Commun. Math. Phys. 209, 517 – 545 (2000)

Communications in
Mathematical Physics
© Springer-Verlag 2000

Percolation, Path Large Deviations and Weakly Gibbs States

Christian Maes[1,⋆], Frank Redig[1,⋆⋆], Senya Shlosman[2,⋆⋆⋆], Annelies Van Moffaert[1,†]

[1] Inst. voor Theoretische Fysica, K.U. Leuven, Celestijnenlaan 200D, 3001 Leuven, Belgium. E-mail: Christian.Maes@fys.kuleuven.ac.be; Frank.Redig@fys.kuleuven.ac.be; Annelies.VanMoffaert@fys.kuleuven.ac.be
[2] CPT/CNRS, Luminy, Case 907, 13288 Marseille, Cedex, France. E-mail: shlosman@cptsu6.univ-mrs.fr

Received: 14 May 1999 / Accepted: 16 September 1999

Abstract: We present a unified approach to establishing the Gibbsian character of a wide class of non-Gibbsian states, arising in the Renormalisation Group theory. Inside the realm of the Pirogov–Sinai theory for lattice spin systems, we prove that RG transformations applied to low temperature phases give rise to weakly Gibbsian measures. In other words, we show that the Griffiths–Pearce–Israel scenario of RG pathologies is carried by atypical configurations. The renormalized measures are described by an effective interaction, with relative energies well-defined on a full measure set of configurations. In this way we complete the first part of the Dobrushin Restoration Program: to give a Gibbsian description to non-Gibbsian states. A disagreement percolation estimate is used in the proof to bound the decay of quenched correlations through which the interaction potential is constructed. The percolation is controlled via a novel type of pathwise large deviation theory.

1. Introduction and the Main Result

1.1. Problem of Gibbsianity for the restrictions of Gibbs random fields. In this paper we continue the study of the Gibbsian nature of certain random fields, arising naturally in the context of statistical mechanics. As it is known by now, not all reasonable random fields are Gibbs fields. One class of examples can be obtained by applying simple Renormalization Group transformations to some of the most usual lattice Gibbs fields of statistical mechanics. A theorem of van Enter, Fernandez, and Sokal [EFS], extending earlier results of Griffiths and Pearce [GP1, GP2] and Israel [I], states that the restriction

⋆ Onderzoeksleider FWO, Flanders.

⋆⋆ Post-doctoraal Onderzoeker FWO, Flanders.

⋆⋆⋆ The work of S.S. is partially supported by the NSF through the grant DMS-9800860 and by the Russian Basic Research Foundation through the grant 99-01-00284.

† Aspirant FWO, Flanders.

of the $(+)$-phase of the two-dimensional low temperature Ising model to any square sublattice is not a Gibbs field. In [Sch] it is proven that the restriction of the same $(+)$-phase to the one-dimensional sublattice is also not a Gibbs field. So, originated by Dobrushin [D2], some efforts were made to generalize the notion of a Gibbsian field so as to bring these restrictions back into some class of generalized Gibbs fields.

One way of doing this is to compromise on the condition that the interaction energy between a finite volume configuration and the outside world is always defined. More precisely, for states of infinite range dependence, this energy has to be represented by an infinite series. However, in some natural cases the absolute convergence of this series can not be satisfied for all configurations, for whatever choice of the interaction function. Dobrushin's idea was that the convergence condition can be sacrificed, and one should be content with only almost everywhere convergence, according to the corresponding probability measure. This idea was implemented in [DS, MvdV1] for the projection of the $(+)$-phase of the two-dimensional low temperature Ising model to the 1D sublattice, and in [BKL] for the projection to a square sublattice. It is worthwhile to mention that the methods of these papers are quite different. In [BKL] the crucial technique is the one used in the study of the behavior of the Ising model in a random field, [BK]. It seems that such a technique can not be applied to the case of projections to lower dimensions. The methods of [DS] can in principle be applied to other situations, but their implementation requires much additional technical work.

Let us remind the reader of the remark in Sect. 4.2 of [EFS] (used in [LMV]) that the image measure under a renormalization group transformation of a Gibbs measure may be viewed as the restriction of (another) Gibbs measure (obtained as the joint distribution of the original Gibbs measure with its RG image). In that way, the study of restrictions of Gibbs measures in fact incorporates a wide class of renormalization group transformations applied to Gibbs measures. For that reason, in the present paper we will concentrate on the case of the simplest renormalization group transformation, that of the restriction **(**≡projection **)** to a sublattice. The generalization to other examples of renormalization group transformations is straightforward. In fact, our restrictions are more general, since we project Gibbs measures on a quite arbitrary infinite countable subset $\mathbb{M}$ of the lattice. On the other hand, the Gibbs measures we treat are the so called pure phases of models satisfying conditions of the Pirogov-Sinai theory [Sin].(The plus and minus phases of the standard Ising model are the best known examples.) We develop a universal approach to the problem, which is insensitive to the geometry of the subset $\mathbb{M}$. In particular, all the above cited results are included. However, the temperatures for which our technique works depends on how sparse the set $\mathbb{M}$ is, and goes to zero when the sparseness increases. Our strategy is the development of the one used in [MvdV1] for the case of projecting the 2D Ising model onto the 1D sublattice. The idea of [MvdV1] to use percolation techniques has to be supplemented in our present more general situation by a certain large deviation theory. The required large deviation estimates are of a novel type, which is developed in Sect. 6 of this paper (see also Sect. 1.2 of the present Introduction).

We now describe briefly our results. Throughout the paper we fix a countable subset $\mathbb{M}$ of the regular d-dimensional lattice $\mathbb{Z}^d$, containing the origin. The only restriction is that the set $\mathbb{M}$ has to be *k-connected*, $k > 0$. It means that the set

$$\left\{x \in \mathbb{Z}^d : \operatorname{dist}(x, \mathbb{M}) \leq k\right\} \tag{1}$$

is connected. The number k is fixed throughout the paper. Let now the random field $\mathbb{P}$ be an extremal low temperature Gibbs state of a model of statistical mechanics on $\mathbb{Z}^d, d \geq 2$, satisfying all the conditions of the Pirogov–Sinai (PS) theory (see [Sin]).

(The reader can think about the (+)-phase of the two-dimensional Ising model.) The field $\mathbb{P}$ is a probability measure on the set $\Omega = S^{\mathbb{Z}^d}$, S finite, of all spin configurations σ on $\mathbb{Z}^d$. We are interested in the projection ($\equiv$restriction) $\mathbb{P}_{\mathbb{M}}$ of $\mathbb{P}$ onto the subset $\Omega_{\mathbb{M}} \subset \Omega$ of all spin configurations $\sigma_{\mathbb{M}}$ on $\mathbb{M}$. We are looking for a *Gibbsian potential* for $\mathbb{P}_{\mathbb{M}}$, i.e. for a system $\mathcal{U} = (U(T, \sigma_T), T \subset \mathbb{M}, 0 < |T| < \infty)$ of real-valued functions $U(T, \sigma_T)$ of $\sigma_T \in S^T$, such that the usual Gibbs formula for the conditional distributions of $\mathbb{P}_{\mathbb{M}}$ holds. However, a potential, which is absolutely summable, does not exist in general, as we already said above. What it is possible to find, is a system $\mathcal{U}$, which makes $\mathbb{P}_{\mathbb{M}}$ into a *weakly Gibbs random field*. That means that one can find a tail measurable subset $\tilde{\Omega}_{\mathbb{M}} \subset \Omega_{\mathbb{M}}$, such that

$$\mathbb{P}_{\mathbb{M}}\left(\tilde{\Omega}_{\mathbb{M}}\right) = 1, \tag{2}$$

and the *relative energy* series

$$\begin{aligned} &E_V^{\mathcal{U}}(\sigma_V | \bar{\sigma}_{\mathbb{M}\setminus V}) \\ &= \sum_{T \subseteq V, T \neq \emptyset} U(T, \sigma_T) + \sum_{T \subset \mathbb{M}: T\cap V \neq \emptyset, T \cap (\mathbb{M}\setminus V) \neq \emptyset, |T| < \infty} U(T, \sigma_{T\cap V} \cup \bar{\sigma}_{T\cap(\mathbb{M}\setminus V)}) \end{aligned} \tag{3}$$

converges absolutely for all boundary conditions $\bar{\sigma}_{\mathbb{M}} \in \tilde{\Omega}_{\mathbb{M}}$. The properties (2), (3) allow to write the Gibbs specification for $\mathbb{P}_{\mathbb{M}}$-almost all configurations, and hence one can also write down the DLR equations, which in turn are satisfied by our measure $\mathbb{P}_{\mathbb{M}}$. We refer to [MRV1,EMS] for further definitions and for a comparison with the notion of an almost Gibbsian field.

Summarizing, our results in a preliminary form are given by the following

Theorem 1. *The projection $\mathbb{P}_{\mathbb{M}}$ of a Gibbs state $\mathbb{P}$, describing a low temperature pure state of the PS model, to a k-connected subset $\mathbb{M} \subset \mathbb{Z}^d$, is a weakly Gibbs random field. The set of configurations $\tilde{\Omega}_{\mathbb{M}}$, for which the Gibbs specifications can be defined, is given by a constructive procedure.*

(A more detailed statement is contained in Theorem 4 below.)

Actually, the construction of the set $\tilde{\Omega}_{\mathbb{M}}$ is an interesting subject in itself, so we conclude the introduction by mentioning our results concerning it.

1.2. Path large deviations. For the sake of simplicity we describe in the introduction the corresponding results in the simplified setting of the (+)-phase of the low temperature 2D Ising model. We want to discuss properties of the typical configurations, or, rather, typical properties of configurations.

One well-known example of a typical property is the property of "*having the right magnetization*". It means the following. Consider the event

$$A(\varepsilon, V) = \left\{ \sigma \in \Omega : \left| \frac{1}{|V|} \sum_{x \in V} \sigma_x - m^*(\beta) \right| > \varepsilon \right\},$$

where $m^*(\beta)$ is the spontaneous magnetization at inverse temperature β. Then for every $\varepsilon > 0$,

$$\mathbb{P}_V^{\beta,+}(A(\varepsilon, V)) \to 0 \text{ as } V \to \mathbb{Z}^2, \tag{4}$$

where $\mathbb{P}_V^{\beta,+}$ is the Gibbs state in the square box V with $(+)$ boundary conditions. In particular, if we put

$$A(\varepsilon) = \bigcap_N \bigcup_{n \geq N} A(\varepsilon, V_n),$$

where V_n is an n-square, then for every $\varepsilon > 0$,

$$\mathbb{P}^{\beta,+}\left[(A(\varepsilon))^c\right] = 1.$$

Here $\mathbb{P}^{\beta,+}$ stands for the $(+)$-phase.

In the present paper we need properties which are valid for almost all configurations not just in the bulk, but along *every* single selfavoiding path. So let $\mathcal{S}_V$ be the collection of all selfavoiding paths in V, connecting the origin to the boundary of V. Is it then true that the following strengthening of (4) holds: for every $\varepsilon > 0$,

$$\mathbb{P}_{V_n}^{\beta,+}\left[\bigcup_{W \in \mathcal{S}_{V_n}} A(\varepsilon, W)\right] \to 0 \text{ as } V_n \to \mathbb{Z}^d. \tag{5}$$

In other words, is the property of "*having the right magnetization along every path*" a typical one? The answer is clearly negative, since for a $\mathbb{P}^{\beta,+}$-typical configuration σ we easily can find a selfavoiding path γ, which avoids essentially all contours of σ, and so the magnetization of σ along the path γ – that is the quantity $\frac{1}{|\gamma|}\sum_{x\in\gamma}\sigma_x$ – can easily be almost equal to 1. So we introduce a smaller event

$$B(\varepsilon, W) = \left\{\sigma \in \Omega : m^*(\beta) - \frac{1}{|W|}\sum_{x \in W}\sigma_x > \varepsilon\right\}. \tag{6}$$

For example, the event $B(\varepsilon, \gamma)$ happens, if the path γ enters too often inside the contours of the configuration σ. Then the following theorem holds:

Theorem 2. *If β is large enough, then*

$$\mathbb{P}_{V_n}^{\beta,+}\left[\bigcup_{W \in \mathcal{S}_{V_n}} B(\varepsilon, W)\right] \to 0 \text{ as } n \to \infty. \tag{7}$$

In words, the above statement means that the magnetization of *any* typical configuration along *every* selfavoiding path is above $m^*(\beta) - \varepsilon$. If a configuration $\sigma \notin \bigcup_{W \in \mathcal{S}_{V_n}} B(\varepsilon, W)$, then we say that σ has a correct *Path Large Deviation* properties, or simply that σ is a *PLD configuration*.

In contrast with (4), where the convergence is exponential in $|V|$, in (7) we only have a stretched exponential decay. This is the content of Theorem 9 below, which in particular proves the claim of Theorem 2 above.

In the next section we introduce the notations. In Sect. 3 we reduce the proof of Theorem 1 to the question of correlation decay in a random (quenched) environment. In Sect. 4 the correlation decay question is reduced to a question about percolation in a random environment. In Sect. 5 the percolation problem is solved under the hypothesis that the random environment has a property of the type described in Theorem 2. Finally, in Sect. 6 the generalization of Theorem 2 is proven, which justifies the use of the hypothesis above.

2. Notation

We fix an arbitrary k-connected (see (1)) infinite subset $\mathbb{M}$ of the regular d-dimensional lattice $\mathbb{Z}^d$, containing the origin. The most interesting cases concern $\mathbb{M} = l\,\mathbb{Z}^r, r = 1, \dots, d; l = 1, ..., 2k-1$, where we keep the invariance under (a subgroup of) translations. In the following, $|A|$ is the cardinality of the set A, while A^c denotes the complement of A in $\mathbb{Z}^d$. For $W \subset \mathbb{M}$ we sometimes denote by W^c the complement $\mathbb{M} \setminus W$. General elements (sites) of $\mathbb{Z}^d$ are written as x, y, z, but we write i, j when referring to sites (elements) in $\mathbb{M}$. The distance between $x = (x^1, \dots, x^d)$ and $y = (y^1, \dots, y^d)$ is

$$|x - y| \equiv \sum_{\alpha=1}^{d} |x^\alpha - y^\alpha|. \tag{8}$$

The distance between two sets A and B is $d(A, B) \equiv \min_{x\in A, y\in B} |x - y|$. The diameter of a set A is diam $A \equiv \max_{x,y\in A} |x-y|$. When dealing with a singleton $A = \{x\}$, we often write $A = x$. Λ_n is the cube $\{x \in \mathbb{Z}^d : |x| \le n\}, n = 1, 2, \dots$. Its intersection with $\mathbb{M}$ is $\Lambda_n \cap \mathbb{M} \equiv V_n$. The boundary of a set Λ is $\partial\Lambda \equiv \{x \in \Lambda^c : \exists y \in \Lambda, |x-y| = 1\}$ and must be distinguished from its internal boundary $\partial_i \Lambda \equiv \{x \in \Lambda : \exists y \in \partial\Lambda, |x - y| = 1\}$. We find it also useful to regard $\mathbb{Z}^d$ as a graph with its sites as vertices and its bonds (nearest-neighbor connections) as edges. x and y are adjacent (nearest-neighbors) if $|x - y| = 1$. A (finite) path (of length n) from x to y is a sequence of consecutive and mutually distinct nearest-neighbors $(x_0 = x, x_1, \dots, x_{n-1} = y)$. Infinite paths are the natural extensions of this. A path from A to B is any path starting in a site $x \in A$ and ending in a site $y \in B$ (its length is at least $d(A, B)$).

We will be using the lexicographic order "$\le$" on $\mathbb{Z}^d$ to say that $x < y$ if $x^1 < y^1$, or $x^1 = y^1$ and $x^2 < y^2$, or ... $x^1 = y^1, x^2 = y^2, \dots, x^{d-1} = y^{d-1}$ and $x^d < y^d$. Dual to paths are surfaces. They are sometimes referred to as $\star$-circuits in two dimensions. A surface around A is any collection of next-nearest-neighbor connected sites in A^c so that by removing them from the lattice, no infinite path can exist starting in A.

We consider lattice spin systems on $\mathbb{Z}^d$. A general spin configuration on $\mathbb{Z}^d$ is denoted by σ or η. They are elements of the configuration space $\Omega \equiv S^{\mathbb{Z}^d}$, where S is the finite set ($|S| \equiv q \ge 2$) of spin-values $a, b, c, \dots$ at a single site. Ising-spins have $S = \{+1, -1\}$. The value of the spin at a site x in the configuration σ is $\sigma(x) \in S$. We will frequently use some reference configuration, denoted by 1 with $1(x) = +1$ everywhere. The restriction of a $\sigma \in \Omega$ to a set A is $\sigma_A \in S^A$; $\sigma_A \eta_{A^c} \equiv \sigma_A \cup \eta_{A^c}$ equals σ on A (i.e., $\sigma_A \eta_{A^c}(x) = \sigma(x)$ for all $x \in A$) and equals η on A^c. We write σ^A for the configuration which equals σ on A and is equal to $+1$ outside A. The restriction of Ω to $\mathbb{M}$ is $\Omega_{\mathbb{M}} \equiv S^{\mathbb{M}}$ and $\Omega_n \equiv \Omega_{\Lambda_n}$. We often consider $\Omega_{\mathbb{M}}$ as a subset of $\Omega_{\mathbb{Z}^d}$ via natural embedding $\sigma \in \Omega_{\mathbb{Z}^d} \to \sigma_{\mathbb{M}} \in \Omega_{\mathbb{M}}$. Therefore the same symbols σ, η, ξ will sometimes appear for configurations in Ω and in $\Omega_{\mathbb{M}}$. All notation is inherited, e.g. ξ^V equals ξ on V and is $+1$ on V^c.

A function f on Ω is local if its dependence set D_f, i.e. the minimal set A such that $f(\sigma) = f(\eta)$ whenever $\sigma_A = \eta_A$, is finite. Continuous functions are uniform limits of local functions with the sup-norm $||f|| \equiv \sup_\sigma |f(\sigma)|$.

The sigma-algebra generated by the evaluations $x \in A \to \sigma(x)$ is denoted by $\mathcal{F}_A$. When $A = \mathbb{Z}^d$, respectively $A = \mathbb{M}$, we simply set $\mathcal{F} = \mathcal{F}_{\mathbb{Z}^d}$, respectively $\mathcal{F}' = \mathcal{F}_{\mathbb{M}}$. The tailfield sigma-algebras are denoted by $\mathcal{F}^\infty = \cap_n \mathcal{F}_{\Lambda_n^c}$ and $\mathcal{F}'^\infty = \mathcal{F}^\infty \cap \mathcal{F}'$ respectively.

In what follows we will be considering probability measures μ on $(\Omega, \mathcal{F})$. Their corresponding random field is denoted by $X \equiv (X(x), x \in \mathbb{Z}^d)$. Expectations are abbreviated as $\int f(\sigma)d\mu(\sigma) = \mu(f)$ and a covariance is written as $\mu(f; g) = \mu(fg) - \mu(f)\mu(g)$. The probability of an event $E \in \mathcal{F}$ is $\mu[E] = \mu[X \in E]$ or also $\mu(I[X \in E])$, where we introduced the indicator function I. The same notation is used for probability measures ν on $(\Omega, \mathcal{F}')$. Such a measure appears as the restriction of μ to $\mathcal{F}'$ (or, to $\mathbb{M}$). We denote by $Y \equiv (Y(i), i \in \mathbb{M})$ the restriction of the random field X to $\mathbb{M}$.

The basic model system we will be dealing with here is defined via a nearest-neighbor interaction

$$\mathcal{H}_{\Lambda_n}(\sigma) = - \sum_{x \in \Lambda_n, y \in \mathbb{Z}^d, |x-y|=1} J(\sigma^{\Lambda_n}(x), \sigma^{\Lambda_n}(y)). \tag{9}$$

The interaction term $J(\cdot, \cdot)$ is a real-valued symmetric function on $S \times S$ (possibly containing a self-energy). We have taken care of putting $+1$ boundary conditions outside the cube Λ_n. The nearest-neighbor aspect ensures that the corresponding Gibbs fields will be Markov random fields but that will not be essential. The partition function is

$$Z_n \equiv \sum_{\sigma \in S^{\Lambda_n}} e^{-\beta \mathcal{H}_{\Lambda_n}(\sigma)}. \tag{10}$$

We suppose that $\mathcal{H}$ satisfies the conditions of the PS theory, and that the configuration 1 with $1(x) = +1$ everywhere is a ground state and gives rise to pure states at *low* temperatures in the usual sense of PS theory. More specifically, we assume that for all $a, b \in S \setminus \{+1\}$,

$$J(+1, +1) = 0, \; J(a, +1) < -1, \; J(a, b) \le 0. \tag{11}$$

We assume further, that all ground state configurations of $\mathcal{H}$ are translation invariant. That implies in particular that $J(a, b) < 0$ for $a \neq b$. We then assume that in such a situation $J(a, b) < -1$ as well.

3. Correlation Decay ⇒ Weak Gibbsianity

We start with the definition of some finite subsets of $\mathbb{M}$. For every $i \in \mathbb{M}$ we put

$$L_{i,m} \equiv \{j \in \mathbb{M} : j \le i, |j - i| \le m\}, m = 0, 1 \ldots . \tag{12}$$

Clearly, $L_{i,m-1} \subset L_{i,m} \subset \mathbb{M}$ and

$$v(i, m) \equiv |L_{i,m} \setminus L_{i,m-1}| \le \frac{2^{d-1}}{(d-1)!} m^{d-1} \le 2m^{d-1}, \; m \ge 1. \tag{13}$$

We can now, following the lexicographic order, add sites one by one to $L_{i,m-1}$ to end up with the set $L_{i,m}$. So let us write $L_{i,m} \setminus L_{i,m-1} = \{j_1, j_2, \ldots , j_v\}$ with $j_1 < j_2 < \ldots < j_v$ and their number $v = v(i, m)$ possibly depending on m and i but not exceeding $2m^{d-1}$. It can happen that $L_{i,m-1} = L_{i,m}$ in which case $v(i, m) = 0$. Define the sets

$$Q_{i,m,r} \equiv L_{i,m-1} \cup \{j_1, j_2, \ldots , j_r\} \tag{14}$$

with $r = 1, 2, \ldots , v$. We have $Q_{i,m,v} = L_{i,m}$ and we put $Q_{i,m,0} \equiv L_{i,m-1}$.

Let μ_n be the Gibbs state in the box Λ_n, defined by our Hamiltonian $\mathcal{H}$ and the configuration 1 as boundary condition. According to our hypothesis, the sequence of probability measures $\mu_n \equiv \mu_{\Lambda_n}$ on $(\Omega, \mathcal{F})$ weakly converges to a measure μ, which is a pure state of our model. Let ν_n denote the probability measures obtained from μ_n by restricting them to $\mathbb{M}$. Then, the limit $\lim_n \nu_n = \nu$ exists and equals μ restricted to $\mathbb{M}$.

For every configuration $\xi \in \Omega_{\mathbb{M}}$ the measure μ_n^ξ on $(\Omega_n, \mathcal{F})$ is defined in the following way:
i) one takes the conditional Gibbs distribution $\bar{\mu}_{n,\bar{\sigma}_{n,\xi}}$ in Λ_n, defined by our Hamiltonian $\mathcal{H}$ and the following boundary conditions $\bar{\sigma}_{n,\xi}$:

$$\bar{\sigma}_{n,\xi}(x) = \begin{cases} \xi(x) & \text{for } x \in \partial\Lambda_n \cap \mathbb{M}, \\ 1 & \text{for } x \in \Lambda_n^c \setminus (\partial\Lambda_n \cap \mathbb{M}), \end{cases}$$

(in our notation, $\bar{\sigma}_{n,\xi} = \xi^{\partial\Lambda_n \cap \mathbb{M}}$),
ii) one defines μ_n^ξ by the conditioning:

$$\mu_n^\xi[X = \sigma \text{ on } A] = \bar{\mu}_{n,\bar{\sigma}_{n,\xi}}[X = \sigma \text{ on } A | X = \xi \text{ on } V_n], \text{ where } A \subset \Lambda_n. \tag{15}$$

We denote by $\Omega_n^\xi \equiv \Omega_{n,\mathbb{M}}^\xi \subset \Omega_{\mathbb{Z}^d}$ the support of the measure μ_n^ξ; by definition, these configurations coincide with $\bar{\sigma}_{n,\xi}$ outside Λ_n, and with ξ on V_n. Also, we define μ^ξ by $\mu^\cdot[A] = \mu\left(1_A|\mathcal{F}'\right)$; by what follows one can show that $\mu^\xi = \lim_n \mu_n^\xi$ on a set of ξ's of $\mathbb{P}_{\mathbb{M}}$-measure one.

Let us define the local observables $\phi_i = \phi_i^\xi(\sigma)$, $i \in \mathbb{M}$ so that for all $i \in V_n$,

$$\mu_n[X = \xi^{Q\setminus i} \text{ on } V_n] = \mu_n[X = \xi^Q \text{ on } V_n]\mu_n^{\xi^Q}(\phi_i). \tag{16}$$

Of course, these functions ϕ_i can be written down, but we do not need explicit expressions for them. Also, for all $i, j \in V_n$ we have

$$\mu_n[X = \xi^{Q\setminus\{i,j\}} \text{ on } V_n] = \mu_n[X = \xi^Q \text{ on } V_n]\mu_n^{\xi^Q}(\Phi_{ij}). \tag{17}$$

Since the interaction is nearest neighbor,

$$\Phi_{ij} = \phi_i\phi_j, \tag{18}$$

provided $|i - j| > 4$.

Next we formulate the Correlation Decay property, which, if valid, implies Weak Gibbsianity.

Definition 3. *We say that the Quenched Correlation Decay (**QCD**) property holds, if there are constants $C < \infty$, $\lambda > 0$, a tail-set $K \in \mathcal{F}'$ with*

$$\mu_n[K] = \mu[K] = 1 \tag{19}$$

and a function $\ell(i,\xi)$, defined for $\xi \in K$, $i \in \mathbb{M}$, such that for every $j \in \mathbb{M}$ the set

$$\text{Bad}(j,\xi) \equiv \{i \in \mathbb{M} : \ell(i,\xi) \geq |i - j|\} \tag{20}$$

is finite, and for all finite $Q \subset \mathbb{M}$, all n and all j with $|i - j| > \min\{\ell(i,\xi), \ell(j,\xi)\}$,

$$|\mu_n^{\xi^Q}(\phi_i;\phi_j)| \leq Ce^{-\lambda|i-j|}. \tag{21}$$

We suppose additionally that for every $\xi \in K$ and every $\bar{Q} \subset \mathbb{M}$ (finite or infinite) $\xi^{\bar{Q}} \in K$ as well, and that for every $j \in \mathbb{M}$ $\text{Bad}\left(j, \xi^{\bar{Q}}\right) \subset \text{Bad}(j,\xi)$.

Theorem 4. *Assume that the condition QCD holds. Then the restricted state, ν, is weakly Gibbsian with interaction potential $\mathcal{U} = \{U(T, \cdot), T \subset \mathbb{M}\}$ vanishing except possibly for the sets $T = L_{i,m}$ with $v(i, m) \geq 1$, where $i \in \mathbb{M}, m \in \mathbb{N}$, and is then given by*

$$U(L_{i,m}, \xi) = -\sum_{r=1}^{v(i,m)} \ln \frac{\mu^{\xi^{Q_{i,m,r}}}(\phi_i \phi_{j_r})}{\mu^{\xi^{Q_{i,m,r}}}(\phi_i)\mu^{\xi^{Q_{i,m,r}}}(\phi_{j_r})} \text{ for } m > 4, \tag{22}$$

$$U(L_{i,m}, \xi) = -\sum_{r=1}^{v(i,m)} \ln \frac{\mu^{\xi^{Q_{i,m,r}}}(\Phi_{ij_r})}{\mu^{\xi^{Q_{i,m,r}}}(\phi_i)\mu^{\xi^{Q_{i,m,r}}}(\phi_{j_r})} \text{ for } 1 \leq m \leq 4, \tag{23}$$

while for $m = 0$

$$U(L_{i,0}, \xi) = \ln \mu^{\xi^i}(\phi_i). \tag{24}$$

This potential is absolutely summable on the tail-set K. Moreover, it satisfies the following bound: there exist constants $C_1, C_2 < \infty, \lambda > 0$ such that for all $m \geq 0, \xi \in K, i \in \mathbb{M}$,

$$|U(L_{i,m}, \xi)| \leq C_1 I[m \leq \ell(i, \xi)] m^{d-1} + C_2 I[m > \ell(i, \xi)] m^{d-1} \exp[-\lambda m], \tag{25}$$

with the $\ell(i, \xi)$ and the set K as in QCD conditions (19–21) above.

Proof. Consider the probability to find the configuration ξ in V_n. In our notation it is

$$\nu_n[Y = \xi \text{ on } V_n] = \mu_n[X = \xi \text{ on } V_n], \tag{26}$$

and we will abbreviate it as $\mu_n[\xi]$. Order the sites in V_n lexicographically as $i_1 < i_2 < \ldots < i_{|V_n|}$ to write

$$\frac{\mu_n[\xi]}{\mu_n[1]} = \prod_{s=1}^{|V_n|} \frac{\mu_n[\xi^{\{i_1,\ldots,i_s\}}]}{\mu_n[\xi^{\{i_1,\ldots,i_{s-1}\}}]}. \tag{27}$$

For every $i \in V_n$ we define $m(i, n) \equiv \max_{j \in V_n : j \leq i} |i - j|$. Then we can rewrite every factor in (27) as

$$\frac{\mu_n[\xi^{\{i_1,\ldots,i_s\}}]}{\mu_n[\xi^{\{i_1,\ldots,i_{s-1}\}}]} = \prod_{m=0}^{m(i_s,n)} \frac{\mu_n[\xi^{L_{i_s,m}}]}{\mu_n[\xi^{L_{i_s,m-1}}]} \frac{\mu_n[\xi^{L_{i_s,m-1} \setminus i_s}]}{\mu_n[\xi^{L_{i_s,m} \setminus i_s}]}.$$

So if we define the family $\mathcal{U}^n = \{U^n(\cdot, \cdot)\}$ by

$$U^n(L_{i,m}, \xi) \equiv -\ln \frac{\mu_n[\xi^{L_{i,m}}]}{\mu_n[\xi^{L_{i,m-1}}]} \frac{\mu_n[\xi^{L_{i,m-1} \setminus i}]}{\mu_n[\xi^{L_{i,m} \setminus i}]}, \tag{28}$$

then we have

$$\mu_n[\xi] = \mu_n[1] \exp\left\{ -\sum_{i \in V_n} \sum_{m=0}^{m(i,n)} U^n(L_{i,m}, \xi^{V_n}) \right\}. \tag{29}$$

Note that $\xi^{\emptyset} \equiv 1$, and so

$$U^n(L_{i,0}, \xi) \equiv -\ln \frac{\mu_n[\xi^i]}{\mu_n[1]} = \ln \mu_n^{\xi^i}(\phi_i). \tag{30}$$

Observe also that $m(i, n) = 0$ when $i = i_1$ is the "first" site in V_n, and that

$$U^n(L_{i,m}, \xi) = 0, \text{ provided } \begin{array}{l} \xi(i) = +1, \\ \text{or} \\ v(i,m) = 0, \\ \text{or} \\ \xi(j) = +1 \text{ for all } j \in L_{i,m} \setminus L_{i,m-1}. \end{array} \tag{31}$$

We can further telescope (28) as

$$U^n(L_{i,m}, \xi) = -\sum_{r=1}^{v(i,m)} \ln \frac{\mu_n[\xi^{Q_{i,m,r}}]\mu_n[\xi^{Q_{i,m,r-1}\setminus i}]}{\mu_n[\xi^{Q_{i,m,r-1}}]\mu_n[\xi^{Q_{i,m,r}\setminus i}]}, \tag{32}$$

provided $m > 0$ and $v(i, m) \geq 1$. If $m > 4$, we can use (16-18) to rewrite (32) as

$$U^n(L_{i,m}, \xi) = -\sum_{r=1}^{v(i,m)} \ln \frac{\mu_n^{\xi^{Q_{i,m,r}}}(\phi_i \phi_{j_r})}{\mu_n^{\xi^{Q_{i,m,r}}}(\phi_i)\mu_n^{\xi^{Q_{i,m,r}}}(\phi_{j_r})}, \tag{33}$$

while for $1 \leq m \leq 4$ we can not use the factoring (18), so we keep the initial local observables Φ_{ij_r} Here we remind the reader that the conditioning $\mu_n^{\xi^Q}$ of the measure μ_n by the configuration ξ^Q means the use of the condition

$$X = \xi \text{ on } Q, \ X = 1 \text{ on } \mathbb{M}\setminus Q, \ |Q| < \infty.$$

So the cluster expansion easily provides us with the existence of the limits $\lim_{n\to\infty} \mu_n^{\xi^{Q_{i,m,r}}}(f)$ for every local observable f. Taking the limits $n \to \infty$ in (33), (30), we arrive to the formulas (22)–(24). The estimate (25) follows from relation (21). The almost sure convergence of the relative energy series (3) follows from the finiteness of the sets Bad (j, ξ) of "bad" points (20), when $\xi \in K$.

What remains is to show that indeed the random field ν is a Gibbs field with the potential $\mathcal{U} = \{U(T, \cdot), T \subset \mathbb{M}\}$, i.e. to prove that the corresponding DLR equations hold. The rest of the present section contains the proof of validity of DLR equations. Since this proof is not used in the rest of the paper, the reader might want to go directly to the next section.

The following proof is an adaptation of the similar statement from [DS], Sect. 8. We begin by introducing for every finite $V \subset \mathbb{M}$ the Gibbs specification

$$p_V^{\mathcal{U}}(\xi_V | \bar{\xi}_{\mathbb{M}\setminus V}) = \frac{\exp\{-E_V^{\mathcal{U}}(\xi_V | \bar{\xi}_{\mathbb{M}\setminus V})\}}{Z_V^{\mathcal{U}}(\bar{\xi}_{\mathbb{M}\setminus V})}, \tag{34}$$

where the partition function

$$Z_V^{\mathcal{U}}(\bar{\xi}_{\mathbb{M}\setminus V}) = \sum_{\xi_V \in X^V} \exp\{-E_V^{\mathcal{U}}(\xi_V | \bar{\xi}_{\mathbb{M}\setminus V})\}. \tag{35}$$

Here the function $E_V^{\mathcal{U}}(\xi_V|\bar{\xi}_{\mathbb{M}\setminus V})$ is defined by the series (3), if the latter converges. Let us show that for $\bar{\xi} = \xi_V \cup \bar{\xi}_{\mathbb{M}\setminus V} \in K$ this convergence follows from that part of Theorem 4 that is already proven.

Indeed, let the set $L_{i,m} \subset \mathbb{M}$ be "bad", in the sense that it intersects the box V, and $\ell(i,\bar{\xi}) \geq m$. In other words, the contribution of $U(L_{i,m}, \xi_V \cup \bar{\xi}_{\mathbb{M}\setminus V})$ to $E_V^{\mathcal{U}}(\xi_V|\bar{\xi}_{\mathbb{M}\setminus V})$ might be big, according to (25). Then necessarily the site i belongs to the set $\bigcup_{j\in V} \mathrm{Bad}\left(j,\bar{\xi}\right)$, which is finite. Note that the number of different sets $L_{i,m}$ with $m \leq \ell(i,\bar{\xi})$ is less than $\left(2\ell(i,\bar{\xi})\right)^d$, so the total number of "bad" $L_{i,m}$-s is bounded by

$$\sum_{i\in\bigcup_{j\in V}\mathrm{Bad}(j,\bar{\xi})} \left(2\ell(i,\bar{\xi})\right)^d ,$$

which is also finite.

The above argument implies that the functions $p_V^{\mathcal{U}}(\xi_V|\bar{\xi}_{\mathbb{M}\setminus V})$ are defined ν-a.s., which makes well-defined the rhs of the DLR equation (36), which follows:

$$\int \phi_W(\xi_W)\nu(d\xi) = \int \sum_{\xi_V\in S^V} \left(\phi_W(\xi_W) p_V^{\mathcal{U}}(\xi_V|\bar{\xi}_{\mathbb{M}\setminus V})\right) \nu_{\mathbb{M}\setminus V}(d\bar{\xi}_{\mathbb{M}\setminus V}). \tag{36}$$

Here $V \subseteq W$ are arbitrary finite subsets of $\mathbb{M}$, $\xi_W = \left(\xi_V \cup \bar{\xi}_{\mathbb{M}\setminus V}\right)|_W$ is (with some abuse of notation) the restriction, the local observable $\phi_W(\xi_W) \equiv \phi_W(\xi_V \cup \xi_{W\setminus V})$ is a function on S^W, while $\nu_{\mathbb{M}\setminus V}$ is the restriction of the measure ν to the σ-algebra $\mathcal{F}'_{\mathbb{M}\setminus V}$. Equation (36) should hold for any choice of $V \subseteq W$.

To prove (36), let us introduce the subsets $X_N \equiv X_N(W) \subset K$ as

$$X_N = \left\{\bar{\xi} \in K : \sum_{j\in W} \left|\mathrm{Bad}\left(j,\bar{\xi}\right)\right| < N\right\},$$

and rewrite the rhs integral of (36) as a sum:

$$\int (\cdot) = \int_{X_N} (\cdot) + \int_{(X_N)^c} (\cdot) . \tag{37}$$

Note that the last integral goes to zero as $N \to \infty$, since the function $\sum_{j\in W} \left|\mathrm{Bad}\left(j,\bar{\xi}\right)\right|$ is finite on K and because the integrand is bounded. We further introduce the subsets $X_{N,R}$ of X_N:

$$X_{N,R} = \left\{\bar{\xi} \in X_N : \forall\, i \in \bigcup_{j\in W} \mathrm{Bad}\left(j,\bar{\xi}\right) \text{ we have dist}\,(i, W) \leq R\right\}.$$

We have that

$$\int_{X_N} (\cdot) = \int_{X_{N,R}} (\cdot) + \int_{X_N\setminus X_{N,R}} (\cdot) . \tag{38}$$

Again, the last integral goes to zero as $R \to \infty$.

The estimate (25) implies the continuity of the function $p_V^{\mathcal{U}}(\xi_V|\bar{\xi}_{\mathbb{M}\setminus V})$ on the subspaces $X_N, X_{N,R}$. Note also that Definition 3 of the QCD property implies that the following implications hold: $\left(\bar{\xi} \in X_N\right) \Rightarrow \left(\bar{\xi}^Q \in X_N\right)$, and $\left(\bar{\xi} \in X_{N,R}\right) \Rightarrow \left(\bar{\xi}^Q \in X_{N,R}\right)$. Together these properties imply that for every $\varepsilon > 0$ there exists a distance $\rho(\varepsilon, N, R)$ big enough, such that for every set Q_ρ containing the set

$$Q(W; \rho(\varepsilon, N, R)) = \{i \in \mathbb{M} : \text{dist}(i, W) \le \rho(\varepsilon, N, R)\}$$

we have:

$$\left| \int_{X_{N,R}} \sum_{\xi_V \in S^V} \left(\phi_W(\xi_W) p_V^{\mathcal{U}}(\xi_V|\bar{\xi}_{\mathbb{M}\setminus V})\right) \nu_{\mathbb{M}\setminus V}(d\bar{\xi}_{\mathbb{M}\setminus V}) - \right. \tag{39}$$
$$\left. - \int_{X_{N,R}} \sum_{\xi_V \in S^V} \left(\phi_W(\xi_W) p_V^{\mathcal{U}}(\xi_V|\bar{\xi}^{Q_\rho}_{\mathbb{M}\setminus V})\right) \nu_{\mathbb{M}\setminus V}(d\bar{\xi}_{\mathbb{M}\setminus V}) \right| < \varepsilon.$$

Repeating the arguments (37), (38) for the last integral, in the reversed order, we can replace it by the integral over the whole space,

$$\int \sum_{\xi_V \in S^V} \left(\phi_W(\xi_W) p_V^{\mathcal{U}}(\xi_V|\bar{\xi}^{Q_\rho}_{\mathbb{M}\setminus V})\right) \nu_{\mathbb{M}\setminus V}(d\bar{\xi}_{\mathbb{M}\setminus V}), \tag{40}$$

again with arbitrary precision. But because of (31) the integrand in the last integral is a local function! Therefore we can approximate it arbitrarily close by the integral with respect to the finite volume measure ν_n,

$$\int \sum_{\xi_V \in S^V} \left(\phi_W(\xi_W) p_V^{\mathcal{U}}(\xi_V|\bar{\xi}^{Q_\rho}_{\Lambda_n\setminus V})\right) (\nu_n)_{\Lambda_n\setminus V}(d\bar{\xi}_{\Lambda_n\setminus V}), \tag{41}$$

provided $n > n\left(Q_\rho\right)$ is large enough.

Note that for any local event A we have the exponential convergence $\mu_n(A) \to \mu(A)$ as $n \to \infty$ (though not uniform in A, of course). Therefore the two functions, $p_V^{\mathcal{U}}(\xi_V|\bar{\xi}^{Q_\rho}_{\Lambda_n\setminus V})$ and $p_V^{\mathcal{U}^n}(\xi_V|\bar{\xi}^{Q_\rho}_{\Lambda_n\setminus V})$ can be made arbitrarily close, uniformly in $\xi_V \cup \bar{\xi}^{Q_\rho}_{\Lambda_n\setminus V}$, provided only that $n\,(= n(\rho))$ is large enough. The last step would be to use again the approximations (37-39) to replace the integral

$$\int \sum_{\xi_V \in S^V} \left(\phi_W(\xi_W) p_V^{\mathcal{U}^n}(\xi_V|\bar{\xi}^{Q_\rho}_{\Lambda_n\setminus V})\right) (\nu_n)_{\Lambda_n\setminus V}(d\bar{\xi}_{\Lambda_n\setminus V})$$

by

$$\int \sum_{\xi_V \in S^V} \left(\phi_W(\xi_W) p_V^{\mathcal{U}^n}(\xi_V|\bar{\xi}_{\Lambda_n\setminus V})\right) (\nu_n)_{\Lambda_n\setminus V}(d\bar{\xi}_{\Lambda_n\setminus V}),$$

and then to observe that because of the validity of the finite volume DLR equation (which is an identity)

$$\int \sum_{\xi_V \in S^V} \left(\phi_W(\xi_W) p_V^{\mathcal{U}^n}(\xi_V | \bar{\xi}^{Q_\rho}_{\Lambda_n \setminus V})\right) (\nu_n)_{\Lambda_n \setminus V} (d\bar{\xi}_{\Lambda_n \setminus V})$$
$$= \int \phi_W(\xi_W) \nu_n(d\bar{\xi}_{\Lambda_n}) \to \int \phi_W(\xi_W) \nu(d\bar{\xi}).$$

This use is possible since the relations (37-39) hold in fact for the integrals

$$\int \sum_{\xi_V \in S^V} \left(\phi_W(\xi_W) p_V^{\mathcal{U}^n}(\xi_V | \bar{\xi}_{\mathbb{M} \setminus V})\right) (\nu_n)_{\mathbb{M} \setminus V} (d\bar{\xi}_{\mathbb{M} \setminus V})$$

as well, uniformly in $n \to \infty$, due to our Definition 3 of QCD. □

Remark. The construction of the potential of the basic lemma above goes back to the paper of Kozlov, [Koz]. This telescoping potential was already used in [MvdV1, MRV2] and it has the nice property that, when the set $\mathbb{M}$ allows it, U is explicitly translation-invariant.

4. Percolation ⇒ Correlation Decay

We continue with the same setup of the previous section. We must estimate the covariances in the left hand side of (21). Using an idea of [B, BM, BS] as applied in [MvdV1, MRV2], we can reduce this to a percolation question. For $m > 4$, the two local observables $f = \phi_i$ and $g = \phi_{j_r}$ have disjoint dependence sets $D_f = F$, $D_g = G \subset \Lambda_n$. We will take n large enough, so that $Q_{i,m,r} \subset \Lambda_n$. To save on notation we fix ξ and $Q_{i,m,r}$ and we write $\mu_n^{\xi^{Q_{i,m,r}}} = \rho_n$. We must estimate $\rho_n(f; g)$.

To proceed we consider the product space of configurations $S^{\Lambda_n} \times S^{\Lambda_n} = (S \times S)^{\Lambda_n}$ on which we put the product coupling $\rho_n \times \rho_n$. So, we just consider two independent copies of the original system. Now, for any two disjoint finite sets B and C, we introduce the event $E(B, C)$ that there is a path from B to C in Λ_n, such that for every site x of this path in $\Lambda_n \setminus Q_{i,m,r}$ we have $(X(x), X'(x)) \neq (+1, +1)$.

Lemma 5. *For the Markov random field* ρ_n *we have*

$$|\rho_n(f; g)| \leq 2\, ||f||\, ||g||\, (\rho_n \times \rho_n)\, [E(F, G)]. \tag{42}$$

Proof. First of all, we can write

$$\begin{aligned} &|\rho_n(fg) - \rho_n(f)\rho_n(g)| \\ &= |\rho_n\, (g(X)[\rho_n(f|X_G) - \rho_n(f(X))])\,| \\ &\leq ||g||\, \rho_n\, (|\rho_n(f|X_G) - \rho_n(f)|) \\ &= ||g||\, \rho_n\, \left(dX \left|\left(\rho_n\, \left(dX'|X_G\right) \times \rho_n\, (\cdot)\right) (f(X) \times 1 - 1 \times f(X'))\right|\right). \end{aligned} \tag{43}$$

To save on notation, let $\Lambda \equiv \Lambda_n \setminus Q_{i,m,r} \setminus G$. Now imagine a fixed configuration σ on $\partial\Lambda$ with $\sigma = +1$ on $\partial\Lambda_n$, $\sigma = \xi$ on $Q_{i,m,r}$ and σ some fixed configuration on G; we must study

$$(\rho_n(\cdot|X = \sigma \text{ on } \partial\Lambda) \times \rho_n\, (\cdot))\, (f \times 1 - 1 \times f).$$

If for every path in Λ_n from F to G there is a point $x \in \Lambda_n \setminus Q_{i,m,r}$ on it where $(X(x), X'(x)) = (+1, +1)$, then there exists a surface (or $\star-$ circuit) around F separating it inside Λ_n from G and on which $X \equiv X'$. In general, this surface has a part in $\Lambda \setminus F \cup \partial\Lambda_n$ on which $(X, X') \equiv (+1, +1)$ and a part in $Q_{i,m,r}$ on which the configuration ξ lives and thus there $(X, X') \equiv (\xi, \xi)$. Hence there exists a maximal surface among these, Δ, inside Λ (maximal in the inclusion sense). Now since $X = X'$ on Δ and by the Markov property,

$$\begin{aligned} &|(\rho_n(\cdot|X = \sigma \text{ on } \partial\Lambda) \times \rho_n(\cdot))(f \times 1 - 1 \times f)| \\ &\le 2||f||\,(\rho_n(\cdot|X = \sigma \text{ on } \partial\Lambda) \times \rho_n(\cdot))\,[E(F, G)]. \end{aligned} \tag{44}$$

Continuing with (43), (44) now yields that

$$|\rho_n(f; g)| \le 2\,||f||\,||g||\,(\rho_n \times \rho_n)\,[E(F, G)], \tag{45}$$

which we wanted to prove. □

Remark. Clearly, from the proof above, an analogous estimate to (42) holds in the case that μ_n is not (strictly) Markovian but becomes Markovian when to each site of the lattice a finite number of edges is attached linking that site with more than just its nearest-neighbors.

5. PLD ⇒ Contour Estimates ⇒ Exponentially Weak Percolation

Let F be a finite subset of $\mathbb{Z}^d$, as in the previous section. Denote by C_F the random set containing all sites $x \in \mathbb{Z}^d \setminus \mathbb{M}$ which are nearest-neighbor connected to ∂F via sites $y \in \mathbb{Z}^d \setminus \mathbb{M}$ for which $(X(y), X'(y)) \neq (+1, +1)$. This is the cluster of F. Our task is now to find a (large) set K of configurations ξ and corresponding lengths $\ell(i, \xi)$ (see(19)), for which

$$\mu_n^{\xi^{Q_{i,m,r}}} \times \mu_n^{\xi^{Q_{i,m,r}}}\,[\text{diam}\,(C_F) > m] \le e^{-\lambda m}, \tag{46}$$

whenever $m > \ell(i, \xi)$. Given the bound (42), this would take care of the assumption (21). This of course is reminiscent of the stochastic-geometric structure of the low-temperature phases in the realm of the Pirogov-Sinai theory.

Since the coupling we are considering is just a product coupling $\mu_n^{\xi^{Q_{i,m,r}}} \times \mu_n^{\xi^{Q_{i,m,r}}}$, it is clear that we have (46) once we know that a suitable analogue of the Peierls estimate holds for the state $\mu_n^{\xi^{Q_{i,m,r}}} \equiv \mu_n^{\xi^{Q}}$ itself. So we will formulate it next, leaving the derivation of (46) for the end of the present section.

Let $X \in \Omega_n^+ = \Omega_n^+(\xi)$, where $\Omega_n^+(\xi)$ is the set of configurations that are $+\mathbf{1}$ outside Λ_n, except at sites in $\mathbb{M} \setminus \Lambda_n$, where they coincide with ξ. We introduce now the contours of configuration X in the usual manner. Namely, we call a face F of the dual lattice a boundary face, iff the values X_y, X_z at the sites y, z, closest to F, are different. The connected components of the boundary faces are called contours. We denote by $\mathcal{G}(X)$ the set of all contours of X, while $\mathcal{G}^{ex}(X)$ is the set of exterior contours of X; these are the contours which are not surrounded by other contours. Finally, we call a contour closed, if it does not contain faces outside Λ_n. Otherwise the contour will be called open.

If the configuration $X \in \Omega_n^+(\xi)$, and $\xi \not\equiv +1$, then X can have both open and closed contours. For $x \in \Lambda_n$ we define the contour $\Theta_x(X)$ by

$$\Theta_x(X) = \begin{cases} \Gamma, & \text{if } \Gamma \in \mathcal{G}^{ex}(X), \; x \in \mathrm{Int}\Gamma, \\ \emptyset & \text{otherwise.} \end{cases}$$

We are interested in showing that the event that the contour $\Theta_0(\cdot)$ is of the size L happens with $\mu_n^{\xi Q}$-probability $\leq \exp\{-c\beta L\}$. However, that cannot possibly be true without putting conditions on the configuration ξ. The condition should be of the type that the set $D(\xi) \subset \mathbb{M}$ of sites where $\xi \neq 1$ is quite sparse. The following generalization of the Path Large Deviation property is a condition of this type. We will still call it **PLD** property.

Definition 6. *Let $\xi \in \Omega_{\mathbb{M}}$, $\lambda > 0$ and $x \in \mathbb{Z}^d$ be given. Put*

$$\ell_\lambda(x,\xi) \equiv \ell_\lambda(x, D(\xi)) = \begin{cases} \min\{l : \text{ for all finite } k\text{-connected} \\ T \subset \mathbb{Z}^d \text{ with } x \in T, \; \mathrm{diam}(T) > l \text{ if such } T\text{-s exist,} \\ \text{we have } |T \cap D(\xi)| \leq \lambda |T|\} \\ \\ \infty & \text{otherwise.} \end{cases}$$

A set D will be called λ-sparse, iff $\ell_\lambda(x, D)$ is finite for all x. We define the set

$$\mathrm{Bad}(y, \xi; \lambda) \equiv \{x \in \mathbb{Z}^d : \ell_\lambda(x,\xi) \geq \frac{1}{2}|x-y|\}, \tag{47}$$

and we put $K(\lambda) \equiv \cap_{y \in \mathbb{Z}^d} K_y(\lambda)$ with

$$K_y(\lambda) \equiv \{\xi \in \Omega_{\mathbb{M}} : |\mathrm{Bad}(y, \xi; \lambda)| < \infty\}.$$

We say that PLD holds iff $\mu(K(\lambda)) = 1$.

Theorem 7. *Let $\xi \in K(\lambda/2)$ with $\lambda \leq \lambda(J)$, where $\lambda(J)$ is small enough. Then for all $Q \subset \mathbb{Z}^d$ finite and uniformly in large n*

$$\mu_n^{\xi Q}\{X : \mathrm{diam}(\Theta_0(X)) > L\} \leq \exp\{-c\beta L\}, \tag{48}$$

provided $L > \ell_\lambda(0,\xi)$. Here c is some constant.

Note. In case the set $D(\xi)$ itself would contain a long contour Γ surrounding the origin, the event $\mathrm{diam}(\Theta_0(X)) \geq \mathrm{diam}(\Gamma)$ happens with $\mu_n^{\xi Q}$-probability one, if $\Gamma \subset Q$, so Theorem 7 has no chance to hold in such a case. Happily, under the condition that $\xi \in K_0(\lambda)$ we have immediately that $D(\xi)$ cannot contain a contour with $\mathrm{diam}(\Gamma) > \ell_\lambda(0,\xi)$ once $\lambda < 1/2$.

Proof. We begin by showing that under our hypothesis the value $\ell_\lambda(0,\xi)$ is finite. (This is the only property of the configuration ξ needed to prove (48)). Since $\xi \in K(\lambda/2)$, we have $\ell_{\lambda/2}(x,\xi) < \infty$ for some x. Let us check now that if for some x the value $\ell_\kappa(x,\xi)$ is finite, then for all y the values $\ell_{2\kappa}(y,\xi)$ are also finite. Indeed, let $y \in T$, $\mathrm{diam}(T) = l$ and $|T \cap D(\xi)| \geq 2\kappa|T|$. Below, we denote by $[x,y] \subset \mathbb{Z}^d$ any n.n. connection between x and y, having $|x-y|$ sites. If $\max_{z \in T} |x-z| > \ell_\kappa(x,\xi)$, then

$$|(T \cup [x,y]) \cap D(\xi)| < \kappa\,|T \cup [x,y]| \leq \kappa\,|T| + \kappa\,|x-y|.$$

On the other hand,

$$|(T \cup [x, y]) \cap D(\xi)| \geq |T \cap D(\xi)| \geq 2\kappa |T| \geq 2\kappa \frac{l}{k}.$$

Therefore $l < k|x - y|$, and $\ell_{2\kappa}(y, \xi) < \max\{k|x - y|, 2\ell_\kappa(x, \xi)\}$.

For convenience, we suppress the index Q throughout the proof. First, let Γ be a contour, surrounding the origin and not attached to the boundary of the box Λ_n, i.e. a closed contour. Then we have to estimate the probability

$$\mu_n^\xi \left\{X : \Gamma \in \mathcal{G}^{ex}(X)\right\}.$$

The first remark is that it does not exceed the ratio

$$\frac{Z'(\operatorname{Int}\Gamma, +1, \beta)}{Z''(\operatorname{Int}\Gamma, +1, \beta)}, \tag{49}$$

where the partition function Z' is calculated over the subset $\Omega'(\operatorname{Int}\Gamma, \mathbb{M}, \xi)$ of configurations in $\operatorname{Int}\Gamma$, where

$$\begin{aligned}\Omega'(\Lambda, \mathbb{M}, \xi) = \{\sigma \in \Omega_\Lambda : \sigma(x) \neq +1 \text{ for every } x \in \partial_i(\Lambda), \\ \sigma(x) = \xi(x) \text{ for all } x \in \mathbb{M} \cap \Lambda\},\end{aligned} \tag{50}$$

while Z'' is calculated over $\Omega''(\operatorname{Int}\Gamma, \mathbb{M}, \xi)$ with

$$\Omega''(\Lambda, \mathbb{M}, \xi) = \{\sigma \in \Omega_\Lambda : \sigma(x) = \xi(x) \text{ for all } x \in \mathbb{M} \cap \Lambda\}. \tag{51}$$

To estimate the ratio (49), we need the following two cluster expansion results for the partition functions in the framework of the PS theory (see [Sin] and [KP] or [D1]). The first one deals with the case when we calculate the partition function $Z(\Lambda, a, \beta)$ in the box Λ with the constant boundary condition a, corresponding to the stable phase. Then

$$\ln(Z(\Lambda, a, \beta)) = |\Lambda| f(\beta) + c(\Lambda, a, \beta), \tag{52}$$

where $f(\beta)$ is the free energy of the model, and the boundary term $c(\Lambda, a, \beta)$ has the property that the ratio $\dfrac{c(\Lambda, a, \beta)}{|\partial\Lambda|}$ is exponentially small in β. The second case is obtained when we restrict the summation in the partition function (52) to the configurations σ, which possess a contour Γ right at the boundary of Λ. In other words, for all points $x \in \Lambda$, adjacent to $\partial_i\Lambda$, $\sigma(x) \neq a$. This partition function will be denoted by $Z(\Lambda, \Gamma, a, \beta)$. Then

$$\ln(Z(\Lambda, \Gamma, a, \beta)) = |\operatorname{Int}\Gamma| f(\beta) + c'(\Lambda, a, \beta) - \beta E(\Gamma). \tag{53}$$

Here $c'(\Lambda, a, \beta)$ is again a boundary term, while $E(\Gamma)$ is the (temperature independent) energy of the contour Γ. It satisfies the *Peierls condition*:

$$E(\Gamma) \geq c_1 |\Gamma| \text{ with } c_1 > 0,$$

which bound is the precondition of the PS theory.

The estimate of (49) proceeds as follows. Note that every configuration σ from the set (50) possesses in addition to the contour Γ a collection of contours $\tilde{\kappa}(\sigma) = \{\gamma_i\}$, separating the set

$$\mathbb{M}_1(\xi, \Gamma) = \{x \in \mathbb{M} \cap \mathrm{Int}\,\Gamma,\ \xi(x) = 1\} \tag{54}$$

from the contour Γ. Their presence is due to the fact that the 1-spins sitting in the set $\mathbb{M}_1(\xi, \Gamma)$ force the 1-phase inside the box $\mathrm{Int}\,\Gamma$. Our notation $\tilde{\kappa}(\sigma)$ refers to the collection of *all* such contours, and we denote by $\kappa(\sigma) \subset \tilde{\kappa}(\sigma)$ a subset of external contours of $\tilde{\kappa}(\sigma)$. The contours from $\kappa(\sigma)$ have the 1-phase inside them and are not separated from Γ by other contours. Another way of defining them is to say that the contours from $\kappa(\sigma)$ are boundaries of maximal 1-surfaces, surrounding points of $\mathbb{M}_1(\xi, \Gamma)$. We now split the partition function $Z'(\mathrm{Int}\,\Gamma, +1, \beta)$ according to what the family $\kappa = \kappa(\sigma)$ is:

$$Z'(\mathrm{Int}\,\Gamma, +1, \beta) = \sum_{\kappa} Z'(\mathrm{Int}\,\Gamma \cap \mathrm{Ext}\,\kappa, +1, \beta)\, Z''(\mathrm{Int}\,\kappa \setminus \partial_i(\mathrm{Int}\,\kappa), +1, \beta);$$

here the partition function $Z'(\mathrm{Int}\,\Gamma \cap \mathrm{Ext}\,\kappa, +1, \beta)$ is calculated over the set $\Omega'(\mathrm{Int}\,\Gamma \cap \mathrm{Ext}\,\kappa, \mathbb{M}, \xi)$. Hence,

$$\begin{aligned} &\frac{Z'(\mathrm{Int}\,\Gamma, +1, \beta)}{Z''(\mathrm{Int}\,\Gamma, +1, \beta)} \\ &\le \sum_{\kappa} \frac{Z'(\mathrm{Int}\,\Gamma \cap \mathrm{Ext}\,\kappa, +1, \beta)\, Z''(\mathrm{Int}\,\kappa \setminus \partial_i(\mathrm{Int}\,\kappa), +1, \beta)}{Z''(\mathrm{Int}\,\Gamma, +1, \beta)} \\ &\le \sum_{\kappa} \frac{Z'(\mathrm{Int}\,\Gamma \cap \mathrm{Ext}\,\kappa, +1, \beta)}{Z''(\mathrm{Int}\,\Gamma \cap \mathrm{Ext}\,\kappa, +1, \beta)}. \end{aligned} \tag{55}$$

We note for clarity that for every κ the set $\mathrm{Int}\,\Gamma \cap \mathrm{Ext}\,\kappa$ contains no sites of the set $\mathbb{M}_1(\xi, \Gamma)$.

To obtain the upper estimate for Z', we use (53). Our upper bound on Z' is

$$\begin{aligned} &\ln Z'(\mathrm{Int}\,\Gamma \cap \mathrm{Ext}\,\kappa, +1, \beta) \\ &\le |(\mathrm{Int}\,\Gamma \cap \mathrm{Ext}\,\kappa) \setminus \mathbb{M}|\, f(\beta) - \tau(|\Gamma| + |\kappa|), \end{aligned} \tag{56}$$

with τ diverging as $\beta \to \infty$. Moreover, $\tau = \tau(\beta)$ satisfies the estimate

$$\tau(\beta) \ge c_2 \beta, \tag{57}$$

with $c_2 = c_2(c_1) > 0$.

To estimate the denominator, we diminish the partition function $Z''(\mathrm{Int}\,\Gamma \cap \mathrm{Ext}\,\kappa, +1, \beta)$, replacing it by the partition function $Z'''(\mathrm{Int}\,\Gamma \cap \mathrm{Ext}\,\kappa, +1, \beta)$ calculated over the subset $\Omega'''(\mathrm{Int}\,\Gamma \cap \mathrm{Ext}\,\kappa, \mathbb{M}, \xi)$ of $\Omega''(\mathrm{Int}\,\Gamma \cap \mathrm{Ext}\,\kappa, \mathbb{M}, \xi)$, where

$$\begin{aligned} \Omega'''(\Lambda, \mathbb{M}, \xi) = \{&\sigma \in \Omega_\Lambda : \sigma(x) = \xi(x) \text{ for all } x \in \mathbb{M} \cap \Lambda, \\ &\sigma(s) = 1 \text{ for all } s \in \partial D(\xi) \cap \Lambda\}. \end{aligned}$$

Here by $\partial D(\xi)$ we denote the set

$$\partial D(\xi) = \left\{x \in \mathbb{Z}^d \setminus \mathbb{M} : \text{ for some } i \in \mathbb{M} \text{ with } \xi(i) \ne 1 \text{ we have } |x - i| = 1\right\}.$$

To estimate the partition function Z''' from below, we use (52):

$$\ln Z''' (\text{Int}\,\Gamma \cap \text{Ext}\,\kappa, +1, \beta) \geq |(\text{Int}\,\Gamma \cap \text{Ext}\,\kappa) \setminus \mathbb{M}|\, f(\beta) \tag{58}$$
$$-c_3(\beta)\left(|\Gamma| + |\kappa| + |\partial D(\xi) \cap (\text{Int}\,\Gamma \cap \text{Ext}\,\kappa)|\right) - R\,|\partial D(\xi) \cap (\text{Int}\,\Gamma \cap \text{Ext}\,\kappa)|,$$

with $c_3(\beta) \to 0$ as $\beta \to \infty$; the last term corresponds to the interaction of the 1-spins σ sitting at points $x \in \partial D(\xi)$ with spins $\xi(i) \neq 1,\ i \in \mathbb{M} \cap (\text{Int}\,\Gamma \cap \text{Ext}\,\kappa)$. Clearly,

$$R \leq \beta C(J(\cdot,\cdot)). \tag{59}$$

The crucial observation now is that the set $\partial(\text{Int}\,\Gamma) \cup [\mathbb{M} \cap \text{Int}\,\Gamma]$ is k-connected, and so the set $\mathbb{M}(\Gamma,\kappa) = \partial(\text{Int}\,\Gamma \cap \text{Ext}\,\kappa) \cup [\mathbb{M} \cap (\text{Int}\,\Gamma \cap \text{Ext}\,\kappa)]$ is k-connected as well. We are going to argue now that it implies that the term $(|\Gamma| + |\kappa|)$ from (56) suppresses the term $|\partial D(\xi) \cap (\text{Int}\,\Gamma \cap \text{Ext}\,\kappa)|$ in (58) provided the constant λ is small. To see it we first need the following lemma. It claims, roughly, that if an Ising contour Γ surrounds a λ-sparse set D, which in turn is a subset of a k-connected set T, then either the length (surface area) $|\Gamma|$ of this contour is of the order of $\lambda^{-1}|D|$, or else the contour surrounds a lot of points belonging to T.

Lemma 8. *Let $D \subset \mathbb{Z}^d$ be a finite λ-sparse set, see Def. 6, and let $T \subset \mathbb{Z}^d$ be a k-connected set, with $D \subset T$. Let κ be a finite collection of mutually external contours, while Γ is a contour surrounding the family κ.*

(The family κ can be empty.) Suppose that $D \subset (\text{Int}\,\Gamma \cap \text{Ext}\,\kappa)$. Then

$$|\Gamma| + |\kappa| + |T \cap (\text{Int}\,\Gamma \cap \text{Ext}\,\kappa)| \geq \lambda^{-1}|D|.$$

Proof. Consider the set $\partial(\text{Int}\,\Gamma \cap \text{Ext}\,\kappa) \cup [T \cap (\text{Int}\,\Gamma \cap \text{Ext}\,\kappa)]$. This set is k-connected and it contains D. Therefore

$$|D| \leq \lambda\,|\partial(\text{Int}\,\Gamma \cap \text{Ext}\,\kappa) \cup [T \cap (\text{Int}\,\Gamma \cap \text{Ext}\,\kappa)]|,$$

so the claim follows. □

We would like now to use for the set $\mathbb{M}(\Gamma,\kappa)$ the information provided by the assumptions of the theorem we are proving: namely, that $\text{diam}(\mathbb{M}(\Gamma,\kappa)) > \ell_\lambda(0,\xi)$. However, that would be of use only if $0 \in \mathbb{M}(\Gamma,\kappa)$. If that is not the case, we can modify the set $\mathbb{M}(\Gamma,\kappa)$ by attaching to it a connection to the origin of length $\leq |\Gamma|/2$. The resulting set will still be denoted by $\mathbb{M}(\Gamma,\kappa)$. Now Lemma 8, under the conditions of the theorem we are proving, provides us with the estimate:

$$|\partial D(\xi) \cap (\text{Int}\,\Gamma \cap \text{Ext}\,\kappa)| \leq \frac{3d\lambda}{1-\lambda}(|\Gamma| + |\kappa|). \tag{60}$$

(To get it we apply Lemma 8 with $D = D(\xi) \cap (\text{Int}\,\Gamma \cap \text{Ext}\,\kappa)$ and $T = \mathbb{M}(\Gamma,\kappa)$.) Using (60) together with the bounds (57), (59) and the estimates (56) and (58), we have for (55),

$$\sum_\kappa \frac{Z'(\text{Int}\,\Gamma \cap \text{Ext}\,\kappa, +1, \beta)}{Z''(\text{Int}\,\Gamma \cap \text{Ext}\,\kappa, +1, \beta)} \leq \sum_\kappa \exp\left\{-\tau'(|\Gamma| + |\kappa|)\right\} \tag{61}$$

with $\tau' = \tau'(\tau,\lambda)$ diverging with τ. We remind the reader that the last summation goes over families κ surrounding the set $\mathbb{M}_1(\xi,\Gamma) \cap \text{Int}\,\Gamma$. Hence (48) follows from (61) by standard combinatorics.

As the last step in the proof of Theorem 7 we have to consider the case when a contour Γ is not a closed contour, but is attached to the boundary of the box Λ_n, due to the impurity of the boundary conditions around the box Λ_n, caused by the presence of the conditioning by ξ. Note however that we can well assume that $n > \ell_\lambda(0, \xi)$, since otherwise the theorem holds trivially, and also that $L > n$. That implies that the set $D(\xi) \cap \partial_i (\text{Int}\,\Gamma)$ is λ-sparse. (Again, we add to Γ the connector to 0, if needed.) So the number of these impurities to which the contour Γ is attached does not exceed $\lambda |\Gamma|$. If now σ is any configuration in Λ_n with the contour Γ, then it can be modified to the value $+1$ in the vicinities of these impurities, so that the resulting configuration σ' has a (closed) contour Γ', which is not attached to the boundary of Λ_n any more. The energy cost of such a modification is no more than $\lambda C(J)|\Gamma|$, while the contour Γ' can be treated by the already proven part of our theorem. So if λ is small, the factor $\exp\left\{-c\beta \left|\Gamma'\right|\right\}$ beats the energy weight $\exp\{\beta\lambda C(J)|\Gamma|\}$, and that completes our argument. □

Proof of Estimate (46). The proof goes essentially along the same lines. Without loss of generality we can suppose that $0 \in C_F$. Let $Y = \{y \in C_F : X(y) \neq +1\}$, $Y' = \left\{y \in C_F : X'(y) \neq +1\right\}$. Then $Y \cup Y' = C_F$. Consider the following collections of contours:

$$\Theta_Y = \left\{\Gamma \in \mathcal{G}^{ex}(X) : \text{Int}\,\Gamma \cap Y \neq \emptyset\right\},$$

$$\Theta_{Y'} = \left\{\Gamma \in \mathcal{G}^{ex}\left(X'\right) : \text{Int}\,\Gamma \cap Y' \neq \emptyset\right\}.$$

Note that the union Υ of all contours from $\Theta_Y \cup \Theta_{Y'}$ is a connected set. Clearly,

$$\mu_n^{\xi^{Q_{i,m,r}}} \times \mu_n^{\xi^{Q_{i,m,r}}}[\text{diam}(C_F) > m] \leq \sum_{\substack{\Theta_Y, \Theta_{Y'}: \\ \text{diam}\left(\Theta_Y \cup \Theta_{Y'}\right) > m}} \mu_n^{\xi^{Q_{i,m,r}}}(\Theta_Y)\, \mu_n^{\xi^{Q_{i,m,r}}}(\Theta_{Y'}).$$

Using the analogue of the estimates (56), (58) and the analogue of Lemma 8 with γ replaced by the (connected) union Υ, we arrive at the following analogue of (61):

$$\mu_n^{\xi^{Q_{i,m,r}}}(\Theta_Y)\, \mu_n^{\xi^{Q_{i,m,r}}}(\Theta_{Y'}) \leq \exp\left\{-\tau'' |\Upsilon|\right\}.$$

The rest of the proof is standard combinatorics. □

6. Proof of PLD

In this last section we finally prove an unconditional statement: the PLD property, as defined in the previous section, holds with probability one. We will start with the low-temperature (+)-phase $\mathbb{P}^{\beta,+}$ of the Ising model. We consider the general case in the last subsection.

The events we are interested in here are

$$B(\varepsilon, W) = \left\{\sigma \in \Omega : m^*(\beta) - \frac{1}{|W|}\sum_{x \in W} \sigma_x > \varepsilon\right\}. \tag{62}$$

We now introduce the notation $\mathcal{S}_V(x; k, N)$ for the family of all k-connected subsets of the box V of diameter between N and $2N$, containing a site x, with $\mathcal{S}_V(k, N) \equiv \mathcal{S}_V(0; k, N)$. Then we have the following theorem:

Theorem 9. *For every $k > 0$, $\varepsilon_1 > 0$ there exists a value β_k, such that for all $\beta > \beta_k$ and all V containing x*

$$\mathbb{P}_V^{\beta,+}\left[\bigcup_{W\in\mathcal{S}_V(x;k,N)} B(\varepsilon, W)\right] \leq \exp\left\{-\beta N^{\frac{d-1}{(2+\varepsilon_1)d}}\right\}. \tag{63}$$

Let us now fix N, and consider the union

$$\overline{\mathcal{S}}_{\mathbb{Z}^d}(y; k, N) = \bigcup_{x\in\mathbb{Z}^d} \mathcal{S}_{\mathbb{Z}^d}\left(x; k, N + \frac{1}{3}|x-y|\right).$$

Note that because of (63)

$$\mathbb{P}^{\beta,+}\left[\bigcup_{W\in\overline{\mathcal{S}}_{\mathbb{Z}^d}(y;k,N)} B(\varepsilon, W)\right] \to 0$$

as $N \to \infty$. On the other hand, if $\sigma \in \Omega \setminus \left(\bigcup_{W\in\overline{\mathcal{S}}_{\mathbb{Z}^d}(0;k,N)} B(\varepsilon, W)\right)$, then the set Bad $\left(0, \sigma; \bar{\lambda}\right)$ (see (20), (47)) with

$$\bar{\lambda} = \bar{\lambda}(\varepsilon, \beta) = \frac{1 + \varepsilon - m^*(\beta)}{2}$$

is contained within the ball of radius $6N$ around the origin and is therefore finite. Hence, the set of configurations

$$\tilde{K}(\varepsilon) = \bigcap_{y\in\mathbb{Z}^d}\left(\Omega \setminus \bigcap_N \left(\bigcup_{W\in\overline{\mathcal{S}}_{\mathbb{Z}^d}(y;k,N)} B(\varepsilon, W)\right)\right)$$

is contained in the set $K\left(\bar{\lambda}\right)$. Since $\tilde{K}(\varepsilon)$ has full measure, PLD is satisfied. Note finally that $\bar{\lambda}(\varepsilon, \beta) \to 0$ as $\varepsilon \to 0$, $\beta \to \infty$, so for ε small and β large $\bar{\lambda}(\varepsilon, \beta) \leq \lambda(J)$ (see Theorem 7).

So what is left is the proof of Theorem 9. We will consider the case $x = 0$, therefore x will disappear from the notation.

6.1. The strategy. The proofs of the above results turn out to be quite non-trivial. To explain the nature of the difficulties, we present below a short account of the straightforward idea of the proof, together with the explanation why it does not work.

Let us try to prove Theorem 9 for the case of $k = 1$; moreover, let us take $V = V_n$ to be a cube, and restrict the set $\mathcal{S}_{V_n}(k)$ to consist of selfavoiding paths only, connecting the origin to the boundary ∂V_n. So, we are talking about the events which may be called "large deviations along paths". If we would be able to prove the estimate

$$\mathbb{P}_{V_n}^{\beta,+}\left[B(\varepsilon, \gamma)\right] \leq \exp\{-c(\beta)|\gamma|\}, \tag{64}$$

valid for every path γ, with the exponent $c(\beta)$ diverging as $\beta \to \infty$, then we would be done. Indeed, the number of paths γ containing the origin and having length l is bounded by 3^l, so we can finish the proof by summation over all paths. The point however is that the estimate (64) does not hold in general. To see it can be violated, let us introduce the set $C(\Gamma)$ of configurations having the contour Γ among their exterior contours, and take Γ to be a boundary of a cubic box centered at the origin and having the volume n. Then

$$\mathbb{P}_{V_n}^{\beta,+}[C(\Gamma)] \geq \exp\left\{-c(\beta)\, n^{\frac{d-1}{d}}\right\}.$$

Now, consider a shortest possible path γ_Γ, which visits all points inside Γ prior to leaving for ∂V. Its length is proportional to n. On the other hand, evidently

$$\mathbb{P}_{V_n}^{\beta,+}\left[B(\varepsilon,\gamma)\right] \geq \mathbb{P}_{V_n}^{\beta,+}\left[B(\varepsilon,\gamma)\,|C(\Gamma)\right]\mathbb{P}_{V_n}^{\beta,+}[C(\Gamma)],$$

and the first factor is larger than $\frac{1}{2}$ for β and n large. So the probability $\mathbb{P}_{V_n}^{\beta,+}\left[B(\varepsilon,\gamma)\right]$ decays in n subexponentially only.

The reason why the above argument fails is that the summation over all paths includes an overcounting; the same set of configurations makes many different paths to have the wrong magnetization along them. On the other hand, to control the contribution of the set $C(\Gamma)$ to the probability $\mathbb{P}_{V}^{\beta,+}\left[\bigcup_{\gamma\in\mathcal{S}_V(k)} B(\varepsilon,\gamma)\right]$, one does not need any extra argument involving path counting, simply because $\mathbb{P}_{V}^{\beta,+}[C(\Gamma)] \leq \exp\left\{-c'(\beta)\, n^{\frac{d-1}{d}}\right\}$ (Peierls estimate).

To proceed with the proof of Theorem 9, we introduce some more definitions. Let a configuration $\sigma \in \Omega^+$ be given, where $\Omega^+ = \cup_n \Omega_n^+$. We denote by $G(\sigma)$ the set of its exterior contours. We call a set $W \in \mathcal{S}_V(k,N)$ a *bad set* for σ, iff

$$\frac{\sum_{\substack{\Gamma\in G(\sigma):\\ \text{Int}\,\Gamma\cap W\neq\emptyset}} |\text{Int}\,\Gamma|}{|W|} \geq \delta. \tag{65}$$

Here $\delta > 0$ is some fixed number. Clearly, if σ belongs to the event $B(\varepsilon, W)$, then W is a bad set for σ, with $\delta = \delta(\varepsilon,\beta)$. We denote the event (65) by $C(\delta, W)$.

As one will see later, the proofs of the above theorems require the introduction of different scales, and these scales are needed to treat contours of different sizes. Anticipating that, we introduce now the events $C_r(\cdot, W)$, $r = 1, 2, ...,$ as follows:

$$\sigma \in C_r\left(\frac{\delta_r}{r}, W\right) \Longleftrightarrow \frac{\sum_{\substack{\Gamma\in G(\sigma):\,|\text{Int}\,\Gamma|=r\\ \text{Int}\,\Gamma\cap W\neq\emptyset}} 1}{|W|} \geq \frac{\delta_r}{r}. \tag{66}$$

If σ belongs to the event $C_r\left(\frac{\delta_r}{r}, W\right)$, then we say that the set W is an *r-bad set* for σ.

Our choice of the parameters δ_r will be the following:

$$\delta_r = \frac{\delta}{K(\varepsilon_1)\, r^{1+\varepsilon_1}}, \tag{67}$$

with any positive $\varepsilon_1 < \frac{1}{2d}$. In this way the inclusion

$$C(\delta, W) \subset \cup_r C_r\left(\frac{\delta_r}{r}, W\right) \tag{68}$$

holds, provided $\sum_1^\infty \frac{1}{r^{1+\varepsilon_1}} < K(\varepsilon_1)$. That means, in words, that if W is a bad set for σ, then W is also r-bad for σ for at least one value of r.

Let now the sequence of integers $x_1 \le x_2 \le \dots$ be given, $x_1 \ge k$ (where k is the value of the connection parameter entering the formulation of our theorems). The values x_r will be used as scales for studying the configurations with r-bad sets. The choice of these scales, depending on δ_r, will also be made later. The sequence x_r does not depend on any other parameter, but for a fixed value of N we will use only the first $r(N)$ terms of it, where $r(N)$ is defined in (80) below. The reason for that is that for a given N the contribution of contours of size $r > r(N)$ to the event (65) can be neglected.

Now we will start to estimate the probability of the event $\cup_{W\in\mathcal{S}_V(k,N)} C_r\left(\frac{\delta_r}{r}, W\right)$. In order to avoid the overcounting, to which we were alluding above, we will make a coarse graining of our system. So we introduce the natural partition $\mathcal{L}_r$ of $\mathbb{Z}^d$ into cubes A_{x_r} of size x_r, with faces parallel to the coordinate planes:

$$\mathcal{L}_r = \left\{ A_{x_r}(y) \equiv A_{x_r} + y,\ y \in x_r\mathbb{Z}^d \right\},$$

and we denote by W^r the x_r-fattening of W, i.e. the union of all those cubes $A_{x_r}(y)$ of $\mathcal{L}_r$, which contain at least one point of the initial set W. Note that the set W^r is always connected, since W is supposed to be k-connected, while $x_r \ge k$. If $\sigma \in C_r\left(\frac{\delta_r}{r}, W\right)$, then evidently

$$\frac{\sum_{\substack{\Gamma\in G(\sigma):\,|\mathrm{Int}\,\Gamma|=r\\ \mathrm{Int}\,\Gamma\cap W^r\neq\emptyset}} 1}{|W|} \ge \frac{\sum_{\substack{\Gamma\in G(\sigma):\,|\mathrm{Int}\,\Gamma|=r\\ \mathrm{Int}\,\Gamma\cap W\neq\emptyset}} 1}{|W|} \ge \frac{\delta_r}{r}, \tag{69}$$

so

$$\frac{\sum_{\substack{\Gamma\in G(\sigma):\,|\mathrm{Int}\,\Gamma|=r\\ \mathrm{Int}\,\Gamma\cap W^r\neq\emptyset}} 1}{|W^r|} \ge \frac{\delta_r}{\kappa r\ (x_r)^{d-1}}, \tag{70}$$

provided the estimate

$$\left|W^r\right| \le \kappa\,(x_r)^{d-1}\,|W| \tag{71}$$

holds for some $\kappa = \kappa(d,k)$, all $W \in \mathcal{S}_V(k,N)$ and all $r = 1, 2, \dots, r(N)$. The estimate (71) is indeed valid under the condition that the size $x_{r(N)}$ of the cubes of the partition $\mathcal{L}_{r(N)}$ is much smaller than N. The reason is that if $\gamma \subset \mathbb{Z}^d$ is a path of length $|\gamma| \ge N$, then the number of cubes of the partition $\mathcal{L}_r$ the path γ can hit is bounded from above by $C(d)\,|\gamma|/x_r$, provided $x_r \ll N$. This last condition is ensured for all N large enough and all $r = 1, 2, \dots, r(N)$ by the choices (80) and (86) made below.

Let us now introduce the family $\mathcal{S}_V^r(N)$ as the collection of all connected subsets of $\mathbb{Z}^d$ of diameter between N and $3N$, containing the origin, which are made from the cubes of the partition $\mathcal{L}_r$, intersecting V. For future use we introduce for the sets $W \in \mathcal{S}_V^r(N)$ the notation $||W||$ for the number of x_r-cubes they are composed of; of course, $||W|| = (x_r)^{-d}\,|W|$. It follows from (69), (70) and the definition (66) that

$$\mathbb{P}_V^{\beta,+}\left\{\bigcup_{W\in\mathcal{S}_V(k,N)} C_r\left(\frac{\delta_r}{r}, W\right)\right\} \le \mathbb{P}_V^{\beta,+}\left\{\bigcup_{W\in\mathcal{S}_V^r(N)} C_r\left(\frac{\delta_r}{\kappa\, r\,(x_r)^{d-1}}, W\right)\right\}$$
$$\le \sum_{W\in\mathcal{S}_V^r(N)} \mathbb{P}_V^{\beta,+}\left\{C_r\left(\frac{\delta_r}{\kappa r\ (x_r)^{d-1}}, W\right)\right\}. \quad (72)$$

A heuristic comment. At first glance the maneuver we performed in (72) could have well been done with the sets W themselves, without passing to the fattenings W^r. The key difference however lies in the amount of events which are of the form of (66) or (70). For example, there are about $C(d,k)^N$ different sets W with $|W| = N$, which may appear in (66), so in order to have the summation over them to converge, we need the estimate of the probability of the event (66) to be exponentially small in N. But such an estimate simply does not hold! On the other hand, the corresponding number of different (connected) sets which can be produced from these W-s via fattening by x_r-cubes is bounded by $\left(3^d\right)^{\frac{2C(d)N}{xr}}$, and that makes the summation possible, as we will see soon.

Note that the Peierls estimate implies immediately, that for every set $\tilde{W}$ and $\tilde{\delta} > 0$ we have

$$\mathbb{P}_V^{\beta,+}\left\{C_r\left(\tilde{\delta}, \tilde{W}\right)\right\} \le \sum_{k=\tilde{\delta}\left|\tilde{W}\right|}^{\left|\tilde{W}\right|} \binom{\left|\tilde{W}\right|}{k}\left(\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\right)^k, \quad (73)$$

with β' diverging together with β. Indeed, we have to have $\tilde{\delta}\left|\tilde{W}\right|$ exterior contours, surrounding points of $\tilde{W}$, which explains the first factor. The surface of every contour contributing to the lhs of (73) is at least $2dr^{\frac{d-1}{d}}$, the number of different Γ-s, containing a given face and having the surface L, is bounded by 3^L, and

$$\sum_{L\ge 2dr^{\frac{d-1}{d}}} 3^L \exp\{-2\beta L\} \le \exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}.$$

We want to interpret the rhs of (73) as a probability. So we introduce the Bernoulli random field, made by i.i.d. Bernoulli random variables $\xi_i^r,\ i\in\mathbb{Z}^d$:

$$\xi_i^r = \begin{cases} 1 \text{ with probability } p_r = \exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\left(1+\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\right)^{-1}, \\ 0 \text{ with probability } q_r = \left(1+\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\right)^{-1}. \end{cases} \quad (74)$$

Then we can rewrite (73) as

$$\mathbb{P}_V^{\beta,+}\left\{C_r\left(\tilde{\delta},\tilde{W}\right)\right\} \le \left(1+\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\right)^{\left|\tilde{W}\right|}\sum_{k=\tilde{\delta}\left|\tilde{W}\right|}^{\left|\tilde{W}\right|}\binom{\left|\tilde{W}\right|}{k}p_r{}^k q_r{}^{\left|\tilde{W}\right|-k}$$
$$\le \left(1+\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\right)^{\left|\tilde{W}\right|}\mathbb{P}\left\{\sum_{i=1}^{\left|\tilde{W}\right|}\xi_i^r \ge \tilde{\delta}\left|\tilde{W}\right|\right\}. \tag{75}$$

That is why we will study now the large deviation properties of the random field ξ_i^r.

6.2. Large deviations for Bernoulli variables. The results in this section are fairly standard. We present them just for completeness.

Lemma 10. *Let ξ_i be a sequence of Bernoulli random variables,*

$$\xi_i = \begin{cases} 1 \text{ with probability } p, \\ 0 \text{ with probability } 1-p. \end{cases}$$

Let $S_K = \sum_{i=1}^K \xi_i$. Then for every k and every $z > 2$ we have the estimate

$$\mathbb{P}\left(S_K > k\right) \le z^{-k}\left(1-p+pz\right)^K. \tag{76}$$

Proof. We use Cramer tilting. We have

$$\mathbb{P}\left(S_K = k\right) = \binom{K}{k}p^k\left(1-p\right)^{K-k} \equiv$$
$$\equiv \binom{K}{k}\left(pz\right)^k\left(1-p\right)^{K-k}z^{-k} \equiv$$
$$\equiv \left[\frac{\binom{K}{k}\left(pz\right)^k\left(1-p\right)^{K-k}}{\left(1-p+pz\right)^K}\right]z^{-k}\left(1-p+pz\right)^K.$$

But the expression in the square brackets is again the probability of the same event, now according to the Bernoulli sequence with $p' = \frac{pz}{1-p+pz}$, $1-p' = \frac{1-p}{1-p+pz}$. Summation over k yields the result. □

Corollary 11. *Let A be a real number, such that $Ap < 1$. Then*

$$\mathbb{P}\left(S_K > ApK\right) \le \varepsilon^K, \tag{77}$$

with

$$\varepsilon = \varepsilon\left(p, A\right) = \left(\frac{1-p}{1-Ap}\right)^{(1-Ap)}\left(\frac{1}{A}\right)^{Ap}. \tag{78}$$

Note. Of course, the notation $\varepsilon\left(p, A\right)$ does not necessarily make the quantity (78) small. This notation just expresses our hope.

Proof. Let us take z to be the solution z_0 of the equation

$$\frac{pz}{1-p+pz} = Ap.$$

(Note that such a choice of z makes the p'-Bernoulli process of the previous proof to have the mean value of the sum of K its element to be ApK.) The solution is given by

$$z_0 = A\frac{1-p}{1-Ap}.$$

Note also the relation $1 - p + pz_0 = \frac{z_0}{A}$. Putting this into (76), we have

$$\mathbb{P}\left(S_K > ApK\right) \leq z_0^{-ApK}\left(\frac{z_0}{A}\right)^K = \frac{\left(z_0\right)^{K(1-Ap)}}{A^K} = \\ = \frac{\left(\frac{1-p}{1-Ap}\right)^{K(1-Ap)}}{A^{ApK}} \equiv \left[\left(\frac{1-p}{1-Ap}\right)^{(1-Ap)}\left(\frac{1}{A}\right)^{Ap}\right]^K. \qquad \square$$

6.3. Proof of Theorem 9. i) Let the number N be fixed. Let us denote by $R\left(>k, N\right)$ the set of all configurations $\sigma = \left(\sigma_x, x \in V\right)$, which have a large contour $\Gamma \in \mathcal{G}\left(\sigma\right)$ in the vicinity of the origin:

$$\text{dist}\left(0, \text{Int}\,\Gamma\right) < 2N,$$

$$|\Gamma| > k.$$

Then it is immediate to see that if

$$k\left(N\right) > \ln N,$$

then

$$\mathbb{P}_V^{\beta,+}\left(R\left(>k\left(N\right), N\right)\right) \leq \exp\left\{-\beta' k\left(N\right)\right\} \to 0 \text{ as } N \to \infty, \tag{79}$$

so we need to study only the intersection $\left[\bigcup_{W\in\mathcal{S}_V(k,N)} B\left(\varepsilon, W\right)\right] \cap R\left(<k\left(N\right), N\right)$. Note that this intersection satisfies the inclusion

$$\left[\bigcup_{W\in\mathcal{S}_V(k,N)} B\left(\varepsilon, W\right)\right] \cap R\left(<k\left(N\right), N\right) \subset \cup_{r=1}^{r(N)} C_r\left(\frac{\delta_r}{r}, W\right)$$

(compare with (68)) if

$$r\left(N\right) = k\left(N\right)^{\frac{d}{d-1}}. \tag{80}$$

For our purposes the optimal choice of the function $k(N)$ turns out to be

$$k(N) = N^{\frac{d-1}{2d(1+\varepsilon_1)}}. \tag{81}$$

ii) To estimate the probability $\mathbb{P}_V^{\beta,+}\left\{C_r\left(\frac{\delta_r}{\kappa r\ (x_r)^{d-1}}, W^r\right)\right\}$ of the event (70) we will use first the estimate (75) and then (77) with

$$p = p_r = \exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\left(1+\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\right)^{-1},$$

see (74), and

$$A = p_r^{-1}\frac{\delta_r}{\kappa r\ \ (x_r)^{d-1}}.$$

We have:

$$\begin{aligned}
&\mathbb{P}_V^{\beta,+}\left\{C_r\left(\frac{\delta_r}{\kappa\, r\ (x_r)^{d-1}}, W^r\right)\right\} \le \\
&\left\{\left[1+\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\right]^{(x_r)^d}\right\}^{||W^r||} \\
&\left\{\left[\left(\frac{1-\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\left(1+\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\right)^{-1}}{1-\dfrac{\delta_r}{\kappa r\ \ (x_r)^{d-1}}}\right)^{\left(1-\frac{\delta_r}{\kappa r\ (x_r)^{d-1}}\right)} \times\right.\right. \\
&\left.\left.\times\left(\frac{\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\left(1+\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\right)^{-1}}{\dfrac{\delta_r}{\kappa r\ \ (x_r)^{d-1}}}\right)^{\frac{\delta_r}{\kappa r\ (x_r)^{d-1}}}\right]^{(x_r)^d}\right\}^{||W^r||}.
\end{aligned} \tag{82}$$

If we are able to show that the product of the curly brackets in the last expression is small for every r, then we would be done, since that would enable us to beat the entropy factor $\left(3^d\right)^{2||W^r||}$. (It is not hard to see that the quantity $\left(3^d\right)^{2k}$ estimates the number of connected sets made from k unit cubes in $\mathbb{R}^d$, containing a given one.) So we will look at the logarithm of $\{\cdot\}$ in (82). It is equal to

$$\begin{aligned}
\ln\ \{\,\cdot\,\} &= (x_r)^d \ln\left[1+\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\right] \\
&+ (x_r)^d\left(1-\frac{\delta_r}{\kappa r\ \ (x_r)^{d-1}}\right)\ln\left(\frac{1-\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\left(1+\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\right)^{-1}}{1-\dfrac{\delta_r}{\kappa r\ \ (x_r)^{d-1}}}\right) \\
&+ (x_r)^d\,\frac{\delta_r}{\kappa r\ \ (x_r)^{d-1}}\ln\left(\frac{\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\left(1+\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\right)^{-1}}{\dfrac{\delta_r}{\kappa r\ \ (x_r)^{d-1}}}\right)
\end{aligned}$$

$$\lesssim (x_r)^d \exp\left\{-\beta' r^{\frac{d-1}{d}}\right\} + (x_r)^d \left(1 - \frac{\delta_r}{\kappa r\ (x_r)^{d-1}}\right)$$
$$\cdot \left[\frac{\delta_r}{\kappa r\ (x_r)^{d-1}} - \exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\left(1+\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}\right)^{-1}\right]$$
$$+\frac{x_r\delta_r}{\kappa r}\left[-\beta' r^{\frac{d-1}{d}} - \exp\left\{-\beta' r^{\frac{d-1}{d}}\right\} - \ln\frac{\delta_r}{\kappa r\ (x_r)^{d-1}}\right]$$
$$\lesssim -\frac{x_r\delta_r}{\kappa r}\left(\beta' r^{\frac{d-1}{d}} + \ln\frac{\delta_r}{\kappa r\ (x_r)^{d-1}} - 1\right) + (x_r)^d \exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}. \quad (83)$$

The only thing needed for the validity of the above estimate is the smallness of $\exp\left\{-\beta' r^{\frac{d-1}{d}}\right\}$ and $\frac{\delta_r}{\kappa r\ (x_r)^{d-1}}$, which is not an extra constraint. From this estimate we see that the following requirements are sufficient for the numbers x_r to make the above logarithm very negative:

$$\frac{x_r\delta_r}{\kappa r} \geq \alpha > 0 \text{ for some } \alpha \text{ and all } r, \quad (84)$$

$$(x_r)^d \ll \alpha \exp\left\{\beta' r^{\frac{d-1}{d}}\right\} \text{ for all } r. \quad (85)$$

Indeed, under (84)

$$\beta' r^{\frac{d-1}{d}} + \ln\frac{\delta_r}{\kappa r\ (x_r)^{d-1}} = \ln\left[\exp\left\{\beta' r^{\frac{d-1}{d}}\right\}\frac{\delta_r}{\kappa r\ (x_r)^{d-1}}\right]$$
$$\geq \ln\left[\alpha \exp\left\{\beta' r^{\frac{d-1}{d}}\right\}(x_r)^{-d}\right] \gg 1,$$

because of (85). For example, any choice

$$x_r \sim r^{2+3\varepsilon_1/2} \quad (86)$$

would go. Then the dominating term in (83) would be

$$-\beta'\frac{x_r\delta_r}{\kappa r} r^{\frac{d-1}{d}} \ll -1$$

uniformly in r, provided β is large.

iii) What remains now is to estimate the sum

$$\sum_{r=1}^{k(N)^{\frac{d}{d-1}}} \sum_{W^r \in \mathcal{S}_V^r(N)} \mathbb{P}_V^{\beta,+}\left\{C_r\left(\frac{\delta_r}{\kappa\, r\, (x_r)^{d-1}}, W^r\right)\right\}.$$

For $W^r \in \mathcal{S}_V^r(N)$ we have

$$||W_r|| \geq \frac{N}{x_r}.$$

Hence, by the estimates (82), (83) the last sum is bounded by

$$\sum_{r=1}^{k(N)^{\frac{d}{d-1}}} \sum_{L \geq \frac{N}{x_r}} \left(3^d\right)^{2L} \exp\left\{-\beta' \frac{x_r \delta_r}{\kappa r} r^{\frac{d-1}{d}} L\right\}.$$

According to the choices made before, see (67), the second exponent beats the first one, and the resulting upper bound is, due to (81),

$$\sum_{r=1}^{k(N)^{\frac{d}{d-1}}} \exp\left\{-\beta' \frac{\delta_r}{\kappa r} r^{\frac{d-1}{d}} N\right\} \leq \exp\left\{-\beta' N^{\frac{d-1}{2d(1+\varepsilon_1)}}\right\}.$$

Together with the estimate (79) and again in view of the choice made in (81), that proves the result. □

Remark 1. The reader might have noted from (85) that the values of β for which our results hold, depend on k and go to infinity as k increases. We believe that this is only due to technical reasons, and in fact the same result should hold for β large enough, uniformly in k.

This belief turned out to be correct, as it is shown in the paper [S] of one of us.

6.4. The general case. In this subsection we use the notation $\mathbb{P}^{\beta,+}$ for the Gibbs state of our Hamiltonian (9) corresponding to the boundary condition $+1$.

For $\mathbb{P}^{\beta,+}$-almost every configuration σ the set of points $x \in \mathbb{Z}^d$, where $\sigma(x) = 1$ contains a unique infinite component $\mathrm{Ext}\,\sigma$, with all connected components of the complement $\mathbb{Z}^d \setminus \mathrm{Ext}\,\sigma$ finite. We will denote these components by $\Delta_i(\sigma)$, their collection by $\mathcal{D}(\sigma)$, and will call them *droplets of* σ. In analogy with (62) we introduce the event

$$B(\varepsilon, W) = \left\{\sigma \in \Omega : \frac{1}{|W|} \sum_{i:\Delta_i(\sigma)\cap W \neq \emptyset} |\Delta_i(\sigma)| > \varepsilon\right\}.$$

With this notation Theorem 9 is valid in the above generality.

Proof. The proof of this statement is the same one as is given for the case of the Ising model. The only extra ingredient needed for the general case is the estimate that for every finite connected set $\Delta \subset \mathbb{Z}^d$,

$$\mathbb{P}^{\beta,+}(\sigma : \Delta \in \mathcal{D}(\sigma)) \leq \exp\{-c(\beta)|\partial\Delta|\},$$

with $c(\beta)$ diverging with β. But this is a standard corollary of the PS theory. □

7. Outlook

We end the paper with a short discussion concerning logical continuations of the present work. First of all, it is clear that not all interesting cases of non-Gibbsian states have been covered. As an example, we have not dealt here with fuzzy descriptions of Gibbs random fields [MvdV2] which can be seen as a coarse graining of the spin space. A well known example is the projection of the massless Gaussian field on the sign variables, see [LM,EFS,ES]. Moreover, even for the projection to a sublattice, our assumption that the original measure is a low temperature phase of the PS theory, is generally violated for the space-time Gibbs measures describing the steady state of a stochastic dynamics. So, we have no results on the Gibbsianness of the projections to spatial layers, hence on the Gibbsian character of stationary measures in the coexistence regime, see [LMS, GKLM].

Finally, the present work should be followed by establishing the standard results of the Gibbs formalism (existence of thermodynamic potentials, variational principle, etc.). The basis for that can be the structure of weakly Gibbsian fields, uncovered in the present work, see the estimate (25) and Definition 6. That would contribute to the second part of the Dobrushin program of Gibbsian Restoration.

References

[B] van den Berg, J.: A uniqueness condition for Gibbs measures, with application to the 2-dimensional Ising antiferromagnet. Commun. Math. Phys. **152**, 161–166 (1993)

[BK] Bricmont, J. and Kupiainen, A.: Phase transition in the 3D random field Ising model. Commun. Math. Phys. **116**, 539–572 (1988)

[BKL] Bricmont, J., Kupiainen, A. and Lefevere, R.: Renormalization group pathologies and the definition of Gibbs states. Commun. Math. Phys. **194**, 359–388 (1998)

[BM] van den Berg, J. and Maes, C.: Disagreement percolation in the study of Markov fields. Ann. Prob. **22**, 749–763 (1994)

[BS] Burton, R.M. and Steif, J.E.: Quite weak Bernoulli with exponential rate and percolation for random fields. Stoch. Process. Appl. **58**, 35 (1995)

[D1] Dobrushin, R.L.: Estimates of semi-invariants for the Ising model at low temperatures. In: *Beresin memorial volume*, AMS Translations Series, vol. **177**(2), 59–81 (1996)

[D2] Dobrushin, R.L.: A Gibbsian representation for non-Gibbsian fields. Lecture given at the workshop “Probability and Physics”, September 1995, Renkum (the Netherlands)

[DS] Dobrushin, R.L. and Shlosman, S.B.: “Non-Gibbsian” states and their Gibbs description. Commun. Math. Phys., **200**, 1, 125–179 (1999)

[EFS] van Enter, A.C.D., Fernandez, R. and Sokal, A.D.: Regularity properties and pathologies of position-space renormalization transformations: scope and limitations of Gibbsian theory. J. Stat. Phys. **72**, 879–1167 (1993)

[EMS] van Enter, A.C.D., Maes, C. and Shlosman, S.B.: Dobrushin’s program on Gibbsianity restoration: Weakly Gibbsian and Almost Gibbsian random fields. In: *On Dobrushin’s way. From Probability Theory to Statistical Mechanics*, R.A. Minlos, S.B. Shlosman, Yu.M. Suhov eds, Providence, RI: AMS, 1999

[ES] van Enter, A.C.D. and Shlosman, S.B.: (Almost) Gibbsian description of the sign-fields of SOS-fields. J. Stat. Phys. **92**, 353–368 (1998)

[GKLM] Goldstein, S., Kuik, R., Lebowitz, J.L. and Maes, Ch.: From PCAs to equilibrium systems and back. Commun. Math. Phys. **125**, 71–79 (1989)

[GP1] Griffiths, R.B. and Pearce, P.A.: Position-space renormalization-group transformations: Some proofs and some problems. Phys. Rev. Lett. **41**, 917–920 (1978)

[GP2] Griffiths, R.B. and Pearce, P.A.: Mathematical problems of position-space renormalization-group transformations. J. Stat. Phys. **20**, 499–545 (1979)

[I] Israel, R.B.: Banach algebras and Kadanoff transformations. In: *Random Fields*, Esztergom, 1979, J. Fritz, J.L. Lebowitz, D. Szasz eds, Amsterdam: North-Holland, 1981, **2**, pp. 593–608

[KP] Kotecky, R. and Preiss, D.: Cluster expansion for abstract polymer models. Commun. Math. Phys. **103**, 491–498 (1986)

[Koz] Kozlov, O.K.: Gibbs description of a system of random variables. Prob. Inform. Transmission **10**, 258–265 (1974)

[LM] Lebowitz, J.L. and Maes, C.: The effect of an external field on an interface, entropic repulsion. J. Stat. Phys. **46**, 39–49 1987

[LMS] Lebowitz, J.L., Maes, C. and Speer, E.: Statistical mechanics of probabilistic cellular automata. J. Stat. Phys. **59**, 117–170 1990

[LMV] Lorinczi, J., Maes, C. and Vande Velde, K.: Transformations of Gibbs measures. Prob. Th. Rel. Fields, **112**, 121–147 (1998)

[MRV1] Maes, K., Redig, F. and Van Moffaert, A.: Almost Gibbsian versus weakly Gibbsian measures. Stoch. Proc. Appl. **79**, 1–15 (1999)

[MRV2] Maes, C., Redig, F. and Van Moffaert, A.: The restriction of the Ising model to a layer. To appear in J. Stat. Phys. (1999)

[MvdV1] Maes, C. and Vande Velde, K.: Relative energies for non-Gibbsian states. Commun. Math. Phys. **189**, 277–286 (1997)

[MvdV2] Maes, C. and Vande Velde, K.: The fuzzy Potts model. J. Phys. A **28**, 4261–4271 (1995)

[Sch] Schonmann, R.H.: Projections of Gibbs measures can be non-Gibbsian. Commun. Math. Phys. **124**, 1–7 (1989)

[S] Shlosman, S.: Path large deviation and other typical properties of the low-temperature models, with applications to the weakly Gibbs states. In preparation

[Sin] Sinai, Ya.G.: *Theory of Phase Transitions: Rigorous Results*. Oxford: Pergamon Press, 1982

Communicated by Ya. G. Sinai