



Spectral Analysis of the Quantum Random Energy Model

Chokri Manai¹ , Simone Warzel^{1,2}

¹ MCQST and Department of Mathematics, Technische Universität München, Munich, Germany.

E-mail: chokri.manai@tum.de

² Department of Physics, Technische Universität München, Munich, Germany

E-mail: simone.warzel@tum.de

Received: 11 February 2022 / Accepted: 20 April 2023

Published online: 12 June 2023 – © The Author(s) 2023

Abstract: The quantum random energy model (QREM) is a random matrix of Anderson-type which describes effects of a transversal magnetic field on Derrida’s spin glass. The model exhibits a glass phase as well as a classical and a quantum paramagnetic phase. We analyze in detail the low-energy spectrum and establish a localization-delocalization transition for the corresponding eigenvectors of the QREM. Based on a combination of random matrix and operator techniques as well as insights in the random geometry, we derive next-to-leading order asymptotics for the ground-state energy and eigenvectors in all regimes of the parameter space. Based on this, we also deduce the next-to-leading order of the free energy, which turns out to be deterministic and on order one in the system size in all phases of the QREM. As a result, we determine the nature of the fluctuations of the free energy in the spin glass regime.

1. Introduction and Main Results

1.1. Quantum random energy model. In the theory of disordered systems the random energy model (REM) is a simple, yet ubiquitous toy model. It assigns to every N -bit or Ising string $\sigma = (\sigma_1, \dots, \sigma_N) \in \{-1, 1\}^N =: \mathcal{Q}_N$ a rescaled Gaussian random variable

$$U(\sigma) := \sqrt{N} \omega(\sigma)$$

with $(\omega(\sigma))$ forming 2^N canonically realized independent and identically distributed (i.i.d.) random variables with standard normal law \mathbb{P} . The Hamming cube \mathcal{Q}_N is rendered a graph by declaring two bit strings connected by an edge if they differ by a single bit flip: introducing the flip operators $F_j \sigma := (\sigma_1, \dots, -\sigma_j, \dots, \sigma_N)$ on components $j \in \{1, \dots, N\}$, the edges of the Hamming cube are formed by all pairs of the form $(\sigma, F_j \sigma)$. The graph’s negative adjacency matrix

$$(T\psi)(\sigma) := - \sum_{j=1}^N \psi(F_j \sigma)$$

is defined on $\psi \in \ell^2(\mathcal{Q}_N)$, the 2^N -dimensional Hilbert space of complex-valued functions on N -bit strings. Since every vertex in \mathcal{Q}_N has a constant degree N , the negative graph Laplacian, $T + N\mathbb{1}$, just differs by N times the identity matrix. We study the quantum random energy model (QREM) which is the random matrix

$$H := \Gamma T + U, \tag{1.1}$$

where $\Gamma \geq 0$ is a parameter, and U is diagonal in the canonical configuration basis (δ_σ) of $\ell^2(\mathcal{Q}_N)$, i.e., $U\delta_\sigma = U(\sigma)\delta_\sigma$ and $\psi(\sigma) = \langle \delta_\sigma | \psi \rangle$. As usual in mathematical physics, we choose the scalar product $\langle \cdot | \cdot \rangle$ on $\ell^2(\mathcal{Q}_N)$ to be linear in its second component.

The QREM is a random matrix of Anderson type—albeit on a quite unconventional graph whose connectivity grows to infinity with the system size N , and with a scaling of the random potential U which enforces the operator norm of both, T and U , to be of the same order N (cf. (1.4) and (1.9)). It is thus natural to investigate the localization properties of its eigenfunctions. The interest in the QREM is however many-faceted. In mathematical biology, the model has received attention under the name REM House-of-Cards model [63] as an element of a simplistic probabilistic model of population genetics, in which \mathcal{Q}_N is the space of gene types and U encodes their fitness [7, 8, 39, 42]. In this interpretation, the operator T implements mutations of the gene type, and one is interested in the long-time limit of the semigroup generated by H (cf. [6], in which the parameter regime $\Gamma = \kappa/N$ with fixed $\kappa > 0$ corresponding to the normalized Laplacian is considered).

The Anderson-perspective has also attracted attention in discussions of many-body or Fock-space localization, where the QREM occasionally serves as an analytically more approachable toy to test ideas about more realistic disordered spin systems [9, 14, 27, 45]. We will comment on some of the conjectures in the physics literature concerning the localization properties of the eigenfunctions after presenting our results on this topic.

In statistical mechanics, the QREM was introduced [37] as a simplified model to investigate the quantum effects caused by a transversal magnetic field on classical mean-field spin-glass models [62]. In this context, the Hilbert space $\ell^2(\mathcal{Q}_N)$ is unitarily identified with the tensor-product Hilbert space $\otimes_{j=1}^N \mathbb{C}^2$ of N spin- $\frac{1}{2}$ quantum objects. A corresponding unitary maps the canonical basis (δ_σ) to the tensor-product basis in which the Pauli- z -matrix is diagonal on each tensor component. The Pauli matrices $\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, $\sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ are naturally lifted to $\otimes_{j=1}^N \mathbb{C}^2$ by their action on the j th tensor component, $\sigma_j^x := \mathbb{1} \otimes \dots \otimes \sigma^x \otimes \dots \otimes \mathbb{1}$. Upon the above unitary equivalence, T corresponds to $-\sum_{j=1}^N \sigma_j^x$, i.e., a constant field in the negative x -direction exerted on all N spin- $\frac{1}{2}$ (cf. [52]). In this interpretation, the random potential U is the energy operator of the spin- $\frac{1}{2}$ -objects, which interact disorderly only through their z -components. Derrida [28, 29] originally invented the classical REM potential U as a simplification to other mean-field spin glasses such as the Sherrington-Kirkpatrick model.

The phenomenon common to such classical spin glass models is a glass freezing transition into a low temperature phase which, due to lack of translation invariance, is described by an order parameter (due to Parisi) more complicated than a global magnetization [54, 57, 58, 65]. In the absence of external fields the latter typically vanishes. These thermodynamic properties are encoded in the (normalized) partition function

$$Z(\beta, \Gamma) := 2^{-N} \text{Tr} e^{-\beta H}$$

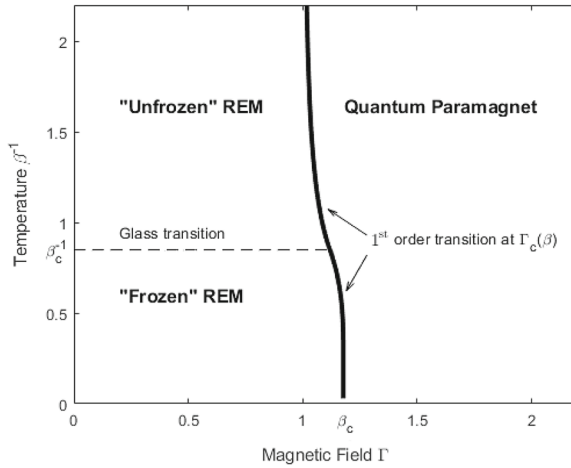


Fig. 1. Phase diagram of the QREM as a function of the transversal magnetic field Γ and the temperature β^{-1} [37, 50]. The first-order transition occurs at fixed β and $\Gamma_c(\beta)$. The freezing transition is found at temperature β_c^{-1} , which is unchanged in the presence of a small magnetic field

at inverse temperature $\beta \in [0, \infty)$, or, equivalently, its pressure

$$\Phi_N(\beta, \Gamma) := \ln Z(\beta, \Gamma). \tag{1.2}$$

Up to a factor of $-\beta^{-1}$, the latter coincides with the free energy. In the thermodynamic limit $N \rightarrow \infty$, the specific pressure of the REM converges almost surely [17, 28, 29, 56],

$$N^{-1} \Phi_N(\beta, 0) \rightarrow p^{\text{REM}}(\beta) = \begin{cases} \frac{1}{2} \beta^2 & \text{if } \beta \leq \beta_c, \\ \frac{1}{2} \beta_c^2 + (\beta - \beta_c) \beta_c & \text{if } \beta > \beta_c. \end{cases} \tag{1.3}$$

It exhibits a freezing transition into a low-temperature phase characterized by the vanishing of the specific entropy above

$$\beta_c := \sqrt{2 \ln 2}.$$

In the presence of the transversal field, the spin-glass phase of the REM disappears for large $\Gamma > 0$ and a first-order phase transition into a quantum paramagnetic phase described by

$$p^{\text{PAR}}(\beta \Gamma) := \ln \cosh(\beta \Gamma)$$

occurs at the critical magnetic field strength

$$\Gamma_c(\beta) := \beta^{-1} \operatorname{arcosh} \left(\exp \left(p^{\text{REM}}(\beta) \right) \right).$$

In particular, $\Gamma_c(0) = 1$ and $\Gamma_c(\beta_c) = \beta_c^{-1} \operatorname{arcosh}(2)$. The precise location of this first-order transition and the shape of the phase diagram of the QREM, which we sketch in Fig. 1, had been predicted by Goldschmidt [37] in the 1990s and was rigorously established in [50].

Proposition 1.1 [50]. *For any $\Gamma, \beta \geq 0$ we have the almost sure convergence as $N \rightarrow \infty$:*

$$N^{-1} \Phi_N(\beta, \Gamma) \rightarrow \max \{ p^{\text{REM}}(\beta), p^{\text{PAR}}(\beta \Gamma) \}.$$

1.2. *Low energy states.* Through the low-temperature limit $\beta \rightarrow \infty$, Proposition 1.1 contains also information on the ground state energy of the QREM,

$$N^{-1} \inf \text{spec } H \rightarrow \begin{cases} -\beta_c & \text{if } \Gamma < \beta_c, \\ -\Gamma & \text{if } \Gamma > \beta_c. \end{cases}$$

The critical coupling for this quantum phase transition is at the endpoint $\lim_{\beta \rightarrow \infty} \Gamma_c(\beta) = \beta_c$ of the first order phase transition. As will be demonstrated below, this ground-state phase transition at $\Gamma = \beta_c$ is manifested by a change of the nature of the corresponding eigenvector from sharply localized to (almost) uniformly delocalized. To provide some heuristics, it is useful to compare the ground-state energy and eigenvectors of the two operators entering $H = \Gamma T + U$:

1. The spectrum of T consists of $N + 1$ eigenvalues,

$$\text{spec } T = \{2n - N \mid n \in \mathbb{N}_0, n \leq N\}, \tag{1.4}$$

with degeneracy given by the binomials $\binom{N}{n}$. The corresponding ℓ^2 -normalized eigenvectors are the natural orthonormal basis for the Hadamard transformation, which diagonalizes T . They are indexed by subsets $A \subset \{1, \dots, N\}$:

$$\Phi_A(\sigma) := \frac{1}{\sqrt{2^N}} \prod_{j \in A} \sigma_j. \tag{1.5}$$

The eigenvalue to Φ_A is $2|A| - N$ with $|A|$ the cardinality of the set. In particular, the unique ground-state of ΓT is Φ_\emptyset with energy $-N\Gamma$. All eigenvectors Φ_A are maximally uniformly delocalized over the Hamming cube.

2. In contrast, all eigenvectors δ_σ of U are maximally localized. The REM’s minimum energy, $\min U$, is roughly at $-N\beta_c$. For $\eta > 0$ the event that $\|U\|_\infty := \max_{\sigma \in \mathcal{Q}_N} |U(\sigma)| > (\beta_c + \eta)N$ has exponentially small probability, i.e.,

$$\begin{aligned} \Omega_{N,\eta}^{\text{REM}} &:= \{\|U\|_\infty \leq (\beta_c + \eta)N\}, \\ \mathbb{P}(\Omega_{N,\eta}^{\text{REM}}) &\geq 1 - 2^{N+1} e^{-\frac{1}{2}(\eta+\beta_c)^2 N} = 1 - 2 e^{-N(\eta\beta_c + \frac{\eta^2}{2})}, \end{aligned} \tag{1.6}$$

where the inequality follows from the union bound and a Markov-Chernoff estimate. A more precise description of the extremal value statistics of $\min U$ is [17, 46]

$$\mathbb{P}(\min U \geq s_N(x)) = \left(1 - 2^{-N} e^{-x+o(1)}\right)^{2^N} \tag{1.7}$$

for any x in terms of the function s_N given by

$$s_N(x) := -\beta_c N + \frac{\ln(N \ln 2) + \ln(4\pi)}{2\beta_c} - \frac{x}{\beta_c}. \tag{1.8}$$

By symmetry of the distribution, a similar expression applies to the maximum.

These limiting cases suggest the following heuristic, perturbative description of the ground-state of $H = \Gamma T + U$ in the regimes of small and large Γ . To our knowledge, it goes back to [40]:

1. For small Γ , second-order perturbation theory starting from the vector $\delta_{\sigma_{\min}}$, which is localized at $\sigma_{\min} := \arg \min U$, reads:

$$\begin{aligned} \inf \text{spec } H &\approx \min U + \Gamma \langle \delta_{\sigma_{\min}} | T \delta_{\sigma_{\min}} \rangle + \Gamma^2 \\ &\sum_{\sigma \neq \sigma_{\min}} \frac{|\langle \delta_{\sigma} | T \delta_{\sigma_{\min}} \rangle|^2}{U(\sigma_{\min}) - U(\sigma)} \approx -N\beta_c - \frac{\Gamma^2}{\beta_c}. \end{aligned} \quad (1.9)$$

The first-order term vanishes. The sum in the second-order term is restricted to the neighbors of the minimum, whose potential term typically is only of the order \sqrt{N} .

2. For large Γ , second-order perturbation theory starting from the ground state Φ_{\emptyset} of T reads:

$$\begin{aligned} \inf \text{spec } H &\approx -N\Gamma + \langle \Phi_{\emptyset} | U \Phi_{\emptyset} \rangle - \sum_{A \neq \emptyset} \frac{|\langle \Phi_{\emptyset} | U \Phi_A \rangle|^2}{2|A| \Gamma} \\ &\approx -N\Gamma - \frac{\langle \Phi_{\emptyset} | U^2 \Phi_{\emptyset} \rangle}{N \Gamma} \approx -N\Gamma - \frac{1}{\Gamma}. \end{aligned} \quad (1.10)$$

The next-to-leading term, $\langle \Phi_{\emptyset} | U \Phi_{\emptyset} \rangle = 2^{-N} \sum_{\sigma \in \mathcal{Q}_N} U(\sigma)$, vanishes by the law of large numbers. In the order Γ^{-1} -term, one uses the approximation that most of the states of T are found near $|A| \approx N/2$. As will be explained in more detail in Sect. 3, one crucial point is that U is exponentially small when restricted to the eigenspace of T outside an interval around $|A| \approx N/2$. By a decomposition of unity one is therefore left with $\langle \Phi_{\emptyset} | U^2 \Phi_{\emptyset} \rangle \approx N$, again by the law of large numbers.

Unlike in a finite-dimensional situation, higher orders in this naive perturbation theory turn out to be of lower order with N^{-1} the relevant parameter. One result of this paper is that these predictions can be confirmed: for $\Gamma < \beta_c$ the ground state is sharply localized near a lowest-energy configuration of the REM. In contrast, for $\Gamma > \beta_c$ the ground state resembles the maximally delocalized state given by the ground state of T . In both cases, the ground state is energetically separated and the ground-state gap only closes exponentially near $\Gamma = \beta_c$, see also [1]. In fact, we do not only restrict attention to the ground state but characterize a macroscopic window of the entire low-energy spectrum in the different parameter regimes.

Before delving into the details, let us emphasize that the localization-delocalization transition at extreme energies presented here relies on the delocalization properties of T on \mathcal{Q}_N , which fundamentally differ from the finite-dimensional situation. The eigenfunctions of T can only form localized states from linear combinations in the center of its spectrum. This is given a precise mathematical formulation in the form of novel estimates on the spectral shift and Green function of Dirichlet restrictions of T to Hamming balls in Sect. 2, and random matrix estimates on projections of multiplication operators in Sect. 3. A separation of the localized versus delocalized parts of the spectrum beyond the extremal energies, on which the subsequent results concerning the finite-size corrections of the free energy rest, is facilitated by a novel detailed description of the geometry of extremal fluctuations the REM in Sect. 5.

Aside from Theorem 1.10, our results pertain to fixed, but arbitrarily large N on the product probability space Ω_N corresponding to 2^N i.i.d. standard normal random variables whose product measure we denote by \mathbb{P} . We suppress its dependence on N . In this setting, the results apply to all realizations ω , aside from exceptional events whose

probability will be estimated and go (exponentially) to zero as $N \rightarrow \infty$. This strong concentration enables in Theorem 1.10 (and also Proposition 1.1) the use of Borel-Cantelli arguments, applied to the product space $\prod_{N=1}^\infty \Omega_N$. To present our results and estimates in a precise, yet reader-friendly, manner, we will make use of an ‘‘indexed’’ version of Landau’s O-notation.

Definition 1.2. Let $\Theta = (\theta_1, \dots, \theta_k)$ be a tuple of parameters, $(a_N)_{N \in \mathbb{N}}$ a real and $(b_N)_{N \in \mathbb{N}}$ a positive sequence. We then write

$$a_N = \mathcal{O}_\Theta(b_N) \quad \text{if} \quad \limsup_{N \rightarrow \infty} \frac{|a_N|}{b_N} \leq C(\Theta), \tag{1.11}$$

for some positive constant $C(\Theta)$, which may depend on Θ . Analogously, we write

$$a_N = o_\Theta(b_N) \quad \text{if} \quad |a_N| \leq c_N(\Theta)|b_N|, \tag{1.12}$$

where $c_N(\Theta)$ denotes a sequence which tends to zero.

In particular, the appearing constant $C(\Theta)$ or, respectively sequence $c_N(\Theta)$, does not depend on any other parameters in question not included in Θ . That is, if a_N is a random sequence and the realization ω of the randomness is not included in the list Θ of parameters, the estimates are understood to hold uniformly on the event of interest.

1.2.1. Paramagnetic regime $\Gamma > \beta_c$. Our first main result shows that in the paramagnetic regime the addition of the REM shifts the eigenvalues (1.4) of T at energies below the minimum of U deterministically.

Theorem 1.3. For $\Gamma > \beta_c$ and any $\tau \in (0, 1)$ there are events $\Omega_{N,\tau}^{\text{par}}$ with probability

$$\mathbb{P}(\Omega_{N,\tau}^{\text{par}}) \geq 1 - e^{-N/C}$$

and $C \in (0, \infty)$ a universal constant such that for all sufficiently large N and any $\eta > 0$ on $\Omega_{N,\tau}^{\text{par}} \cap \Omega_{N,\eta}^{\text{REM}}$ (cf. (1.6)) all eigenvalues of $H = \Gamma T + U$ below $-(\beta_c + 2\eta)N$ are found in the union of intervals of radius $\mathcal{O}_{\Gamma,\eta}(N^{\frac{\tau-1}{2}})$ centered at

$$(2n - N)\Gamma + \frac{N}{(2n - N)\Gamma} \tag{1.13}$$

with $n \in \{m \in \mathbb{N}_0 \mid (2m - N)\Gamma < -(\beta_c + 2\eta)N\}$. Moreover, the interval centered at (1.13) contains exactly $\binom{N}{n}$ eigenvalues of H .

For the ground-state in the regime $\Gamma > \beta_c$, Theorem 1.3 implies that with overwhelming probability

$$\inf \text{spec } H = -\Gamma N - \frac{1}{\Gamma} + o_\Gamma(1). \tag{1.14}$$

The energy shift with respect to the ground state of ΓT is as predicted by naive second-order perturbation theory (1.10). Second-order perturbation theory for the eigenvalues corresponds to first-order perturbation theory for the eigenvectors: the eigenvectors are well approximated by their first order corrections. In particular, the ground state in the paramagnetic phase is close to the fully paramagnetic state Φ_\emptyset . This is made more precise in our next main result, whose proof alongside that of Theorem 1.3 can be found in Sect. 3.

Theorem 1.4. *In the situation of Theorem 1.3 on the event $\Omega_{N,\tau}^{\text{par}} \cap \Omega_{N,\eta}^{\text{REM}}$ with $0 < \eta < (\Gamma - \beta_c)/4$ the ℓ^2 -normalized ground state ψ of $H = \Gamma T + U$ satisfies:*

1. *The ℓ^2 -distance of ψ and Φ_\emptyset is $\|\psi - \Phi_\emptyset\| = \mathcal{O}_\Gamma(N^{\frac{\tau-1}{2}})$.*
2. *The ground state ψ is exponentially delocalized in the maximum norm, i.e.*

$$\|\psi\|_\infty^2 \leq 2^{-N} e^{N\gamma((\beta_c+\eta)/(2\Gamma))+o_\Gamma(N)}, \tag{1.15}$$

where $\gamma : [0, 1] \rightarrow \mathbb{R}$ denotes the binary entropy

$$\gamma(x) := -x \ln x - (1 - x) \ln(1 - x). \tag{1.16}$$

The true ℓ^2 -distance of the ground-state function to the fully delocalized state Φ_\emptyset is presumably of order $N^{-\frac{1}{2}}$ up to a logarithmic correction in N . The norm estimate (1.15) is not expected to be sharp: we conjecture a delocalization bound of the form $\|\psi\|_\infty^2 \leq 2^{-N+o(N)}$. Section 3, in which the proofs of Theorems 1.3 and 1.4 can be found, also contains (non-optimal) ℓ^∞ -delocalization estimates for all eigenvalues strictly below the threshold $-\beta_c N$ in the paramagnetic regime. The optimal decay rates for excited states are not known.

1.2.2. Spin glass regime $\Gamma < \beta_c$. In the spin glass phase the low-energy configurations of the REM, which occur on the extremal sites

$$\mathcal{L}_\varepsilon := \{\sigma \mid U(\sigma) \leq -\varepsilon N\} \quad \text{with } \varepsilon \in (0, \beta_c), \tag{1.17}$$

are also shifted by a deterministic, order-one correction by the transverse field as predicted by second-order perturbation theory. To characterize localization properties of the corresponding eigenvectors in the canonical z -basis, i.e., the configuration basis (δ_σ) of $\ell^2(\mathcal{Q}_N)$, we let

$$B_R(\sigma) := \{\sigma' \mid d(\sigma, \sigma') \leq R\}, \quad S_R(\sigma) := \{\sigma' \mid d(\sigma, \sigma') = R\}$$

stand for the Hamming ball and sphere of radius R , which are defined in terms of the Hamming distance

$$d(\sigma, \sigma') := \frac{1}{2} \sum_{i=1}^N |\sigma_i - \sigma'_i|$$

of two configurations $\sigma, \sigma' \in \mathcal{Q}_N$.

Theorem 1.5. *For $\Gamma < \beta_c$ and $\delta > 0$ small enough there are events $\Omega_{N,\Gamma,\delta}^{\text{loc}}$ with probability*

$$\mathbb{P}(\Omega_{N,\Gamma,\delta}^{\text{loc}}) \geq 1 - e^{-cN}$$

for some $c = c(\Gamma, \delta)$ such that the following applies for sufficiently large N on $\Omega_{N,\Gamma,\delta}^{\text{loc}}$:

1. *The eigenvalues E of $H = \Gamma T + U$ below $-(\beta_c - \delta)N$ and low-energy configurations $U(\sigma)$ are in a one-to-one correspondence such that*

$$E = U(\sigma) + \frac{\Gamma^2 N}{U(\sigma)} + \mathcal{O}_{\Gamma,\delta}(N^{-1/4}). \tag{1.18}$$

In particular, the estimate $\mathcal{O}_{\Gamma,\delta}(N^{-1/4})$ is independent of $\sigma \in \mathcal{L}_{\beta_c-\delta}$.

2. The ℓ^2 -normalized eigenvector ψ corresponding to E and σ concentrates near this configuration in the sense that:

(a) Close to extremum: For any $K \in \mathbb{N}$ and for all $\sigma' \in S_K(\sigma)$:

$$|\psi(\sigma')| = \mathcal{O}_{\Gamma, \delta, K}(N^{-K}), \quad \text{and} \quad \sum_{\sigma' \notin B_K(\sigma)} |\psi(\sigma')|^2 = \mathcal{O}_{\Gamma, \delta, K}(N^{-(K+1)}).$$

(b) Far from extremum: For any $0 < \alpha < 1$, there is $c_\alpha \in (0, \infty)$ such that

$$\sum_{\sigma' \notin B_{\alpha N}(\sigma)} |\psi(\sigma')|^2 \leq e^{-c_\alpha N}. \tag{1.19}$$

This theorem covers states in the extreme localization regime in which the eigenvectors are sharply localized—each in its own extremal site of the potential. In this regime, the estimates on the decay rate of the eigenvectors close to the extremum are optimal and far from the extremum they are optimal up to determining the decay rate c_α . Concrete, non-optimized values of the energy threshold $-N(\beta_c - \delta)$ as well as more precise values of the error terms can be found in the proof of Theorem 1.5 in Sect. 4. In essence, the localization analysis in Sect. 4 proves that resonances and tunneling among different large deviation sites does not play a role in this energy regime. An upper bound for our technique to fail is at $\delta = \beta_c/2$. The energy threshold at which eigenvectors are believed [9, 14] to occupy a positive fraction of \mathcal{Q}_N is strictly larger than $-N\beta_c/2$ and for small fields yet smaller than $-N\Gamma$.

The precise low energy statistics of the REM U beyond the location of the minimum (1.7) is well known. Utilizing the rescaling (1.8) around its minimal value, the point process

$$\sum_{\sigma \in \mathcal{Q}_N} \delta_{s_N^{-1}(U(\sigma))} \rightarrow \text{PPP}(e^{-x} dx) \tag{1.20}$$

converges weakly to the Poisson point process with intensity measure $e^{-x} dx$ on \mathbb{R} (i.e, when integrating the random measure against a continuous compactly supported function, the resulting random variables converge weakly, see e.g. [17, Thm 9.2.2] or [46]). Theorem 1.5 implies a similar result for the low energy statistics in the QREM.

Corollary 1.6. *Let $\Gamma < \beta_c$ and let*

$$s_N(x; \Gamma) := -\beta_c N + \frac{\ln(N \ln 2) + \ln(4\pi)}{2\beta_c} - \frac{\Gamma^2}{\beta_c} - \frac{x}{\beta_c}. \tag{1.21}$$

Then, the rescaled eigenvalue process $\text{spec } H$ of the QREM $H = \Gamma T + U$ converges weakly

$$\sum_{E \in \text{spec } H} \delta_{s_N^{-1}(E; \Gamma)} \rightarrow \text{PPP}(e^{-x} dx). \tag{1.22}$$

In particular, the ground state energy converges weakly

$$\inf \text{spec } H - \left(-\beta_c N + \frac{\ln(N \ln 2) + \ln(4\pi)}{2\beta_c} - \frac{\Gamma^2}{\beta_c} \right) \rightarrow -\frac{X}{\beta_c}, \tag{1.23}$$

where X is a random variable distributed according to the law of the maximum of a Poisson point process $\text{PPP}(e^{-x} dx)$ with intensity $e^{-x} dx$ on the real line.

Proof. Corollary 1.6 is a straightforward consequence of Theorem 1.5 combined with (1.20). \square

Theorem 1.5 in particular covers the ground-state of the QREM and thus extends the result [6, Lemma 2.1] on the leading asymptotics of the ground-state energy in the parameter regime $\Gamma = \kappa/N$ with $\kappa > 0$. The proof already contains more information on the ℓ^2 -properties of the ground-state eigenvector, which we record next. More can be said on its ℓ^1 -localization properties. The latter is of interest in the context of the interpretation of the QREM in population genetics [7, 8, 39].

Theorem 1.7. *For $\Gamma < \beta_c$ there are events $\hat{\Omega}_{N,\Gamma}^{\text{loc}}$ with probability*

$$\mathbb{P}(\hat{\Omega}_{N,\Gamma}^{\text{loc}}) \geq 1 - e^{-cN}$$

for some $c = c(\Gamma)$ such that on $\hat{\Omega}_{N,\Gamma}^{\text{loc}}$, for all N large enough there is $\delta > 0$ and $\sigma_0 \in \mathcal{L}_{\beta_c - \delta}$ such that the positive ℓ^2 -normalized ground state ψ of the QREM Hamiltonian is concentrated near σ_0 in the sense that:

1. the ℓ^2 -distance of ψ and δ_{σ_0} is $\|\psi - \delta_{\sigma_0}\|^2 = \mathcal{O}_\Gamma\left(\frac{1}{N}\right)$, and its first order correction

$$\xi := \sqrt{1 - \frac{\Gamma^2}{\beta_c^2 N}} \delta_{\sigma_0} + \frac{\Gamma}{\beta_c N} \sum_{\sigma \in S_1} \delta_\sigma \tag{1.24}$$

has the same energy as ψ up to order one, and $\|\psi - \xi\|^2 = \mathcal{O}_\Gamma\left(\frac{1}{N^2}\right)$.

2. the ℓ^1 -norm of ψ converges to a bounded constant:

$$\|\psi\|_1 = \sum_{\sigma} \psi(\sigma) = \frac{\beta_c}{\beta_c - \Gamma} + o_\Gamma(1), \tag{1.25}$$

and, for any $1 < p < \infty$: $\|\psi\|_p^p = \sum_{\sigma} |\psi(\sigma)|^p = 1 + o_{\Gamma,p}(1)$.

It is natural to assume that the configuration σ_0 on which the ground-state is asymptotically localized and the classical minimal configuration $\sigma_{\min} := \arg \min U$ agree. While this is true with high probability, it does not hold almost surely. In the situation of Theorem 1.7 one may show that there are two constants $C \geq c > 0$ such that for N large enough:

$$\frac{c}{N} \leq \mathbb{P}(\sigma_0 \neq \sigma_{\min}) \leq \frac{C}{N}. \tag{1.26}$$

The reason for this is found in the following description of low-energy eigenvalues,

$$E_\sigma = U(\sigma) - \frac{\Gamma^2}{\beta_c} + \frac{\Gamma^2}{\beta_c^2 N} Z_\sigma + \mathcal{O}_\Gamma(N^{-5/4}), \quad Z_\sigma := \frac{1}{N} \sum_{\sigma' \in S_1(\sigma)} U(\sigma'),$$

which is proved in Lemma 4.3 below and which takes into account the next leading term in comparison to (1.18). The random variables Z_σ are standard normal distributed and independent of the large deviations $U(\sigma)$ with $\sigma \in \mathcal{L}_{\beta_c - \delta}$ and $\delta > 0$ small enough. Since the extremal energies form a Poisson process with mean density of order one, the normal fluctuations in the energy-correction of order $\mathcal{O}(1/N)$ are able to cause the event (1.26). More generally, the method presented in this paper allows for a systematic control of subleading corrections in an expansion of the energy eigenvalues. As we will see, they are determined by potential fluctuations on increasing spheres around the extremal sites.

1.2.3. *Critical case* $\Gamma = \beta_c$. We complete the picture on the ground state by describing the situation in the critical case $\Gamma = \beta_c$, where the quantum phase transition occurs. Adapting techniques, one may also prove that typically one observes a paramagnetic behavior at criticality.

Proposition 1.8. *Let $\Gamma = \beta_c$. On an event of probability $1 - \mathcal{O}(N^{-1/2})$ the ground state is at $\inf \text{spec } H = -\Gamma N - \Gamma^{-1} + \mathcal{O}(N^{-1/4})$ and the eigenvector ψ is paramagnetic in the sense that $\|\psi - \Phi_\emptyset\| = \mathcal{O}(N^{-1/4})$. On an event of probability $\mathcal{O}(N^{-1/2})$ the ground state is at $\inf \text{spec } H = \min U - \Gamma + \mathcal{O}(N^{-1/4})$, and the eigenvector ψ is localized in the sense that $\|\psi - \delta_{\sigma_0}\| = \mathcal{O}(N^{-1/4})$.*

The heuristics explanation for this is the following. For $\Gamma = \beta_c$ the ground state energy of ΓT is given by $-\beta_c N$, whereas the classical minimal energy is given by $\min U = -\beta_c N + C \ln(N) + \mathcal{O}(1)$ with $C > 0$. The logarithmic correction in this expression ensures that the paramagnetic behavior is dominant. This argument also suggests that the phase transition should be observed at the N -dependent field strength Γ_N , where the energy predictions of Theorems 1.3 and 1.5 agree,

$$-\Gamma_N N - \frac{1}{\Gamma_N} = \min U + \frac{\Gamma_N^2 N}{\min U}$$

which leads to

$$\Gamma_N = -\frac{\min U}{N} + \frac{1}{N} \left(\frac{N}{\min U} - \frac{\min U}{N} \right) + o(N^{-1}). \tag{1.27}$$

Indeed, in an $o(N^{-1})$ neighborhood of Γ_N one can observe a sign of critical behavior, the exponential vanishing gap of the Hamiltonian.

Proposition 1.9. *Let $\Delta_N(\Gamma) > 0$ denote the energy gap of the QREM Hamiltonian. Then, for some $c > 0$ and N large enough*

$$\min_{\Gamma \geq 0} \Delta_N(\Gamma) \leq e^{-cN} \tag{1.28}$$

except for a exponentially small event. The minimum is attained at some Γ_N^* satisfying (1.27).

The proof of both Proposition 1.8 and 1.9 are found in the extended arXive version [49]. It relies on a spectral analysis of H and is completely different from the derivation in [1].

1.3. *Free energy and partition function.* The spectral techniques presented here also allow to pin down the pressure Φ_N and its fluctuations up to order one in N in all three phases of the QREM: the spin-glass phase as well as the classical ('unfrozen REM') and quantum paramagnetic phase, cf. Fig. 1.

Theorem 1.10. 1. *If $\Gamma > \Gamma_c(\beta)$ the pressure $\Phi_N(\beta, \Gamma)$ is up to order one deterministic and one has the almost sure convergence*

$$\Phi_N(\beta, \Gamma) - (\ln \cosh(\beta\Gamma))N \rightarrow \frac{\beta}{\Gamma \tanh(\beta\Gamma)}. \tag{1.29}$$

2. If $\Gamma < \Gamma_c(\beta)$ and $\beta \leq \beta_c$, the pressure $\Phi_N(\beta, \Gamma)$ differs from the REM's pressure $\Phi_N(\beta, 0)$ by a deterministic β -independent shift of order one, i.e., one has the almost sure convergence

$$\Phi_N(\beta, \Gamma) - \Phi_N(\beta, 0) \rightarrow \Gamma^2. \quad (1.30)$$

3. If $\Gamma < \Gamma_c(\beta)$ and $\beta > \beta_c$, the pressure $\Phi_N(\beta, \Gamma)$ differs from the REM's pressure by a deterministic β -dependent shift of order one, i.e., one has the almost sure convergence

$$\Phi_N(\beta, \Gamma) - \Phi_N(\beta, 0) \rightarrow \frac{\Gamma^2 \beta}{\beta_c}. \quad (1.31)$$

The proof of the almost-sure convergence, for which the probability space is the product $\prod_{N=1}^{\infty} \Omega_N$ of independently redrawn variables for every single N , is based on a Borel-Cantelli argument and contained in Sect. 5.

At all values of $\beta > 0$, the fluctuations of the REM's pressure $\Phi_N(\beta, 0)$ below its deterministic leading term $Np^{\text{REM}}(\beta)$ have been determined in [20] (see also [17, Thm. 9.2.1]). Their nature and scale changes from normal fluctuations on the scale $\exp(-\frac{N}{2}(\ln 2 - \beta^2))$ for $\beta \leq \beta_c/2$ into a more interesting form of exponentially small fluctuations in the regime $\beta \in (\beta_c/2, \beta_c)$. In the spin glass phase $\beta > \beta_c$, the fluctuations are of order one [34] and asymptotically described by Ruelle's partition function of the REM [59]. More precisely, one has the weak convergence [20, Thm. 1.6]:

$$e^{-N[\beta\beta_c - \ln 2] + \frac{\beta}{2\beta_c}[\ln(N \ln 2) + \ln 4\pi]} Z_N(\beta, 0) \rightarrow \int_{-\infty}^{\infty} e^{x\beta/\beta_c} \text{PPP}(e^{-x} dx). \quad (1.32)$$

As a consequence of Theorem 1.10, we thus obtain the analogous result for the QREM.

Corollary 1.11. *If $\Gamma < \Gamma_c(\beta)$ and $\beta > \beta_c$, we have the weak convergence:*

$$e^{-N[\beta\beta_c - \ln 2] + \frac{\beta}{2\beta_c}[\ln(N \ln 2) + \ln 4\pi] - \frac{\beta\Gamma^2}{\beta_c}} Z_N(\beta, \Gamma) \rightarrow \int_{-\infty}^{\infty} e^{x\beta/\beta_c} \text{PPP}(e^{-x} dx).$$

Proof. By the continuity of the exponential function, this follows immediately from (1.31) and (1.32). \square

The fluctuations of the QREM's partition function outside the spin glass phase are expected to be much smaller—for $\Gamma < \Gamma_c(\beta)$ and $\beta < \beta_c$ most likely on a similar scale as in the REM and for the paramagnetic regime presumably even smaller. The methods in this paper do not allow to determine fluctuations on an exponential scale.

1.4. Comments. We close this introduction by putting our main results into the broader context of related questions discussed in the physics and mathematics literature.

In the past years, the QREM has attained interest in the physics community as basic testing ground for quantum annealing algorithms [40, 41] and, somewhat related, physicist have started to investigate many-body localization in the QREM [9, 14, 22, 32, 45]. Based on numerical computations and non-rigorous methods such as the forward-scattering approximation and the replica trick, they predict a dynamical phase transition between ergodic and localized behavior in the parameter region $\Gamma < \Gamma_c(\beta)$, $\beta < \beta_c$. This transition is expected to be reflected in a change in the spread of eigenfunctions at the correspond energies, which in the ergodic regime is neither uniform nor localized. It is an interesting

mathematical challenge to investigate this. As this requires a good understanding of the eigenfunctions far away from the spectral edges, the methods presented in this paper are not yet sharp enough to tackle those problems.

In simplified models of Rosenzweig-Porter type, such non-ergodic delocalization regimes have been predicted [44,61] and confirmed by a rigorous analysis [68]. In an even more simplified model in which one replaces T by the orthogonal projection onto its ground-state $-|\Phi_\emptyset\rangle\langle\Phi_\emptyset|$, a fully detailed description of the localization-delocalization transition has been worked out in [5].

Focusing on the physics of spin glasses, the independence of the REM is an oversimplification. This was the main motivation for Derrida to introduce the Generalized Random Energy Model (GREM) [30,31], in which the basic random variables are correlated, but still with a prescribed hierarchical structure. The free energy of the GREM has been studied extensively [18,19,23,59]. On the quantum side, the specific free energy of the QGREM has been determined in [52]; and in [53] the effects of an additional longitudinal field have been considered. We expect that our methods can be adapted to the case of a finite-level QGREM to derive analogous results as in Theorems 1.3, 1.5 and 1.10. More precisely, we conjecture that the multiple phase transitions in the QGREM are reflected in the behavior of the ground state wavefunction, i.e., at the critical field strengths Γ_k the wavefunction undergoes a transition from being localized in the block σ_k to a delocalized states in the respective part of the spin components. The infinite-level case might require substantially new ideas, as standard interpolation techniques do not reveal order-one corrections. Our methods, however, are strong enough to cover non-Gaussian REM type models, i.e., a centered square integrable i.i.d. process, whose distribution satisfies a large deviation principle (see also [52, Assumption 2.1]). Clearly, explicit expressions in analogous versions of Theorem 1.3, 1.5 and 1.10 will depend on the distribution of the process as already the parameter β_c is specific to Gaussians.

Among spin glass models with a transversal field, the Quantum Sherrington-Kirkpatrick (QSK) model, in which one substitutes in (1.1) for U the classical SK potential, is of particular interest [62]. In contrast to the classical SK model, which is solved by Parisi's celebrated formula, such an explicit expression for the free energy of the QSK is lacking, and its analysis remains a physical and mathematical challenge. So far, the universality of the limit of the free energy has been settled in [25], and in [2] the limit of the free energy was expressed as a limit of Parisi-type formulas for high-dimensional vector spin glass models. Unfortunately, despite the knowledge of a Parisi-type formula, the qualitative features of the phase transition in the QSK could only be analyzed by other means, adapting the methods of [4,21]. In terms of the glass behavior, the analysis in [48] shows that the glass parameter vanishes uniformly in Γ for all $\beta \leq 1$. This is complemented by [47], where the existence of a glass phase has been established for $\beta > 1$ and weak magnetic fields Γ .

The localization-delocalization transition for the QREM differs drastically from related results on a finite-dimensional graph such as \mathbb{Z}^d (see e.g. [3,43] and references). Unlike on \mathbb{Z}^d , all low-energy eigenvectors on \mathcal{Q}_N are delocalized in a regime of large Γ (a regime, which is also absent if one takes $\Gamma = \kappa/N$ as in [6]). The localized states appear only for small Γ . Although the norm of the adjacency matrix T is on the same scale N as the random potential U , which is not the case for the any of the variety of unbounded distributions studied on subsets of \mathbb{Z}^d , the localization of eigenvectors for extremal energies is even stronger on \mathcal{Q}_N . For the Gaussian distribution studied here, the mass of the eigenvectors sharply concentrates not only for a finite number of eigenvalues in one of the extremal sites of U , but rather for all eigenvalues below a threshold (cp. [38])

with Theorem 1.5). In the finite-dimensional setting, the ground state and the first few excited states concentrate on a small, but growing subdomain of \mathbb{Z}^d and, hence, a finite ℓ^1 -norm for the ground state is specific to the QREM. This seemingly contradictory strong localization property compared to \mathbb{Z}^d can be traced to the adjacency matrix's T bad localization properties to balls, on which we elaborate in Sect. 2: the spectral shift due to localization on a ball of radius K is of order N and not K^{-2} as on \mathbb{Z}^d . This together with the sparseness of the potential's extremal sites does not allow for resonances (cf. Lemma 4.4). In this sense, our proof is in fact somewhat simpler (and hence also stronger) than existing proofs of localization in the extremal sites of a random potential on \mathbb{Z}^d . E.g. most recently and notably, in [15] the statistics of a finite number of eigenvalues above the ground state and the localization properties of their eigenvectors were studied for single-site distributions with doubly exponential tails (see also [43] for more references). While the degree of localization in the $\Gamma < \beta$ phase is significantly stronger than in the models studied in [15], we observe a similar exponential decay of the localized states for larger distances, and in both cases the extremal statistics is governed by a Poisson process. In the study of the parabolic Anderson model, an interesting question is how the shape of the localized eigenstates and the speed of convergence depend on the underlying distribution of the random potential [43]. For the sake of concreteness, we only study the most prominent case of a Gaussian distribution. Although several quantities such as the constant β_c depend crucially on the Gaussian nature, we expect the qualitative aspects of the localization–delocalization transition to be persistent even with other unbounded distributions (e.g. those which meet [52, Ass 2.1]).

The operator T coincides up to a diagonal shift N with the Laplacian, i.e., the generator of a simple clock process on \mathcal{Q}_N . This correspondence gives rise to yet another link with the parabolic Anderson model on \mathbb{Z}^d . The dynamics of the Anderson model is a vast research topic and its study has revealed many interesting phenomena such as ageing. The spin glass nature is believed to be reflected in non-equilibrium properties and a slow relaxation to equilibrium. However, aging in spin glasses is typically not studied under an unbiased random walk, but rather under the Glauber dynamics for which the transition rates depend on the sites' energies. In the case of the REM, the related Glauber dynamics has drawn considerable interest as a well treatable case for metastability and aging [10, 11, 24, 35, 36]. Our spectral methods might provide some further insights into the dynamics of REM-type clock processes.

2. Adjacency Matrix on Hamming Balls

This section collects results on the spectral properties of the restriction of T to Hamming balls. We focus on the analysis of the Green's function, which by rank-one perturbation theory, is closely related to the ground state for potentials corresponding to a narrow deep hole - a situation typically encountered in potentials of REM type. Most of the spectral analysis in the literature related to T is motivated by the theory of error corrections (see e.g. [16, 26, 33] and references therein). The methods we use are rather different and neither rely on elaborate combinatorics nor a Hadamard transformation, which is applicable on a full Hamming cube only.

2.1. Norm estimates. In the following, we fix $\sigma_0 \in \mathcal{Q}_N$ and $0 \leq K \leq N \in \mathbb{N}$. The restriction T_K of T to the Hamming ball $B_K(\sigma_0)$ is defined through its matrix elements in the canonical orthonormal basis on $\ell^2(B_K(\sigma_0))$, which is naturally embedded in

$\ell^2(\mathcal{Q}_N)$:

$$\langle \delta_\sigma | T_K \delta_{\sigma'} \rangle = \begin{cases} \langle \delta_\sigma | T \delta_{\sigma'} \rangle & \text{if } \sigma, \sigma' \in B_K(\sigma_0) \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

We start with two known results on T_K . The first part of the following lemma has been already proved in [33] in case $K = \varrho N$, and a simpler proof is included in [49]. The second part is just a special case of the spectral symmetry of any bipartite graph’s adjacency matrix (cf. [26]).

Proposition 2.1 (cf. [33]). *For the restriction T_K to balls $B_K(\sigma_0)$ of radius $K \leq N/2$:*

1. *The operator norm is bounded according to*

$$\|T_K\| \leq 2\sqrt{K(N - K + 1)}, \tag{2.2}$$

and for any radius ϱN with $0 < \varrho < 1/2$:

$$E_N(\varrho) := \inf \text{spec } T_{\varrho N} = -\|T_{\varrho N}\| = -2\sqrt{\varrho(1 - \varrho)N} + o_\varrho(N). \tag{2.3}$$

2. *If φ is an eigenvector of T_K , then $\hat{\varphi}$ given by $\hat{\varphi}(\sigma) := (-1)^{d(\sigma, \sigma_0)}\varphi(\sigma)$ is also an eigenvector of T_K with $\langle \hat{\varphi} | T_K \hat{\varphi} \rangle = -\langle \varphi | T_K \varphi \rangle$. Consequently, the spectrum is symmetric, $\text{spec}(T_K) = -\text{spec}(T_K)$.*

If K is of order one as a function of N , we have $\|T_K\| = \mathcal{O}_K(\sqrt{N})$. This drastic shift of the operator norm due to confinement should be compared to the finite-dimensional situation where this shift for a ball of radius K is proportional to K^{-2} .

In the remaining part of this section, we will analyze T_K and its Green function in the two extreme cases in relation to N : (1) fixed-size balls in Sect. 2.2, and (2) growing balls with radius $K = \varrho N$ with some $0 < \varrho < 1/2$ in Sect. 2.3.

2.2. Green function for balls of fixed size. The Green’s function of the operator T_K on $\ell^2(B_K(\sigma_0))$ is defined by

$$G_K(\sigma, \sigma_0; E) := \left\langle \delta_\sigma | (-T_K - E)^{-1} \delta_{\sigma_0} \right\rangle. \tag{2.4}$$

Before we derive decay estimates in case $E \notin [-\|T_K\|, \|T_K\|]$, we recall some general facts:

1. By radial symmetry, $G_K(\sigma, \sigma_0; E)$ only depends on the distance $d(\sigma, \sigma_0)$.
2. All ℓ^2 -normalized eigenvectors (φ_j) of T_K with eigenvalues (E_j) can chosen to be real, and we have

$$\begin{aligned} G_K(\sigma, \sigma_0; E) &= \sum_j \frac{\varphi_j(\sigma)\varphi_j(\sigma_0)}{E_j - E} = \sum_j (-1)^{d(\sigma, \sigma_0)} \frac{\varphi_j(\sigma)\varphi_j(\sigma_0)}{-E_j - E} \\ &= (-1)^{d(\sigma, \sigma_0)+1} G_K(\sigma, \sigma_0; -E), \end{aligned}$$

where the second equality follows from the symmetry of the spectrum stated in Lemma 2.1. Thus, it is sufficient to derive decay estimates for $E < -\|T_K\|$.

3. The Green function at $E < -\|T_K\|$ is related to the ground-state φ of the rank-one perturbation

$$H^{(E)} := T_K - \alpha^{(E)}|\delta_{\sigma_0}\rangle\langle\delta_{\sigma_0}| \tag{2.5}$$

on $\ell^2(B_K(\sigma_0))$. More precisely, by rank-one perturbation theory $\alpha^{(E)} := G_K(\sigma_0, \sigma_0; E)^{-1}$ is the unique value at which $H^{(E)}$ has a ground-state at $E < -\|T_K\|$, and

$$G_K(\sigma, \sigma_0; E) = \frac{1}{\alpha^{(E)}} \frac{\varphi(\sigma)}{\varphi(\sigma_0)}, \tag{2.6}$$

cf. [3, Theorem 5.3]. By the Perron-Frobenius theorem, φ and hence the Green function is strictly positive on $B_K(\sigma_0)$. A decay estimate for $G_K(\cdot, \sigma_0; E)$ translates to a bound on the ground state φ of $H^{(E)}$ and vice versa. Our proof of the localization results in Sect. 4 will make use of this relation.

In order to establish decay estimates, we employ the radial symmetry and write the Green function as a telescopic product

$$G_K(\sigma, \sigma_0; E) = \prod_{d=0}^{\text{dist}(\sigma, \sigma_0)} \Gamma_K(d; E) \tag{2.7}$$

with factors $\Gamma_K(0; E) := G_K(\sigma_0, \sigma_0; E)$ and

$$\Gamma_K(d; E) := \frac{G_K(\sigma, \sigma_0; E)}{G_K(\sigma', \sigma_0; E)}, \quad \text{if } 1 \leq d = \text{dist}(\sigma, \sigma_0) = \text{dist}(\sigma', \sigma_0) - 1.$$

The choice of $\sigma \in S_d(\sigma_0)$ and $\sigma' \in S_{d-1}(\sigma_0)$ in the last definition is irrelevant due to the radial symmetry.

The fundamental equation $(-T_K - E)G_K(\cdot, \sigma_0; E) = \delta_{\cdot, \sigma_0}$ yields for a configuration σ with $1 \leq d = \text{dist}(\sigma, \sigma_0) \leq K$

$$\begin{aligned} 0 &= [(T_K - E)G_K(\cdot, \sigma_0; E)](\sigma) \\ &= -d \prod_{j=0}^{d-1} \Gamma_K(j; E) - E \prod_{j=0}^d \Gamma_K(j; E) - (N - d) \prod_{j=0}^{d+1} \Gamma_K(j; E) \\ &= \left(\frac{d}{\Gamma_K(d; E)} - E + (N - d)\Gamma_K(d + 1; E) \right) \prod_{j=0}^d \Gamma_K(j; E), \end{aligned}$$

where we use the convention $\Gamma_K(K + 1; E) := 0$. In the case $d = 0$, we have $1 = (-N\Gamma_K(1; E) - E)\Gamma_K(0; E)$. That translates to the following recursive relation of Riccati type:

$$\Gamma_K(d; E) = \mathcal{M}_{d,E}(\Gamma_K(d + 1; E)), \quad 0 \leq d \leq K. \tag{2.8}$$

with the fractional linear transformation acting on \mathbb{C} :

$$\mathcal{M}_{d,E}(\Gamma) = \frac{\max\{d, 1\}}{-E - (N - d)\Gamma}. \tag{2.9}$$

We now analyze the behavior of solutions to the recursive relation in the various regimes of interest.

Proposition 2.2. *For any $K \in \mathbb{N}$ there is some $C_K < \infty$ such that for any $N > 2K$ and $E < -\|T_K\|$ we have*

$$G_K(\sigma, \sigma_0; E) \leq \frac{C_K}{|E + \|T_K\||} \binom{N}{d(\sigma, \sigma_0)}^{-1/2} \left(\frac{\sqrt{N}}{|E|}\right)^{d(\sigma, \sigma_0)}. \tag{2.10}$$

Since the proof of Proposition 2.2 is arguably simpler than the one of Proposition 2.4, we refer to [49] for the arguments.

2.3. Green function for growing balls. We now turn to the behavior of the Green’s function on balls, which grow with N . This will require a more detailed analysis of the recursion relation (2.8). To see what to expect, we first derive an estimate on the Green’s function of the full Hamming cube.

Lemma 2.3. *For any $N \in \mathbb{N}$, $E < -N = -\|T\|$ and $\sigma, \sigma_0 \in \mathcal{Q}_N$:*

$$G_N(\sigma, \sigma_0; E) \leq \frac{1}{|E + N|} \left(\frac{N}{|E|}\right)^d \binom{N}{d(\sigma, \sigma_0)}^{-1}. \tag{2.11}$$

Proof. The Neumann series formula readily implies the operator identity

$$\frac{1}{1 - X} = \sum_{k=0}^{d-1} X^k + X^d \frac{1}{1 - X} \tag{2.12}$$

for any operator with $\|X\| < 1$. Setting $d = d(\sigma, \sigma_0)$, we thus obtain

$$\left\langle \delta_\sigma | (T - E)^{-1} \delta_{\sigma_0} \right\rangle = \frac{-1}{E} \left\langle \delta_\sigma | (1 - T/E)^{-1} \delta_{\sigma_0} \right\rangle = \frac{1}{E^d} \left\langle \delta_\sigma \left| \frac{T^d}{T - E} \delta_{\sigma_0} \right. \right\rangle,$$

since terms in (2.12) corresponding to $k < d$ vanish. Radial symmetry of the Green function yields

$$\begin{aligned} \left\langle \delta_\sigma | (T - E)^{-1} \delta_{\sigma_0} \right\rangle &= \binom{N}{d}^{-1} \sum_{\sigma \in S_d(\sigma_0)} \left\langle \delta_\sigma | (T - E)^{-1} \delta_{\sigma_0} \right\rangle \\ &\leq \binom{N}{d}^{-1} \frac{\sqrt{2^N}}{E^d} \left\langle \Phi_\emptyset | T^d \frac{1}{T - E} \delta_{\sigma_0} \right\rangle = \binom{N}{d}^{-1} \left(\frac{N}{|E|}\right)^d \frac{1}{|E| - N}, \end{aligned}$$

where $\Phi_\emptyset(\sigma) = 2^{-N/2}$ denotes the lowest energy eigenfunction of T , and we applied the eigenfunction equation, $T\Phi_\emptyset = -N\Phi_\emptyset$, in the last step. \square

A main difference between the small versus large ball behavior of the Green’s function is in the factor $(\sqrt{N}/|E|)^d$ in (2.10) versus $(N/|E|)^d$ in (2.11). In the case of interest where $|E|$ is of order N , we arrive at a decay of the order $N^{-d/2}$ versus e^{-Cd} .

There are at least two strategies to derive upper bounds on the Green function $G_{\varrho N}(\sigma, \sigma_0; E)$ for $E < E_N(\varrho) = -2\sqrt{\varrho(1 - \varrho)}N + o(N)$ and $0 < \varrho < 1/2$, cf. (2.3). The first strategy is to apply the arguments, which led to (2.11) and which yield

$$G_{\varrho N}(\sigma, \sigma_0; E) \leq \frac{1}{E_N(\varrho) - E} \left(\frac{E_N(\varrho)}{E}\right)^d \frac{\Psi_\varrho(\sigma_0)}{\Psi_\varrho(\sigma)} \binom{N}{d}^{-1}, \tag{2.13}$$

with $\Psi_\varrho \in \ell^2(B_{\varrho N}(\sigma_0))$ the ℓ^2 -normalized, positive eigenfunction corresponding to $E_N(\varrho)$. It then remains to establish a bound on the ratio $\Psi_\varrho(\sigma_0)/\Psi_\varrho(\sigma)$. We, however, will instead proceed by an analysis of the factors $\Gamma_{\rho N}$ defined in (2.7).

Proposition 2.4. *Let $0 < \varrho < 1/2$, and $\varepsilon > 0$. Then for $E \leq E_N(\varrho) - \varepsilon N$, all $\sigma \in B_{\rho N}(\sigma_0)$ and all N large enough:*

$$G_{\varrho N}(\sigma, \sigma_0; E) \leq \frac{1}{\varepsilon N} \binom{N}{d(\sigma_0, \sigma)}^{-1/2} 2^{-\min\{d(\sigma_0, \sigma), \rho_0(\varrho)N\}} \tag{2.14}$$

where $0 < \varrho_0(\varrho) < \varrho$ is the unique solution of the equation $2\sqrt{\varrho(1-\varrho)} = 3\sqrt{\varrho_0(1-\varrho_0)}$. Moreover, for any fixed $K \in \mathbb{N}$ there is some $C_K < \infty$ such that for all N large enough:

1. for all $\sigma \in S_K(\sigma_0)$: $G_{\varrho N}(\sigma, \sigma_0; E) \leq \frac{1}{\varepsilon N} \frac{C_K}{\sqrt{N^K}} \binom{N}{d(\sigma_0, \sigma)}^{-1/2}$.
2. $\sum_{\sigma \notin B_K(\sigma_0)} G_{\varrho N}(\sigma, \sigma_0; E)^2 \leq \frac{C_K}{\varepsilon^2 N^{K+2}}$.

Proof. It is convenient to separate the combinatorial factor $\binom{N}{d(\sigma_0, \sigma)}^{-1/2}$ and study

$$\hat{G}_{\varrho N}(\sigma, \sigma_0; E) := \binom{N}{d(\sigma_0, \sigma)}^{1/2} G_{\varrho N}(\sigma, \sigma_0; E) = \prod_{d=0}^{d(\sigma, \sigma_0)} \hat{\Gamma}_{\varrho N}(d; E). \tag{2.15}$$

By direct inspection of (2.15) one obtains the relation $\hat{\Gamma}_{\varrho N}(d; E) := \sqrt{\frac{N-d}{d}} \Gamma_{\varrho N}(d; E)$ for $d \geq 1$, which in turn implies the recursive relation

$$\begin{aligned} \hat{\Gamma}_{\varrho N}(d; E) &= \frac{1}{\frac{|E|}{V(d)} - m(d)\hat{\Gamma}_{\varrho N}(d; E)} \quad \text{for } 1 \leq d \leq \varrho N \\ \text{with } V(d) &:= \sqrt{d(N-d)}, \quad m(d) := \sqrt{\frac{(d+1)(N-d)}{d(N-d+1)}} \quad \text{and} \\ \hat{\Gamma}_{\varrho N}(\varrho N + 1; E) &= 0, \quad \hat{\Gamma}_{\varrho N}(0; E) = \Gamma_{\varrho N}(0; E) = G_{\varrho N}(\sigma_0, \sigma_0; E). \end{aligned} \tag{2.16}$$

We will now analyze the solution of these recursive relations.

We first claim that for all N large enough:

$$\hat{\Gamma}_{\varrho N}(d; E) \leq 1 \text{ for all } d \in [\varrho_0 N, \varrho N]. \tag{2.17}$$

This is proven by induction on d starting from $d = \varrho N + 1$, where it trivially holds. For the induction step from $d + 1$ to d , we recall that $E_N(\varrho) = -2\sqrt{\varrho(1-\varrho)} + o_\varrho(N)$ from (2.3). The monotonicity of $V(d)$ and $m(d)$ then implies that for all $\varrho_0 N \leq d \leq \varrho N$ and all N large enough:

$$\frac{|E|}{V(d)} \geq 2 + \frac{\varepsilon}{2\sqrt{\varrho(1-\varrho)}}, \quad m(d) \leq m(\varrho_0 N) = \sqrt{\frac{1 + 1/(\varrho_0 N)}{1 - 1/(\varrho N)}} = 1 + \mathcal{O}_\varrho(N^{-1}).$$

Inserting these estimates into the recursion relation (2.16), the claimed inequality (2.17) follows.

We now control the recursion relation in the regime $1 \leq d \leq \varrho_0 N$. To this end, note that the definition of ϱ_0 implies that for any $d \leq \varrho_0 N$ and N large enough: $|E|/V(d) \geq 3+\varepsilon/(2\sqrt{\rho(1-\rho)})$. Using $\hat{\Gamma}_{\varrho_0 N}(\varrho_0 N+1; E) \leq 1$ one readily establishes $\hat{\Gamma}_{\varrho_0 N}(d; E) \leq \frac{1}{2}$ inductively as long as $m(d) \leq 2$. The monotonicity $m(d) \leq m(1) = \sqrt{2}(1 + \mathcal{O}(N^{-1}))$ implies that this is true for any $d \geq 1$ at sufficiently large N . The proof of the claimed exponential decay (2.14) is then completed using the trivial norm bound

$$\hat{\Gamma}_{\varrho_0 N}(0; E) = G_{\varrho_0 N}(\sigma_0, \sigma_0; E) \leq \left\| (T_{\varrho_0 N} - E)^{-1} \right\| = \text{dist}(E, \text{spec}(T_{\varrho_0 N}))^{-1} \leq \frac{1}{\varepsilon N}.$$

Let us finally consider the case of fixed integers K . Note that for any $K \geq 1$ we know by the above $\hat{\Gamma}_{\varrho_0 N}(K+1; E) \leq 1/2$. The recursion relation (2.16) then yields for any $1 \leq d \leq K$

$$\hat{\Gamma}_{\varrho_0 N}(d; E) \leq \frac{d_K}{\sqrt{N}}$$

with some constants $d_K = d_K(\rho)$. This completes the proof of the first item. For the second item we organize the summation into sums over spheres of radius greater or equal to $K+1$:

$$\begin{aligned} & \sum_{\sigma \notin B_K(\sigma_0)} G_{\varrho_0 N}(\sigma, \sigma_0; E)^2 \\ &= \prod_{d=0}^K \hat{\Gamma}_{\varrho_0 N}(d; E)^2 \left(\sum_{D=K+1}^{\varrho_0 N} \prod_{d=K+1}^D \hat{\Gamma}_{\varrho_0 N}(d; E)^2 + \sum_{D=\varrho_0 N} \prod_{d=K+1}^D \hat{\Gamma}_{\varrho_0 N}(d; E)^2 \right). \end{aligned}$$

The product in the prefactor is estimated by $C_K/(\varepsilon^2 N^{K+2})$ using the first item. The second product is dominated by 4^{K-D} such that the summation over $D \geq K+1$ is bounded by a geometric series. The last product is bounded by $4^{K-\varrho_0 N}$ such that the sum is bounded trivially by this exponential factor times $\varrho_0 N$. This completes the proof. \square

The decay established in Proposition 2.4 for fixed distance K to the center of the ball agrees in its dependence on N with the result of Proposition 2.2. Moreover, the rough decay estimate (2.14) is ‘qualitatively correct’ in the sense that we expect an estimate of the form

$$G_{\varrho_0 N}(\sigma, \sigma_0; E) \leq \frac{1}{\varepsilon N} \binom{N}{d(\sigma_0, \sigma)}^{-1/2} e^{-L(E, \varrho, d(\sigma_0, \sigma))N}$$

with some positive function $L(E, \varrho, d(\sigma_0, \sigma))$. However, it is clear from the proof of Proposition 2.4 that we did not attempt to derive a sharp bound for L as it requires a more elaborate analysis of the factors $\hat{\Gamma}_{\varrho_0 N}(d; E)$.

3. Delocalization Regime

3.1. Spectral concentration. The analysis of the low-energy spectrum in the paramagnetic phase is based on the Schur complement method [3, Theorem 5.10] for which we define the spectral projections for $\varepsilon \in (0, 1)$

$$Q_\varepsilon := \mathbb{1}_{(-\varepsilon N, \varepsilon N)}(T) \quad P_\varepsilon := 1 - Q_\varepsilon, \tag{3.1}$$

which separate eigenstates of T with energies at the center of its spectrum from the edges. Here and in the following, $\mathbb{1}(\cdot)$ stands for the indicator function. A Chernoff bound shows that the dimension of the range of P_ε is only an exponential fraction of the total dimension of the Hilbert space:

$$\dim P_\varepsilon = \sum_{|k - \frac{N}{2}| > \frac{\varepsilon N}{2}} \binom{N}{k} \leq 2^{N+1} e^{-\varepsilon^2 N/2}. \quad (3.2)$$

The exact asymptotics of $\dim P_\varepsilon$ is in fact well-known, $\ln \dim P_\varepsilon = (\gamma(\frac{1-\varepsilon}{2}) + o(1))N$, in terms of the binary entropy γ defined in (1.16).

The following spectral concentration bound expresses the exponential smallness of the projection of symmetric random multiplication operators to the above subspace. It will be our main working horse in the paramagnetic phase.

Proposition 3.1. *Let $\varepsilon > 0$ and $W(\sigma)$, $\sigma \in Q_N$, be independent and identically distributed random variables such that*

- i. *the mean is zero, $\mathbb{E}[W(\sigma)] = 0$,*
- ii. *the variance of $W(\sigma)$ is bounded by one, i.e. $\mathbb{E}[W(\sigma)^2] \leq 1$, and*
- iii. *W is bounded, i.e. $\|W\|_\infty \leq M_N$ with some $M_N < \infty$, and $M_N^2 N \dim P_\varepsilon / 2^N \leq 1$.*

Then there are (universal) constants $c, C \in (0, \infty)$ such for any $\lambda > 0$:

$$\mathbb{P} \left(\|P_\varepsilon W P_\varepsilon\| - \mathbb{E}[\|P_\varepsilon W P_\varepsilon\|] > \lambda \sqrt{\frac{\dim P_\varepsilon}{2^N}} \right) \leq C e^{-c\lambda^2}. \quad (3.3)$$

Moreover, we have the following bound:

$$\mathbb{E}[\|P_\varepsilon W P_\varepsilon\|] \leq C \sqrt{N} \sqrt{\frac{\dim P_\varepsilon}{2^N}}. \quad (3.4)$$

Proof. The first statement follows from Talagrand's concentration inequality [64] (see also [66, Thm. 2.1.13]) by considering $F : \mathbb{R}^{Q_N} \rightarrow \mathbb{R}$ given by $F(W) := \|P_\varepsilon W P_\varepsilon\|$. We need to show that F is Lipschitz continuous and convex. Convexity, i.e., $F(\alpha W + (1-\alpha)W') \leq \alpha F(W) + (1-\alpha)F(W')$ for all $\alpha \in [0, 1]$, is evident from the triangle inequality. To establish the Lipschitz continuity, let $W, W' \in \mathbb{R}^{Q_N}$ and $\psi \in P_\varepsilon \ell^2(Q_N)$ with $\|\psi\| = 1$ be such that $\|P_\varepsilon(W - W')P_\varepsilon\| = \langle \psi, (W - W')\psi \rangle$. Then, one has

$$\begin{aligned} |F(W) - F(W')| &\leq \langle \psi, (W - W')\psi \rangle = \sum_{\sigma} |\psi(\sigma)|^2 (W(\sigma) - W'(\sigma)) \\ &\leq \|W - W'\|_2 \|\psi\|_4^2 \leq \|W - W'\|_2 \|\psi\|_\infty \leq \max_{\sigma} \sqrt{\langle \delta_\sigma | P_\varepsilon \delta_\sigma \rangle} \|W - W'\|_2. \end{aligned}$$

The first estimate is the triangle inequality. The next two estimates are special cases of Hölder's inequality, in which we also use $\|\psi\| = 1$. The last estimate results from the Cauchy-Schwarz inequality applied to $\|\psi\|_\infty = \max_{\sigma} |\langle P_\varepsilon \delta_\sigma | \psi \rangle|$ and the fact that $\|P_\varepsilon \delta_\sigma\| = \sqrt{[\langle P_\varepsilon \delta_\sigma | \delta_\sigma \rangle]}$. Since by symmetry for any $\sigma \in Q_N$:

$$\langle \delta_\sigma | P_\varepsilon \delta_\sigma \rangle = \frac{\dim P_\varepsilon}{2^N}, \quad (3.5)$$

we conclude that F is Lipschitz with constant $2^{-N/2} \sqrt{\dim P_\varepsilon}$. This finishes the proof of (3.3).

The second statement is derived from the matrix Bernstein inequality [55,67]. For its application, we note that the matrix under consideration is a sum of independent random matrices,

$$P_\varepsilon W P_\varepsilon = \sum_{\sigma} S(\sigma), \quad \text{with } S(\sigma) := \frac{\dim P_\varepsilon}{2^N} W(\sigma) |\psi(\sigma)\rangle\langle\psi(\sigma)|,$$

where $|\psi(\sigma)\rangle\langle\psi(\sigma)|$ denotes the rank-one projection onto $\psi(\sigma) := \sqrt{\frac{2^N}{\dim P_\varepsilon}} P_\varepsilon \delta_\sigma$, which in view of (3.5) is a normalised vector. By assumption the matrices $S(\sigma)$ are centred, $\mathbb{E}[S(\sigma)] = 0$, and bounded

$$\|S(\sigma)\| \leq M_N \frac{\dim P_\varepsilon}{2^N} \leq \sqrt{\frac{\dim P_\varepsilon}{N 2^N}}.$$

The mean variance matrix of $P_\varepsilon W P_\varepsilon$ is

$$\sum_{\sigma} \mathbb{E} [S(\sigma)^2] = \left(\frac{\dim P_\varepsilon}{2^N}\right)^2 \sum_{\sigma} \mathbb{E} [W(\sigma)^2] |\psi(\sigma)\rangle\langle\psi(\sigma)| \leq \frac{\dim P_\varepsilon}{2^N} P_\varepsilon.$$

The last inequality follows from the assumption, $\mathbb{E} [W(\sigma)^2] \leq 1$, as well as the fact that (δ_σ) form an orthonormal basis. Consequently, [67, Thm. 6.6.1] together with the trivial bound, $\dim P_\varepsilon \leq 2^N$, on the dimension of the matrices implies

$$\mathbb{E} [\|P_\varepsilon W P_\varepsilon\|] \leq \left(\sqrt{2 \ln 2^{N+1}} + \frac{\ln 2^{N+1}}{3\sqrt{N}}\right) \sqrt{\frac{\dim P_\varepsilon}{2^N}},$$

which completes the proof. □

Alternatively to Talagrand’s concentration inequality, the concentration of measure part of the matrix Bernstein inequality [67, Thm. 6.6.1] would also have been sufficient for proving a slightly less sharp upper bound on the upper tail of the large-deviation probability (3.3).

As an application, we state the following straightforward corollary. Its assumptions are tailored to fit in particular the case of the REM.

Corollary 3.2. *Suppose that $W(\sigma)$, $\sigma \in Q_N$ are i.i.d. random variables which are*

- i. *mean zero with variance $w_N := \mathbb{E} [W(\sigma)^2] \leq N$ and obey a moment bound $\mathbb{E} [W(\sigma)^8] \leq c N^4$ for some $c < \infty$.*
- ii. *linearly bounded in the sense that there is some $c < \infty$ such that $\|W\|_\infty \leq c N$.*

Then, there is some $C \in (0, \infty)$ such that for any $\tau \in (0, 1)$ there are events $\Omega_{N,\tau}$ with

$$\mathbb{P}(\Omega_{N,\tau}) \geq 1 - e^{-N/C} \tag{3.6}$$

such that for all sufficiently large N and at $\varepsilon = N^{\frac{\tau-1}{2}}$:

$$\|P_\varepsilon W P_\varepsilon\| \leq C N e^{-N^\tau/4}, \tag{3.7}$$

$$\left\|P_\varepsilon (W^2 - w_N) P_\varepsilon\right\| \leq C N^{\frac{3}{2}} e^{-N^\tau/4}, \tag{3.8}$$

$$\|P_\varepsilon W^p P_\varepsilon\| \leq C N^{\frac{p}{2}} \text{ for all } p \in [1, 4]. \tag{3.9}$$

Proof. The proof of these inequalities follows by three applications of Proposition 3.1 with different W' always at the same $\lambda = \sqrt{N}$. We note that by (3.2) our choice $\varepsilon = N^{\frac{\tau-1}{2}}$ implies $\dim P_\varepsilon \leq 2^{N+1} e^{-N^\tau/2}$. This in turn yields for any polynomial M_N and N large enough $M_N^2 N \dim P_\varepsilon/2^N \leq 1$, which indeed checks one of the assumptions of Proposition 3.1. We then construct three events $\Omega_{N,\tau}^{(j)}$ with $j \in \{1, 2, 3\}$ each with probability $\mathbb{P}(\Omega_{N,\tau}^{(j)}) \geq 1 - 3^{-1} e^{-N/C}$ with some (universal) $C < \infty$ and all N large enough. Their intersection $\Omega_{N,\tau} := \Omega_{N,\tau}^{(1)} \cap \Omega_{N,\tau}^{(2)} \cap \Omega_{N,\tau}^{(3)}$ then defines the required events.

More specifically, for a proof of (3.7), we take $W'(\sigma) = W(\sigma)/\sqrt{N}$. The event $\Omega_{N,\tau}^{(1)}$ on which (3.7) holds then satisfies the required probability estimate.

The proof of (3.8) follows again from Proposition 3.1 with $W'(\sigma) = c^{-1/4} (W(\sigma)^2 - w_N)/N$ and the prefactor ensuring $\mathbb{E} [W'(\sigma)^2] \leq 1$. In this way, we construct $\Omega_{N,\tau}^{(2)}$.

By Jensen's inequality $\langle \psi, W^p \psi \rangle^{4/p} \leq \langle \psi, W^4 \psi \rangle$ for any $p \in [1, 4]$, it suffices to establish (3.9) for $p = 4$. We choose $W'(\sigma) = c^{-1/2} (W(\sigma)^4 - \mathbb{E} [W(\sigma)^4])/N^2$ to define $\Omega_{N,\tau}^{(3)}$. □

3.2. Proof of Theorem 1.3. We now use the estimates of the preceding subsection in our Schur's complement analysis for the proofs of Theorem 1.3 and 1.4. These results will actually follow from a slightly more general theorem on operators $H = \Gamma T + W$ of QREM-type. As a preparation and motivation of the following lemma, we collect some basic facts about these operators. The kinetic part of the block component $Q_\varepsilon H Q_\varepsilon = \Gamma T Q_\varepsilon + Q_\varepsilon W Q_\varepsilon$ is estimated by

$$\|T Q_\varepsilon\| \leq \varepsilon N, \tag{3.10}$$

which implies

$$-\|W\|_\infty - \Gamma \varepsilon N \leq \inf \text{spec } Q_\varepsilon H Q_\varepsilon. \tag{3.11}$$

For any $z \in \mathbb{C}$ with $\text{Re } z < \|W\|_\infty - \Gamma \varepsilon N$, the operator $Q_\varepsilon H Q_\varepsilon - z$ is hence invertible on $Q_\varepsilon \ell^2(Q_N)$ with inverse denoted by $R_\varepsilon(z) := (Q_\varepsilon H Q_\varepsilon - z Q_\varepsilon)^{-1}$. The latter features in Schur's complement formula for the resolvent of H projected onto the subspace $P_\varepsilon \ell^2(Q_N)$:

$$P_\varepsilon (H - z)^{-1} P_\varepsilon = (P_\varepsilon (H - z) P_\varepsilon - P_\varepsilon W Q_\varepsilon R_\varepsilon(z) Q_\varepsilon W P_\varepsilon)^{-1}. \tag{3.12}$$

Our main observation is that Schur's complement is approximated by an operator proportional to the identity.

Lemma 3.3. *Consider the operator $H := \Gamma T + W$ on $\ell^2(Q_N)$ with W satisfying the assumptions in Corollary 3.2 and let $\Omega_{N,\tau}$ with $\tau \in (0, 1)$ be the events constructed there. Then on $\Omega_{N,\tau}$ and at $\varepsilon = N^{\frac{\tau-1}{2}}$ for all N large enough:*

$$\left\| P_\varepsilon W R_\varepsilon(z) W P_\varepsilon + P_\varepsilon \frac{w_N}{z} \right\| \leq \max\{1, \Gamma\} \frac{C}{d^2} N^{\frac{\tau-1}{2}}, \quad R_\varepsilon(z) := (Q_\varepsilon H Q_\varepsilon - z Q_\varepsilon)^{-1}, \tag{3.13}$$

for all $z \in \mathbb{C}$ such that $\min\{|z|, \text{dist}(\text{spec } Q_\varepsilon H Q_\varepsilon, z)\} \geq d N$ with $d \in (0, 1]$.

Proof. We use the resolvent equation to write

$$\begin{aligned}
 P_\varepsilon \left(W R_\varepsilon(z) W + \frac{w_N}{z} \right) P_\varepsilon &= \frac{1}{z} P_\varepsilon (w_N - W Q_\varepsilon W + W R_\varepsilon(z) Q_\varepsilon H Q_\varepsilon W) P_\varepsilon \\
 &= \frac{1}{z} P_\varepsilon (w_N - W Q_\varepsilon W) P_\varepsilon + \frac{1}{z} P_\varepsilon (W R_\varepsilon(z) Q_\varepsilon H Q_\varepsilon W) P_\varepsilon,
 \end{aligned}
 \tag{3.14}$$

and estimate both terms in the second line separately. For the first expression we rewrite

$$P_\varepsilon (w_N - W Q_\varepsilon W) P_\varepsilon = P_\varepsilon (w_N - W^2) P_\varepsilon + P_\varepsilon W P_\varepsilon W P_\varepsilon.
 \tag{3.15}$$

According to (3.7) and (3.8), the norm of the two terms in the right side is negligible in comparison to $N^{\frac{\tau-1}{2}}$ for all N large enough. It hence remains to estimate the norm of the second term in the right side of (3.14). To do so, we split the terms as follows

$$\frac{1}{z} P_\varepsilon W R_\varepsilon(z) Q_\varepsilon H Q_\varepsilon W P_\varepsilon = \frac{1}{z} P_\varepsilon W R_\varepsilon(z) Q_\varepsilon \Gamma T Q_\varepsilon W P_\varepsilon + \frac{1}{z} P_\varepsilon W R_\varepsilon(z) Q_\varepsilon W Q_\varepsilon W P_\varepsilon$$

and use (3.10) together with $\|R_\varepsilon(z)\| \leq (dN)^{-1}$ (since $\text{dist}(\text{spec } Q_\varepsilon H Q_\varepsilon, z) \geq dN$) and $\|P_\varepsilon W\|^2 = \|P_\varepsilon W^2 P_\varepsilon\| \leq CN$ by (3.8). On $\Omega_{N,\tau}$ for all N large enough, we thus conclude:

$$|z|^{-1} \|P_\varepsilon W R_\varepsilon(z) \Gamma T Q_\varepsilon W P_\varepsilon\| \leq \frac{C}{d^2 N} \|T Q_\varepsilon\| \leq \frac{C}{d^2} N^{\frac{\tau-1}{2}}.
 \tag{3.16}$$

Similarly, we estimate

$$\begin{aligned}
 |z|^{-1} \|P_\varepsilon W R_\varepsilon(z) Q_\varepsilon W Q_\varepsilon W P_\varepsilon\| &\leq |z|^{-1} \|P_\varepsilon W\| \|R_\varepsilon(z)\| \|W Q_\varepsilon W P_\varepsilon\| \\
 &\leq \frac{C}{d^2 N^{3/2}} \sqrt{\|P_\varepsilon W Q_\varepsilon W^2 Q_\varepsilon W P_\varepsilon\|}.
 \end{aligned}
 \tag{3.17}$$

In order to estimate the norm in the right side with the help of (3.9), we rewrite

$$P_\varepsilon W Q_\varepsilon W^2 Q_\varepsilon W P_\varepsilon = P_\varepsilon W^4 P_\varepsilon - P_\varepsilon W^3 P_\varepsilon W P_\varepsilon - P_\varepsilon W P_\varepsilon W^3 P_\varepsilon + P_\varepsilon W P_\varepsilon W^2 P_\varepsilon W P_\varepsilon.
 \tag{3.18}$$

On $\Omega_{N,\tau}$ the norm of this operator is bounded by $C N^2$ for all N large enough by (3.9). This concludes the proof. \square

These preparations enable us to prove the following general result.

Theorem 3.4. *Consider the operator $H = \Gamma T + W$ on $\ell^2(\mathcal{Q}_N)$ with W satisfying the assumptions in Corollary 3.2 and let $\Omega_{N,\tau}$ with $\tau \in (0, 1)$ arbitrary be the events constructed there. Then on $\Omega_{N,\tau}$ and for all N large enough the eigenvalues of H below $-\|W\|_\infty - \eta N$ with $\eta > 0$ are found in the union of intervals of radius $\mathcal{O}_{\Gamma,\eta}(N^{\frac{\tau-1}{2}})$ centered at*

$$(2n - N)\Gamma + \frac{w_N}{(2n - N)\Gamma}
 \tag{3.19}$$

with $n \in \{m \in \mathbb{N}_0 \mid (2m - N)\Gamma < -\|W\|_\infty - \eta N\}$. Moreover, the ball centered at (3.19) contains exactly $\binom{N}{n}$ eigenvalues of H if $\Gamma > \eta + \|W\|_\infty/N$.

Proof. We write H using the block decomposition of $\ell^2(\mathcal{Q}_N)$ induced by P_ε and employ the Schur complement method. Since the \mathcal{Q}_ε block is lower bounded according to (3.11), all eigenvalues E of H strictly below $-\|W\|_\infty - \Gamma\varepsilon N$ can be read from the equation

$$0 \in \text{spec}(T_\varepsilon(E)) \quad \text{with} \quad T_\varepsilon(E) := P_\varepsilon \left(\Gamma T + \frac{N}{E} \right) - E + Y_\varepsilon(E),$$

$$Y_\varepsilon(E) := P_\varepsilon W P_\varepsilon - \left(P_\varepsilon \frac{N}{E} + P_\varepsilon W R_\varepsilon(E) W P_\varepsilon \right). \quad (3.20)$$

Lemma 3.3 combined with (3.11) and (3.7) implies that for any $\eta > 0$ at $\varepsilon = N^{(\tau-1)/2}$ and on the event $\Omega_{N,\tau}$ in Corollary 3.2

$$\sup_{E < -\|W\|_\infty - \eta N} \|Y_\varepsilon(E)\| \leq C \max\{1, \Gamma\} \eta^{-2} N^{\frac{\tau-1}{2}}, \quad (3.21)$$

for all N large enough. As a consequence of standard perturbation theory [13, Corollary 3.2.6] and using the explicit values (1.4) of the spectrum of T , within this energy region the solution of (3.20) are found within the union of intervals of radius at most $C \max\{1, \Gamma\} \eta^{-2} N^{(\tau-1)/2}$ from the solutions to the equation

$$(2n - N)\Gamma + \frac{w_N}{z} - z = 0$$

with integers $2n < N(\Gamma - \|W\|_\infty - \eta)/\Gamma$. This leads to

$$z = \frac{2n - N}{2} \Gamma - \sqrt{\frac{1}{4}(2n - N)^2 \Gamma^2 + w_N} = (2n - N)\Gamma + \frac{w_N}{(2n - N)\Gamma} + \mathcal{O}_\Gamma(N^{-1}),$$

which completes the proof of (3.19). The assertion concerning the range of the spectral projections on the small intervals around the above points follows from the monotonicity of $T_\varepsilon(E)$ and the fact that the eigenvalue $2n - N$ of T has multiplicity $\binom{N}{n}$. \square

Theorem 1.3 now immediately follows.

Proof of Theorem 1.3. On $\Omega_{N,\eta/2}^{\text{REM}}$ the REM’s extremal values are bounded by $\|U\|_\infty \leq N(\beta_c + \eta)$. Moreover, $\mathbb{E}[U(\sigma)^2] = N$ and $\mathbb{E}[U(\sigma)^8] = 105 N^4$, so that U satisfies all requirements on W in Corollary 3.2. The claim is thus a straightforward consequence of Theorem 3.4 with $W = U$. \square

3.3. Proof of Theorem 1.4. The proof of our second main result, Theorem 1.4, is based on delocalization properties of the eigenprojection of T , which will be derived using the semigroup properties of T . More generally, let $B \subset \mathcal{Q}_N$ be any subset of the Hamming cube and $T(B)$ the corresponding restriction, i.e, the operator with matrix elements $\langle \delta_\sigma | T(B) \delta_{\sigma'} \rangle := -\mathbb{1}_{d(\sigma,\sigma')=1} \mathbb{1}_B(\sigma) \mathbb{1}_B(\sigma')$. For Hamming balls $B_K(\sigma_0)$ the operator $T(B_K(\sigma_0))$ was studied in Section 2 and abbreviated there by T_K . Standard semigroup techniques may be used to obtain for any $V : B \rightarrow \mathbb{R}$ the bound

$$0 \leq \langle \delta_\sigma | e^{-\beta(T(B)+V)} \delta_{\sigma'} \rangle \leq e^{-\min \beta V} \langle \delta_\sigma | e^{-\beta T(B)} \delta_{\sigma'} \rangle. \quad (3.22)$$

for all $\beta \geq 0$ and $\sigma, \sigma' \in \mathcal{Q}_N$, cf. [49]. Since $-T(B)$ and $-T$ have nonnegative matrix elements and $\langle \delta_\sigma | (-T(B)) \delta_{\sigma'} \rangle \leq \langle \delta_\sigma | (-T) \delta_{\sigma'} \rangle$ for any σ, σ' , we also conclude

$$\langle \delta_\sigma | e^{-\beta T(B)} \delta_{\sigma'} \rangle \leq \langle \delta_\sigma | e^{-\beta T} \delta_{\sigma'} \rangle = (\cosh \beta)^N (\tanh \beta)^{d(\sigma,\sigma')}, \quad (3.23)$$

where the last equality is by an explicit calculation using the Hadamard transformation, i.e., the representation of T in terms of Pauli matrices.

Proposition 3.5. *Let $B \subset \mathcal{Q}_N$ and $V : B \rightarrow \mathbb{R}$ a potential with $V \geq -vN$ for some $0 \leq v < 1$. Then the eigenprojection $P_E := \mathbb{1}_{(-\infty, E)}(T(B) + V)$ onto eigenvalues $E \in [-N(1 + v), -vN]$ satisfies:*

$$\max_{\sigma} \langle \delta_{\sigma} | P_E \delta_{\sigma} \rangle \leq 2^{-N} \exp \left(N \gamma \left(\frac{1 + v(E)}{2} \right) \right) \tag{3.24}$$

with the binary entropy γ from (1.16) and $v(E) := \frac{E}{N} + v$. Moreover, for all normalised states $\psi \in \ell^2(B)$:

$$\|P_E \psi\|_{\infty}^2 \leq 2^{-N} \exp \left(N \gamma \left(\frac{1 + v(E)}{2} \right) \right). \tag{3.25}$$

Proof. The spectral theorem combined with an exponential Markov inequality implies for any $\beta \geq 0$: $\langle \delta_{\sigma} | \mathbb{1}_{(-\infty, E)}(T(B) + V) \delta_{\sigma} \rangle \leq e^{\beta E} \langle \delta_{\sigma} | e^{-\beta(T(B)+V)} \delta_{\sigma} \rangle \leq e^{\beta v(E)N} (\cosh \beta)^N$. An elementary optimization with respect to β concludes the proof (cf. [49]). \square

We are now ready to complete the proofs of the main results in the paramagnetic regime.

Proof of Theorem 1.4. We pick $\tau \in (0, 1)$ and $0 < \eta < (\Gamma - \beta_c)/4$ arbitrary and restrict our attention to the event $\Omega_{N, \tau}^{\text{per}} \cap \Omega_{N, \eta}^{\text{REM}}$ on which the assertions of Corollary 3.2 for $W = U$ and Theorem 1.3 are valid.

For a proof of the first assertion, we apply Schur’s complement formula to the ground state $\psi = \psi_1 + \psi_2$ of $H = \Gamma T + U$. We split ψ into $\psi_1 \in P_{\varepsilon} \ell^2(\mathcal{Q}_N)$ and $\psi_2 \in Q_{\varepsilon} \ell^2(\mathcal{Q}_N)$ such that:

$$\begin{aligned} (P_{\varepsilon} H P_{\varepsilon} - E - P_{\varepsilon} H R_{\varepsilon}(E) H P_{\varepsilon}) \psi_1 &= 0 \\ \psi_2 &= -R_{\varepsilon}(E) Q_{\varepsilon} H P_{\varepsilon} \psi_1, \end{aligned}$$

where $E = \inf \text{spec } H = -\Gamma N - \frac{1}{\Gamma} + \mathcal{O}_{\Gamma}(N^{\frac{\tau-1}{2}})$ is the ground-state energy according to Theorem 1.3 since $\Gamma N - \|U\|_{\infty} > \frac{1}{2}(\Gamma - \beta_c)N > \eta N$ on $\Omega_{N, \eta}^{\text{REM}}$ by the choice for η . Sticking to the notation (3.20), from the proof of Theorem 1.3 we conclude that the first equation can be rewritten in terms of

$$P_{\varepsilon} H P_{\varepsilon} - E - P_{\varepsilon} H R_{\varepsilon}(E) H P_{\varepsilon} = P_{\varepsilon} \Gamma T P_{\varepsilon} + (N E^{-1} - E) P_{\varepsilon} + Y_{\varepsilon}(E),$$

with $\|Y_{\varepsilon}(E)\| \leq \mathcal{O}_{\Gamma}(N^{\frac{\tau-1}{2}})$. Since T has an energy gap 2 above its unique ground state Φ_{\emptyset} (cf. (1.4)), we thus conclude

$$\|(\mathbb{1} - |\Phi_{\emptyset}\rangle\langle\Phi_{\emptyset}|)\psi_1\| \leq \mathcal{O}_{\Gamma} \left(N^{\frac{\tau-1}{2}} \right).$$

To further estimate the norm of $\psi_2 = -R_{\varepsilon}(E) Q_{\varepsilon} U \psi_1$, we recall that $\|R_{\varepsilon}(E)\| \leq \frac{C_{\Gamma}}{N}$ and $\|U \psi_1\|^2 \leq \|P_{\varepsilon} U^2 P_{\varepsilon}\| \leq \mathcal{O}(N)$ by Corollary 3.2. Hence, $\|\psi_2\|^2 \leq \mathcal{O}_{\Gamma} \left(\frac{1}{N} \right)$. We thus arrive at

$$\|\psi - \Phi_{\emptyset}\|^2 = \mathcal{O}_{\Gamma} \left(N^{\tau-1} \right). \tag{3.26}$$

For the second part, we recall the bound (1.7), and write $H = \Gamma(T + U/\Gamma)$. The claim now follows directly from Proposition 3.5. \square

4. Extreme Localization Regime

4.1. *Deep-hole geometry.* The proof of our main results in the spin-glass regime are based on the deep-hole geometry of the REM. They rest on the fact that the large extremal sites $\mathcal{L}_{\beta_c - \delta}$ of the REM, which were defined in (1.17), are well separated on \mathcal{Q}_N at least if $\delta \in (0, \beta_c)$ is not too large.

Definition 4.1. Let $\varepsilon, \delta > 0$ and $\alpha \in (0, \frac{1}{2})$. Then $U : \mathcal{Q}_N \rightarrow \mathbb{R}$ is said to satisfy:

1. a *local* $(\varepsilon, \delta, \alpha)$ -*deep hole scenario* on $B_{\alpha N}(\sigma)$ with $\sigma \in \mathcal{L}_{\beta_c - \delta}$ if:
 - (a) $|U(\sigma')| \leq \varepsilon N$ for all $\sigma' \in B_{\alpha N}(\sigma)$ with $\sigma' \neq \sigma$,
 - (b) $u(\sigma) := \frac{1}{N^2} \sum_{\sigma' \in S_1(\sigma)} |U(\sigma')| \leq N^{-1/4}$.
2. a *global* $(\varepsilon, \delta, \alpha)$ -*deep hole scenario* if:
 - (a) U satisfies a local $(\varepsilon, \delta, \alpha)$ -deep hole scenario on $B_{\alpha N}(\sigma)$ for all $\sigma \in \mathcal{L}_{\beta_c - \delta}$,
 - (b) $B_{\alpha N}(\sigma) \cap B_{\alpha N}(\sigma') = \emptyset$ for all pairs $\sigma, \sigma' \in \mathcal{L}_{\beta_c - \delta}$ with $\sigma \neq \sigma'$.

The probabilistic estimate for the occurrence of a global deep-hole scenario in the REM is the subject of the following lemma.

Lemma 4.2. Let $\varepsilon, \delta > 0$ and $\alpha \in (0, 1/2)$ be such that

$$2\gamma(3\alpha) + \delta(2\beta_c - \delta) < \varepsilon^2. \tag{4.1}$$

The event $\Omega_N(\varepsilon, \delta, \alpha) := \{U \text{ satisfies a global } (\varepsilon, \delta, \alpha) \text{ - deep hole scenario}\}$ occurs with probability exponentially close to one, i.e., there is some $c(\varepsilon, \delta, \alpha) > 0$ such that for all N sufficiently large:

$$\mathbb{P}(\Omega_N(\varepsilon, \delta, \alpha)) \geq 1 - e^{-c(\varepsilon, \delta, \alpha)N}. \tag{4.2}$$

Proof. We first bound the probability of the event

$$\widehat{\Omega}_N(\varepsilon, \delta, \alpha) := \{\exists \sigma \in \mathcal{L}_{\beta_c - \delta}, \sigma' \in B_{3\alpha N}(\sigma) \setminus \{\sigma\} \text{ s.t. } |U(\sigma')| > \varepsilon N\}.$$

On its complement, all $\sigma \in \mathcal{L}_{\beta_c - \delta}$ satisfy the first requirement in the local deep-hole definition on $B_{\alpha N}(\sigma) \subset B_{3\alpha N}(\sigma)$, and the balls of radius αN around the large deviation sites are disjoint., i.e., the second requirement in the global deep-hole definition is also checked. By a union bound and independence, we conclude:

$$\begin{aligned} \mathbb{P}(\widehat{\Omega}_N(\varepsilon, \delta, \alpha)) &\leq \sum_{\sigma \in \mathcal{Q}_N} \sum_{\sigma' \in B_{3\alpha N}(\sigma) \setminus \{\sigma\}} \mathbb{P}(U(\sigma) \leq -(\beta_c - \delta)N) \mathbb{P}(|U(\sigma')| \geq \varepsilon N) \\ &\leq 2^{N+1} |B_{3\alpha N}| e^{-(\beta_c - \delta)^2 N/2} e^{-\varepsilon^2 N/2} \leq e^{(\gamma(3\alpha) + \beta_c \delta - \frac{\delta^2 + \varepsilon^2}{2} + o(1))N}. \end{aligned}$$

The second line is a result of the usual Gaussian-tail estimates and the fact that the volume of a Hamming ball of radius $\alpha N < N/2$ is asymptotically given in terms of the binary entropy, $\ln |B_{\alpha N}| = N(\gamma(\alpha) + o(1))$ as $N \rightarrow \infty$. Using assumption (4.1), we see that the above probability is exponentially small in N .

The proof is concluded by showing that the event

$$\Omega_N^u := \left\{ \max_{\sigma \in \mathcal{Q}_N} u(\sigma) \leq N^{-1/4} \right\} \tag{4.3}$$

occurs with a probability, which is exponentially close to one, i.e.

$$\mathbb{P}(\exists \sigma \in \mathcal{Q}_N \text{ s.t. } u(\sigma) > N^{-1/4}) \leq 2^{2N} e^{-N^{3/2}/2}. \tag{4.4}$$

For a proof of this bound, we rewrite the moment-generating function of $u(\sigma)$ for any $t > 0$ in terms of a standard normal variable g :

$$\mathbb{E}[e^{tu(\sigma)}] = \mathbb{E}[e^{tN^{-3/2}|g|}]^N \leq 2^N \mathbb{E}[e^{tN^{-3/2}g}]^N = 2^N e^{t^2/(2N^2)}.$$

By an exponential Chebychev-Markov estimate with $t = N^{7/4}$, this then yields $\mathbb{P}(u(\sigma) > N^{-1/4}) \leq 2^N e^{-N^{3/2}/2}$, and hence the claim by a union bound using $|\mathcal{Q}_N| = 2^N$. \square

4.2. Rank-one analysis. If U satisfies a local $(\varepsilon, \delta, \alpha)$ -deep hole scenario on $B_{\alpha N}(\sigma)$ at some fixed $\sigma \in \mathcal{L}_{\beta_c - \delta}$, it is natural to consider the Hamiltonian $H_{\alpha N}(\sigma) = \Gamma T_{\alpha N} + U$ restricted to $\ell^2(B_{\alpha N}(\sigma))$, i.e.

$$\langle \delta_\tau | H_{\alpha N}(\sigma) \delta_{\tau'} \rangle = \langle \delta_\tau | H \delta_{\tau'} \rangle \mathbb{1}_{B_{\alpha N}(\sigma)}(\tau) \mathbb{1}_{B_{\alpha N}(\sigma)}(\tau'),$$

A spectral analysis of these self-adjoint matrices is facilitated by rank-one perturbation theory. Since δ_σ is a cyclic vector for $H_{\alpha N}(\sigma)$, the spectrum can read from zeros of the meromorphic function given by

$$\begin{aligned} \langle \delta_\sigma | (H_{\alpha N}(\sigma) - z)^{-1} \delta_\sigma \rangle^{-1} &= U(\sigma) - \Sigma(\sigma, z), \\ \Sigma(\sigma, z) &:= -\langle \delta_\sigma | (H'_{\alpha N}(\sigma) - z)^{-1} \delta_\sigma \rangle^{-1}, \end{aligned} \tag{4.5}$$

where $H'_{\alpha N}(\sigma)$ coincides with the matrix $H'_{\alpha N}(\sigma)$ when setting $U(\sigma) = 0$. Moreover, an ℓ^2 -normalized eigenvector φ_E corresponding to $E \in \text{spec } H_{\alpha N}(\sigma)$ is given in terms of the free resolvent, i.e.,

$$\varphi_E(\tau) = -U(\sigma) \varphi_E(\sigma) \langle \delta_\tau | (H'_{\alpha N}(\sigma) - E_\sigma)^{-1} \delta_\sigma \rangle, \tag{4.6}$$

for any $\tau \in B_{\alpha N}(\sigma)$, cf. [3, Theorem 5.3]. The deep-hole scenario then entails the following information about the low-energy part of the spectrum.

Lemma 4.3. *Suppose U satisfies a local $(\varepsilon, \delta, \alpha)$ -deep hole scenario on $B_{\alpha N}(\sigma)$ at some $\sigma \in \mathcal{L}_{\beta_c - \delta}$ with*

$$2\Gamma\sqrt{\alpha(1-\alpha)} + \varepsilon < \beta_c - 2\delta. \tag{4.7}$$

Then for all sufficiently large N , the spectrum $\text{spec}_{E_\delta} H_{\alpha N}(\sigma) := \text{spec } H_{\alpha N}(\sigma) \cap (-\infty, E_\delta)$ below $E_\delta := -N(\beta_c - \delta)$ consists only of one simple eigenvalue E_σ which satisfies

$$\begin{aligned} E_\sigma &= U(\sigma) + \frac{\Gamma^2 N}{E_\sigma} + \frac{\Gamma^2}{E_\sigma^2} \sum_{\sigma' \in S_1(\sigma)} U(\sigma') + \mathcal{O}_{\Gamma, \delta, \varepsilon}(N^{-5/4}) \\ &= U(\sigma) + \frac{\Gamma^2 N}{U(\sigma)} + \mathcal{O}_{\Gamma, \delta}(N^{-1/4}). \end{aligned} \tag{4.8}$$

The ℓ^2 -normalized eigenfunction ψ_σ corresponding to E_σ satisfies:

1. for any $K \in \mathbb{N}$ and for all $\sigma' \in S_K(\sigma)$

$$|\psi(\sigma')| = \mathcal{O}_{\Gamma, \delta, K}(N^{-K}), \quad \text{and} \quad \sum_{\sigma' \notin B_K(\sigma)} |\psi(\sigma')|^2 = \mathcal{O}_{\Gamma, \delta, K}(N^{-(K+1)}). \tag{4.9}$$

2. for any $\alpha' \in (0, \alpha]$ there are $C = C(\Gamma, \delta)$, $c = c(\alpha, \alpha') \in (0, \infty)$, such that

$$\sum_{\sigma' \notin B_{\alpha'N}(\sigma)} |\psi_\sigma(\sigma')|^2 \leq CN \exp(-Nc). \tag{4.10}$$

Proof. The deep-hole scenario together with (2.3) and (4.7) implies that for all sufficiently large N :

$$H'_{\alpha N}(\sigma) \geq \Gamma T_{\alpha N} - \varepsilon N \geq -(\beta_c - 2\delta)N > E_\delta. \tag{4.11}$$

By rank-one perturbation theory, there is exactly one zero of (4.5) and hence one simple eigenvalue E_σ of $H_{\alpha N}(\sigma)$ below $\inf \text{spec } H'_{\alpha N}(\sigma)$. A Rayleigh-Ritz bound

$$E_\sigma \leq \langle \delta_\sigma | H_{\alpha N}(\sigma) \delta_\sigma \rangle = U(\sigma) \leq E_\delta \tag{4.12}$$

provides a first, crude estimate on this eigenvalue. According to (4.6) the corresponding ℓ^2 -normalized eigenvector ψ_σ satisfies for all $\sigma' \in B_{\alpha N}(\sigma)$:

$$\begin{aligned} \psi_\sigma(\sigma') &= -U(\sigma) \psi_\sigma(\sigma) \langle \delta_{\sigma'} | (H'_{\alpha N}(\sigma) - E_\sigma)^{-1} \delta_\sigma \rangle \\ &\leq -U(\sigma) \langle \delta_{\sigma'} | (\Gamma T_{\alpha N} - (E_\sigma + \varepsilon N))^{-1} \delta_\sigma \rangle \\ &\leq -U(\sigma) \Gamma^{-1} \langle \delta_{\sigma'} | (T_{\alpha N} - (U(\sigma) + \varepsilon N)\Gamma^{-1})^{-1} \delta_\sigma \rangle. \end{aligned} \tag{4.13}$$

As in (4.11), these inequalities are consequence of the deep-hole scenario, the crude bound (4.12) combined with the positivity of the semigroup, cf. (3.22). The assertions (4.9) and (4.10) concerning the decay rates of the eigenfunction are now a straightforward consequence of Proposition 2.4. For its application, we note that the assumption (4.7) ensure that $\text{dist}(\Gamma^{-1} \text{spec } T_{\alpha N}, U(\sigma) + \varepsilon N) \geq \Gamma^{-1}(E_\delta - U(\sigma) + \delta)N \geq \frac{\delta}{\Gamma}N$. The first inequality in Proposition 2.4 then yields

$$|\psi_\sigma(\sigma')| \leq \frac{\beta_c - \delta}{\delta} \binom{N}{d(\sigma_0, \sigma')}^{-1/2} 2^{-\min\{d(\sigma, \sigma'), \rho_0(\alpha)N\}}, \tag{4.14}$$

where we also used that the function $U \mapsto \frac{-U}{x-U}$ is monotone increasing in U on $(-\infty, x)$. Hence, (4.10) follows after a summation over the spheres $S_d(\sigma)$ with $d \in (\alpha'N, \alpha N]$. The above binomial decay factor is thereby exactly compensated by the volume $|S_d(\sigma)| = \binom{N}{d}$. The claimed bounds (4.9) follow analogously from the respective bounds in Proposition 2.4.

For a proof of the asymptotics (4.8), we first consider the eigenvalue equation at any $\sigma' \in S_1(\sigma)$:

$$\begin{aligned} E_\sigma \psi_\sigma(\sigma') &= U(\sigma') \psi_\sigma(\sigma') - \Gamma \psi_\sigma(\sigma) - \Gamma \sum_{\sigma'' \in S_1(\sigma') \setminus \{\sigma\}} \psi_\sigma(\sigma'') \\ &= U(\sigma') \psi_\sigma(\sigma') - \Gamma \psi_\sigma(\sigma) + \mathcal{O}_{\Gamma, \delta}(N^{-1}). \end{aligned} \tag{4.15}$$

The uniform $\mathcal{O}_{\Gamma, \delta}(N^{-1})$ estimate is a direct consequence of (4.9). This equation can be rewritten as

$$\psi_\sigma(\sigma') = -\frac{\Gamma}{E_\sigma - U(\sigma')} \left(\psi_\sigma(\sigma) + \mathcal{O}_{\Gamma, \delta}(N^{-1}) \right), \tag{4.16}$$

which we insert into the eigenvalue equation at σ :

$$\begin{aligned}
 E_\sigma \psi_\sigma(\sigma) &= U(\sigma) \psi_\sigma(\sigma) - \Gamma \sum_{\sigma' \in S_1(\sigma)} \psi_\sigma(\sigma') \\
 &= U(\sigma) \psi_\sigma(\sigma) + \frac{\Gamma^2}{E_\sigma} \left(\sum_{\sigma' \in S_1(\sigma)} \frac{\psi_\sigma(\sigma) + \mathcal{O}_{\Gamma, \delta}(N^{-1})}{1 - U(\sigma')/E_\sigma} \right) \\
 &= \left[U(\sigma) + \frac{\Gamma^2 N}{E_\sigma} + \frac{\Gamma^2}{E_\sigma} \left(\sum_{\sigma' \in S_1(\sigma)} \frac{U(\sigma')}{E_\sigma} \right) \right] \psi_\sigma(\sigma) + \mathcal{O}_{\Gamma, \delta, \varepsilon}(N^{-5/4}).
 \end{aligned}
 \tag{4.17}$$

The third equality follow from a second-order Taylor expansion with an error estimate using $|U(\sigma')|^2 \leq \varepsilon N |U(\sigma')|$ as well as the bound on $u(\sigma)$ in the deep-hole assumption in Definition 4.1. Since $\psi_\sigma(\sigma) = 1 + \mathcal{O}(N^{-1})$, the first identity in (4.8) follows. For a proof of the second identity, we again use the bound on $u(\sigma)$ as well as our crude estimate (4.12) to estimate the last term in the above square brackets by $\mathcal{O}_{\Gamma, \delta}(N^{-1/4})$. This concludes the proof. \square

4.3. Spectral averaging. In order to control the probability of resonances between distinct extremal sites, we will use the spectral averaging technique from the theory of random operators [3, Chapter 4.1].

Lemma 4.4. *Let $\varepsilon, \delta > 0$ and $\alpha \in (0, 1/2)$ be such that (4.1) and (4.7) holds. Then, there is some $c = c(\varepsilon, \delta, \alpha) > 0$ such that for all N sufficiently large and*

1. *for any real interval I :*

$$\mathbb{P}(\exists \sigma \in \mathcal{L}_{\beta_c - \delta} \text{ s.t. } \text{spec}_{E_\delta} H_{\alpha N}(\sigma) \cap I \neq \emptyset) \leq 2|I| e^{\beta_c \delta N - \delta^2 N/2} + e^{-cN}.
 \tag{4.18}$$

2. *for any $r > 0$:*

$$\begin{aligned}
 &\mathbb{P}(\exists \sigma, \sigma' \in \mathcal{L}_{\beta_c - \delta}, \sigma \neq \sigma' \text{ s.t. } \text{dist}(\text{spec}_{E_\delta} H_{\alpha N}(\sigma), \text{spec}_{E_\delta} H_{\alpha N}(\sigma')) \leq r) \\
 &\leq 4r e^{(2\beta_c \delta - \delta^2)N} + e^{-cN}.
 \end{aligned}
 \tag{4.19}$$

Proof. For a proof of the above estimates, we may thus restrict attention to events in $\Omega_N(\varepsilon, \delta, \alpha)$, cf. Lemma 4.2.

1. According to Lemma 4.3, under the deep-hole scenario $\text{spec}_{E_\delta} H_{\alpha N}(\sigma) \cap I \neq \emptyset$ if and only if $E_\sigma = \inf \text{spec } H_{\alpha N}(\sigma) \in I$. Since $\psi_\sigma(\sigma)^2 \geq 1/2$ by Lemma 4.3 for sufficiently large N and all $\sigma \in \mathcal{L}_{\beta_c - \delta}$, the latter implies $\langle \delta_\sigma | P_I \delta_\sigma \rangle \geq 1/2$, where P_I denotes the spectral projection of $H_{\alpha N}(\sigma)$ onto I . A union bound hence enables to estimate the probability of the event in the left side of (4.18) and its intersection with $\Omega_N(\varepsilon, \delta, \alpha)$ by

$$\sum_{\sigma \in \mathcal{Q}_N} \mathbb{P}(\sigma \in \mathcal{L}_{\beta_c - \delta} \text{ and } \langle \delta_\sigma | P_I \delta_\sigma \rangle \geq 1/2) \leq 2 \mathbb{E}[\mathbb{1}[\sigma \in \mathcal{L}_{\beta_c - \delta}] \langle \delta_\sigma | P_I \delta_\sigma \rangle].$$

The inequality is a Chebychev-Markov estimate. Conditioning on all random variables aside from $U(\sigma)$, the integration of $p_I(U(\sigma)) := \langle \delta_\sigma | P_I \delta_\sigma \rangle$ with respect to the random variable $U(\sigma)$ is bounded with the help of the spectral averaging lemma (also referred to as Wegner estimate, cf. [3, Thm. 4.1]).

- On $\Omega_N(\varepsilon, \delta, \alpha)$, we may assume that $B_{\alpha N}(\sigma) \cap B_{\alpha N}(\sigma') = \emptyset$ for all pairs $\sigma, \sigma' \in \mathcal{L}_{\beta_c - \delta}$. This ensures that the random variables $E_{\sigma'} = \inf \text{spec } H_{\alpha N}(\sigma')$ and $U(\sigma')$ are independent of all random variables in $B_{\alpha N}(\sigma)$. Using the strategy as in 1., we thus bound the probability of the event in the left side of (4.19) and its intersection with $\Omega_N(\varepsilon, \delta, \alpha)$ by

$$\sum_{\sigma, \sigma' \in \mathcal{Q}_N} \mathbb{E} [\mathbb{1}[\sigma' \in \mathcal{L}_{\beta_c - \delta} \text{ and } B_{\alpha N}(\sigma) \cap B_{\alpha N}(\sigma') = \emptyset] \times \mathbb{P} \left(\sigma \in \mathcal{L}_{\beta_c - \delta} \text{ and } \langle \delta_\sigma | P_{(E_{\sigma'} - r, E_{\sigma'} + r)} \delta_\sigma \rangle \geq 1/2 \mid B_{\alpha N}(\sigma)^c \right)] \leq 2^{2N+2} e^{-(\beta_c - \delta)^2 N} r.$$

where $\mathbb{P}(\cdot | B_{\alpha N}(\sigma)^c)$ denotes the conditional expectation, conditioned on all random variables aside from those in $B_{\alpha N}(\sigma)$ and P_I is still the spectral projection of $H_{\alpha N}(\sigma)$ onto I . The last inequality resulted from an application of the bound from 1. to the conditional expectation. \square

4.4. *Proof of Theorem 1.5.* The proof of Theorem 1.5 makes use of the deep-hole geometry of the REM. If U satisfies a global $(\varepsilon, \delta, \alpha)$ -deep hole scenario, we study the auxiliary Hamiltonian

$$H' := \left(\bigoplus_{\sigma \in \mathcal{L}_{\beta_c - \delta}} H_{\alpha N}(\sigma) \right) \bigoplus H_r, \tag{4.20}$$

with operators $H_{\alpha N}(\sigma)$, whose action is restricted to the non-intersecting balls $B_{\alpha N}(\sigma)$ around extremal sites $\sigma \in \mathcal{L}_{\beta_c - \delta}$. These operator have been introduced and studied in Sect. 4.2. The remainder H_r is that part of H which purely belongs to the complement of the union of balls,

$$\langle \delta_\tau | H_r \delta_{\tau'} \rangle = \langle \delta_\tau | H \delta_{\tau'} \rangle \left(1 - \sum_{\sigma \in \mathcal{L}_{\beta_c - \delta}} \mathbb{1}_{B_{\alpha N}(\sigma)}(\tau) \right) \left(1 - \sum_{\sigma \in \mathcal{L}_{\beta_c - \delta}} \mathbb{1}_{B_{\alpha N}(\sigma)}(\tau') \right).$$

The difference between the Hamiltonian of interest $H = \Gamma T + U$ and the auxiliary H' is

$$H - H' =: -\Gamma A =: -\Gamma \bigoplus_{\sigma \in \mathcal{L}_{\beta_c - \delta}} A_\sigma.$$

It describes the hopping between the balls and the complementary configuration space, i.e.,

$$\langle \delta_\tau | A_\sigma \delta_{\tau'} \rangle = \mathbb{1}_{d(\tau, \tau')=1} (\mathbb{1}_{d(\tau, \sigma)=\alpha N} \mathbb{1}_{d(\tau', \sigma)=\alpha N+1} + \mathbb{1}_{d(\tau, \sigma)=\alpha N+1} \mathbb{1}_{d(\tau', \sigma)=\alpha N}).$$

The norm of A can be bounded as follows

$$\|A\| = \max_{\sigma \in \mathcal{L}_{\beta_c - \delta}} \|A_\sigma\| \leq \|T_{\alpha N+1}\| = 2N\sqrt{\alpha(\alpha - 1)} + o_\alpha(N), \tag{4.21}$$

where the last equality is (2.2). It is easy to see that $\|A\|$ is indeed of order N . However, for energies below $E_\delta = -N(\beta_c - \delta)$, the perturbation is of a much smaller magnitude. This is the basic idea in the proofs of our main results for the localization regime. As a preparation, we also need the following result, which is implicitly contained in [52].

Proposition 4.5 (cf. [52]). *For all $\Gamma, \delta > 0$ the truncated Hamiltonian $H := \Gamma T + U \mathbb{1}_{U \geq -(\beta_c - \delta)N}$ acting on $\ell^2(\mathcal{Q}_N)$ is lower bounded by*

$$\inf \text{spec } H \geq -N \max\{\Gamma, \beta_c - \delta\} + o_{\Gamma, \delta}(N)$$

except for an event of exponentially small probability.

Proof of Theorem 1.5. We only study the joint event $\Omega_N(\Gamma, \delta, \alpha)$ on which i) the bound in Proposition 4.5 applies, and ii) U satisfies a global $(\varepsilon, \delta, \alpha)$ -deep hole scenario with parameters

$$\varepsilon = \frac{\beta_c}{2} \quad \text{and} \quad \delta \in (0, \min\{\beta_c - \Gamma, \beta_c/8\}),$$

and $\alpha > 0$ small enough such that (4.1) and $2\Gamma\sqrt{\alpha(1-\alpha)} < \delta/8$, and hence in particular (4.7) is satisfied. Together with Lemma 4.2 this ensures that $\Omega_N(\Gamma, \delta, \alpha)$ occurs with a probability of at least $1 - e^{-cN}$ with at some $c \equiv c(\Gamma, \delta, \alpha) > 0$. Moreover:

1. From Lemma 4.3 we learn that for any $\sigma \in \mathcal{L}_{\beta_c - \delta}$ the spectrum $\text{spec } H_{\alpha N}(\sigma)$ below $E_\delta = -N(\beta_c - \delta)$ consists of just one eigenvalue $E_\sigma = \inf \text{spec } H_{\alpha N}(\sigma)$, which is given by (4.8) with an error term $\mathcal{O}_{\Gamma, \delta}(N^{-1/4})$ uniformly for all $\sigma \in \mathcal{L}_{\beta_c - \delta}$.
2. By the variational principle and the natural embedding of Hilbert spaces, the ground state energy of H_r is bounded from below by that of $\Gamma T + U \mathbb{1}_{U \geq -(\beta_c - \delta)N}$ on $\ell^2(\mathcal{Q}_N)$. The lower bound in Proposition 4.5 then shows that

$$\inf \text{spec } H_r \geq -N(\beta_c - \delta + o_{\Gamma, \delta}(1)).$$

Hence, H_r does not contribute to the low-energy spectrum of H' below $E_{\delta/2} = -N(\beta_c - \delta/2)$ for all N large enough. Moreover, the spectral projection $P_\delta := \mathbb{1}_{(-\infty, E_{\delta/2})}(H')$ can be written as

$$P_\delta = \sum_{\sigma \in \mathcal{L}_{\beta_c - \delta}, E_\sigma < E_{\delta/2}} |\psi_\sigma\rangle\langle\psi_\sigma| \tag{4.22}$$

in terms of rank-one projections of the ℓ^2 -normalized ground states ψ_σ of $H_{\alpha N}(\sigma)$. We thus conclude for some $C = C(\Gamma, \delta) < \infty$, and $c = c(\alpha) > 0$

$$\|AP_\delta\| = \max_{\sigma \in \mathcal{L}_{\beta_c - \delta}} \|A\sigma\psi_\sigma\| \leq \|A\| \max_{\sigma \in \mathcal{L}_{\beta_c - \delta}} \left(\sum_{\sigma' \in S_{\alpha N}(\sigma)} |\psi_\sigma(\sigma')|^2 \right)^{1/2} \leq C N^2 e^{-cN}, \tag{4.23}$$

where the inequalities follow from (4.21) and (4.10) together with the fact that $A\sigma$ only acts on the part of ψ_σ on $S_{\alpha N}(\sigma)$.

We then rewrite H using the block decomposition of $\ell^2(\mathcal{Q}_N)$ induced by P_δ and $Q_\delta := 1 - P_\delta$ and again employ the Schur complement method. Since H' is diagonal in this decomposition and its Q_δ projection has a spectrum above the threshold energy $E_{\delta/2}$, it remains to investigate the blocks of the perturbation ΓA :

1. Since P_δ is supported entirely on the balls, the first diagonal term vanishes, i.e. $P_\delta A P_\delta = 0$. The operator norms of the off-diagonals $\|P_\delta A(1 - P_\delta)\| \leq \|A P_\delta\|$ are exponentially small by (4.23).
2. The operator $Q_\delta A Q_\delta$ is bounded from below by $-\|A\|$ which is estimated in (4.21). We thus conclude that for all N large enough:

$$\begin{aligned} Q_\delta H' + Q_\delta A Q_\delta &\geq E_{\delta/2} - \|A\| \geq -N \left(\beta_c - \delta/2 + 2\Gamma \sqrt{\alpha(1 - \alpha)} + o_{\Gamma, \alpha}(1) \right) \\ &\geq -N(\beta_c - \delta/4). \end{aligned}$$

Consequently, the Schur complement matrix

$$S_\delta(E) := (Q_\delta H' + Q_\delta A Q_\delta - E)^{-1}$$

is well defined on $Q_\delta \ell^2(Q_N)$ and bounded, $\|S_\delta(E)\| \leq (E_{\delta/4} - E)^{-1}$ for any $E < E_{\delta/4}$.

The spectrum of H below $E_{\delta/4} = -N(\beta_c - \delta/4)$ is thus characterized using Schur’s method, which yields:

1. $E < E_{\delta/4}$ is an eigenvalue of H if and only if $E \in \text{spec}(P_\delta H' - P_\delta A S_\delta(E) A P_\delta)$.
2. The ℓ^2 -normalized eigenvector ψ corresponding to E and H satisfies:

$$\begin{aligned} (P_\delta H' - E P_\delta)\psi &= P_\delta A S_\delta(E) A P_\delta \psi \\ Q_\delta \psi &= -S_\delta(E) A P_\delta \psi. \end{aligned} \tag{4.24}$$

We now proceed with the completion of the proof of the assertion on the spectrum and eigenvectors separately.

Spectrum: The spectrum of H below $E_{\delta/8}$ is determined through the above Schur complement method. Since for all $E \leq E_{\delta/8}$ at at some $C = C(\Gamma, \delta) < \infty$ and $c = c(\alpha) > 0$

$$\|P_\delta A S_\delta(E) A P_\delta\| \leq \|S_\delta(E)\| \|A P_\delta\|^2 \leq C N^3 e^{-2cN}, \tag{4.25}$$

the eigenvalues below $E_{\delta/8}$ thus coincide with the eigenvalues of $P_\delta H'$ below this energy up to an error, which is exponentially small in N [13, Corollary 3.2.6]. Since the eigenvalues of $P_\delta H'$ are given by (4.8), the assertion in Theorem 1.5 follows.

Eigenvectors: We concentrate our attention on energies below $E_s = -N(\beta_c - s)$ with $s \in (0, \delta/8]$ small enough such that $2\beta_c s < c$ with the decay rate $c > 0$ from (4.25). This ensures that $e^{-\alpha c N} < e^{-2\beta_c s N} =: r(s)$ for all sufficiently large N . According to the spectral averaging Lemma 4.4, since $s \leq \delta/8$ and the condition (4.7) is monotone in δ , the event

$$\{\forall \sigma, \sigma' \in \mathcal{L}_{\beta_c - s}, \sigma \neq \sigma' : \text{dist}(\text{spec}_{E_s} H_{\alpha N}(\sigma), \text{spec}_{E_s} H_{\alpha N}(\sigma')) > r(s)\} \tag{4.26}$$

has probability of at least $1 - 4e^{-s^2 N} - e^{-cN}$ for some $c > 0$. We may therefore assume its occurrence.

Perturbation theory based on the above Schur complement analysis and (4.25) (combined with the characterization of eigenvalues established in Theorem 1.5) then guarantees that the eigenvector ψ of H corresponding to the eigenvalue $E = U(\sigma) +$

$\Gamma^2 N/U(\sigma) + \mathcal{O}(N^{-1/4})$, which is uniquely characterized by $\sigma \in \mathcal{L}_{\beta_c - s}$, is norm-close to the ground-state eigenvector ψ_σ of $H_{\alpha N}(\sigma)$, i.e.,

$$\begin{aligned} \|\psi - \psi_\sigma\| &\leq \|P_\delta \psi - \psi_\sigma\| + \|Q_\delta \psi\| \\ &\leq \frac{\|P_\delta A S_\delta(E) A P_\delta\|}{r(s)} + \|S_\delta(E)\| \|A P_\delta\| \leq C e^{-cN}. \end{aligned} \tag{4.27}$$

Here, the inequalities combine (4.24)–(4.26). The rest of the claim on the ℓ^2 -estimates of the eigenvectors then follows from the respective properties of ψ_σ established in Lemma 4.3. The event $\Omega_{N,\Gamma,\delta}^{\text{loc}}$ is then defined by specifying a value for $\alpha_0 = \alpha_0(\Gamma, \delta)$ and intersecting $\Omega_N(\Gamma, \delta, \alpha_0)$ with (4.26). \square

4.5. *Proof of Theorem 1.7.* All assertions concerning the ℓ^2 -properties of the ground-state can easily be collected from the proof of Theorem 1.5.

Proof of Theorem 1.7 (ℓ^2 -properties). According to Theorem 1.5 for all events aside from one of exponentially small probability, there is some $\sigma_0 \in \mathcal{Q}_N$ such that the ground state eigenvector is approximated by $\|\psi - \delta_{\sigma_0}\|_{\ell^2}^2 = \mathcal{O}_\Gamma\left(\frac{1}{N}\right)$. The estimate $\mathcal{O}_\Gamma\left(\frac{1}{N}\right)$ does not depend on δ anymore as we may fix δ , if we only consider the ground state. This will be always assumed in the following. Moreover, the ground-state energy is $E = U(\sigma_0) + \frac{\Gamma^2 N}{U(\sigma_0)} + \mathcal{O}_\Gamma(N^{-1/4})$, where $U(\sigma_0)$ is one of the REM’s extremal energies for which we may assume that

$$|U(\sigma_0) + \beta_c N| \leq \mathcal{O}(\sqrt{N}), \quad \text{and hence} \quad |E + \beta_c N| \leq \mathcal{O}(\sqrt{N}) \tag{4.28}$$

at the expense of excluding another event of exponentially small probability stemming from deviations to the known extremal statistics of the REM, cf. (1.7).

It thus remains to establish the assertion on the first order perturbation $\xi \in \ell^2(\mathcal{Q}_N)$. That $(\xi | H \xi)$ agrees with the ground state energy up to order $\mathcal{O}_\Gamma(1)$ is a result of a simple calculation and a comparison with the above formula for E . It remains to prove $\|\psi - \xi\|^2 = \mathcal{O}_\Gamma(N^{-2})$. To this end, we revisit the proof of Theorem 1.5. From the validity of the global $(\beta_c/2, \delta, \alpha)$ -deep hole scenario specified there and in view of (4.9), it suffices to show

$$\left| \psi(\sigma_0) - \sqrt{1 - \frac{\Gamma^2}{\beta_c^2 N}} \right|^2 = \mathcal{O}_\Gamma(N^{-2}) \quad \text{and} \quad \sum_{\sigma \in S_1(\sigma_0)} \left| \psi(\sigma) - \frac{\Gamma}{\beta_c N} \right|^2 = \mathcal{O}_\Gamma(N^{-2}). \tag{4.29}$$

For a proof of these assertions, we use the eigenvalue equation (4.16) on $S_1(\sigma_0)$ together with $\psi(\sigma_0) = 1 + \mathcal{O}_\Gamma(N^{-1})$. If we pick $\sigma \in S_1(\sigma_0)$, this yields

$$\begin{aligned} \psi(\sigma) - \frac{\Gamma}{\beta_c N} &= -\frac{\Gamma}{E - U(\sigma)} \left(1 + \mathcal{O}_\Gamma(N^{-1})\right) - \frac{\Gamma}{\beta_c N} \\ &= \frac{\Gamma U(\sigma)}{\beta_c N(E - U(\sigma))} + \mathcal{O}_\Gamma(N^{-3/2}). \end{aligned}$$

Here the last step also relied on the estimate $|U(\sigma)| \leq \varepsilon N$ valid in the $(\varepsilon, \delta, \alpha)$ -deep hole scenario, as well as (4.28). With a suitable constant $C = C(\Gamma) < \infty$, we then have

$$\sum_{\sigma \in S_1(\sigma_0)} \left| \psi(\sigma) - \frac{\Gamma}{\beta_c N} \right|^2 \leq \frac{C}{N^4} \sum_{\sigma \in S_1(\sigma_0)} U(\sigma)^2 + \mathcal{O}_\Gamma(N^{-2})$$

An exponential Chebychev-Markov estimate leads to $\mathbb{P}(N^{-2} \sum_{\sigma \in S_1(\sigma_0)} (U(\sigma)^2 - N)) \geq 1) \leq e^{-cN}$ for some $c > 0$. Thus, except for an event of exponentially small probability the second claim in (4.29) holds. Since ψ is ℓ^2 -normalized, this leads to

$$\psi(\sigma_0)^2 = 1 - \sum_{\sigma \in S_1(\sigma_0)} \psi(\sigma)^2 + \mathcal{O}_\Gamma(N^{-2}) = 1 - \frac{\Gamma^2}{\beta_c^2 N} + \mathcal{O}_\Gamma(N^{-2}),$$

which readily implies the first claim in (4.29).

For a proof of the ℓ^1 -estimate on the ground state eigenfunction, we need to sharpen estimates on the large-deviation geometry of the REM. To this end we define for $\varepsilon, \delta > 0$ the following tripartition of the Hamming cube:

$$\begin{aligned} A_1(\varepsilon) &:= \{\sigma \in \mathcal{Q}_N \mid |U(\sigma)| \leq \varepsilon N\} \\ A_2(\varepsilon, \delta) &:= \{\sigma \in \mathcal{Q}_N \mid \varepsilon N < |U(\sigma)| \leq (\beta_c - \delta)N\} \\ A_3(\delta) &:= \{\sigma \in \mathcal{Q}_N \mid |U(\sigma)| > (\beta_c - \delta)N\}. \end{aligned}$$

A modification of ideas used in the proof of Lemma 4.2 and [51, Lemma 2] yields:

Lemma 4.6. *For any $\varepsilon > 0$ there exist $K = K(\varepsilon) \in \mathbb{N}$ and a family of events $\Omega_{\varepsilon, N}$ such that for N large enough*

- (i) For any $\sigma \in A_2(\varepsilon, \delta) \cup A_3(\delta)$: $|B_4(\sigma) \cap (A_2(\varepsilon, \delta) \cup A_3(\delta))| \leq K$ on $\Omega_{\varepsilon, N}$.
- (ii) $\mathbb{P}(\Omega_{\varepsilon, N}) \geq 1 - 2^{-N}$.

Proof. Let $\Omega_{\varepsilon, N, K}$ be the event, where the assertion (i) holds true with constant K . It remains to show that the complement satisfies $\mathbb{P}(\Omega_{\varepsilon, N, K}^c) \leq 2^{-N}$ for an appropriate choice for K and N large enough. To this end we estimate

$$\begin{aligned} \mathbb{P}(\Omega_{\varepsilon, N, K}^c) &= \mathbb{P}(\exists \sigma \in A_2(\varepsilon, \delta) \cup A_3(\delta) \text{ s.t. } |B_4(\sigma) \cap (A_2(\varepsilon, \delta) \cup A_3(\delta))| \geq K) \\ &\leq \sum_{\sigma_0 \in \mathcal{Q}_N} \mathbb{P}(|U(\sigma_0)| \geq \varepsilon N) \mathbb{P}(\exists K - 1 \text{ different } \sigma_1, \dots, \sigma_{K-1} \in B_4(\sigma_0) \setminus \{\sigma_0\} \text{ s.t.} \\ &\hspace{15em} |U(\sigma_j)| \geq \varepsilon N \text{ for } j = 1, \dots, K - 1) \\ &\leq \binom{N^4}{K - 1} 2^N \mathbb{P}(|U(\sigma_0)| \geq \varepsilon N)^K \leq N^{4K} 2^N e^{-KN\varepsilon^2/2}. \end{aligned}$$

Here the second line is a consequence of the union bound and the third line follows from the independence and a simple counting argument. Choosing $K > 4 \ln 2/\varepsilon^2$, we see that $\mathbb{P}(\Omega_{\varepsilon, N, K}^c) < 2^{-N}$ for N large enough. \square

As a final preparation, we also need the following elementary observation on the size of large deviation sites.

Lemma 4.7. *For any $\delta \in (0, \beta_c)$ and all N :*

$$\mathbb{P}\left(|A_3(\delta)| \geq 2e^{\beta_c \delta N}\right) \leq e^{-N\delta^2/2}.$$

Proof. The cardinality $|A_3(\delta)|$ is a sum of 2^N independent Bernoulli variables with success probability $p = \mathbb{P}(|U(\sigma)| > (\beta_c - \delta)N) \leq 2e^{-\frac{1}{2}(\beta_c - \delta)^2 N}$ such that $\mathbb{E}[|A_3(\delta)|] = 2^N p$. The claim thus follows from a standard Markov estimate. \square

We are finally ready to finish the proof of our main result in the localization regime.

Proof of Theorem 1.7 (ℓ^p -properties). We first observe that the claims on the ℓ^p -norms immediately follow from the ℓ^1 -norm asymptotics (1.25). To see this, recall that $\psi(\sigma_0) = 1 + o_\Gamma(1)$ for some $\sigma_0 \in \mathcal{Q}_N$, and that $\psi(\sigma) \leq cN^{-1}$ for all $\sigma \neq \sigma_0$. Hence for any $1 < p < \infty$:

$$1 + o_\Gamma(1) \leq \|\psi\|_{\ell^p}^p \leq 1 + \frac{c^{p-1}}{N^{p-1}} \sum_{\sigma \neq \sigma_0} \psi(\sigma) \leq 1 + \frac{c^{p-1}}{N^{p-1}} \|\psi\|_{\ell^1} = 1 + o_{\Gamma,p}(1).$$

It therefore remains to establish (1.25).

Recalling that the ground state wavefunction ψ is positive, we can write $\|\psi\|_{\ell^1} = \sum_{\sigma} \psi(\sigma)$. The eigenvalue equation for ψ leads to

$$\begin{aligned} E \sum_{\sigma} \psi(\sigma) &= +\Gamma \sum_{\sigma} (T\psi)(\sigma) + \sum_{\sigma} (U\psi)(\sigma) = -\Gamma N \sum_{\sigma} \psi(\sigma) - \sum_{\sigma} U(\sigma)\psi(\sigma) \\ &= -\Gamma N \sum_{\sigma} \psi(\sigma) + U(\sigma_0)\psi(\sigma_0) + \sum_{\sigma \neq \sigma_0} U(\sigma)\psi(\sigma). \end{aligned} \tag{4.30}$$

The second equality follows from the fact that each σ has N neighbors. The main idea is now to show that the remainder term $\sum_{\sigma \neq \sigma_0} U(\sigma)\psi(\sigma)$ can be controlled by the other two terms on the right side. Here, we use the tripartition $A_1(\varepsilon), A_2(\varepsilon, \delta), A_3(\delta)$ of the configuration space and bound the contribution of each A_i separately.

In the following, we fix $\delta, \alpha > 0$ small enough, such that the REM satisfies a global $(\beta_c/2, \delta, \alpha)$ -deep hole scenario with a probability which is exponentially close to one. Moreover, we pick $\varepsilon > 0$ arbitrary and fix $K = K(\varepsilon) \in \mathbb{N}$ the assertions of Lemma 4.6 hold on a joint event on which the global $(\beta_c/2, \delta_0, \alpha_0)$ -deep hole scenario applies as well. This event still has a probability of at least $1 - e^{-c(\delta, \alpha)N}$ with some $c(\delta, \alpha) > 0$, which is independent of ε .

Contribution of $A_1(\varepsilon)$: In this case we use the trivial estimate, $|\sum_{\sigma \in A_1} U(\sigma)\psi(\sigma)| \leq \varepsilon N \|\psi\|_{\ell^1}$.

Contribution of $A_3(\delta)$: We only consider $\delta \leq \delta_0$, such that sites $\sigma \in A_3(\delta) \setminus \{\sigma_0\}$ lie outside the ball $B_{\alpha_0 N}(\sigma_0)$. In particular, there is some $c > 0$ such that for all N large enough and all $\sigma \in A_3(\delta) \setminus \{\sigma_0\}$ the ground state is uniformly bounded, $|\psi(\sigma)| \leq e^{-cN}$. We now pick $\delta := \min\{\delta_0, c/(4\beta_c)\}$ and shrink the considered event such that $|A_3(\delta)| \leq 2e^{Nc/4}$. According to Lemma 4.7 this event still has a probability greater than $1 - e^{-c(\delta, \alpha)N}$ with some $c(\delta, \alpha) > 0$. On this event, we conclude for all N large enough

$$\sum_{\sigma \in A_3(\delta) \setminus \{\sigma_0\}} |U(\sigma)|\psi(\sigma) \leq e^{-Nc/2}.$$

Contribution of $A_2(\varepsilon, \delta)$: We first consider the configurations in $A_2(\varepsilon, \delta)$ close to the center σ_0 , which we estimate for N large by

$$\sum_{\sigma \in A_2(\varepsilon, \delta) \cap B_4(\sigma_0)} |U(\sigma)|\psi(\sigma) \leq |A_2(\varepsilon, \delta) \cap B_4(\sigma_0)| \max_{\sigma \in B_4(\sigma_0) \setminus \{\sigma_0\}} |U(\sigma)|\psi(\sigma) \leq CK$$

with some $C = C(\Gamma)$. We use $|U(\sigma)| \leq \beta_c N/2$ due to the validity of the global $(\beta_c/2, \delta, \alpha)$ -deep hole scenario as well as the pointwise bound $\psi(\sigma) \leq CN^{-1}$ for all $\sigma \in B_4(\sigma_0)$ with $\sigma \neq \sigma_0$.

It remains to consider $\sigma \in A_2(\varepsilon, \delta) \setminus B_4(\sigma_0)$. The eigenvalue equation reads

$$(E - U(\sigma))\psi(\sigma) = \Gamma \sum_{\sigma' \in S_1(\sigma)} \psi(\sigma').$$

Since $E \leq (\beta_c - \delta/2)N$ for N large enough, we obtain for $\sigma \in A_2(\varepsilon, \delta)$ the bound

$$\psi(\sigma) \leq \frac{2\Gamma}{\delta N} \sum_{\sigma' \in S_1(\sigma)} \psi(\sigma').$$

The essence of the following argument is that the value of any $\psi(\sigma)$ is comparable to the mean on the corresponding $S_1(\sigma)$ sphere and, thus, ψ cannot take especially large values on $A_2(\varepsilon, \delta)$. To make this intuition precise, we separate the $A_3(\delta)$ configurations, which we possibly encounter in the spherical mean and repeat the procedure for the remaining $\sigma' \in S_1(\sigma)$. This leads to

$$\begin{aligned} \psi(\sigma) &\leq \frac{2\Gamma}{\delta N} \sum_{\sigma' \in S_1(\sigma) \cap A_3(\delta)} \psi(\sigma') + \frac{4\Gamma^2}{\delta^2 N^2} \sum_{\sigma' \in S_1(\sigma) \setminus A_3(\delta)} \sum_{\sigma'' \in S_1(\sigma')} \psi(\sigma'') \\ &\leq \frac{2\Gamma}{\delta N} \sum_{\sigma' \in S_1(\sigma) \cap A_3(\delta)} \psi(\sigma') + \frac{4\Gamma^2}{\delta^2 N} \psi(\sigma) + \frac{8\Gamma^2}{\delta^2 N^2} \sum_{\sigma' \in S_2(\sigma)} \psi(\sigma'), \end{aligned}$$

which for N large enough implies

$$\psi(\sigma) \leq \frac{4\Gamma}{\delta N} \sum_{\sigma' \in S_1(\sigma) \cap A_3(\delta)} \psi(\sigma') + \frac{16\Gamma^2}{\delta^2 N^2} \sum_{\sigma' \in S_2(\sigma)} \psi(\sigma').$$

We now further shrink the considered event to ensure that $\|U\|_\infty \leq 2\beta_c N$ holds true. This happens for all but an event of exponentially probability, cf. (1.6). Thus, for N large enough

$$\begin{aligned} \left| \sum_{\sigma \in A_2 \setminus B_4(\sigma_0)} U(\sigma)\psi(\sigma) \right| &\leq 2\beta_c N \sum_{\sigma \in A_2 \setminus B_4(\sigma_0)} \psi(\sigma) \\ &\leq \sum_{\sigma \in A_2 \setminus B_4(\sigma_0)} \left(\frac{8\beta_c \Gamma}{\delta} \sum_{\sigma' \in S_1(\sigma) \cap A_3(\delta)} \psi(\sigma') + \frac{16\beta_c \Gamma^2}{\delta^2 N} \sum_{\sigma' \in S_2(\sigma)} \psi(\sigma') \right) \\ &\leq \frac{8\beta_c K \Gamma}{\delta} \sum_{\sigma \in A_3(\delta) \setminus \{\sigma_0\}} \psi(\sigma) + \frac{16\beta_c \Gamma^2 K}{\delta^2 N} \|\psi\|_{\ell^1} \leq e^{-Nc/4} + \frac{16\beta_c \Gamma^2 K}{\delta^2 N} \|\psi\|_{\ell^1}. \end{aligned}$$

In the third line we used the observation that each configuration $\sigma \in \mathcal{Q}_N$ appears in the summation at most K times due to Lemma 4.6. The last step is a consequence of our exponential bound on $\psi(\sigma)$ on the $A_3(\delta)$ -configurations.

Combining the partial results on each A_J , we arrive with some $C = C(\Gamma)$ at the bound

$$\left| \sum_{\sigma \neq \sigma_0} U(\sigma)\psi(\sigma) \right| \leq (2\varepsilon + \mathcal{O}_{K,\Gamma}(N^{-1}))N\|\psi\|_{\ell^1} + 4CK$$

which is valid on with probability of at least $1 - e^{-cN}$ with some $c > 0$ which in independent of ε . Since $\varepsilon > 0$ was arbitrary, the claimed convergence now follows from (4.30).

5. Free Energy Asymptotics

For our proof of Theorem 1.10 we exploit that the partition function is determined by the eigenvalues close to the thermal averages

$$\langle U \rangle_\beta^{\text{cl}} := \frac{\text{Tr } U e^{-\beta U}}{\text{Tr } e^{-\beta U}} \quad \text{or} \quad \langle T \rangle_\beta^{\text{pm}} := \frac{\text{Tr } T e^{-\beta T}}{\text{Tr } e^{-\beta T}}, \tag{5.1}$$

depending on the phase. To determine their behavior we consider the local region around $\sigma \in \mathcal{L}_\varepsilon$, where $\varepsilon > 0$ has to be allowed to be arbitrarily small. In this case, we cannot guarantee anymore that all balls $B_R(\sigma)$ are disjoint. However, we will show that this is still true for isolated extremal sites $\sigma \in \mathcal{L}_\varepsilon$, which are in the majority. Then, we establish the order-one corrections of Theorem 1.5 for those isolated large deviations. Based on these results, we prove Theorem 1.10 via a suitable approximation argument using auxiliary operators on cut domains of the configuration space.

5.1. Basic large deviations. We first record some standard facts in the statistical mechanics of the pure REM and pure paramagnet.

Proposition 5.1. *1. For any $\beta \geq 0$ we have $\langle T \rangle_\beta^{\text{pm}} = -N \tanh \beta$. Moreover, for any $\beta \geq 0, \delta > 0$ there exists some $c = c(\beta, \delta) > 0$ such that*

$$\frac{\text{Tr } \mathbb{1}_{[-N(\tanh \beta + \delta), -N(\tanh \beta - \delta)]}(T) e^{-\beta T}}{\text{Tr } e^{-\beta T}} \geq 1 - e^{-cN}. \tag{5.2}$$

2. For $\beta < \beta_c$ we have almost surely $\langle U \rangle_\beta^{\text{cl}} = -(\beta + o(1))N$. Moreover, for $\beta < \beta_c$ and $\delta > 0$ there exists some $c = c(\beta, \delta) > 0$ such that

$$\frac{\text{Tr } \mathbb{1}_{[-N(\beta + \delta), -N(\beta - \delta)]}(U) e^{-\beta U}}{\text{Tr } e^{-\beta U}} \geq 1 - e^{-cN} \tag{5.3}$$

except for an exponentially small event.

3. For $\beta > \beta_c$ we have almost surely $\langle U \rangle_\beta^{\text{cl}} = -(\beta_c + o(1))N$. Moreover, for $\beta > \beta_c$ and $\delta > 0$ there exists some $c = c(\delta) > 0$ such that

$$\frac{\text{Tr } \mathbb{1}_{(-\infty, -N(\beta_c - \delta)]}(U) e^{-\beta U}}{\text{Tr } e^{-\beta U}} \geq 1 - e^{-cN}. \tag{5.4}$$

The proof for the expressions of the thermal averages $\langle T \rangle_\beta^{\text{pm}}$, $\langle U \rangle_\beta^{\text{cl}}$ is by differentiating the explicit formulas for the pressure with respect to β . The results on the concentration of the Gibbs measure are then of Cramér type and follow from the usual convexity estimates of the (explicit) free energy.

5.2. *Spectral analysis on clusters.* In the proofs of Theorem 1.5 and 1.7 we derived the order-one correction of the energy levels $U(\sigma)$ caused by extremal sites $\sigma \in \mathcal{L}_\varepsilon$ with $\varepsilon \approx \beta_c$ from a local analysis on non-overlapping balls $B_R(\sigma)$ of some radius R . For the proof of Theorem 1.10 however, we need good control on all eigenvalues with energy below $-\varepsilon N$ with $\varepsilon > 0$ arbitrary and, thus, the large deviation set \mathcal{L}_ε has to be considered for any $\varepsilon > 0$. The balls $B_R(\sigma)$ then have a nonempty intersection and the aim of this subsection is to deal with this modified situation. Let us introduce some definitions and notation.

Definition 5.2. Let $\varepsilon > 0$ and $k \in \mathbb{N}_0$. We denote $\sigma \overset{k}{\sim} \sigma' \iff d(\sigma, \sigma') \leq 2k + 2$. We call a set $G \subset \mathcal{L}_\varepsilon$ (k, ε) -connected (with respect to $\overset{k}{\sim}$) if for any $\sigma, \sigma' \in G$ there exists a sequence $\sigma = \sigma^{(0)}, \sigma^{(1)}, \dots, \sigma^{(m)} = \sigma'$ such that $\sigma^{(i)} \in G$ and $\sigma^{(i)} \overset{k}{\sim} \sigma^{(i+1)}$ for all $0 \leq i \leq m - 1$. If $G \subset \mathcal{L}_\varepsilon$ is (k, ε) -connected and for any (k, ε) -connected G' with $G \subset G' \subset \mathcal{L}_\varepsilon$ it follows $G = G'$, we call G a (k, ε) -component. We denote the family of (k, ε) -components of \mathcal{L}_ε by $\mathcal{G}_{k,\varepsilon}$.

We call $\sigma \in \mathcal{Q}_N(k, \varepsilon)$ -isolated if $G = \{\sigma\} \in \mathcal{G}_{k,\varepsilon}$ and $I_{k,\varepsilon}$ denotes the collection of (k, ε) -isolated configurations.

The case $k = 0$ coincides with the notion of 'gap-connected' used in [50–52].

The extremal set \mathcal{L}_ε naturally decomposes in its components, i.e., $\mathcal{L}_\varepsilon = \cup_{G \in \mathcal{G}_{k,\varepsilon}} G$. We define for each (k, ε) -component G the corresponding cluster

$$C_k(G) := \bigcup_{\sigma \in G} B_k(\sigma).$$

By construction $d(C_k(G), C_k(G')) \geq 2$ for different k -components $G \neq G'$.

We start with a combinatorial lemma which shows that the size of (k, ε) -components remains bounded and that most (k, ε) -components are isolated.

Lemma 5.3. Let $\varepsilon > 0$ and $k \in \mathbb{N}_0$ be fixed, but arbitrary.

1. There exists an $M = M(k, \varepsilon) \in \mathbb{N}$ such that

$$\Omega_{N,M}(\varepsilon, k) := \left\{ \max_{G \in \mathcal{G}_{k,\varepsilon}} |G| \leq M \right\}. \tag{5.5}$$

occurs with probability $\mathbb{P}(\Omega_{N,M}(k, \varepsilon)) \geq 1 - e^{-cN}$ for some $c > 0$.

2. Let $\varepsilon < a < b$ and $b < \beta_c$. Then for all events, but one of exponentially small probability:

$$\frac{|\mathcal{L}_{a,b} \cap I_{k,\varepsilon}^c|}{|\mathcal{L}_{a,b}|} \leq e^{-\varepsilon^2 N/4}, \tag{5.6}$$

where $\mathcal{L}_{a,b} := \mathcal{L}_a \cap \mathcal{L}_b^c$ and $(\cdot)^c$ indicates the complement of that set.

3. Let $a = \beta_c - \delta$ and suppose that $\delta(2\beta_c - \delta) < \varepsilon^2$. Then, besides of an exponentially small event

$$\mathcal{L}_a \cap I_{k,\varepsilon}^c = \emptyset. \tag{5.7}$$

Proof. For a proof of the first assertion, we estimate for any $M \in \mathbb{N}$ using a union bound

$$\mathbb{P}\left(\max_{G \in \mathcal{G}_{k,\varepsilon}} |G| \geq M\right) \leq 2^N \binom{|B_{(2k+2)M}|}{M} e^{-\frac{1}{2}\varepsilon^2 MN}.$$

As the binomial coefficient is a polynomial in N the claim follows for $M > 2 \ln 2/\varepsilon^2$.

For a proof of the second assertion, we rewrite $|\mathcal{L}_{a,b}| = \sum_{\sigma} Z_{\sigma}$, where Z_{σ} are i.i.d. Bernoulli variables with success probability

$$p_N := \mathbb{P}(\sigma \in \mathcal{L}_{a,b}) \geq \frac{\sqrt{N}(b-a)}{\sqrt{2\pi}} e^{-Nb^2/2}. \tag{5.8}$$

Since $b < \beta_c$, the average size $\mathbb{E}[|\mathcal{L}_{a,b}|] = 2^N p_N$ is exponentially large and, by a Markov estimate, the same applies to all events aside from one of super-exponentially small probability, i.e., $\mathbb{P}(|\mathcal{L}_{a,b}| \leq 2^{N-1} p_N) \leq e^{-e^{CN}}$ for some $C > 0$. Similarly, the conditional probability $\mathbb{P}_{\sigma}(\cdot | \{\sigma\}^c)$ of the configuration to not be (ε, k) -isolated equals the probability to find on $B_{2k+2}^o := B_{2k+2}(\sigma) \setminus \{\sigma\}$ another large deviation in $\mathcal{L}_{\varepsilon}$ and hence $\mathbb{P}_{\sigma}(\exists \sigma' \in I_{k,\varepsilon}^c \cap B_{2k+2}^o) \leq |B_{2k+2}^o| e^{-N\varepsilon^2/2} \leq N^{2k+2} e^{-N\varepsilon^2/2}$ by the union bound and the Gaussian-tail estimate. This allows us to estimate

$$\mathbb{E}[|\mathcal{L}_{a,b} \cap I_{k,\varepsilon}^c|] = \sum_{\sigma \in \mathcal{Q}_N} \mathbb{E}[\mathbb{1}[\sigma \in \mathcal{L}_{a,b}] \mathbb{P}_{\sigma}(\sigma \in I_{k,\varepsilon}^c)] \leq 2^N p_N N^{2k+2} e^{-N\varepsilon^2/2} \tag{5.9}$$

with p_N from (5.8). Excluding the event on which $|\mathcal{L}_{a,b}| \leq 2^{N-1} p_N$, we thus arrive at

$$\mathbb{P}\left(|\mathcal{L}_{a,b} \cap I_{k,\varepsilon}^c| \geq e^{-\varepsilon^2 N/4} |\mathcal{L}_{a,b}|\right) \leq \frac{e^{\varepsilon^2 N/4}}{2^{N-1} p_N} \mathbb{E}[|\mathcal{L}_{a,b} \cap I_{k,\varepsilon}^c|] + e^{-e^{CN}}$$

by a Chebychev-Markov estimate. Inserting the bound (5.9) completes the proof.

For the last assertion, we note that by Lemma 4.2 the condition on δ implies that for $\alpha > 0$ small enough a global $(\varepsilon, \delta, \alpha)$ -deep hole scenario occurs with probability exponentially close to one. \square

The next lemma establishes the spectral properties of the restriction $H_{C_k(G)}$ of the QREM Hamiltonian to the Hilbert space $\ell^2(C_k(G))$ of a cluster corresponding to $G \in \mathcal{G}_{k,\varepsilon}$. For its formulation, we define for $\delta > 0$ the spectral projections

$$P_{\delta}(G) := \mathbb{1}_{(-\infty, -\delta N)}(H_{C_k(G)}), \quad Q_{\delta}(G) := \mathbb{1} - P_{\delta}(G).$$

Recall the events Ω_N^u defined in (4.3) and $\Omega_{N,M}(k, \varepsilon)$ defined in (5.5).

Lemma 5.4. *Let $\varepsilon > 0$ and $k \geq 2$. On the event $\Omega_N^u \cap \Omega_{N,M}(\varepsilon, k)$ the following assertions are valid for all N large enough:*

1. $\max_{G \in \mathcal{G}_{k,\varepsilon}} \|H_{C_k(G)} - U_{C_k(G)}\| = \mathcal{O}_{\Gamma,k,M}(\sqrt{N})$.

2. If ψ is an ℓ^2 -normalized eigenfunction of $H_{C_k(G)}$ with $\langle \psi, H_{C_k(G)}\psi \rangle \leq -\frac{3}{2}\varepsilon N$, we have

$$|\psi(\sigma)| = \mathcal{O}_{\Gamma,k,M,\varepsilon}(N^{-\text{dist}(G,\sigma)}), \quad (5.10)$$

and

$$\|\mathbb{1}_{C_k(G)\setminus G}\psi\|^2 = \mathcal{O}_{\Gamma,k,M,\varepsilon}(N^{-1}), \quad \|\mathbb{1}_{\partial C_k(G)}\psi\|^2 = \mathcal{O}_{\Gamma,k,M,\varepsilon}(N^{-k}), \quad (5.11)$$

where $\mathbb{1}_{\partial C_k(G)}$ is the natural projection onto the boundary of $C_k(G)$. In particular, all estimates are independent of ψ and G .

3. $\sup_{G \in \mathcal{G}_{k,\varepsilon}} \sup_{\sigma \in \mathcal{L}_{2\varepsilon} \cap G} \langle \delta_\sigma | Q_{3\varepsilon/2}(G) \delta_\sigma \rangle = \mathcal{O}_{\Gamma,k,M,\varepsilon}(N^{-1})$.
4. If $G = \{\sigma_0\}$ is (k, ε) -isolated and $U(\sigma_0) \leq -2\varepsilon N$, then the ground state energy of $H_{C_k(G)}$ is given by

$$E_{\sigma_0} := \inf \text{spec } H_{C_k(G)} = U(\sigma_0) + \frac{\Gamma^2 N}{U(\sigma_0)} + \mathcal{O}_{\Gamma,k,\varepsilon}(N^{-1/4}). \quad (5.12)$$

Proof. 1. We write $H_{C_k(G)} = U_{C_k(G)} + \Gamma T_{C_k(G)}$ and recall that $C_k(G)$ is a union of at most M Hamming balls $B_k(\sigma)$ with $\sigma \in G$. Thus, by the triangle inequality and Proposition 2.1 we obtain $\|T_{C_k(G)}\| \leq M c_k \sqrt{N}$, and hence the claim.

2. We introduce the modified spheres $S_r(G)$ for $0 \leq r \leq k$,

$$S_r(G) := C_r(G) \setminus C_{r-1}(G) = \{\sigma \in C_k(G) \mid \text{dist}(\sigma, G) = r\}$$

and for the eigenvector ψ the maximal values on the spheres, $s_r := \max_{\sigma \in S_r(G)} |\psi(\sigma)|$. We use the convention $S_0(G) = G$ and note that $S_k(G) = \partial C_k(G)$. Moreover, we observe that for any $\sigma \in S_r(G)$ and $1 \leq r \leq k$:

$$\begin{aligned} |S_1(\sigma) \cap S_r(G)| &\leq M, & r \leq |S_1(\sigma) \cap S_{r-1}(G)| &\leq rM, \\ N - (r+1)M &\leq |S_1(\sigma) \cap S_{r+1}(G)| &\leq N - r. \end{aligned}$$

We now use the eigenvalue equation

$$-E\psi(\sigma) = \Gamma \sum_{\sigma' \in S_1(\sigma)} \psi(\sigma') - U(\sigma)\psi(\sigma),$$

to derive the claimed decay estimate. Inserting the above geometric bounds into the eigenvalue equation, we obtain for all $1 \leq r \leq k$ with the convention $s_{k+1} = 0$:

$$-Es_r \leq \Gamma r M s_{r-1} + \Gamma(N-r)s_{r+1} + (\varepsilon N + \Gamma M)s_r. \quad (5.13)$$

We claim that for all $1 \leq r \leq k$ and N large enough

$$s_r \leq \frac{2M\Gamma k}{|E| - \varepsilon N - \Gamma M} s_{r-1}.$$

This is immediate from (5.13) in case $r = k$ (even without the factor 2). In case $1 \leq r < k$, the bound is proven recursively. If the inequality holds for $r+1$, then (5.13) implies

$$|E|s_r \leq \Gamma r M s_{r-1} + \left(\frac{2N\Gamma^2 k}{|E| - \varepsilon N - \Gamma M} + \varepsilon N + \Gamma M \right) s_r,$$

and hence the claimed inequality for all N large enough. Since $s_0 \leq 1$, this establishes (5.10) by iteration.

The first claim in (5.11) follows from (5.10) using $|G| \leq M$. Indeed, for some $C = C(\Gamma, k, M, \varepsilon)$

$$\sum_{\sigma \neq \sigma_0} |\psi(\sigma)|^2 \leq C \sum_{r=1}^K N^{-2r} |\{\sigma \mid \text{dist}(\sigma, G) = r\}| \leq CM \sum_{r=1}^K N^{-r} = \frac{CM}{N-1}.$$

Since $|\partial C_k(G)| \leq MN^k$ the second inequality in (5.11) follows similarly.

3. We will repeatedly make use of a coupling principle which follows from 1., namely the fact that the eigenvalues of $H_{C_k(G)}$ and $U_{C_k(G)}$ agree up to a uniform error of order $\mathcal{O}_{\Gamma, k, M}(N^{1/2})$. Since $|\mathcal{L}_\varepsilon \cap G| \leq |G| \leq M$, this implies that $\dim P_{3\varepsilon/2}(G) \leq M$ for any component $G \in \mathcal{G}_{k, \varepsilon}$ if N is chosen large enough. By the pigeon-hole principle, for any component G we find some $a = a(G) \in [3\varepsilon/2, 2\varepsilon]$ such that

$$P_{a-\varepsilon/(2M)}(G) - P_{a+\varepsilon/(2M)}(G) = \mathbb{1}_{(-(a+\varepsilon/(2M))N, -(a-\varepsilon/(2M))N)}(H_{C_k(G)}) = 0. \tag{5.14}$$

Since $Q_{3\varepsilon/2}(G) \leq Q_a(G)$, it is enough to prove the assertion with $Q_{3\varepsilon/2}(G)$ replaced by $Q_a(G)$.

To this end, we fix a component $G \in \mathcal{G}_{k, \varepsilon}$ and observe that the coupling principle and (5.14) yield

$$|\mathcal{L}_a \cap G| = \dim P_a(G) =: m_a \tag{5.15}$$

with a natural number $m_a \leq M$. We denote by $\psi_1, \dots, \psi_{m_a}$ the normalized low energy eigenfunctions of $H_{C_k(G)}$ corresponding to $P_a(G)$, which form an orthonormal basis for this subspace. The first inequality in (5.11) bounds the contribution of each eigenfunction to $C_k(G) \setminus G$. Moreover, the eigenvalue equation readily implies for $\sigma \in G \setminus \mathcal{L}_a$:

$$|\psi_j(\sigma)| \leq \frac{\Gamma}{|E_j - U(\sigma)|} \sum_{\sigma' \in S_1(\sigma)} |\psi_j(\sigma')| \leq \frac{2M\Gamma}{\varepsilon N} \sum_{\sigma' \in S_1(\sigma)} |\psi_j(\sigma')| \leq \frac{2M\Gamma}{\varepsilon \sqrt{N}},$$

with $E_j = \langle \psi_j, H_{C_k(G)} \psi_j \rangle \leq -aN$ the eigenvalue corresponding to ψ_j . The second inequality follows from (5.14) and the last step is a consequence of the Cauchy-Schwarz inequality and $\|\psi_j\| = 1$. As $|G \setminus \mathcal{L}_a| \leq M$, we also conclude that

$$\sum_{\sigma \in C_k(G) \setminus \mathcal{L}_a} |\psi_j(\sigma)|^2 \leq \frac{C}{N}$$

with a uniform $C = C(\Gamma, k, M, \varepsilon) < \infty$. We thus learn that $\sup_j \|\mathbb{1}_{\mathcal{L}_a \cap G} \psi_j - \psi_j\|^2 \leq C/N$. Lemma 5.5 below shows that (with $P = \mathbb{1}_{\mathcal{L}_a \cap G}$ and $F = P_a(G)$)

$$\sup_{\sigma \in \mathcal{L}_a \cap G} \langle \delta\sigma | Q_a(G) \delta\sigma \rangle \leq \|Q_a(G) \mathbb{1}_{\mathcal{L}_a \cap G}\| \leq 4(\sqrt{m_a} + m_a)^2 \frac{C}{N} \leq 16M^2 \frac{C}{N}.$$

Since $a \leq 2\varepsilon$, this proves the claim.

4. By the Rayleigh-Ritz variational principle, we have $E_{\sigma_0} \leq -2\varepsilon N$, and hence the results of 2. apply to the corresponding ground state wavefunction $\psi \in \ell^2(C_k(G))$. By (5.11) this ensures $\psi(\sigma_0) = 1 + O_{\Gamma,k,\varepsilon}(N^{-1/2})$. Following the steps in analysis (4.15)–(4.17) of the eigenfunction equation, in which we use (5.10) and the assumed bound on u , we thus conclude that (4.8) remains valid. This concludes the proof of (5.12). \square

In the proof of Lemma 5.4 we used the following result on finite-rank projections:

Lemma 5.5. *Suppose \mathcal{H} is a finite-dimensional Hilbert space, P an orthogonal projection of rank m and f_1, f_2, \dots, f_m a sequence of m orthonormal vectors in \mathcal{H} , which span the projection F . If for some $c < \infty$*

$$\max_{j=1, \dots, m} \|Pf_j - f_j\| \leq c, \tag{5.16}$$

then $\|P - F\| \leq (m + 2\sqrt{m})c$.

Proof. We employ the triangle inequality $\|P - F\| \leq \|PF - F\| + \|PF - PFP\| + \|P - PFP\|$ and bound the three terms on the right-hand side individually. For the first term we invoke that $PF - F$ vanishes on the orthogonal complement $\text{Im } F^\perp$ and, thus, a Frobenius norm estimate yields

$$\|PF - F\| \leq \sqrt{\sum_{j=1}^m \|(P - F)f_j\|^2} = \sqrt{\sum_{j=1}^m \|Pf_j - f_j\|^2} \leq \sqrt{m}c.$$

Our bound on the second term, relies on the norm estimate for the first term, $\|PF - PFP\| = \|P(F - FP)\| \leq \|F - FP\| = \|PF - F\| \leq \sqrt{m}c$, where we used that $\|P\| = 1$ for the first bound and applied the elementary identity $\|A\| = \|A^*\|$. For the last term, we employ the operator inequality $0 \leq PFP \leq P$ and the fact that the operator norm is bounded by the trace norm $\|\cdot\|_1$:

$$\|P - PFP\| \leq \|P - PFP\|_1 = \text{Tr } P - \text{Tr } PFP = \sum_{j=1}^m \langle \psi_j, (1 - P)\psi_j \rangle \leq mc.$$

This completes the proof. \square

5.3. Proof of Theorem 1.10. Before we dive into the details of the proof, we fix some notation. For $k \in \mathbb{N}$ and $\varepsilon > 0$, we will use the restricted Hamiltonian corresponding to the collection of all clusters $C_k(G)$,

$$H^{(c)} := \bigoplus_{G \in \mathcal{G}_{k,\varepsilon}} H_{C_k(G)}$$

acting on the complete Hilbert space $\ell^2(\mathcal{Q}_N)$. We further denote by

$$P_\varepsilon^{(c)} := \mathbb{1}_{(-\infty, -3N\varepsilon/2)}(H^{(c)}) = \bigoplus_{G \in \mathcal{G}_{k,\varepsilon}} P_{3\varepsilon/2}(G), \quad Q_\varepsilon^{(c)} := \mathbb{1} - P_\varepsilon^{(c)}$$

the spectral projections of $H^{(c)}$. The factor $3/2$ is motivated by the third assertion of Lemma 5.4. The subspace corresponding to $P_\delta^{(c)}$ represents the "localized" part of the

QREM and $Q_\delta^{(c)}$ corresponds to the "delocalized" part. Corresponding to this block decomposition, we set the diagonal parts of H as well as their partition functions:

$$H^{(1)} := P_\varepsilon^{(c)} H P_\varepsilon^{(c)}, \quad H^{(2)} := Q_\varepsilon^{(c)} H Q_\varepsilon^{(c)}, \tag{5.17}$$

$$Z_N^{(j)}(\beta, \Gamma) := 2^{-N} \text{Tr}_j e^{-\beta H^{(j)}}, \quad j = 1, 2. \tag{5.18}$$

Here the traces $\text{Tr}_j(\cdot)$ run over the natural subspaces $P_\varepsilon^{(c)} \ell^2(Q_N)$ in case $j = 1$, or $Q_\varepsilon^{(c)} \ell^2(Q_N)$ in case $j = 2$, on which $H^{(j)}$ acts non-trivially.

The key observation is now that $P_\varepsilon^{(c)}$ commutes with the restriction of H to the clusters $H^{(c)}$. If we denote by A the adjacency matrix between the inner and outer boundaries of the clusters $C_k(G)$, we see

$$P_\varepsilon^{(c)} H Q_\varepsilon^{(c)} = \Gamma P_\varepsilon^{(c)} A Q_\varepsilon^{(c)}. \tag{5.19}$$

We recall that $d(C_k(G), C_k(G')) \geq 2$ for two different components $G \neq G' \in \mathcal{G}_{k,\varepsilon}$, which implies that the adjacency matrix A is a direct sum of operators $A_{C_k(G)}$ corresponding to each cluster $C_k(G)$. This in turn yields

$$\|A\| \leq M c_k \sqrt{N} \tag{5.20}$$

by Proposition 2.1 on the event on which the assertions in Lemma 5.3 apply. We further observe that A only acts nontrivially on the boundaries $\partial C_k(G)$. Exploiting the decay estimate (5.11) from Lemma 5.4, we arrive at

$$\|P_\varepsilon^{(c)} H Q_\varepsilon^{(c)}\| = \mathcal{O}_{\Gamma,k,M,\varepsilon}(N^{-(k-1)/2}).$$

We conclude that for any $k \geq 2$:

$$Z_N(\beta, \Gamma) = e^{\mathcal{O}_{\Gamma,k,M,\varepsilon}(1)} (Z_N^{(1)} + Z_N^{(2)}). \tag{5.21}$$

The proof of Theorem 1.10 now reduces to an analysis of $Z_N^{(1)}$ and $Z_N^{(2)}$.

Proof of Theorem 1.10. Since our claims in case $\beta = 0$ and $\Gamma = 0$ are trivial, we fix $\beta, \Gamma > 0$ away from the phase transition, and pick

$$0 < \varepsilon < \frac{1}{8} \min\{\beta, \beta_c, \Gamma \tanh \beta \Gamma, \min\{1, \beta^{-1}\} \ln \cosh \beta \Gamma\}.$$

In the following, we will only work on the event $\Omega_{N,\beta_c}^{\text{REM}} \cap \Omega_N^u \cap \Omega_{N,M}(\varepsilon, k)$, where the conditions of Lemma 5.3 are valid at $k \geq 2$ and some M . According to Lemma 5.4 and (4.4) as well as (1.7), this event can be chosen to have a probability of at least $1 - e^{-cN}$.

We now proceed in four steps. We first analyze the localized part $Z_N^{(1)}$. As a second and third step we derive an upper and lower bound for $Z_N^{(2)}$. The last part then collects these estimates.

Step 1—Analysis of $Z_N^{(1)}$: Let us first remark that $H^{(1)} = P_\varepsilon^{(c)} H P_\varepsilon^{(c)} = P_\varepsilon^{(c)} H^{(c)} P_\varepsilon^{(c)}$ and hence $H^{(1)} P_\varepsilon^{(c)} = H^{(c)} P_\varepsilon^{(c)}$. It thus remains to consider the low energy spectrum of $H^{(c)}$. We abbreviate

$$u := u(\beta) := \lim_{N \rightarrow \infty} \langle U \rangle_\beta^{\text{cl}} / N = \begin{cases} -\beta, & \beta \leq \beta_c, \\ -\beta_c, & \beta > \beta_c, \end{cases}$$

by Proposition 5.1. Since $8\varepsilon < -u$, the dominant energy levels of U are not effected by the projection $P_\varepsilon^{(c)}$. We now split $Z_N^{(1)}$ into the contribution arising from energy levels within $J_\delta := [(u - \delta)N, (u + \delta)N]$ with $0 < \delta < \min\{-u/2 - 3\varepsilon/4, \varepsilon^2/(16\beta)\}$ arbitrarily small, and a remainder:

$$Z_N^{(1)}(\beta, \Gamma) = 2^{-N} \left(\text{Tr} e^{-\beta H^{(1)}} \mathbb{1}_{J_\delta}(H^{(1)}) + \text{Tr} e^{-\beta H^{(1)}} \mathbb{1}_{(-\infty, -3\varepsilon N/2) \setminus J_\delta}(H^{(1)}) \right).$$

The second term is estimated using Lemma 5.4 and subsequently Proposition 5.1, which yields some $c = c(\beta, \delta) > 0$ such that for all sufficiently large N :

$$\text{Tr} e^{-\beta H^{(1)}} \mathbb{1}_{\mathbb{R} \setminus J_\delta}(H^{(1)}) \leq e^{\beta \mathcal{O}_{\Gamma, k, M}(\sqrt{N})} \text{Tr} e^{-\beta U} \mathbb{1}_{\mathbb{R} \setminus J_{\delta/2}}(U) \leq e^{-cN} 2^N Z_N(\beta, 0). \tag{5.22}$$

The remaining term is decomposed further into the contribution of isolated and non-isolated clusters:

$$\begin{aligned} \text{Tr} e^{-\beta H^{(1)}} \mathbb{1}_{J_\delta}(H^{(1)}) &= \sum_{G \in \mathcal{I}_{k, \varepsilon}} \text{Tr} e^{-\beta H_{C_k(G)}} \mathbb{1}_{J_\delta}(H_{C_k(G)}) \\ &+ \sum_{G \in \mathcal{G}_{k, \varepsilon} \setminus \mathcal{I}_{k, \varepsilon}^c} \text{Tr} e^{-\beta H_{C_k(G)}} \mathbb{1}_{J_\delta}(H_{C_k(G)}). \end{aligned} \tag{5.23}$$

Since $\sup_{G \in \mathcal{G}_{k, \varepsilon}} \|H_{C_k(G)} - U_{C_k(G)}\| \leq \mathcal{O}_{\Gamma, k, M}(\sqrt{N})$ by Lemma 5.4, we bound the second term for all sufficiently large N as follows:

$$\begin{aligned} \sum_{G \in \mathcal{G}_{k, \varepsilon} \setminus \mathcal{I}_{k, \varepsilon}^c} \text{Tr} e^{-\beta H_{C_k(G)}} \mathbb{1}_{J_\delta}(H_{C_k(G)}) &\leq e^{-\beta N(u-\delta)} \sum_{G \in \mathcal{G}_{k, \varepsilon} \setminus \mathcal{I}_{k, \varepsilon}^c} \text{Tr} \mathbb{1}_{J_{2\delta}}(U_{C_k(G)}) \\ &\leq e^{-N\varepsilon^2/4} e^{3\beta N\delta} \sum_{G \in \mathcal{G}_{k, \varepsilon}} \text{Tr} e^{-\beta U_{C_k(G)}} \mathbb{1}_{J_{2\delta}}(U_{C_k(G)}) \leq e^{-cN} 2^N Z_N(\beta, 0). \end{aligned} \tag{5.24}$$

At the expense of throwing out another event of exponentially small probability, we consult Lemma 5.3 and assume in case $u > -\beta_c$ the validity of (5.6) with $a = -u - 2\delta$ and $b = -u + 2\delta$ and in case $u = -\beta_c$ the validity of (5.7) with $a = -u - 2\delta$. This guarantees that non-isolated clusters are exponentially rare. The last inequality is a consequence of the choice of δ and of the fact our definition of the partition function Z_N includes a normalisation by 2^{-N} .

The first term on the right side of (5.23) can be expressed using the energy correction formula (5.12) for isolated extremal sites. At the expense of excluding or including small subintervals at the boundary of J_δ , which are negligible in comparison to the main term by Proposition 5.1, this first term is of the form

$$\sum_{\sigma \in I_{k, \varepsilon} \cap \mathcal{L}_{u-\delta, u+\delta}} e^{-\beta(U(\sigma) + \frac{\Gamma^2 N}{U(\sigma)} + o(1))} = S - R,$$

where, similarly to (5.24), the remainder is again bounded using (5.6):

$$R := \sum_{\sigma \in I_{k, \varepsilon}^c \cap \mathcal{L}_{u-\delta, u+\delta}} e^{-\beta(U(\sigma) + \frac{\Gamma^2 N}{U(\sigma)} + o(1))} \leq e^{-cN} 2^N Z_N(\beta, 0).$$

The main term is

$$S := \sum_{\sigma \in \mathcal{Q}_N} e^{-\beta U(\sigma) + \frac{\Gamma^2 N}{U(\sigma)} + o(1)} \mathbb{1}[U(\sigma) \in J_\delta].$$

By definition of J_δ and since the REM's partition function concentrates around u by Proposition 5.1, S equals $2^N Z(\beta, 0)$ plus an error which is bounded by $e^{-cN} 2^N Z(\beta, 0)$.

In summary, in this first step we have shown that for any $\delta > 0$ small and N large enough:

$$e^{-\frac{\beta \Gamma^2}{u-\delta} + o(1)} Z_N(\beta, 0) \leq Z_N^{(1)}(\beta, \Gamma) \leq e^{-\frac{\beta \Gamma^2}{u+\delta} + o(1)} Z_N(\beta, 0). \tag{5.25}$$

Step 2—Upper bound on $Z_N^{(2)}$: We write $U_\varepsilon^< := U \mathbb{1}_{U \geq -2\varepsilon N}$ and $U_\varepsilon^> := U \mathbb{1}_{U > -2\varepsilon N}$ as well as $U_\varepsilon := U \mathbb{1}_{|U| \leq 2\varepsilon N}$, and estimate using the Jensen-Peierls inequality [12]:

$$\begin{aligned} 2^N Z_N^{(2)}(\beta, \Gamma) &= \text{Tr} 2e^{-\beta Q_\varepsilon^{(c)}[\Gamma T + U_\varepsilon^< + Q_\varepsilon^{(c)} U_\varepsilon^> Q_\varepsilon^{(c)}] Q_\varepsilon^{(c)}} \\ &\leq \text{Tr} Q_\varepsilon^{(c)} e^{-\beta[\Gamma T + U_\varepsilon^< + Q_\varepsilon^{(c)} U_\varepsilon^> Q_\varepsilon^{(c)}]} \leq \text{Tr} e^{-\beta[\Gamma T + U_\varepsilon + Q_\varepsilon^{(c)} U_\varepsilon^> Q_\varepsilon^{(c)}]}. \end{aligned}$$

The last inequality follows from a trivial extension of the trace and the monotonicity of eigenvalues in the potential, $U_\varepsilon^< \geq U_\varepsilon$. From Lemma 5.4 we learn that $\max_{\sigma \in \mathcal{L}_{2\varepsilon}} \|Q_\varepsilon^{(c)} \delta_\sigma\|^2 \leq CN^{-1}$. Moreover, if $\sigma \in C_k(G)$ for some component G , the projection $Q_\varepsilon^{(c)} \delta_\sigma$ has only support on $C_k(G)$. As any cluster $C_k(G)$ has at most M configurations $\sigma \in \mathcal{L}_{2\varepsilon}$, these observations result in the norm estimate

$$\|Q_\varepsilon^{(c)} U_\varepsilon^> Q_\varepsilon^{(c)}\| \leq CM \|U\|_\infty N^{-1}.$$

Since on the event considered we also have $\|U\|_\infty \leq 2\beta_c$ and the operator $Q_\varepsilon^{(c)} U_\varepsilon^> Q_\varepsilon^{(c)}$ only acts non trivially on the clusters $C_k(G)$, we thus conclude that for some $D \in (0, \infty)$:

$$Q_\varepsilon^{(c)} U_\varepsilon^> Q_\varepsilon^{(c)} \geq V := -D \mathbb{1}_C = -D \sum_{G \in \mathcal{G}_{k,\varepsilon}} \mathbb{1}_{C_k(G)}, \quad C := \bigcup_{G \in \mathcal{G}_{k,\varepsilon}} C_k(G)$$

To summarize, we have thus shown that $Z_N^{(2)}(\beta, \Gamma) \leq 2^{-N} \text{Tr} e^{-\beta[\Gamma T + U_\varepsilon + V]}$.

From here, there are at least two possible ways to continue the proof. One could show that the potential $U_\varepsilon + V$ meets the requirements of Theorem 3.4. Then, one needs to control V , which is a little bit technical. Instead, we will employ a convexity argument. To this end, we introduce for $\lambda \in \mathbb{R}$ the family of pressures and corresponding Hamiltonians on $\ell^2(\mathcal{Q}_N)$:

$$\Phi_N(\beta, \Gamma, \lambda) := \ln 2^{-N} \text{Tr} e^{-\beta H(\lambda)}, \quad H(\lambda) := \Gamma T + U_\varepsilon + \lambda V \tag{5.26}$$

The pressure $\Phi_N(\beta, \Gamma, \lambda)$ is convex [60] in λ , and $\lambda = 1$ is the case of interest.

Let us first discuss the case $\lambda = 0$ in which case Theorem 3.4 is applicable with $W = U_\varepsilon$. Since $\|U_\varepsilon\|_\infty \leq 2\varepsilon N$ and $\mathbb{E}[U_\varepsilon(\sigma)^2] \leq N(1 - e^{-2\varepsilon^2 N}) \leq N$, Theorem 3.4 guarantees that all eigenvalues of $\Gamma T + U_\varepsilon$ below $E < -4\varepsilon N$, counted with multiplicity, are shifted with respect to the eigenvalues E of ΓT to $E + \frac{N}{E} + o(1)$. Since $\langle \Gamma T \rangle_\beta^{\text{pm}} = -N\Gamma \tanh \beta \Gamma \leq -8\varepsilon N$, Proposition 5.1 allows to spectrally focus the partition function onto an interval around $\langle \Gamma T \rangle_\beta^{\text{pm}}$ of arbitrarily small size $0 < \delta < \Gamma \tanh \beta \Gamma - 4\varepsilon$. A similar argument as in Step 1, then yields for all sufficiently large N :

$$\Phi_N(\beta, \Gamma, 0) \leq N \ln \cosh \beta\Gamma + \frac{\beta}{\Gamma \tanh \beta\Gamma - \delta} + o(1).$$

Next, we consider general parameters λ . Recall that $\mathbb{1}_C$ stands for the orthogonal projection onto the subspace of the union of all clusters, and that $\mathbb{1}_{C^c}$ is the orthogonal complement. In terms of the operator A introduced in (5.19), the norm estimate (5.20) yields: $\|H(\lambda) - \mathbb{1}_C H(\lambda) \mathbb{1}_C - \mathbb{1}_{C^c} H(\lambda) \mathbb{1}_{C^c}\| \leq \|A\| \leq C\sqrt{N}$ and hence

$$\begin{aligned} \text{Tr } e^{-\beta H(\lambda)} &\leq e^{C\sqrt{N}} \text{Tr } e^{-\beta(\mathbb{1}_C H(\lambda) \mathbb{1}_C + \mathbb{1}_{C^c} H(\lambda) \mathbb{1}_{C^c})} \\ &= e^{C\sqrt{N}} \left[\text{Tr } \mathbb{1}_C e^{-\beta(\mathbb{1}_C H(\lambda) \mathbb{1}_C)} + \text{Tr } \mathbb{1}_{C^c} e^{-\beta(\mathbb{1}_{C^c} H(\lambda) \mathbb{1}_{C^c})} \right] \end{aligned}$$

at some $C < \infty$, which is independent of N and λ . Each of the traces in the right side is now estimated separately:

$$2^{-N} \text{Tr } \mathbb{1}_C e^{-\beta(\mathbb{1}_C H(\lambda) \mathbb{1}_C)} \leq e^{\beta \|\mathbb{1}_C H(\lambda) \mathbb{1}_C\|} \leq \exp\left(\beta \left(C\Gamma\sqrt{N} + 2\varepsilon N + \lambda D\right)\right)$$

where we used the triangle inequality for the operator norm as well as (5.20) again. Since $H(\lambda)$ and $H(0)$ agree on $\mathbb{1}_{C^c}$ $\ell^2(\mathcal{Q}_N)$, we also have

$$\begin{aligned} 2^{-N} \text{Tr } \mathbb{1}_{C^c} e^{-\beta \mathbb{1}_{C^c} H(\lambda) \mathbb{1}_{C^c}} &= 2^{-N} \text{Tr } \mathbb{1}_{C^c} e^{-\beta \mathbb{1}_{C^c} H(0) \mathbb{1}_{C^c}} \\ &\leq 2^{-N} \text{Tr } \mathbb{1}_{C^c} e^{-\beta H(0)} \leq e^{\Phi_N(\beta, \Gamma, 0)}. \end{aligned}$$

The first inequality relied on the Jensen–Peierls estimate, which allows to pull down the projections [12].

Since $\Phi_N(\beta, \Gamma, 0) > 4\beta\varepsilon N$, the correction to the pressure at $\lambda_0 := \frac{\varepsilon N}{D}$, is still of order $\mathcal{O}(\sqrt{N})$:

$$\Phi_N(\beta, \Gamma, \lambda_0) \leq \Phi_N(\beta, \Gamma, 0) + C\sqrt{N}.$$

We are now in the situation to exploit convexity:

$$\begin{aligned} \Phi_N(\beta, \Gamma, 1) &\leq (1 - \lambda_0^{-1}) \Phi_N(\beta, \Gamma, 0) + \lambda_0^{-1} \Phi_N(\beta, \Gamma, \lambda_0) \\ &\leq N \ln \cosh \beta\Gamma + \frac{\beta}{\Gamma \tanh \beta\Gamma - \delta} + o(1). \end{aligned}$$

Step 3—Paramagnetic lower bound: To show that the upper bound of Step 2 is also an asymptotic lower bound, it is more convenient to work with the full partition function Z_N , which by (5.21) is a lower bound on $Z_N^{(2)}$ up to a multiplicative error of $e^{o(1)}$.

For an estimate on Z_N , we split the potential $U = U_\varepsilon + V_\varepsilon$, where $V_\varepsilon := U \mathbb{1}[|U| > 2\varepsilon]$. The pressure $\Phi_N(\beta, \Gamma, 0)$ of $H(0) = \Gamma T + U_\varepsilon$, which is defined in (5.26), was already analyzed in Step 2. Here, we now consider the following family of Hamiltonians $H(\lambda) = \Gamma T + U_\varepsilon + \lambda V_\varepsilon$, which differs from the one in Step 2. By a slight abuse of notation, we nevertheless denote the corresponding pressure again by $\Phi_N(\beta, \Gamma, \lambda) := \ln \text{Tr } 2^{-N} \text{Tr } e^{-\beta H(\lambda)}$. The convexity of the pressure in λ is again the basis for our argument.

Since on the event considered, we may assume $\|U\| \leq 2\beta_c N$, the potential $W = U_\varepsilon + \lambda V_\varepsilon$ meets the requirements of Theorem 3.4 with $\|W\|_\infty \leq N \max\{2\varepsilon, 2\lambda\beta_c\}$ and $\mathbb{E}[W(\sigma)^2] \leq N(1 - e^{-2\varepsilon^2 N}) \leq N$. Thus if $\lambda < \varepsilon/\beta_c$ the eigenvalues of $H(\lambda)$ below $E < -4\varepsilon N$, counted with multiplicity, are shifted with respect to the eigenvalues

E of ΓT to $E + \frac{N}{E} + o(1)$ and we have $\Phi_N(\beta, \Gamma, \lambda) = \Phi_N(\beta, \Gamma, 0) + o(1)$. Fixing $\lambda_0 := \frac{\varepsilon}{2\beta_c} < 1$, convexity implies:

$$\begin{aligned}\Phi_N(\beta, \Gamma, 1) &\geq \Phi_N(\beta, \Gamma, 0) + \frac{1}{\lambda_0} (\Phi_N(\beta, \Gamma, \lambda_0) - \Phi_N(\beta, \Gamma, 0)) \\ &= N \ln \cosh(\beta\Gamma) + \frac{\beta}{\Gamma \tanh \beta\Gamma + \delta} + o(1).\end{aligned}$$

The last step, which holds for all $\delta > 0$ sufficiently small, is the desired lower bound, again relied on an explicit estimate based on the concentration of the partition function of ΓT around energies near $-\Gamma \tanh \beta\Gamma$, cf. Proposition 5.1.

Step 4—Completing the proof. Away from the first-order phase transition at $\Gamma = \Gamma_c(\beta)$ described in Proposition 1.1 and on the event on which Step 1–3 are valid, the partition function (5.21) is either dominated by the REM-term $Z_N^{(1)}$ in case $\Gamma < \Gamma_c(\beta)$, or by the paramagnetic term $Z_N^{(2)}$ in case $\Gamma > \Gamma_c(\beta)$.

More precisely, in case $\Gamma < \Gamma_c(\beta)$ and since the probability of the event, which is excluded in Step 1–3, is exponentially small in N and hence summable, we conclude for any $\varepsilon > 0$:

$$\sum_{N \geq 1} \mathbb{P} \left(\left| \Phi_N(\beta, \Gamma) - \Phi_N(\beta, 0) + \frac{\beta\Gamma^2}{u(\beta)} \right| > \varepsilon \right) < \infty. \quad (5.27)$$

The claimed almost-sure convergence then follows by a Borel-Cantelli argument. The analogous argument establishes the claim in case $\Gamma > \Gamma_c(\beta)$. \square

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

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Adame, J., Warzel, S.: Exponential vanishing of the ground-state gap of the QREM via adiabatic quantum computing. *J. Math. Phys.* **56**, 113301 (2015)
2. Adhikari, A., Brennecke, C.: Free energy of the quantum Sherrington–Kirkpatrick spin-glass model with transverse field. *J. Math. Phys.* **61**, 083302 (2020)
3. Aizenman, M., Warzel, S.: *Random Operators: Disorder Effects on Quantum Spectra and Dynamics*. AMS, Providence (2015)
4. Aizenman, M., Lebowitz, J.L., Ruelle, D.: Some rigorous results on the Sherrington–Kirkpatrick spin glass model. *Commun. Math. Phys.* **112**, 3–20 (1987)
5. Aizenman, M., Shamis, M., Warzel, S.: Resonances and partial delocalization on the complete graph. *Ann. Henri Poincaré* **16**, 1969–2003 (2015)

6. Avena, L., Gün, O., Hesse, M.: The parabolic Anderson model on the hypercube. *Stoch. Proc. Appl.* **130**, 3369–3393 (2020)
7. Baake, E., Baake, M., Wagner, H.: Ising quantum chain is equivalent to a model of biological evolution. *Phys. Rev. Lett.* **78**, 559–562 (1997). Erratum: *Phys. Rev. Lett.* **79**, 1782 (1997)
8. Baake, E., Wagner, H.: Mutation-selection models solved exactly with methods of statistical mechanics. *Genet. Res.* **78**, 93–117 (2001)
9. Baldwin, C.L., Laumann, C.R., Pal, A., Scardicchio, A.: The many-body localized phase of the quantum random energy model. *Phys. Rev. B* **93**, 024202 (2016)
10. Ben Arous, G., Bovier, A., Gayrard, V.: Glauber dynamics of the random energy model. I. Metastable motion on the extreme states. *Commun. Math. Phys.* **235**, 379–425 (2003)
11. Ben Arous, G., Bovier, A., Gayrard, V.: Glauber dynamics of the random energy model. II. Aging below the critical temperature. *Commun. Math. Phys.* **236**, 1–54 (2003)
12. Berezin, F.A.: Convex operator functions. *Math. USSR. Sbornik* **17**, 269–277 (1972)
13. Bhatia, R.: *Matrix Analysis*. Springer, New York (1997)
14. Biroli, G., Facchetti, D., Schiró, M., Tarzia, M., Vivo, P.: Out-of-equilibrium phase diagram of the quantum random energy model. *Phys. Rev. B* **103**, 014204 (2021)
15. Biskup, M., König, W.: Eigenvalue order statistics for random Schrödinger operators with doubly-exponential tails. *Commun. Math. Phys.* **341**, 179–218 (2016)
16. Bollobás, B., Leed, J., Letzter, S.: Eigenvalues of subgraphs of the cube. *Eur. J. Comb.* **70**, 125–148 (2018)
17. Bovier, A.: *Statistical Mechanics of Disordered Systems. A Mathematical Perspective*. Cambridge University Press, Cambridge (2012)
18. Bovier, A., Kurkova, I.: Derrida’s generalised random energy models. I. Models with finitely many hierarchies. *Ann. Inst. H. Poincaré Probab. Stat.* **40**, 439–480 (2004)
19. Bovier, A., Kurkova, I.: Derrida’s generalized random energy models. II. Models with continuous hierarchies. *Ann. Inst. H. Poincaré Probab. Stat.* **40**, 481–495 (2004)
20. Bovier, A., Kurkova, I., Löwe, M.: Fluctuations of the free energy in the REM and the p-spin SK models. *Ann. Probab.* **30**, 605–651 (2002)
21. Bray, A.J., Moore, M.A.: Replica theory of quantum spin glasses. *J. Phys. C Solid State Phys.* **13**, L655 (1980)
22. Burin, A.: 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. *Ann. Phys.* **529**, 1600292 (2017)
23. Capocaccia, D., Cassandro, M., Picco, P.: On the existence of thermodynamics for the generalized random energy model. *J. Stat. Phys.* **46**, 493–505 (1987)
24. Cerný, J., Wassmer, T.: Aging of the metropolis dynamics on the random energy model. *Probab. Theory Relat. Fields* **167**, 1–51 (2015)
25. Crawford, N.: Thermodynamics and universality for mean field quantum spin glasses. *Commun. Math. Phys.* **274**, 821–839 (2007)
26. Cvetković, D.M., Doob, M., Sachs, H.: *Spectra of Graphs: Theory and Applications*, 3rd edn. Johann Ambrosius Barth, Heidelberg (1995)
27. De Tomasi, G., Khaymovich, I.M., Pollmann, F., Warzel, S.: Rare thermal bubbles at the many-body localization transition from the Fock space point of view. *Phys. Rev. B* **104**, 024202 (2021)
28. Derrida, B.: Random energy model: limit of a family of disordered models. *Phys. Rev. Lett.* **45**, 79–82 (1980)
29. Derrida, B.: Random energy model: an exactly solvable model of disordered systems. *Phys. Rev. B* **24**, 2613–2326 (1981)
30. Derrida, B.: A generalization of the random energy model that includes correlations between the energies. *J. Phys. Lett.* **46**, 401–407 (1985)
31. Derrida, B., Gardner, E.: Solution of the generalized random energy model. *J. Phys. C* **19**, 2253–2274 (1986)
32. Faoro, L., Feigelman, M.V., Ioffe, L.: Non-ergodic extended phase of the quantum random energy model. *Ann. Phys.* **409**, 167916 (2019)
33. Friedman, J., Tillich, J.P.: Generalized Alon–Boppana theorems and error-correcting codes. *SIAM J. Discrete Math.* **19**, 700–718 (2005)
34. Galves, A., Martinez, S., Picco, P.: Fluctuations in Derrida’s random energy and generalized random energy models. *J. Stat. Phys.* **54**, 515–529 (1989)
35. Gayrard, V.: Aging in Metropolis dynamics of the REM: a proof. *Probab. Theory Relat. Fields* **174**, 501–551 (2019)
36. Gayrard, V., Hartung, L.: Dynamic phase diagram of the REM. In: Gayrard, V., Arguin, L.P., Kistler, N., Kourkova, I. (eds.) *Statistical Mechanics of Classical and Disordered Systems*. StaMeClaDys 2018. Springer Proceedings in Mathematics & Statistics, vol. 293. Springer, Cham (2019)
37. Goldschmidt, Y.Y.: Solvable model of the quantum spin glass in a transverse field. *Phys. Rev. B* **41**, 4858 (1990)

38. Grenkova, L.N., Molchanov, S.A., Sudarev, Yu.N.: Structure of spectrum edge for multidimensional Anderson model. *Theor. Math. Phys.* **85**, 1033–1039 (1990)
39. Hermisson, J., Redner, O., Wagner, H., Baake, E.: Mutation-selection balance: ancestry, load, and maximum principle. *Theor. Popul. Biol.* **62**, 9–46 (2002)
40. Jörg, T., Krzakala, F., Kurchan, J., Maggs, A.C.: Simple glass models and their quantum annealing. *Phys. Rev. Lett.* **101**, 147204 (2008)
41. Jörg, T., Krzakala, F., Semerjian, G., Zamponi, F.: First-order transitions and the performance of quantum algorithms in random optimization problems. *Phys. Rev. Lett.* **104**, 207206 (2010)
42. Kingman, J.F.C.: A simple model for the balance between selection and mutation. *J. Appl. Probab.* **15**, 1–12 (1978)
43. König, W.: *The Parabolic Anderson Model*. Birkhäuser, Basel (2016)
44. Kravtsov, V.E., Khaymovich, I.M., Cuevas, E., Amini, M.: A random matrix model with localization and ergodic transitions. *New J. Phys.* **17**, 122002 (2015)
45. Laumann, C.R., Pal, A., Scardicchio, A.: Many-body mobility edge in a mean-field quantum spin glass. *Phys. Rev. Lett.* **113**, 200405 (2014)
46. Leadbetter, M.R., Lindgren, G., Rootzén, H.: *Extremes and Related Properties of Random Sequences and Processes*. Springer Series in Statistics. Springer, New York (1983)
47. Leschke, H., Manai, C., Ruder, R., Warzel, S.: Existence of replica-symmetry breaking in quantum glasses. *Phys. Rev. Lett.* **127**, 207204 (2021)
48. Leschke, H., Rothlauf, S., Ruder, R., Spitzer, W.: The free energy of a quantum Sherrington–Kirkpatrick spin-glass model for weak disorder. *J. Stat. Phys.* **182**, 55 (2021)
49. Manai, C., Warzel, S.: Spectral Analysis of the Quantum Random Energy Model. Extended version [arXiv:2202.00334](https://arxiv.org/abs/2202.00334) (2022)
50. Manai, C., Warzel, S.: Phase diagram of the quantum random energy model. *J. Stat. Phys.* **180**, 654–664 (2020)
51. Manai, C., Warzel, S.: The quantum random energy model as a limit of p-spin interactions. *Rev. Math. Phys.* **33**, 2060013 (2020)
52. Manai, C., Warzel, S.: Generalized random energy models in a transversal magnetic field: free energy and phase diagrams. *Probab. Math. Phys.* **3**, 215–245 (2022)
53. Manai, C., Warzel, S.: The de Almeida–Thouless line in hierarchical quantum spin glasses. *J. Stat. Phys.* **186**, 14 (2022)
54. Mézard, M., Parisi, G., Virasoro, M.A.: *Spin Glass Theory and Beyond*. World Scientific, Singapore (1987)
55. Oliveira, R. I.: Concentration of the adjacency matrix and of the Laplacian in random graphs with independent edges. Preprint [arXiv:0911.0600](https://arxiv.org/abs/0911.0600) (2010)
56. Olivieri, E., Picco, P.: On the existence of thermodynamics for the random energy model. *Commun. Math. Phys.* **96**, 125–144 (1984)
57. Panchenko, D.: *The Sherrington–Kirkpatrick Model*. Springer, New York (2013)
58. Parisi, G.: The order parameter for spin glasses: a function on the interval 0–1. *J. Phys. A Math. Gen.* **13**, 1101–1112 (1980)
59. Ruelle, D.: A mathematical reformulation of Derrida’s REM and GREM. *Commun. Math. Phys.* **108**, 225–239 (1987)
60. Simon, B.: *The Statistical Mechanics of Lattice Gases*. Princeton University Press, Princeton (1993)
61. Smelyanskiy, V.N., Kechedzhi, K., Boixo, S., Isakov, S.V., Neven, H., Altshuler, B.: Nonergodic delocalized states for efficient population transfer within a narrow band of the energy landscape. *Phys. Rev. X* **10**, 011017 (2020)
62. Suzuki, S., Inoue, J., Chakrabarti, B.K.: *Quantum Ising Phases and Transitions in Transverse Ising Models*, 2nd edn. Springer, New York (2013)
63. Szendro, I.G., Schenk, M.F., Franke, J., Krug, J., de Visser, J.A.G.M.: Quantitative analyses of empirical fitness landscapes. *J. Stat. Mech. Theory Exp.* **2013**, P01005 (2013)
64. Talagrand, M.: Concentration of measure and isoperimetric inequalities in product spaces. *Publ. Math. IHES* **81**, 73–205 (1995)
65. Talagrand, M.: *Mean Field Models for Spin Glasses*, vols. I+II. Springer, New York (2011)
66. Tao, T.: *Topics in Random Matrix Theory*. AMS, Providence (2012)
67. Tropp, J.A.: An introduction to matrix concentration inequalities. *Found. Trends Mach. Learn.* **8**, 1–230 (2015)
68. von Soosten, P., Warzel, S.: Non-ergodic delocalization in the Rosenzweig–Porter model. *Lett. Math. Phys.* **109**, 905–922 (2019)