\]

In this situation, we may rephrase (1-13) as follows:

**Corollary 1.3.** In the situation of Theorem 1.2 with \( b = \Gamma > 0 \),

\[
\Phi(\beta, \Gamma) = \sum_{l=1}^{m} \left( \varphi^{(l)}(\beta) 1_{\Gamma < \Gamma^{(l)}_c} + L_l \ln(2 \cosh(\beta \Gamma)) 1_{\Gamma \geq \Gamma^{(l)}_c} \right). \tag{1-15}
\]

The proof is again found in Section 3A. The free energy coincides with the sum of \( m \) weighted and rescaled QREM terms; see [21; 27]. In particular, there are as many magnetic first-order transitions as second-order glass transitions. The glass transitions continue to occur at the (classical) critical inverse
Figure 1. Phase diagram of the quantum GREM as a function of the transversal constant magnetic field $\Gamma$ and the temperature $\beta^{-1}$. The figure shows an example with three second-order glass transitions (dotted lines) and three first-order magnetic transitions (bold lines). If $\Gamma < \Gamma_c^{(3)}(\beta_l)$, the free energy coincides with the classical one ($\Gamma = 0$), whereas, for $\Gamma > \Gamma_c^{(l)}(\beta_l)$, the system becomes a pure quantum paramagnet. In between mixed quantum-classical phases appear.

Temperatures $\beta_l = \sqrt{(2 \ln 2)\gamma^{-1}_{l}}$ as long as $\Gamma < \Gamma_c^{(l)}(\beta_l)$ and disappear for field strengths $\Gamma > \Gamma_c^{(l)}(\beta_l)$; see Figure 1. The specific magnetization in the $x$-direction

$$m_x(\beta, \Gamma) := \frac{1}{\beta} \frac{\partial}{\partial \Gamma} \Phi(\beta, \Gamma)$$

takes discontinuously at $\Gamma = \Gamma_c^{(l)}$. The physics described by (1-15) is that of the block or types of spins corresponding to $l$ flipping into transversal order at $\Gamma = \Gamma_c^{(l)}$. At temperatures below $\beta_l^{-1}$, the transition is from spin-glass order to a quantum paramagnet in that block. At zero-temperature we find quantum phase transitions at $\Gamma = \lim_{\beta \to \infty} \Gamma_c^{(l)} = \sqrt{(2 \ln 2)\gamma^{(l)}}$. It is an interesting question to determine the fate of Parisi’s order parameter as well as the structure of the eigenvectors in the different low-temperature regimes. The picture suggested by our results on the free energy would point to functions which violate ergodicity partially in a mixed low-temperature phase in that they will be extended on a fraction of the Hamming cube only. The rigorous classical analysis of Parisi’s order parameter for the GREM, which partially captures the geometric structure of the Gibbs measure, can be found in [10]. An extension of this analysis will be the subject of a future work.
Moving on to the more general CREM potentials, it is convenient to introduce truncated versions of the free energy in (1-7). For any \( z \in [0, 1] \), we define

\[
\Phi(\beta, z) := \sqrt{2 \ln 2} \beta \int_0^{\min\{x(\beta), z\}} \sqrt{\tilde{a}(x)} \, dx + \mathbb{1}_{z > x(\beta)} \left( \frac{\beta^2}{2} (\tilde{A}(z) - \tilde{A}(x(\beta))) + (z - x(\beta)) \ln 2 \right). \tag{1-16}
\]

As in the quantum GREM, the free energy of the quantum CREM converges almost surely and the limit may be expressed as a variational formula involving \( \Phi(\beta, z) \):

**Theorem 1.4.** For the CREM specified by \( U \) as in (1-1) in terms of its distribution function \( A \), let \( \tilde{A} \) be the concave hull of \( A \), \( \tilde{a} \) the right-derivative of \( \tilde{A} \), \( \Phi(\beta, z) \) as in (1-16) and \( b \) an absolutely integrable random variable. Then, the quantum pressure \( \Phi_N(\beta, b) \) converges almost surely,

\[
\lim_{N \to \infty} \Phi_N(\beta, b) = \Phi(\beta, b) := \sup_{0 \leq z \leq 1} \left[ \Phi(\beta, z) + (1 - z) \mathbb{E} [\ln 2 \cosh(\beta b)] \right]. \tag{1-17}
\]

The proof is found in Section 3A.

The free energy \( \Phi_N(\beta, b) \) does not only converge almost surely, but also in mean. This is a consequence of the usual Gaussian concentration of measure estimate, i.e., a special case of Proposition 2.9. If \( b \) is even an \( L^r \)-random variable for some \( r > 1 \), the pressure even converges in \( r \)-th mean.

In order to determine the order of occurring magnetic phase transitions, we will replace the variational formula (1-17) in the case \( b = 0 \) by a more explicit expression. To this end, we assume from now that the concave hull \( \tilde{A} \) is a continuously differentiable function different from the identity (in order to exclude the QREM situation). Since \( \tilde{a}(x(\beta)) = 2 \ln 2 / \beta^2 \), \( \Phi(\beta, z) \) is differentiable with respect to \( z \) with derivative

\[
\frac{\partial \Phi(\beta, z)}{\partial z} = \sqrt{2 \ln 2} \tilde{a}(z) \beta \mathbb{1}_{z < x(\beta)} + \left( \ln 2 + \frac{\beta^2}{2} \tilde{a}(z) \right) \mathbb{1}_{z \geq x(\beta)}.
\]

We note that \( \frac{\partial \Phi(\beta, \cdot)}{\partial z} : [0, 1] \to [s(\beta), t(\beta)] \) is a nondecreasing continuous function with values in the closed interval specified by its boundary values

\[
s(\beta) := \left. \frac{\partial \Phi(\beta, z)}{\partial z} \right|_{z = 1} \quad \text{and} \quad t(\beta) := \left. \frac{\partial \Phi(\beta, z)}{\partial z} \right|_{z = 0}.
\]

**Corollary 1.5.** Let \( g(\beta, \cdot) : [s(\beta), t(\beta)] \to [0, 1] \) be a (generalized) inverse of the derivative \( \frac{\partial \Phi(\beta, z)}{\partial z} \) as a function of \( z \). Then,

\[
\Phi(\beta, \Gamma) = \begin{cases} 
\Phi(\beta, 1), & p(\beta \Gamma) \leq s(\beta), \\
\Phi(\beta, g_{\beta}(p(\beta \Gamma))) + (1 - g_{\beta}(p(\beta \Gamma))) p(\beta \Gamma), & s(\beta) < p(\beta \Gamma) < t(\beta), \\
p(\beta \Gamma), & t(\beta) \leq p(\beta \Gamma),
\end{cases} \tag{1-18}
\]

with the paramagnetic pressure \( p(\beta \Gamma) = \ln 2 \cosh(\beta \Gamma) \).

**Corollary 1.5** implies that there are either one or two magnetic phase transitions, depending on \( s(\beta) \). If \( s(\beta) = \ln 2 \) or, equivalently, \( \tilde{a}(1) = 0 \), we find a single magnetic phase transition at the critical magnetization

\[
\Gamma^{(r)}_c(\beta) = \frac{1}{\beta} \arccosh\left( \frac{1}{2} e^{f(\beta)} \right).
\]
Figure 2. Both figures illustrate examples for the phase diagram of the quantum CREM as a function of the transversal magnetic field $\Gamma$ and the temperature $\beta^{-1}$. The first plot contains two magnetic phase transitions (bold lines) into transversal magnetic order. The second plot shows the case of one magnetic phase transition. The dotted line corresponds to the glass transition at $\beta_c = \sqrt{(2 \ln 2)/\bar{a}(0)}$. If $\bar{A}$ is continuously differentiable, the magnetic transitions are second order.

Otherwise, there is a second phase transition at

$$\Gamma^{(l)}_c(\beta) = \frac{1}{\beta} \text{arcosh} \left( \frac{1}{2} e^{s(\beta)} \right).$$

An explicit computation using (1-18) shows that the specific magnetization in the transversal direction

$$m_x(\beta, \Gamma) = \frac{1}{\beta} \frac{\partial}{\partial \Gamma} \Phi(\beta, \Gamma) = \begin{cases} 0, & p(\beta \Gamma) \leq s(\beta), \\ (1 - g_\beta(p(\beta \Gamma))) \tanh(\beta \Gamma), & s(\beta) < p(\beta \Gamma) < t(\beta), \\ \tanh(\beta \Gamma), & t(\beta) \leq p(\beta \Gamma), \end{cases}$$

is continuous. This transversal magnetic order does not vanish over the line $\Gamma^{(r)}_c(\beta)$ but rather only at $\Gamma^{(l)}_c(\beta)$ (which is absent in the case $\tilde{a}(1) = 0$). If the derivative of $\tilde{a}(x)$ exists at $x = 0$ or $x = 1$, the second derivative of $\Phi(\beta, \Gamma)$ has a jump at the respective critical magnetic fields and we have a second-order magnetic transition and not first order as in the quantum GREM; see Figure 2. In the classical model, the low-temperature glass phase is described by a random probability measure which captures the distribution of the spin overlaps \[8; 11\]. As with the GREM, it is an interesting question, which will be postponed to a future work, to study the influence of the transversal field on these quantities as well as on the nature of the eigenstates.

**1C. Quantum Parisi formula for the nonhierarchical GREM.** The nonhierarchical GREM was introduced in [7] to illustrate Parisi’s ultrametricity conjecture in an explicitly solvable model. We study the nonhierarchical GREM with a transverse field, since this is a basic test of whether our results in Section 2...
are only strictly valid for hierarchical models or if one might hope that they still hold true to a certain extent for more complicated models. Indeed, we are again able to explicitly determine the free energy. As in the GREM we write \( \sigma = \sigma_1 \cdots \sigma_n \) with \( \sigma_k \in \mathcal{Q}_N^{(k)} \) and \( L_k = x_k - x_{k-1} \) are the corresponding interval lengths. We denote by \( \mathcal{P} \) the power set of \( \{1, \ldots, n\} \). To each subset \( J = \{j_1, \ldots, j_m\} \in \mathcal{P} \) we assign the spin vector \( \sigma_J = \sigma_{j_1} \cdots \sigma_{j_m} \) and a nonnegative number \( a_J \geq 0 \) with \( a_\emptyset := 0 \). We further assume that the numbers \( a_J \) add up to 1, \( \sum_{J \in \mathcal{P}} a_J = 1 \). For each \( J \in \mathcal{P} \) we denote by \( X^J_{\sigma_J} \) independent standard Gaussian variables, which are also independent from each other for different superscripts \( J_1 \neq J_2 \). The total Gaussian process \( U \) of the nonhierarchical GREM on \( \mathcal{Q}_N \) is then given by

\[
U(\sigma) := \sqrt{N} \sum_{J \in \mathcal{P}} \sqrt{a_J} X^J_{\sigma_J}.
\]

The GREM is the special case where \( a_J \neq 0 \) only if \( J = \{1, \ldots, k\} \) for \( k \in \{1, \ldots, n\} \).

In contrast to the GREM the induced metric \( \mathbb{E}[|U(\sigma) - U(\sigma')|^2]^{1/2} \) does in general not satisfy ultrametricity. The hierarchical structure only emerges in the limit. For the formulation of the limiting free energy, we recall from [7] that a chain \( S = \{A_0, A_1, A_2, \ldots, A_n\} \subset \mathcal{P} \) consists of nested sets \( A_i \), i.e., \( \emptyset = A_0 \subset A_1 \subset A_2 \subset \cdots \subset A_n \), with cardinality \( |A_i| = i \). To each chain \( S = \{A_1, A_2, \ldots, A_n\} \) we assign a hierarchical GREM with weights

\[
a^S_k := \sum_{D \subset A_k, D \not\subset A_{k-1}} a_D,
\]

which satisfy \( \sum_{k=1}^n a^S_k = 1 \) for any chain \( S \) by construction, and endpoints

\[
y^S_k := \sum_{j \in A_k} L_j.
\]

The corresponding hierarchical GREM’s pressure converges and we denote the limit by \( \Phi(\beta, S) \). In [7], Bolthausen and Kistler showed that the limit of the pressure in the nonhierarchical GREM converges to a minimum of such GREMs,

\[
\lim_{N \to \infty} \frac{1}{N} \ln \text{Tr} e^{-\beta V} = \min_{S \in \mathcal{C}} \Phi(\beta, S),
\]

where the minimum is taken over the set \( \mathcal{C} \) of all chains.

After these preparations, we are able to consider the nonhierarchical GREM with a transverse magnetic field whose weights are independent copies of some random variable \( b \). We define the pressure as before,

\[
\Phi_N(\beta, b) := \frac{1}{N} \ln \text{Tr} e^{-\beta (U - B)}.
\]

The general theory developed in Section 2 also applies to the nonhierarchical GREM and yields:
Theorem 1.6. Let $\beta \geq 0$ and $b$ an absolutely integrable random variable. Then, the pressure $\Phi_N(\beta, b)$ converges almost surely and the limit is given by

$$\Phi(\beta, b) := \lim_{N \to \infty} \Phi_N(\beta, b) = \max_{D \in \mathcal{P}} \min_{S \in \mathcal{C}^D} \left[ \Phi_D(\beta, S) + \sum_{k \in D'} L_k \mathbb{E}[\ln 2 \cosh(\beta b)] \right].$$

(1-20)

Here, $\mathcal{C}^D$ denotes the set of chains which end at $D$, i.e., $S = \{A_0, A_1, \ldots, A_m\} \in \mathcal{C}^D$ if and only if

$$\emptyset = A_0 \subset A_1 \subset \cdots \subset A_m = D$$

and $|A_i| = i$ for any $1 \leq i \leq m$. Moreover, $\Phi_D(\beta, S)$ is the pressure of the corresponding GREM on the reduced hypercube associated to $D$.

The proof of this theorem is found in Section 3B.

The max-min structure of the limit in (1-20) seems to be quite universal as it also appears in the Parisi’s formula for vector spin glasses [32]. This formula was used in [2] to obtain an expression for the limit in the quantum SK-model. However, there the maximum is essentially taken over the infinite-dimensional path overlap, which makes it hard to analyze. One might hope to find a less involved parametrization of the overlap distribution which is easier to access.

In fact, (1-20) can be further simplified, since the limit does only depend on a single chain.

Corollary 1.7. There exists a chain $S \in \mathcal{C}$ such that for any $\beta \geq 0$ and any absolutely integrable variable $b$,

$$\Phi(\beta, b) = \Phi(\beta, b, S).$$

(1-21)

Here $\Phi(\beta, b, S)$ denotes the pressure of quantum GREM assigned to $S$; see (1-13).

Corollary 1.7, whose proof is also found in Section 3B, shows that the nonhierarchical GREM in a transversal field is at least on a thermodynamical level equivalent to an ordinary quantum GREM.

2. Hierarchical peeling

In this section, we present the general principle, which we dubbed hierarchical peeling, from which the main results presented in the previous section will follow. We first describe the core of this idea in the binary setup.

2A. Peeling principle. We start by describing the general setting. Picking a parameter $0 < x \leq 1$, we will decompose the hypercube $Q_N$ into two reduced hypercubes of spin arrays of length $[xN]$ and $N - [xN]$. Accordingly, we write

$$\sigma = (\sigma_1, \sigma_2) \in Q_N, \text{ where } \sigma_1 \in Q_N^{(1)} := Q_{[xN]} \text{ and } \sigma_2 \in Q_N^{(2)} := Q_{N-[xN]}.$$

We consider Hamiltonians $H = U - B$, where $U$ is a random potential on $Q_N$ and $B$ is a random transversal field, which satisfy several assumptions. We start with $U$:

Assumption 2.1 (assumptions on $U$). The random potential $U$ on $Q_N$ takes the form

$$U(\sigma) = V_N(\sigma) + X_{\sigma_1}$$

(2-1)
with some random potential $V_N$ which is independent of the random process $X_{\sigma_1}$. The random variables $X_{\sigma_1}$ with $\sigma_1 \in Q_N^{(1)}$ are absolutely integrable, centered, and satisfy:

1. $X_{\sigma_1}$ are independent and identically distributed for each fixed $N \in \mathbb{N}$.
2. The pushforward measures $\mu_N$ of the negative parts $X_{\sigma_1}^- / N$ satisfy a large deviation principle (LDP) with a lower semicontinuous rate function $I : \mathbb{R} \to [0, \infty]$, i.e., for any Borel set $A \subset \mathbb{R}$,

$$- \inf_{x \in \text{int}(A)} I(x) \leq \liminf_{N \to \infty} \frac{1}{N} \ln \mu_N(A) \leq \limsup_{N \to \infty} \frac{1}{N} \ln \mu_N(A) \leq - \inf_{x \in \text{clos}(A)} I(x).$$

Moreover, we assume

$$\inf_{x \in (-\infty, -\epsilon]} I(x) > 0$$

for any $\epsilon > 0$.

3. For any random weights $w_{\sigma_1}$ which are independent from $X_{\sigma_1}$ and further fulfill almost surely

$$w_{\sigma_1} \geq 0, \quad \sum_{\sigma_1 \in Q_N^{(1)}} w_{\sigma_1} = 1,$$

a generalized strong law holds true almost surely,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{\sigma_1 \in Q_N^{(1)}} w_{\sigma_1} X_{\sigma_1} = 0.$$

As will be discussed in the next subsection, we are mostly interested in hierarchical $V_N$ as in the GREM or CREM, but our results also apply to the more general situation. An important example where $V_N$ is not of CREM type is the case of a nonvanishing longitudinal magnetic field. The addition of a pure longitudinal field to a classical hierarchical glass technically remains in the realm of probability theory and has been studied in [4; 5; 9]. As is further discussed in [28], which deals with the application of the peeling principle for a study of the combined effects of transversal and longitudinal fields, the two choices of field direction not only differ in the mathematics involved, but also cause different physical behavior.

The LDP (2-2) with (2-3) ensure that probabilities of the type $\mathbb{P}(X_{\sigma_1} < -\epsilon N)$ decay exponentially in $N$ for any $\epsilon > 0$. The assumption (2-4) is a technical condition needed for our proof of Theorem 2.3. The following examples of random variables $X_{\sigma_1}$ meet Assumption 2.1, which can be seen by the sufficient criterion Lemma A.1 that we present in the Appendix:

1. $X_{\sigma_1} = \sqrt{N a} Y_{\sigma_1}$ with independent standard Gaussian $Y_{\sigma_1}$ and some $a > 0$. The rate function of the negative part is $I(x) = \frac{a}{2} x^2 \mathbb{1}_{x < 0}$.
2. Another example is $X_{\sigma_1} = -NY_{\sigma_1}$, where $Y_{\sigma_1}$ are independent and follow an exponential distribution with parameter $N$. The rate function of the negative part is $I(x) = |x| \mathbb{1}_{x < 0}$.
3. More generally, let $Y \leq 0$ be a random variable with a decay of the form $\exp(-Ct^\alpha)$ for some $\alpha, C > 0$, i.e.,

$$- \lim_{t \to \infty} t^{-\alpha} \ln \mathbb{P}(Y < t) = C.$$
Then, we define $X_{\sigma_1} = N^{1-1/\alpha} Y_{\sigma_1}$, where $Y_{\sigma_1}$ are independent copies of $Y$. The corresponding rate function is given by $I(x) = C|\sigma|^{\alpha} \mathbb{1}_{x<0}$.

We consider a not necessarily constant transversal magnetic field $B = \sum_{j=1}^{N} b_j s_j$ as in (1-10) with random variables $(b_j)$ which do not need to be independent from each other. The transversal field $B$ splits into two parts $B^{1,x}$ and $B^{2,x}$, which act exclusively on the respective part of the array,

$$B^{1,x} := \sum_{i=1}^{\lfloor xN \rfloor} b_i s_i, \quad B^{2,x} := \sum_{i=\lfloor xN \rfloor + 1}^{N} b_i s_i.$$ 

If $x = 1$, we simply set $B^{2,1} = 0$. Subsequently, we assume the following on the transversal field $B$:

**Assumption 2.2** (assumptions on $B$). The random weights $(b_j)$ are independent of the potential $U$ and satisfy almost surely

$$\limsup_{N \to \infty} N^{-1} \sqrt{\sum_{j=1}^{N} |b_j|^2} = 0. \quad (2-5)$$

Let us discuss some sufficient conditions on $(b_j)$ which ensure the validity of **Assumption 2.2**:

1. **Assumption 2.2** obviously covers the constant field case $b_j = \Gamma \geq 0$.
2. If the weights are almost surely dominated by $\sqrt{N}$, that is,

$$\limsup_{N \to \infty} N^{-\frac{1}{2}} \max_{1 \leq j \leq N} |b_j| = 0, \quad (2-6)$$

then (2-5) holds true.
3. In view of the framework in Section 1, we are mostly interested in weights $(b_j)$ forming independent copies of an absolute integrable random variable $b$. Then, (2-5) is satisfied and this result is presented as Lemma A.2 in the Appendix. If we additionally assume that $E[|b|^r]$ is finite for some $r > 1$, **Assumption 2.2** is easily verified. Namely if $r \in (1, 2]$, then

$$N^{-1} \sqrt{\sum_{i=1}^{N} |b_i|^2} \leq N^{-(1-1/r)} \left( N^{-1} \sum_{i=1}^{N} |b_i|^r \right)^{1/r}.$$ 

The term in the bracket converges almost surely to a constant by the strong law of large numbers. So (2-5) is fulfilled.

If Assumptions 2.1 and 2.2 hold true, our main results state that the pressure

$$\Phi_N(\beta) := \frac{1}{N} \ln \text{Tr} [e^{-\beta(U-B)}]$$

asymptotically agrees with the maximum of the pressures of partially quantum or classical type

$$\Phi_{N}^{(\text{qm})}(\beta) := \frac{1}{N} \ln \text{Tr} [e^{-\beta(V_N-B)}] \quad \text{and} \quad \Phi_{N}^{(\text{cl})}(\beta) := \frac{1}{N} \ln \text{Tr} [e^{-\beta(U_{N}^{2,x})}]$$

even if $\Phi_{N}^{(\text{qm})}(\beta)$ and $\Phi_{N}^{(\text{cl})}(\beta)$ do not converge:
Theorem 2.3. Under Assumptions 2.1 and 2.2, for any \( x \in (0, 1) \) we have the almost sure convergence
\[
\limsup_{N \to \infty} |\Phi_N(\beta) - \max\{\Phi_N^{(qm)}(\beta), \Phi_N^{(cl)}(\beta)\}| = 0. \tag{2-7}
\]

Roughly speaking, the Gaussian variables \( (X_{\sigma}) \) and the partial magnetic term \( B^{1,x} \) only contribute separately from each other to the free energy. This result may be regarded as a generalization of the limit theorem for the QREM in [27]. If the almost-sure limits

\[
\Phi^{(qm)}(\beta) := \lim_{N \to \infty} \Phi_N^{(qm)}(\beta) \quad \text{and} \quad \Phi^{(cl)}(\beta) := \lim_{N \to \infty} \Phi_N^{(cl)}(\beta)
\]

exist for any \( \beta \geq 0 \), we immediately obtain
\[
\lim_{N \to \infty} \Phi_N(\beta) = \max\{\Phi^{(qm)}(\beta), \Phi^{(cl)}(\beta)\}. \tag{2-8}
\]

For a proof of Theorem 2.3 the methods in [27] are robust enough to be extended. We briefly recall some notation and results necessary for doing so. For \( \epsilon > 0 \) we denote the large deviation set of \( X_{\sigma_1} \) by
\[
\mathcal{L}_\epsilon := \{\sigma_1 \in Q_N^{(1)} \mid X_{\sigma_1} \leq -\epsilon N\}. \tag{2-9}
\]

The operator \( B_{\mathcal{L}_\epsilon}^{1,x} \) is the Dirichlet restriction of \( B^{1,x} \) to the complement \( \mathcal{L}_\epsilon^c \), that is, \( B_{\mathcal{L}_\epsilon}^{1,x} = P_{\mathcal{L}_\epsilon} B^{1,x} P_{\mathcal{L}_\epsilon} \) with the natural orthogonal projection \( P_{\mathcal{L}_\epsilon} \) induced by the complement of the set \( \mathcal{L}_\epsilon \) on the first component \( \mathcal{H} := \ell^2(Q_N^{(1)}) \) of the tensor-product Hilbert space \( \ell^2(Q_N) = \mathcal{H}^1 \otimes \mathcal{H}^2 \) with \( \mathcal{H}^2 := \ell^2(Q_N^{(2)}) \). We further introduce
\[
A_{\mathcal{L}_\epsilon} := A_{\mathcal{L}_\epsilon}^1 \otimes \mathbb{1}_{\mathcal{H}^2} := B^{1,x} - B_{\mathcal{L}_\epsilon}^{1,x},
\]
where \( A_{\mathcal{L}_\epsilon}^1 \) only acts on \( \mathcal{H}^1 \) and \( A_{\mathcal{L}_\epsilon} \) is its trivial extension to the full Hilbert space. More precisely, the matrix elements of \( A_{\mathcal{L}_\epsilon}^1 \) are given by
\[
\langle \sigma_1 | A_{\mathcal{L}_\epsilon}^1 | \sigma'_1 \rangle = \mathbb{1}_{[\sigma_1, \sigma'_1] \cap \mathcal{L}_\epsilon \neq \emptyset} \sum_{j=1}^{[xN]} b_j \delta_{s_j \sigma_1, \sigma'_1}, \tag{2-10}
\]
where in a slight abuse of notation we extend the operators \( s_j \) to configurations by setting \( s_j \sigma := (\sigma_1, \sigma_2, \ldots, -\sigma_j, \ldots, \sigma_N) \). Moreover, \( \delta_{. .} \) denotes Kronecker’s delta, so that the sum in the right side of (2-10) reduces to one term only.

We will need the following generalization of Lemma 2 and Lemma 3 in [27]:

Proposition 2.4. Under Assumptions 2.1 and 2.2, for any \( \epsilon > 0 \) and \( x \in (0, 1] \) the operator norm of \( A_{\mathcal{L}_\epsilon} \) satisfies almost surely:
\[
\limsup_{N \to \infty} N^{-1} \|A_{\mathcal{L}_\epsilon}\| = 0. \tag{2-11}
\]

The proof of Proposition 2.4 is based on an estimate for the maximal size of the so-called gap-connected components of \( \mathcal{L}_\epsilon \), which are defined as follows:

Definition 2.5. Let \( \mathcal{Q}_N^{(1)} \) be the supergraph of the Hamming cube \( Q_N^{(1)} \), which one obtains by adding all edges \( \{\sigma_1, \sigma'_1\} \), where \( \sigma_1, \sigma'_1 \) are any two vertices at distance \( d(\sigma_1, \sigma'_1) = 2 \). We call \( \mathcal{L}_\epsilon \subset \mathcal{L}_\epsilon \) a...
gap-connected component if $C_\varepsilon$ is connected as a subset of $\tilde{Q}_N^{(1)}$. A gap-connected component $C_\varepsilon$ is maximal if there is no other vertex $\sigma_1 \in L_\varepsilon \setminus C_\varepsilon$ such that $C_\varepsilon \cup \{\sigma_1\}$ forms a gap-connected component. We denote by $(C_\varepsilon^\alpha)_\alpha$ the collection of maximal gap-connected components of $L_\varepsilon$.

We claim that the maximum of the cardinality, $\max_\alpha |C_\varepsilon^\alpha|$, is almost surely of order one:

**Lemma 2.6.** Under Assumption 2.1, for any $\varepsilon > 0$ and $x \in (0, 1]$ there is $K > 0$ such that

$$\limsup_{N \to \infty} \max_\alpha |C_\varepsilon^\alpha| \leq K$$

holds almost surely.

**Proof.** We follow the lines of the proof of Lemma 2 in [27]. We fix $K \in \mathbb{N}$ and introduce the event

$$\Omega_{\varepsilon,K,N} := \bigcap_{\sigma_1 \in Q_N^{(1)}} \{ |B_{4K}(\sigma_1) \cap L_\varepsilon| < K \}.$$ 

We note that for $\omega \in \Omega_{\varepsilon,K,N}$ we always have $\max_\sigma |C_\varepsilon^\sigma| < K$, as any gap-connected component with $K$ vertices is contained in some ball $B_{4K}(\sigma_1) \subset Q_N^{(1)}$ of radius $4K$, which is centered at some $\sigma_1$. We estimate the probability of the complement $\Omega_{\varepsilon,K,N}^c$ using the union bound:

$$\mathbb{P}(\Omega_{\varepsilon,K,N}^c) \leq \sum_{\sigma_1 \in Q_N^{(1)}} \mathbb{P}( |B_{4K}(\sigma_1) \cap L_\varepsilon| \geq K ) \leq 2^{[xN]} \left( \frac{|B_{4K}|}{K} \right) \mathbb{P}( X_{\sigma_1} < -\varepsilon N ) K.$$

The second inequality follows from independence of the random variables $X_{\sigma_1}$ and an estimate on the number of subsets of a given size in terms of the binomial coefficients. The rate function $I$ of $X_{\sigma_1}/N$ satisfies $\inf_{-\infty < z \leq -\varepsilon} I(z) = \delta_\varepsilon > 0$, from which we conclude

$$\mathbb{P}(\Omega_{\varepsilon,K,N}^c) \leq 2^{[xN]} \frac{|B_{4K}|}{K!} e^{-K N (\delta_\varepsilon + o(1))}.$$ 

Since $|B_{4K}| \leq e N^{4K}$, we may choose $K = K(\varepsilon)$ large enough such that this probability decays exponentially fast. A Borel–Cantelli argument then yields the almost-sure bound

$$\limsup_{N \to \infty} \max_\alpha |C_\varepsilon^\alpha| \leq K.$$

**Proposition 2.4** is now a simple consequence of Assumption 2.2 and Lemma 2.6:

**Proof of Proposition 2.4.** The operator $A_{L_\varepsilon}$ exhibits a natural decomposition as the direct sum

$$A_{L_\varepsilon} = \bigoplus_\alpha A_{C_\varepsilon^{\alpha}}^1 \otimes 1_{H^2},$$

where $A_{C_\varepsilon^{\alpha}}^1$ denotes the restriction of $A_{L_\varepsilon}^1$ to the subspace of vertices with nonempty intersection with the maximal gap-connected component $C_\varepsilon^{\alpha}$. Estimating the operator norm on every block in terms of the
Frobenius norm using (2-10), we conclude

\[ \|A_{c_2}^1\|^2 \leq \max_{\sigma_1, \sigma_1'} \sum_{\sigma_1, \sigma_1' \in Q_N^{(1)}} |\langle \sigma_1 | A_{c_2}^1 | \sigma_1' \rangle|^2 \leq \max_{\sigma_1, \sigma_1' \in Q_N^{(1)}} \sum_{j=1}^{[xN]} b_j \delta_{\sigma_1, \sigma_1'} \leq 2 \max_{\alpha} |C_\alpha^x| \sum_{j=1}^{N} |b_j|^2, \]

which together with Assumption 2.2 and Lemma 2.6 completes the proof. □

We finally spell out the proof of Theorem 2.3:

Proof of Theorem 2.3. We separately establish an asymptotically sharp lower and upper bound.

Lower bound: The lower bound rests on a twofold application of Gibbs’ variational principle [34; 35]. First, let \( \rho_\beta^{(qm)} \) be the Gibbs state of \( H^{(qm)} = V_N - B \). An application of the Gibbs variational principle with \( \rho = \rho_\beta^{(qm)} \) and \( H = H^{(qm)} + X_{\sigma_1} \) yields

\[ \Phi_N(\beta) = N^{-1} \sup_{\rho}[ -\beta \text{Tr}(H\rho) - \text{Tr}(\rho \ln \rho) ] \geq \Phi_N^{(qm)}(\beta) - \beta N^{-1} \sum_{\sigma_1} X_{\sigma_1} w_{\sigma_1}. \]

The weights \( w_{\sigma_1} := \sum_{\sigma_2} \langle \sigma_1 \sigma_2 | \rho_\beta^{(qm)} | \sigma_1 \sigma_2 \rangle \) are nonnegative, add up to 1, and are independent of \( X_{\sigma_1} \). By Assumption 2.1 we conclude that almost surely

\[ \lim_{N \to \infty} \inf (\Phi_N(\beta) - \Phi_N^{(qm)}(\beta)) \geq 0. \]

Next, the eigenstates \( |\psi\rangle \in \ell^2(Q_N) \) of \( H^{(cl)} = U - B_{1,x} \) take the form of tensor products \( |\psi\rangle = |\sigma_1\rangle \otimes |\phi\rangle \) with a certain \( |\phi\rangle \in \mathcal{H}_2 \). As the matrix elements \( \langle \psi | B_{1,x} | \psi \rangle \) vanish for these eigenstates, the Gibbs state \( \rho_\beta^{(cl)} := e^{-\beta H^{(cl)}} / \text{Tr} e^{-\beta H^{(cl)}} \) satisfies

\[ \text{Tr} B_{1,x} \rho_\beta^{(cl)} = 0. \quad (2-13) \]

The Gibbs variational principle with \( \rho = \rho_\beta^{(cl)} \) and \( H = H^{(cl)} - B_{1,x} \) then yields

\[ \Phi_N(\beta) \geq \Phi_N^{(cl)}(\beta). \quad (2-14) \]

Combining both lower bounds, we obtain almost surely

\[ \lim_{N \to \infty} \inf (\Phi_N(\beta) - \max\{ \Phi_N^{(qm)}(\beta), \Phi_N^{(cl)}(\beta) \}) \geq 0. \]

Upper bound: Let \( \varepsilon > 0 \) be arbitrary and consider the direct-sum decomposition of the Hilbert space \( \ell^2(Q_N) = (\ell^2(L_c) \otimes \mathcal{H}_2) \oplus (\ell^2(L^*_c) \otimes \mathcal{H}_2) \). The only term in \( H \) connecting the two subspaces is \( A_{L_c} \). The Golden–Thompson inequality (see [35, Corollary I.4.13]) together with the positivity of \( e^H \) for \( H = H^* \) and trivial norm estimates thus yield

\[ \text{Tr} e^{-\beta(U - B)} \leq \text{Tr} e^{-\beta A_{L_c}} e^{-\beta(U - B_{1,x} - B_{2,x})} \leq e^{\beta \| A_{L_c} \|} \text{Tr} e^{-\beta(U - B_{1,x} - B_{2,x})} \]

\[ \leq e^{\beta \| A_{L_c} \|} (\text{Tr} |\ell^2(L_c) \otimes \mathcal{H}_2 e^{-\beta(U - B_{2,x})} + e^{\beta \varepsilon N} \text{Tr} |\ell^2(L^*_c) \otimes \mathcal{H}_2 e^{-\beta(V_N - B_{1,x} - B_{2,x})} ). \]
In the last term we additionally used the fact that $X_{\sigma_1} \geq -\varepsilon N$ for all $\sigma_1 \in \mathcal{L}_c$. The first term is bounded by

$$\text{Tr}_{\{\ell^2(\mathcal{L}_c) \otimes \mathcal{H}\}} e^{-\beta(U-B^{2,x})} \leq \text{Tr} e^{-\beta(U-B^{2,x})}.$$ 

The second term is estimated using the nonnegativity of the diagonal matrix elements of the semigroups generated by $B$ and the Golden–Thompson inequality again:

$$\text{Tr}_{\{\ell^2(\mathcal{L}_c) \otimes \mathcal{H}\}} e^{-\beta(V_N-B^{1,x}-B^{2,x})} \leq \text{Tr} e^{-\beta(V_N-B^{1,x}-B^{2,x})} \leq e^{\beta \|A_{\mathcal{L}_c}\|} \text{Tr} e^{-\beta(V_N-B)}$$

Since $\varepsilon > 0$ was arbitrary and $\|A_{\mathcal{L}_c}\| = o(N)$ by Proposition 2.4, we conclude the almost-sure inequality

$$\limsup_{N \to \infty} \left( \Phi_N(\beta) - \max \{ \Phi_N^{(qm)}(\beta), \Phi_N^{(c)}(\beta) \} \right) \leq 0. \tag{2-15}$$

2B. Application to QGREM and QCREM. Since we are free in the choice of $V_N$ in Theorem 2.3, we obtain the following corollary for GREM type potentials:

**Corollary 2.7.** Let $X = \sqrt{a_1} X_{\sigma_1} + \sqrt{a_2} X_{\sigma_1\sigma_2} + \cdots + \sqrt{a_n} X_{\sigma_1\cdots\sigma_n}$ be a Gaussian vector as in (1-4). Then, we have the almost sure convergence

$$\limsup_{N \to \infty} \frac{1}{N} \ln \text{Tr} e^{-\beta(\sqrt{N}X-B)} = \max_{0 \leq k \leq n} \frac{1}{N} \ln \text{Tr} e^{-\beta(\sqrt{N}(\sqrt{a_1} X_{\sigma_1} + \sqrt{a_2} X_{\sigma_1\sigma_2} + \cdots + \sqrt{a_{n-1}} X_{\sigma_1\cdots\sigma_{n-1}}) - B^{k,n})} = 0. \tag{2-15}$$

**Proof:** We apply Theorem 2.3 iteratively backwards. We start with peeling off $\sqrt{a_n} X_{\sigma_1\cdots\sigma_n}$ (which takes the role of $X_{\sigma_1}$ in Theorem 2.3) setting $V_N = V_N^{(n)} := \sqrt{N}a_n^{-1} X_{\sigma_1} + \sqrt{N}a_2^{-1} X_{\sigma_1\sigma_2} + \cdots + \sqrt{N}a_{n-1}^{-1} X_{\sigma_1\cdots\sigma_{n-1}}$. The peeling principle (2-7) yields

$$\limsup_{N \to \infty} \frac{1}{N} \ln \text{Tr} e^{-\beta(\sqrt{N}X-B)} = \max \{ \ln \text{Tr} e^{-\beta(\sqrt{N}a_n^{-1} X_{\sigma_1\cdots\sigma_n} + V_N^{(n)})}, \ln \text{Tr} e^{-\beta(V_N^{(n)}-B)} \} = 0.$$

As a next step, we write $V_N^{(n)} =: V_N^{(n-1)} + \sqrt{N}a_{n-1}^{-1} X_{\sigma_1\cdots\sigma_{n-1}}$ and again apply Theorem 2.3 to the second expression in the maximum. Thus,

$$\limsup_{N \to \infty} \frac{1}{N} \ln \text{Tr} e^{-\beta(V_N^{(n)}-B)} = \max \{ \ln \text{Tr} e^{-\beta(V_N^{(n)}-B^{2,x}_{n-1})}, \ln \text{Tr} e^{-\beta(V_N^{(n-1)}-B)} \} = 0.$$

Proceeding like this, we arrive after $n$ steps at (2-15).

Theorem 1.2 and Corollary 2.7 look alike. However, in Theorem 1.2 we further evaluate the trace and claim that the maximum in (2-15) is attained at some endpoint $\gamma_l$ of the concave hull $\tilde{A}$. We postpone this remaining part of the proof of Theorem 1.2 to Section 3A.

Now, instead, we will extend Corollary 2.7 to CREM type potentials. To this end, we introduce a useful shorthand notation. If $X$ is a centered Gaussian vector with hierarchical distribution function $A$, we define for $0 \leq \varepsilon \leq 1$ the centered Gaussian vector $X^{(\varepsilon)}$ with hierarchical distribution function given by

$$A^{(\varepsilon)}(x) := \begin{cases} A(x) & \text{if } x \leq \varepsilon, \\ A(\varepsilon) & \text{else}. \end{cases}$$

We are now ready to formulate the following theorem:
We observe, however, that an interpolation argument only controls the expectation value with respect to $H$ where

$$V$$ variable due to the randomness of $X$ as well. For $X, Y$ be two independent centered Gaussian processes on $Q_N$, which are independent of $V_N$ as well. For $t \in [0, 1]$ we set the interpolated pressure $\Psi$,

$$\Psi(t) = \frac{1}{N} \ln \left[ \frac{1}{\sqrt{1-t}} \right] e^{-\beta(\sqrt{1-t}NY + V_N - B)}$$

where by Lemma 1.1 we may assume without loss of generality that $b_j \geq 0$ for all $j$. By standard Gaussian interpolation (see, e.g., [39, Lemma 1.3.1]), we obtain

$$E_{X,Y}[\Psi(1) - \Psi(0)] = \frac{1}{2} \sum_{\sigma, \sigma'} \int_0^1 \left[ E[Y(\sigma)Y(\sigma')] - E[X(\sigma)X(\sigma')] \right] \left[ \frac{\partial^2 \Psi(t)}{\partial X_\sigma \partial X_{\sigma'}} + \frac{\partial^2 \Psi(t)}{\partial Y_\sigma \partial Y_{\sigma'}} \right] dt,$$

where $E_{X,Y}$ denotes the expectation with respect to $X$ and $Y$. In general, $E_{X,Y}[\Psi(t)]$ is still a random variable due to the randomness of $V_N$ and $B$. The second partial derivatives of $\Psi(t)$ differ from their classical expression but can still be computed using the Duhamel formula:

$$\frac{\partial^2 \Psi(t)}{\partial X_\sigma \partial X_{\sigma'}} + \frac{\partial^2 \Psi(t)}{\partial Y_\sigma \partial Y_{\sigma'}} = -\beta^2 \frac{\langle \sigma \rvert e^{H_t} \rvert \sigma' \rangle \langle \sigma' \rvert e^{H_t} \rvert \sigma \rangle}{(\text{Tr} \ e^{H_t})^2} + \beta^2 \int_0^1 \frac{\langle \sigma \rvert e^{sH_t} \rvert \sigma' \rangle \langle \sigma' \rvert e^{(1-s)H_t} \rvert \sigma \rangle}{\text{Tr} \ e^{H_t}} ds$$

with the abbreviation $H_t := -\beta(\sqrt{tNY + (1-t)NY + V_N - B})$. Since we assumed without loss of generality that $b_j \geq 0$, the matrix elements $\langle \sigma \rvert e^{H_t} \rvert \sigma' \rangle$ are nonnegative for any $\sigma, \sigma'$. Moreover, we know that

$$\sum_{\sigma, \sigma'} \frac{\langle \sigma \rvert e^{H_t} \rvert \sigma' \rangle \langle \sigma' \rvert e^{H_t} \rvert \sigma \rangle}{(\text{Tr} \ e^{H_t})^2} = 1 = \sum_{\sigma, \sigma'} \int_0^1 \frac{\langle \sigma \rvert e^{sH_t} \rvert \sigma' \rangle \langle \sigma' \rvert e^{(1-s)H_t} \rvert \sigma \rangle}{\text{Tr} \ e^{H_t}} ds.$$

Consequently, we arrive at the bound

$$|E_{X,Y}[\Psi(1) - \Psi(0)]| \leq \beta^2 \max_{\sigma, \sigma'} |E[X(\sigma)X(\sigma')] - E[Y(\sigma)Y(\sigma')]|.$$

In case $X$ and $Y$ are of CREM-type with distribution functions $A_X$ and $A_Y$, respectively, we thus conclude

$$|E_{X,Y}[\Psi(1) - \Psi(0)]| \leq \beta^2 \| A_X - A_Y \|_\infty. \quad (2-17)$$

Analogously, we get

$$\frac{1}{N} \left| E_{X,Y} \left[ \ln \text{Tr} e^{-\beta(\sqrt{1-t}NY + V_N - B)} - \ln \text{Tr} e^{-\beta(\sqrt{1-t}NY + V_N - B)} \right] \right| \leq \beta^2 \| A_X - A_Y \|_\infty \quad (2-18)$$

for any $z \in [0, 1]$. The bounds (2-17) and (2-18) are our first main ingredients for the proof of Theorem 2.8. We observe, however, that an interpolation argument only controls the expectation value with respect to
the Gaussian variables. The following Gaussian concentration inequality is a convenient method to lift the convergence of expectation values to almost sure statements and vice versa.

**Proposition 2.9.** Let \(X\) be a Gaussian vector of CREM-type, \(V_N\) a random potential, and \(B\) a random transversal field, all independent from each other. The corresponding pressure

\[\Phi_N(\beta) = \frac{1}{N} \ln \text{Tr} e^{-\beta(\sqrt{N}X + V_N - B)}\]

exhibits a Gaussian concentration estimate with respect to probability measure of \(X\), i.e., for any \(t > 0\) and \(N \in \mathbb{N}\),

\[
P_X \left( |\Phi_N(\beta) - \mathbb{E}_X[\Phi_N(\beta)]| > \frac{t\beta}{\sqrt{N}} \right) \leq 2 \exp \left( -\frac{t^2}{4} \right). \tag{2-19} \]

The same bounds hold true for \(\Phi_N^{(z)}(\beta) = \frac{1}{N} \ln \text{Tr} e^{-\beta(\sqrt{N}X^{(z)} + V_N - B^{(z)})}\).

**Proof.** Since the lexicographic overlap (1-2) can only take values \(k/N\) with \(k = 0, 1, \ldots, N\) for every fixed \(N \in \mathbb{N}\), the CREM-type Gaussian vector \(X\) may be represented as a GREM-type distribution:

\[X(\sigma) = \sqrt{a}_1 X_{\sigma_1} + \sqrt{a}_2 X_{\sigma_1\sigma_2} + \cdots + \sqrt{a}_n X_{\sigma_1\sigma_2\cdots\sigma_n}\]

with independent standard Gaussian variables \(X_{\sigma_1}, \ldots, X_{\sigma_1\sigma_2\cdots\sigma_n}\) and some \(n = n(N)\). We calculate the free energy’s variation with respect to the i.i.d Gaussian variables \(X_{\sigma_1,\ldots,\sigma_k}\),

\[
-\frac{\partial \Phi_N(\beta)}{\partial X_{\sigma_1,\ldots,\sigma_k}} = \frac{\beta \sqrt{a}_k}{\sqrt{N} \text{Tr} e^{-\beta(X + V_N - B)}} \sum_{\hat{\sigma}_k} \langle \sigma_1 \cdots \sigma_k \sigma \rangle \langle e^{-\beta(X + V_N - B)} | \sigma_1 \cdots \sigma_k \hat{\sigma}_k \rangle.
\]

Here, \(\hat{\sigma}_k\) is an abbreviation for the remaining entries of the element \(\sigma \in Q_N\). Consequently, the square of the pressure’s Lipschitz constant is bounded by

\[
\sum_k \sum_{\sigma_1,\ldots,\sigma_k} \left( \frac{\partial \Phi_N(\beta)}{\partial X_{\sigma_1,\ldots,\sigma_k}} \right)^2 \leq \frac{\beta^2}{N},
\]

where we used that the weights \(a_k\) add up to 1. If we condition on \(V_N\) and \(B\), the Gaussian concentration inequality for Lipschitz functions (see [39, Theorem 1.3.4]) yields

\[
P_X \left( |\Phi_N(\beta, B) - \mathbb{E}_X[\Phi_N(\beta)]| > \frac{t\beta}{\sqrt{N}} \right) \leq 2 \exp \left( -\frac{t^2}{4} \right).
\]

A similar argument using the fact that the sum of the weights \(a_k^{(z)}\) add up to at most 1, shows that we have the same concentration inequality for \(\Phi_N^{(z)}(\beta)\). \(\square\)

Let us remark that a Gaussian concentration estimate still holds true if the weights \((a_k)\) do not add up to 1. Only the multiplicative constant in front of the exponential term changes. We move on to the proof of **Theorem 2.8:**

**Proof of Theorem 2.8.** We pick some \(\varepsilon > 0\) and an independent Gaussian vector \(Y\) of GREM-type with distribution (step-)function \(\tilde{A}\) such that

\[\|A - \tilde{A}\|_\infty \leq \varepsilon.\]
This is possible since $A$ is an increasing, right-continuous function and therefore a uniform limit of increasing step functions. We denote by $0 = x_0 < x_1 < x_2 < \cdots x_n = 1$ the points supporting $\tilde{A}$.

We now exploit the estimates in (2-17), (2-18) and Proposition 2.9 in order to obtain the almost sure bounds

$$\limsup_{N \to \infty} \frac{1}{N} \left| \ln \text{Tr} e^{-\beta(\sqrt{N}X + V_N - B)} - \ln \text{Tr} e^{-\beta(\sqrt{N}Y + V_N - B)} \right| \leq \beta^2 \varepsilon$$

and, since we also have $\|A^{(z)} - \tilde{A}^{(z)}\|_\infty \leq \varepsilon$,

$$\limsup_{N \to \infty} \sup_{z \in [0, 1]} \frac{1}{N} \left| \ln \text{Tr} e^{-\beta(\sqrt{N}X^{(z)} + V_N - B^{2,z})} - \ln \text{Tr} e^{-\beta(\sqrt{N}Y^{(z)} + V_N - B^{2,z})} \right| \leq \beta^2 \varepsilon.$$

The expressions depending on $Y$ do not necessarily converge. Nevertheless, we have almost surely

$$\limsup_{N \to \infty} \frac{1}{N} \left| \ln \text{Tr} e^{-\beta(\sqrt{N}Y + V_N - B)} - \sup_{0 \leq z \leq 1} \ln \text{Tr} e^{-\beta(\sqrt{N}Y^{(z)} + V_N - B^{2,z})} \right| = \limsup_{N \to \infty} \frac{1}{N} \left| \ln \text{Tr} e^{-\beta(\sqrt{N}Y + V_N - B)} - \max_{k=0,1,\ldots,n} \ln \text{Tr} e^{-\beta(\sqrt{N}Y^{(z_k)} + V_N - B^{2,z_k})} \right| = 0.$$

For the first equality we recall that for any $x_k \leq z < x_{k+1}$ the processes agree, i.e., $Y^{(z)} = Y^{(x_k)}$. Consequently, the Gibbs’ variational principle with $H = H' - (B^{2,x_k} - B^{2,z})$ and $H' = \sqrt{N}Y^{(x_k)} + V_N - B^{2,z}$ and an argument similar to (2-13)–(2-14) show that the maximum is attained at some $x_k$. The second equality follows from Corollary 2.7. Combining all these estimates, we arrive at

$$\limsup_{N \to \infty} \frac{1}{N} \left| \ln \text{Tr} e^{-\beta(\sqrt{N}X + V_N - B)} - \sup_{0 \leq z \leq 1} \ln \text{Tr} e^{-\beta(\sqrt{N}X^{(z)} + V_N - B^{2,z})} \right| \leq 2\beta^2 \varepsilon.$$

As $\varepsilon > 0$ is arbitrary, the proof of (2-16) is completed. \hfill \Box

3. Proofs of the main results

3A. The Quantum GREM and CREM. We first aim to prove Theorem 1.2, i.e.,

$$\lim_{N \to \infty} \frac{1}{N} \ln \text{Tr} e^{-\beta(\sqrt{N}X - B)} = \max_{0 \leq l \leq m} \left[ \sum_{j=1}^{l} \varphi^{(j)}(\beta) + (1 - y_l) \mathbb{E} [\ln 2 \cosh(\beta b)] \right]$$

for a GREM type variable $X$ and transversal field $B$ consisting of independent weights $(b_j)$ with the same distribution as $b$. We recall that $x_1, \ldots, x_n$ denote the jump points of the function $A$, the points $y_1, \ldots, y_m$, over which the above maximum is taken, are the endpoints of the concave hull’s $\tilde{A}$ linear segments and $\varphi^{(j)}(\beta)$ are the partial free energies from (1-6). For the remainder of this subsection and since we are interested in the limit $N \to \infty$, we also assume without loss of generality that $x_k N \in \mathbb{N}$ for all $k \in \{0, \ldots, n\}$. 

GENERALIZED RANDOM ENERGY MODELS IN A TRANSVERSAL MAGNETIC FIELD 233
Our starting point is Corollary 2.7 which for any GREM-type vector $X$ yields

$$\frac{1}{N} \ln \text{Tr} e^{-\beta(\sqrt{N}X + B)} = \max_{0 \leq k \leq n} \frac{1}{N} \ln \text{Tr} e^{-\beta(\sqrt{N}X^{(s)}_k - B^{2,x_k})} + o(1)$$

$$= \max_{0 \leq k \leq n} \left[ \frac{1}{N} \ln \text{Tr} \left| Q_{x_N} e^{-\beta\sqrt{N}X^{(s)}_k} \right| + \frac{1}{N} \sum_{i=|x_N x_k|}^{N} \ln 2 \cosh(\beta b_i) \right] + o(1),$$

since

$$X^{(s)}_k = \sqrt{a_1}X_1 + \sqrt{a_2}X_2\sigma_2 + \cdots + \sqrt{a_k}X_2\sigma_k$$

only acts nontrivially on the configurations $\sigma_1 \cdots \sigma_k$ and $B^{2,x_k}$ on $\sigma_{k+1} \cdots \sigma_n$ so that the total trace is simply the product of the partition functions of $X^{(s)}_k$ and $B^{2,x_k}$ on the corresponding reduced Hilbert space; see (1-14). The limit $N \to \infty$ of the bracket on the right side exists for any $k \in \{0, \ldots, n\}$. More precisely, the strong law of large numbers implies that the second term almost surely tends to $(1 - x_k)E[\ln 2 \cosh(\beta b)]$. Moreover, the first term converges since $X^{(s)}_k$ is still a GREM-type Gaussian vector on $Q_{x_N}$. The only difference is that the weights $a_1, \ldots, a_k$ do not add up to 1. This minor obstacle can be easily done away with by rescaling the inverse temperature $\beta$. In particular, if $x_k$ coincides with an endpoint $y_l$ of the concave hull’s segments,

$$\lim_{N \to \infty} \frac{1}{N} \ln \text{Tr} \left| Q_{x_N} e^{-\beta\sqrt{N}X^{(s)}_k} \right| = \sum_{j=1}^{l} \varphi^{(j)}(\beta),$$

where the partial free energies $\varphi^{(j)}(\beta)$ remain unchanged, i.e., they are still given by (1-6). This follows from the observation that $X$ and $X^{(y)}$ have the same concave hull up to the point $y_l$.

Since the limit $N \to \infty$ exists for each $k$, we may exchange the limit with the maximum. In order to prove Theorem 1.2, it therefore suffices to check that in

$$\lim_{N \to \infty} \frac{1}{N} \ln \text{Tr} e^{-\beta(\sqrt{N}X + B)} = \max_{k=0, \ldots, n} \left[ \lim_{N \to \infty} \frac{1}{N} \ln \text{Tr} \left| Q_{x_N} e^{-\beta\sqrt{N}X^{(s)}_k} \right| + (1 - x_k)E[\ln 2 \cosh(\beta b)] \right],$$

the maximum of the limit is always attained at some $y_l$. This is the content of the next lemma:

**Lemma 3.1.** If $X$ is a Gaussian vector of GREM type, we have

$$\max_{k=0, \ldots, n} \left[ \lim_{N \to \infty} \frac{1}{N} \ln \text{Tr} \left| Q_{x_N} e^{-\beta\sqrt{N}X^{(s)}_k} \right| + (1 - x_k)E[\ln 2 \cosh(\beta b)] \right]$$

$$= \max_{l=0, \ldots, m} \left[ \lim_{N \to \infty} \frac{1}{N} \ln \text{Tr} \left| Q_{y_N} e^{-\beta\sqrt{N}X^{(y)}_l} \right| + (1 - y_l)E[\ln 2 \cosh(\beta b)] \right].$$

**Proof.** If $\{x_0, \ldots, x_n\} = \{y_0, \ldots, y_m\}$, the statement is trivial. We thus consider one of the terms on the left side of (3-2) corresponding to $x_k$ with $y_l < x_k < y_{l+1}$. We recall that the distribution function $A^{(x_k)}$ of $X^{(x_k)}$ is given by

$$A^{(x_k)} = \begin{cases} A(x) & \text{if } x \leq x_k, \\ A(x_k) & \text{else,} \end{cases}$$
and introduce the Gaussian processes \( Y \) and \( Z \) of GREM type with the distribution functions

\[
A_Y(x) := \begin{cases} 
A(x) & \text{if } x \leq y_l, \\
A(y_l) & \text{if } y_l < x < x_k, \\
A(x_k) & \text{if } x \geq x_k,
\end{cases}
\]

\[
A_Z(x) := \begin{cases} 
A(x) & \text{if } x \leq y_l, \\
A(y_l) + \frac{x_k - y_l}{y_{l+1} - y_l} (A(y_{l+1}) - A(y_l)) & \text{if } y_l < x < x_k, \\
& \text{if } x \geq x_k,
\end{cases}
\]

respectively. Slepian’s lemma \cite[Lemma 10.2.1]{8} states that the less correlated a classical system is, the higher is the pressure, which yields the first inequality in

\[
\lim_{N \to \infty} \frac{1}{N} \ln \text{Tr} [Q_{X,N} e^{-\beta \sqrt{N} X^{(r_x)}(x_k)}] \leq \lim_{N \to \infty} \frac{1}{N} \ln \text{Tr} [Q_{X,N} e^{-\beta \sqrt{N} Y}] \leq \lim_{N \to \infty} \frac{1}{N} \ln \text{Tr} [Q_{X,N} e^{-\beta \sqrt{N} Z}].
\tag{3-3}
\]

For the second inequality, we recall that \( A \) is majorized by its concave hull \( \tilde{A} \) and agrees with \( \tilde{A} \) at \( y_l \) and \( y_{l+1} \), i.e.,

\[
A(x_k) \leq A(y_l) + \frac{x_k - y_l}{y_{l+1} - y_l} (A(y_{l+1}) - A(y_l)).
\]

Since the classical pressure is an increasing function of the jump heights, we arrive at the second bound in (3-3). As for (1-5), the classical free energy of \( Z \) is given in terms of the partial pressure (1-6) corresponding to the concave hull of \( A_Z \), which agrees with \( \tilde{A} \) up to \( x_k \),

\[
\lim_{N \to \infty} \frac{1}{N} \ln \text{Tr} [Q_{X,N} e^{-\beta \sqrt{N} Z}] = \sum_{j=1}^{l} \varphi^{(j)}(\beta) + \frac{x_k - y_l}{y_{l+1} - y_l} \varphi^{(l+1)}(\beta).
\]

Plugging this into (3-3), we obtain

\[
\lim_{N \to \infty} \frac{1}{N} \ln \text{Tr} [Q_{X,N} e^{-\beta \sqrt{N} X^{(r_x)}}] + (1 - x_k) \mathbb{E} [\ln \cosh(\beta b)] \\
\leq \sum_{j=1}^{l} \varphi^{(j)}(\beta) + (1 - y_l) \mathbb{E} [\ln \cosh(\beta b)] + \frac{x_k - y_l}{y_{l+1} - y_l} \varphi^{(l+1)}(\beta) - (y_{l+1} - y_l) \mathbb{E} [\ln \cosh(\beta b)]).
\]

Depending on the sign of the term in the last bracket we then arrive at

\[
\lim_{N \to \infty} \frac{1}{N} \ln \text{Tr} [Q_{X,N} e^{-\beta \sqrt{N} X^{(r_x)}}] + (1 - x_k) \mathbb{E} [\ln \cosh(\beta b)] \leq \sum_{j=1}^{l} \varphi^{(j)}(\beta) + (1 - y_l) \mathbb{E} [\ln \cosh(\beta b)]
\]

in case of a negative sign, or in case of a positive sign because of \( x_k - y_l \leq y_{l+1} - y_l \), at

\[
\lim_{N \to \infty} \frac{1}{N} \ln \text{Tr} [Q_{X,N} e^{-\beta \sqrt{N} X^{(r_x)}}] + (1 - x_k) \mathbb{E} [\ln \cosh(\beta b)] \leq \sum_{j=1}^{l+1} \varphi^{(j)}(\beta) + (1 - y_{l+1}) \mathbb{E} [\ln \cosh(\beta b)].
\]

Consequently, the term corresponding to any \( x_k \) on the left side of (3-2) is bounded by one of the terms on its right side.

The following observation is useful for the proof of Corollary 1.3:
Lemma 3.2. Let $\phi^{(j)}(\beta)$ be the partial pressures from (1-6) and $L_j := y_j - y_{j-1}$ the interval lengths. Then, the discrete concavity estimate

$$\frac{\phi^{(1)}(\beta)}{L_1} > \frac{\phi^{(2)}(\beta)}{L_2} > \cdots > \frac{\phi^{(m)}(\beta)}{L_m} \quad (3-4)$$

holds for any inverse temperature $\beta > 0$.

Proof. We call $\phi^{(j)}(\beta)$ “frozen” if $\beta > \beta_j$, i.e., $\phi^{(j)}(\beta)$ is given by the linear expression in (1-6). Otherwise we say $\phi^{(j)}(\beta)$ is “unfrozen”. By construction of the concave hull $A$, we know that the slopes $\gamma_j = \tilde{a}_j/L_j$ are strictly decreasing in $j$. The inequalities in (3-4), where two consecutive partial free energies are either both frozen or both unfrozen, are thus obvious. It remains to consider the case where $\phi^{(j)}(\beta)$ is frozen, but $\phi^{(j+1)}(\beta)$ is unfrozen. By (1-6) we then have

$$\frac{\phi^{(j)}(\beta)}{L_j} = \beta \sqrt{2 \ln 2} \gamma_j \quad \text{and} \quad \frac{\phi^{(j+1)}(\beta)}{L_{j+1}} = \frac{\beta^2}{2} \gamma_{j+1} + \ln 2.$$ 

Moreover, as $\phi^{(j)}(\beta)$ is frozen and $\phi^{(j+1)}(\beta)$ is unfrozen, the inverse temperatures satisfy

$$\beta_j = \sqrt{(2 \ln 2) \gamma_j^{-1}} < \beta \leq \sqrt{(2 \ln 2) \gamma_{j+1}^{-1}} = \beta_{j+1}.$$ 

We thus conclude that

$$\frac{\phi^{(j)}(\beta)}{L_j} = \frac{\beta}{\beta_j} 2 \ln 2 > 2 \ln 2 \geq \ln 2 + \frac{\beta^2}{\beta_{j+1}^2} \ln 2 = \ln 2 + \frac{\beta^2}{2} \gamma_{j+1} = \frac{\phi^{(j+1)}(\beta)}{L_{j+1}}. \quad \square$$

Remark. If $f$ denotes the function on $[0, 1]$ which linearly interpolates between the restricted classical free energies $\lim_{N \to \infty} \frac{1}{N} \ln \text{Tr}_{\Omega_{\beta_N}} e^{-\beta \sqrt{N} X^{(\alpha)}}$, then Lemmas 3.1 and 3.2 in particular show that $F$, the convex hull of the graph of $f$, is a polygon with the extreme points $(y_1, f(y_1))$. On the other hand, if this fact is assumed to be known, the assertion of Lemma 3.1 can be derived as follows. Consider the function $g(x, y) := (x, y + a(1 - x))$ with some $a > 0$. It is then easy to show that the set of extreme points of $g(F)$ coincides with $g(\text{ext}(F))$. Since the maximum in (3-1) is attained at an extreme point of $g(K)$, the claim follows.

Proof of Theorem 1.2 and Corollary 1.3. Theorem 1.2 is an immediate consequence of (3-1) and Lemma 3.1.

It remains to show Corollary 1.3. To this end, let us introduce the energy differences

$$\Delta^{(j)}(\beta, \Gamma) := (y_j - y_{j-1}) \ln 2 \cosh(\beta \Gamma) - \phi^{(j)}(\beta).$$

In view of Lemma 3.2, we conclude:

1. If $\Delta^{(j)}(\beta, \Gamma) < 0$ for some $j \geq 1$, then $\Delta^{(i)}(\beta, \Gamma) < 0$ for all $0 < i \leq j$.
2. If $\Delta^{(j)}(\beta, \Gamma) \geq 0$ for some $j \geq 1$, then $\Delta^{(i)}(\beta, \Gamma) \geq 0$ for all $j \leq i \leq m$.

Consequently, the maximum in (1-13) is attained at $m$ if all energy differences $\Delta^{(j)}$ are negative for $0 < j \leq m$ and, otherwise at the minimal integer $k < m$ such that $\Delta^{(k+1)} \geq 0$. We may thus rewrite the
We note that the condition \( \Delta^{(l)} > 0 \) is equivalent to \( \Gamma > \Gamma^{(l)}_{c} (\beta) \). This concludes the proof of (1-15). □

Our next goal is to prove Theorem 1.4 and Corollary 1.5. It is convenient to use Theorem 1.2 and the interpolation estimate (2-17) rather than the general Theorem 2.8. To do so, we first establish some continuity properties of the functions

\[
\Phi(\beta, A, z) := \sqrt{2 \ln 2} \beta \int_{0}^{\min \{ x(\beta), z \}} \sqrt{\bar{a}(x)} \, dx + \mathbb{1}_{z > x(\beta)} \left( \frac{\beta^2}{2} (\bar{A}(z) - \bar{A}(x(\beta))) + \ln 2(z - x(\beta)) \right)
\]

with respect to the distribution function \( A \). Therefore, we emphasize here the dependence on \( A \) in notation.

**Lemma 3.3.** Let \( A \) and \( (A_n)_{n \in \mathbb{N}} \) be distribution functions on \([0, 1]\) such that \( A_n \) converges uniformly to \( A \) as \( n \to \infty \). Then:

1. The concave hulls \( \bar{A}_n \) converge uniformly to \( \bar{A} \) as \( n \to \infty \), i.e.,
   \[
   \lim_{n \to \infty} \| \bar{A} - \bar{A}_n \|_{\infty} = 0.
   \]
2. The right derivatives \( \bar{a}_n(x) \) converge to \( \bar{a}(x) \) at any \( x \) where \( \bar{a} \) is continuous.
3. For any \( \beta \geq 0 \), the functions \( \Phi(\beta, A_n, z) \) converge uniformly to \( \Phi(\beta, A, z) \) as a function of \( z \), i.e.,
   \[
   \lim_{n \to \infty} \| \Phi(\beta, A_n, \cdot) - \Phi(\beta, A, \cdot) \|_{\infty} = 0.
   \]

**Proof.**

1. The function \( \bar{A} + \| \bar{A} - \bar{A}_n \|_{\infty} \) is a concave function which majorizes \( A_n \); that is, \( \bar{A}_n \leq \bar{A} + \| \bar{A} - \bar{A}_n \|_{\infty} \). Similarly, one shows that \( \bar{A} \leq \bar{A}_n + \| \bar{A} - \bar{A}_n \|_{\infty} \). The first assertion is a direct consequence of these bounds.

2. Since \( \bar{A}_n \) is a sequence of concave functions converging uniformly to \( \bar{A} \), the second claim follows from standard convex analysis (see, e.g., [36]).

3. We first recall that \( x(\beta, A) = \sup \{ x \mid \bar{a}(x) > 2 \ln 2/\beta^2 \} \). Since \( \bar{a} \) is a decreasing function, \( \bar{a} \) is continuous except for an at most countable set. The second statement implies then that \( x(\beta, A_n) \) converges to \( x(\beta, A) \). Next, we rewrite
   \[
   \Phi(\beta, A, z) = \int_{0}^{z} \varphi(\beta, A, x) \, dx
   \]
   with the function
   \[
   \varphi(\beta, A, x) := \beta \sqrt{2 \ln 2} \bar{a}(x) \mathbb{1}_{x < x(\beta, A)} + \left( \frac{\beta^2}{2} \bar{a}(x) + \ln 2 \right) \mathbb{1}_{x \geq x(\beta)}.
   \]
   Therefore, it suffices to show
   \[
   \lim_{n \to \infty} \int_{0}^{1} |\varphi(\beta, A, x) - \varphi(\beta, A_n, x)| \, dx = 0.
   \]
Due to our previous considerations, we know that \( \varphi(\beta, A_n, x) \) converges almost everywhere (with respect to the Lebesgue measure and \( x \)) to \( \varphi(\beta, A_n, x) \). Moreover, the functions \( \varphi(\beta, A_n, \cdot) \) are uniformly bounded at \([\delta, 1]\) for any \( \delta > 0 \), since \( \varphi(\beta, A_n, x) \) is decreasing in \( x \). By dominated convergence we then obtain

\[
\lim_{n \to \infty} \int_{\delta}^{1} |\varphi(\beta, A, x) - \varphi(\beta, A_n, x)| \, dx = 0
\]

for any \( \delta > 0 \). On the other hand,

\[
\int_{0}^{\delta} |\varphi(\beta, A, x) - \varphi(\beta, A_n, x)| \, dx \leq \int_{0}^{\delta} \varphi(\beta, A, x) + \varphi(\beta, A_n, x) \, dx \leq \frac{\beta^2}{2} (\bar{A}(\delta) + \bar{A}_n(\delta)) + 2\delta \ln 2.
\]

Since \( \bar{A} \) is continuous, \( \bar{A}(0) = 0 \) and the sequence \( \bar{A}_n \) converges uniformly, the third assertion follows as \( \delta \to 0 \).

We are now ready to show Theorem 1.4 and Corollary 1.5.

**Proof of Theorem 1.4 and Corollary 1.5**. We pick a sequence of step functions \( A_n \), which are also distribution functions and converge uniformly to the distribution function \( A \). By Theorem 2.8, the expression for \( \Phi(\beta, b, A_n) \) may be rewritten as

\[
\Phi(\beta, b, A_n) = \sup_{0 \leq z \leq 1} \left[ \Phi(\beta, A_n, z) + (1 - z) E[\ln 2 \cosh(\beta b)] \right].
\]

By the interpolation estimate (2-17), the left side converges to the corresponding limit of the quantum CREM’s pressure \( \Phi(\beta, b, A) \), whereas the right side converges to

\[
\lim_{n \to \infty} \sup_{0 \leq z \leq 1} \left[ \Phi(\beta, A_n, z) + (1 - z) E[\ln 2 \cosh(\beta b)] \right] = \sup_{0 \leq z \leq 1} \left[ \Phi(\beta, A, z) + (1 - z) E[\ln 2 \cosh(\beta b)] \right]
\]

by Lemma 3.3. This completes the proof of Theorem 1.4.

In the case where \( \bar{A} \) is continuously differentiable and \( b = \Gamma \), the convex function \([0, 1] \ni z \mapsto \Phi(\beta, A, z) + (1 - z) \ln(2 \cosh(\beta \Gamma)) \) possesses a maximum in the interior of its domain if and only if there exists a solution \( z \in (0, 1) \) of

\[
\frac{\partial \Phi(\beta, A, z)}{\partial z} - \ln 2 \cosh(\beta \Gamma) = 0.
\]

Otherwise the maximum is attained at \( z = 0 \) or \( z = 1 \). A straightforward calculation then leads to the formula in Corollary 1.5.

\[ \square \]

### 3B. The nonhierarchical GREM in a transversal field

We start with the proof of Theorem 1.6. In the following we will use the notation introduced in Section 1.3.

**Proof of Theorem 1.6.** Our strategy is to adapt the proof of Corollary 2.7. To be more precise, we introduce for any subset \( J \in \mathcal{P} \) the restriction \( B^J \) of \( B \) to the subgraph spanned by the spins \( \sigma_J \),

\[
B^J := \sum_{k \in J} B^{(k)}, \quad B^{(k)} := B^{1, x_k} - B^{1, x_{k-1}}.
\]
and set $B^\emptyset := 0$. For any two subsets $I, J \in \mathcal{P}$ and any potential $V_N$ independent of $X_{\sigma_I}^J$, we have
\[
\lim_{N \to \infty} \frac{1}{N} \ln \mathrm{Tr} e^{-\beta (\sqrt{N a_I} X_{\sigma_I}^J + V_N - B^J)} - \max_{\mathcal{F} \subseteq \mathcal{P}, D \in \mathcal{P}, D^c \cap \mathcal{F} = \emptyset} \ln \mathrm{Tr} e^{-\beta (\sqrt{a_F} N X_{\sigma_F}^D - B^{D^c})} + \max_{\mathcal{F} \subseteq \mathcal{P}, D \in \mathcal{P}, D^c \cap \mathcal{F} = \emptyset} \ln \mathrm{Tr} e^{-\beta (\sqrt{a_F} N X_{\sigma_F}^D - B^{D^c})} \bigg| = 0. \tag{3-5}
\]

We note that $B^J$ can be represented as a transversal magnetic field whose weights corresponding to the complement $J^c$ are set zero. Thus, (3-5) follows from Theorem 2.3 after possibly rearranging the spin components. Using (3-5) successively for each subset $J \in \mathcal{P}$ (where the remaining potential $V_N$ might change from step to step), we finally arrive at
\[
\lim_{N \to \infty} \frac{1}{N} \ln \mathrm{Tr} e^{-\beta (U - B)} - \max_{\mathcal{F} \subseteq \mathcal{P}, D \in \mathcal{P}, D^c \cap \mathcal{F} = \emptyset} \ln \mathrm{Tr} e^{-\beta (\sum_{F \in \mathcal{P}(D)} \sqrt{a_F} N X_{\sigma_F}^D - B^{D^c})} \bigg| = 0
\]
where $D^c \cap \mathcal{F} = \emptyset$ is understood elementwise, that is, $D^c \cap F = \emptyset$ for any $F \in \mathcal{F}$. We note that the convexity of the exponential and the variables $X_{\sigma_J}$ being centered Gaussians, implies (e.g., by (3-5)) for any $D^c \cap \mathcal{F} = \emptyset$,
\[
\lim_{N \to \infty} \frac{1}{N} \ln \mathrm{Tr} e^{-\beta (\sum_{F \in \mathcal{P}(D)} \sqrt{a_F} N X_{\sigma_F}^D - B^{D^c})} - \ln \mathrm{Tr} e^{-\beta (\sum_{F \in \mathcal{F}} \sqrt{a_F} N X_{\sigma_F}^D - B^{D^c})} \bigg| \geq 0,
\]
where $\mathcal{P}(D)$ is the power set of $D$. On the other hand, the limit of the left term exists almost surely and is given by
\[
\lim_{N \to \infty} \frac{1}{N} \ln \mathrm{Tr} e^{-\beta (\sum_{F \in \mathcal{P}(D)} \sqrt{a_F} N X_{\sigma_F}^D - B^{D^c})} = \min_{S \in D} \Phi_D(\beta, S) + \min_{k \in D^c} \sum_{I \subseteq D^c} \Phi_D(\beta, S) + \sum_{k \in D^c} \mathbb{E} [\ln 2 \cosh(\beta b)],
\]
where we used the strong law of large numbers for the expression involving $B^{D^c}$ and the known convergence [7] of the classical nonhierarchical GREM. We in fact need a slightly more generalized version of [7] which is also applicable to the reduced model on the subgraph generated by $\sigma_{D^c}$. However, this can be proved in the exactly same manner as the result on the whole graph $Q_N$. Combining our findings, we arrive at the claim (1-20). \hfill \Box

It remains to show Corollary 1.7. To this end, we need the corresponding result in the case $b = 0$ for the classical nonhierarchical GREM, which is a simple consequence of the derivation in [7] (see Remark 7 in that paper). For completeness, we spell out this classical result as Lemma A.3 in the Appendix.

**Proof of Corollary 1.7.** Let $S_0$ be the minimizing chain of the classical problem as spelled out in Lemma A.3. After relabelling the components of $\sigma$, we may assume that
\[
S_0 = \{ \emptyset, \{1\}, \{1, 2\}, \ldots, \{1, \ldots, n\} \}. \tag{3-6}
\]
We will show that
\[
\max_{D \in \mathcal{P}} \left[ \min_{S \in C^D} \Phi(\beta, S) + \sum_{k \in D^c} L_k \mathbb{E} [\ln 2 \cosh(\beta b)] \right] = \Phi(\beta, b, S_0)
\]
by establishing two inequalities. First, abbreviating
\[
D_k := \{1, \ldots, k\}
\]
with $D_0 := \emptyset$, we have
\[
\max_{D \in \mathcal{P}} \left[\min_{S \in \mathcal{C}^D} \Phi(\beta, S) + \sum_{k \in D^c} L_k \mathbb{E}[\ln 2 \cosh(\beta b)]\right] \\
\geq \max_{0 \leq k \leq n} \left[\min_{S \in \mathcal{C}^{D_k}} \Phi(\beta, S) + \sum_{k \in D_k} L_k \mathbb{E}[\ln 2 \cosh(\beta b)]\right] \\
= \max_{0 \leq k \leq n} \left[\Phi(\beta, S_0^{D_k}) + \sum_{k \in D_k} L_k \mathbb{E}[\ln 2 \cosh(\beta b)]\right] = \Phi(\beta, b, S_0).
\]

Here $S_0^D$ denotes the chain which coincides with $S_0$ but ends at $D$. The last line follows from Lemma A.3 as it implies that even in the constrained setting the cut versions of $S_0$ are indeed minimizing chains.

For the reverse inequality, let $I_1, \ldots, I_m$ be the sets associated to the concave hull $\overline{A}_{S_0}$ and let $\varphi_{S_0}(\beta)$ be the partial pressure corresponding to the GREM assigned to the chain $S_0$; see (1-6). Moreover, for any $D \in \mathcal{P}$ we define the ordered-restriction chain $S_0^D$,
\[
S_0^D := \{\emptyset, \{j_1\}, \{j_1, j_2\}, \ldots, \{j_1, \ldots, j_{k_D}\}\},
\]
where $j_1 < j_2 < \cdots < j_{k_D} \in D$ and $\{j_1, \ldots, j_{k_D}\} = D$. Then for any $D \in \mathcal{P}$,
\[
\min_{S \in \mathcal{C}^D} \Phi(\beta, S) + \sum_{j \in D^c} L_j \mathbb{E}[\ln 2 \cosh(\beta b)] \leq \Phi(\beta, S_0^D) + \sum_{j \in D^c} L_j \mathbb{E}[\ln 2 \cosh(\beta b)]
\]
\[
= \sum_{l=1}^{m_D} \varphi_{S_0}^{(l)}(\beta) + \sum_{j \in D^c} L_j \mathbb{E}[\ln 2 \cosh(\beta b)] \leq \sum_{l=1}^{m} \sum_{k \in I_l \cap D} L_k \varphi_{S_0}^{(l)}(\beta) + \sum_{j \in D^c} L_j \mathbb{E}[\ln 2 \cosh(\beta b)].
\]

The last inequality, follows from three observations. First, we recall that the weights $\varphi_{S_0}^{(l)}$ assigned to the chain $S_0^D$ are less than or equal to the weights $\varphi_{S_0}^{(l)}$ of the chain (3-6). Secondly, we note that the increments $\Delta_l \overline{A}_{S_0}$ on the segments $D \cap I_l \neq \emptyset$ can be bounded,
\[
\frac{\Delta_l \overline{A}_{S_0}}{\sum_{k \in I_l \cap D} L_k} \leq \frac{\Delta_l \overline{A}_{S_0}}{\sum_{k \in I_l} L_k},
\]
since otherwise we may construct a chain $S'$ violating Lemma A.3 using the first observation. Thirdly, an application of Slepian’s lemma as in the proof of Lemma 3.1 extends the summation to $m$ and yields the claimed inequality. We thus obtain
\[
\sum_{l=1}^{m} \left[\sum_{k \in I_l \cap D} \frac{L_k}{L_k} \varphi_{S_0}^{(l)}(\beta) + \sum_{j \in D^c \cap I_l} L_j \mathbb{E}[\ln 2 \cosh(\beta b)]\right] \leq \sum_{l=1}^{m} \max_{k \in I_l} \left\{\frac{\varphi_{S_0}^{(l)}(\beta)}{\sum_{k \in I_l} L_k}, \mathbb{E}[\ln 2 \cosh(\beta b)]\right\} = \Phi(\beta, b, S_0),
\]
where the last equality is based on the concavity, Lemma 3.2 and the explicit expression (1-13) for the pressure of the quantum GREM. This completes the proof as $D$ was chosen arbitrarily. \qed
Appendix: Supplementary results

**Sufficient condition for Assumption 2.1.** We want to present a quite general condition on the distribution of $X_{\sigma_1}$ which implies the third point in Assumption 2.1:

**Lemma A.1.** Let $X_{\sigma_1}$ be independent and identically distributed centered random variables which satisfy an LDP with good rate function $I$, i.e., the sets $\{x \mid I(x) \leq a\}$ are compact for any $a \geq 0$. Moreover, the rate function shall satisfy

$$\inf_{|x| > \varepsilon} I(x) > 0$$

for any $\varepsilon > 0$. Then, $(X_{\sigma_1})$ fulfills the conditions (1), (2) and (3) in Assumption 2.1.

**Proof.** The points (1) and (2) are clear and it remains to check (3). Let $w_{\sigma_1}$ be random weights which are independent of $X_{\sigma_1}$ and satisfy almost surely $w_{\sigma_1} \geq 0$ and $\mathbb{P}(\sigma_1) = 1$. We introduce the sets

$$A_N := \{\sigma_1 \in \mathcal{Q}_N^{(1)} \mid w_{\sigma_1} \geq 1/N^2\}$$

and show separately

$$\lim_{N \to \infty} \frac{1}{N} \sum_{\sigma_1 \in A_N} w_{\sigma_1} X_{\sigma_1} = 0 \quad (A-1)$$

and

$$\lim_{N \to \infty} \frac{1}{N} \sum_{\sigma \in A_N^c} w_{\sigma_1} X_{\sigma_1} = 0. \quad (A-2)$$

**Proof of (A-1).** We apply the trivial bound

$$\left| \frac{1}{N} \sum_{\sigma_1 \in A_N} w_{\sigma_1} X_{\sigma_1} \right| \leq \frac{1}{N} \sup_{\sigma_1 \in A_N} |X_{\sigma_1}|,$$

and note that the cardinality of $A_N$ is bounded by $N^2$. The independence of $w_{\sigma_1}$ and $X_{\sigma_1}$ then implies, for any $\delta > 0$,

$$\mathbb{P}(\sup_{\sigma_1 \in A_N} |X_{\sigma_1}| \geq \delta N) \leq N^2 \mathbb{P}(|X_{\sigma_1}| \geq \delta N) \leq N^2 e^{-(c_\delta + o(1))N}$$

with $c_\delta = \inf_{|x| \geq \delta} I(x) > 0$.

Therefore, the bound on the probability is summable in $N$ for any $\delta > 0$ and a Borel–Cantelli argument finishes the proof of (A-1). \qed

**Proof of (A-2).** As $I$ is a good rate function, we find a constant $C > 0$ such that

$$\inf_{|x| \geq C} I(x) \geq 2 \ln 2,$$

and hence

$$\mathbb{P}(\sup_{\sigma_1 \in \mathcal{Q}_N^{(1)}} |X_{\sigma_1}| \geq CN) \leq 2^N e^{-(2N \ln 2 + o(1))N} = (2 + o(1))^{-N}.$$
By a Borel–Cantelli argument we may assume without loss of generality that $|X_{\sigma_1}| \leq CN$ holds true for all sufficiently large $N$ with probability one. Conditioning on $w_{\sigma_1}$, the independence of $X_{\sigma_1}$ and $\mathbb{E}[X_{\sigma_1}] = 0$ implies

$$\mathbb{E}\left[ \left( \frac{1}{N} \sum_{\sigma_1 \in A_N^c} w_{\sigma_1} X_{\sigma_1} \right)^2 \right] \leq \frac{1}{N^2} \mathbb{E}\left[ \sum_{\sigma_1 \in A_N^c} w_{\sigma_1}^2 X_{\sigma_1}^2 \right] \leq C^2 \mathbb{E}\left[ \sum_{\sigma_1 \in A_N^c} w_{\sigma_1}^2 \right] \leq C^2 \mathbb{E}\left[ \sum_{\sigma_1 \in A_N^c} w_{\sigma_1} \right] \leq \frac{C^2}{N^2}.$$  

In the second and third inequalities, we use $|X_{\sigma_1}| \leq CN$ and $w_{\sigma_1} \leq N^{-2}$ for $\sigma_1 \in A_N^c$. The Borel–Cantelli lemma again completes the proof. \hfill \Box

**Assumption 2.2 for independent $L^1$ weights.** The aim of this section is to verify that Assumption 2.2 is satisfied for independent copies $(b_j)$ of an absolutely integrable variable $b$:

**Lemma A.2.** If the weights $b_i$ are independent copies of an absolutely integrable variable $b$, we almost surely have

$$\limsup_{N \to \infty} N^{-1} \sqrt{\sum_{i=1}^{N} |b_i|^2} = 0. \quad \text{(A-3)}$$

**Proof:** Our proof relies on a thinning and truncation argument and is similar to the proof of the strong law of large numbers in the $L^1$-case.

Let us abbreviate the partial sums $S_N := \sum_{i=1}^{N} |b_i|^2$. We pick some $\varepsilon > 0$ and introduce the sequence $N_m := 2^m$. Suppose we have already shown the almost sure convergence

$$\lim_{m \to \infty} (N_m)^{-1} S_{N_m} = 0. \quad \text{(A-4)}$$

Since $S_N$ is an increasing sequence we conclude that

$$\limsup_{N \to \infty} \frac{S_N}{N^2} \leq \limsup_{m \to \infty} \frac{S_{N_m}}{N^2_{m-1}} = 4 \limsup_{m \to \infty} \frac{S_{N_m}}{N^2_{m}} = 0.$$  

So it suffices to show (A-4). To this end, let $K_m$ be a nonnegative sequence which we will fix later and $S_{N_m}^{<}$, $S_{N_m}^{>}$ be the truncated sums given by

$$S_{N_m}^{<} := \sum_{i=1}^{N_m} |b_i|^2 1_{|b_i| \leq K_m} \quad \text{and} \quad S_{N_m}^{>} := S_{N_m} - S_{N_m}^{<}.$$  

For any $\varepsilon > 0$, a Markov-type estimate yields

$$\mathbb{P}(S_{N_m}^{<} > \varepsilon N_m^2) \leq \frac{\mathbb{E}[|b|^2 1_{|b| \leq K_m}]}{\varepsilon N_m}.$$  

We also have

$$\mathbb{P}(S_{N_m}^{>} \neq 0) \leq N_m \mathbb{P}(|b| > K_m).$$
The assertion follows by a Borel–Cantelli argument if we can choose \( K_m \) such that
\[
\sum_{m=1}^{\infty} \frac{\mathbb{E}[|b|^2 \mathbb{1}_{|b| \leq K_m}]}{N_m} + \sum_{m=1}^{\infty} N_m \mathbb{P}(|b| > K_m) < \infty.
\]
We claim that this can be accomplished by setting \( N_m = K_m \). To this end, we note that the second sum is finite as \( b \) is absolutely integrable. Moreover,
\[
\sum_{m=1}^{\infty} \frac{\mathbb{E}[|b|^2 \mathbb{1}_{|b| \leq N_m}]}{N_m} \leq 2 \sum_{m=1}^{\infty} N_m^2 \mathbb{P}(|b| \geq N_m) \sum_{k \geq m} N_m^{-1} \leq 4 \sum_{m=1}^{\infty} N_m \mathbb{P}(|b| \geq N_m) < \infty,
\]
where the first inequality is a consequence of the layer-cake representation and the last bound is again a consequence of \( b \) being absolutely integrable.

\[\square\]

**The pressure of the classical nonhierarchical GREM.** The following lemma is contained in the analysis of [7]. We provide a proof for the convenience of the reader.

**Lemma A.3.** Let \( X \) be a Gaussian vector of nonhierarchical GREM type. Then, there exists a chain \( S_0 \) such that for any chain \( S \) the pointwise estimate
\[
\tilde{A}_S(x) \leq \tilde{A}_{S_0}(x)
\]
holds true, where \( \tilde{A}_S \) and \( \tilde{A}_{S_0} \) are the concave hulls of the ordinary GREM vectors assigned to \( S \) and \( S_0 \), respectively. Moreover, we have, for any \( \beta \geq 0 \),
\[
\Phi(\beta) = \min_{S \in \mathcal{C}} \Phi(\beta, S) = \Phi(\beta, S_0).
\]

**Proof.** For any \( \emptyset \neq J \subset \{1, \ldots, n\} \) we define the corresponding slope \( \gamma_J \):
\[
\gamma_J := \frac{\tilde{a}_J}{L_J} := \frac{\sum_{I \subset J} a_I}{\sum_{k \in J} L_k}.
\]
We now construct a (possibly incomplete) chain \( J_1 \subset J_2 \subset \cdots \subset J_m = \{1, \ldots, n\} \) as follows. We first pick a subset \( J_1 \) with maximal slope \( \gamma_{J_1} \). If \( J_1 = \{1, \ldots, n\} \), we are done. Otherwise we pick a subset \( J_2 \) such that
\[
\gamma_{J_2} = \max_{I \subset \{1, \ldots, n\}: I \not\subset J_1} a_I.
\]
One easily checks that \( \gamma_{J_2} \leq \gamma_{J_1 \cup J_2} \), so we may assume that \( J_1 \subset J_2 \). We stop if \( J_2 = \{1, \ldots, n\} \) and continue the procedure otherwise. After at most \( n \) steps we arrive at a (possibly incomplete) chain as claimed. We set \( S_0 \) to be a completion of \( J_1, \ldots, J_m \), that is, \( S_0 \) is a chain which contains \( J_1, \ldots, J_m \). Clearly, \( S_0 \) does not depend on \( \beta \).

Both assertions follow now easily. We see that the concave hull \( \tilde{A}_{S_0} \) assigned to \( S_0 \) is the unique piecewise linear function satisfying \( \tilde{A}_{S_0}(L_{J_k}) = \tilde{a}_{J_k} \) for any \( k \). By construction, \( \tilde{A}_{S_0} \) is pointwise maximal as we iteratively pick the subset \( J_k \) leading to the maximal mean slope. On the other hand, the bound \( \tilde{A}_{S_0} \geq \tilde{A}_S \) for any chain \( S \), by Slepian’s lemma, yields \( \Phi(\beta, S_0) \leq \Phi(\beta, S) \) from which the second statement follows. \[\square\]
Acknowledgments

This work was supported by the DFG under EXC-2111–390814868.

References

[1] J. Adame and S. Warzel, “Exponential vanishing of the ground-state gap of the quantum random energy model via adiabatic quantum computing”, J. Math. Phys. 56:11 (2015), art. id. 113301. MR Zbl

[2] A. Adhikari and C. Brennecke, “Free energy of the quantum Sherrington–Kirkpatrick spin-glass model with transverse field”, J. Math. Phys. 61:8 (2020), art. id. 083302. MR Zbl

[3] M. Aizenman, J. L. Lebowitz, and D. Ruelle, “Some rigorous results on the Sherrington–Kirkpatrick spin glass model”, Comm. Math. Phys. 112:1 (1987), 3–20. MR Zbl

[4] L.-P. Arguin and N. Kistler, “Microcanonical analysis of the random energy model in a random magnetic field”, J. Stat. Phys. 157:1 (2014), 1–16. MR Zbl

[5] L.-P. Arguin and R. Persechino, “The free energy of the GREM with random magnetic field”, pp. 37–61 in Statistical mechanics of classical and disordered systems, edited by V. Gayrard et al., Springer Proc. Math. Stat. 293, Springer, 2019. MR Zbl

[6] V. Bapst, L. Foini, F. Krzakala, G. Semerjian, and F. Zamponi, “The quantum adiabatic algorithm applied to random optimization problems: the quantum spin glass perspective”, Phys. Rep. 523:3 (2013), 127–205. MR

[7] E. Bolthausen and N. Kistler, “On a nonhierarchical version of the generalized random energy model”, Ann. Appl. Probab. 16:1 (2006), 1–14. MR Zbl

[8] A. Bovier, Statistical mechanics of disordered systems: A mathematical perspective, Cambridge Series in Statistical and Probabilistic Mathematics 18, Cambridge University Press, 2006. MR Zbl

[9] A. Bovier and A. Klimovsky, “Fluctuations of the partition function in the generalized random energy model with external field”, J. Math. Phys. 49:12 (2008), art. id. 125202. MR Zbl

[10] A. Bovier and I. Kurkova, “Derrida’s generalised random energy models, I: Models with finitely many hierarchies”, Ann. Inst. H. Poincaré Probab. Statist. 40:4 (2004), 439–480. MR Zbl

[11] A. Bovier and I. Kurkova, “Derrida’s generalized random energy models, II: Models with continuous hierarchies”, Ann. Inst. H. Poincaré Probab. Statist. 40:4 (2004), 481–495. MR Zbl

[12] A. J. Bray and M. A. Moore, “Replica theory of quantum spin glasses”, Journal of Physics C: Solid State Physics 13:24 (1980), L655–L660.

[13] A. Burin, “Localization and chaos in a quantum spin glass model in random longitudinal fields: Mapping to the localization problem in a Bethe lattice with a correlated disorder”, Annalen der Physik 529:7 (2017), art. id. 1600292.

[14] D. Capocaccia, M. Cassandro, and P. Picco, “On the existence of thermodynamics for the generalized random energy model”, J. Statist. Phys. 46:3-4 (1987), 493–505. MR

[15] N. Crawford, “Thermodynamics and universality for mean field quantum spin glasses”, Comm. Math. Phys. 274:3 (2007), 821–839. MR Zbl

[16] B. Derrida, “Random-energy model: limit of a family of disordered models”, Phys. Rev. Lett. 45:2 (1980), 79–82. MR

[17] B. Derrida, “Random-energy model: an exactly solvable model of disordered systems”, Phys. Rev. B (3) 24:5 (1981), 2613–2626. MR Zbl

[18] B. Derrida, “A generalization of the random energy model which includes correlations between energies”, Journal de Physique Lettres 46 (1985), 401–407.

[19] B. Derrida and E. Gardner, “Solution of the generalised random energy model”, Journal of Physics C: Solid State Physics 19:13 (may 1986), 2253–2274.

[20] Y. Fedorov and E. Shender, “Quantum spin glasses in the Ising model with a transverse field”, JETP Lett. 43 (1986), 681–684.

[21] Y. Y. Goldschmidt, “Solvable model of the quantum spin glass in a transverse field”, Phys. Rev. B 41 (Mar 1990), 4858–4861.
[22] F. Guerra and F. L. Toninelli, “The thermodynamic limit in mean field spin glass models”, *Comm. Math. Phys.* **230**:1 (2002), 71–79. MR Zbl

[23] C. R. Laumann, A. Pal, and A. Scardicchio, “Many-body mobility edge in a mean-field quantum spin glass”, *Phys. Rev. Lett.* **113** (Nov 2014), art. id. 200405.

[24] H. Leschke, C. Manai, R. Ruder, and S. Warzel, “Existence of replica-symmetry breaking in quantum glasses”, *Phys. Rev. Lett.* **127**:20 (2021), art. id. 207204. MR Zbl

[25] H. Leschke, S. Rothlauf, R. Ruder, and W. Spitzer, “The free energy of a quantum Sherrington–Kirkpatrick spin-glass model for weak disorder”, *J. Stat. Phys.* **182**:3 (2021), art. id. 55. MR Zbl

[26] C. Manai and S. Warzel, “Phase diagram of the quantum random energy model”, *J. Stat. Phys.* **180**:1-6 (2020), 654–664. MR Zbl

[27] C. Manai and S. Warzel, “The quantum random energy model as a limit of $p$-spin interactions”, *Rev. Math. Phys.* **33**:1 (2021), art. id. 2060013. MR Zbl

[28] C. Manai and S. Warzel, “The de Almeida–Thouless line in hierarchical quantum spin glasses”, *J. Stat. Phys.* **186**:1 (2022), art. id. 14. MR

[29] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin glass theory and beyond*, World Scientific Lecture Notes in Physics 9, World Scientific Publishing Co., Teaneck, NJ, 1987. MR Zbl

[30] T. Obuchi, H. Nishimori, and D. Sherrington, “Phase diagram of the $p$-spin-interacting spin glass with ferromagnetic bias and a transverse field in the infinite-$p$ limit”, *Journal of the Physical Society of Japan* **76**:5 (2007), art. id. 054002.

[31] D. Panchenko, “The Sherrington–Kirkpatrick model: an overview”, *J. Stat. Phys.* **149**:2 (2012), 362–383. MR Zbl

[32] D. Panchenko, “The Parisi formula for mixed $p$-spin models”, *Ann. Probab.* **42**:3 (2014), 946–958. MR Zbl

[33] G. Parisi, “The order parameter for spin glasses: a function on the interval 0-1”, *Journal of Physics A: Mathematical and General* **13**:3 (mar 1980), 1101–1112.

[34] D. W. Robinson, “Statistical mechanics of quantum spin systems”, *Comm. Math. Phys.* **6** (1967), 151–160. MR Zbl

[35] B. Simon, *The statistical mechanics of lattice gases, Vol. I*, Princeton University Press, 1993. MR Zbl

[36] B. Simon, *Convexity: An analytic viewpoint*, Cambridge Tracts in Mathematics **187**, Cambridge University Press, 2011. MR Zbl

[37] S. Suzuki, J.-i. Inoue, and B. K. Chakrabarti, *Quantum Ising phases and transitions in transverse Ising models*, 2nd ed., Lecture Notes in Physics **862**, Springer, 2013. MR Zbl

[38] M. Talagrand, “The Parisi formula”, *Ann. of Math.* (2) **163**:1 (2006), 221–263. MR

[39] M. Talagrand, *Mean field models for spin glasses, Volume I: Basic examples*, Ergeb. Math. Grenzgeb. **54**, Springer, 2011. MR

[40] M. Talagrand, *Mean field models for spin glasses, Volume II: Advanced replica-symmetry and low temperature*, Ergeb. Math. Grenzgeb. **55**, Springer, 2011. MR

[41] K. Usadel, “Spin glass transition in an Ising spin system with transverse field”, *Solid State Communications* **58**:9 (1986), 629–630.

Received 15 Jul 2020. Accepted 27 Jun 2021.

**Chokri Manai:** chokri.manai@tum.de
Zentrum Mathematik, Technische Universität München, Garching bei München, Germany

**Simone Warzel:** warzel@ma.tum.de
Zentrum Mathematik, Technische Universität München, Garching bei München, Germany

---

**Source:** Generalized Random Energy Models in a Transversal Magnetic Field, Mathematische Zeitschrift. Mathematical Sciences Publishers.