HAL open science

# Quasi-diffusion in a 3D Supersymmetric Hyperbolic Sigma Model

Margherita Disertori, Tom Spencer, Martin R. Zirnbauer

**To cite this version:**

Margherita Disertori, Tom Spencer, Martin R. Zirnbauer. Quasi-diffusion in a 3D Supersymmetric Hyperbolic Sigma Model. Communications in Mathematical Physics, 2010, 300 (2), pp.435-486. 10.1007/s00220-010-1117-5 . hal-00352357

**HAL Id: hal-00352357**

**https://hal.science/hal-00352357v1**

Submitted on 12 Jan 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Quasi-diffusion in a 3D Supersymmetric Hyperbolic Sigma Model

M. Disertori[a]*, T. Spencer[b], M.R. Zirnbauer[c]
*a*) Laboratoire de Mathématiques Raphaël Salem, UMR CNRS 6085
Université de Rouen, 76801, France
*b*) Institute for Advanced Study, Einstein Drive,
Princeton, NJ 08540, USA
*c*) Institut für Theoretische Physik, Universität zu Köln,
Zülpicher Straße 77, 50937 Köln, Germany

January 12, 2009

**Abstract**

Abstract: We study a lattice field model which qualitatively reflects the phenomenon of Anderson localization and delocalization for real symmetric band matrices. In this statistical mechanics model, the field takes values in a supermanifold based on the hyperbolic plane. Correlations in this model may be described in terms of a random walk in a highly correlated random environment. We prove that in three or more dimensions the model has a 'diffusive' phase at low temperatures. Localization is expected at high temperatures. Our analysis uses estimates on non-uniformly elliptic Green's functions and a family of Ward identities coming from internal supersymmetry.

## 1 Introduction

### 1.1 Some history and motivation

It has been known since the pioneering work of Wegner [20, 19] that information about the spectral and transport properties of random band matrices and random Schrödinger operators can be inferred from the correlation functions of statistical mechanical models of a certain kind. These models have a hyperbolic symmetry, typically a noncompact group such as $\mathrm{O}(p,q)$ or $\mathrm{U}(p,q)$, and were originally studied in the limit of $p = q = 0$ replicas.

*e-mail: Margherita.Disertori@univ-rouen.fr

The connection between random Schrödinger operators and statistical mechanics models was made more precise by Efetov [7], who introduced the so-called supersymmetry method to avoid the use of replicas. In Efetov's formulation one employs both commuting (or bosonic) and anticommuting (or Grassmann) integration variables, and these are related by a natural symmetry that makes the emerging statistical mechanics system supersymmetric (SUSY). The simplest class of these models has a $\mathrm{U}(1,1|2)$ symmetry. This means that for the bosonic variables there exists a hyperbolic symmetry $\mathrm{U}(1,1)$ preserving an indefinite Hermitian form on $\mathbb{C}^2$, and the Grassmann variables are governed by a compact $\mathrm{U}(2)$ symmetry. Moreover, there exist odd symmetries mixing Grassmann and bosonic variables.

The fields $\mathbb{Z}^d \ni j \mapsto Q_j$ of the supersymmetric models introduced by Efetov are 4 by 4 supermatrices built from bosonic as well as Grassmann entries. In the physics literature one usually assumes the sigma model approximation, which is believed to capture the essential features of the energy correlations and transport properties of the underlying quantum system. The sigma model approximation constrains the matrix field $Q$ by $Q_j^2 = \mathrm{Id}$ for all $j$. This constraint is similar to the constraints appearing in the Ising or Heisenberg models, where $S_j \cdot S_j = 1$. We refer the reader to [6, 14, 9, 3, 11] for an introduction to these ideas.

The models described above are difficult to analyse with mathematical rigor in more than one dimension. In this paper we study a simpler SUSY model. Our exposition will be essentially self-contained and the full supersymmetric formalism alluded to here will serve primarily as a source of motivation.

## 1.2 Probabilistic representation of our model

In this paper we analyze a lattice field model which may be thought of as a simplified version of one of Efetov's nonlinear sigma models. More precisely, it is related to the model that derives from real symmetric matrices, see Section 3. In this statistical mechanics model the field at site $j$ has four degrees of freedom. Two of these, $t_j$ and $s_j$, parametrize a hyperboloid and the other two, $\bar{\psi}_j$ and $\psi_j$, are Grassmann (i.e., anticommuting) variables. Technically speaking, the field takes values in a target space denoted by $\mathrm{H}^{2|2}$, which is a supermanifold extension of the hyperbolic plane $\mathrm{H}^2$; see Section 2. This model was introduced by one of us in [21, 5], and localization was established in one dimension (1D) in the sense that the conductance was proven to decay exponentially in the system size [21]. The model is expected to reflect the qualitative behavior of random band matrices – namely localization and diffusion – in any dimension. However, most of our discussion will be restricted to three dimensions.

Our supersymmetric hyperbolic nonlinear sigma model, called the $\mathrm{H}^{2|2}$ model for short, will be formulated on a lattice cube $\Lambda \subset \mathbb{Z}^d$ of side $L$. We shall see (in Section 2.2) that the action of the field variables is quadratic in $\psi$, $\bar{\psi}$, and $s$. This feature is special to the horospherical coordinate system that we use. It enables us to reduce the $\mathrm{H}^{2|2}$ model to the statistical mechanics of a *single* field $t : \Lambda \to \mathbb{R}$, $j \mapsto t_j$. Its free energy or effective action, $F(t)$, is real, so the resulting statistical mechanical model has a probabilistic interpretation.

In order to specify $F(t)$, first consider the finite-difference elliptic operator $D_{\beta,\varepsilon}(t)$ defined by the quadratic form

$$[v\,;D_{\beta,\varepsilon}(t)\,v]_\Lambda = \beta\sum\nolimits_{(ij)} \mathrm{e}^{t_i+t_j}(v_i-v_j)^2 + \varepsilon\sum\nolimits_{k\in\Lambda}\mathrm{e}^{t_k}v_k^2\ . \tag{1.1}$$

This operator plays a central role in our analysis. The first sum is over nearest neighbor pairs in $\Lambda$, and $[\,;\,]_\Lambda$ denotes the usual scalar product in $\ell^2(\Lambda)$. We see that $D_{1,0}(\mathbf{0})$ is the finite-difference Laplacian. The regularization parameter $\varepsilon>0$ will serve to make the theory well-defined. One may interpret $D_{\beta,\varepsilon}(t)$ as the generator of a random walk in an environment given by the fluctuating field $t$, with a death rate of $\varepsilon\,\mathrm{e}^{t_j}$ at site $j$. Note that the operator $D$ is elliptic but not uniformly so, as $t_j\in\mathbb{R}$ has unbounded range.

The free energy or effective action $F_{\beta,\varepsilon}(t)$ is now expressed by

$$\begin{aligned} F_{\beta,\varepsilon}(t) &= \beta\sum\nolimits_{(ij)}(\cosh(t_i-t_j)-1) \\ &\quad -\ln\mathrm{Det}^{1/2}D_{\beta,\varepsilon}(t) + \sum\nolimits_{k\in\Lambda}(t_k-\varepsilon+\varepsilon\,\cosh t_k)\ . \end{aligned} \tag{1.2}$$

If $d\mu_\Lambda$ denotes the product measure

$$d\mu_\Lambda = \prod_{k\in\Lambda}\frac{dt_k}{\sqrt{2\pi}} \tag{1.3}$$

on $\mathbb{R}^{|\Lambda|}$, then the partition function is

$$Z_\Lambda(\beta,\varepsilon) = \int_{\mathbb{R}^{|\Lambda|}}\mathrm{e}^{-F_{\beta,\varepsilon}}\,d\mu_\Lambda = 1\ . \tag{1.4}$$

The partition function is identically equal to unity independent of $\beta$, $\varepsilon$ even when $\beta$ depends on the edge $(ij)$ and $\varepsilon$ depends on the lattice point $k$; see (5.1). This is a reflection of an internal supersymmetry which will be explained in later sections. There exist many variants of this identity. One of them gives us easy control of nearest neighbor fluctuations of the field $t$ (cf. Section 6).

The expectation of an observable function $t\mapsto f(t)$ is defined by

$$\langle f\rangle_{\Lambda,\beta,\varepsilon} = \int f\,\mathrm{e}^{-F_{\beta,\varepsilon}}d\mu_\Lambda\ . \tag{1.5}$$

Let us make a few comments on these expository definitions.

1. The action or free energy $F_{\beta,\varepsilon}(t)$ is nonlocal due to the presence of the term $-\ln\mathrm{Det}^{1/2}D(t)$. This nonlocality arises from integrating out three massless free fields, one $(s)$ of bosonic and two $(\bar\psi,\psi)$ of Grassmann type.

2. $F_{\beta,\varepsilon}(t)$ is not convex as a function of $t$ and therefore the Brascamp-Lieb estimates used in earlier work on a related model [18] do not apply. The lack of convexity is an important feature and opens the possibility for a localization-delocalization transition to occur.

3. When $\varepsilon = 0$, $F_{\beta,0}(t)$ is invariant under shifts $t_j \to t_j + c$ by any constant $c \in \mathbb{R}$. To see this, note that for $\varepsilon = 0$ we have $D_{\beta,0}(t+c) = \mathrm{e}^{2c} D_{\beta,0}(t)$ by (1.1). The resulting additional term $-|\Lambda|c$ from $-\ln \mathrm{Det}^{1/2} D_{\beta,0}(t)$ in (1.2) is canceled by another such term, which arises from shifting $\sum_{k\in\Lambda} t_k$. This symmetry (which is a formal one, since the integral is ill-defined for $\varepsilon = 0$) is associated with the presence of a massless mode. The importance of the regularization $\varepsilon$, which was omitted from the present argument, becomes evident from the saddle point discussed below.

4. The model at hand describes a disordered quantum system at *zero temperature*. Nevertheless, adopting the familiar language of statistical mechanics and thermodynamics, we refer to the field stiffness $\beta$ as the inverse 'temperature'. ($\beta$ is actually the dimensionless conductance for an Ohmic system of size $L = 1$ as measured in lattice units.)

## 1.3 Main result

The main goal of this paper is to estimate the fluctuations of the field $t$ for large values of the parameter $\beta$ and dimension $d = 3$. This will enable us to prove that the random walk in the random environment drawn from $F(t)$ is transient. More precisely, we will prove the following. (Similar estimates hold for all dimensions $d \geq 3$.)

**Theorem 1** *For $d = 3$, there is a $\bar{\beta} \geq 1$ such that if $\beta \geq \bar{\beta}$, the fluctuations of the field $t$ are uniformly bounded in $x$, $y$, and $\Lambda$ :*

$$\langle \cosh^m (t_x - t_y) \rangle_{\Lambda,\beta,\,\varepsilon} \leq 2 \,, \tag{1.6}$$

*provided that $m \leq \beta^{1/8}$.*

This theorem implies that for any $x$ and $y$, $|t_x - t_y|$ is very unlikely to be large. A stronger version of (1.6) is given in (9.3). We will use this result to prove

**Theorem 2** *Under the hypothesis of Theorem 1 the average field is bounded:*

$$\langle \cosh^p (t_x) \rangle_{\Lambda,\beta,\,\varepsilon} \leq \frac{5}{2} \,, \tag{1.7}$$

*provided $p \leq 10$ and $|\Lambda|^{1-\alpha/3}\,\varepsilon \geq 1$ with $\alpha \geq 1/\ln\beta$. Thus in the thermodynamic limit $|\Lambda| \to \infty$ we may send $\varepsilon \to 0$ while maintaining the bound on $\langle \cosh^p t_x \rangle$.*

To investigate the localized or extended nature of the energy eigenstates of a disordered quantum system with Hamiltonian $H$, one looks at the average square of the quantum Green's function, $|(H - E + \mathrm{i}\varepsilon)^{-1}(x,y)|^2$. The analog of this Green's function in the $\mathrm{H}^{2|2}$ model is the two-point correlation function

$$C_{xy} = \left\langle \mathrm{e}^{t_x} s_x \, \mathrm{e}^{t_y} s_y \right\rangle , \tag{1.8}$$

where the expectation is given by the full functional integral defined in Sections 2.1, 2.2. After integration over the fields $\bar{\psi}$, $\psi$, and $s$, we have

$$C_{xy} = \left\langle \mathrm{e}^{t_x+t_y} D_{\beta,\varepsilon}(t)^{-1}(x,y) \right\rangle_{\Lambda,\beta,\varepsilon} \equiv \left\langle \tilde{D}_{\beta,\varepsilon}(t)^{-1}(x,y) \right\rangle_{\Lambda,\beta,\varepsilon} , \tag{1.9}$$

where $\tilde{D} = \mathrm{e}^{-t} D \circ \mathrm{e}^{-t}$. Note that $C_{xy}$ is positive both pointwise and as a quadratic form. A simple calculation shows that

$$\tilde{D}_{\beta,\varepsilon}(t) = -\beta\Delta + \beta V(t) + \varepsilon\, \mathrm{e}^{-t} , \tag{1.10}$$

where $V(t)$ is a diagonal matrix (or 'potential') given by

$$V_{jj}(t) = \sum\nolimits_{|i-j|=1} (\mathrm{e}^{t_i - t_j} - 1)$$

(sum over nearest neighbors) and $\mathrm{e}^{-t}$ is the diagonal matrix with $(\mathrm{e}^{-t})_{jj} = \mathrm{e}^{-t_j}$. In Appendix B we establish the sum rule $\varepsilon \sum_{y\in\Lambda} C_{xy} = 1$, reflecting conservation of probability for the quantum dynamics generated by a Hamiltonian $H$.

Note that if $t$ were bounded, then $D(t)$ (given by (1.1)) would be uniformly elliptic and we could establish good diffusive bounds on the two-point function $C$ (1.9). However, Theorems 1 and 2 only say that large field values are unlikely. To get optimal bounds on $C$ we would need to prove uniform ellipticity on a percolating set. The set on which $|t_j + t_{j'}| < M$, is presumably a percolating set but this does not readily follow from our estimates.

Our next theorem states a quasi-diffusive estimate on $C$. More precisely let $G_0 = (-\beta\Delta + \varepsilon)^{-1}$ be the Green's function for the discrete Laplacian (with a regularization term $\varepsilon$) and $\tilde{G}_0 = (-\beta\Delta + \varepsilon/2)^{-1}$. In 3 dimensions $G_\varepsilon(x,y) \leq \beta^{-1}(1+|x-y|)^{-1}$ (and the same is true for $\tilde{G}_0$). Then we have

**Theorem 3** *Let $f : \Lambda \to \mathbb{R}$ be non-negative. Then assuming the hypotheses of Theorems 1 and 2 we have*

$$\frac{1}{K'}[\tilde{f}; G_0 \tilde{f}] \leq [f; Cf] = \sum_{ij} C_{ij}\, f(i) f(j) \leq K[f; \tilde{G}_0 f] , \tag{1.11}$$

*where $\tilde{f}(j) = (1+|j-x|^\alpha)^{-1} f(j)$, $x \in \Lambda$ is any fixed point, and $K$ and $K'$ are constants independent of $f$.*

**Remark.** In this paper we always use periodic boundary conditions on $\Lambda \subset \mathbb{Z}^3$. The distance $|x-y|$ between two points is always the distance on $\Lambda$ with periodic boundary conditions.

## 1.4 Saddle point

One may try to gain a crude understanding of the behavior of the $\mathrm{H}^{2|2}$ sigma model via a simple saddle-point analysis. Let $t^{(0)}$ be the configuration of $t = \{t_j\}$

which minimizes the effective action $F_{\beta,\varepsilon}(t)$ defined in (1.2). In Appendix A we prove that $t^{(0)}$ is unique and $t_j^{(0)} = t^*$ independent of $j$. For large $\beta$ we find

$$\text{1D:} \quad \varepsilon\, \mathrm{e}^{-t^*} \simeq \beta^{-1}, \qquad \text{2D:} \quad \varepsilon\, \mathrm{e}^{-t^*} \simeq \mathrm{e}^{-\beta}\,, \tag{1.12}$$

in one and two dimensions, respectively. Thus in 1D or 2D the saddle point depends sensitively on the regularization parameter $\varepsilon$. The value of $t^*$ suggests a strong asymmetry of the field favoring negative values of $t$. On the other hand, in 3D at low temperatures, we find $t^* = 0$ independent of $\varepsilon$. Our estimates (1.7) confirm this value by controlling fluctuations about the saddle. For $\beta$ small, in 3D, the saddle $t^*$ is again strongly $\varepsilon$-sensitive, suggesting localization.

The bias to negative values of the field $t$ is expected to be closely related to localization. Note that since $-\Delta + V(t) \geq 0$, the additional term $\varepsilon\, \mathrm{e}^{-t}$ makes $\tilde{D}_{\beta,\varepsilon}$ strictly positive at the saddle suggesting that $C_{xy}$ decays roughly like $\mathrm{e}^{-m|x-y|}$ with $m^2 = \varepsilon\, \mathrm{e}^{-t^*}/\beta = \beta^{-2}$ and $\mathrm{e}^{-\beta}$ in 1D and 2D respectively. There are important fluctuations away from this saddle but we do not expect them to spoil the exponential decay. For the 1D chain this has been proved [21].

## 1.5 Edge reinforced random walk

A number of mathematicians (Kozma, Heydenreich, Sznitmann) have noted that our random walk looks similar to a linearly edge reinforced random walk (ERRW). ERRW is a history-dependent walk which prefers to visit edges it has visited in the past. Let $n(e)$ denote the number of times the walk has visited the edge $e$. Then the probability that the walk at vertex $v$ will visit a neighboring edge $e$ equals $(a+n(e))/S_a(v)$ where $S$ is the sum of $a+n(e')$ over all the edges $e'$ touching $v$. The parameter $a$ is analogous to our $\beta$. Coppersmith and Diaconis [2] proved that this history-dependent walk can be expressed as a random walk in a random environment; see also more recent work by Merkl and Rolles [13] in which recurrence of the walk is established on a 2D lattice for small $\beta$. This is analogous to localization in our model. The environment of ERRW is very similar to the environment in $\mathrm{H}^{2|2}$. In fact, both environments have nonlocal actions arising from the square root of a determinant. Although the two models do not seem to be identical, they may have similar properties.

## 1.6 Outline of the paper

The remainder of this paper is organized as follows. In the next section we give a precise definition of the full $\mathrm{H}^{2|2}$ model and introduce the horospherical coordinate system. The effective action defined in (1.2) is then derived by integration of the field $s$ and the Grassmann fields $\bar{\psi}$ and $\psi$. Section 3 provides a physical motivation for the study of this model. In Section 4 we explain the symmetries of the model and briefly discuss its perturbative renormalization group flow. The basic Ward identities we shall need are given in Section 5 and are derived in Appendix C. Section 5 ends with a rough outline of our proof and a description of the remaining sections of this paper.

**Acknowledgments.** It is our pleasure to thank S. Varadhan and J. Lebowitz for discussions and suggestions related to this paper. Very special thanks go to D. Brydges for sharing his many insights on the model and for many comments on an early version of this paper. We wish to thank the Newton Institute (Cambridge) for its suppport and hospitality during the completion of this article.

# 2 Definition of the model

We now fill in the details of the definition the $\mathrm{H}^{2|2}$ model and derive the free energy $F_{\beta,\varepsilon}(t)$ given above.

## 2.1 Full supersymmetric model

As in Section 1.2, let $\Lambda \subset \mathbb{Z}^d$ be a cube of size $L$. For each lattice site $j \in \Lambda$ we introduce a supervector $u_j \in \mathbb{R}^{3|2}$,

$$u_j = (z_j, x_j, y_j, \xi_j, \eta_j) \ , \tag{2.1}$$

with 3 real components $x_j, y_j, z_j$ and 2 Grassmann variable components $\xi_j, \eta_j$ . We then define an inner product on $\mathbb{R}^{3|2}$ by

$$(u, u') = -zz' + xx' + yy' + \xi\eta' - \eta\xi' \tag{2.2}$$

and constrain $u_j$ by the quadratic equation

$$\forall j \in \Lambda \ : \quad (u_j \, , u_j) = -1 \ , \tag{2.3}$$

which is solved by

$$z_j = \pm\sqrt{1 + x_j^2 + y_j^2 + 2\xi_j\eta_j} \ . \tag{2.4}$$

In the following we take the positive square root for all $j \in \Lambda$ . This singles out a choice of connected subspace, $\mathrm{H}^{2|2}$, parametrized by two bosonic variables $x_j, y_j$ and two fermionic variables $\xi_j, \eta_j$ .

On the product space $(\mathrm{H}^{2|2})^{|\Lambda|}$ we introduce a 'measure' (more accurately, a Berezin superintegration form)

$$D\mu_\Lambda = \prod_{k\in\Lambda} (2\pi)^{-1} dx_k dy_k \, \partial_{\xi_k}\partial_{\eta_k} \circ (1 + x_k^2 + y_k^2 + 2\xi_k\eta_k)^{-1/2} \ . \tag{2.5}$$

We use the notation $\partial_\xi \equiv \partial/\partial\xi$ for the partial derivative. The statistical measure then is of the Gibbs form $D\mu_\Lambda \, \mathrm{e}^{-A_{\beta,\varepsilon}}$ with action

$$\begin{aligned} A_{\beta,\varepsilon} &= \frac{\beta}{2} \sum_{i,j} J_{ij} \, (u_i - u_j \, , u_i - u_j) + \varepsilon \sum_{k\in\Lambda} (z_k - 1) \\ &= \beta \sum_{i,j} J_{ij} \, (z_i z_j - (1 + x_i x_j + y_i y_j + \xi_i\eta_j - \eta_i\xi_j)) + \varepsilon \sum_{k\in\Lambda} (z_k - 1) \ . \end{aligned} \tag{2.6}$$

Here $J_{ij} = 1$ if $i, j$ are nearest neighbors (NN) and $J_{ij} = 0$ otherwise. As will be discussed later, the action $A_{\beta,0}$ is $\mathrm{SO}(1,2)$-invariant. The $\varepsilon$-term breaks this noncompact symmetry and makes the integral $\int D\mu_\Lambda \, \mathrm{e}^{-A_{\beta,\varepsilon}}$ converge.

## 2.2 Horospherical coordinates

As with [18], it is very helpful to switch to horospherical coordinates — it is only in this coordinate system that we can obtain the probabilistic interpretation of Section 1.2. We thus use the following parametrization of the supermanifold:

$$x = \sinh t - \mathrm{e}^t \left(\tfrac{1}{2}s^2 + \bar{\psi}\psi\right) \ , \quad y = \mathrm{e}^t s \ , \quad \xi = \mathrm{e}^t \bar{\psi} \ , \quad \eta = \mathrm{e}^t \psi \ , \tag{2.7}$$

where $t$ and $s$ range over the real numbers. Note that $(t, s\,; \bar{\psi}, \psi)$ are globally defined coordinates and

$$(t, s\,; \bar{\psi}, \psi) = (0, 0; 0, 0) \Leftrightarrow (x, y\,; \xi, \eta) = (0, 0; 0, 0) \ .$$

The expression for the action in them is

$$A_{\beta,\varepsilon} = \beta \sum\nolimits_{(ij)} (S_{ij} - 1) + \varepsilon \sum\nolimits_{k\in\Lambda} (z_k - 1) \ , \tag{2.8}$$

where $(ij)$ are NN pairs and

$$S_{ij} = B_{ij} + (\bar{\psi}_i - \bar{\psi}_j)(\psi_i - \psi_j)\, \mathrm{e}^{t_i + t_j} \ , \tag{2.9}$$

$$B_{ij} = \cosh(t_i - t_j) + \tfrac{1}{2}(s_i - s_j)^2\, \mathrm{e}^{t_i + t_j} \ , \tag{2.10}$$

$$z_k = \cosh t_k + \left(\tfrac{1}{2}s_k^2 + \bar{\psi}_k \psi_k\right) \mathrm{e}^{t_k} \ . \tag{2.11}$$

We also need the expression for the measure $D\mu_\Lambda$ in horospherical coordinates. By applying Berezin's transformation formula [1] for changing variables in a (super-)integral, one finds that

$$D\mu_\Lambda = \prod\nolimits_{j\in\Lambda} (2\pi)^{-1} \mathrm{e}^{-t_j}\, dt_j ds_j\, \partial_{\bar{\psi}_j} \partial_{\psi_j} \ . \tag{2.12}$$

For any function $f$ of the lattice field variables $\{t_j\,, s_j\,, \bar{\psi}_j\,, \psi_j\}_{j\in\Lambda}$ we now define its expectation as

$$\langle f \rangle_{\beta,\varepsilon} = \int D\mu_\Lambda\, \mathrm{e}^{-A_{\beta,\varepsilon}} f \ , \tag{2.13}$$

whenever this integral exists.

## 2.3 Effective bosonic field theory

Since the action (2.8) is quadratic in the fields $\bar{\psi}$, $\psi$, and $s$, each with covariance $D_{\beta,\varepsilon}(t)^{-1}$, we know from standard free-field calculus that integration over $s$ yields a factor of $\mathrm{Det}^{-1/2}(D_{\beta,\varepsilon}(t))$ while integration over $\bar{\psi}, \psi$ yields $\mathrm{Det}\, D_{\beta,\varepsilon}(t)$. By performing these integrations, we arrive at the nonlocal free energy functional $F_{\beta,\varepsilon}(t)$ given by (1.2). Moreover, the basic two-point functions are

$$\begin{aligned} \left\langle s(v)^2 \right\rangle &= +\left\langle [v\,; D_{\beta,\varepsilon}(t)^{-1} v] \right\rangle \ , \\ \left\langle \bar{\psi}(v)\psi(v) \right\rangle &= -\left\langle [v\,; D_{\beta,\varepsilon}(t)^{-1} v] \right\rangle \ , \end{aligned} \tag{2.14}$$

where

$$s(v) = \sum\nolimits_{j\in\Lambda} s_j\, v(j)\ , \quad \psi(v) = \sum\nolimits_{j\in\Lambda} \psi_j\, v(j)\ ,$$

and the expectations on the left-hand and right-hand side are defined by (2.13) and (1.5), respectively. We will often use the formula (2.14) as well as its generalization

$$\big\langle \mathrm{e}^{\sum_{\lambda=1}^{n} \bar\psi(v_\lambda)\psi(v_\lambda)} \big\rangle = \big\langle \mathrm{Det}(1-\mathcal{A}) \big\rangle\ , \tag{2.15}$$

where $\mathcal{A}$ is the $n\times n$ matrix given by

$$\mathcal{A}_{\lambda\lambda'}(t) = [v_\lambda\,; D_{\beta,\varepsilon}(t)^{-1} v_{\lambda'}]\ . \tag{2.16}$$

**Remark 2.1.** If the Grassmann fields $\bar\psi, \psi$ were absent, then $\mathrm{Det}^{1/2}$ in (1.2) would be replaced by $\mathrm{Det}^{-1/2}$ (and $\prod_k \mathrm{e}^{-t_k} dt_k$ by $\prod_k \mathrm{e}^{t_k} dt_k$) and $Z_\Lambda$ would be the partition function of the hyperbolic sigma model studied in [18].

**Remark 2.2.** If we integrate only over the fields $\bar\psi, \psi$ (but not over $s$) we produce a positive integrand depending on $t$ and $s$. The square root of the determinant is then replaced by $\mathrm{Det}\, D_{\beta,\varepsilon}(t) > 0$.

**Remark 2.3.** The logarithm of $\mathrm{Det}\, D_{\beta,\varepsilon}(t)$ is convex in $t$.

**Proof (D. Brydges).** By the matrix tree theorem we have

$$\mathrm{Det}\, D_{\beta,\varepsilon}(t) = \sum_{\mathcal{F}} \beta^{|\Lambda|-|R|}\, \varepsilon^{|R|} \prod_{\ell\in\mathcal{F}} \mathrm{e}^{t_{j_\ell}+t_{j'_\ell}} \prod_{k\in R} \mathrm{e}^{t_k}\ , \tag{2.17}$$

where $\mathcal{F}$ denotes the spanning rooted forests, $R$ the set of roots, $|R|$ the cardinality of this set, and $\ell = (j_\ell, j'_\ell)$ denotes an edge in the forest. The proof is now immediate since any positive sum of exponentials in $t$ is log convex.

Note that the logarithm of $\mathrm{Det}\, D_{\beta,\varepsilon}(t)$ competes with the other factor, $\mathrm{e}^{-\beta\sum_{(ij)}\cosh(t_i-t_j)}$, which is log concave.

# 3 Microscopic origin of the model

In this subsection we use the language and heuristic ideas of physics to sketch the origin of our field theory model from a microscopic model of disorder. Consider real symmetric random band matrices, $H$, say with independent Gaussian distributed entries, of band width $W$ in $d$ dimensions. (Such a band matrix model possesses a time-reversal symmetry and belongs to symmetry class $A$I – traditionally referred to as the Wigner-Dyson class of orthogonal symmetry – of the 10-way classification of disordered fermion systems [10]).

Now suppose that we wish to compute the disorder average of

$$\sqrt{\mathrm{Det}(E+\mathrm{i}\eta-H)/\mathrm{Det}(E+\mathrm{i}\varepsilon-H)} \times \big|(E+\mathrm{i}\varepsilon-H)^{-1}(x,y)\big|^2 \tag{3.1}$$

for real energy parameters $E$ and $\varepsilon, \eta > 0$. The unconventional feature here is that the square $|(E + \mathrm{i}\varepsilon - H)^{-1}(x, y)|^2$ of the Green's function is weighted by the square root of a ratio of one determinant taken at energy $E + \mathrm{i}\varepsilon$ and another one at energy $E + \mathrm{i}\eta$. Although one might think that the presence of these extra factors complicates the problem, quite the opposite is true; it will actually lead to simplifications when $\eta$ is taken to be large.

First of all, the combination

$$(E + \mathrm{i}\varepsilon - H)^{-1}(x, y)\ \mathrm{Det}^{-1/2}(E + \mathrm{i}\varepsilon - H)$$

can be generated by Gaussian integration over a single real boson field, $\phi_1^+$ . Second, writing the complex conjugate $\overline{(E + \mathrm{i}\varepsilon - H)^{-1}(x, y)}$ of the Green's function as a Gaussian integral requires two real boson fields $\phi_\alpha^-$ and two anticommuting fields $\psi_\alpha^-$ $(\alpha = 1, 2)$. Third, to express the square root of $\mathrm{Det}(E + \mathrm{i}\eta - H)$ as a Gaussian integral, we need another real boson $\phi_2^+$ and two more anticommuting fields $\psi_\alpha^+$. Altogether, we then have four bosonic fields $\phi_\alpha^\sigma$ and four fermionic fields $\psi_\alpha^\sigma$ $(\sigma = \pm,\ \alpha = 1, 2)$.

Now assume for the moment that $\eta = \varepsilon$, in which case the two determinants in (3.1) cancel each other. If the band width $W$ is large enough, then the standard steps of disorder averaging followed by Hubbard-Stratonovich transformation and elimination of the massive modes, take us to Efetov's nonlinear sigma model for systems with orthogonal symmetry (class $A$I).

Physically speaking, the order parameter fields of retarded (+) and advanced (−) type acquire different expectation values:

$$\langle \phi_\alpha^\sigma \phi_\beta^\sigma \rangle = \delta_{\alpha\beta} \langle G^\sigma \rangle\ , \quad \langle \psi_\alpha^\sigma \psi_\beta^\sigma \rangle = \epsilon_{\alpha\beta} \langle G^\sigma \rangle \qquad (\sigma = \pm\ ;\ \alpha, \beta = 1, 2)\ ,$$

where we are using the abbreviations $\langle \phi_\alpha^+(x) \phi_\beta^+(x) \rangle = \langle \phi_\alpha^+ \phi_\beta^+ \rangle$,

$$G^+ = (E + \mathrm{i}\varepsilon - H)^{-1}(x, x)\ , \quad G^- = \overline{G^+}\ ,$$

and $\epsilon_{\alpha\beta} = -\epsilon_{\beta\alpha}$ is the antisymmetric tensor for two degrees of freedom. In the region of nonzero average density of states, where $\langle G^+ \rangle \neq \langle G^- \rangle$, these expectation values break a continuous symmetry of the Gaussian integrand at $\varepsilon = 0$. The components of Efetov's sigma model field have the physical meaning of being the Goldstone modes associated with this broken symmetry. There are 4 bosonic Goldstone modes due to the symmetry breaking $\langle \phi_\alpha^+ \phi_\alpha^+ \rangle \neq \langle \phi_\beta^- \phi_\beta^- \rangle$ and four more such modes due to $\langle \psi_1^+ \psi_2^+ \rangle = -\langle \psi_2^+ \psi_1^+ \rangle$ not being equal to $\langle \psi_1^- \psi_2^- \rangle = -\langle \psi_2^- \psi_1^- \rangle$. There also exist 8 fermionic Goldstone modes due to the breaking of the odd symmetries connecting the boson-boson sector $\langle \phi_1^\sigma \phi_1^\sigma \rangle = \langle \phi_2^\sigma \phi_2^\sigma \rangle$ with the fermion-fermion sector $\langle \psi_1^\tau \psi_2^\tau \rangle = -\langle \psi_2^\tau \psi_1^\tau \rangle$ of opposite type $\tau = -\sigma$. All these modes organize into a supermanifold with tangent space $\mathbb{R}^{8|8}$ over a symmetric space $(\mathrm{H}^2 \times \mathrm{H}^2) \times \mathrm{S}^4$.

Now let $\eta \gg \varepsilon > 0$, so that the two determinants in the expression (3.1) no longer cancel. The difference $\eta - \varepsilon \approx \eta$ then acts as a mass term for the Goldstone modes connecting the advanced sector (−) with the $\eta$ retarded sector $\langle \phi_2^+ \phi_2^+ \rangle = \langle \psi_1^+ \psi_2^+ \rangle = -\langle \psi_2^+ \psi_1^+ \rangle$. By a Thouless-type argument, these massive

Goldstone modes do not affect the renormalized physics at length scales much greater than the length $L'$ determined by the equation

$$\eta = 2\pi\hbar D/L'^2 \ ,$$

where $D \propto W^2$ is the bare diffusion constant of the system.

Thus at large length scales $L \gg L'$ we may simply drop the massive Goldstone modes from the theory or, in a more careful treatment, integrate them out perturbatively. What we are left with, then, are the $2+2=4$ massless bosonic and fermionic Goldstone modes connecting the retarded component $\langle \phi_1^+ \phi_1^+ \rangle$ of the order parameter with its four components $\langle \phi_1^- \phi_1^- \rangle = \langle \phi_2^- \phi_2^- \rangle = \langle \psi_1^- \psi_2^- \rangle = -\langle \psi_2^- \psi_1^- \rangle$ in the advanced sector. These four residual Goldstone modes organize into a supermanifold with tangent space $\mathbb{R}^{2|2}$ and base manifold $\mathrm{H}^2$ — we thus arrive at the field space $\mathrm{H}^{2|2}$ of the model we are going to study.

# 4 Symmetries and their consequences

As an effective theory derived by reduction from an underlying sigma model, the statistical mechanics problem posed by (1.1)–(1.5) enjoys a number of symmetries. First among these is a hidden supersymmetry which ensures that the partition function is always equal to unity,

$$Z_\Lambda(\beta, \varepsilon) = 1 \ ,$$

independent of the inverse temperature $\beta$ and regularization parameter $\varepsilon$. Thus the reduced statistical measure $\mathrm{e}^{-F_{\beta,\varepsilon}} d\mu_\Lambda$ can be regarded as a probability measure, and the physical observables of the model are given as expectations

$$\langle f \rangle = \int f \, \mathrm{e}^{-F_{\beta,\varepsilon}} d\mu_\Lambda \ .$$

In the following subsection we provide some background to the normalization property $Z_\Lambda(\beta, \varepsilon) = 1$.

## 4.1 $Q$-symmetry

We start by observing that, for any $\varepsilon$, the full action $A_{\beta,\varepsilon}$ defined in (2.6) is invariant under transformations that preserve the short inner product

$$x_i x_j + y_i y_j + \xi_i \eta_j - \eta_i \xi_j \tag{4.1}$$

for all $i, j \in \Lambda$. Such transformations are given, at the infinitesimal level, by even and odd derivations (i.e., first-order differential operators) with the property that they annihilate the expression (4.1) for all $i, j$ and their coefficients are linear functions of the coordinates $x_k, y_k, \xi_k, \eta_k$. These differential operators form a representation of the orthosymplectic Lie superalgebra $\mathfrak{osp}_{2|2}$. An important example of an odd operator $Q \in \mathfrak{osp}_{2|2}$ is

$$Q = \sum\nolimits_{j \in \Lambda} \left( x_j \partial_{\eta_j} - y_j \partial_{\xi_j} + \xi_j \partial_{x_j} + \eta_j \partial_{y_j} \right) \ . \tag{4.2}$$

Since $\prod_j dx_j dy_j\, \partial_{\xi_j} \partial_{\eta_j}$ is the Berezin superintegration form given by the inner product (4.1), it is immediate that $D\mu_\Lambda$ is $\mathfrak{osp}_{2|2}$-invariant, which implies that $\int D\mu_\Lambda\, Qf = 0$ whenever the function $f$ is differentiable and $Qf$ is integrable.

For present use, let us record here the explicit expression for the $\mathfrak{osp}_{2|2}$ generator $Q$ in horospherical coordinates: a straightforward computation starting from (4.2) gives $Q = \sum_{j\in\Lambda} q_j$ with single-site generator (index $j$ omitted)

$$q = \bar\psi\partial_t + (\psi - s\bar\psi)\partial_s - s\partial_{\bar\psi} + \tfrac{1}{2}(1 - \mathrm{e}^{-2t} - s^2 - 4\bar\psi\psi)\partial_\psi \ . \tag{4.3}$$

Now consider any differentiable integrand $f$ which is invariant by $Q$, i.e., $Qf = 0$. This invariance property has strong consequences for the integral of $f$ (if it exists): in Appendix C, Proposition 2, we prove that the integral of such $f$ equals $f$ evaluated on the zero-field configuration (i.e., on $t_j = s_j = \bar\psi_j = \psi_j = 0$ or equivalently, $x_j = y_j = \xi_j = \eta_j = 0$, for all $j \in \Lambda$):

$$\int D\mu_\Lambda\, f = f(o) \ . \tag{4.4}$$

The idea of the proof is easy to state: one shows that the integral of $f$ remains unchanged by the replacement $f \to \mathrm{e}^{-\tau h} f$ with $h = \sum_{j\in\Lambda}(x_j^2 + y_j^2 + 2\xi_j\eta_j)$ and $\tau \geq 0$, and then deduces the result (4.4) by sending the deformation parameter $\tau \to +\infty$ to localize the integral at the zero-field configuration.

Using the explicit expression (4.3) it is easy to check that the action $A_{\beta,\varepsilon}$ is $Q$-invariant. Since the differential operator $Q$ is of first order, one directly infers the relation $Q\,\mathrm{e}^{-A_{\beta,\varepsilon}} = 0$. Therefore, as a particular consequence of (4.4) and $A_{\beta,\varepsilon}(o) = 0$ it follows that the partition function equals unity,

$$Z_\Lambda(\beta,\varepsilon) = \int D\mu_\Lambda\, \mathrm{e}^{-A_{\beta,\varepsilon}} = \mathrm{e}^{-A_{\beta,\varepsilon}(o)} = 1 \ , \tag{4.5}$$

for all values of $\beta \geq 0$ and $\varepsilon > 0$.

Further consequences of (4.4) will be elaborated below.

## 4.2 Hyperbolic symmetry

While $Q$ is a symmetry of our action $A_{\beta,\varepsilon}$ for *all* values of $\varepsilon$, further symmetries emerge in the limit of vanishing regularization $\varepsilon \to 0+$. Relegating a more detailed discussion to Appendix B, we here gather the crucial facts.

The model (1.1)–(1.5) for $\varepsilon \to 0+$ acquires a global symmetry by the Lorentz group $\mathrm{SO}(1,2)$ – the isometry group of the hyperbolic plane $\mathrm{H}^2$ viewed as a noncompact symmetric space $\mathrm{H}^2 \simeq \mathrm{SO}(1,2)/\mathrm{SO}(2)$. This global symmetry entails a number of conserved currents and associated Ward identities. Of these let us mention here the most important one,

$$\sum_{y\in\Lambda} \big\langle \mathrm{e}^{t_x + t_y} D_{\beta,\varepsilon}(t)^{-1}(x,y) \big\rangle = \frac{1}{\varepsilon} \ , \tag{4.6}$$

which is the sigma model version of the quantum sum rule

$$\sum_{y\in\Lambda} \left\langle \left|(E-\mathrm{i}\varepsilon-H)^{-1}(x,y)\right|^2 \right\rangle = \frac{1}{\varepsilon}\left\langle \mathrm{Im}\,(E-\mathrm{i}\varepsilon-H)^{-1}(x,x)\right\rangle = \frac{\pi}{\varepsilon}\rho(E)\ ,$$

where $\rho(E)$ is the mean local density of states. In the sigma model approximation one sets $\pi\rho(E)=1$. The above relation reflects the unitarity of the quantum theory. Its classical interpretation is conservation of probability.

Notice that the right-hand side of (4.6) diverges in the limit of vanishing regularization $\varepsilon\to 0$. For an infinite lattice $\Lambda$ there exist two principal scenarios [12] by which to realize this divergence. In the first one, the correlation function $C_{xy}=\langle \mathrm{e}^{t_x+t_y}D^{-1}_{\beta,\varepsilon}(t)(x,y)\rangle$, while bounded in $\varepsilon$, becomes of long range and thus ceases to be summable in the limit $\varepsilon\to 0$. In this case the $\mathrm{SO}(1,2)$ symmetry is spontaneously broken and the system is in a phase of extended states. On the other hand, $C_{xy}$ may already diverge for any *fixed* pair of lattice sites $x,y$, signaling strong field fluctuations and restoration of the noncompact symmetry $\mathrm{SO}(1,2)$ as $\varepsilon\to 0$. Exponential decay of $C_{xy}$ with distance $|x-y|$ then corresponds to exponential localization of the energy eigenstates. Thus the question of extended versus localized states of the disordered quantum system translates to the question of the Lorentzian symmetry $\mathrm{SO}(1,2)$ of the statistical mechanical model with free energy (1.1) being spontaneously broken or not.

At this stage, a remark is called for. Niedermaier and Seiler have recently shown [15, 16] for a large class of sigma models that if the symmetry group of the sigma model is non-amenable – this includes in particular the case of the Lorentz group $\mathrm{SO}(1,2)$ – then spontaneous symmetry breaking occurs in *all* dimensions $d\geq 1$ and for all $\beta>0$. It must therefore be emphasized that, although our sigma model does acquire the non-amenable symmetry $\mathrm{SO}(1,2)$ in the limit $\varepsilon\to 0$, it does *not* belong to the class of models where the arguments of [15, 16] apply. (The culprit is the nonlocal part of the free energy due to integration over the Grassmann fields.) In fact, the non-amenable symmetry $\mathrm{SO}(1,2)$ of our model is known to be unbroken in $d=1$. This follows from the work of [21] where the conductance of the one-dimensional system was shown to exhibit exponential decay with increasing length of the system.

## 4.3 Perturbative renormalization group

We now sketch a perturbative result from Wilsonian renormalization theory by which our model is expected to be in a symmetry-unbroken phase also for $d=2$ and all values of the inverse temperature $\beta$, and thus to exhibit Anderson localization of all electronic states.

This result follows from Friedan's work [8] on renormalization for the general class of nonlinear sigma models. According to it, the RG flow of the temperature $T=\beta^{-1}$ with increasing renormalization scale $a$ is given by

$$a\frac{dT}{da} = (2-d)T + RT^2 + \mathcal{O}(T^3)\ , \tag{4.7}$$

where $R$ is the target space curvature – more precisely, the multiplicative constant $R$ by which the Ricci tensor of the target space differs from its metric tensor. For both the $\mathrm{H}^{2|2}$ model and Efetov's sigma model of class $A$I a quick computation shows the curvature $R$ to be positive. In contrast, $R = 0$ and $R < 0$ for Efetov's sigma models of class $A$ (broken time-reversal symmetry) and class $A$II (spin-orbit scattering), respectively.

According to (4.7), a positive value of $R$ implies that a small initial value of the temperature $T$ increases under renormalization in dimension $d = 2$. By extrapolation, one therefore expects the existence of a mass gap (or, equivalently, localization of all states) in this case. For the localization length $\xi = \xi(a, T(a))$, which is a physical observable and hence a renormalization group invariant, one obtains the formula

$$\xi \propto a\, \mathrm{e}^{1/(R\,T)} \quad (d = 2)$$

by direct integration of the RG equation (4.7).

In dimension $d = 3$, equation (4.7) predicts the localizing tendency of positive target space curvature to become irrelevant at small enough temperatures and hence the RG flow to be attracted to the fixed point $T = 0$ corresponding to extended states. As was remarked above, the Lorentzian symmetry $\mathrm{SO}(1, 2)$ is spontaneously broken at this fixed point.

With increasing temperature $T$ (or decreasing field stiffness $\beta$) the model in $d = 3$ is expected to undergo an Anderson-type transition to the phase of unbroken symmetry. This phase transition was studied numerically in [4], where the critical value of $\beta$ was found to be $\beta_c \approx 0.04$. The transition has also been investigated in detail using the Migdal-Kadanoff renormalization scheme [5].

# 5 Ward identities and outline of proof

In order to control fluctuations of the field $t$ at low temperatures $T = \beta^{-1}$ we rely on a family of Ward identities due to the internal supersymmetries of the model. These Ward identities are naturally expressed in terms of both the real variables $t_j$ , $s_j$ and the Grassmann variables $\bar{\psi}_j$ , $\psi_j$. In order to obtain probabilistic information we integrate out the Grassmann variables using (2.14) and (2.15), thereby producing a Green's function.

As was already mentioned, our partition function always equals unity even when the temperature varies in space. By using this fact, we show that gradients of the field $t$ between neighboring sites are strongly suppressed for small $T$. There also exist Ward identities at larger scales, and information may be extracted from them by using information on previous length scales.

In addition to Ward identities, there are two other crucial ingredients of our proof. The first one is a basic estimate on Green's functions which are non-uniformly elliptic. The second one is the use of SUSY characteristic functions, which help to control large-scale field fluctuations.

A more detailed outline of our proof is given below, where the notation and the needed Ward identities are explained. Once the Ward identities are established, most of our proof is very classical.

## 5.1 Ward identities due to $Q$-symmetry

We recall the formula (4.4) for the integral of a $Q$-invariant function. It is easy to check that the functions $S_{ij}$ and $z_k$ given in (2.9) and (2.11) satisfy the invariance conditions $QS_{ij} = 0$ and $Qz_k = 0$. Therefore, using $S_{ij}(o) = 1$ and $z_k(o) = 1$ we have the identity

$$\int D\mu_\Lambda \, \mathrm{e}^{-\beta \sum_{x,y} J_{xy}(S_{xy}-1) - \sum_{x\in\Lambda} \varepsilon_x (z_x - 1)} = 1 \tag{5.1}$$

for all values of $\beta \geq 0$, $J_{xy} \geq 0$, and $\varepsilon_x > 0$. Note that in order for this statement to be true, $J_{xy}$ does not have to be nearest neighbor. Moreover, for $m \in \mathbb{R}$ and any pair $x, y \in \Lambda$ we have

$$1 = \langle S_{xy}^m \rangle_{\beta,\varepsilon} = \big\langle B_{xy}^m + m B_{xy}^{m-1} (\bar\psi_x - \bar\psi_y)(\psi_x - \psi_y) \, \mathrm{e}^{t_x + t_y} \big\rangle_{\beta,\varepsilon} \ , \tag{5.2}$$

where the expectation $\langle \cdot \rangle_{\beta,\varepsilon}$ was defined in (2.13), and we used the nilpotency $(\psi_x - \psi_y)^2 = (\bar\psi_x - \bar\psi_y)^2 = 0$. By integrating over the Grassmann fields $\bar\psi$ and $\psi$ as in (2.14) we obtain our basic identity,

$$1 = \langle S_{xy}^m \rangle_{\beta,\varepsilon} = \langle B_{xy}^m \big(1 - m G_{xy}\big) \rangle \ . \tag{5.3}$$

The last expectation is taken with respect to the effective action for the fields $t$ and $s$, and the Green's function $G_{xy}$ is

$$G_{xy} = \frac{\mathrm{e}^{t_x + t_y}}{B_{xy}} \big[ (\delta_x - \delta_y); \, D_{\beta,\varepsilon}(t)^{-1} (\delta_x - \delta_y) \big]_\Lambda \ . \tag{5.4}$$

More generally if $(x_i, y_i)$ are $n$ pairs of points, then

$$1 = \Big\langle \prod_{i=1}^n S_{x_i y_i}^m \Big\rangle_{\beta,\varepsilon} = \Big\langle \prod_{i=1}^n B_{x_i y_i}^m \mathrm{Det}\big(1 - m\mathcal{G}\big) \Big\rangle \ , \tag{5.5}$$

where $\mathcal{G}$ is an $n \times n$ matrix of Green's functions

$$\mathcal{G}_{ij} = [g_i \, ; D_{\beta,\varepsilon}(t)^{-1} g_j]_\Lambda \tag{5.6}$$

and

$$g_i = B_{x_i y_i}^{-1/2} \, \mathrm{e}^{(t_{x_i} + t_{y_i})/2} (\delta_{x_i} - \delta_{y_i}) \ . \tag{5.7}$$

The matrix $\mathcal{G}$ is real symmetric and positive. It will be important later that we are choosing $g_i$ to be orthogonal to the zero mode (i.e., the constant functions).

## 5.2 Outline of proof

Our proof of Theorem 1 relies on the Ward identity (5.3), (5.5) and an induction on length scales. The basic idea is quite simple: suppose $m > 0$ and we had a uniform bound $|G_{xy}| \leq C/\beta < 1/m$ on the Green's function (5.4), for all configurations of $t$. Then we could conclude from (5.3) that

$$\langle \cosh^m (t_x - t_y) \rangle \leq \langle B_{xy}^m \rangle \leq (1 - mC/\beta)^{-1} \ , \tag{5.8}$$

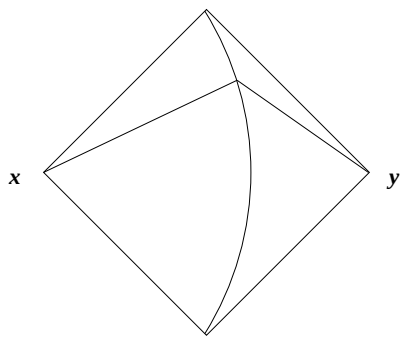


Figure 1: "diamond" region: a double cone in 3 dimensions

and this would imply Theorem 1.

In Section 6 we prove that if $|x-y|=1$ then indeed $0 \le G_{xy} \le 1/\beta$, and we establish an even stronger version of (5.8). This proves that nearest neighbor fluctuations of the field $t$ are very unlikely for large $\beta$ (see Lemma 4).

For distances $|x-y|>1$, however, there is no uniform bound on $G_{xy}$. In Section 7 we study the Green's function (5.4) and establish sufficient conditions on the field $t$ to obtain the desired bound on $G_{xy}$. In 3D these conditions are roughly given as follows (where $|j-x| \ge 1$):

$$\cosh(t_j - t_x) \le B_{jx} \le a\,|j-x|^{\alpha}\ , \quad 0<\alpha<1/2\ , \tag{5.9}$$

and the same for $\cosh(t_j - t_y)$. The number $a$ is a constant, say $a>10$. It will turn out that these estimates are needed only for the sites $j$ in a 3D diamond-type region, $R_{xy}$, containing $x$ and $y$; see Fig. 1. Notice that since the exponent $\alpha$ is positive, we are allowing larger fluctuations at larger scales. The probability that such a condition is violated will be shown to be small by induction.

Section 8 uses the conditions described above to prove conditional estimates on the fluctuations of the field $t$ at all scales. These conditions are initially expressed in terms of $Q$-invariant characteristic functions $\chi$. Later we show that the nilpotent (or Grassmann) part of $\chi$ is not important, so we may think of $\chi$ in the usual classical sense.

The remaining problem is to obtain unconditional estimates on the fluctuations and thereby prove Theorem 1. This is first done for short scales in Section 10. For larger scales we use induction. Our induction hypothesis is formulated in Section 11. Roughly speaking it asserts that

$$\left\langle \prod\nolimits_{i=1}^{n} B^{m}_{x_i y_i} \right\rangle \le 2^n \tag{5.10}$$

holds under the assumption that the diamond-type regions $R_{x_i y_i}$ (Fig. 1) associated with $i=1,\dots,n$ have disjoint interiors. The induction is in $\ell$, defined as the maximal separation $|x_i - y_i|$ in the product over $i=1,\dots,n$. For $\ell=1$ this hypothesis was verified in Section 6.

Section 12 contains the technical core of our paper. There we prove unconditional estimates on the fluctuations and thus obtain Theorem 1. The main idea is to consider a site $b$ in $R_{xy}$ closest to $x$ or $y$ such that condition (5.9) is violated for $j = b$. We shall then prove by induction that the probability for such an event to occur is small. The inequality $B^m_{xy} < 2^m B^m_{xc} B^m_{cy}$ (see Lemma 2 below) is used for a point $c$ near $b$. Since the distances $|x - c|$ and $|c - y|$ are less than $|x - y|$, induction can be applied. The factor $2^m$ is offset by the small probability of the event when $\beta$ is large.

Theorem 2 is proved in Section 13. Here we must estimate the contribution of the zero mode and at this stage $\varepsilon > 0$ plays a key role. Finally, Theorem 3 follows from the estimates of Theorem 2; its proof is given in Section 14.

## 5.3 Two simple lemmas

We conclude this section with two simple lemmas which will be frequently used below. The first lemma is useful for estimates on Green's functions. To state it, let $V$ be a finite-dimensional Euclidean vector space with scalar product $[\,;\,]_V$.

**Lemma 1** *Let $M : V \to V$ be a positive real symmetric operator. Then for any set of $n$ vectors $v_i \in V$ we have*

$$M - \sum_{i=1}^{n} v_i \, [v_i \, ; \, \cdot\,]_V \geq 0 \tag{5.11}$$

*if and only if the symmetric $n \times n$ matrix $K$ with matrix elements*

$$K_{ij} = [v_i \, ; M^{-1} v_j]_V \tag{5.12}$$

*satisfies* $0 \leq K \leq \mathrm{Id}$.

**Proof.** Letting $w_i = M^{-1/2} v_i \in V$ we observe that $K_{ij} = [w_i \, ; w_j]_V$ and (5.11) is equivalent to

$$\mathrm{Id} - \sum_{i=1}^{n} w_i \, [w_i \, ; \, \cdot\,]_V \geq 0 \, .$$

By evaluating this quadratic form at $w = \sum \lambda_i w_i$ for any real numbers $\lambda_i$ we see that (5.11) is equivalent to

$$\sum_{i,j=1}^{n} \lambda_i \lambda_j \, (K - K^2)_{ij} \geq 0 \, , \tag{5.13}$$

or $0 \leq K^2 \leq K$, from which our assertion follows. ∎

The second lemma will be used in our induction process of Section 12.

**Lemma 2** *If $B_{xy}$ , $B_{cy}$ , $B_{xc}$ are defined as in* (2.10) *then*

$$B_{xy} < 2 B_{xc} B_{cy} \, . \tag{5.14}$$

The inequality (5.14) can be verified by direct computation (proof omitted).

**Remark.** The *raison d'etre* behind (5.14) is easy to state: $B_{xy}$ has an interpretation as the hyperbolic cosine of the geodesic distance on $\mathrm{H}^2$. Therefore, if $x, y, c$ are three points on $\mathrm{H}^2$, then since the geodesic distance $\mathrm{dist}(x, y)$ is the minimal length of any curve connecting $x$ and $y$, the triple of geodesic distances satisfy the triangle inequality $\mathrm{dist}(x, y) \leq \mathrm{dist}(x, c) + \mathrm{dist}(c, y)$. Given this, the inequality (5.14) follows by taking the hyperbolic cosine of both sides and using that $\cosh(a + b) < 2\cosh(a)\cosh(b)$ holds for any two real numbers $a, b$.

# 6 Bounds on NN fluctuations

As was already mentioned, for nearest neighbor (NN) pairs we can obtain a result stronger than (5.8). Recall that we have now fixed $J_{xy} = 1$ for all $xy$ that are NN pairs, and $J_{xy} = 0$ otherwise. This fact is essential in the next lemma.

**Lemma 3** *Let $x, y$ be an NN pair and suppose that $0 < \gamma < 1$. Then*

$$\left\langle \mathrm{e}^{\beta\gamma\,(B_{xy}-1)} \right\rangle \leq (1-\gamma)^{-1} \ . \tag{6.1}$$

*More generally, if $(x_j, y_j)$, $j = 1, \ldots, n$ is a set of $n$ different NN pairs, then*

$$\left\langle \mathrm{e}^{\beta\gamma \sum_{j=1}^{n}(B_{x_j y_j}-1)} \right\rangle \leq (1-\gamma)^{-n} \ . \tag{6.2}$$

This shows that NN fluctuations are strongly suppressed.

**Remark.** Since $J_{x_j y_j} = 1$ and $\gamma < 1$ the integrals in (6.1)–(6.2) are well defined. This would not be true if $x_j, y_j$ were not NN, or if $\gamma > 1$, or if two or more NN pairs were allowed to be identical without further restrictions on the value of $\gamma$.

**Proof.** For $x, y$ an NN pair let

$$F_{xy}(j) = \mathrm{e}^{(t_x+t_y)/2}(\delta_x(j) - \delta_y(j)) \ , \tag{6.3}$$

and introduce the Green's function

$$G^0_{xy}(t) = [F_{xy}\,; D^{-1}_{\beta,\varepsilon}(t)\, F_{xy}] = B_{xy} G_{xy} \ . \tag{6.4}$$

Since $S_{xy}$ is $Q$-invariant, Proposition 2 of Appendix C implies

$$\begin{aligned} \mathrm{e}^{\beta\gamma} &= \left\langle \mathrm{e}^{\beta\gamma S_{xy}} \right\rangle = \left\langle \mathrm{e}^{\beta\gamma(B_{xy}+\bar\psi(F_{xy})\psi(F_{xy}))} \right\rangle \\ &= \left\langle \mathrm{e}^{\beta\gamma B_{xy}}(1 + \beta\gamma\, \bar\psi(F_{xy})\psi(F_{xy})) \right\rangle \\ &= \left\langle \mathrm{e}^{\beta\gamma B_{xy}}(1 - \beta\gamma\, G^0_{xy}) \right\rangle \ , \end{aligned} \tag{6.5}$$

where we have used $\psi^2 = \bar\psi^2 = 0$ and (2.14). Now from (1.1) we have

$$[v\,; D_{\beta,\,\varepsilon}(t)\, v] = \beta \sum\nolimits_{(ij)} [v\,; F_{ij}]^2 + \varepsilon \sum\nolimits_k \mathrm{e}^{t_k} v_k^2 \geq \beta\, [v\,; F_{xy}]^2 \ . \tag{6.6}$$

Therefore Lemma 1 implies that $0 \le \beta\, G^0_{xy}(t) \le 1$ for all $t$, and (6.1) follows.

Similarly, for $n > 1$ we have

$$\begin{aligned} \mathrm{e}^{n\beta\gamma} = \left\langle \mathrm{e}^{\beta\gamma \sum_{j=1}^n S_{x_j y_j}} \right\rangle &= \left\langle \mathrm{e}^{\beta\gamma \sum_{j=1}^n B_{x_j y_j}} \mathrm{e}^{\beta\gamma \sum_{j=1}^n \bar\psi(F_{x_j y_j})\psi(F_{x_j y_j})} \right\rangle \\ &= \left\langle \mathrm{e}^{\beta\gamma \sum_{j=1}^n B_{x_j y_j}} \mathrm{Det}(1-\gamma K) \right\rangle , \end{aligned} \tag{6.7}$$

where $K$ is the $n \times n$ matrix

$$K_{ij} = \beta\, [F_{x_j y_j}\, ; D_{\beta,\varepsilon}(t)^{-1}\, F_{x_i y_i}] \tag{6.8}$$

given by $n$ different NN pairs $x_i, y_i$. From (6.6) and Lemma 1 it follows that $\|K\| \le 1$. This implies $|\mathrm{Det}(1-\gamma K)| \ge (1-\gamma)^n$ and the lemma follows. ∎

As a corollary, since $1 \le B^m_{xy} \le \mathrm{e}^{m(B_{xy}-1)}$ for $m \ge 0$, we have the bound

$$\left\langle \prod_{j=1}^n B^m_{x_j y_j} \right\rangle \le (1 - m/\beta)^{-n} \le 2^n \tag{6.9}$$

for any $m$ in the range $m \le \beta/2$.

A first important consequence of Lemma 3 is the following statement.

**Lemma 4** *Let $x_j, y_j$ be a set of $n$ different nearest neighbor pairs. Then*

$$\mathrm{Prob}\left(\forall j = 1, \ldots, n :\ B_{x_j y_j} > 1 + \delta\right) \le (1-\gamma)^{-n} \mathrm{e}^{-n(\beta\gamma)\delta} \tag{6.10}$$

*for any $0 < \gamma < 1$.*

**Proof.** Let $n = 1$. By the Chebyshev inequality [1] ,

$$\mathrm{Prob}\,(B_{xy} > 1 + \delta) = \langle \chi(B_{xy} > 1 + \delta) \rangle \le \mathrm{e}^{-\beta\gamma(1+\delta)} \left\langle \mathrm{e}^{\beta\gamma\, B_{xy}} \right\rangle , \tag{6.11}$$

where $\chi(B_{xy} > 1+\delta)$ is the characteristic function for $B_{xy} > 1+\delta$ to hold. The desired inequality for $n = 1$ now follows directly from Lemma 3.

The proof for $n$ pairs is no different. ∎

# 7 Conditional estimates on Green's functions

For general $x, y$ (not NN) we do not have the option of considering $\langle \mathrm{e}^{\beta\gamma\, S_{xy}} \rangle$, as the underlying integral need not exist. Nevertheless, $\langle S^m_{xy} \rangle$ does exist and from (5.3) we have

$$1 = \left\langle B^m_{xy} \left(1 - m G_{xy}\right) \right\rangle , \tag{7.1}$$

with $G_{xy}$ defined by (5.4),

$$G_{xy} = \frac{\mathrm{e}^{t_x + t_y}}{B_{xy}} \left[ (\delta_x - \delta_y)\, ; D_{\beta,\varepsilon}(t)^{-1} (\delta_x - \delta_y) \right] .$$

[1] Actually, the Chebyshev inequality states that for any random variable $X$ with average $X_0$ we have $\mathrm{Prob}[(X - X_0)^2 > a^2] \le a^{-2} \left\langle (X - X_0)^2 \right\rangle$. Here we are using the same principle.

Now, as was explained in Section 5.2, *if we knew that* $G_{xy} \leq C/\beta$ for all configurations of $t$, then we could conclude that

$$\langle B_{xy}^m \rangle \leq (1 - mC/\beta)^{-1}. \tag{7.2}$$

While we have seen that this estimate is true for $|x-y| = 1$ (with $C = 1$), it is *false* in general, as there are rare configurations with large negative $t$ surrounding $x$ or $y$. Nonetheless, in 3D we can get an upper bound on $G_{xy}$ by estimating the local 'conductance' at an edge $(ij)$ from below. This conductance is

$$\begin{aligned} A_{xy}(ij) &\equiv B_{xy}\, \mathrm{e}^{-t_x - t_y} \mathrm{e}^{t_i + t_j} \\ &\geq \tfrac{1}{2} \max\left(\mathrm{e}^{t_i + t_j - 2t_x}, \mathrm{e}^{t_i + t_j - 2t_y}\right), \end{aligned} \tag{7.3}$$

where we have used $B_{xy} \geq \cosh(t_x - t_y)$. It will suffice to estimate the expression (7.3) for NN pairs $(ij)$ in a region $R_{xy}$ which is like a 3D double cone with vertices at $x$ and $y$. Note that Neumann boundary conditions increase $G_{xy}$ and $\delta_x - \delta_y$ is orthogonal to the zero mode. We will have to require that $R_{xy}$ be essentially three-dimensional in the following sense:

**Definition 1** *A region $R_{xy} \subset \Lambda$ containing $x$ and $y$ is called $\delta$-admissible if it is connected by nearest-neighbor bonds and the two one-parameter families of intersections $R_z(r) \equiv R_{xy} \cap B_z^r$ with the ball $B_z^r$ of radius $r$ centered at $z = x, y$ satisfy*

$$\mathrm{vol}\{R_z(r)\} \geq r^3 \delta \quad \textit{for} \quad r \leq |x-y|/\sqrt{2} \qquad (z = x, y)\,.$$

*In addition we require that the following Poincaré inequality:*

$$\sum_{j \in R_z(r)} f(j)^2 \leq \mathrm{const}\; r^2 \sum_{j \in R_z(r)} (\nabla f)^2(j)\,,$$

*holds for all functions $f : \Lambda \to \mathbb{R}$ subject to the condition $\sum_{j \in R_z(r)} f(j) = 0$.*

We observe that by the choice of maximal radius $r = |x-y|/\sqrt{2}$ the scaling of volume is monitored up to the full side length of a rectangular diamond $R_{xy}$ (or a double cone $R_{xy}$, see Fig. 1) with opposite corners placed at $x$ and $y$.

In the continuum limit this definition is satisfied by a double cone obtained by rotating (around the line $\overline{xy}$ connecting $x$ and $y$) a 2D diamond with vertices on $x$ and $y$ and angle $\theta \geq \theta_0(\delta) \geq \pi/10$ (see Fig. 2a). Since we are on a lattice we may have to add a few lattice points near $x$ and $y$ to ensure connectedness (see Fig. 2b). The Poincaré inequality is straightforward to prove in such convex regions.

**Definition 2** *Given a $\delta$-admissible region $R_{xy}$, we define the regions $R_{xy}^z$ for $z = x$ and $z = y$ by*

$$R_{xy}^z = \{j \in R_{xy} \mid 1 \leq |j - z| \leq |x-y|/\sqrt{2}\}\,. \tag{7.4}$$

For the case of a diamond, $R_{xy}^x \cup R_{xy}^y = R_{xy} \setminus \{x, y\}$.

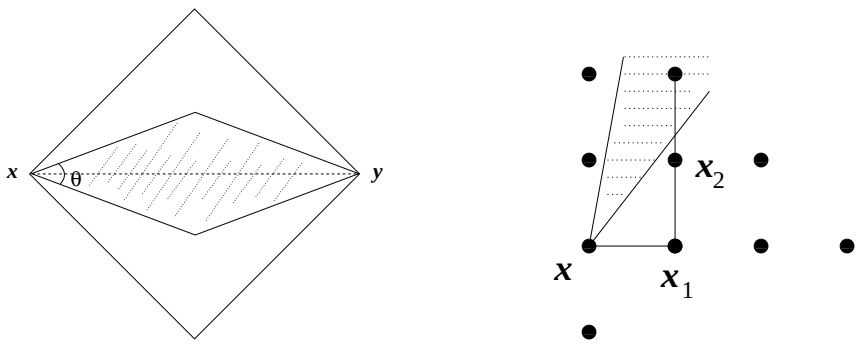


Figure 2: a) The region $R_{xy}$ in the continuum limit. b) On the lattice, the points $x_1$ and $x_2$ must be added to ensure connectedness.

**Remark 7.1.** The values of the field $t$ outside the region $R_{xy}$ are not important, as we can use Neumann boundary conditions to eliminate the exterior of $R_{xy}$. Indeed, in the subspace orthogonal to the constant functions the Laplacian on $R_{xy}$ with Neumann boundary conditions is bounded (by the Poincaré inequality) from below by some number, say $c$, times the inverse square of the linear size $L$ of $R_{xy}$. By this token, since the vector $\delta_x - \delta_y$ used in the definition of $G_{xy}$ lies in that subspace, we may utilize the bound on the inverse of the Neumann Laplacian by $c^{-1}L^2$ and in this way eventually obtain an upper bound on $G_{xy}$ (see Lemma 5).

**Remark 7.2.** From (7.3) we have

$$A_{xy}(ij)^{-1} \leq 8\cosh(t_i - t_z)\cosh(t_j - t_z) \tag{7.5}$$

for both $z = x$ and $z = y$. The main result of this section is that we can get an upper bound on $G_{xy}$ even without imposing a uniform upper bound on $A_{xy}(ij)^{-1}$, as long as some growth restriction on the fluctuations of $t$ is met for a $\delta$-admissible region $R_{xy}$ :

**Lemma 5** *Fix two constants $a > 1$ and $1/2 > \alpha > 0$. If $R_{xy}$ is a $\delta$-admissible region in the sense of Def. 1 and the statement*

$$\forall j \in R^z_{xy} \ : \quad \cosh(t_j - t_z) \leq \ a\,|j - z|^\alpha \tag{7.6}$$

*holds for both $z = x$ and $z = y$, then we have*

$$0 \leq G_{xy} \leq G^N_{xy} \leq C(a, \alpha, \delta)/\beta\ , \tag{7.7}$$

*where $C(a, \alpha, \delta)$ is some constant depending on the parameters $a$, $\alpha$ and the geometry of the region $R_{xy}$ (encoded in the parameter $\delta$). The notation $G^N$ means Neumann boundary conditions on $\partial R_{xy}$.*

**Proof.** The following is a variation on an argument presented in [18]. For each $k \in \mathbb{N}$ consider two cubes of side $2^k$ centered at $x$ and $y$. (For concreteness, imagine the edges of the two cubes to be parallel to the vector $x - y$.) Let $R_x^k, R_y^k$ denote the corresponding intersections with $R_{xy}$ and let $I_k$, $\tilde{I}_k$ be the indicator functions of $R_x^k$ and $R_y^k$, respectively, normalized so that for each $k$

$$\sum\nolimits_j I_k(j) = 1 = \sum\nolimits_j \tilde{I}_k(j) \ . \tag{7.8}$$

We observe that

$$\forall k \geq k_m \ : \quad R_x^k = R_y^k = R_{xy} \ , \tag{7.9}$$

where $k_m$ is the smallest number $k \in \mathbb{N}$ such that $2^{k-1} \geq |x - y|$. Since $R_{xy}$ is $\delta$-admissible, $R_z^k$ has the same properties as $R_z (r = 2^k)$ in Def. 1 and we therefore have

$$\operatorname{vol} R_x^k \geq 2^{kd}\delta \qquad (k \leq k_m - 1 \ , \quad d = 3) \tag{7.10}$$

for all $2^k > 10$. For $2^k \leq 10$ this is not true (see Def. 1) but the corresponding volume is no less than unity, as $R_{xy}$ is connected.

Now we express $\delta_x - \delta_y$ as a telescopic sum:

$$\delta_x - \delta_y = \sum_{k=1}^{k_m} (\rho_k - \tilde{\rho}_k) \ , \tag{7.11}$$

where $\rho_k = I_{k-1} - I_k$, $\tilde{\rho}_k = \tilde{I}_{k-1} - \tilde{I}_k$, $I_0 = \delta_x$, and $\tilde{I}_0 = \delta_y$. This sum terminates at $k_m$ because by (7.9) we have $I_k = \tilde{I}_k$ for $k \geq k_m$. Note that $\rho_k$, $\tilde{\rho}_k$ are orthogonal to the constant functions: $\sum_{j\in\Lambda} \rho_k(j) = 0 = \sum_{j\in\Lambda} \tilde{\rho}_k(j)$.

Next we put the telescopic sum to use by the following computation:

$$\begin{aligned} \big[\delta_x - \delta_y \, ; D_{\beta,\varepsilon}(t)^{-1}(\delta_x - \delta_y)\big] &= \sum_{k,l=1}^{k_m} \big[\rho_k - \tilde{\rho}_k \, ; D_{\beta,\varepsilon}(t)^{-1}(\rho_l - \tilde{\rho}_l)\big] \\ &\leq \left( \sum_{k=1}^{k_m} \Big( \big[\rho_k \, ; D_{\beta,\varepsilon}(t)^{-1}\rho_k\big]^{1/2} + \big[\tilde{\rho}_k \, ; D_{\beta,\varepsilon}(t)^{-1}\tilde{\rho}_k\big]^{1/2} \Big) \right)^2 \ , \end{aligned} \tag{7.12}$$

where the Cauchy-Schwarz inequality was employed. Hence we need to estimate $[\rho_k \, ; D_{\beta,\varepsilon}(t)^{-1}\rho_k]$ and $[\tilde{\rho}_k \, ; D_{\beta,\varepsilon}(t)^{-1}\tilde{\rho}_k]$. This is done, say for the former, by the inequality

$$[\rho_k \, ; D_{\beta,\varepsilon}(t)^{-1} \, \rho_k] \leq \|D_{R_x^k}^{-1}\| \, \|\rho_k\|_2^2 \ , \tag{7.13}$$

where $D_R$ (for a region $R$) stands for the operator (1.3) with Neumann boundary conditions on $R$. In view of $\sum \rho_k(j) = 0$ the operator norm is to be taken on the orthogonal complement of the constant functions.

The square of the $L^2$-norm of $\rho_{k+1}$ is bounded by $(\operatorname{vol} R_x^k)^{-1}$. Thus by (7.10)

$$\|\rho_{k+1}\|_2^2 \leq (\operatorname{vol} R_x^k)^{-1} \leq 2^{-kd}\delta^{-1} \ . \tag{7.14}$$

The corresponding inequality also holds for $\tilde{\rho}_{k+1}$.

We must still bound the operator norm $\|D^{-1}_{R^k_x}\|$. For this we observe that the conditions (7.5) ensure that

$$A_{xy}(jj')^{-1} \le 2^3(a\,|j-x|^\alpha)^2 \le 8a^2 2^{2k\alpha}$$

for all $j, j' \in R^k_x$ and $k \le k_m - 1$, since in that case $jj' \in R^x_{xy}$ and we apply (7.6) for $z = x$. For $k = k_m$ we are looking at pairs that belong to $R^y_{xy}$ but not to $R^x_{xy}$. In that case we apply (7.6) for $z = y$ and still have $A_{xy}(jj')^{-1} \le 8a^2 2^{2k\alpha}$. Therefore, since $R_{xy}$ is $\delta$-admissible and the lowest nonzero eigenvalue of the Neumann Laplacian on $R^k_x$ is of the order of $(2^k)^{-2}$, we obtain

$$\mathrm{e}^{t_x+t_y} B^{-1}_{xy} \|D^{-1}_{R^k_x}\| \le c(\delta)\, a^2 \beta^{-1} 2^{2k} 2^{2k\alpha} \tag{7.15}$$

for some $c(\delta)$ and all $k \le k_m$. For $2^k \le 10$ the connectedness of $R_{xy}$ ensures that $\|D^{-1}_{R^k_x}\| \le$ const. The same bounds apply for $\|D^{-1}_{R^k_y}\|$.

Finally, by combining (7.12) with (7.13), (7.14), and (7.15), we arrive at

$$G_{xy} = \frac{\mathrm{e}^{t_x+t_y}}{B_{xy}} \left[\delta_x - \delta_y\,; D_{\beta,\varepsilon}(t)^{-1}(\delta_x - \delta_y)\right] \le c(\delta)\,\frac{a^2}{\beta}\left(2\sum_{k=1}^{k_m} \sqrt{2}^{\,k(2\alpha+2-d)}\right)^2 .$$

For $2\alpha < d - 2 = 1$ the value of this sum is bounded uniformly in $k_m$. ∎

**Remark.** The bound (7.7) also applies when the definition of $R^x_{xy}$ and $R^y_{xy}$ is modified in the following way (for $z = x, y$ as before):

$$R^z_{xy} = \{j \in R_{xy} : |j - z| \le |x - y|\, f_z\}\,, \tag{7.16}$$

where $f_x$, $f_y$ are a pair of positive numbers which add up to (at least) unity and neither of which is too small. It is easy to see that the relevant scales involved are the ones for $k$ near $k_m$ and we can get the same bound but with a change of overall factor. This remark will become important in Section 12, Lemma 11, where we will need this estimate with $f_y \simeq 1/5$.

# 8 Conditional estimates on fluctuations

In this section we establish bounds on the fluctuations of the field $t$ by bounding $\langle B^m_{xy}\,\bar\chi_{xy}\rangle$ where $\bar\chi_{xy}$ has the property that $\bar\chi_{xy} = 0$ whenever the hypothesis (7.6) of Lemma 5 fails.

**Definition 3 (characteristic function)**: *As before, fix two constants $a > 1$ and $1/2 > \alpha > 0$, and let $r_{j-k} := (a\,|j-k|^\alpha)^{-1}$ for $j$, $k \in \Lambda$, $j \ne k$. Let $\chi : \mathbb{R}_+ \to \mathbb{R}$ be the characteristic function of the interval $[0,1]$, i.e., $\chi(t) = 1$ for $0 \le t \le 1$ and $\chi(t) = 0$ for $t > 1$. Moreover, let $R_{xy}$ be $\delta$-admissible and choose the regions $R^x_{xy}$, $R^y_{xy}$ as in (7.4). In this setting we define*

$$\bar\chi_{xy} = \prod_{j\in R^x_{xy}} \chi_{xj} \prod_{j\in R^y_{xy}} \chi_{yj}\,, \qquad \chi_{zj} = \chi(r_{j-z}B_{zj}) \quad (z = x, y)\,. \tag{8.1}$$

Here the constants $a$, $\alpha$ are taken to coincide with those in Lemma 5.

With these definitions we have

**Lemma 6** *Let $R_{xy}$ be a $\delta$-admissible region, and let $C = C(a, \alpha, \delta)$ be the constant that appears in Lemma 5. Then for $0 \le m < \beta/C$ we have*

$$\left\langle B^m_{xy}\, \bar{\chi}_{xy} \right\rangle \le (1 - mC/\beta)^{-1} \ . \tag{8.2}$$

**Proof.** Our proof uses the identity $\langle S^m_{xy}\, \chi^S_{xy} \rangle = 1$ where $\chi^S_{xy}$ is a supersymmetric version of $\bar{\chi}_{xy}$ defined above. After integrating out the Grassmann fields we shall show that this identity implies

$$\left\langle B^m_{xy}\, \bar{\chi}_{xy}\, (1 - mG_{xy}) \right\rangle \le 1 \ .$$

Lemma 5 and the presence of $\bar{\chi}_{xy}$ then yield (8.2).

More precisely, let $\chi_\gamma \in \mathrm{C}^\infty(\mathbb{R}_+)$ with $\frac{d}{dt}\chi_\gamma(t) \le 0$ and

$$\chi_\gamma(t) = \begin{cases} 1 & t \le 1 - \gamma \ , \\ 0 & t \ge 1 \ , \end{cases}$$

be a smooth regularization of $\chi = \lim_{\gamma \to 0} \chi_\gamma$ . We fix a small value of $\gamma > 0$ and write $\tilde{\chi} \equiv \chi_\gamma$ for short. Then, recalling the definition (6.3) of $F_{xy}$ we introduce

$$\chi^S_{xj} = \tilde{\chi}(r_{j-x}\, S_{xj}) = \tilde{\chi}_{xj} + r_{j-x}\, \tilde{\chi}'_{xj}\, \bar{\psi}(F_{xj})\psi(F_{xj}) \ , \tag{8.3}$$

where $\tilde{\chi}_{xj} = \tilde{\chi}(r_{j-x} B_{xj})$. Since $\chi^S_{xj}$ is $Q$-invariant and $\chi^S_{xj}(0) = 1$ we have

$$1 = \left\langle S^m_{xy} \prod_{j \in R^x_{xy}} \chi^S_{xj} \prod_{j \in R^y_{xy}} \chi^S_{yj} \right\rangle \equiv \left\langle S^m_{xy}\, \tilde{\chi}^S_{xy} \right\rangle \ . \tag{8.4}$$

Now, we express

$$\left\langle S^m_{xy}\, \tilde{\chi}^S_{xy} \right\rangle = \left\langle S^m_{xy}\, \tilde{\chi}_{xy}\, \exp -[\bar{\psi}\, ; A\psi] \right\rangle, \tag{8.5}$$

where the symmetric operator $A$ is given by

$$[f; Af] = -\sum_{j \in R^x_{xy}} \frac{r_{j-x}\tilde{\chi}'_{xj}}{\tilde{\chi}_{xj}} [f; F_{xj}]^2 - \sum_{j \in R^y_{xy}} \frac{r_{j-y}\tilde{\chi}'_{yj}}{\tilde{\chi}_{yj}} [f; F_{yj}]^2 \ . \tag{8.6}$$

Clearly $A \ge 0$ as a quadratic form since $\tilde{\chi}' \le 0$ . The total $\bar{\psi}\psi$ contribution to (8.5) including the fermionic part of the action is

$$[\bar{\psi}\, ; (D_{\beta,\varepsilon}(t) + A)\psi] - mB^{-1}_{xy} \bar{\psi}(F_{xy})\, \psi(F_{xy}) \ ,$$

where the second summand stems from $S^m_{xy}$ , see (5.2)-(5.3). Thus, integration over the Grassmann fields $\bar{\psi}, \psi$ gives $\mathrm{Det}(Q + A)$ where

$$Q = D_{\beta,\varepsilon}(t) - mB^{-1}_{xy} F_{xy}\, [F_{xy}\, ; \cdot\,] \, .$$

Since we are taking $m$ to be less than $\beta/C$, the presence of the factor $\tilde{\chi}_{xy}$ in (8.5) ensures (by Lemma 5) that $mG_{xy} < 1$. Now by Lemma 1 the inequality $1 \geq mG_{xy} = mB_{xy}^{-1}[F_{xy}; D_{\beta,\varepsilon}(t)^{-1}F_{xy}]$ is equivalent to $Q \geq 0$. Therefore the result $\mathrm{Det}(Q+A)$ of integrating over $\bar{\psi}, \psi$ is bounded from below by

$$\mathrm{Det}(Q+A) \geq \mathrm{Det}(Q) = \mathrm{Det}(D_{\beta,\varepsilon}(t))\,(1 - mG_{xy}) \geq 0\ ,$$

and we obtain the estimate

$$1 = \left\langle B_{xy}^m\, \tilde{\chi}_{xy}\, \mathrm{e}^{-[\bar{\psi}\,;(Q+A)\psi]} \right\rangle \geq \left\langle B_{xy}^m\, \tilde{\chi}_{xy}\, (1 - mG_{xy}) \right\rangle\ .$$

We finally take the limit $\gamma \to 0$. The smooth function $\tilde{\chi}_{xy}$ then converges to the characteristic function $\bar{\chi}_{xy}$. Hence

$$1 \geq \left\langle B_{xy}^m\, \bar{\chi}_{xy}\, (1 - mG_{xy}) \right\rangle \geq \left\langle B_{xy}^m\, \bar{\chi}_{xy} \right\rangle (1 - mC/\beta)\ ,$$

which is the desired result. ∎

**Lemma 7** *If all of the regions $R_{x_1y_1}, R_{x_2y_2}, \ldots, R_{x_ny_n}$ are $\delta$-admissible and disjoint (meaning they have disjoint interiors), then we still have*

$$\left\langle \prod_{j=1}^{n} B_{x_jy_j}^m \bar{\chi}_{x_jy_j} \right\rangle \leq (1 - mC/\beta)^{-n}\ . \tag{8.7}$$

**Proof.** As before we use the fact that the supersymmetrized observable, which here results from replacing $B_{x_jy_j}$ by $S_{x_jy_j}$, has expectation one.

Consider first the simpler problem of computing the expectation of the product $\prod S_{x_iy_i}^m \bar{\chi}_{x_iy_i}$. After integrating over $\psi$ and $\bar{\psi}$ we see that

$$\left\langle \prod_{i=1}^{n} S_{x_iy_i}^m\, \bar{\chi}_{x_iy_i} \right\rangle = \left\langle \prod_{i=1}^{n} B_{x_iy_i}^m\, \bar{\chi}_{x_iy_i}\, \mathrm{Det}(1 - m\mathcal{G}) \right\rangle\ , \tag{8.8}$$

where $\mathcal{G}$ is an $n \times n$ matrix of Green's functions

$$\mathcal{G}_{ij} = \beta^{-1}[g_i\,; D_{\beta,\varepsilon}(t)^{-1} g_j]\ , \qquad g_i = B_{x_iy_i}^{-1/2}\, \mathrm{e}^{(t_{x_i}+t_{y_i})/2}\, (\delta_{x_i} - \delta_{y_i})\ . \tag{8.9}$$

The matrix $\mathcal{G}$ is positive as a quadratic form. In order to reduce the problem to the previous case (of just a single region) note that $\mathcal{G} \leq \mathcal{G}_N$, where the subscript denotes Neumann boundary conditions on the boundaries of the disjoint regions $R_{x_iy_i}$. The presence of the factors $\bar{\chi}_{x_iy_i}$ implies bounds on the $G_{x_iy_i}$ so that

$$\mathrm{Det}(1 - m\mathcal{G}) \geq \mathrm{Det}(1 - m\mathcal{G}_N) = \prod_{i=1}^{n} (1 - mG_{x_iy_i}) \geq (1 - mC/\beta)^{-n}\ . \tag{8.10}$$

The proof of the lemma is completed by introducing the effects of $\tilde{\chi}'$ as before. Since there are no new aspects to this argument, we omit it.

**Remark.** From this lemma one obtains estimates for conditional probabilities only. Yet, in order to bound $C_{xy}$ in Theorem 3 we need probability estimates without any conditions, which is why we now have to develop an inductive argument.

# 9 Unconditional estimates on fluctuations

We are now going to remove the constraints enforced by insertion of $\bar{\chi}$. In order to do so, we have to consider $\chi^c_{xj} = 1 - \chi_{xj}$ for $\chi_{xj}$ defined by (8.1). Short scales (given by $0 < |j - x| < \beta^{1/4}$) will be treated separately by monitoring, in Section 10, only the size of nearest neighbor gradients inside the region $R_{xy}$. At the very large scales of $|j - x| \geq \beta^{1/4}$, however, looking only at NN fluctuations is not enough. There, in order to remove the $\bar{\chi}$ constraints we will show by induction on the distance $|j - x|$ that the corresponding contribution is small.

We will distinguish between two types of geometry: diamonds and deformed diamonds. For deformed diamonds we will quantify the bounds given by (7.7) and call such regions $C$-admissible.

**Definition 4** *Let $R_{xy} \subset \Lambda$ be $\delta$-admissible in the sense of Def. 1.*

1. *We call $R_{xy}$ a* diamond *if it is the set of lattice points which is contained in a 3-dimensional double cone obtained in the following way: we take a 2-dimensional rectangular diamond with opposite vertices placed on $x$ and $y$ and edges of length $|x - y|/\sqrt{2}$ (see Fig. 3) and rotate it around the line $\overline{xy}$. In order to ensure connectedness we may have to add a few lattice points near $x$ and $y$ (see Fig. 2 a,b).*

2. *We call $R_{xy}$ a $C$-*admissible *region (or deformed diamond) if*

$$0 \leq G^N_{xy}\, \bar{\chi}_{xy} \leq C/\beta \quad \textit{for } |x - y| > \beta^{1/4}, \tag{9.1}$$

$$0 \leq G^N_{xy} \prod\nolimits'_{pq} \chi_{pq} \leq C/\beta \quad \textit{for } |x - y| \leq \beta^{1/4}, \tag{9.2}$$

*where $\bar{\chi}_{xy}$ is defined in* (8.1)*, the superscript $N$ stands for Neumann boundary conditions on $R_{xy}$ and $\prod'$ denotes the product over all nearest neighbor pairs in $R_{xy}$.*

Note that for short scales, dealt with in (9.2), instead of using $\bar{\chi}_{xy}$ we impose constraints on *all* NN pairs in the region $R_{xy}$.

With these definitions we can state the main result of this paper.

**Theorem 4** *Let $m = \beta^{1/8}$, and let $R_{x_i y_i}$ for $i = 1, \ldots, n_1$ be diamonds with disjoint interiors and $|x_i - y_i| > \beta^{1/4}$. Then we have*

$$\left\langle \prod_{i=1}^{n_1} B^m_{x_i y_i} \right\rangle \leq 2^{n_1} \tag{9.3}$$

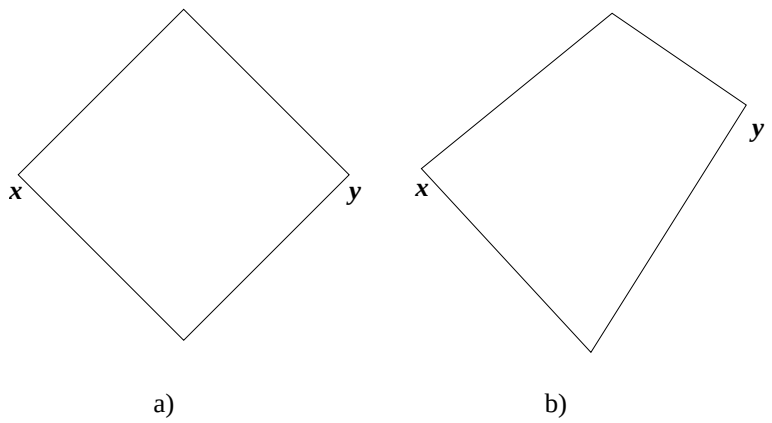


Figure 3: a) diamond region, b) $C$-admissible region (deformed diamond)

*for all* $n_1 \geq 0$. *Moreover if* $p_j, q_j$ *for* $j = 1, \ldots, n_2$ *are such that* $|p_j - q_j| > \beta^{1/4}$, *the regions* $R_{p_j q_j}$ *are* $C$*-admissible, have disjoint interiors and do not overlap with any of the regions* $R_{x_i y_i}$ *, then there exists a constant* $\rho \leq 1/2$ *such that*

$$\left\langle \prod_{i=1}^{n_1} B^m_{x_i y_i} \prod_{j=1}^{n_2} B^{3m}_{p_j q_j} \bar{\chi}_{p_j q_j} \right\rangle \leq 2^{n_1} (1+\rho)^{n_2} \tag{9.4}$$

*for all* $n_1 \geq 0$ *and* $n_2 \geq 0$. *Finally, let* $r_k, s_k$ *for* $k = 1, \ldots, n_3$ *be such that* $|r_k - s_k| \leq \beta^{1/4}$, $R_{r_k s_k}$ *are* $C$*-admissible, have disjoint interiors and do not overlap with any of the regions* $R_{x_i y_i}$ *or* $R_{p_j q_j}$ *. Then for all* $n_1 \geq 0$, $n_2 \geq 0$ *and* $n_3 \geq 0$ *we have*

$$\left\langle \prod_{i=1}^{n_1} B^m_{x_i y_i} \prod_{j=1}^{n_2} B^{3m}_{p_j q_j} \bar{\chi}_{p_j q_j} \prod_{k=1}^{n_3} B^{3m}_{r_k s_k} \right\rangle \leq 2^{n_1} (1+\rho)^{n_2}\, 2^{n_3} \tag{9.5}$$

*with* $\rho \leq 1/2$.

The proof of the theorem is carried out in Sections 10–12. We will need to distinguish between three situations, which we refer to as *classes*.

**Class 1.** $|x - y| > \beta^{1/4}$ and the pair is not protected by a factor of $\bar{\chi}_{xy}$. In this case we need an inductive argument on scales to prove a bound on the expectation of $B^m_{xy}$. The induction will be done on $\ell = \max_j |x_j - y_j|$ and is carried out in Sections 11 and 12. We will need to inductively select non-overlapping smaller diamonds inside the region $R_{xy}$ while making sure that these remain $\delta$-admissible. To arrange for all geometrical details to work out, we take $R_{xy}$ to be a perfect diamond.

**Class 2.** $|x - y| > \beta^{1/4}$ but the pair is protected by a factor of $\bar{\chi}_{xy}$. In this case we can apply the results of Section 8, thereby obviating the need for any

induction. $R_{xy}$ is then allowed to be a deformed diamond and the bound we can get is stronger than in Class 1 (power $3m$ instead of $m$).

**Class 3.** $|x-y| \leq \beta^{1/4}$. This includes short scales and the NN case, which was already treated in Section 6. We will show in Section 10 that these scales do not require any factor of $\bar{\chi}_{xy}$ to ensure a good bound. No induction is needed, and we can therefore take $R_{xy}$ to be a deformed diamond.

Note that the larger exponent $3m$ appearing in (9.4) and (9.5) is important for the inductive proof to go through. The enlarged exponent can be handled either because of the presence of $\bar{\chi}$ or because the pair is of Class 3.

### 9.1 Fixing the different parameters

We have introduced a certain number of parameters: $m$, $a$, $\rho$, $\delta$, $C$, $\alpha$. Before going on, we briefly review why they appeared and how to choose their values.

1. The parameter $m$ is ubiquitous in this paper as the power of $B_{xy}$. Since the probability of large deviations will be bounded by $K^{-m}$ with $K > 1$, we want $m$ to be as large as possible. On the other hand, to apply the SUSY argument of Section 8 we must have $(1 - 3mC/\beta)^{-1} < 1$, where the factor $3m$ in this inequality comes from the power of $B$ in (9.4), (9.5). Therefore the magnitude of $m$ is limited by $\beta$. To arrange for all the conditions to be met, we fix $m = \beta^{1/8}$. The factors $m$ and $3m$ will be kept fixed in the whole course of proof.

2. The constants $C$ and $\delta$ appearing in the definition of the region $R_{xy}$ (see Def. 1) are not subject to any special requirements, but their values do constrain the other parameters. They will be fixed throughout.

3. To prove the induction hypothesis we need $0 \leq \rho \leq 1$. More precisely (see Eq. (12.6)) we need $\rho + \mathcal{R}(x,y) \leq 1$. Since we prove $\mathcal{R}(x,y) \leq 1/2$ we will take $\rho \leq 1/2$.

4. The constant $a$ in Lemma 5 in Section 7 plays a key role in bounding the entropy for small scales; see Section 12, Case 1, Eq. (12.16). It will become clear there that $a > 10$ is sufficient.

5. We need to take $\alpha > 0$ in Lemma 5 in order to control entropy factors for large deviations (see Section 12). On the other hand, the result of Theorem 2 would be optimal for $\alpha = 0$. Therefore we wish to make $\alpha$ as small as possible. We will see in Section 12 (Case 2b eq. (12.21) and Case 2c eq. (12.23)) that $\alpha \geq O(1/\ln\beta)$ is a requirement for our analysis to go through.

## 10 Short-scale fluctuations

We now prove Theorem 4 for $\ell \leq \beta^{1/4}$, i.e., for Class 3 pairs. These estimates will follow from the bounds on NN fluctuations established in Section 6.

**Lemma 8** *There is a constant $\beta_0$ such that for $\beta \geq \beta_0$ , $|x-y| = \ell \leq \beta^{1/4}$ and $3m \leq \beta^{1/8}$, we have*

$$\langle B_{xy}^{3m} \rangle \leq 2 \; . \tag{10.1}$$

*More generally let $(x_1, y_1), \ldots, (x_n, y_n)$ be $n$ pairs with $|x_j - y_j| \leq \ell$ for all $j$, and let the interiors of the corresponding $C$-admissible regions $R_{x_1 y_1}$, $R_{x_2 y_2}, \ldots, R_{x_n y_n}$ be disjoint. Then*

$$\left\langle \prod_{j=1}^{n} B_{x_j y_j}^{3m} \right\rangle \leq 2^n. \tag{10.2}$$

**Proof.** As in Def. 3, let $\chi$ be the characteristic function of the interval $[0,1]$ and let (with a parameter $\delta$ to be defined shortly)

$$\chi_{pq} = \chi((1+\delta)^{-1} B_{pq}) \; , \quad |p-q| = 1 \; , \tag{10.3}$$

and $\chi_{pq}^c = 1 - \chi_{pq}$ . Using $\chi_{pq} \leq 1$ we have $1 \leq \prod_{(pq)} \chi_{pq} + \sum_{(pq)} \chi_{pq}^c$ and

$$\langle B_{xy}^{3m} \rangle \leq \langle B_{xy}^{3m} \prod\nolimits_{(pq)} \chi_{pq} \rangle + \sum\nolimits_{(pq)} \langle B_{xy}^{3m} \, \chi_{pq}^c \rangle \; , \tag{10.4}$$

where the product and the sum are over all nearest neighbor pairs $(pq)$ in $R_{xy}$ .

We estimate the first term on the right-hand side of (10.4) by applying the strategy of the proof of Lemma 6 to show that

$$\langle B_{xy}^{3m} \big(1 - 3m \, G_{xy}\big) \prod\nolimits_{(pq)} \chi_{pq} \rangle \leq 1 \; . \tag{10.5}$$

To bound $G_{xy}$ , note that on the support of $\chi_{pq}$ we have

$$0 \leq \tfrac{1}{2}(t_p - t_q)^2 \leq \cosh(t_p - t_q) - 1 \leq \delta \; . \tag{10.6}$$

Thus $|t_p - t_q| \leq \sqrt{2\delta}$ and $|t_z - t_j| \leq \ell\sqrt{2\delta}$ for $z = x, y$ and all $j \in R_{xy}$ . Now let us require

$$\ell\sqrt{2\delta} = 1 \; , \quad \text{or} \quad \delta = \tfrac{1}{2}\beta^{-1/2} \; , \tag{10.7}$$

since $\ell \leq \beta^{1/4}$. Thus we have a uniform lower bound on the conductance (7.5). It then follows that $0 \leq G_{xy} \leq C/\beta$ with $C$ independent of $\beta$, and (10.5) gives

$$\langle B_{xy}^{3m} \prod\nolimits_{(pq)} \chi_{pq} \rangle \leq (1 - 3m \, C/\beta)^{-1} \leq 3/2 \tag{10.8}$$

for $\beta$ large.

For the second summand of (10.4) we use

$$\chi_{pq}^c = 1 - \chi\big((1+\delta)^{-1} B_{pq}\big) \leq \left( \frac{B_{pq}}{1+\delta} \right)^{\beta/2} \; . \tag{10.9}$$

The factor $B_{xy}^{3m}$ is estimated by repeated application of (5.14):

$$2B_{xy} \leq \prod\nolimits_j 2B_{p_j q_j} \; , \tag{10.10}$$

where the product ranges over a set of NN pairs connecting $x$ and $y$. By combining (10.9) and (10.10) and then using the result (6.9) for NN pairs we have

$$\begin{aligned}\left\langle B_{xy}^{3m}\,\chi_{pq}^{c}\right\rangle &\le \frac{2^{3m(\ell-1)}}{(1+\delta)^{\beta/2}}\left\langle B_{pq}^{\beta/2}\prod\nolimits_j B_{p_jq_j}^{3m}\right\rangle \\ &\le \frac{2^{3m(\ell-1)}}{(1+\delta)^{\beta/2}}\left(\tfrac{1}{2}-3m/\beta\right)^{-1}(1-3m/\beta)^{-\ell} \\ &\le \mathrm{e}^{3m\ell}\mathrm{e}^{-\beta\delta/3}\,. \end{aligned} \tag{10.11}$$

Since $3m\ell \le \beta^{1/8}\beta^{1/4}$ by hypothesis, and $\delta = \frac{1}{2}\beta^{-1/2}$ by (10.7), we see that the expression (10.11) is less than $\exp(\beta^{3/8} - \beta^{1/2}/6)$.

Combining our estimates on the two terms on the r.h.s. of (10.4) we have

$$3/2 + 3\ell^3\mathrm{e}^{\beta^{3/8}-\beta^{1/2}/6} \le 2$$

for large enough $\beta$. The factor $3\ell^3 \le 3\beta^{3/4}$ comes from the sum over all NN pairs in $R_{xy}$.

When several disjoint regions are present, the bounds over disjoint regions factor, and we can get the same result using the same argument. ■

# 11 Induction hypothesis and some preliminary estimates

The argument in the last section cannot be repeated for all values of $\ell$. In order to control all scales we need an inductive argument.

**Induction Hypothesis:** Let $x_i, y_i$ $(i = 1, \ldots, n_1)$, $p_j, q_j$ $(j = 1, \ldots, n_2)$, and $r_k, s_k$ $(k = 1, \ldots, n_3)$ be pairs of Class 1, 2, resp. 3, in the sense of Theorem 3. Then the bounds (9.3), (9.4), (9.5) hold when $|x_i - y_i| \le \ell$ for all $i \le n_1$. ■

The induction is on $\ell = \max_i |x_i - y_i|$. The said bounds were already established for $\ell = 1$ (NN case, Section 6) and $\ell \le \beta^{1/4}$ (Section 10). Assuming that the Induction Hypothesis holds up to scale $\ell$, we shall prove (in Section 12) that it holds up to scale $\ell + 1$. This will complete the proof of Theorem 4 and, as an immediate consequence, Theorem 1.

The idea of the proof is the same as in Section 10. If the pair $xy$ is protected by a $\bar{\chi}_{xy}$ factor (Class 2), then we apply Lemma 6 in Section 8.

To get the unconditional estimates we must study the situation when $\bar{\chi}_{xy}$ is violated. This violation may happen at any scale from 1 up to $\ell$. To quantify this we introduce the following definition.

**Definition 5** *A point $x \in \Lambda$ is called $n$-good if*

$$B_{xy} \le a\,|x-y|^{\alpha} \tag{11.1}$$

*for all $y \in \Lambda$ with distance $1 \le |x-y| \le 4^n$ from $x$.*

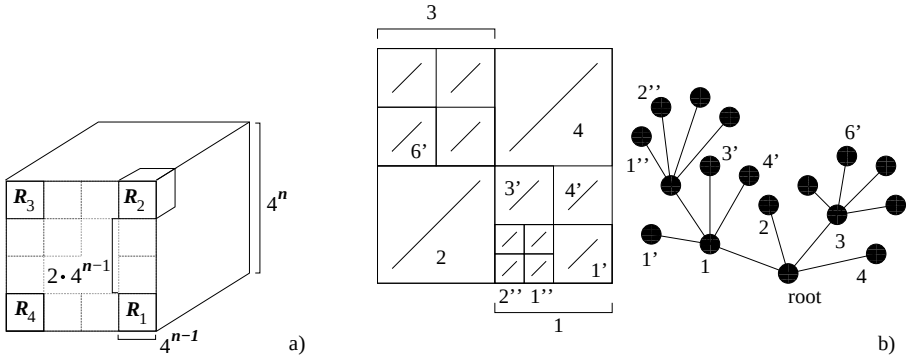


Figure 4: a) In a 3D cube of side $4^n$ we select 8 cubes of side $4^{n-1}$. b) Here we see an example of a rooted tree (on a 2D square) with coordination number 5 or 1 at each vertex, and the corresponding set of subsquares. The root corresponds to the large square.

**Definition 6** *For a cube $R_n$ of side $4^n$ we define $\chi^c_{R_n}$ to be the indicator function of the event that there exists no $n$-good point in $R_n$ .*

Our goal in the present section is to bound the expectation of the indicator function $\chi^c_{R_n}$. In brief we will achieve this by estimating $\chi^c_{R_n}$ by a sum of products of factors of $B_{xy}$ and then using (9.5). The details are as follows.

A 3D cube $R_n$ of side $4^n$ can be expressed as a union of $4^3$ disjoint subcubes of side $4^{n-1}$. It is clear by inspection of Fig. 4 that we can select $2^3 = 8$ of these subcubes, say $R^i_{n-1}$ $(i = 1, \ldots, 8)$, so that $\mathrm{dist}(R^i_{n-1}, R^j_{n-1}) > 4^{n-1}$ $(i \neq j)$. Our approach now rests on the following simple observation: if there is no $n$-good point in $R_n$ , then there is either no $(n-1)$-good point in any of the 8 subcubes $R^i_{n-1}$ , or else there exists at least one bad pair $(x, y) \in R_n \times \Lambda$ at scale $4^{n-1} < |x - y| \leq 4^n$. Thus, $\chi^c_{R_n}$ is bounded by the inequality

$$\chi^c_{R_n} \leq S^c_{R_n} + \prod_{i=1}^{8} \chi^c_{R^i_{n-1}}, \tag{11.2}$$

where

$$S^c_{R_n} = \sum_{\substack{x \in R_n,\, y \in \Lambda \\ 4^{n-1} < |x-y| \leq 4^n}} \chi^c_{xy} \,. \tag{11.3}$$

We iterate (11.2) inside each cube $R^j_{n-1}$ , thus selecting $8^2$ subcubes of side $4^{n-2}$, and we keep repeating this procedure until we reach cubes of side $4^0 = 1$ (i.e. points). We denote by $\tilde{R}_{n-k}$ $(k = 0, \ldots, n)$ the set of $8^k$ cubes of side $4^{n-k}$ obtained in this way. $\tilde{R}_n = \{R_n\}$ is the starting cube. Moreover let $\tilde{R} = \cup_{k=0}^{n} \tilde{R}_{n-k}$. In this way we bound $\chi^c_{R_n}$ by a positive sum of products of $\chi^c_{xy}$ , which in turn are bounded by $B^m_{xy}/(a\,|x-y|^\alpha)^m$. The resulting expression

can be organized as a sum over rooted trees picturing the hierarchy of inclusion relations of the subcubes.

The following set of definitions serves to prepare the statement of Lemma 9 below. Let $V$ be an abstract set of vertices such that $|V| = |\tilde{R}|$. We associate by a fixed bijective map each vertex $v$ in $V$ to a cube $R_v \in \tilde{R}$. We denote by $k_v$ the scale of the corresponding cube: $R_v \in \tilde{R}_{n-k_v}$. The vertex associated with the largest cube $R_n$ is denoted by $r$ (root). Finally let $\mathcal{A}(v)$ (ancestor of $v$) be the unique vertex in $V$ such that $R_v \subset R_{\mathcal{A}(v)}$ (see vertex 1 and 1' in Fig. 4b).

With these definitions we can introduce $\mathcal{T}_n$ the set of labelled rooted trees on some subset of $V$ with root $r$, such that the root has coordination number $d_r = 8$ or $d_r = 0$ (in which case the tree is reduced to a single vertex) and the other vertices have coordination number $d_v = 9$ or $d_v = 1$. Moreover if $v$ belongs to the tree then there must be a tree line connecting $v$ to its ancestor $\mathcal{A}(v)$. The maximal distance of a vertex from the root is $n$. Let $L_T$ denote the set of vertices in $T$ with $d_v = 1$ (the leaves) or $d_r = 0$ (then $L_T$ contains only the root). Let $V_k$ be the set of vertices in $T$ at distance $k$ from the root. With these definitions the tree is completely fixed by the leaves $L_T$ (or equivalently by the choice of the coordination numbers for each vertex). See Fig. 4b for an example in the case of $d_v = 5$ instead of 9.

**Lemma 9** *With the definitions above we have the inequality*

$$\chi^c_{R_n} \leq \sum_{T \in \mathcal{T}_n} \prod_{v \in L_T} S^c_{R_v} \ , \tag{11.4}$$

*where $R_v$ is a cube of side $4^{n-d_v}$ and $S^c_{R_v}$ is defined as in* (11.3)*, with $n$ replaced by $n - d_v$ .*

**Proof.** Our trees $T \in \mathcal{T}_n$ are constructed by iterating (11.2). In each iteration we get to choose between the first and second term of the r.h.s. of (11.2).

The construction starts with the root of the tree. In the first step of the iterative scheme, if we pick the first term of (11.2) then the construction ends and we have produced nothing but the trivial tree (the root). If we pick the second term, we have a product of 8 different indicator functions $\chi^c_R$, one for each subcube in $\tilde{R}_{n-1}$. We represent them in the tree by attaching 8 vertices to the root. Each vertex is then associated to a subcube by lexicographical order (see Fig. 4b). Now we repeat this procedure at the end of each branch and, continuing in this way, construct a tree. Whenever we pick the first term in (11.2), the corresponding branch of the tree terminates and we produce a terminal vertex or leaf. If we pick the second term, we generate a vertex of coordination number 9. The iteration stops when we reach scale $n$. ■

We are now in a position to state and prove the main result of this section.

**Proposition 1** *Let $d = 3$ and let the parameters $a, m, \alpha$ be chosen such that $m\alpha \geq 4d$ and $a \geq 2^{3\alpha+(d+2)/m}$. Assume that the induction hypothesis* (9.5) *holds up to scale $\ell = 4^{n+1}$. Then*

$$\left\langle \chi^c_{R_n} \right\rangle \leq 2^{-(n+1)\alpha m} \ , \tag{11.5}$$

*and if $R_{n(k)}$, $k = 1, \ldots, N$, denotes a family of cubes of side $4^{n(k)} \le \ell$ such that*

$$\mathrm{dist}(R_{n(k)}, R_{n(k')}) \ge \max(4^{n(k)}, 4^{n(k')}) \tag{11.6}$$

*then*

$$\left\langle \prod_{k=1}^{N} \chi^c_{R_{n(k)}} \right\rangle \le \prod_{k=1}^{N} 2^{-(n(k)+1)\alpha m} \ . \tag{11.7}$$

**Proof.** Note that the average of (11.3) is bounded by

$$\begin{aligned} \langle S^c_{R_n} \rangle &\le \sum_{\substack{x \in R_n,\, y \in \Lambda \\ 4^{n-1} < |x-y| \le 4^n}} \frac{\langle B^m_{xy} \rangle}{a^m |x-y|^{\alpha m}} \\ &\le (4^n)^d (4^n 2)^d \frac{2}{a^m 4^{(n-1)\alpha m}} \le \tfrac{1}{2}\, 2^{-(n+1)\alpha m} \end{aligned}$$

for all $n \ge 1$. For $n = 0$ the sum is over $y \in \Lambda$ at distance $|x - y| = 1$ and we obtain the same bound (see below).

First, consider the minimal scale $n = 0$. In this case $R_n = R_0$ is just a single point, the sum over $y$ is a sum over $2d$ nearest neighbors, and we simply have

$$\langle \chi^c_{R_0} \rangle \le \langle S^c_{R_0} \rangle \le 2d\, \frac{2}{a^m} \le 2^{-\alpha m} \ .$$

Thus (11.5) holds for $n = 0$. Now let $n \ge 1$. By (11.2) we have

$$\langle \chi^c_{R_n} \rangle \le \langle S^c_{R_n} \rangle + \left\langle \prod_{i=1}^{8} \chi^c_{R^i_{n-1}} \right\rangle \ .$$

Assuming that $\langle \chi^c_{R_{n-1}} \rangle \le 2^{-n\alpha m}$ holds, we conclude that

$$\langle \chi^c_{R_n} \rangle \le \tfrac{1}{2}\, 2^{-(n+1)\alpha m} + (2^{-n\alpha m})^8 < 2^{-(n+1)\alpha m} \ .$$

Hence the result is true for all $n$ by induction.

To prove (11.7) we apply (11.4) to each term $\chi^c_{R_{n(k)}}$. Then we can again apply (9.5) as long as the corresponding diamonds are disjoint – this is ensured by the procedure for choosing subcubes and by the constraint $\mathrm{dist}(R_{n(k)}, R_{n(k')}) \ge \max(4^{n(k)}, 4^{n(k')})$. This concludes the proof of Proposition 1. ∎

# 12 Proof of the Induction Hypothesis

In this section we shall establish the induction hypothesis of Section 11 at scale $\ell$ assuming that it holds up to scale $\ell - 1$. Since our regions $R_{x_i y_i}$ are disjoint by assumption, we will be able to re-express each factor $B_{x_i y_i}$ as a sum over non-overlapping regions where our induction hypothesis applies. To simplify

the notation let $x_i = x$ and $y_i = y$. We will assume that $\ell = |x - y|$ is large, i.e., $|\ell| \geq \beta^{1/4}$. (The case of small $\ell < \beta^{1/4}$ was dealt with in Section 10.)

For $z = x, y$ we recall the meaning of the regions $R^z_{xy}$ from Def. 2, Section 7. To ensure that the new regions produced by the analysis below remain inside the original region $R_{xy}$ we need to introduce the following subsets.

**Definition 7** *Let $R_{xy}$ be a diamond region as described in Def. 4. Then we define $\tilde{R}^z_{xy}$ for $z = x, y$ as*

$$\tilde{R}^z_{xy} = \{j \in R^z_{xy} : \angle(jz, xy) \leq \pi/8 \text{ for } |j - z| > 10\} , \tag{12.1}$$

*where $\angle(jz, xy)$ is the angle between the lines $\overline{jz}$ and $\overline{xy}$.*

This definition roughly selects (at distances larger than 10) a double cone which is obtained by rotating around $\overline{xy}$ a 2D diamond with vertices on $x$ and $y$ and opening angle $\theta = \pi/8$ (see Fig. 2a). The condition $|j - k| > 10$ ensures that $\tilde{R}^z_{xy} \cup \{z\}$ is connected.

We also define

$$u_{xy} = \prod_{j \in \tilde{R}^x_{xy}} \chi_{xj} \prod_{j \in \tilde{R}^y_{xy}} \chi_{yj} \tag{12.2}$$

for $\chi_{xj}$ , $\chi_{yj}$ as defined in (8.1). Note that $\tilde{R}^x_{xy} \cup \tilde{R}^y_{xy} \cup \{x, y\}$ is a $\delta$-admissible region in the sense of Def. 1 (Section 7), so Lemma 5 and Lemma 6 can be applied to give

$$\langle B^m_{xy}\, u_{xy} \rangle \leq (1 - mC/\beta)^{-1} . \tag{12.3}$$

Now let $\chi^c_{xj} = 1 - \chi_{xj}$ and $\chi^c_{yj} = 1 - \chi_{yj}$ . The next lemma is nothing but a combinatorial identity based on the following partitions of unity:

$$1 = \prod_{j \in \tilde{R}^x_{xy}} (\chi_{xj} + \chi^c_{xj}) , \qquad 1 = \prod_{j \in \tilde{R}^y_{xy}} (\chi_{yj} + \chi^c_{yj}) , \tag{12.4}$$

where $\tilde{R}^x_{xy}$ and $\tilde{R}^y_{xy}$ are the regions defined above.

**Lemma 10** *The identity function can be written as*

$$1 = u_{xy} + \sum_{b \in \tilde{R}^x_{xy}} \chi^c_{xb} \prod_{j, |j-x| < |b-x|} \chi_{xj} + \sum_{b \in \tilde{R}^y_{xy}} \chi^c_{yb} \prod_{i \in \tilde{R}^x_{xy}} \chi_{xi} \prod_{j, |j-y| < |b-y|} \chi_{yj} . \tag{12.5}$$

**Proof.** We start from (12.4) and expand the first product over $j$, beginning with small $|j - x|$. For each factor $\chi_{xj} + \chi^c_{xj}$ we have two possibilities: either we pick $\chi_{xj}$ , in which case we proceed to the next factor and repeat, or else we pick $\chi^c_{xj}$ and then we stop expanding and leave all the other factors (with larger $|j - x|$) in summed form $\chi_{xj} + \chi^c_{xj} = 1$. In the resulting sum there is the term $\prod_{j \in \tilde{R}^x_{xy}} \chi_{xj}$ . This we multiply by the other product (over $j \in \tilde{R}^y_{xy}$) in (12.4), which we expand in the same way.

In total, we have either picked a factor $\chi$ for all $j \in \tilde{R}^x_{xy}$ and $j \in \tilde{R}^y_{xy}$ (this results in the term $u_{xy}$), or we have picked a term $\chi^c_{xj}$ or $\chi^c_{yj}$ somewhere during the course of the expansion process (this gives all the other terms). The point where we stopped is denoted by $b$ (where $b$ stands for 'bad') because $\chi^c_{xb} > 0$ or $\chi^c_{yb} > 0$ means that there is a large deviation at that point. ∎

Using the equality (12.5) of Lemma 10 we can rewrite $B^m_{xy}$ as

$$B^m_{xy} = B^m_{xy}\, u_{xy} + \mathcal{R}(x,y)\ , \tag{12.6}$$

where $\mathcal{R}(x,y)$ is defined as

$$\begin{aligned}\mathcal{R}(x,y) = &\sum_{b\in\tilde{R}^x_{xy}} B^m_{xy}\,\chi^c_{xb} \prod\nolimits_{j:\,|j-x|<|b-x|} \chi_{xj} \\ &+ \sum_{b\in\tilde{R}^y_{xy}} B^m_{xy}\,\chi^c_{yb} \prod\nolimits_{i\in\tilde{R}^x_{xy}} \chi_{xi} \prod\nolimits_{j:\,|j-y|<|b-y|} \chi_{yj}\ .\end{aligned} \tag{12.7}$$

We have $\langle B^m_{xy}\, u_{xy}\rangle \le (1-mC/\beta)^{-1}$ by (12.3), without any need for an inductive argument. Thus if we can prove that $\langle \mathcal{R}(x,y)\rangle \le 2-(1-mC/\beta)^{-1}$ our proof will be complete. The desired statement is formulated in the next lemma.

**Lemma 11** *For large $\beta$ the remainder* (12.7) *is bounded in average by*

$$\langle \mathcal{R}(x,y)\rangle \le 1/2\ . \tag{12.8}$$

*Moreover, $\mathcal{R}(x,y)$ can be written as a sum over products of $B_{x'y'}$ with $|x'-y'| \le \ell-1$ in such a way that the corresponding regions $R_{x'y'}$ are disjoint.*

Now, using this lemma and arranging for $(1-mC/\beta)^{-1}$ not to exceed $3/2$, we have

$$\langle B^m_{xy}\rangle \le 3/2 + 1/2 = 2\ , \tag{12.9}$$

thus completing the proof of the Induction Hypothesis for the case $n_1 = 1$ and $n_2\,, n_3 = 0$. The general case – (9.5) – is done in the same way.

Before starting the proof of Lemma 11 we provide some orientation and motivation. To bound the expectation of $B^m_{xy}\,\chi^c_{xb}$ we could try to use (5.14):

$$B^m_{xy} \le 2^m B^m_{xb} B^m_{by}\ , \tag{12.10}$$

while $\chi^c_{xb} = 1-\chi_{xb}$ can be bounded by

$$\chi^c_{xb} \le B^p_{xb}\, a^{-p} |b-x|^{-\alpha p} \tag{12.11}$$

where we used Def. 3 of Section 8. Since both $|b-x|$ and $|b-y|$ are smaller than $\ell$ it is natural to try to apply the Induction Hypothesis. However we face at least two problems:

**[A].** The Induction Hypothesis does not cover $B_{xb}^{m+p}$ when $p > 0$. Indeed, a factor $\bar{\chi}_{xb} = \prod \chi_{xj} \prod \chi_{bj}$ must be present in order for (9.4) to apply in such a case with $0 < p \leq 2m$. Notice, however, that while we have no immediate control without the missing factors $\chi_{bj}$, the factors $\chi_{xj}$ for $j \in \tilde{R}_{xy}^x$ and $|j-x| < |b-x|$ are already in place. To overcome the problem, we shall introduce the needed factors $\prod \chi_{bj}$ by the same partition of unity scheme that was used above.

Before embarking on that scheme, let us quickly evaluate the situation which emerges after insertion of $\bar{\chi}_{xb}$. We can then choose $p = 2m$, the induction (9.4) applies, and we get a small contribution

$$\begin{aligned} \langle B_{xy}^m \chi_{xb}^c \bar{\chi}_{xb} \rangle &\leq 2^m a^{-p} |b-x|^{-\alpha p} \langle B_{xb}^{m+p} \bar{\chi}_{xb} B_{by}^m \rangle \\ &\leq 4 \cdot 2^m a^{-2m} |b-x|^{-2\alpha m} \end{aligned} \tag{12.12}$$

for $|b-x|$ large. Here we used (9.4) since $xb$ is of Class 2 and $by$ is of Class 1. Note that the expression (12.12) is summable in $b$ for $m\alpha$ large. Moreover, the factor $a^{-2m}$ ensures that also the contributions for $|b-x| = O(1)$ are small.

**[B].** The second problem is that, in order for (9.4) to apply we must make sure that we can find inside $R_{xy}$ two non-overlapping regions $R_{xb}$ and $R_{by}$ of which the former is $C$-admissible and the latter of diamond type. Moreover, since we have $\prod \chi_{xj}$ only for $j \in \tilde{R}_{xy}^x$ we must ensure that $R_{xb}$ is inside the reduced region $\tilde{R}_{xy}^x \cup \tilde{R}_{xy}^y \cup \{x, y\}$ in $R_{xy}$. Since this might have to be repeated many times at smaller and smaller scales, we must be sure that all regions remain $\delta$-admissible (or, put differently, we do not want $\delta$ to be scale-dependent). We will see in the next lemma that this can be arranged.

**Lemma 12** *For a diamond $R_{xy}$ consider the subsets $R_{xy}^x$ and $R_{xy}^y$ of* (7.4).

*[1] Let $w$ be any point in $\tilde{R}_{xy}^x$. Then we can always find a point $a \in R_{xy}$ and regions $R_{xw}$, $R_{wa}$, and $R_{ay}$, such that $R_{wa}$ and $R_{ay}$ are diamonds inside $R_{xy}$, $R_{xw}$ lies inside $\tilde{R}_{xy}^x \cup \{x\}$ and is $C$-admissible, and the three regions have disjoint interiors (see Fig. 5).*

*[2] Let $w$ be any point in $\tilde{R}_{xy}^x$ such that $|w-x| > \beta^{1/4}$ ($xw$ not of Class 3). Let $w_1 w_2$ be a pair in $\tilde{R}_{xy}^x \cup \tilde{R}_{xy}^y \cup \{x, y\}$ such that $|w_1 - w| \leq |w-x|^{1/2}$ and $|w-x|^{1/2} \leq |w_1 - w_2| < |w-x|/5$ (see* (12.18)*). Let $R_{w_1 w_2}$ be the corresponding diamond region. Then we can always find 4 points $a_i \in R_{xy}$ ($i = 1, \ldots, 4$), such that all of the regions $R_{xa_1}$, $R_{a_j a_{j+1}}$ ($j = 1, \ldots, 3$), and $R_{a_4 y}$, are diamonds with disjoint interiors and do not overlap with $R_{w_1 w_2}$ (see Fig. 6). The same can be done for $x \in R_{xy}^y$.*

**Proof.** The most dangerous situations are shown in Figures 5 and 6. It is a simple geometrical argument to see that the region $R_{xw}$ in Fig. 5 is $C$-admissible, as the angles $\theta_1$ and $\theta_2$ are never smaller than $\pi/8$, see (12.1). In the cases shown in Figs. 6a and 6b one has to check that the diamonds do not transgress $R_{xy}$. This never happens since $|w_1 - x| \leq |x-y|/\sqrt{2}$ and $|w_1 - w_2| \leq |w_1 - x|/5$. ∎

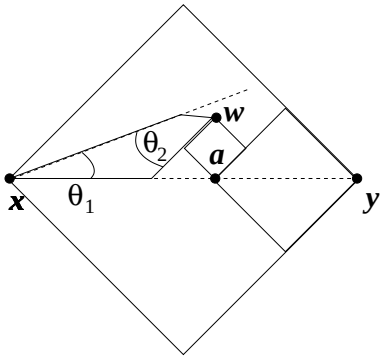


Figure 5: We need one intermediate point $a$. The two angles $\theta_1$ and $\theta_2$ are never smaller than $\pi/8$.

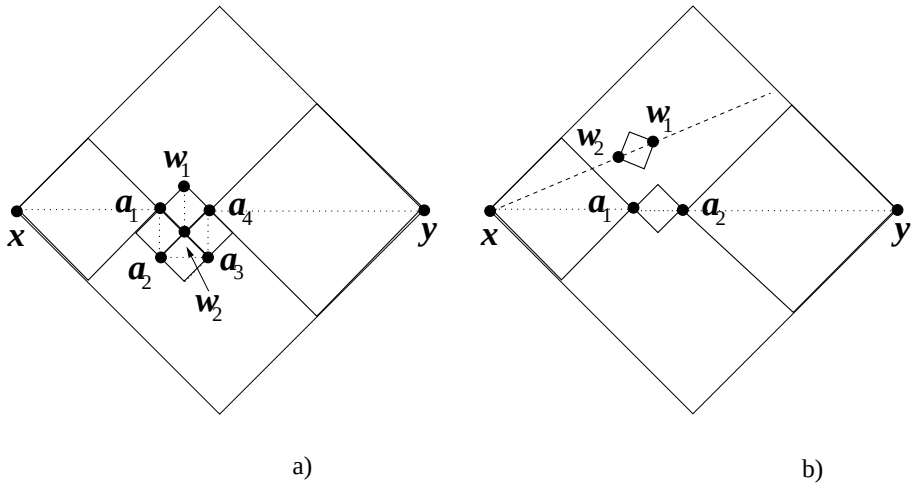


Figure 6: a) If the pair $w_1w_2$ is right in the middle, then we need to add four intermediate points $a_1, \ldots, a_4$. b) Even if the pair $w_1w_2$ is located on the boundary of $\tilde{R}^x_{xy}$, the region $R_{w_1w_2}$ still lies inside $R_{xy}$.

**Proof of Lemma 11.** We split the sum over bad points $b$ in (12.7) into several groups of terms.

**Case 1.** The bad point $b$ is located close to $x$, i.e., $|b-x| \le \beta^{1/4}$. Then we can bound $B^m_{xy}$ by (5.14) and $\chi^c_{xb}$ by (12.11), which gives

$$B^m_{xy}\chi^c_{xb} \;\le\; 2^m a^{-p}|b-x|^{-\alpha p} B^{m+p}_{xb} B^m_{by} \;. \tag{12.13}$$

To apply the Induction Hypothesis we need to select inside the diamond $R_{xy}$ two regions $R_{xb}$ and $R_{by}$. The first one, $R_{xb}$, need only be $C$-admissible (since $xb$ is of Class 3, see Section 9), so it may be a deformed diamond (Fig. 3b). On the other hand, $R_{by}$ has to be diamond-shaped, since $|b-y| > \beta^{1/4}$ (Class 1). To make the requirement of diamond shape conform with our constrained geometry, we must add an intermediate point $a$ as in Fig. 5 with $w = b$, and use

$$B^m_{by} \le 2^m B^m_{ba} B^m_{ay} \;. \tag{12.14}$$

We have seen in Lemma 12 that we can always find such a point $a$, so the induction (9.4) does apply. Note that since $|b-x| < \beta^{1/4}$ there will be no additional induction on $R_{xb}$. Therefore there is no risk that the region might get more and more deformed by the induction steps and $\delta$-admissibility might finally be lost. Thus we have

$$B^m_{xy}\chi^c_{xb} \le 2^{2m} a^{-p}|b-x|^{-\alpha p} B^{m+p}_{xb} B^m_{ba} B^m_{ay} \;. \tag{12.15}$$

The situation for $b$ near $y$ is analogous. Summing the contributions from $b$ near $x$ or $y$ we obtain

$$\begin{aligned}
&\sum_{z=x,y} \sum_{|b-z|\le\beta^{1/4}} \Big\langle B^m_{xy}\, \chi^c_{zb} \prod\nolimits_{j:\,|j-z|<|b-z|} \chi_{zj} \Big\rangle \\
&\le \sum_{z=x,y} \sum_{|b-z|\le\beta^{1/4}} 4^m a^{-2m}|b-x|^{-2\alpha m}\, 2^2(1+\rho) \\
&\le 2^3(1+\rho)\,\frac{4^m}{a^{2m}} K_1 \sum_{|b-x|=1}^{\beta^{1/4}} |b-x|^{2-2\alpha m} \le (4/a^2)^m K_1' \le \frac{\rho}{10}\;,
\end{aligned} \tag{12.16}$$

where in the second line we used (9.4) and $p = 2m$. We can accommodate $m+p = 3m > m$ without any protection factor $\bar\chi$ since $bx$ is Class 3. In the third line, $K_1|b-x|^2$ is the entropy factor for the 3D sum over bad points at distance $|b-x|$, the factor $K_1'$ is a constant of order unity, and we used that $4/a^2 < 1$ and $m > 4d/\alpha$ is large. We bounded the expression by $\rho/10$ for convenience; since both $a$ and $m$ are large, the factor $(4/a^2)^m K_1'$ is in fact very small.

**Case 2.** The first bad point $b$ is far from $x$ (i.e., $|b-x| > \beta^{1/4}$) and also far from $y$. Let us consider the case $b \in \tilde{R}^x_{xy}$ for definiteness. (The other case, $b \in \tilde{R}^y_{xy}$, is treated in the same way.) Again, we have to estimate

$$\left\langle B^m_{xy}\, \chi^c_{xb} \prod\nolimits_{j:\, |j-x|<|b-x|} \chi_{xj} \right\rangle \ . \tag{12.17}$$

As was observed above, if we succeeded in promoting the last product in the average to a complete factor $\bar{\chi}_{xb}$, then we could apply the Induction Hypothesis as in (12.12). In order to satisfy the hypothesis of Lemma 5, Eq. (7.6), we should have a constraint $\chi_{jb}$ for all $|j-b| \le |b-x|/\sqrt{2}$. Actually, from the remark after the proof of that lemma we only require $\chi_{jb}$ for $|j-b| < |b-x|/5$ since we know that all $t_x - t_j$ fluctuations are good up to $|j-x| \approx |b-x|$.

Guided by the idea of partition of unity (cf. (12.4)–(12.5)), we will first check whether there is some large fluctuation $\chi^c > 0$ at large scale near $b$. If no such event occurs, we proceed to the step of checking fluctuations at intermediate distance scales. Then either all intermediate distance fluctuations are good too (and we have the desired factor $\bar{\chi}$), or there must be some bad event at intermediate scale. In this last case we will see that many bad events must happen. We will now make this more precise.

**Case 2a.** The nearest bad point $b$ is far from $x$ (and $y$), $|b-x| > \beta^{1/4}$, and there is a large scale bad event near $b$. This means that $B_{jk} \ge a\,|j-k|^\alpha$ for some pair $j,k \in \tilde{R}^x_{xy} \cup \tilde{R}^y_{xy}$ such that

$$|j-b| \le |b-x|^{1/2} \quad \text{and} \quad |b-x|^{1/2} \le |j-k| \le |b-x|/5 \ . \tag{12.18}$$

Now, using (5.14) and (12.11),

$$B^m_{xy}\chi^c_{jk} \le 2^{4m} B^m_{xa_1} \prod_{i=1}^{3} B^m_{a_i a_{i+1}} B^m_{a_4 y} B^m_{jk}\, |j-k|^{-\alpha m} a^{-m} \ . \tag{12.19}$$

To apply the Induction Hypothesis the corresponding regions must all be diamonds (all pairs are Class 1). By the assumptions made on the pair $jk$, Lemma 12 guarantees that we can choose the four intermediate points $a_i \in R_{xy}$ ($i = 1, \ldots, 4$) so that all of the regions $R_{xa_1}$, $R_{a_j a_{j+1}}$ ($j = 1, \ldots, 3$), and $R_{a_4 y}$, are diamonds with disjoint interiors and do not overlap with $R_{jk}$ (see Fig. 6). Since the regions are non-overlapping and $\ell > |j-k| \ge |b-x|^{1/2}$, our induction hypothesis yields

$$\langle B^m_{xy}\chi^c_{jk}\rangle \le 2^{4m} 2^6 |b-x|^{-\alpha m/2} a^{-m} \ .$$

For large $m$ the value of the sum over $b$ is small.

To estimate the entropy factor, note that there are less than $|b-x|^{d+d/2}$ pairs $jk$ satisfying (12.18). Altogether then, the present partial sum of contributions from $r \equiv |b-x| \ge \beta^{1/4}$ is bounded by

$$(4/a)^m K_2 \sum_{r>\beta^{1/4}} r^{(d-1)+d+d/2-\alpha m/2} = O\big(\beta^{-1/4}\big) < \frac{\rho}{10} \ .$$

Note that there is nothing special or optimal about the exponent $1/4$ of $1/\beta$ – it is just convenient.

**Case 2b.** We now suppose that $|b-x| \geq \beta^{1/4}$ and there is *no* large deviation near $b$, i.e., $B_{jk} \leq a\,|j-k|^\alpha$ holds for all $j,k$ subject to (12.18). This implies that at long scales $|j-k| \geq |b-x|^{1/2}$ we have $\chi_{jk}=1$. It remains to check whether $\chi_{jk}$ holds also at shorter scales $|j-k| \leq |b-x|^{1/2}$.

First we consider the case of there being a point $g$ ($g$ stands for good) in $R_{xy}$ with $|g-b| \leq |b-x|^{1/2}$ such that $\chi_{gh}=1$ holds for all $h$ with $|g-h| \leq |b-x|^{1/2}$. We then have in particular that $\chi_{gb}=1$, and so by Def. 3

$$B_{gb} \leq a\,|b-g|^\alpha \leq a\,(|b-x|^{1/2})^\alpha = a\,|b-x|^{\alpha/2}\ .$$

This inequality combined with the constraint $\chi^c_{xb}=1$ and (5.14) yields

$$2B_{xg} \geq \frac{B_{xb}}{B_{gb}} \geq \frac{a\,|b-x|^\alpha}{a\,|b-x|^{\alpha/2}} = |b-x|^{\alpha/2}\ .$$

Thus we have

$$B^m_{xy}\,\chi^c_{xb}\,\bar\chi_{xg} \leq 2^m (B^m_{xg}\bar\chi_{xg}) B^m_{gy}\,\chi^c_{xb} \leq 2^{3m}(B^{3m}_{xg}\bar\chi_{xg})B^m_{gy}\,|b-x|^{-\alpha m}\ . \quad (12.20)$$

Now we have to be somewhat careful about the choice of the regions $R_{xg}$ and $R_{gy}$, as they may not have the canonical diamond shape. For $R_{xg}$ this is not a problem, because of the presence of $\bar\chi_{xg}$ ($xg$ is of Class 2). All we need is that $R_{xg}$ be $C$-admissible. On the other hand, $R_{gy}$ is (as in Case 1) slightly more delicate. To be sure that we deal with diamond-shaped regions, we add an intermediate point $a$ as in Fig. 5 and use $B^m_{gy} \leq 2^m B^m_{ga} B^m_{ay}$. We have seen in Lemma 12 that it is always possible to find such a point $a$.

The regions $R_{ga}$ and $R_{ay}$ are of diamond type, so induction applies. Note that $R_{xg}$ comes with a $\bar\chi_{xg}$ factor ($xg$ is of Class 2) so no additional induction is required for it. Therefore, as in Case 1, there is no risk that the region might get more and more deformed by the induction steps. It should be emphasized, however, that $\bar\chi_{xg}$ is not exactly the same as in (8.1), but rather is given by

$$\bar\chi_{xg} = \prod_{\tilde R^x_{xy}\ni j\,:\,|j-x|<|b-x|} \chi_{xj} \prod_{\tilde R^x_{xy}\ni j\,:\,|j-g|\leq|b-x|/5} \chi_{gj}\ .$$

Since $|b-g| \leq |b-x|^{1/2}$ and $|b-x| > \beta^{1/4}$ we have $|g-x| \simeq |b-x|$ up to a correction factor of order $O(|b-x|^{-1/2}) \leq O(\beta^{-1/8}) \ll 1$. Therefore $\bar\chi_{xg}$ is equivalent to the following constraints:

$$\begin{aligned} &\forall j \in R_{xg}\,,\ |j-x| \leq |g-x| f_1\ :\quad B_{xj} \leq a\,|j-x|^\alpha\ ,\\ \text{and}\quad &\forall j \in R_{xg}\,,\ |j-g| \leq |g-x| f_2\ :\quad B_{gj} \leq a\,|j-g|^\alpha\ , \end{aligned}$$

with $f_2 = 1/5$ and $f_1 = 1 - O(\beta^{-1/8})$. From Remark 7.3 we know that Lemma 5 and hence Lemma 6 still hold, so we can apply the induction and

$$\left\langle B^m_{xy}\chi^c_{xb}\bar\chi_{xg}\right\rangle \leq 2^{3m}(1+\rho)\,2^2\,|b-x|^{-\alpha m}.$$

There are $O(|b-x|^{d/2})$ choices for $g$, so the sum over these contributions is bounded by

$$2^{3m} K_3 \sum_{r>\beta^{1/4}} r^{d-1+d/2-\alpha m} = O(\beta^{-1/4}) < \frac{\rho}{10} \,. \tag{12.21}$$

**Case 2c.** The last case to consider is the situation where no such point $g$ exists. In that case we can always find a cube $R_n \subset \tilde{R}^x_{xy} \cup \tilde{R}^y_{xy}$ which contains the point $b$ and has side $4^n = |b-x|^{1/2}$ such that $R_n$ is at least at distance $\beta^{1/4} \le 4^n \ll \ell$ from the boundary of $R_{xy}$ and contains no $n$-good point (see Def. 5 in Section 11). Then by (11.4) and (12.11) with $p = m$, we have

$$\chi^c_{R_n} \le \sum_{T\in\mathcal{T}_n} \sum_{\{j_v k_v\}_{v\in L_T}} \prod\nolimits_{v\in L_T} \chi^c_{j_v k_v} \le \sum_{T\in\mathcal{T}_n} \sum_{\{j_v k_v\}_{v\in L_T}} \prod\nolimits_{v\in L_T} \frac{B^m_{j_v k_v}}{a^m |j_v - k_v|^{\alpha m}} \,,$$

where according to (11.3) the sum over configurations of pairs $(j_v, k_v) \in R_v \times \Lambda$ is constrained by $4^{n_v - 1} < |j_v - k_v| \le 4^{n_v}$, with $n_v$ the scale of the leaf $v$. Since all cubes $R_v$ are inside the small region $R_n$ and all pairs $j_v, k_v$ satisfy the conditions of Lemma 12 for the pair $w_1 w_2$, we can proceed as in Case 2b and select 4 intermediate points $a_j$, $j = 1, \ldots, 4$ such that the corresponding regions are diamonds and do not overlap with any $R_{j_v k_v}$ (see Fig. 6). Then by (5.14) we have

$$B^m_{xy} \chi^c_{R_n} \le 2^{4m} B^m_{xa_1} \cdots B^m_{a_4 y} \sum_{T\in\mathcal{T}_n} \sum_{\{j_v k_v\}_{v\in L_T}} \prod\nolimits_{v\in L_T} \frac{B^m_{j_v k_v}}{a^m |j_v - k_v|^{\alpha m}} \,, \tag{12.22}$$

and we can apply the Induction Hypothesis. By Proposition 1 in Section 11 we have

$$\left\langle B^m_{xy} \chi^c_{R_n} \right\rangle \le 2^{4m} 2^{-n\alpha m} \,, \quad n \approx \ln|b-x| \ge \ln(\beta^{1/4}) \,. \tag{12.23}$$

Therefore we have enough decay to control the entropy factors:

$$2^{4m} K_4 \sum_{r>\beta^{1/4}} |b-x|^{-\alpha m} < O(\beta^{-1/4}) < \frac{\rho}{10} \,.$$

This concludes the proof of (12.8). From Eqs. (12.15) and (12.19)–(12.22), we see that $\mathcal{R}(x,y)$ can be written as a sum over products of such $B_{x'y'}$ with $|x'-y'| \le \ell - 1$ that the corresponding regions $R_{x'y'}$ are disjoint. This concludes the proof of Lemma 11 and the Induction Hypothesis. ∎

# 13 Proof of Theorem 2

Now that we have estimated $\langle B^m_{xy} \rangle$ for all $x, y$ we need to estimate $\langle \cosh^p(t_x) \rangle$ for moderate values of $p \le 10$. If we suppose that the field $t$ is pinned at some point $j_0$, so that $t_{j_0} = 0$, then Theorem 2 follows directly from Theorem 1:

$$\langle \cosh^p t_x \rangle = \langle \cosh^p (t_x - t_{j_0}) \rangle \le 2 \,,$$

for any $x$ in the lattice (since Theorem 1 does not require bounds on $\varepsilon$). When the field is not pinned, we need $\varepsilon > 0$ and some conditions on the volume. The rest of this section is devoted to this case. As in the proof of Theorem 1 we will first prove bounds on conditional expectations.

**Definition 8** *A point $x \in \Lambda$ is called 'good at all scales' if*

$$\forall j \in \Lambda \setminus \{x\} \ : \quad B_{xj} \leq a\,|j-x|^{\alpha}$$

*(see also Def. 5 in Section 11). The corresponding characteristic function is*

$$\bar{\chi}_x := \prod\nolimits_{j\in\Lambda\setminus\{x\}} \chi_{xj} \ , \tag{13.1}$$

*where the factors $\chi_{xj}$ are those of Def. 3 (Section 8).*

**Lemma 13** *Let $x$ be good at all scales, and let*

$$B_x = \cosh t_x + \tfrac{1}{2}\mathrm{e}^{t_x} s_x^2 \ . \tag{13.2}$$

*If $\beta \gg 1$ and $\varepsilon \geq 8pa\,L^{-d+\alpha}$, then for any $0 < p \leq O(\beta)$ we have*

$$\langle B_x^p\,\bar{\chi}_x\rangle \leq 2 \ . \tag{13.3}$$

**Proof.** The proof uses a combination of ideas already present in the proofs of Lemma 5 (Section 7) and Lemma 6 (Section 8). By supersymmetry (Proposition 2, Appendix C) we have

$$1 = \left\langle z_x^p \prod\nolimits_{j\in\Lambda\setminus\{x\}} \chi_{xj}^S \right\rangle \ , \tag{13.4}$$

where $z_x$ is defined in (2.11) and $\chi_{xj}^S$ in (8.3). Following exactly the same steps as in the proof of Lemma 6, we obtain the inequality

$$1 \geq \langle B_x^p\,\bar{\chi}_x(1-p\,G_x)\rangle \ , \qquad G_x = \frac{\mathrm{e}^{t_x}}{B_x}[\delta_x\,;D_{\beta,\varepsilon}(t)^{-1}\delta_x] \ , \tag{13.5}$$

if $p\,G_x < 1$. We must now bound the Green's function $G_x$ using the constraint $\bar{\chi}_x$ (as we did in Lemma 5). For this purpose define $\tilde{D} = \mathrm{e}^{-t_x}B_x D_{\beta,\varepsilon}(t)$ by

$$[v\,;\tilde{D}v] = \mathrm{e}^{t_x}B_x\beta\sum\nolimits_{(ij)} \mathrm{e}^{t_i+t_j-2t_x}(v_i-v_j)^2 + \varepsilon B_x \sum\nolimits_k \mathrm{e}^{t_k-t_x}v_k^2 \ ,$$

and note that $B_{xj} \leq a\,|j-x|^{\alpha}$ implies the bound

$$\mathrm{e}^{t_j-t_x} \geq (2a\,|j-x|^{\alpha})^{-1} \ .$$

We then follow the proof of Lemma 5 and introduce a telescopic sum

$$\delta_x = (\delta_x - I_1) + (I_1 - I_2) + \ldots + (I_{N-1} - I_N) + I_N = \sum_{n=0}^{N} \rho_n \ , \tag{13.6}$$

where $I_n$ is the (normalized) indicator function of a cube of center $x$ and side $2^n$, and $\rho_n = I_n - I_{n+1}$. There is no need to introduce $\tilde{I}$ as we did in the proof Lemma 5, as we are now working not on $R_{xy}$ but on the whole volume. The sum terminates on reaching the system size $2^N$. Note that for $n < N$ we have $\sum_j \rho_n(j) = 0$ and $\|\rho_n\|_2^2 \leq 2^{-nd} = 2^{-3n}$. The function $\rho_n$ for $n = N$ is constant: $\rho_N(j) = I_N(j) = |\Lambda|^{-1}$ for all $j \in \Lambda$.

Now, by the Cauchy-Schwarz inequality,

$$[\delta_x\,;\tilde{D}^{-1}\delta_x] \leq \left(\sum_{n=0}^{N}[\rho_n\,;\tilde{D}^{-1}\rho_n]^{1/2}\right)^2 . \tag{13.7}$$

For $n < N$ we use the bound $[\rho_n\,;\tilde{D}^{-1}\rho_n] \leq \|(\tilde{D}^{-1})_{\rho_n}\|\,\|\rho_n\|_2^2$, where $\|(\tilde{D}^{-1})_{\rho_n}\| \leq 2^{2n+2n\alpha}c_1$. Then for the sum of terms with $n < N$ we have

$$\sum_{n=0}^{N-1}[\rho_n\,;\tilde{D}^{-1}\rho_n]^{1/2} \leq \frac{c_1}{\sqrt{\beta}}\sum_{n=0}^{N-1}\sqrt{2}^{\,-n(d-2-2\alpha)} < \frac{\gamma_1}{\sqrt{\beta}} \tag{13.8}$$

uniformly in $N$ since $d = 3$ and $2\alpha \ll 1$. For $n = N$, on the other hand, we no longer have orthogonality to the constant functions ('zero mode') and therefore must take recourse to the $\varepsilon$-term in $\tilde{D}$ :

$$[\rho_N\,;\tilde{D}^{-1}\rho_N] \leq \left(\varepsilon|\Lambda|\min_{j\in\Lambda}\mathrm{e}^{t_j-t_x}\right)^{-1} \leq 2a\frac{L^\alpha}{\varepsilon L^d} \leq \frac{1}{4p} . \tag{13.9}$$

Hence $G_x \leq (\gamma_1/\sqrt{\beta} + 1/\sqrt{4p}\,)^2 < 1/(2p)$ for $\beta \gg 1$. So,

$$\tfrac{1}{2}\langle B_x^p\,\bar{\chi}_x\rangle \leq \langle B_x^p\,\bar{\chi}_x(1 - p\,G_x)\rangle \leq 1 \tag{13.10}$$

by (13.5), and the lemma is proved. ∎

With this lemma we can finally complete the proof of Theorem 2, i.e. the bound on the unconditional expectation of $\cosh^p t_x$ .

**Proof of Theorem 2.** We recall from Def. 5 (Section 11) that a point $x$ is said to be $n$-good if $B_{xj} \leq a\,|j-x|^\alpha$ for all $j \in \Lambda$ subject to $1 \leq |j-x| \leq 4^n$. A point $x$ is good at all scales if $B_{xj} \leq a\,|j-x|^\alpha$ for all $j \in \Lambda \setminus \{x\}$; we then say that $x$ is $N$-good.

We proceed as in Lemma 11 (Section 12):

$$\langle\cosh^p t_x\rangle = \langle\bar{\chi}_x\cosh^p t_x\rangle + \langle\bar{\chi}_x^c\cosh^p t_x\rangle\ , \tag{13.11}$$

where $\bar{\chi}_x$ ensures that the point $x$ is $N$-good. Then by Lemma 13 we have

$$\langle\bar{\chi}_x\cosh^p t_x\rangle \leq 2\ . \tag{13.12}$$

It remains to estimate the second term, $\langle\bar{\chi}_x^c\cosh^p t_x\rangle$. We prove in Lemma 14 below that this term is bounded by a constant. Once this has been accomplished, the proof of Theorem 2 will be finished. ∎

**Lemma 14** *Let $\bar{\chi}_x^c = 1 - \bar{\chi}_x$ , with $\bar{\chi}_x$ defined by* (13.1). *Let $\beta \gg 1$ and $\varepsilon \geq 8\cdot 4\cdot 10\,a\,L^{\alpha-d}$. Then for any $0 \leq p \leq 10$ we have*

$$\langle\bar{\chi}_x^c\cosh^p t_x\rangle \leq 1/2\ . \tag{13.13}$$

**Proof.** While $\bar\chi^c_x$ means that $x$ is not $N$-good, it is still possible for other points in $\Lambda$ to be $N$-good. If $g \neq x$ is the nearest such point (as seen from $x$), then none of the points inside the ball $K^x_{|g-x|}$ of radius $|g-x|$ and center $x$ is $N$-good. Denoting the indicator function for the latter event by $\chi^c_{K^x_{|g-x|}}$ we have the identity

$$\bar\chi^c_x = \sum_{\Lambda \ni g \neq x} \bar\chi_g \, \chi^c_{K^x_{|g-x|}} + \prod_{j\in\Lambda} \bar\chi^c_j \, , \tag{13.14}$$

where the last term accounts for the possibility that there is no $N$-good point in $\Lambda$ at all. Thus we obtain the decomposition

$$\langle \bar\chi^c_x \cosh^p t_x \rangle = \sum_{g\neq x} \Big\langle \bar\chi_g \, \chi^c_{K^x_{|g-x|}} \cosh^p t_x \Big\rangle + \Big\langle \prod\nolimits_j \bar\chi^c_j \cosh^p t_x \Big\rangle \, . \tag{13.15}$$

We will prove that both of these two terms are bounded by 1/4.

**1.** We consider the first sum. Using $\cosh t_x \leq 2\cosh(t_x - t_g)\cosh t_g$ and applying the Cauchy-Schwarz inequality twice, we obtain

$$\begin{aligned}
&\big\langle \chi^c_{K^x_{|g-x|}} \bar\chi_g \cosh^p t_x \big\rangle \leq \big\langle \chi^c_{K^x_{|g-x|}} \bar\chi_g \big\rangle^{1/2} \big\langle \bar\chi_g \cosh^{2p} t_x \big\rangle^{1/2} \\
&\leq \big\langle \chi^c_{K^x_{|g-x|}} \bar\chi_g \big\rangle^{1/2} \, 2^p \big\langle \bar\chi_g \cosh^{2p}(t_x - t_g) \cosh^{2p} t_g \big\rangle^{1/2} \\
&\leq 2^p \, \big\langle \chi^c_{K^x_{|g-x|}} \bar\chi_g \big\rangle^{1/2} \big\langle \bar\chi_g \cosh^{4p} t_g \big\rangle^{1/4} \big\langle \cosh^{4p}(t_x - t_g) \big\rangle^{1/4} \\
&\leq 2^p \, c_p \, \big\langle \chi^c_{K^x_{|g-x|}} \bar\chi_g \big\rangle^{1/2} ,
\end{aligned}$$

where in the last step we used Lemma 13 and Theorem 1, and we introduced $c_0 = 1$ and $c_p = 2^{1/2}$ for $p \geq 1$. It remains to bound

$$2^p c_p \sum_g \big\langle \chi^c_{K^x_{|g-x|}} \bar\chi_g \big\rangle^{1/2} = 2^p c_p \sum_{n\geq 0} \sum_{4^n \leq |g-x| < 4^{n+1}} \big\langle \chi^c_{K^x_{|g-x|}} \bar\chi_g \big\rangle^{1/2} \, . \tag{13.16}$$

Let $R^x_n$ be the cube centered at $x$ of side $4^n$. Now fixing a point $g$ with $4^n \leq |g-x| < 4^{n+1}$ we have $R^x_n \subseteq K^x_{|g-x|}$ , and we distinguish between two cases:

**1a.** The interior of $K^x_{|g-x|}$ is void not only of $N$-good points but also of $n$-good points. Let $\chi^c_n$ denote the corresponding indicator function. Then for $n \geq 1$, using Proposition 1 we have

$$\big\langle \chi^c_{K^x_{|g-x|}} \bar\chi_g \, \chi^c_n \big\rangle \leq \big\langle \chi^c_{R^x_n} \big\rangle \leq 2^{-(n+1)\alpha m} \, , \tag{13.17}$$

where $\chi^c_{R^x_n}$ is given in Def. 6, Section 11. For $n = 0$ the cube $R^x_0$ contains only the point $x$, so

$$\big\langle \chi^c_{R^x_0} \big\rangle \leq \sum_{|z-x|=1} \big\langle \chi_{xz} \big\rangle \leq \sum_{|z-x|=1} \frac{\big\langle B^m_{xz} \big\rangle}{a^m} \leq 2\frac{2d}{a^m} < 2^{-\alpha m} \, .$$

**1b.** There is at least one $n$-good point $y$ inside $K^x_{|g-x|}$. Let $\chi_n(y)$ be the corresponding indicator function. Because the point $y$ cannot be $N$-good, there must be a first scale $q > n$ so that $y$ is $q$-bad. Thus there exists a first point $b$ at distance $|b-y| > 4^n$ with $B_{yb} > a\,|b-y|^\alpha$. It follows that

$$\sum_{y\in K^x_{|g-x|}} \big\langle \bar\chi_g \chi^c_{K^x_{|g-x|}} \chi_n(y)\big\rangle \le \sum_{y\in K^x_{|g-x|}} \sum_{b:\,|b-y|>4^n} \frac{\langle B^m_{yb}\rangle}{a^m|b-y|^{\alpha m}}$$
$$\le \frac{4^{nd}k_1}{a^m} \sum_{r>4^n} \frac{2\,r^{d-1}k_2}{r^{\alpha m}} \le 2^{-n\alpha m} a^{-m} \le 2^{-(n+1)\alpha m}\ ,$$

where the factor $4^{nd}k_1$ comes from the sum over $y$ and $r^{d-1}k_2$ comes from the sum over $b$. Inserting these results into (13.16) we obtain

$$2^p c_p \sum_{n\ge 0} \sum_{4^n\le|g-x|\le 4^{n+1}} \big\langle \chi^c_{K^x_{|g-x|}} \bar\chi_g\big\rangle^{1/2} \le 2^p c_p \sum_{n\ge 0} 4^{(n+1)d} k_3 \big(2\cdot 2^{-(n+1)\alpha m}\big)^{1/2}\ ,$$

where $4^{(n+1)d}k_3$ comes from the sum over $g$. This will be no greater than $1/4$ provided that $\alpha m$ is large enough.

**2.** To complete the proof, we have to estimate the last term $\langle \prod_j \bar\chi^c_j \cosh^p t_x\rangle$ in (13.15). By Proposition 1 the probability for no $N$-good point to be found in a cube $\Lambda$ of side $L = 4^N$ is bounded by $2^{-N\alpha m} = L^{-\alpha m/2}$. Hence

$$\big\langle \prod\nolimits_j \bar\chi^c_j \cosh^p t_x\big\rangle \le \big\langle \cosh^{2p} t_x\big\rangle^{1/2} \big\langle \prod\nolimits_j \bar\chi^c_j\big\rangle^{1/2} \le \big\langle \cosh^{2p} t_x\big\rangle^{1/2} L^{-\alpha m/4}\ .$$

To get a bound on the expected value of $\cosh^{2p} t_x$ we once again use supersymmetry (Proposition 2), as follows:

$$\mathrm{e}^{\gamma\varepsilon} = \langle \mathrm{e}^{\gamma\varepsilon z_x}\rangle = \langle \mathrm{e}^{\gamma\varepsilon B_x}(1-\gamma\varepsilon\, G'_x)\rangle\ ,$$

where we choose $0<\gamma<1/2$, and $G'_x = \mathrm{e}^{t_x}[\delta_x\,; D_{\beta,\varepsilon}(t)^{-1}\delta_x]$. Since the operator $D_{\beta,\varepsilon}(t) - \varepsilon\,\mathrm{e}^{t_x}\delta_x[\delta_x\,;\cdot]$ is non-negative, by Lemma 1 we have $\varepsilon\, G'_x \le 1$, so

$$\langle \mathrm{e}^{\gamma\varepsilon(B_x-1)}\rangle \le (1-\gamma)^{-1}\ .$$

Also, $\cosh^{2p} t_x \le (2p)!\,(\gamma\varepsilon)^{-2p}\,\mathrm{e}^{\gamma\varepsilon B_x}$ by an elementary computation, and hence

$$\langle \cosh^{2p} t_x\rangle \le (2p)!\,(\gamma\varepsilon)^{-2p}\langle \mathrm{e}^{\gamma\varepsilon B_x}\rangle \le O(\varepsilon^{-2p})\ .$$

We thus finally obtain

$$\big\langle \prod\nolimits_j \bar\chi^c_j \cosh^p t_x\big\rangle \le \big\langle \cosh^{2p} t_x\big\rangle^{1/2} L^{-\alpha m/4} \le O(\varepsilon^{-p}) L^{-\alpha m/4} < 1/4\ , \quad (13.18)$$

since $\alpha m$ is large and $\varepsilon \ge L^{\alpha-d}$. This concludes the proof of Lemma 14. ∎

**Remark.** In the proof of Theorem 2 the $\varepsilon$ term (zero mode) appears only in two places: (13.9) of Lemma 13 (the last term in the telescopic sum) and (13.18) (when no $N$ good point is present). The inequality (13.9) is the reason why we cannot take $\varepsilon = O(L^{-d})$ but must take $\varepsilon = O(L^{\alpha-d})$.

# 14 Proof of Theorem 3

Finally we can prove the bound on the Green's function $C_{xy}$ of (1.9). Let $f$ be such that $f(j) \geq 0$ for all $j \in \Lambda$. We need to estimate

$$[f; Cf] = \big\langle [\mathrm{e}^t f; D_{\beta,\varepsilon}(t)^{-1} \mathrm{e}^t f] \big\rangle = \langle [W; G_t W] \rangle \ , \tag{14.1}$$

where $D_{\beta,\varepsilon}(t)^{-1} = G_t$ was defined in (1.1), and $W(j) = \mathrm{e}^{t_j} f(j)$.

## 14.1 Upper bound

Let $L_0 = -\beta\Delta + \varepsilon$ and $G_0 = L_0^{-1}$ (as defined in the statement of the Theorem). Now

$$\begin{aligned}
[W; G_t W] &= [L_0 G_0 W; G_t W] = \beta\, [\nabla(G_0 W); \nabla(G_t W)] + \varepsilon\, [G_0 W; G_t W] \\
&= \beta \sum_{(jj')} \nabla_{jj'}(G_0 W) \nabla_{jj'}(G_t W) + \varepsilon \sum_j (G_0 W)(j)(G_t W)(j) \\
&= \beta \sum_{(jj')} \frac{\nabla_{jj'}(G_0 W)}{\mathrm{e}^{(t_j+t_{j'})/2}} \frac{\nabla_{jj'}(G_t W)}{\mathrm{e}^{-(t_j+t_{j'})/2}} + \varepsilon \sum_j \left( \frac{(G_0 W)(j)}{\mathrm{e}^{+t_j/2}} \right) \left( \frac{(G_t W)(j)}{\mathrm{e}^{-t_j/2}} \right) .
\end{aligned}$$

Since $|a \cdot b + c \cdot d| \leq (a \cdot a + c \cdot c)^{1/2} (b \cdot b + d \cdot d)^{1/2}$ we have

$$[W; G_t W] \leq \left( \beta \sum_{(jj')} \frac{|\nabla_{jj'}(G_0 W)|^2}{\mathrm{e}^{t_j + t_{j'}}} + \varepsilon \sum_j \frac{|(G_0 W)(j)|^2}{\mathrm{e}^{t_j}} \right)^{1/2} [W; G_t W]^{1/2}.$$

Therefore[2]

$$[W; G_t W] \leq \beta \sum_{(jj')} \frac{|\nabla_{jj'}(G_0 W)|^2}{\mathrm{e}^{t_j + t_{j'}}} + \varepsilon \sum_j \frac{|(G_0 W)(j)|^2}{\mathrm{e}^{t_j}} \ . \tag{14.2}$$

Now

$$|\nabla_{jj'}(G_0 W)| \leq \sum_k |(G_0(j,k) - G_0(j',k))|\, W(k) \leq \mathrm{const} \sum_k H_{jk} W(k) \ , \tag{14.3}$$

where we defined $H_{jk} = \beta^{-1}(|j-k|^2 + 1)^{-1} \mathrm{e}^{-\tilde{\varepsilon}|j-k|}$, $\tilde{\varepsilon} = (\varepsilon/2\beta)^{1/2}$, and we used

$$|(G_0(j,k) - G_0(j',k))| \leq \mathrm{const}\, H_{jk} \ .$$

[2] We thank S.R.S. Varadhan for explaining the inequality (14.2) to us.

By inserting (14.3) into (14.2) we get

$$[f;Cf] \le \text{const}\,\beta \sum_{(j,j'),k,l} H_{jk}H_{jl}\, f(k)f(l)\left\langle \mathrm{e}^{(t_k+t_l-t_j-t_{j'})}\right\rangle$$
$$+\,\varepsilon \sum_{j,k,l} G_0(j,k)\, G_0(j,l)\, f(k)f(l)\left\langle \mathrm{e}^{t_k+t_l-t_j}\right\rangle . \tag{14.4}$$

By Theorems 1 and 2 the expectation over the field $t$ is uniformly bounded. Now we can sum over $j$ :

$$\sum_j H_{jk}H_{jl} \le \text{const}\ \tilde{G}_0(k,l), \qquad \sum_j G_0(j,k)\, G_0(j,l) = G_0^2(k,l), \tag{14.5}$$

where $\tilde{G}_0 = (-\beta\Delta + \varepsilon/2)^{-1}$. Note that $G_0 \le \tilde{G}_0$ . We finally obtain

$$[f;Cf] \le \text{const}\ [f;\tilde{G}_0 f] + \varepsilon[f;G_0^2 f] \le 2\,\text{const}\,[f;\tilde{G}_0 f]\ . \tag{14.6}$$

This completes our proof of the upper bound. ■

## 14.2 Lower bound

Let $\bar{\chi}_x$ be the characteristic function ensuring that $x \in \Lambda$ is good at all scales (see (13.1)). Recall that if $\bar{\chi}_x > 0$ then

$$\mathrm{e}^{t_j - t_x} \ge (2a(1+|j-x|^\alpha))^{-1} \tag{14.7}$$

for all $j \in \Lambda$. We have the inequality $1 = \bar{\chi}_x + \bar{\chi}_x^c \ge \bar{\chi}_x$ . Inserting it into (14.1) we obtain

$$\langle [W;G_tW]\rangle \ge \langle \bar{\chi}_x[W;G_tW]\rangle = \sum_{jk} \left\langle \bar{\chi}_x W(j)W(k)\, G_t(j,k)\right\rangle \tag{14.8}$$
$$\ge \frac{1}{4a^2}\sum_{jk} \tilde{f}(j)\tilde{f}(k)\left\langle \bar{\chi}_x \bar{D}_t^{-1}(j,k)\right\rangle = \frac{1}{4a^2}\left\langle \bar{\chi}_x[\tilde{f};\bar{D}_t^{-1}\tilde{f}]\right\rangle ,$$

where $\tilde{f}(j) = (1+|j-x|^\alpha)^{-1}f(j)$ and $\bar{D}_t^{-1} = \mathrm{e}^{2t_x}G_t = (\mathrm{e}^{-2t_x}D_{\beta,\varepsilon}(t))^{-1}$. In the first line we used the fact that $G_t$ is positive as a quadratic form for each configuration of $t$. In the second line we used the fact that this is a sum of positive terms since $W(j) \ge 0$ and $G_t$ is pointwise positive. Furthermore, we applied (14.7) to estimate $W(j)$. Now,

$$\left\langle \bar{\chi}_x[\tilde{f};\bar{D}_t^{-1}\tilde{f}]\right\rangle = \langle \bar{\chi}_x\rangle\, \mathrm{E}\big([\tilde{f};\bar{D}_t^{-1}\tilde{f}]\big) \ge \langle \bar{\chi}_x\rangle\, [\tilde{f};\mathrm{E}(\bar{D}_t)^{-1}\tilde{f}] \tag{14.9}$$

where

$$\mathrm{E}(\cdot) = \frac{\langle \bar{\chi}_x\, \cdot\rangle}{\langle \bar{\chi}_x\rangle}$$

is a probability distribution and we used Jensen's inequality.

In order to complete the proof we need to estimate $\langle \bar{\chi}_x \rangle$ and $\langle \bar{\chi}_x \bar{D}_t \rangle$. From Lemma 14 (in the previous section) with $p = 0$ we know that

$$\langle \bar{\chi}_x \rangle = 1 - \langle \bar{\chi}_x^c \rangle \geq 1/2 \, .$$

Moreover $\langle \bar{\chi}_x \bar{D} \rangle \leq \langle \bar{D} \rangle$ as a quadratic form and for any function $u$ we have

$$\begin{aligned} [u; \langle \bar{D} \rangle \, u] &= \beta \sum_{(jk)} (u(j) - u(k))^2 \, \langle \mathrm{e}^{t_j + t_k - 2t_x} \rangle + \varepsilon \sum_j u(j)^2 \, \langle \mathrm{e}^{t_j - 2t_x} \rangle \\ &\leq \beta c \sum_{(jk)} (u(j) - u(k))^2 + \varepsilon c' \sum_j u(j)^2 \leq c_1 [u; G_0 u] \, , \end{aligned} \tag{14.10}$$

where we applied Theorems 1 and 2, and $c_1 = \sup\{c, c'\}$. Thus $\langle \bar{D} \rangle \leq c_1 G_0$. By applying these relations we see that $\mathrm{E}(\bar{D}_t) \leq 2c_1 G_0$ and hence

$$[f; Cf] \geq \frac{1}{4a^2 c_1} \, [\tilde{f}; G_0 \tilde{f}] \, . \tag{14.11}$$

This concludes the proof of Theorem 3. ■

**Remark.** If $W$ did not depend on $t$ we would have the quadratic form estimate

$$[W; G_0 W] \, c_1 \leq \langle [W; G_t W] \rangle \leq c_2 \, [W; G_0 W]$$

with

$$c_2 = \sup_{(jj'),k} \left( \langle \mathrm{e}^{-t_j - t_{j'}} \rangle \, , \, \langle \mathrm{e}^{-t_k} \rangle \right) \, , \quad c_2 = \sup_{(jj'),k} \left( \langle \mathrm{e}^{t_j + t_{j'}} \rangle \, , \, \langle \mathrm{e}^{t_k} \rangle \right) \, .$$

The upper bound follows directly from (14.2), the lower bound from Jensen's inequality.

# Appendices

## A Minimum of the effective action

Let $j \mapsto t_j \equiv \bar{t} \in \mathbb{R}$ (for all $j \in \Lambda$) be a constant field configuration. Evaluating the statistical weight function on it we get

$$\mathrm{e}^{-\varepsilon |\Lambda| (\cosh \bar{t} - 1)} \, \mathrm{Det}^{1/2} (-\beta \Delta + \varepsilon \, \mathrm{e}^{-\bar{t}}) \, . \tag{A.1}$$

Let $t^*$ be the number that maximizes this statistical weight. The condition for the first derivative to vanish at $t^*$ is

$$2 \sinh t^* = -\mathrm{e}^{-t^*} G_0(x, x) \qquad (x \in \Lambda) \, , \tag{A.2}$$

where $G_0 \equiv (-\beta\Delta + \varepsilon\, \mathrm{e}^{-t^*})^{-1} \geq 0$. Equivalently, $1 - \mathrm{e}^{2t^*} = G_0(x,x)$, and since $G_0$ is non-negative, it follows that $t^* \leq 0$. We thus infer that

$$0 \leq 1 - \mathrm{e}^{2t^*} = G_0(x,x) \leq 1 \,. \tag{A.3}$$

Next, we show that the constant field $t^*$ maximizes the integrand over the full set of all field configurations $t = \{t_j\}$. For this, we recall the definition (1.2) of the effective action or free energy $F_{\beta,\varepsilon}$ in combination with (1.10):

$$\begin{aligned} F_{\beta,\varepsilon}(t) &= \beta \sum\nolimits_{(ij)} (\cosh(t_i - t_j) - 1) + \varepsilon \sum\nolimits_k (\cosh t_k - 1) \\ &- \ln \mathrm{Det}^{1/2}\big( -\beta\Delta + \beta V(t) + \varepsilon\, \mathrm{e}^{-t} \big) \,. \end{aligned}$$

Now we introduce $A := G_0^{1/2}\big(\beta V(t) + \varepsilon\, (\mathrm{e}^{-t} - \mathrm{e}^{-t^* \mathrm{Id}})\big)\, G_0^{1/2}$ and write

$$\mathrm{Det}\big( -\beta\Delta + \beta V(t) + \varepsilon\, \mathrm{e}^{-t} \big) = \mathrm{Det}(G_0^{-1})\, \mathrm{Det}(\mathrm{Id} + A) \,.$$

Using $\ln \mathrm{Det}(\mathrm{Id} + A) \leq \mathrm{Tr}\, A$ we then obtain

$$\begin{aligned} F_{\beta,\varepsilon}(t) &\geq -\ln \mathrm{Det}(G_0^{-1/2}) + \varepsilon |\Lambda|\, (\cosh t^* - 1) \\ &+ \beta \sum\nolimits_{(ij)} (\cosh(t_i - t_j) - 1) - G_0(x,x)\, \tfrac{1}{2} \sum\nolimits_k \beta V_{kk} \\ &+ \varepsilon \sum\nolimits_k \big( \cosh t_k - \cosh t^* - \tfrac{1}{2}(\mathrm{e}^{-t_k} - \mathrm{e}^{-t^*})\, G_0(x,x) \big) \,. \end{aligned}$$

The second line of the r.h.s. is non-negative by $\frac{1}{2}\sum_j V_{jj} = \sum_{(ij)}(\cosh(t_i - t_j) - 1)$ and $G_0(x,x) \leq 1$, and so is the third line by the identity $G_0(x,x) = 1 - \mathrm{e}^{2t^*}$ and a trivial computation. This proves that $F_{\beta,\varepsilon}(t)$ is bounded from below by $F_{\beta,\varepsilon}(t^*) = -\ln \mathrm{Det}(G_0^{-1/2}) + \varepsilon|\Lambda|\, (\cosh t^* - 1)$.

# B Hyperbolic symmetry

In Section 4 we explained that the $\mathrm{H}^{2|2}$ nonlinear sigma model in the limit of vanishing regularization $\varepsilon \to 0+$ acquires a global symmetry by the Lorentz group $\mathrm{SO}(1,2)$. We will now exhibit the Ward identities due to this Lorentzian symmetry $\mathrm{SO}(1,2)$. (Consequences due to the supersymmetries of model will be explored in Appendix C.) To prepare the discussion, the reader is invited to recall the expressions (2.7) for the functions $x, y, \xi, \eta$ in horospherical coordinates. We also recall that $z = \cosh t + \mathrm{e}^t(\frac{1}{2}s^2 + \bar\psi\psi)$.

We now seek the first-order differential operator, $L_1$, generating Lorentz boosts in the $zx$-plane, i.e.,

$$L_1\, z = x \,, \quad L_1\, x = z \,, \quad L_1\, y = L_1\, \xi = L_1 \eta = 0 \,.$$

It is easy to verify that the unique operator with these properties is

$$L_1 = \partial_t - \bar\psi\partial_{\bar\psi} - \psi\partial_\psi - s\partial_s \,. \tag{B.1}$$

Similarly, the generator $L_2$ of Lorentz boosts in the $zy$-plane and the generator $L_0$ of Euclidean rotations in the $xy$-plane, are expressed by

$$L_2 = s\left(\partial_t - \bar\psi\partial_{\bar\psi} - \psi\partial_\psi\right) + \tfrac{1}{2}\left(1 + \mathrm{e}^{-2t} - s^2 + 2\bar\psi\psi\right)\partial_s \ ,$$
$$L_0 = s\left(-\partial_t + \bar\psi\partial_{\bar\psi} + \psi\partial_\psi\right) + \tfrac{1}{2}\left(1 - \mathrm{e}^{-2t} + s^2 - 2\bar\psi\psi\right)\partial_s \ .$$

Being the generators of the Lie algebra $\mathfrak{so}_{1,2}$ of the Lorentz group, the operators $L_0, L_1, L_2$ satisfy the commutation relations:

$$[L_0, L_1] = -L_2 \ , \quad [L_0, L_2] = L_1 \ , \quad [L_1, L_2] = L_0 \ .$$

In particular, the generator $L_0 + L_2 = [L_1, L_0 + L_2]$ is the generator of translations of the coordinate $s$.

So far, we have been concerned with the case of a single site. To pass to a lattice $\Lambda$ with many sites, we take the sum

$$L_a = \sum\nolimits_{j\in\Lambda} L_a(j) \quad (a = 0, 1, 2)$$

of differential operators over all sites.

By construction, the $\mathfrak{so}_{1,2}$ operators $L_a = \sum L_a(j)$ are symmetries of the Berezin measure $D\mu_\Lambda$. Therefore, they give rise to Ward identities:

$$0 = \int D\mu_\Lambda\, L_a\left(\mathrm{e}^{-A_{\beta,\varepsilon}} F\right) = \langle L_a F - F L_a A_{\beta,\varepsilon}\rangle \quad (a = 0, 1, 2) \ , \tag{B.2}$$

which hold for any observable $F$ as long as these expectations exist. By computing the symmetry-breaking terms from the formula $L_a A_{\beta,\varepsilon} = \varepsilon \sum_j L_a(j)\, z_j$ one obtains these Ward identities in the more explicit form

$$\langle L_1 F\rangle = \varepsilon \sum\nolimits_j \left\langle (\sinh t_j - \tfrac{1}{2}\mathrm{e}^{t_j} s_j^2 - \mathrm{e}^{t_j}\bar\psi_j\psi_j) F\right\rangle \ ,$$
$$\langle L_2 F\rangle = \varepsilon \sum\nolimits_j \left\langle \mathrm{e}^{t_j} s_j F\right\rangle \ , \qquad \langle L_0 F\rangle = 0 \ .$$

The sum rule (4.6) now follows from the identity for $\langle L_2 F\rangle$ by taking $F = \mathrm{e}^{t_i} s_i$ and performing the Gaussian integrals over the fields $\psi$, $\bar\psi$, and $s$.

Another important consequence results from making the choice $F = \mathrm{e}^{t_j} s_j$. Since $L_0 F = -\sinh t_j + \mathrm{e}^{t_j}(\frac{1}{2}s_j^2 + \bar\psi_j\psi_j)$, it follows from $\langle L_0 F\rangle = 0$ that

$$\langle \mathrm{e}^{t_j}\rangle = \langle \cosh t_j + \sinh t_j\rangle = \langle \cosh t_j + \mathrm{e}^{t_j}(\tfrac{1}{2}s_j^2 + \bar\psi_j\psi_j)\rangle = \langle z_j\rangle = 1 \ . \tag{B.3}$$

The last step, $\langle z_j\rangle = 1$, is by Proposition 2 of Appendix C.

# C SUSY Ward identities

The action function of our $\mathrm{H}^{2|2}$ model has a global symmetry w.r.t. the Lie superalgebra $\mathfrak{g} := \mathfrak{osp}_{2|2}$ (for any $\varepsilon \geq 0$). As a result, there exist supersymmetric

Ward identities for suitable ($\mathfrak{osp}_{2|2}$ invariant) observables. Although such identities are standard material from the theory of localization of supersymmetric integrals [17], we nonetheless give their derivation for completeness here, as the said identities play a central role in our analysis. The essence of the argument can already be understood at the very special example of a lattice $\Lambda$ consisting of just a single site. For pedagogical reasons we first handle this simple situation and then, in a second step, give the generalization to arbitrary lattices.

The treatment will be most transparent if we do all calculations using the coordinates $x, y, \xi, \eta$ described at the beginning of Section 2. As stated there, for our purposes we may view $\mathfrak{osp}_{2|2}$ as the space of first-order differential operators $D$ with coefficients that are linear in the variables $x, y, \xi, \eta$ and the property

$$DH = 0$$

of annihilating the quadratic polynomial

$$H = x^2 + y^2 + 2\xi\eta \ .$$

Let $Q$ be the distinguished first-order differential operator defined by

$$Q = x\partial_\eta - y\partial_\xi + \xi\partial_x + \eta\partial_y \ . \tag{C.1}$$

Clearly $Q$ is odd, converting even coordinate generators $x, y$ into odd generators $\xi, \eta$ and vice versa. $Q$ is also seen to annihilate $H$, and thus represents an element of $\mathfrak{osp}_{2|2}$ . Notice that $Q$ squares to

$$Q^2 = x\partial_y - y\partial_x + \xi\partial_\eta - \eta\partial_\xi \ ,$$

which is a generator from the Lie algebra part $\mathfrak{o}_2 \oplus \mathfrak{sp}_2$ of $\mathfrak{osp}_{2|2}$ .

Now recall from Section 2 that our Berezin superintegration form is

$$D\mu = (2\pi)^{-1} dxdy\, \partial_\xi\partial_\eta \circ (1+H)^{-1/2} \ .$$

**Lemma 15** *The Berezin superintegration form $D\mu$ is $Q$-invariant, i.e.,*

$$\int_{\mathbb{R}^2} D\mu\, Qf = 0$$

*for any bounded smooth superfunction $f = f(x, y, \xi, \eta)$.*

**Proof.** Since $Q$ is a first-order differential operator, we have from $QH = 0$ that $Q(1+H)^{-1/2} = 0$. Therefore, $D\mu\, Qf = D\mu\, (1+H)^{1/2} Q (1+H)^{-1/2} f$ and

$$\int_{\mathbb{R}^2} D\mu\, Qf = (2\pi)^{-1} \int_{\mathbb{R}^2} dxdy\, \partial_\xi\partial_\eta\, (x\partial_\eta - y\partial_\xi + \xi\partial_x + \eta\partial_y)\, (1+H)^{-1/2} f \ .$$

The desired result now follows because $\partial_\xi^2 = \partial_\eta^2 = 0$ and the integral over $\mathbb{R}^2$ of the total derivatives $\partial_x (1+H^2)^{-1/2} f$ and $\partial_y (1+H^2)^{-1/2} f$ vanishes. ∎

An important property of the differential operator $Q$ is that the joint zero locus of its coefficients is the origin $x = y = 0$ and $\xi = \eta = 0$. Denoting the origin by $o$ we write $f(x = 0, y = 0, \xi = 0, \eta = 0) \equiv f(o)$.

**Lemma 16** *Let $f = f(x, y, \xi, \eta)$ be a smooth superfunction which satisfies the invariance condition $Qf = 0$ and decreases sufficiently fast at infinity in order for the integral $\int_{\mathbb{R}^2} D\mu\, f$ to exist. Then*

$$\int_{\mathbb{R}^2} D\mu\, f = f(o)\ .$$

**Proof.** The idea is to 'deform' the integrand $f$ (without changing the integral) by a factor that localizes the integral at $o$. We will do this deformation by multiplication with $\mathrm{e}^{-\tau H}$ for some positive real parameter $\tau$. Thus we are going to show that

$$\int D\mu\, f = \int D\mu\, \mathrm{e}^{-\tau H} f\ , \tag{C.2}$$

independent of $\tau \geq 0$. The desired result will then follow by taking $\tau \to +\infty$.

We begin by observing that the localizing function $H$ is $Q$-exact: it can be written as $H = Q\lambda$ with

$$\lambda := x\,\eta - y\,\xi$$

an odd superfunction. Next, using the relation $QH = Q^2\lambda = 0$ we do the following calculation:

$$\mathrm{e}^{-\tau H} = 1 + \left(\mathrm{e}^{-\tau(Q\lambda)} - 1\right) = 1 + Q\left(\lambda\,\frac{\mathrm{e}^{-\tau(Q\lambda)} - 1}{Q\lambda}\right)\ .$$

Here the term in parentheses stands for

$$\frac{\mathrm{e}^{-\tau(Q\lambda)} - 1}{Q\lambda} := \sum_{n=0}^{\infty} \frac{(-\tau)^{n+1}}{(n+1)!}\,(Q\lambda)^n\ .$$

Inserting this decomposition into the integral we obtain

$$\int D\mu\, \mathrm{e}^{-\tau H} f = \int D\mu\, f + \int D\mu\, f\, Q\left(\lambda\,\frac{\mathrm{e}^{-\tau(Q\lambda)} - 1}{Q\lambda}\right)\ .$$

Since our integrand $f$ is $Q$-invariant by assumption ($Qf = 0$), the second integral can also be written as

$$\int D\mu\, f\, Q\left(\lambda\,\frac{\mathrm{e}^{-\tau(Q\lambda)} - 1}{Q\lambda}\right) = \int D\mu\, Q\left(f\,\lambda\,\frac{\mathrm{e}^{-\tau(Q\lambda)} - 1}{Q\lambda}\right) = 0\ ,$$

which vanishes by Lemma 15. This already proves (C.2).

To complete the proof, we consider the effect of a scale transformation $\phi_\tau^*$: $x \mapsto x/\sqrt{\tau}$, $y \mapsto y/\sqrt{\tau}$, $\xi \mapsto \xi/\sqrt{\tau}$, $\eta \mapsto \eta/\sqrt{\tau}$. Note that $\phi_\tau^* H = H/\tau$ and the Berezin superintegration form $dxdy\,\partial_\xi\partial_\eta = D\mu \circ (1+H)^{1/2}$ is invariant by $\phi_\tau^*$. The statement of the Lemma now results from taking the limit

$$\begin{aligned}\int_{\mathbb{R}^2} D\mu\, f &= \lim_{\tau\to\infty} \int D\mu\, \mathrm{e}^{-\tau H} f = \lim_{\tau\to\infty} \int \phi_\tau^*\left(D\mu\, \mathrm{e}^{-\tau H} f\right)\\ &= \lim_{\tau\to\infty} \int D\mu\,(1+H)^{1/2}(1+H/\tau)^{-1/2}\mathrm{e}^{-H}\phi_\tau^* f = f(o)\ ,\end{aligned}$$

where the last step is done by verifying the normalization integral

$$\int_{\mathbb{R}^2} D\mu\,(1+H)^{1/2}\mathrm{e}^{-H} = (2\pi)^{-1}\int_{\mathbb{R}^2} dxdy\,\partial_\xi\partial_\eta\,\mathrm{e}^{-x^2-y^2-2\xi\eta} = 1\ ,$$

and observing that $\lim_{\tau\to\infty}\phi_\tau^* f$ is the constant function of value $f(o)$. ■

We finally turn to the setting of an arbitrary lattice $\Lambda$. We have a first-order differential operator $Q_j$ for every site $j\in\Lambda$ and we now take the symmetry generator $Q$ to be the sum of all of these:

$$Q = \sum_{j\in\Lambda} Q_j = \sum_{j\in\Lambda}\left(x_j\partial_{\eta_j} - y_j\partial_{\xi_j} + \xi_j\partial_{x_j} + \eta_j\partial_{y_j}\right)\ .$$

By the same argument as before, one sees that $D\mu_\Lambda$ is $Q_j$-invariant for all $j$ and hence $Q$-invariant. There still exists $H = \sum_{j\in\Lambda}(x_j^2+y_j^2+2\xi_j\eta_j)$ and $\lambda = \sum_{j\in\Lambda}(x_j\,\eta_j - y_j\,\xi_j)$ with $Q\lambda = H$. Hence we can still localize the integral $\int D\mu_\Lambda\,F$ for any $Q$-invariant function $F$ by deforming with $\mathrm{e}^{-\tau H}$ and sending $\tau\to\infty$. Thus we arrive at the following result which, though valid for any choice of coordinate system, will be stated in terms of the horospherical coordinates $t_j, s_j, \bar\psi_j, \psi_j$ used in the body of the paper.

**Proposition 2** *For any $Q$-invariant, smooth and integrable function $F$ of the lattice variables $t_j, s_j, \bar\psi_j, \psi_j$ the integral of $F$ localizes at the zero-field configuration $t_j = s_j = \bar\psi_j = \psi_j = 0$ (for all $j\in\Lambda$):*

$$\int_{(\mathbb{R}^2)^{|\Lambda|}} D\mu_\Lambda\,F = F(o)\ .$$

*In particular, for the partition function* (4.5) *we have*

$$Z(\beta,\varepsilon) = \int_{(\mathbb{R}^2)^{|\Lambda|}} D\mu_\Lambda\,\mathrm{e}^{-A_{\beta,\varepsilon}} = 1\ .$$

# References


[1] F.A. Berezin, *Introduction to Superanalysis* (Reidel Publishing Co., Dordrecht, 1987)

[2] D. Coppersmith and P. Diaconis, *Random walk with reinforcement*, unpublished manuscript, 1986

[3] M. Disertori, *Density of states for GUE through supersymmetric approach*, Rev. Math. Phys. **16** (2004) 1191-1225

[4] T. Dupré, *Localization transition in three dimensions: Monte Carlo simulation of a nonlinear sigma model*, Phys. Rev. B **54** (1996) 12763-12774

[5] W. Drunk, D. Fuchs, M.R. Zirnbauer, *Migdal-Kadanoff renormalization of a nonlinear supervector model with hyperbolic symmetry*, Ann. Physik **1** (1992) 134-150

[6] K.B. Efetov, *Supersymmetry in disorder and chaos* (Cambridge University Press, Cambridge, 1997)

[7] K.B. Efetov, *Supersymmetry and theory of disordered metals*, Adv. Phys. **32** (1983) 874

[8] D.H. Friedan, *Nonlinear models in 2 + epsilon dimensions*, Ann. Phys. **163** (1985) 318-419

[9] Y.V. Fyodorov, *Negative moments of characteristic polynomials of random matrices: Ingham-Siegel integral as an alternative to Hubbard-Stratonovich transformation*, Nucl. Phys. B **621** (2002) 643-674

[10] P. Heinzner, A. Huckleberry, M.R. Zirnbauer, *Symmetry classes of disordered fermions*, Commun. Math. Phys. **257** (2005) 725-771

[11] P. Littelmann, H.-J. Sommers, M.R. Zirnbauer, *Superbosonization of invariant random matrix ensembles*, Commun. Math. Phys. **283** (2008) 343-395

[12] A.J. McKane, M. Stone, *Localization as an alternative to Goldstone theorem*, Ann. Phys. **131** (1981) 36-55

[13] F. Merkl, S.W.W. Rolles, *Asymptotic behavior of edge-reinforced random walks*, Ann. Prob. **35** (2007) 115-140

[14] A.D. Mirlin, *Statistics of energy levels and eigenfunctions in disordered systems*, Phys. Rep. **326** (2000) 260-382

[15] M. Niedermaier, E. Seiler, *Structure of the space of ground states in systems with non-amenable symmetries*, Commun. Math. Phys. **270** (2007) 373-443

[16] A. Duncan, M. Niedermaier, E. Seiler, *Vacuum orbit and spontaneous symmetry breaking in hyperbolic sigma-models*, Nucl. Phys. B **720** (2005) 235-288

[17] A. Schwarz, O. Zaboronsky, *Supersymmetry and localization*, Commun. Math. Phys. **183** (1997) 463-476

[18] T. Spencer, M.R. Zirnbauer, *Spontaneous symmetry breaking of a hyperbolic sigma model in three dimensions*, Commun. Math. Phys. **252** (2004) 167-187

[19] L. Schäfer, F. Wegner, *Disordered system with n orbitals per site: Lagrange formulation, hyperbolic symmetry, and Goldstone modes*, Z. Phys. B **38** (1980) 113-126

[20] F. Wegner, *The mobility edge problem: continuous symmetry and a conjecture*, Z. Phys. B **35** (1979) 207-210

[21] M.R. Zirnbauer, *Fourier analysis on a hyperbolic supermanifold with constant curvature*, Commun. Math. Phys. **141** (1991) 503-522