Pinning and disorder relevance for the lattice Gaussian free field

Giambattista Giacomin, Hubert Lacoin

PINNING AND DISORDER RELEVANCE FOR THE LATTICE GAUSSIAN FREE FIELD

GIAMBATTISTA GIACOMIN AND HUBERT LACOIN

ABSTRACT. This paper provides a rigorous study of the localization transition for a Gaussian free field on $\mathbb{Z}^d$ interacting with a quenched disordered substrate that acts on the interface when its height is close to zero. The substrate has the tendency to localize or repel the interface at different sites and one can show that a localization-delocalization transition takes place when varying the average pinning potential h: the free energy density is zero in the delocalized regime, that is for h smaller than a threshold h_c, and it is positive for $h > h_c$. For $d \geq 3$ we compute h_c and we show that the transition happens at the same value as for the annealed model. However we can show that the critical behavior of the quenched model differs from the one of the annealed one. While the phase transition of the annealed model is of first order, we show that the quenched free energy is bounded above by $((h-h_c)_+)^2$ times a positive constant and that, for Gaussian disorder, the quadratic behavior is sharp. Therefore this provides an example in which a *relevant disorder critical exponent* can be made explicit: in theoretical physics disorder is said to be *relevant* when the disorder changes the critical behavior of a system and, while there are cases in which it is known that disorder is relevant, the exact critical behavior is typically unknown. For $d = 2$ we are not able to decide whether the quenched and annealed critical points coincide, but we provide an upper bound for the difference between them.

2010 *Mathematics Subject Classification: 60K35, 60K37, 82B27, 82B44*

Keywords: Lattice Gaussian Free Field, Disordered Pinning Model, Localization Transition, Critical Behavior, Disorder Relevance, Co-membrane Model

CONTENTS

1. Introduction

A central question in statistical mechanics is understanding the effect of disorder on phase transitions and critical phenomena. This issue has been raised soon after Lars Onsager's celebrated solution of the two dimensional Ising model with zero external field, see [28, Section 5.3] for a historical overview and references. The model solved by Onsager has constant couplings – Onsager's solution actually allows couplings that are different in the horizontal and vertical directions, but not more than that – and the question of whether this result withstands, and to which extent, the introduction of impurities emerged as a compelling *stability* issue. Modeling systems with impurities naturally led to considering systems in which the interaction terms, for example the potentials between nearest neighbor spins, are chosen by sampling a random field – which we call *disorder* – with good ergodic properties, often even a field of independent identically distributed random variables. One then tries to understand the properties of the arising statistical mechanics system which is no longer translation invariant, even if it retains some translation invariance in a statistical sense. Some basic results like existence of the thermodynamic limit and the fact that observables are self-averaging (i.e., independent on the sample of the disorder) can be established [12]. When the variance of the disorder tends to zero the system approaches the non disordered, or *pure*, case but transferring a result proven for the pure system to the disordered case, even if the disorder is very weak, is far from being straightforward.

As a matter of fact, the first arguments set forth pointed toward predicting that even a very low amount of disorder would wash out completely the phase transition [28, Section 5.3]. Only later on a substantially richer picture emerged. Since the Ising model has to a certain extent driven the progress, it is worthwhile recalling that a disorder in the form of an external random field makes the Ising transition disappear in two dimensions [1], while the transition persists in $d \geq 3$ [13]. On the other hand, it is not difficult to realize that introducing a disorder in the coupling potentials, for example by introducing a dilution, may in general modify the precise value of the critical point, but preserves the existence of a transition: the nature of the transition (for example, the critical exponents) is however still an open question (at least in low dimensions) [28, Section 5.3]. Giving more details on this beautiful issue is beyond our scope, but what interests us the most is that A. B. Harris [33] introduced an intriguing way of looking at the problem and proposed a surprisingly easy criterion to predict whether a small amount of disorder modifies the critical behavior with respect to that of the pure system (assuming that the transition persists). Essentially, Harris criterion says that if the transition for the pure system is sufficiently smooth, a small or a moderate amount of disorder does not modify the transition: this is the case of *irrelevant disorder*. When the Harris criterion fails, one expects to be in a *relevant disorder* case, except possibly at the boundary between these two situations where the analysis is trickier (*marginal disorder*). These notions of relevant, irrelevant and marginal disorder are connected to the framework in which the theory has been developed, that is *renormalization* [12] and the idea behind Harris' approach is that disorder may be downsized or enhanced by the renormalization transformation, leading, after repeated application of the transformation, in the first case to the pure system fixed point and, in the second case, to a different one or to no fixed point at all [22, 28, 33].

One of the substantial obstacles to the mathematical exploration of the Harris criterion is that a good understanding of the critical properties of pure systems is limited to very special cases. But in the last twenty years this question has been addressed, first by

theoretical physicists (e.g. [23] and references therein) and then by mathematicians, for a simple model of one dimensional interface interacting with a substrate: the random walk (RW) disordered pinning model (see [27, 28]). For this model the interface is given by the graph of a random walk which takes random energy rewards when it touches a defect line. The random walk can be very general and the full class of RW pinning models is better apprehended if viewed in terms of *renewal pinning*: we refer the interested reader to the introductions of [28, 27]. The pure system has the remarkable quality of being what physicists call *solvable*, meaning that there exists an explicit expression for the free energy [26]. All the results which have been obtained confirm the validity of the Harris criterion and its interpretation for the RW pinning model [2, 3, 7, 16, 17, 21, 28, 30, 31, 35, 39].

A natural generalization of the RW pinning model is obtained by replacing the graph of the random walk by a random surface, and one of the first natural choices is the Lattice Gaussian Free Field (LGFF) – recently called also Discrete Gaussian Free Field – on a subset, for example an (hyper)-cube, of $\mathbb{Z}^d$, $d \geq 2$. While the pure model is not exactly solvable in that case, it has been studied and the nature of the phase transition is well known [11, 41]. However, until now very few attempts have been made to understand the quenched behavior of the system (see [19, 20]).

Our model has two parameters: the noise intensity $\beta \geq 0$ and the average pinning strength $h \in \mathbb{R}$. In this paper, we describe completely the characteristics of the phase transition in the case $d \geq 3$ and the results can be summed up as follows:

(1) We identify the disordered critical point $h_c = h_c(\beta)$. More precisely, with the choice of the parameters we make, which is the same as the one adopted for the RW pinning in the mathematical literature, the critical point of the disordered (i.e. quenched) model and the one of the annealed model coincide. However, the critical behaviors do not and this contrasts sharply with what happens for RW pinning where, except for the marginal disorder case for which the question is open, coincidence of critical points happens if and only if the critical behaviors coincide. We stress also that, with our choice, the annealed model coincides with the one in which we simply switch off the disorder by setting its intensity β to zero and this is what we call *pure* model.

(2) The free energy density, or just free energy for conciseness, is zero for $h \leq h_c(\beta)$ and it is positive for $h > h_c(\beta)$. We prove in full generality (in the choice of the disorder) that the free energy is $O\left((h - h_c(\beta))^2\right)$ as $h \searrow h_c(\beta)$, which implies that the first derivative of the free energy is continuous at $h_c(\beta)$: this is what is usually called a *second order* transition. The transition for the pure system instead is of first order, i.e. the first derivative of the free energy is discontinuous (it has a jump) at $h_c(\beta)$, hence, in the Harris' sense, disorder is relevant.

(3) When disorder is Gaussian we show that the behavior of the free energy at criticality is precisely quadratic and the critical exponent associated to the free energy is therefore precisely identified. Harris' theory yields no prediction of how the critical properties are modified when disorder is relevant. As a matter of fact capturing the critical exponent of transitions in disorder relevant cases appears to be a major challenge and the authors do not know of any rigorous results in this direction when the disorder is weakly correlated (for strongly correlated environment see [5, 6]). Even in the RW pinning models several contrasting conjectures have been set forth, but a certain consensus appears to emerge in favor of an infinite order transition, i.e. C^∞ regularity of the the free energy at the critical point (see the

review of the literature in [28, Section 5.3] to which one should add the recent contribution [22]).

We present also results for $d = 2$, but we are unable to show that disordered and pure critical points coincide and, as a consequence, we are unable to establish results on the critical behavior. Finally, we have also a quick look at the higher dimensional analog of the problem of a *copolymer near an interface between selective solvents*.

Note added in proof. Our two dimensional result has been substantially improved in the recent preprint [36], where it is proved that, for $d = 2$, the critical point for the disordered system coincides with the pure one, with yet a different critical behavior: near the critical point the free-energy is shown to grow slower than any power of $(h-h_c(\beta))$, which indicates a phase transition of infinite order.

2. Model and results

2.1. **The disordered model.** Given Λ be a finite subset of $\mathbb{Z}^d$, we let $\partial\Lambda$ denote the internal boundary of Λ, $\mathring{\Lambda}$ the set of interior points of Λ and and $\partial^-\Lambda$ the set of interior points that are in contact with the boundary.

$$\begin{aligned} \partial\Lambda &:= \{x \in \Lambda : \exists y \notin \Lambda,\ x \sim y\},\\ \mathring{\Lambda} &:= \Lambda \setminus \partial\Lambda,\\ \partial^-\Lambda &:= \{x \in \mathring{\Lambda} : \exists y \in \partial\Lambda,\ x \sim y\}. \end{aligned} \tag{2.1}$$

In general some of these sets could be empty, but throughout this work Λ is going to be a large (hyper-)cube.

Given $(\widehat{\phi}_x)_{x\in\mathbb{Z}^d}$ a real valued field, one defines $\mathbf{P}_\Lambda^{\widehat{\phi}}$ to be the law of the lattice Gaussian free field on Λ (denoted by $\phi = (\phi_x)_{x\in\Lambda}$) with boundary conditions $\widehat{\phi}$ on $\partial\Lambda$. Formally we set

$$\phi_x := \widehat{\phi}_x \quad \text{for every } x \in \partial\Lambda_N, \tag{2.2}$$

and consider $\mathbf{P}_\Lambda^{\widehat{\phi}}$ as a measure on $\mathbb{R}^{\mathring{\Lambda}}$ whose density is given by

$$\mathbf{P}_\Lambda^{\widehat{\phi}}(\mathrm{d}\phi) = \frac{1}{\mathcal{Z}_\Lambda^{\widehat{\phi}}} \exp\left(-\frac{1}{2}\sum_{\substack{(x,y)\in(\Lambda)^2\setminus(\partial\Lambda)^2\\ x\sim y}} \frac{(\phi_x-\phi_y)^2}{2}\right)\prod_{x\in\mathring{\Lambda}}\mathrm{d}\phi_x\,, \tag{2.3}$$

where $\prod_{x\in\mathring{\Lambda}}\mathrm{d}\phi_x$ denotes the Lebesgue measure on $\mathbb{R}^{\mathring{\Lambda}}$ and

$$\mathcal{Z}_\Lambda^{\widehat{\phi}} := \int_{\mathbb{R}^{\mathring{\Lambda}}} \exp\left(-\frac{1}{2}\sum_{\substack{(x,y)\in(\Lambda)^2\setminus(\partial\Lambda)^2\\ x\sim y}} \frac{(\phi_x-\phi_y)^2}{2}\right)\prod_{x\in\mathring{\Lambda}}\mathrm{d}\phi_x\,. \tag{2.4}$$

For the particular case $\widehat{\phi} \equiv u$ we write $\mathbf{P}_\Lambda^u$, and $\mathbf{P}_\Lambda$ when $u = 0$. One of the factors $1/2$ in the exponential is present to compensate for the fact that each edge is counted twice.

In what follows we consider mostly the case $\Lambda = \Lambda_N := \{0,\dots,N\}^d$, for some $N \in \mathbb{N}$. Note that in this case we have $\mathring{\Lambda}_N := \{1,\dots,N-1\}^d$. We also introduce the notation $\widetilde{\Lambda}_N := \{1,\dots,N\}^d$. We simply write $\mathcal{Z}_N$ and $\mathbf{P}_N$ for $\mathcal{Z}_{\Lambda_N}$ and $\mathbf{P}_{\Lambda_N}$.

Given $\omega = \{\omega_x\}_{x\in\mathbb{Z}^d}$ a family of IID square integrable centered random variables (of law $\mathbb{P}$), we set

$$\lambda(\beta) := \log \mathbb{E}[e^{\beta\omega_x}], \tag{2.5}$$

and assume that there exists $\overline{\beta} \in (0,\infty]$ such that

$$\max(\lambda(2\beta), \lambda(-\beta)) < \infty \text{ for every } \beta \in (0,\overline{\beta}). \tag{2.6}$$

Many of the arguments rely only on $\lambda(\beta) < \infty$: $\lambda(2\beta) < \infty$ is related to two replica arguments (lower bounds) and $\lambda(-\beta) < \infty$ is exploited when fractional moments estimates are performed (upper bounds) and a look at the proof of Proposition 3.1 suffices to see that this second requirement can be relaxed. Moreover, a part of the results are given for Gaussian ω and in that case $\overline{\beta} = \infty$. Note that (2.6) implies smoothness of $\lambda(\cdot)$ for $\beta \in (-\overline{\beta}, 2\overline{\beta})$ and around zero

$$\lambda(\beta) = \frac{\beta^2}{2} + O(\beta^3). \tag{2.7}$$

For $x \in \Lambda_N$ set $\delta_x := \mathbf{1}_{[-1,1]}(\phi(x))$. For $\beta > 0$ and $h \in \mathbb{R}$, we define a modified measure $\mathbf{P}_{N,h}^{\omega,\beta,\widehat{\phi}}$ via

$$\frac{\mathrm{d}\mathbf{P}_{N,h}^{\omega,\beta,\widehat{\phi}}}{\mathrm{d}\mathbf{P}_N^{\widehat{\phi}}} = \frac{1}{Z_{N,h}^{\beta,\omega,\widehat{\phi}}} \exp\left(\sum_{x\in\widetilde{\Lambda}_N} (\beta\omega_x - \lambda(\beta) + h)\delta_x\right), \tag{2.8}$$

where

$$Z_{N,h}^{\beta,\omega,\widehat{\phi}} := \mathbf{E}_N^{\widehat{\phi}}\left[\exp\left(\sum_{x\in\widetilde{\Lambda}_N} (\beta\omega_x - \lambda(\beta) + h)\delta_x\right)\right]. \tag{2.9}$$

Note that in the definition of $\mathbf{P}_{N,h}^{\omega,\beta,\widehat{\phi}}$, the sum $\left(\sum_{x\in\widetilde{\Lambda}_N}\right)$ can be replaced by either $\left(\sum_{x\in\Lambda_N}\right)$ or $\left(\sum_{x\in\mathring{\Lambda}_N}\right)$ as these changes affect only the partition function. We have chosen to sum over $\widetilde{\Lambda}_N$ for super-additivity reasons (see Proposition 4.2). The superscript $\widehat{\phi}$ is dropped when 0 boundary conditions are considered and replaced by u when $\widehat{\phi} \equiv u$.

2.2. **The pure model.** The natural homogeneous model associated to the disordered model $\mathbf{P}_{N,h}^{\omega,\beta,\widehat{\phi}}$ can be obtained by switching off the disorder: the pure model is therefore precisely $\mathbf{P}_{N,h}^{\omega,0,\widehat{\phi}}$ but the notation is heavy and a bit misleading because the measure does not depend on ω. Moreover our choice of the parametrization is such that the pure model coincides with the annealed model, that is with the model with partition function $\mathbb{E}\left[Z_{N,h}^{\beta,\omega,\widehat{\phi}}\right]$. For the pure model we use the notation $\mathbf{P}_{N,h}$ and we limit ourselves to the case $\widehat{\phi} \equiv 0$:

$$\frac{\mathrm{d}\mathbf{P}_{N,h}}{\mathrm{d}\mathbf{P}_N} = \frac{1}{Z_{N,h}} \exp\left(h \sum_{x\in\mathring{\Lambda}_N} \delta_x\right). \tag{2.10}$$

It is very easy to see – the proof is detailed just below – that this model has a transition at $h = 0$, in the sense that the free energy density

$$\mathrm{F}(h) = \lim_{N\to\infty} \frac{1}{N^d} \log Z_{N,h}, \tag{2.11}$$

satisfies

$$\mathrm{F}(h)\begin{cases}=0 & \text{for } h\le 0\,,\\ >0 & \text{for } h>0\,,\end{cases} \tag{2.12}$$

and therefore it is not analytic at $h=0$. Moreover, by standard convexity arguments $\mathrm{F}(h)$ is differentiable everywhere except, possibly, at countably many values of h. When it exists the derivative of $\mathrm{F}(h)$ is equal to the *asymptotic contact fraction* defined by

$$\lim_{N\to\infty}\frac{1}{N^d}\mathbf{E}_{N,h}\left[\sum_{x\in\widetilde{\Lambda}_N}\delta_x\right]. \tag{2.13}$$

It is obvious from (2.12) that the asymptotic contact fraction is 0 for $h<0$. Moreover, since $\mathrm{F}(\cdot)$ is convex, the asymptotic contact fraction is non decreasing and, again because of (2.12), it is positive for every $h>0$.

The existence of the limit (2.11) is standard: the argument can be recovered from the proof in Section 4.2 (it is a particular easy case). It is non negative because

$$Z_{N,h}\ \ge\ \mathbf{P}_N(\varphi_x>1 \text{ for every } x\in\mathring{\Lambda}_N)\,, \tag{2.14}$$

and it is not difficult to show that the logarithm of the latter expression is $o(N^d)$: this is a (rough) entropic repulsion type estimate and it is an easy consequence of the continuum symmetry of the interaction that is broken only at the boundary [38]. On the other hand $Z_{N,h}\le 1$ for $h\le 0$, and hence $\mathrm{F}(h)=0$ for $h\le 0$.

The fact that $\mathrm{F}(h)>0$ for every $h>0$ can be established in a number of elementary ways (see Section 2.3 for $d\ge 3$ and Remark 7.10 for $d=2$), but here we mention the more refined estimate (see [19, Fact 2.4]): for every $d=2,3,\dots$ there exists $c_d>0$ such that

$$\mathrm{F}(h)\overset{h\searrow 0}{\sim}\begin{cases}c_d\,h & \text{for } d\ge 3\,,\\ c_2\frac{h}{\sqrt{|\log h|}} & \text{for } d=2\,.\end{cases} \tag{2.15}$$

Therefore the transition is of first order for $d\ge 3$ and the contact fraction has a jump discontinuity. Note that the transition is of second order for $d=2$: the contact fraction is continuous at the transition, even if the continuity modulus vanishes (hence matching the behavior in the delocalized phase) only logarithmically.

2.3. **Some more details about the phase transition for $d\ge 3$.** The result in $d\ge 3$ is going to be particularly relevant for us and we want to stress that a rougher version of (2.15) is trivially established and that even the sharp statement isn't much harder. Note that

$$\frac{1}{N^d}\partial_h\log Z_{N,h}|_{h=0}\ =\ \frac{1}{N^d}\mathbf{E}_N\sum_x\delta_x\,. \tag{2.16}$$

Now notice that $\mathbf{P}_N$ is a centered Gaussian measure and the variance of ϕ_x under $\mathbf{P}_N$ is bounded uniformly by the variance of the infinite volume free field which we denote by σ_d^2 (see Section 2.9). Hence

$$\frac{1}{N^d}\partial_h\log Z_{N,h}|_{h=0}\ \ge\ P(\sigma_d\mathcal{N}\in[-1,1])\ =:\ C_d\,, \tag{2.17}$$

where $\mathcal{N}$ is a standard normal variable.

On the other hand the same derivative is bounded above by one and therefore (using convexity and the fact that $Z_{N,0} = 1$) for all N

$$C_d h \,\le\, \frac{1}{N^d} \log Z_{N,h} \,\le\, h\,, \tag{2.18}$$

and therefore for every $h \ge 0$

$$C_d h \,\le\, \mathrm{F}(h) \,\le\, h\,. \tag{2.19}$$

This establishes a rougher version of (2.15) for $d \ge 3$ (which is however a statement only for h small).

In fact, we have $c_d = C_d$. For this observe that if we go back to the partition function in (2.9), but setting $\beta = 0$, that is $Z^{\widehat{\phi}}_{N,h} := Z^{0,\omega,\widehat{\phi}}_{N,h}$, we readily check that for every N

$$\frac{1}{(2N)^d} \log \sup_{\widehat{\phi}} Z^{\widehat{\phi}}_{2N,h} \,\le\, \frac{1}{N^d} \log \sup_{\widehat{\phi}} Z^{\widehat{\phi}}_{N,h}\,, \tag{2.20}$$

from which one infers that $\mathrm{F}(h) \le \frac{1}{N^d} \log \sup_{\widehat{\phi}} Z^{\widehat{\phi}}_{N,h}$ for every N. Now remark that for every $h > 0$, we have

$$\partial_h \log Z^{\widehat{\phi}}_{N,h} \,=\, \mathbf{E}^{\widehat{\phi}}_{N,h} \left[\sum_{x \in \widetilde{\Lambda}_N} \delta_x \right] \,\le\, e^{N^d h} \mathbf{E}^{\widehat{\phi}}_{N,0} \left[\sum_{x \in \widetilde{\Lambda}_N} \delta_x \right] \le e^{N^d h} \mathbf{E}_N \left[\sum_{x \in \widetilde{\Lambda}_N} \delta_x \right]. \tag{2.21}$$

In the last step we have used that that $\sup_m P(\sigma\mathcal{N} + m \in [-1,1]) = P(\sigma\mathcal{N} \in [-1,1])$. Integrating the above inequalty on the interval $[0,h]$ we obtain

$$\frac{\mathrm{F}(h)}{h} \le \frac{1}{N^d} e^{N^d h} \mathbf{E}_N \left[\sum_{x \in \widetilde{\Lambda}_N} \delta_x \right]. \tag{2.22}$$

and hence that for every N,

$$\limsup_{h \searrow 0} \frac{\mathrm{F}(h)}{h} \le \frac{1}{N^d} \mathbf{E}_N \left[\sum_{x \in \widetilde{\Lambda}_N} \delta_x \right]. \tag{2.23}$$

Now using the fact that the variance of ϕ_x is close to σ_d^2 when the distance of x to the boundary is large (see Section 2.9), it is standard to check that

$$\lim_{N \to \infty} \frac{1}{N^d} \mathbf{E}_N \left[\sum_{x \in \widetilde{\Lambda}_N} \delta_x \right] = C_d, \tag{2.24}$$

which is sufficient to conclude.

The proof of (2.15) for $d = 2$ is substantially more involved and it is less related to our results because in any case for $d = 2$ we are unable to address the issue of the order of the transition when disorder is present. However, the reader can check that the above method gives, for $d = 2$, an upper-bound on $\mathrm{F}(h)$ of the right order of magnitude (that is $h(\log h)^{-1/2}$). See also Remark 7.10 for a proof of a lower-bound of the same order (which also implies that the transition is at $h = 0$ for $d = 2$ too).

Before moving toward the disordered case it is worth recalling that the phase transition we have just described is a localization transition and the localized LGFF is profoundly different from the LGFF since the continuum invariance of the latter is broken by the localizing potential. In particular, correlations decay exponentially with the distance for

the localized measure [9, 34, 11, 41], while the decay of correlations for the LGFF is power law (see Section 2.9). Moreover a directly related issue for an akin model is the one of wetting [14, 10, 41]: in this case, added to the pinning potential, the LGFF is constrained not to enter the lower half plane. This constraint generates a repulsion, but the transition is still at $h = 0$.

2.4. **Free energy and transition for the disordered model.** The existence of quenched free energy for the disordered model has been proved in [19, Theorem 2.1]. We recall the result here:

Proposition 2.1. *The free energy*

$$\mathrm{F}(\beta,h) := \lim_{N\to\infty} \frac{1}{N^d}\mathbb{E}\left[\log Z^{\beta,\omega}_{N,h}\right] \overset{\mathbb{P}(\mathrm{d}\omega)-a.s.}{=} \lim_{N\to\infty} \frac{1}{N^d}\log Z^{\beta,\omega}_{N,h}\,, \tag{2.25}$$

exists (and is self-averaging).

Note that $\mathrm{F}(0,h) = \mathrm{F}(h)$. Moreover it is easy to observe that $\mathrm{F}(\beta,h)$ is non-decreasing and convex in h and we have (cf. (2.13))

$$\partial_h \mathrm{F}(\beta,h) \,=\, \lim_{N\to\infty} \frac{1}{N^d}\mathbf{E}^{\beta,\omega}_{N,h}\left[\sum_{x\in\widetilde{\Lambda}_N}\delta_x\right]\,, \tag{2.26}$$

as soon as the left-hand side is defined.

Furthermore, from Jensen's inequality and convexity (we refer to the proof of [27, Proposition 5.1] for more details) we have

$$\mathrm{F}(0,h-\lambda(\beta)) \le \mathrm{F}(\beta,h) \le \mathrm{F}(0,h)\,, \tag{2.27}$$

which implies that $\mathrm{F}(\beta,h) \ge 0$ for every $h \in \mathbb{R}$. This elementary but important lower bound can be established in a direct fashion precisely in the same way as for the non disordered case (cf. (2.14)). But (2.27) guarantees also that $F(\beta,h) = 0$ for $h \le 0$ and that $F(\beta,h) > 0$ if $h > \lambda(\beta)$. Hence we have established the existence of a localization transition and the critical value

$$h_c(\beta) \,:=\, \inf\{h :\ \mathrm{F}(\beta,h) > 0\}\,, \tag{2.28}$$

satisfies

$$0 \,\le\, h_c(\beta) \,\le\, \lambda(\beta)\,. \tag{2.29}$$

2.5. **The main results.** The aim of this paper is to investigate if the inequalities (2.27) and (2.29) are sharp, and to compare the behavior of the model with respect to the pure, i.e. annealed, one.

The result we obtain are the following.

Theorem 2.2. *When $d \ge 3$, we have*

(i) *For all $\beta \in (0,\overline{\beta})$, $\overline{\beta}$ defined in (2.6), there exists a constant C (depending on β, d and the law of ω) such that for $h \in (0,1)$*

$$h^{66d} \,\le\, \mathrm{F}(\beta,h) \,\le\, Ch^2\,. \tag{2.30}$$

(ii) *When ω is Gaussian, for every $\beta > 0$ there exists a constant $c(\beta, d)$ such that for $h \in (0, 1)$*

$$c(\beta, d)h^2 \le \mathrm{F}(\beta, h) \le \frac{h^2}{\beta^2}. \tag{2.31}$$

Moreover one can find a constant $C(d)$ such that $c(\beta, d) \ge C(d)/\beta^2$ for every $\beta \in (0, 1]$.

A trivial consequence of the Theorem is that $h_c(\beta) = 0$ for all $\beta > 0$.

For $d = 2$ we are yet unable to decide whether there is a critical-point shift. However, in the Gaussian case, we are able to get a much better upper bound on $h_c(\beta)$ than the annealed one in (2.29).

Theorem 2.3. *When $d = 2$ and ω is Gaussian, every $\varepsilon > 0$ there exists $c_\varepsilon > 0$ such that for $\beta \in (0, 1)$*

$$0 \le h_c(\beta) \le c_\varepsilon \beta^{3-\varepsilon}. \tag{2.32}$$

2.6. **Behavior of the field under $\mathbf{P}^{\beta,\omega}_{N,h}$.** The main focus of this paper is the free energy, but let us briefly discuss the properties of the trajectories in the case $d \ge 3$. The basic remark is that the behavior of the free energy directly yields that the (asymptotic) contact fraction is zero for $h < h_c(\beta)$ and it is positive and increasing for $h > h_c(\beta)$: strictly speaking the existence of the contact fraction is guaranteed by convexity only out of a countable subset of $\{h : h > h_c(\beta)\}$ but one can extend the definition by taking limits (for example) from the right. For what concerns $h = h_c(\beta)$ the smoothing of the phase transition (due to the disorder) directly yields that the contact fraction is zero at the critical point ($h = 0$), and this is in strong contrast with what happens in the pure case.

These are all issues that are directly related to convexity and free energy estimates, but a number of sharper questions are very natural, notably the precise nature of the delocalized phase, that is when $\mathrm{F}(\beta, h) = 0$: is it true that the total number of contacts is $O(N^{d-1})$ and they are all close to the boundary? The analogous question even in the one dimensional set-up is not trivial even if by now rather sharp results are available [4]. Precise path description in the localized phase raises a number of issues too, in particular those treated, not always with complete success, in the one dimensional set-up (see [28, Ch. 8] and references therein), but the situation in the higher dimensional set-up may be richer and harder to tackle.

Nevertheless we want to observe that the results that we prove suggest the following typical behavior of ϕ for $h > 0$ small, so in the localized phase but close to criticality: ϕ typically stands at a large but finite (depending on h) distance of the interaction zone (the proof seems to indicate that $|\phi_x|$ should be of order $u(h) \sim \sqrt{2\sigma_d^2 \log(1/h)}$) since otherwise it should be difficult to avoid having a larger density of contact. The contacts with the interaction zone are typically produced by atypical peaks off the typical height (since we are talking of peaks of finite height, there is a positive but small density of them).

Here is a statement that goes in the direction of this conjecture, without getting close to the constant we conjecture (in fact $\sigma_d^2 > 1/(2d)$, as one can directly extract for example from the random walk representation that we recall in Section 2.9).

Proposition 2.4. *For every $\varepsilon > 0$ there exists $h_0 = h_0(\varepsilon)$ such that for all $h \in (0, h_0)$*

$$\lim_{N\to\infty} \mathbb{E}\left[\mathbf{P}_{N,h}^{\beta,\omega}\left(\sum_{x\in\widetilde{\Lambda}_N} \mathbf{1}_{\{|\phi_x|\le\sqrt{\frac{1}{4d}\log(1/h)}\}} \ge \varepsilon N^d\right)\right] = 0\,. \tag{2.33}$$

One is then tempted to conjecture that the interface chooses one side where to lie entirely, close to criticality, but we make no claim about this. Proposition 2.4 is proven in Appendix B.

2.7. **Co-membranes and selective solvents.** It is worthwhile stating the generalization of the results to a model in which the localization mechanism is somewhat different, but for which the technics can be adapted in a straightforward way. It is the analog of the model of a copolymer in the proximity of the interface between selective solvents, see [8, 15] and references therein. The model is defined by

$$\frac{\mathrm{d}\check{\mathbf{P}}_{N,h}^{\omega,\varrho}}{\mathrm{d}\mathbf{P}_N} \propto \exp\left(\varrho\sum_{x\in\widetilde{\Lambda}_N}(\omega_x + h)\,\mathrm{sign}\,(\phi_x)\right), \tag{2.34}$$

where without loss of generality we can assume both h and ϱ non negative and $\mathrm{sign}(0) = +1$. There is a rather natural way of understanding the model: imagine that the free field models a membrane made up by portions, say the unit box around x, that have an affinity for solvent A (if $\omega_x + h > 0$) or for solvent B (if $\omega_x + h < 0$). And that the solvent A fills in the upper half plane, and in the lower one there is solvent B. When h is positive there is an overall preference, since ω_x is centered, for solvent A, and the membrane in a average sense prefers to fluctuate in the upper half plane. However, there are membrane trajectories that, staying close to the A-B interface, can collect more energetic rewards and the localization transition is between a regime in which the membrane trajectories stay close to the A-B interface and a regime in which the membranes prefer to stay in the A solvent ($h \ge 0$, so if there is a globally preferred solvent, it has to be A).

A direct link with the pinning measure (2.8) can be made by observing that we can write

$$\frac{\mathrm{d}\check{\mathbf{P}}_{N,h}^{\omega,\varrho}}{\mathrm{d}\mathbf{P}_N} = \frac{1}{\check{Z}_{N,h}^{\omega,\varrho}}\exp\left(-2\varrho\sum_{x\in\widetilde{\Lambda}_N}(\omega_x + h)\Delta_x\right), \tag{2.35}$$

where $\Delta_x := (1 - \mathrm{sign}(\phi_x))/2$, that is Δ_x is the indicator function that ϕ_x is in the lower half plane. It is probably worth stressing that from (2.34) to (2.35) there is a non-trivial (but rather simple) change in energy (and free energy), but this change does not affect the measure, hence the model[1]. And in the form (2.35) the analogy with the pinning case is evident. In particular, the strict analog of Proposition 2.1 holds – the free energy in this case is denoted by $\check{\mathrm{f}}(\varrho, h)$ – and, precisely like for the pinning case, one sees that $\check{\mathrm{f}}(\varrho, h) \ge 0$. We then set $\check{h}_c(\varrho) := \inf\{h > 0 : \check{\mathrm{f}}(\varrho, h) = 0\}$.

Theorem 2.5. *For $d \ge 3$ and under the most general assumptions on the IID field ω (i.e. bounded exponential moments, centered and unit variance) we have that for every $\varrho \in (0, \overline{\beta}]$*

$$\check{h}_c(\varrho) = \frac{1}{2\varrho}\lambda(-2\varrho)\,. \tag{2.36}$$

[1]It is however straightforward to see that the annealed models associated to (2.34) and (2.35) are substantially different [15, 27]

Moreover (2.30)*, with* $\mathrm{F}(\beta,h)$ *replaced by* $\check{\mathrm{F}}(\varrho,h_c(\varrho)-h)$*, holds true and, if* ω *is Gaussian, then also* (2.31) *holds once the same replacement is made.*

For $d=2$ *and assuming* ω *to be Gaussian we have that* $\lim_{\varrho\searrow 0}\check{h}_c(\varrho)/\varrho=1$.

We have preferred to put the emphasis on the critical curve and on (2.36) because that is the same formula that appears for the copolymer, but as a strict upper bound, except for the very particular case of inter-arrival laws of the form $L(n)/n$, $L(\cdot)$ slowly varying, in which the upper bound (2.36) is achieved [8, 15, 27]. Moreover, a substantial emphasis for the copolymer has been put on the slope at the origin of $h_c(\varrho)$: in this case the slope is simply one.

Theorem 2.5 also provides a smoothing result for $d\geq 3$ and, which is most interesting, when the disorder is Gaussian we have again a model in which the disorder is relevant – in fact also for the co-membrane the pure model has a first order transition – and we can compute the critical exponent of the free energy.

We will not give a detailed proof of Theorem 2.5, because the arguments are really close to the ones for the pinning model and we limit ourselves to Remarks 6.2 and 7.9.

2.8. **Discussion of the results, sketch of proofs and structure of the paper.**

On the upper bound (and smoothing). The upper-bound in (2.30) and (2.31) is quite easy to prove and is valid in any dimension. Its proof can be read independently of the rest of the paper: it relies on the *disorder tilt and fractional moment bound* introduced in [29, 21]. However, here the implementation of the idea is remarkably straightforward: no coarse graining procedure is needed (see [28, Section 6] for a review of various coarse graining procedures). The reason why things here are simpler is that the method is not used to show that the free energy density is zero, like in the papers we have just mentioned, but simply to have a positive upper bound on it.

Note that, on its own, the inequality $\mathrm{F}(\beta,h)\leq Ch^2$ does not imply a rounding or smoothing of the free energy function. It does only if one can prove that $h_c(\beta)=0$, and this is precisely what we prove for the disordered LGFF pinning. Nevertheless such a bound recalls the smoothing inequality in [31], proven for RW pinning models. As a matter of fact the upper-bound in (2.30), that is Proposition 3.1, applies to RW pinning models too, but in this case $h_c(\beta)=0$ only if disorder is irrelevant and, even if the smoothing inequality in [31] and Proposition 3.1 are essentially the same result in this case, both of them end up having little importance because a direct application of Jensen inequality (annealed bound) and explicit computations lead to a better result (the exponent is larger than 2! [27]). Of course the smoothing inequality for RW pinning holds with respect to the correct critical point also when disorder is relevant and $h_c(\beta)\neq 0$. Generalizing the rare stretch approach in [31] to LGFF in order to establish a quadratic bound on the critical behavior does not appear to be straightforward and such a result would be in any case sensibly weaker than what we prove here.

On the lower bound ($d\geq 3$)*.* But how can we match the upper bound? That is, how can we show that $h_c(\beta)=0$ and find a lower bound on the free energy of quadratic type? We try to sketch here an answer to this question in a few steps:

(1) We show in Section 4 that one can *raise* the boundary conditions from 0 to an arbitrary u (that, conventionally, we choose positive). The reason why this is true is the continuum symmetry enjoyed by the LGFF: the term in the exponent in (2.3) is formally invariant when ϕ_x is mapped to ϕ_x+u for all x, if one chooses to

neglect the effect of the boundary, which is irrelevant for free energy computations (the reason for this is that the volume of the boundary is negligible with respect to that of the whole box).

We then choose $h > 0$ close to zero and a box of side-length N (N will be chosen as a function of h and it will be somewhat large, see below). We then choose $u = u(h)$ such that the probability that $\phi_x \sim \mathcal{N}(u, \sigma_d^2)$ is in $[-1, 1]$ (the *contact probability*) is ah, where a is a positive constant to be chosen. We have in particular $\lim_{h \searrow 0} u(h) = \infty$.

(2) We make now a bold proposal: we ask the reader to think of the variables ϕ_x's as independent. Of course they are not, but it is well known (see [18] for a quantitative result) that extrema and large excursions of the LGFF in $d \geq 3$ are close to what we would get forgetting the correlations and we are now rather far from the region where the pinning acts ($u = u(h)$ is large!). We stress that in the previous steps we have invoked the continuum symmetry of the model, that leads to power law correlations, so this step is a delicate one. If we accept this bold replacement we are now dealing with a model which is exactly solvable:

$$\widetilde{\mathrm{F}}_N(\beta, h) := \frac{1}{\Lambda_N} \mathbb{E} \log \mathbf{E} \left[\exp \left(\sum_{x \in \Lambda_N} (\beta \omega_x - \lambda(\beta) + h) \widetilde{\delta}_x \right) \right], \tag{2.37}$$

where $\widetilde{\delta}_x = \mathbf{1}_{[-1,1]}(\widetilde{\phi}_x)$ and the $\widetilde{\phi}_x$'s are IID $\mathcal{N}(u, \sigma_d^2)$ random variables. Recall that we have chosen u so that $\mathbf{E}[\widetilde{\delta}_x] = \mathbf{P}(\widetilde{\delta}_x = 1) = ah$ so that it is straightforward to see that

$$\widetilde{\mathrm{F}}_N(\beta, h) = \mathbb{E} \log \mathbf{E} \left[\exp \left((\beta \omega - \lambda(\beta) + h) \widetilde{\delta}_x \right) \right] = \mathbb{E} \log \left(1 + ah\xi \right), \tag{2.38}$$

where x is arbitrary (the variables are IID) and $\xi := \exp(\beta\omega - \lambda(\beta) + h) - 1 > -1$. If we assume that $\mathbb{E}[\exp(3\beta\omega)] < \infty$ (with some more effort one can generalize the argument to $\beta < \overline{\beta}$), for $h \searrow 0$ we have $\mathbb{E}[\xi] = e^h - 1 = h + O(h^2)$ and $\mathbb{E}[\xi^2] = c_\beta + O(h)$, with $c_\beta := e^{\lambda(2\beta) - 2\lambda(\beta)} - 1 > 0$ and $\mathbb{E}[\xi_+^3]$ is bounded. By putting all this together with the elementary bound

$$x^3 \mathbf{1}_{[-1/2,0]}(x) \overset{x \geq -1/2}{\leq} \log(1 + x) - x + \frac{1}{2} x^2 \overset{x > -1}{\leq} \frac{1}{3} x^3 \mathbf{1}_{[0,\infty)}(x), \tag{2.39}$$

one sees that

$$\widetilde{\mathrm{F}}_N(\beta, h) = ah^2 - c_\beta \frac{a^2 h^2}{2} + O(h^3), \tag{2.40}$$

and setting $a = 1/c_\beta$ yields the quadratic behavior in h we were looking for. Note that this gives a justification *a posteriori* for choosing $\mathbf{E}[\widetilde{\delta}_x]$ proportional to h: any other choice would give a smaller, if not negative, lower bound on the free energy.

(3) It appears that N can be chosen arbitrarily up to now (and this is quite troublesome!). However a closer look suggest that N has to be chosen large – at least like a power of $\frac{1}{h}$ – because boundary effects have to be taken care of. In fact, in order to deal with a super-additive model we do not choose boundary conditions equal to u, but boundary conditions that are a sampled from an infinite volume free field of mean u. Therefore the value of the field at the boundary (hence also close to it) can occasionally be also rather different from u and that the contact probability is ah – that we have used above – can be rather far from the truth. We need therefore to be able to neglect a fairly large portion of sites close to the

boundary in order to be sufficiently far so that an averaging effect – the mean on a LGFF is the solution of a Dirichlet problem for the discrete Laplacian – takes place. It is not too difficult to get convinced that one needs to take N to grow like a power of $\frac{1}{h}$: even if we imagine that we are able to make sufficiently sharp estimates for sites that are at a finite (large) distance from the boundary, hence gaining in the bulk a contribution in any case not larger (annealed bound!) than $h\mathbf{P}^u(\phi_0 \in [-1,1])N^d = ah^2N^d$ (we are assuming that $\mathbf{P}^u(\phi_x \in [-1,1])$ essentially does not feel the boundary), when one is on the boundary it is not evident how to argue that one does not get a negative contribution. Actually in (5.4), but this is taken up in a more informal fashion in Section 5.3 and notably in (5.36), it is argued that the boundary gives a contribution smaller, i.e. more negative, than a β-dependent negative constant times $\mathbf{P}^u(\phi_0 \in [-1,1])N^{d-1}$, which is hence of the order of hN^{d-1} and we have therefore to choose $N \gg h^{-1}$ to have a chance that the bulk prevails on the boundary term.

(4) At this point we get back with a last consideration on the *bold replacement* at step (2). The structure of the result we got using this replacement, that is (2.40), is quite clear: we have an energetic gain (the first term in the right-hand side) that is what we would get by Jensen's inequality (annealed bound) even without independence assumption. We have then a quadratic loss, that is the second herm in the right-hand side. So one needs to implement an efficient second moment method and to do this we resort to Gaussian interpolation techniques [32, 39], which limits our result to Gaussian disorder. Still, even exploiting the interpolation formula, the result is not straightforward because the *quadratic coupling term* grows too fast. So what we do is to apply the interpolation after having restricted the model to trajectories of the LGFF that have only a bounded number of contacts on suitably chosen intermediate scale boxes (for example, if the box has volume smaller than $\frac{1}{ah}$ then in average there will be less than one contact). We do not explain this procedure in detail here, but we just remark that the event that the number of contacts is suitably limited becomes improbable if the region in which this requirement is made is too large, but boxes of edge-length that is a power of $\frac{1}{h}$ turn out to be fine.

(5) All of this targets the quadratic behavior. We can be much rougher if we just target h to some (large) positive power, see the lower bound in (2.30). In this case, once N is chosen to grow like a power of $\frac{1}{h}$, we can choose $u(h)$ growing so that the contact probability is h to some power larger than one and we choose a power so large that the probability of having a contact in the whole box vanishes with h. Of course this way we will not get close to the quadratic behavior, but the boundary control, since the field at the boundary is very high, is easier and the second moment procedure is much less delicate because there are so little contacts in the underlying measure. The whole argument then goes through using less sophisticated techniques which are however helpful in understanding the argument leading to the quadratic lower bound.

Structure of the paper. The rest of the paper is organized as follows:

- We conclude Section 2 by mentioning classical results for the lattice free field which we will us throughout the paper.
- In Section 3, we use a very simple fractional moment method to show that $\mathrm{F}(\beta,h) \leq Ch^2$ (in any dimension).

- In Section 4 we show that the free energy is not sensitive to mild modifications of the boundary conditions, and use this information to get a lower bound on $\mathrm{F}(\beta, h)$ which is the free energy of a system of finite size (see (4.11) this is what we call a finite volume criterion). This criterion is used in all the next sections.
- Sections 5 and 6 are dedicated to the lower-bound on the free energy for $d \geq 3$: In Section 5, we establish the non-optimal lower bound in the non Gaussian case 2.30. In Section 6, we establish the sharp bound in the Gaussian case, which is the most technical result of the paper. We advise the reader to go through Section 5 before reading 6.
- Finally Section 7 is dedicated to the case $d = 2$ and the proof of Theorem 2.3: this last section adapts and uses tools of Section 4.

2.9. **A few fun facts about the free field.** Let $(X_t)_{t\geq 0}$ denote the continuous time simple random walk on $\mathbb{Z}^d$ (let P^x denote its law starting from x) whose transition rates are one along $\mathbb{Z}^d$-edges (see [37] for a complete reference on the subject). We let Δ denote the generator of X

$$\Delta f(x) := \sum_{y\sim x} (f(y) - f(x))\,. \tag{2.41}$$

Let us stress that the simple random walk in [37] is generated by $\Delta/2d$ – the walk jumps at rate one and chooses one of the $2d$ neighborhoods at random – but our choice (2.3) requires speeding up the walk by a factor $2d$ to have that the covariance of the ϕ field is the random walk Green function, see (2.42) and (2.43). For a set $B \subset \mathbb{Z}$ let τ_B be the first hitting time of the set B by X. Note that the Gaussian free field is a Gaussian process. Its covariance under measure $\mathbf{E}^{\widehat{\phi}}_{\Lambda}$ does not depend on the boundary conditions and is given by

$$G_\Lambda(x, y) := E^x \left[\int_0^{\tau_{\partial\Lambda}} \mathbf{1}_{\{X_t = y\}} \mathrm{d}t \right]. \tag{2.42}$$

Note that for $d \geq 3$, $G_\Lambda(x, y)$ is uniformly bounded (in Λ). This is the reason for which in this case, there exists a (unique) centered infinite volume version of the field whose covariance function is given by

$$G(x, y) := E^x \left[\int_0^{\infty} \mathbf{1}_{\{X_t = y\}} \mathrm{d}t \right]. \tag{2.43}$$

We have already introduced σ_d^2 for the variance the one dimensional marginals of the infinite volume field and therefore σ_d^2 coincides with $G(x, x)$, independent of x by translation invariance: more generally, $G(x, y) = G(0, y - x)$. We the standard bound

$$G(0, x) \;\leq\; \frac{\mathsf{c}_d}{(1 + |x|^{d-2})}\,, \tag{2.44}$$

where c_d is a constant that can be made explicit if one desires and $|\cdot|$ denotes the Euclidean norm. We call $\mathbf{P}$ resp. $\mathbf{P}^u$ the law of the field with covariance given by (2.43) with mean 0 resp. u. We use the notation $\widehat{\mathbf{P}}^u$ when the field is denoted by $\widehat{\phi}$ instead of ϕ.

In $d = 2$ the infinite volume field does not exist and we will make use of the following estimate [37, Prop. 6.3.2]

$$G_{\{x\in\mathbb{Z}^2 :\, |x|\leq N\}}(0, 0) \;=\; \frac{1}{2\pi} \log N + O(1)\,. \tag{2.45}$$

One can easily extract a number of results from (2.45) by the mean of comparison arguments (use $G_\Lambda(x,y) \geq G_{\Lambda'}(x,y)$ if $\Lambda' \subset \Lambda$), notably that for every $\varepsilon > 0$ we can find $d_\varepsilon > 0$ such that if $N > 2d_\varepsilon$ and if $x \in \Lambda_N$ is such that $\mathrm{dist}(x, \partial\Lambda_N) > d_\varepsilon$ then

$$G_{\Lambda_N}(x,x) \;\geq\; (1-\varepsilon)\frac{1}{2\pi}\log \mathrm{dist}(x,\partial\Lambda_N)\,. \tag{2.46}$$

where

$$\mathrm{dist}(x,A) \;:=\; \min_{y\in A}|y-x|\,. \tag{2.47}$$

For $m > 0$, the massive free field with mass m is defined by adding an harmonic confinement for each x:

$$\mathbf{P}_\Lambda^{\widehat{\phi},m}(\mathrm{d}\phi) = \frac{1}{\mathcal{Z}_\Lambda^{\widehat{\phi},m}}\exp\left(-\frac{1}{2}\sum_{\substack{(x,y)\in(\Lambda)^2\setminus(\partial\Lambda)^2\\ x\sim y}}\frac{(\phi_x-\phi_y)^2}{2}\right)\prod_{x\in\mathring{\Lambda}}\exp\left(-\frac{m^2}{2}\phi_x^2\right)\mathrm{d}\phi_x\,. \tag{2.48}$$

Its covariance fonction is given by the Green function of the operator $\Delta - m^2$, or

$$G_\Lambda^m(x,y) := E^x\left[\int_0^{\max\left(\tau_{\partial\Lambda},m^{-2}\mathcal{T}\right)}\mathbf{1}_{\{X_t=y\}}\mathrm{d}t\right]. \tag{2.49}$$

where X is a simple random walk and $\mathcal{T}$ is an exponential variable, of parameter one, independent of X. The infinite volume massive free field exists in any dimension $d \geq 1$ and the covariance is given by

$$G^m(x,y) := E^x\left[\int_0^{m^{-2}\mathcal{T}}\mathbf{1}_{\{X_t=y\}}\mathrm{d}t\right]. \tag{2.50}$$

It follows from the expression (2.3) that the Gaussian free field (and the massive one) satisfies the spatial Markov property. If $\Gamma \subset \Lambda$ (or $\subset \mathbb{Z}^d$ for the infinite volume case) the law of $\phi_{|\Gamma}$ knowing ϕ ouside of $\mathring{\Gamma}$ is given by $\mathbf{P}_\Gamma^{\phi_{|\partial\Gamma}}$ ($\mathbf{P}_\Gamma^{\phi_{|\partial\Gamma},m}$ in the massive case).

Moreover for $m \geq 0$ the mean of ϕ under $\mathbf{P}_\Lambda^{\widehat{\phi},m}$ is given by $H_\Lambda^{\widehat{\phi},m}$ the solution of

$$\begin{cases}(\Delta - m^2)H(x) = 0 & \text{if } x \in \mathring{\Lambda},\\ H(x) = \widehat{\phi}_x & \text{if } x \in \partial\Lambda\,.\end{cases} \tag{2.51}$$

We will exploit the random walk (or Poisson kernel) representation of this solution

$$H_\Lambda^{\widehat{\phi},m}(x) \;=\; E^x\left[\widehat{\phi}_{X_{\tau_{\partial\Lambda}}};\,\tau_{\partial\Lambda} < m^{-2}\mathcal{T}\right], \tag{2.52}$$

with $\tau_A = \inf\{t:\, X_t \in A\}$. If $m = 0$, we just drop it from the notation.

3. Fractional moment: upper-bound on the free energy

Proposition 3.1. *Choose $\beta < \overline{\beta}$ (cf. (2.6)). For every $c > 1$ there exists $h_0 > 0$ such that for $h \in (0, h_0]$*

$$\mathrm{F}(\beta,h) \;\leq\; \frac{ch^2}{\lambda'(\beta)^2}\,, \tag{3.1}$$

where $\lambda'(\cdot)$ is the derivative of $\lambda(\cdot)$ defined in (2.5). In the Gaussian case we can choose $c = 1$ and the result is valid for all h.

Proof. Let us fist observe that by Jensen's inequality

$$\mathbb{E}\left[\log Z^{\beta,\omega}_{N,h}\right] = 2\mathbb{E}\left[\log\sqrt{Z^{\beta,\omega}_{N,h}}\right] \le 2\log\mathbb{E}\left[\sqrt{Z^{\beta,\omega}_{N,h}}\right]. \tag{3.2}$$

This implies that

$$\mathrm{F}(\beta,h) = \limsup_{N\to\infty}\frac{2}{N^d}\log\mathbb{E}\left[\sqrt{Z^{\beta,\omega}_{N,h}}\right]. \tag{3.3}$$

We are going to estimate $\mathbb{E}\left[\sqrt{Z^{\beta,\omega}_{N,h}}\right]$ by making a change of measure on the environment. Let us start by making the preliminary observation that for every $\beta>0$ and $h\in(0,\lambda(\beta)+\lambda(-\beta))$ there exists a unique solution $\alpha(\beta,h)\in(0,\beta)$ to

$$\lambda(\beta-\alpha)-\lambda(-\alpha)-\lambda(\beta)+h \,=\, 0\,, \tag{3.4}$$

which follows by observing that the left-hand side is positive for $\alpha=0$, negative for $\alpha=\beta$ and decreasing in α in the interval $(0,\beta)$. Moreover, when ω is Gaussian $\alpha(\beta,h)=h/\beta$ and, in general, we have

$$\alpha(\beta,h) \overset{h\searrow 0}{\sim} \frac{h}{\lambda'(\beta)}\,. \tag{3.5}$$

Now let $\widetilde{\mathbb{P}}=\widetilde{\mathbb{P}}_N$ be a new measure on $\mathbb{R}^{\mathbb{Z}^d}$ (we are changing the law of the disorder keeping its independent character) defined by

$$\frac{\mathrm{d}\widetilde{\mathbb{P}}}{\mathrm{d}\mathbb{P}}(\omega) := \exp\left(\sum_{x\in\widetilde{\Lambda}_N}\left(-\alpha\omega_x-\lambda(-\alpha)\right)\right), \tag{3.6}$$

and, by the definition of α, cf. (3.4), one has for $x\in\widetilde{\Lambda}_N$

$$\widetilde{\mathbb{E}}\left[e^{\beta\omega_x-\lambda(\beta)+h}\right] \,=\, 1\,. \tag{3.7}$$

From the Cauchy-Schwartz inequality we obtain

$$\left(\mathbb{E}\left[\sqrt{Z^{\beta,\omega}_{N,h}}\right]\right)^2 \,\le\, \widetilde{\mathbb{E}}\left[Z^{\beta,\omega}_{N,h}\right]\mathbb{E}\left[\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\widetilde{\mathbb{P}}}\right], \tag{3.8}$$

and the first factor in the right hand side is equal to one because of (3.7). For the second one we have instead

$$\mathbb{E}\left[\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\widetilde{\mathbb{P}}}\right] = \exp\left(N^d\left(\lambda(\alpha)+\lambda(-\alpha)\right)\right). \tag{3.9}$$

Hence one can deduce from it that

$$\limsup_{N\to\infty}\frac{2}{N^d}\log\mathbb{E}\left[\sqrt{Z^{\beta,\omega}_{N,h}}\right] \le \lambda(\alpha)+\lambda(-\alpha) \overset{\alpha\searrow 0}{\sim} \alpha^2\,, \tag{3.10}$$

and by (3.3) and (3.5) the proof is complete. □

4. Elevated boundary conditions, stationary boundary conditions and finite volume criterion

In this section we manage to get a comparison between $\mathrm{F}(\beta,h)$ and the free energy per unit site of a finite system. To obtain this inequality, we need to change a bit the boundary conditions: instead of $\phi\equiv 0$ on the boundary of Λ_N, we choose to take ϕ to be distributed as an infinite volume LGFF (this requires $d\geq 3$). We will also play on taking *elevated* boundary conditions, in the sense that the infinite volume LGFF is centered at a non zero value u that then will be chosen suitably large (and will depend on h). For ease of exposition we first show that replacing 0 boundary conditions ($\mathbf{P}_N=\mathbf{P}_N^0$) with u boundary conditions ($\mathbf{P}_N^u$) does not change the free energy. We then show that the boundary conditions u can be replaced by a typical realization of the infinite volume LGFF of mean u.

In this section the only requirement on β is $\lambda(\beta)<\infty$.

4.1. Elevated boundary conditions.

Proposition 4.1. *For any $u\in\mathbb{R}$*

$$\lim_{N\to\infty}\frac{1}{N^d}\mathbb{E}\left[\log Z_{N,h}^{\beta,\omega,u}\right]=\mathrm{F}(\beta,h). \tag{4.1}$$

Proof. We are going to prove almost sure convergence to $\mathrm{F}(\beta,h)$ rather than convergence of the expectation: since $|N^{-d}\log Z_{N,h}^{\beta,\omega,u}|$ is bounded by $N^{-d}\sum_{x\in\widetilde{\Lambda}_N}|\beta\omega_x-\lambda(\beta)+h|$ and the latter forms a uniformly integrable sequence, almost sure convergence implies L^1 convergence.

We now start the proof of the a.s. convergence by observing that for all u

$$\log Z_{N,h}^{\beta,\omega,u}=-\mathbf{1}_{\{|u|>1\}}\left(\sum_{x\in\widetilde{\Lambda}_N\cap\partial\Lambda_N}(\beta\omega_x-\lambda(\beta)+h)\right)$$
$$+\log\mathbf{E}_N\left[\exp\left(\sum_{\substack{x\in\partial\Lambda_N,y\in\partial^-\Lambda_N\\ x\sim y}}\left(u\phi_y-\frac{u^2}{2}\right)\right)\exp\left(\sum_{x\in\widetilde{\Lambda}_N}(\beta\omega_x-\lambda(\beta)+h)\delta_x\right)\right], \tag{4.2}$$

where we used $\mathcal{Z}_N^u=\mathcal{Z}_N^0$ (recall (2.4)). The first term in the right-hand side, we call it $-b_{N,u}(\omega)$, yields a contribution which is $o(N^d)$ and thus has no influence on the limit. What one has to check is that the second term compares well with $\log Z_{N,h}^{\beta,\omega,0}$. For this we first remark that if we choose a $C>\mathbb{E}[|\beta\omega_x-\lambda(\beta)+h|]$ there exists $N_0(\omega)$, with $\mathbb{P}(N_0(\omega)<\infty)=1$, such that for all $N\geq N_0(\omega)$ we have

$$\sup_{\phi\in\mathbb{R}^{\widetilde{\Lambda}_N}}\left|\sum_{x\in\widetilde{\Lambda}_N}(\beta\omega_x-\lambda(\beta)+h)\delta_x\right|\leq CN^d. \tag{4.3}$$

Then one can check that under the probability law $\mathbf{P}_N$ (recall the definition (2.1)),

$$T(\phi):=\sum_{\substack{x\in\partial\Lambda_N,y\in\partial^-\Lambda_N\\ x\sim y}}\phi_y, \tag{4.4}$$

is a centered Gaussian. Its variance is equal to $2d(N-1)^{d-1}$ which is the number of edges linking $\partial\Lambda_N$ to $\partial^-\Lambda_N$ because

$$1 = \frac{\mathcal{Z}_N^u}{\mathcal{Z}_N^0} = \mathbf{E}_N\left[\exp\left(\sum_{\substack{x\in\partial\Lambda_N, y\in\partial^-\Lambda_N \\ x\sim y}} \left(u\phi_y - \frac{u^2}{2}\right)\right)\right]. \tag{4.5}$$

Hence there exists $c>0$ such that for N sufficiently large

$$\begin{aligned} \mathbf{P}_N\left(|T(\phi)| \ge N^{d-1/4}\right) &\le \exp(-cN^{d+1/2})\,, \\ \mathbf{E}_N\left[e^{uT(\phi)}\mathbf{1}_{\{|T(\phi)|\ge N^{d-1/4}\}}\right] &\le \exp(-cN^{d+1/2})\,, \end{aligned} \tag{4.6}$$

where, for the second inequality, how large N should be chosen may depend on u. We now set $A_N := \{|T(\phi)| \le N^{d-1/4}\}$. We observe (by (4.6) and (4.3) for the first line, and by the law of large number for the second one) that

$$\begin{aligned} &\lim_N N^{-d}\log Z_{N,h}^{\beta,\omega,u}(A_N^{\complement}) \;=\; -\infty\,, \\ &\liminf_{N\to\infty} N^{-d}\log Z_{N,h}^{\beta,\omega,u} \;\ge\; -\mathbb{E}|\beta\omega - \lambda(\beta) + h|. \end{aligned} \tag{4.7}$$

One can also easily show also that the inferior limit in the second line is non-negative, but here this bound suffices and we use it, coupled with the first inequality in (4.7), to establish the first of the inequalities, that holds for N sufficiently large, in the next formula

$$\frac{1}{2}Z_{N,h}^{\beta,\omega,u} \le Z_{N,h}^{\beta,\omega,u}(A_N) = Z_{N,h}^{\beta,\omega,u} - Z_{N,h}^{\beta,\omega,u}\left(A_N^{\complement}\right) \le Z_{N,h}^{\beta,\omega,u}, \tag{4.8}$$

and hence that $Z_{N,h}^{\beta,\omega,u}(A_N)$ and $Z_{N,h}^{\beta,\omega,u}$ are equivalent for computing the free energy. Moreover (recall that $b_{N,u}(\omega)$ is defined right after (4.2))

$$\begin{aligned} &e^{-uN^{d-1/4}} Z_{N,h}^{\beta,\omega,0}\left(A_N\right) = e^{-uN^{d-1/4}}\mathbf{E}_N\left[e^{\sum_{x\in\tilde\Lambda_N}(\beta\omega_x - \lambda(\beta)+h)\delta_x};\, A_N\right] \\ &\le \mathbf{E}_N\left[e^{uT(\phi)}e^{\sum_{x\in\tilde\Lambda_N}(\beta\omega_x-\lambda(\beta)+h)\delta_x};\, A_N\right] = e^{b_{N,u}(\omega)} Z_{N,h}^{\beta,\omega,u}\left(A_N\right) \le e^{uN^{d-1/4}} Z_{N,h}^{\beta,\omega,0}\,, \end{aligned} \tag{4.9}$$

which is enough to conclude since the result for $Z_{N,h}^{\beta,\omega} = Z_{N,h}^{\beta,\omega,0}$ is known (cf. Proposition 2.1). □

4.2. **Stationary boundary conditions.** When $d \ge 3$, (recall Section 2.9) $\mathbf{P}^u$ is the law of the infinite volume free field $(\widehat{\phi}_x)_{x\in\mathbb{Z}^d}$ with mean u. We have seen that we can approach the free energy by considering the size N approximation of the free energy $\mathbb{E}[\log Z_{N,h}^{\beta,\omega,u}]/N^d$ instead of the original one $\mathbb{E}[\log Z_{N,h}^{\beta,\omega}]/N^d$. Now we want to make the further step of replacing u at the boundary by a typical configuration of the LGFF with mean u. We do this to recover a sharp super-additive property that in turn guarantees that, for every N, the new size N approximation bounds is a lower bound for the free energy.

Proposition 4.2. *For any value of u one has*

$$\lim_{N\to\infty}\frac{1}{N^d}\mathbb{E}\widehat{\mathbf{E}}^u\left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi}}\right] \;=\; \mathrm{F}(\beta,h)\,. \tag{4.10}$$

Moreover for any u and N one has

$$\frac{1}{N^d}\mathbb{E}\widehat{\mathbf{E}}^u\left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi}}\right] \le \mathrm{F}(\beta,h)\,. \tag{4.11}$$

The result (4.10) is easy to believe because replacing u by a sequence of Gaussian variables, of mean u and variance σ_d, in the boundary conditions does not look a very drastic change: we are in the same framework of ideas as of Proposition 4.1. However, because of the random nature of the boundary values makes the proof more technical. The second result, that is (4.11), just follows from the the Markov property of the LGFF and Jensen's inequality.

Proof. As for Proposition 4.1, (4.10) follows if we can show that

$$\lim_{N\to\infty}\frac{1}{N^d}\widehat{\mathbf{E}}^u\left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi}}\right] = \mathrm{F}(\beta,h), \quad \mathbb{P}\text{ a.s.} \tag{4.12}$$

On the other hand, precisely by the same bound used at the beginning of the proof of Proposition 4.1 we see that also $N^{-d}\log Z_{N,h}^{\beta,\omega,\widehat{\phi}}$ forms a uniformly integrable sequence (this time the measure is $\mathbb{P}\otimes\widehat{\mathbf{P}}^u$). Therefore it suffices to show that

$$\lim_{N\to\infty}\frac{1}{N^d}\log Z_{N,h}^{\beta,\omega,\widehat{\phi}} = \mathrm{F}(\beta,h), \quad \mathbb{P}\otimes\widehat{\mathbf{P}}^u\text{ a.s.} \tag{4.13}$$

For this we first note that

$$\begin{aligned}\log Z_{N,h}^{\beta,\omega,\widehat{\phi}} = &-\left(\sum_{x\in\widetilde{\Lambda}_N\cap\partial\Lambda_N}(\beta\omega_x-\lambda(\beta)+h)\mathbf{1}_{\widehat{\phi}_x\notin[-1,1]}\right)+\log\left(\mathcal{Z}_N^0/\mathcal{Z}_N^{\widehat{\phi}}\right)\\ &+\log\mathbf{E}_N\left[\exp\left(\sum_{\substack{x\in\partial\Lambda_N,\,y\in\partial^-\Lambda_N\\ x\sim y}}\left(\widehat{\phi}_x\phi_y-\frac{\widehat{\phi}_x^2}{2}\right)\right)\exp\left(\sum_{x\in\widetilde{\Lambda}_N}(\beta\omega_x-\lambda(\beta)+h)\delta_x\right)\right].\end{aligned} \tag{4.14}$$

The right-hand side is of the form $T_1+T_2+T_3$: for T_1 we observe that

$$|T_1| \le \sum_{x\in\widetilde{\Lambda}_N\cap\partial\Lambda_N}|\beta\omega_x-\lambda(\beta)+h| = O(N^{d-1})\,, \tag{4.15}$$

$\mathbb{P}\otimes\widehat{\mathbf{P}}^u$-a.s. and thus we can neglect T_1.

Let us now examine T_3: first of all the term $\frac{1}{2}\sum_{\dots}\widehat{\phi}_x^2$ is a constant with respect to $\mathbf{P}_N(\mathrm{d}\phi)$ and drops out of the expectation and one can easily show that it yields $\widehat{P}^u$-a.s. an additive contribution to T_3 of order $O(N^{d-1}\log N)$ and hence plays no role in the limit. Let us then control the influence of the term in the exponential. Set

$$T(\widehat{\phi},\phi) := \sum_{\substack{x\in\partial\Lambda_N,\,y\in\partial^-\Lambda_N\\ x\sim y}}\widehat{\phi}_x\phi_y\,, \tag{4.16}$$

Let us call $M_N = M_N(\widehat{\phi})$ the maximal value of $|\widehat{\phi}_x|$ in $\partial\Lambda_N$ (Note that M_N is $O(\sqrt{\log N})$ $\widehat{\mathbf{P}}^u$-a.s.). Since the correlations are positive, the variance of $T(\widehat{\phi},\phi)$ under $\mathbf{P}_N$ is smaller

than that of $M_N(\widehat{\phi})T(\phi)$ (recall (4.4)). Hence similarly to (4.6) one obtains that there exists $c>0$ and N_0 (not depending on $\widehat{\phi}$) such that for $N\ge N_0$ we have

$$\begin{aligned}&\mathbf{P}_N\left(\left|T(\widehat{\phi},\phi)\right|\ge N^{d-1/4}M_N\right)\le\exp(-cN^{d+1/2})\,,\\&\mathbf{E}_N\left[e^{T(\widehat{\phi},\phi)}\mathbf{1}_{\{|T(\widehat{\phi},\phi)|\ge N^{d-1/4}M_N\}}\right]\le\exp(-cN^{d+1/2})\,.\end{aligned}\tag{4.17}$$

This together with (4.3) guarantees that if we set

$$A_N:=\left\{\left|T(\widehat{\phi},\phi)\right|\ge N^{d-1/4}M_N\right\}\,,\tag{4.18}$$

then, like for for Lemma 4.1, we readily see that $\mathbb{P}\otimes\widehat{\mathbf{P}}^u-a.s.$

$$\begin{aligned}&\lim_{N\to\infty}N^{-d}\log Z^{\beta,\omega,\widehat{\phi}}_{N,h}(A_N)\;=\;-\infty\,,\\&\liminf_{N\to\infty}N^{-d}\log Z^{\beta,\omega,\widehat{\phi}}_{N,h}(A_N)\;\ge\;\mathbb{E}|\beta\omega-\lambda(\beta)+h|\,,\end{aligned}\tag{4.19}$$

and therefore we obtain that there exists a random variable N_0, with $\mathbb{P}\otimes\widehat{\mathbf{P}}^u(N_0<\infty)=1$ such that for $N\ge N_0$

$$\frac{1}{2}Z^{\beta,\omega,\widehat{\phi}}_{N,h}\le Z^{\beta,\omega,\widehat{\phi}}_{N,h}(A_N)\le Z^{\beta,\omega,\widehat{\phi}}_{N,h}\,,\tag{4.20}$$

and an analogous statement for $Z^{\beta,\omega,0}_{N,h}$. Then one concludes similarly to what we have done for Lemma 4.1: we have

$$\begin{aligned}Z^{\beta,\omega,0}_{N,h}(A_N)e^{-M_NN^{d-1/4}}\le\mathbf{E}_N\left[e^{T(\widehat{\phi},\phi)}e^{\sum_{x\in\widetilde{\Lambda}_N}(\beta\omega_x-\lambda(\beta)+h)\delta_x};\,A_N\right]\\\le e^{M_NN^{d-1/4}}Z^{\beta,\omega,0}_{N,h}\,,\end{aligned}\tag{4.21}$$

therefore $\lim_N N^{-d}T_3=\lim_N N^{-d}\log Z^{\beta,\omega,0}_{N,h}$, $\mathbb{P}\otimes\widehat{\mathbf{P}}^u$-a.s., and the latter is just $\mathrm{F}(\beta,h)$. Similarly (and even in a slightly easier way) one shows that

$$|\log\left(\mathcal{Z}^0_N/\mathcal{Z}^{\widehat{\phi}}_N\right)|\le M_NN^{d-1/4}\,,\tag{4.22}$$

and therefore T_2 is negligible and the proof of (4.13) (hence (4.10)) is complete.

To prove (4.11) it is sufficient to show that (see (4.29))

$$\frac{1}{(2N)^d}\mathbb{E}\widehat{\mathbf{E}}^u\left[\log Z^{\beta,\omega,\widehat{\phi}}_{2N,h}\right]\ge\frac{1}{N^d}\mathbb{E}\widehat{\mathbf{E}}^u\left[\log Z^{\beta,\omega,\widehat{\phi}}_{N,h}\right].\tag{4.23}$$

Let us divide the box Λ_{2N} into 2^d boxes, Λ^i_N, $i=1,\dots,2^d$. Set

$$\begin{aligned}\Lambda^i_N&:=\Lambda_N+(\alpha_1(i),\dots,\alpha_d(i))N\,,\\\widetilde{\Lambda}^i_N&:=\widetilde{\Lambda}_N+(\alpha_1(i),\dots,\alpha_d(i))N\,,\end{aligned}\tag{4.24}$$

where $\alpha_j(i)\in\{0,1\}$ is the j-th digit of the dyadic development of $i-1$. Let $\mathbf{P}^{\widehat{\phi},i}_N$ be the law of the free field on Λ^i_N with boundary conditions $\widehat{\phi}$ and set

$$Z^{\beta,\omega,\widehat{\phi},i}_{N,h}:=\mathbf{E}^{\widehat{\phi},i}_N\left[\exp\left(\sum_{x\in\widetilde{\Lambda}^i_N}(\beta\omega_x-\lambda(\beta)+h)\delta_x\right)\right].\tag{4.25}$$

We define

$$\Gamma_N := \left(\bigcup_{i=1}^{2^d d} \partial \Lambda_N^i \right) \setminus \partial \Lambda_{2N}. \tag{4.26}$$

Now we notice by that if one conditions on the realization on ϕ in Γ_N, the partition functions of the system of size $2N$ factors into 2^d partition functions of systemes of size N, whose boundary conditions are determined by $\widehat{\phi}$ and $\phi|_{\Gamma_N}$.

$$\begin{aligned} &\mathbf{E}_{2N}^{\widehat{\phi}} \left[\exp\left(\sum_{x \in \widetilde{\Lambda}_{2N}} (\beta\omega_x - \lambda(\beta) + h)\delta_x \right) \;\middle|\; \phi|_{\Gamma_N} \right] \\ &\quad = \prod_{i=1}^{2^d} \mathbf{E}_{2N}^{\widehat{\phi}} \left[\exp\left(\sum_{x \in \widetilde{\Lambda}_N^i} (\beta\omega_x - \lambda(\beta) + h)\delta_x \right) \;\middle|\; \phi|_{\Gamma_N} \right] =: \prod_{i=1}^{2^d} \widetilde{Z}^i(\widehat{\phi}, \phi|_{\Gamma_N}, \omega). \end{aligned} \tag{4.27}$$

Note that by the spatial Markov property for the infinite volume field, each $\widetilde{Z}^i(\widehat{\phi}, \phi|_{\Gamma_N}, \omega)$ has the same distribution as $Z_{N,h}^{\beta,\omega,\widehat{\phi}}$ (if $\widehat{\phi}$ and $\phi|_{\Gamma_N}$ have distribution $\widehat{\mathbf{E}}^u$ and $\mathbf{E}_{2N}^{\widehat{\phi}}$ respectively and the ω are IID). By Jensen's inequality for $\mathbf{E}_{2N}^{\widehat{\phi}}[\cdot \mid \phi|_{\Gamma_N}]$ we have

$$\mathbb{E}\widehat{\mathbf{E}}^u \left[\log Z_{2N,h}^{\beta,\omega,\widehat{\phi}} \right] \geq \sum_{i=1}^{2d} \mathbb{E}\widehat{\mathbf{E}}^u \mathbf{E}_{2N}^{\widehat{\phi}} \left[\log \widetilde{Z}^i(\widehat{\phi}, \phi|_{\Gamma_N}\ ) \right] = 2^d \mathbb{E}\widehat{\mathbf{E}}^u \left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi}} \right]. \tag{4.28}$$

Iterating this inequality we obtains that

$$\mathrm{F}(\beta, h) = \lim_{k \to \infty} \frac{1}{2^{dk}} \frac{1}{N^d} \mathbb{E}\widehat{\mathbf{E}}^u \left[\log Z_{2^k N,h}^{\beta,\omega,\widehat{\phi}} \right] \geq \frac{1}{N^d} \mathbb{E}\widehat{\mathbf{E}}^u \left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi}} \right]. \tag{4.29}$$

□

5. A lower bound on the free energy

In this section we prove the lower bound in part (i) of Theorem 2.2. The statement is:

Proposition 5.1. *For $d \geq 3$, for any $\beta \in (0, \overline{\beta})$, there exists a constant $h_0 > 0$ (which depends on the dimension and on the law of ω) such that for any $h \in (0, h_0)$*

$$\mathrm{F}(\beta, h) \;\geq\; h^{66d}. \tag{5.1}$$

Remark 5.2. *While the constant 66 is quite arbitrary and is the consequence of some rough approximations made in the proof, there is a more serious reason why our bound gets worse when the dimension increases. This is due to boundary effects which are more important in high dimension (cf. the isoperimetric inequality). See Section 5.3 for more on this.*

We assume in this section that β is a fixed positive number and h is close to zero. Let us set (recall from Section 2.9 that σ_d is the standard deviation of the infinite volume free field)

$$u := 8\sigma_d \sqrt{d \log N} \quad \text{and} \quad N = h^{-2}, \tag{5.2}$$

where, without true loss of generality, we are assuming h^{-2} to be integer. We define the event

$$\mathcal{E}_u := \left\{\phi \in \mathbb{R}^{\mathbb{Z}^d} : \phi_x > u/2 \text{ for } x \in \partial\Lambda_N\right\}. \tag{5.3}$$

The set $\mathcal{E}_u$ plays the role of the set of good boundary conditions. What we are going to show is that $\mathcal{E}_u^c$ has a very small probability and use it to bound its contribution to the partition function.

Proposition 5.1 follows from the following two lemmas. The first takes care of the case of bad boundary conditions:

Lemma 5.3. *For every $\beta > 0$ such that $\lambda(\beta) < \infty$, there exists h_0 such that for every $h \in (0, h_0)$*

$$\mathbb{E}\widehat{\mathbf{E}}^u\left[\log\left(Z_{N,h}^{\beta,\omega,\widehat{\phi}}\right)\mathbf{1}_{\mathcal{E}_u^c}\right] \geq -C\lambda(\beta)N^{d-1}\mathbf{E}^u\left[\delta_0\right], \tag{5.4}$$

where $C > 0$ is a constant that depends only on the dimension.

The second lemma gives a lower bound on (a suitable expectation of) $\log Z_{N,h}^{\beta,\omega,\widehat{\phi}}$ for good boundary conditions and it is obtained by considering only the contribution of the realizations of ϕ which have at most one contact in the box $\widetilde{\Lambda}_N$:

Lemma 5.4. *For every $\beta \in (0, \overline{\beta})$, there exists h_0 such that for every $h \in (0, h_0)$*

$$\mathbb{E}\widehat{\mathbf{E}}^u\left[\log\left(Z_{N,h}^{\beta,\omega,\widehat{\phi}}\right)\mathbf{1}_{\mathcal{E}_u}\right] \geq \frac{h}{2}N^d\mathbf{E}^u\left[\delta_0\right]. \tag{5.5}$$

Proof of Proposition 5.1. From Lemma 5.3, Lemma 5.4 and the choice $N = h^{-2}$ we have for h small

$$\frac{1}{N^d}\mathbb{E}\widehat{\mathbf{E}}^u\left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi}}\right] \geq \left(\frac{h}{2} - C\frac{\lambda(\beta)}{N}\right)\mathbf{E}^u\left[\delta_0\right] \geq \frac{h}{4}\mathbf{E}^u\left[\delta_0\right]. \tag{5.6}$$

Since for u sufficiently large

$$\begin{aligned}\mathbf{E}^u\left[\delta_0\right] &= \frac{1}{\sqrt{2\pi\sigma_d^2}}\int_{u-1}^{u+1}\exp\left(-\frac{z^2}{2\sigma_d^2}\right)\mathrm{d}z \geq \frac{1}{2\sqrt{2\pi\sigma_d^2}}\exp\left(-\frac{\left(u-\frac{1}{2}\right)^2}{2\sigma_d^2}\right)\\ &\geq \exp\left(-\frac{u^2}{2\sigma_d^2}\right) = N^{-32d} = h^{64d}, \end{aligned}\tag{5.7}$$

we obtain as a consequence of Proposition 4.2, and (5.6)

$$\mathrm{F}(\beta, h) \geq \frac{h}{4}\mathbf{E}^u\left[\delta_0\right] \geq h^{65d}/4, \tag{5.8}$$

provided h is small enough. □

5.1. **Proof of Lemma 5.3.** By Jensen's inequality one has

$$\mathbb{E}\widehat{\mathbf{E}}^u\left[\log\left(Z_{N,h}^{\beta,\omega,\widehat{\phi}}\right)\mathbf{1}_{\mathcal{E}_u^c}\right] \geq (-\lambda(\beta)+h)\mathbf{E}^u\left[\sum_{x\in\widetilde{\Lambda}_N}\delta_x\mathbf{1}_{\mathcal{E}_u^c}\right], \tag{5.9}$$

and therefore it suffices to show that

$$\mathbf{E}^u\left[\sum_{x\in\widetilde{\Lambda}_N}\delta_x\mathbf{1}_{\mathcal{E}_u^c}\right] \leq CN^{d-1}\mathbf{E}^u\left[\delta_0\right]. \tag{5.10}$$

For every constant $c \geq 1$ we have

$$\begin{aligned}\mathbf{E}^u\left[\sum_{x\in\widetilde{\Lambda}_N}\delta_x\mathbf{1}_{\mathcal{E}_u^c}\right] &\leq \sum_{\substack{x\in\widetilde{\Lambda}_N\\ y\in\partial\Lambda_N}}\mathbf{E}^u\left[\delta_x\mathbf{1}_{\{\phi_y\leq u/2\}}\right]\\ &= \sum_{\substack{x\in\widetilde{\Lambda}_N, y\in\partial\Lambda_N\\ |x-y|\leq c}}\mathbf{E}^u\left[\delta_x\mathbf{1}_{\{\phi_y\leq u/2\}}\right] + \sum_{\substack{x\in\widetilde{\Lambda}_N, y\in\partial\Lambda_N\\ |x-y|> c}}\mathbf{E}^u\left[\delta_x\mathbf{1}_{\{\phi_y\leq u/2\}}\right] \qquad (5.11)\\ &\leq 2dcN^{d-1}\mathbf{E}^u\left[\delta_0\right] + \sum_{\substack{x\in\widetilde{\Lambda}_N, y\in\partial\Lambda_N\\ |x-y|> c}}\mathbf{E}^u\left[\delta_x\mathbf{1}_{\{\phi_y\leq u/2\}}\right],\end{aligned}$$

where in the first step we have used the union bound and in the third we have replaced, in the obvious way, the expectation in the first sum with an upper bound independent of x and y and we have then estimated the cardinality of the set over which the sum is performed.

Now we claim that for c sufficiently large, how large depends only on the dimension d, we have

$$\mathbf{E}^u\left[\delta_x\mathbf{1}_{\{\phi_y\leq u/2\}}\right] \leq N^{-2d}\mathbf{E}^u\left[\delta_0\right], \qquad (5.12)$$

for every $x, y \in \mathbb{Z}^d$ such that $|x-y| > c$. By putting (5.11) and (5.12) together we obtain

$$\mathbf{E}^u\left[\sum_{x\in\widetilde{\Lambda}_N}\delta_x\mathbf{1}_{\mathcal{E}_u^c}\right] \leq \left(2dcN^{d-1} + 2dN^{2d-1}N^{-2d}\right)\mathbf{E}^u\left[\delta_0\right] \leq 4dcN^{d-1}\mathbf{E}^u\left[\delta_0\right]. \qquad (5.13)$$

Therefore to complete the proof Lemma 5.3 it suffices to establish (5.12) and this is what we do now.

We set $x = 0$ for notational simplicity and we observe that

$$\mathbf{P}^u\left(\phi_y \leq u/2 \mid \delta_0 = 1\right) \leq \max_{z\in[u+1,u-1]}\mathbf{P}^0\left(\phi_y \geq u/2 \mid \phi_0 = z\right). \qquad (5.14)$$

Under $\mathbf{P}^0\left(\cdot \mid \phi_0 = z\right)$, ϕ_y is a Gaussian random variable of mean equal to $zG(0,y)/\sigma_d^2$ and variance is equal to $G(0,0) - G(0,y) \leq \sigma_d^2$. If c is chosen appropriately we have

$$\frac{zG(0,y)}{\sigma_d^2} \leq \frac{u}{4} \qquad \text{for every } y \text{ such that } |y| \geq c. \qquad (5.15)$$

More explicitly, for $u \geq 2$, it suffices to have $G(0,y) \leq \frac{1}{6}\sigma_d^2 = \frac{1}{6}G(0,0)$. We can then apply standard Gaussian bounds to obtain

$$\mathbf{P}^0\left(\phi_y \geq u/2 \mid \phi_0 = z\right) \leq P\left(\mathcal{N} \geq u/(4\sigma_d)\right) \leq e^{-\frac{u^2}{32\sigma_d^2}} = N^{-2d}, \qquad (5.16)$$

and the proof of (5.12), hence of Lemma 5.3, is complete. □

5.2. **Proof of Lemma 5.4.** As a first step, we are going to prove that

Lemma 5.5. *For any $\widehat{\phi} \in \mathcal{E}_u$ and $x, y \in \mathring{\Lambda}_N$, $x \neq y$, we have for every N larger than a constant that depends only on d*

$$\mathbf{E}_N^{\widehat{\phi}}[\delta_x] \leq N^{-2d} \qquad \textit{and} \qquad \mathbf{E}_N^{\widehat{\phi}}[\delta_y \mid \delta_x = 1] \leq N^{-2d}. \qquad (5.17)$$

Proof. From the maximum principle for the discrete harmonic equation (2.51) we have

$$\mathbf{E}_N^{\widehat{\phi}}(\phi_x) \,=\, H_{\Lambda_N}^{\widehat{\phi}}(x) \,\geq\, u/2\,. \tag{5.18}$$

The variance of ϕ_x is $G_N(x,x) \leq \sigma_d^2$. Hence we have

$$\mathbf{E}_N^{\widehat{\phi}}[\delta_x] \leq \mathbf{P}_N^{\widehat{\phi}}\left(\phi_x \leq 1\right) \,\leq\, P\left(\mathcal{N} \geq (u/2-1)/\sigma_d\right) \,\leq\, e^{-\frac{(u-2)^2}{8\sigma_d^2}} \,\leq\, N^{-2d}. \tag{5.19}$$

The second inequality is proved in the same manner: conditioned to the value of ϕ_y (which we set equal to some arbitrary $z \in [-1,1]$) we want to estimate the variance and expectation of ϕ_y. By monotonicity of the solution of the harmonic equation (2.51), we may as well restrict to the case $\widehat{\phi} \equiv u/2$. We notice here that, as the escape probability of the simple random walk in $\mathbb{Z}^d$ $d \geq 3$ is always larger than $3/5$ (see [25, Section 5.9]) we have for any $N > 0$ and any $y \neq x$ in $\mathring{\Lambda}_N$,

$$\frac{G_N(x,y)}{G_N(x,x)} \,=\, P^y\left(\tau_{\Lambda_N} < \tau_x\right) \,\leq\, \frac{2}{5}. \tag{5.20}$$

As a consequence we can write and bound the mean of ϕ_y conditioned to $\phi_x = z$ by

$$\mathbf{E}_N^{u/2}\left[\phi_y \middle|\, \phi_x = z\right] \,=\, z\frac{G_N(x,y)}{G_N(x,x)} + \frac{u}{2}\left(1 - \frac{G_N(x,y)}{G_N(x,x)}\right) \,\geq\, \frac{u}{4} + 1\,. \tag{5.21}$$

Therefore

$$\mathbf{E}_N^{u/2}\left[\delta_y \mid \phi_x = z\right] \leq \mathbf{E}_N^{u/2}\left[\phi_y \leq 1 \mid \phi_x = z\right] \,\leq P\left[\mathcal{N} \geq u/(4\sigma_d)\right] \leq e^{-\frac{u^2}{32\sigma^2}} = N^{-2d}\,, \tag{5.22}$$

which is enough to conclude. □

We now go back to the proof of Lemma 5.4: till the end of the proof we will assume $\widehat{\phi} \in \mathcal{E}^u$. Set

$$\xi(x) := \exp(\beta\omega_x - \lambda(\beta) + h) - 1. \tag{5.23}$$

Let $\mathcal{A}_0$ be the event that the field ϕ has no contact with the defect band and $\mathcal{A}_1(x)$ the event that it has only one contact at x, $\mathcal{A}_1$ the event that there is a unique contact and $\mathcal{A}_2$ the event that there are two contacts or more.

$$\begin{aligned}
\mathcal{A}_0 &:= \left\{\phi_x \notin [-1,1] \text{ for every } x \in \widetilde{\Lambda}_N\right\},\\
\mathcal{A}_1(x) &:= \{\phi_x \in [-1,1]\} \cap \left\{\phi_y \notin [-1,1] \text{ for every } y \in \widetilde{\Lambda}_N \setminus \{x\}\right\},\\
\mathcal{A}_1 &:= \bigcup_{x \in \widetilde{\Lambda}_N} \mathcal{A}_1(x),\\
\mathcal{A}_2 &:= \mathbb{R}^{\widetilde{\Lambda}_N} \setminus (\mathcal{A}_0 \cup \mathcal{A}_1)\,.
\end{aligned} \tag{5.24}$$

One has from Lemma 5.5 for any $x \in \mathring{\Lambda}_N$,

$$\mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_1(x)\right) \,\geq\, \mathbf{E}_N^{\widehat{\phi}}[\delta_x] - \sum_{y \in \mathring{\Lambda}_N \setminus \{x\}} \mathbf{E}_N^{\widehat{\phi}}[\delta_x\delta_y] \geq (1 - N^{-d})\mathbf{E}_N^{\widehat{\phi}}[\delta_x]\,, \tag{5.25}$$

where the first inequality is obtained by applying the union bound to $\{\delta_x = 1\} = \mathcal{A}_1 \cup \bigcup_{y \neq x}\{\delta_x = \delta_y = 1\}$. From (5.25) we directly have

$$\mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_1\right) \,\geq\, (1 - N^{-d}) \sum_{x \in \mathring{\Lambda}_N} \mathbf{E}_N^{\widehat{\phi}}[\delta_x]\,. \tag{5.26}$$

Using again the union bound, Lemma 5.5 and (5.26), we also have

$$\mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_2\right) \le \frac{1}{2} \sum_{\substack{(x,y)\in(\mathring{\Lambda}_N)^2 \\ x\neq y}} \mathbf{E}_N^{\widehat{\phi}}[\delta_x\delta_y] \le \frac{N^{-d}}{2} \sum_{x\in\mathring{\Lambda}_N} \mathbf{E}_N^{\widehat{\phi}}[\delta_x] \le N^{-d}\mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_1\right) . \tag{5.27}$$

Taking only into account the contribution of $\mathcal{A}_1$ and $\mathcal{A}_0$ to the partition function we obtain

$$\begin{aligned} Z_{N,h}^{\beta,\omega,\widehat{\phi}} &\ge \mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_0\cup\mathcal{A}_1\right) + \sum_{x\in\mathring{\Lambda}_N} \xi(x)\mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_1(x)\right) \\ &= 1 + \sum_{x\in\mathring{\Lambda}_N} \xi(x)\mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_1(x)\right) - \mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_2\right) =: Z' . \end{aligned} \tag{5.28}$$

But by (5.17)

$$Z' \ge \mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_0\right) \ge 1 - \frac{1}{N^d} \ge \frac{1}{2} , \tag{5.29}$$

and hence

$$\log Z' \ge (Z'-1) - (Z'-1)^2 . \tag{5.30}$$

Therefore from (5.27) one has

$$\mathbb{E}[Z'-1] = (e^h-1)\mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_1\right) - \mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_2\right) \ge (e^h-1-N^{-d})\mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_1\right) . \tag{5.31}$$

We also have (using (5.17))

$$\mathrm{Var}_{\mathbb{P}}\left(Z'\right) = e^{2h}\left(e^{\lambda(2\beta)-2\lambda(\beta)}-1\right) \sum_{x\in\mathring{\Lambda}_N} \mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_1(x)\right)^2 \le CN^{-2d}\mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_1\right) , \tag{5.32}$$

which is much smaller than $\mathbb{E}[Z']-1$ (recall $N=h^{-2}$). Overall (combining (5.31) (5.32) and (5.26)) one has for all $\widehat{\phi}$ in $\mathcal{E}_u$

$$\begin{aligned} \mathbb{E}\left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi}}\right] &\ge \mathbb{E}\left[\log Z'\right] \ge \mathbf{E}_N^{\widehat{\phi}}\left[Z'-1\right] - \left(\mathbf{E}_N^{\widehat{\phi}}\left[Z'-1\right]\right)^2 - \mathrm{Var}_{\mathbb{P}}\left(Z'\right) \\ &\ge (e^h-1-N^{-d})\mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_1\right) - \left(\left(e^h-1-N^{-d}\right)\mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_1\right)\right)^2 - CN^{-2d}\mathbf{P}_N^{\widehat{\phi}}\left(\mathcal{A}_1\right) , \end{aligned} \tag{5.33}$$

By using again $N=h^{-2}$, as well as (5.26), we deduce that for h sufficiently small

$$\mathbb{E}\left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi}}\right] \ge \frac{3h}{4}\left(\sum_{x\in\mathring{\Lambda}_N} \mathbf{E}_N^{\widehat{\phi}}[\delta_x]\right) . \tag{5.34}$$

Hence

$$\begin{aligned} \mathbb{E}\widehat{\mathbf{E}}^u\left[\log\left(Z_{N,h}^{\beta,\omega,\widehat{\phi}}\right)\mathbf{1}_{\mathcal{E}^u}\right] &\ge \frac{3h}{4}\mathbf{E}^u\left[\sum_{x\in\mathring{\Lambda}_N}\delta_x\mathbf{1}_{\mathcal{E}^u}\right] \\ &= \frac{3h}{4}\left((N-1)^d\mathbf{E}^u\left[\delta_0\right] - \mathbf{E}^u\left[\sum_{x\in\tilde{\Lambda}_N}\delta_x\mathbf{1}_{\mathcal{E}_u^c}\right]\right) \ge \frac{h}{2}N^d\mathbf{E}^u\left[\delta_0\right], \end{aligned} \tag{5.35}$$

where in the last inequality we used h small and (5.10). □

5.3. **Why is this not optimal?** The main idea of the proof above is to change the boundary conditions $\widehat{\phi}$ so that there are only a few contacts (the main contribution to the partition function is given by $\mathcal{A}_0 \cup \mathcal{A}_1$). In that case the partition function (or at least Z') has a very small variance and for this reason, Jensen's inequality for log is essentially sharp. The strategy could in principle (and with a lot of efforts) extend if one has typically a bounded number of contact, or possibly if one allows it to grow logarithmically , but the variance estimates would clearly blow up beyond this point.

So for this strategy to work we need that $\mathbf{E}^u[\delta_0] \le cN^{-d}$ for some positive constant c (possibly large or even growing very slowly with N but this latter possibility does not add much to the discussion). Now let us notice that since δ_x on the boundary is completely determined by the boundary conditions one can easily take away from the log partition function the contribution of the boundary

$$\mathbb{E}\widehat{\mathbf{E}}^u\left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi}}\right] =$$

$$(h-\lambda(\beta)) \sum_{x\in\tilde{\Lambda}_N\cap\partial\Lambda_N} \mathbf{E}^u[\delta_x] + \mathbb{E}\widehat{\mathbf{E}}^u \log \mathbf{E}_N^{\widehat{\phi}}\left[\exp\left(\sum_{x\in\mathring{\Lambda}_N}(\beta\omega_x-\lambda(\beta)+h)\delta_x\right)\right]. \quad (5.36)$$

The first term which is the boundary effect is negative (h is small!) and of order $N^{d-1}\mathbf{E}^u[\delta_0]$. Our best hope for the second term is to get something positive which is of order $hN^d\mathbf{E}^u[\delta_0]$ (this is what we get with an annealed bound). Hence for the second term in (5.36) to be dominant, we need h to be larger than N^{-1}. The best we can hope as a lower bound for the free energy density is then

$$h\mathbf{E}^u[\delta_0] \le c\,hN^{-d} = O\left(h^{d+1}\right). \quad (5.37)$$

To reduce the influence of boundary effects, one has to work with larger boxes, but in this case the total number of contact in the box will be large and one has to try an find other means of controlling $\mathbb{E}[\log Z]$ than only the variance. This is the aim of the coarse graining and replica-coupling approach of the next section.

6. The coarse graining procedure for the critical behavior (lower bound)

For $\sigma = \sigma_d$, $a > 0$ and $h > 0$ small we set

$$u = u(a,h) := \sigma\sqrt{2\log(1/h)} + 1 - \frac{\sigma}{2}\frac{\log\log(1/h)}{\sqrt{2\log(1/h)}} - \sigma\frac{\log\left(2a\sqrt{\pi}\right)}{\sqrt{2\log(1/h)}}. \quad (6.1)$$

Such a choice has been made to guarantee that the contact probability is (essentially) $a\,h$ for h small. The choice (6.1) is clearly connected to the following lemma on a standard Gaussian variable $\mathcal{N}$:

Lemma 6.1. *If $v : (0,\infty) \to \mathbb{R}$ is such that $\lim_{h\searrow 0} v(h)\sqrt{\log(1/h)} = 0$, then*

$$\mathbf{P}\left(u(a,h) + v(h) + \sigma\mathcal{N} \in [-1,1]\right) \overset{h\searrow 0}{\sim} ah. \quad (6.2)$$

Proof. The result follows, via a lengthy computation, from the well known asymptotic ($x \nearrow \infty$) estimate

$$\mathbf{P}(\mathcal{N} > x) = \frac{1}{x\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)\left(1+O\left(\frac{1}{x^2}\right)\right). \quad (6.3)$$

□

Remark 6.2. *It is easy to see that the statement of Lemma 6.1 holds also if we replace* $[-1,1]$ *with* $[c,1]$*, any* $c<1$*. More interestingly for us, it holds also for* $(-\infty,1]$ *(and thus also for* $(-\infty,0]$ *provided that* $u(a,h)$ *is replaced by* $u(a,h)-1$*). This remark is important because it ultimately means that the strategy of the proof that we are going to present below works also if we replace* δ_x *with* Δ_x *(recall* (2.35)*), that is if we pass from disordered pinning to the co-membrane model. This is true also for the rougher proof of Section 5 and in a more evident way since no estimate is sharp in that case.*

Let us introduce

$$\rho_h := \exp\left(-\sqrt{\log(1/h)}\right) \quad \text{and} \quad N_0 := \frac{1}{\rho_h}, \tag{6.4}$$

and without true loss of generality we will assume $N_0 \in 2\mathbb{N}$. For $h \searrow 0$ and arbitrary $b>0$ we have

$$h^b \ll \rho_h \ll |\log h|^{-1/b}.$$

We then choose N_1 such that $N_1/N_0 \in 2\mathbb{N}$, $N_1/N_0 \geq 4$ and

$$N_1 \in \left[\frac{1}{2}h^{-3}, h^{-3}\right]. \tag{6.5}$$

We aim at showing

Proposition 6.3. *Choose* $\beta>0$*. There exist* $a(\beta,d)>0$ *and* $c(\beta,d)>0$ *such that for* $h>0$ *sufficiently small we have*

$$\frac{1}{N_1^d}\mathbb{E}\widehat{\mathbf{E}}^{u(a,h)}\left[\log Z^{\omega}_{N_1,\beta,h}\right] \geq c(\beta,d)\,h^2\,. \tag{6.6}$$

Moreover one find a constant $C(d)$ *such that forall* $\beta\in(0,1]$ *we have*

$$c(\beta,d) \geq C(d)\beta^{-2} \tag{6.7}$$

Proof. The proof is done in several steps: a number of lemmas will be stated and proved after the main body of the proof.

Step 1: Smoothing the roughness of $\widehat{\phi}$*.* We start off by selecting a subset of the $\widehat{\phi}$ configuration (of $\widehat{\mathbf{P}}^u$ probability very close to one) that guarantees that harmonic averages of the boundary value are extremely close to u, at least when we are not too close to the boundary. We do this by introducing the event

$$B_u := \left\{\widehat{\phi}\in\mathbb{R}^{\mathbb{Z}^d} : \left|H^{\widehat{\phi}}_{\Lambda_{N_1}}(x)-u\right| \leq \frac{\rho_h^{\frac{1}{8}}}{2} \text{ for every } x\in\Lambda_{N_1} \text{ s.t. } d\,(x,\partial\Lambda_{N_1}) \geq \frac{N_0}{2}\right\}, \tag{6.8}$$

where $H^{\phi}_{\Lambda}(x)$ is the solution to the harmonic equation (2.51) in Λ_N, with ϕ boundary conditions. We prove the following estimate at the end of the section.

Lemma 6.4. *For* h *sufficiently small*

$$\widehat{\mathbf{P}}^u\left(B_u^{\complement}\right) \leq \exp\left(-\rho_h^{-\frac{1}{5}}\right). \tag{6.9}$$

In other terms, $\widehat{\mathbf{P}}^u\left(B_u^{\complement}\right)$ is smaller than any power of h. As it stands, Lemma 6.4 is stated and will be used for the value of u given in (6.1), but it is easy to realize that $\widehat{\mathbf{P}}^u\left(B_u\right)$ does not depend on u and therefore Lemma 6.4 holds uniformly in u.

Since by Jensen's inequality for $h \ge 0$

$$\mathbb{E}\left[\log Z^{\omega,\widehat{\phi}}_{N_1,\beta,h}\right] \ge \sum_{x\in\widetilde{\Lambda}_{N_1}} \left(h - \frac{\beta^2}{2}\right) \mathbf{E}^{\widehat{\phi}}_{N_1}[\delta_x] \ge -\frac{\beta^2}{2} N_1^d\,, \tag{6.10}$$

we readily see that

$$\frac{1}{N_1^d}\mathbb{E}\widehat{\mathbf{E}}^u\left[\log Z^{\omega}_{N_1,\beta,h};\, B_u^{\complement}\right] \ge -\frac{\beta^2}{2}\widehat{\mathbf{P}}^u\left(B_u^{\complement}\right)\,. \tag{6.11}$$

Therefore, in view of Lemma 6.4 and of the result (6.6) we are after, it suffices to show that

$$\frac{1}{N_1^d}\mathbb{E}\widehat{\mathbf{E}}^u\left[\log Z^{\omega}_{N_1,\beta,h};\, B_u\right] \ge 2c\,h^2\,. \tag{6.12}$$

Step 2: Neglecting the energy contribution near $\partial\Lambda_{N_1}$. We show now that we can neglect the energy contribution coming from the sites on which we do not have a control on the harmonic average of the boundary. This is done by introducing

$$\Lambda^-_{N_1,N_0} := \{N_0+1, N_0+2, \dots, N_1-N_0\}^d\,, \tag{6.13}$$

by restricting the sum in the energy term to sites in $\Lambda^-_{N_1,N_0}$: as we are going to show right away is that this introduces an error in the free energy computation that is $o(h^2)$, hence irrelevant. This space between the boundary of the box $\widetilde{\Lambda}_{N_1}$ and the sites that contribute to the energy has been introduced to allow some averaging of the boundary conditions $\widehat{\phi}$. In fact $\widehat{\phi}$ has fluctuations of order one and therefore the field ϕ close to the boundary has a mean that inherits this incertitude, while, as shown in Lemma 6.4, sufficiently far away – a distance N_0 suffices – the mean will be u up to an error $O(\rho_h^{1/8})$ which is smaller than any negative power of $\log(1/h)$. The estimate for the error introduced by restricting the energy contribution to sites in $\Lambda^-_{N_1,N_0}$ goes as follows: start by observing that

$$\begin{aligned}\sum_{x\in\widetilde{\Lambda}_{N_1}} \left(\beta\omega_x - \frac{\beta^2}{2} + h\right)\delta_x \ge &- \sum_{x\in\widetilde{\Lambda}_{N_1}\setminus\Lambda^-_{N_1,N_0}} \left|\beta\omega_x - \frac{\beta^2}{2} + h\right| \\ &+ \sum_{x\in\Lambda^-_{N_1,N_0}} \left(\beta\omega_x - \frac{\beta^2}{2} + h\right)\delta_x\,,\end{aligned} \tag{6.14}$$

so that,

$$\begin{aligned}\frac{1}{N_1^d}\mathbb{E}\widehat{\mathbf{E}}^u\left[\log Z^{\omega,\widehat{\phi}}_{N_1,\beta,h};\, B_u\right] \ge &-\frac{\left|\widetilde{\Lambda}_{N_1}\setminus\Lambda^-_{N_1,N_0}\right|}{N_1^d}\mathbb{E}\left[\left|\beta\omega_0 - \frac{\beta^2}{2} + h\right|\right] \\ &+ \frac{1}{N_1^d}\mathbb{E}\widehat{\mathbf{E}}^u\left[\log \mathbf{E}^{\widehat{\phi}}_{N_1}\left[\exp\left(\sum_{x\in\Lambda^-_{N_1,N_0}}\left(\beta\omega_x - \frac{\beta^2}{2} + h\right)\delta_x\right)\right];\, B_u\right]\,.\end{aligned} \tag{6.15}$$

Therefore the first term on the right-hand side is $O(N_0/N_1) \ll h^2$ (this has determined our choice of N_1), so that we can effectively neglect the energy contribution of the sites outside $\Lambda^-_{N_1,N_0}$ and Proposition 6.3 reduces to showing

$$\frac{1}{N_1^d}\mathbb{E}\widehat{\mathbf{E}}^u\left[\log \mathbf{E}^{\widehat{\phi}}_{N_1}\left[\exp\left(\sum_{x\in\Lambda^-_{N_1,N_0}}\left(\beta\omega_x-\frac{\beta^2}{2}+h\right)\delta_x\right)\right];\, B_{u,}\right] \,\geq\, 3c\,h^2\,. \tag{6.16}$$

This estimate will be obtained by restricting the $\mathbf{P}^{\widehat{\phi}}_{N_1}$-expectation to an event A_κ, κ a positive integer (given explicitly just below) that depends only on d: κ is a constraint that we are going to introduce on the number of contacts in *intermediate* scale boxes.

Step 3: the coarse graining grid and the event A_κ. To define A_κ we first introduce a decomposition of $\Lambda^-_{N_1,N_0}$ (Figure 1 may be of help in following the construction). For $w\in\{0,1\}^d$ we set

$$\Lambda^w_{N_1,N_0} \,:=\, \left\{x\in\Lambda^-_{N_1,N_0}\,:\,\left\lceil\frac{x_i}{N_0}\right\rceil \overset{\text{mod }2}{=} w_i \text{ for } i=1,2,\dots,d\right\}\,. \tag{6.17}$$

Actually $\Lambda^w_{N_1,N_0}$ can be seen as a disjoint union of $\left(\frac{N_1}{2N_0}-1\right)^d$ (hyper)cubes of edge length

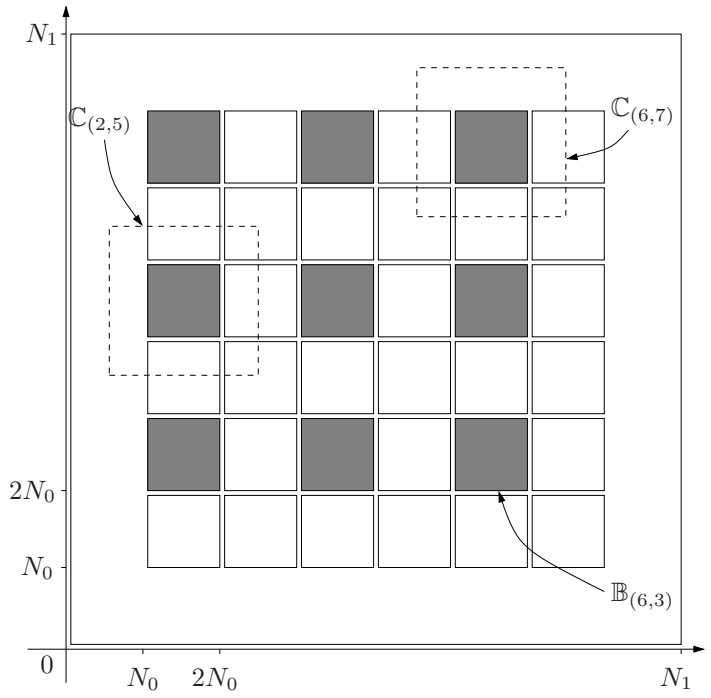

FIGURE 1. The set $\widetilde{\Lambda}_{N_1}$ is drawn for $d=2$ (we need it only for $d\geq 3$, but for illustration purposes $d=2$ is enough) and $N_1=8N_0$. The set $\Lambda^-_{N_1,N_2}$ is the (disjoint union) of the B_j's boxes, $j\in\{2,3,4,5,6,7\}^2=\cup_{w\in\{0,1\}^2}\mathcal{J}_w$. We have singled out $\Lambda^{(0,1)}_{N_1,N_0}$ by making it darker: observe that $\Lambda^-_{N_1,N_0}$ is the disjoint union of $\Lambda^w_{N_1,N_0}, w\in\{0,1\}^2$. We have also drawn, with dashed boundaries, some of the $\mathbb{C}_j$ boxes.

N_0:

$$\Lambda^w_{N_1,N_0} \,=\, \cup_{j\in\mathcal{J}_w}\mathbb{B}_j\,, \tag{6.18}$$

where

$$\mathbb{B}_j := \widetilde{\Lambda}_{N_0} + N_0 j \tag{6.19}$$

Even if it is probably not necessary, we can make $\mathcal{J}_w$ explicit:

$$\mathcal{J}_w \,=\, \left\{ \left(\left\lceil \frac{x_1}{N_0} \right\rceil, \dots, \left\lceil \frac{x_d}{N_0} \right\rceil \right) :\, x\in\Lambda^-_{N_1,N_0} \text{ and } \left\lceil \frac{x_i}{N_0} \right\rceil \overset{\text{mod}\,2}{=} w_i \text{ for } i=1,2,\dots,d \right\}\,, \tag{6.20}$$

Note that

$$\bigcup_{w\in\{0,1\}^d} \Lambda^w_{N_1,N_0} = \Lambda_{N_1,N_0} \tag{6.21}$$

and therefore

$$\Lambda_{N_1,N_0} = \cup_{j\in\mathcal{J}}\mathbb{B}_j \tag{6.22}$$

where $\mathcal{J} = \cup_w \mathcal{J}_w = \{2,\dots,N_1/N_0-1\}^d$. Last, for $j\in\mathcal{J}$ we set

$$\mathbb{C}_j := \{x :\, \text{dist}(x,\mathbb{B}_j) \le N_0/2\}. \tag{6.23}$$

We are now ready to introduce $A_\kappa := A^{(1)} \cap A^{(2)}_\kappa$, where

$$A^{(1)} \,:=\, \left\{ \phi :\, \max_{j\in\mathcal{J}} \max_{x\in\mathbb{B}_j} \left| H^\phi_{\mathbb{C}_j}(x) - u \right| \,\le\, \rho_h^{-\frac{1}{8}} \right\}\,, \tag{6.24}$$

and

$$A^{(2)}_\kappa \,:=\, \left\{ \phi :\, \max_{j\in\mathcal{J}} \sum_{x\in\mathbb{B}_j} \delta_x \,\le\, \kappa \right\}\,. \tag{6.25}$$

Recall that u is chosen in (6.1) and B_u is given in (6.8). The argument that we are going to present works for κ sufficiently large and how large just depends on the dimension. For definiteness we make the choice of κ explicit:

$$\kappa := \lceil d^3 3^{12} 2^{12(d+5)} \mathsf{c}_d^{12} \rceil,$$

where $\mathsf{c}_d \ge 1$ is the constant in (2.44). Such a choice stems out of several arbitrary and lazy choices in the chain of rough bounds that constitutes the proof.

We introduce now an important technical estimate whose proof is postponed to the end of the section.

Lemma 6.5. *For h sufficiently small*

$$\sup_{\widehat{\phi}\in B_u} \mathbf{P}^{\widehat{\phi}}_{N_1}\left(A^{\complement}_\kappa \right) \,\le\, h^{k^{1/4}}\,. \tag{6.26}$$

Step 4: Replica coupling. Going back to (6.16), it is clear that it suffices to show that for h sufficiently small

$$\frac{1}{N_1^d}\mathbb{E}\widehat{\mathbf{E}}^u\left[\log \mathbf{E}_{N_1}^{\widehat{\phi}}\left[\exp\left(\sum_{x\in\Lambda_{N_1,N_0}}\left(\beta\omega_x-\frac{\beta^2}{2}+h\right)\delta_x\right);\,A_\kappa\right];\,B_u\right]\,\geq\,3c\,h^2\,. \tag{6.27}$$

For this we will exploit a replica coupling argument bound. The following result, which we prove in the Appendix, is inspired by and very close to the computations made in [39] for renewal pinning. The proof exploits interpolation technics similar to those found in spin-glass literature [32].

Lemma 6.6.

$$\frac{1}{N_1^d}\mathbb{E}\left[\log \mathbf{E}_{N_1}^{\widehat{\phi}}\left[\exp\left(\sum_{x\in\Lambda_{N_1,N_0}}\left(\beta\omega_x-\frac{\beta^2}{2}+h\right)\delta_x\right);\,A_\kappa\right]\right]\,\geq\,T_1-T_2\,, \tag{6.28}$$

with

$$T_1\,:=\,\frac{1}{N_1^d}\log \mathbf{E}_{N_1}^{\widehat{\phi}}\left[\exp\left(h\sum_{x\in\Lambda_{N_1,N_0}}\delta_x\right);\,A_\kappa\right]\,, \tag{6.29}$$

and

$$T_2\,:=\,\frac{1}{N_1^d}\log\left\langle\exp\left(2\beta^2\sum_{x\in\Lambda_{N_1,N_0}}\delta_x^{(1)}\delta_x^{(2)}\right);\,A_\kappa^2\right\rangle_{N_1,h,\widehat{\phi}}^{\otimes 2}\,, \tag{6.30}$$

with

$$\langle\,\cdot\,\rangle_{N_1,h,\widehat{\phi}}\,:=\,\frac{\mathbf{E}_{N_1}^{\widehat{\phi}}\left[\,\cdot\,\exp\left(h\sum_{x\in\Lambda_{N_1,N_0}}\delta_x\right)\right]}{\mathbf{E}_{N_1}^{\widehat{\phi}}\left[\exp\left(h\sum_{x\in\Lambda_{N_1,N_0}}\delta_x\right)\right]}\,. \tag{6.31}$$

Remark 6.7. *It is obvious from the proof that the above Lemma remains valid without restriction to the event A_κ (or with a restriction to another event). However it is not too difficult to check that without this restriction, the quantity T_2 would be of order β^2 and hence the result completely useless (the right hand side of* (6.28) *would be negative). We have designed the event A_κ to be of small probability so that T_1 is almost equal to to the value it would have with no conditioning, but such that T_2 becomes much smaller with the conditioning.*

We now need a lower bound on $\mathbf{E}^u[T_1;\,B_u]$ and an upper bound on $\mathbf{E}^u[T_2;\,B_u]$.

Step 5: lower bound on $\mathbf{E}^u[T_1;\,B_u]$. We apply Jensen's inequality after a rearrangement

$$\begin{aligned}T_1\,&=\,\frac{1}{N_1^d}\log \mathbf{E}_{N_1}^{\widehat{\phi}}\left[\exp\left(h\sum_{x\in\Lambda_{N_1,N_0}}\delta_x\right)\,\middle|\,A_\kappa\right]+\frac{1}{N_1^d}\log \mathbf{P}_{N_1}^{\widehat{\phi}}\left(A_\kappa\right)\\&\geq\,\frac{h}{N_1^d}\mathbf{E}_{N_1}^{\widehat{\phi}}\left[\sum_{x\in\Lambda_{N_1,N_0}}\delta_x\,\middle|\,A_\kappa\right]+\frac{1}{N_1^d}\log \mathbf{P}_{N_1}^{\widehat{\phi}}\left(A_\kappa\right)\,=:\,T_{1,1}+T_{1,2}.\end{aligned} \tag{6.32}$$

We have

$$\begin{aligned}\widehat{\mathbf{E}}^u[T_{1,1};\, B_u] \,&\geq\, \frac{h}{N_1^d}\widehat{\mathbf{E}}^u\left[\mathbf{E}_{N_1}^{\widehat{\phi}}\left[\sum_{x\in\Lambda_{N_1,N_0}}\delta_x\,;\, A_\kappa\right];\, B_u\right]\\ &=\, \frac{h}{N_1^d}\mathbf{E}^u\left[\sum_{x\in\Lambda_{N_1,N_0}}\delta_x\,;\, A_\kappa\cap B_u\right]\\ &\geq\, \frac{h}{N_1^d}\mathbf{E}^u\left[\sum_{x\in\Lambda_{N_1,N_0}}\delta_x\right]-h\left(\mathbf{P}^u\left((A_\kappa\cap B_u)^{\complement}\right)\right),\end{aligned}\tag{6.33}$$

so that by using Lemma 6.1, Lemma 6.4 and Lemma 6.5 (recall the choice of κ) we readily see that

$$\widehat{\mathbf{E}}^u[T_{1,1};\, B_u] \,\geq\, \frac{2a}{3}h^2\,,\tag{6.34}$$

for h sufficiently small.

On the other hand, by Lemma 6.5 we see that $\mathbf{P}^{\widehat{\phi}}\left(A_\kappa^{\complement}\right)\geq 1/2$ for h small, uniformly in $\widehat{\phi}\in B_u$ and this entails $\widehat{\mathbf{E}}^u[T_{1,2};\, B_u]\geq -8(\log 2)h^{3d}$ so

$$\widehat{\mathbf{E}}^u[T_1;\, B_u] \,\geq\, \frac{a}{2}h^2\,.\tag{6.35}$$

Step 6: upper bound on $\mathbf{E}^u[T_2;\, B_u]$. We start with the preliminary observation that $\phi\in A_\kappa^{(2)}$ means that there are at most κ contacts in $\mathbb{B}_j$ for every $j\in\mathcal{J}$ and this implies that there are at most $3^d\kappa$ contacts in $\mathbb{C}_j\cap\Lambda_{N_1,N_0}$ (and a fortiori in $\mathring{\mathbb{C}}_j\cap\Lambda_{N_1,N_0}$) because if $\mathbb{C}_j\cap\Lambda_{N_1,N_0}=\mathbb{C}_j$ then $\mathbb{C}_j$ is covered by 3^d $\mathbb{B}_j$'s (this is the typical case: in the bulk). When $\mathbb{C}_j\cap\Lambda_{N_1,N_0}$ is (strictly) contained in $\mathbb{C}_j$ (the boundary case), fewer $\mathbb{B}_j$ boxes suffice. Therefore for every $j\in\mathcal{J}$ we introduce the event

$$A_\kappa^{(3)}(j) \,:=\, \left\{\sum_{x\in\mathring{\mathbb{C}}_j\cap\Lambda_{N_1,N_0}}\delta_x\leq 3^d\kappa \;\text{ and }\; \sum_{x\in\mathbb{B}_j}\delta_x\leq\kappa.\right\},\tag{6.36}$$

One can check that

$$A_\kappa := A^{(1)}\cap A_\kappa^{(2)} \,=\, A^{(1)}\cap\left(\cap_{j\in\mathcal{J}}A_\kappa^{(3)}(j)\right),\tag{6.37}$$

Now we apply the Hölder inequality $|\mathbf{E}\prod_{i=1}^k X_i|\leq\prod_{i=1}^k(\mathbf{E}|X_i|^k)^{1/k}$

$$\begin{aligned}T_2 \,&=\, \frac{1}{N_1^d}\log\left\langle\prod_{w\in\{0,1\}^d}\exp\left(2\beta^2\sum_{x\in\Lambda_{N_1,N_0}^w}\delta_x^{(1)}\delta_x^{(2)}\right);\, A_\kappa^2\right\rangle_{N_1,h,\widehat{\phi}}^{\otimes 2}\\ &\leq\, \frac{1}{(2N_1)^d}\sum_{w\in\{0,1\}^d}\log\left\langle\exp\left(2^{1+d}\beta^2\sum_{x\in\Lambda_{N_1,N_0}^w}\delta_x^{(1)}\delta_x^{(2)}\right);\, A_\kappa^2\right\rangle_{N_1,h,\widehat{\phi}}^{\otimes 2}.\end{aligned}\tag{6.38}$$

Let us focus on the argument of the logarithm and condition the measure $\langle\cdot\rangle_{N_1,h,\widehat{\phi}}$ to the σ-algebra generated by $\{\phi_x\}_{x\in\cup_{j\in\mathcal{J}_w}\partial\mathbb{C}_j}$. By the spatial Markov property we obtain

$$\left\langle \exp\left(2^{1+d}\beta^2\sum_{x\in\Lambda^w_{N_1,N_0}}\delta^{(1)}_x\delta^{(2)}_x\right);\left(\phi^{(1)},\phi^{(2)}\right)\in A^2_\kappa\right\rangle^{\otimes 2}_{N_1,h,\widehat{\phi}}\leq$$
$$\left\langle\prod_{j\in\mathcal{J}_w}\mathbf{E}^{\otimes 2}_{\mathbb{C}_j,h,(\phi^{(1)},\phi^{(2)})}\left[\exp\left(2^{1+d}\beta^2\sum_{x\in\mathbb{B}_j}\delta^{(1)}_x\delta^{(2)}_x\right);\left(A^{(3)}_\kappa(j)\right)^2\right];\right.$$
$$\left.\left(\phi^{(1)},\phi^{(2)}\right)\in\left(A^{(1)}_\kappa\right)^2\right\rangle^{\otimes 2}_{N_1,h,\widehat{\phi}},\quad(6.39)$$

where

$$\mathbf{E}^{\otimes 2}_{\mathbb{C}_j,h,(\phi^{(1)},\phi^{(2)})}=\mathbf{E}_{\mathbb{C}_j,h,\phi^{(1)}}\mathbf{E}_{\mathbb{C}_j,h,\phi^{(2)}},\quad(6.40)$$

and

$$\mathbf{E}_{\mathbb{C}_j,h,\phi}[\,\cdot\,]:=\mathbf{E}_{\mathbb{C}_j,\phi}\left[\cdot\,\exp\left(h\sum_{x\in\mathring{\mathbb{C}}_j}\delta_x\right)\right]\Big/\mathbf{E}_{\mathbb{C}_j,\phi}\left[\exp\left(h\sum_{x\in\mathring{\mathbb{C}}_j}\delta_x\right)\right]\quad(6.41)$$

and $\mathbf{P}_{\mathbb{C}_j,\phi}$ is the free field on the set $\mathbb{C}_j$ with boundary conditions ϕ (remark that this notation is a bit improper since ϕ is used for the boundary conditions and in the definition of δ, but we believed this is more readable that introducing $\widetilde{\delta}$ and should not generate confusion). Of course if $F:\mathbb{R}^{\mathring{\mathbb{C}}_j}\mapsto[0,\infty)$ is measurable then $\mathbf{E}_{\mathbb{C}_j,h,\phi}[F]$ is measurable with respect to $\{\phi_x\}_{x\in\partial\mathbb{C}_j}$.

We now recall that we need to bound from above $\widehat{\mathbf{E}}^u[T_2;\,B_u]$ and we will do this by taking the supremum over $\phi\in A^{(1)}$ (and applying $\widehat{\mathbf{E}}^u[\,\cdot\,;\,B_u]$ will be irrelevant), that is by (6.38) and (6.39) we have

$$\widehat{\mathbf{E}}^u[T_2;\,B_u]\leq$$
$$\frac{1}{N_1^d}\sum_{j\in\mathcal{J}}\sup_{\phi^{(1)},\phi^{(2)}\in A^{(1)}}\log\mathbf{E}^{\otimes 2}_{\mathbb{C}_j,h,(\phi^{(1)},\phi^{(2)})}\left[\exp\left(2^{1+d}\beta^2\sum_{x\in\mathbb{B}_j}\delta^{(1)}_x\delta^{(2)}_x\right);\left(A^{(3)}_\kappa(j)\right)^2\right],\quad(6.42)$$

and we are left with estimating the terms in the sum in the right-hand side. These terms are actually identical, except for the boundary cases (but they can be bounded in the very same way). Let us record the first part of the argument as a lemma, that we will also use in the next sections.

Lemma 6.8. *Let X be a positive random variable such that $X\leq\gamma$ with probability 1. Then*

$$\log\mathbb{E}\left[e^X\right]\leq(e^\gamma-1)\,\mathbb{E}[X].\quad(6.43)$$

Proof. We simply use convexity to show that

$$e^X\leq 1+(e^\gamma-1)X,$$

and the inequality $\log(1+u)\leq u$. □

Now applying the Lemma to $X := 2^{1+d}\beta^2 \sum_{x\in\mathbb{B}_j} \delta_x^{(1)}\delta_x^{(2)}$, and remarking that on $A_\kappa^{(3)}(j)$, we have $X \le 2^{1+d}\beta^2\kappa$, we obtain

$$\mathbf{E}^{\otimes 2}_{\mathbb{C}_j,h,(\phi^{(1)},\phi^{(2)})}\left[\exp\left(2^{1+d}\beta^2\sum_{x\in\mathbb{B}_j}\delta_x^{(1)}\delta_x^{(2)}\right);\left(A_\kappa^{(3)}(j)\right)^2\right] \le 1+\eta(\beta,\kappa,d)\sum_{x\in\mathbb{B}_j}\sup_{\phi\in A^{(1)}}\mathbf{E}_{\mathbb{C}_j,h,\phi}\left[\delta_x;A_\kappa^{(3)}(j)\right]^2, \quad (6.44)$$

with

$$\eta(\beta,\kappa,d) := \frac{\exp\left(2^{1+d}\beta^2\kappa\right)-1}{\kappa}. \quad (6.45)$$

In the case $\beta\in(0,1]$ notice that we have

$$\eta(\beta,\kappa,d) \le \frac{\exp\left(2^{1+d}\kappa\right)}{\kappa}\beta^2 := \eta(\kappa,d)\beta^2. \quad (6.46)$$

But by using the definition of $\mathbf{E}_{\mathbb{C}_j,h,\phi}[\cdot]$, since $\mathbf{E}_{\mathbb{C}_j,0,\phi}[\exp(h\sum_{x\in\mathring{\mathbb{C}}_j}\delta_x)] \ge 1$ and since $\sum_{x\in\mathring{\mathbb{C}}_j}\delta_x$ is bounded by $3^d\kappa$ on $A_\kappa^{(3)}(j)$ we have

$$\mathbf{E}_{\mathbb{C}_j,h,\phi}\left[\delta_x;A_\kappa^{(3)}(j)\right] \le \exp\left(3^d\kappa h\right)\mathbf{E}_{\mathbb{C}_j,\phi}\left[\delta_x\right] \le \frac{3}{2}\mathbf{E}_{\mathbb{C}_j,\phi}\left[\delta_x\right] \le 2ah, \quad (6.47)$$

where we have chosen h sufficiently small. The last inequality is due to Lemma 6.1 applied with $v(h,x) := H^\phi_{\mathbb{C}_j}(x) - u$ (since $\phi\in A_\kappa^{(1)}$, $v(h,x)$ satisfies the assumption of the Lemma uniformly in x). Therefore

$$\widehat{\mathbf{E}}^u[T_2;\,B_u] \le 4a^2\eta(\beta,\kappa,d)h^2. \quad (6.48)$$

Step 7: conclusion. By putting (6.35) and (6.48) together and recalling Lemma 6.6 we find that $((a/2)-4a^2\eta(\kappa,d,\beta))h^2$ is a lower bound for the expression in the left-hand side of (6.27) and once a is chosen smaller $1/(8\eta(\beta,\kappa,d))$ the proof of Proposition 6.3 is complete, the case $\beta\in[0,1]$ follow from (6.46). □

6.1. **Proof of Lemmas 6.4 and 6.5.** We start off with an elementary (rough) estimate on the variance of the harmonic extension. Throughout this section we use the short-cut notation for $x\in\Lambda$ and $y\in\partial\Lambda$,

$$p_\Lambda(x,y) := P^x\left(X_{\tau_{\partial\Lambda}}=y\right), \quad (6.49)$$

with which (2.52) becomes

$$H^\phi_\Lambda(x) = \sum_{y\in\partial\Lambda} p_\Lambda(x,y)\phi_y. \quad (6.50)$$

Lemma 6.9. *Let N and M be two integer numbers such that $N>2M>0$. As usual $\Lambda_N=\{0,1,\dots,N\}^d$ and we introduce $\Lambda_{N,M}:=\{M,\dots,N-M\}^d$. Let $\{\phi_x\}_{x\in\Lambda_N}$, with law $\mathbf{P}_N$, be a centered Gaussian field with covariance $G_{\Lambda_N}(x,y)\le G(x,y)$ for every $x,y\in\partial\Lambda_N$ (we recall that $\widehat{\mathbf{P}}^0$ is the law of the infinite volume Gaussian lattice free field). Then there exists $C_d>0$ (depending only on the dimension) such that for every M we have*

$$\sup_{N:\,N>2M}\max_{x\in\Lambda_{N,M}}\mathrm{var}_{\mathbf{P}_N}\left(H^\phi_{\Lambda_N}(x)\right) \le C_dM^{-\frac{d-2}{2}}. \quad (6.51)$$

Remark 6.10. *More advanced computations could show that the l.h.s of* (6.51) *is truly of order* M^{2-d}*, but the bound presented above is much easier to obtain and sufficient for our purposes.*

Proof. By recalling (2.44), we observe that $G_N(x,y)$ is bounded above by $\mathsf{c}_d/(1+|x-y|)^{d-2}$ for every x and y. Hence, we obtain that for $x \in \Lambda_{N,M}$

$$\begin{aligned}
\mathrm{var}_{\mathbf{P}_N}\left(H^{\phi}_{\Lambda_N}(x)\right) &= \sum_{y,y'\in\partial\Lambda_N} p_{\Lambda_N}(x,y)p_{\Lambda_N}(x,y')G_{\Lambda_N}(y,y') \\
&\le \frac{\mathsf{c}_d}{M^{\frac{d-2}{2}}} \sum_{\substack{y,y'\in\partial\Lambda_N:\\ |y-y'|\ge M^{\frac12}}} p_{\Lambda_N}(x,y)p_{\Lambda_N}(x,y') + \mathsf{c}_d \sum_{\substack{y,y'\in\partial\Lambda_N:\\ |y-y'|< M^{\frac12}}} p_{\Lambda_N}(x,y)p_{\Lambda_N}(x,y') \\
&\le \frac{\mathsf{c}_d}{M^{\frac{d-2}{2}}} + \mathsf{c}_d \left(\max_{\substack{y\in\partial\Lambda_N\\ x\in\Lambda_{N,M}}} p_{\Lambda_N}(x,y)\right) \sum_{y\in\partial\Lambda_N} p_{\Lambda_N}(x,y)\left|\left\{y'\in\partial\Lambda_N:\, |y'-y|<M^{\frac12}\right\}\right| \\
&\le \frac{\mathsf{c}_d}{M^{\frac{d-2}{2}}} + d\,2^{d-1}\mathsf{c}_d\, M^{\frac{d-1}{2}} \left(\max_{\substack{y\in\partial\Lambda_N\\ x\in\Lambda_{N,M}}} p_{\Lambda_N}(x,y)\right). \qquad (6.52)
\end{aligned}$$

We are now going to bound the term in the parentheses in the last line by M^{-d+1} times a constant that depends only on the dimension: and once this is done the proof of Lemma 6.9 is complete. This can be achieved by using the explicit expression for $p_{\Lambda_N}(x,y)$ – the exit probability from a cube – that one finds in [37, Prop. 8.1.3], but this expression is rather involved and we prefer to perform some steps to bound $p_{\Lambda_N}(x,y)$ with an exit probability from a half-space. For this we observe that without loss of generality we can assume that y belongs to the hyperplane $\mathbb{H}_d := \{z\in\mathbb{Z}^d:\, z_d = 0\}$ and, by elementary considerations,

$$p_{\Lambda_N}(x,y) \le p_{\mathbb{H}^+_d}(x,y),$$

where

$$\mathbb{H}^+_d := \{z:\, z_d \ge 0\}.$$

In order to simplify further the expression let us point out that we are left with estimating $\max_y p_{\mathbb{H}^+_d}(x,y)$ and we can therefore simply consider the case of $x = (0,\dots,0,L)$, with $L \ge M$. We have

$$\begin{aligned}
p_{\mathbb{H}^+_d}\big((0,\dots,0,L),(y_1,\dots,y_{d-1},0)\big) &= p_{\mathbb{H}^+_d}\big((-y_1,\dots,-y_{d-1},L),(0,\dots,0)\big) \\
&= \frac{2L}{\Sigma_d|z|^d}\left(1+O\left(\frac{L}{|z|^2}\right)\right) + O\left(\frac{1}{|z|^{d+1}}\right), \qquad (6.53)
\end{aligned}$$

where $z = (-y_1,\dots,-y_{d-1},L)$ and in the second step we have used [37, Th. 8.1.2] (Σ_d is the measure of the $d-1$ surface of the unit ball in $\mathbb{R}^d$). This last estimate suffices to conclude the proof of Lemma 6.9. □

Proof of Lemma 6.4. As remarked right after the statement, we can assume $u = 0$. We use Lemma 6.9 with $N = N_1$ and $M = N_0/2$. Therefore (using $d \ge 3$)

$$\max_{\substack{x\in\Lambda_{N_1}:\\ \mathrm{d}\left(x,\Lambda^{\complement}_{N_1}\right)\ge N_0/2}} \mathrm{var}_{\widehat{\mathbf{P}}^0}\left(H^{\widehat{\phi}}_{\Lambda_{N_1}}(x)\right) \le C_d N_0^{-\frac{d-2}{2}} \le C_d\rho_h^{\frac12}. \qquad (6.54)$$

Thus, by exponential Chebychev bounds, for the x's we are dealing with we have

$$\widehat{\mathbf{P}}^0\left(\left|H^{\widehat{\phi}}_{\Lambda_{N_1}}(x)\right| > \frac{\rho_h^{\frac{1}{8}}}{2}\right) \le 2\exp\left(-\frac{\rho_h^{\frac{1}{4}-\frac{1}{2}}}{8C_d}\right) = 2\exp\left(-\frac{\rho_h^{-\frac{1}{4}}}{8C_d}\right). \tag{6.55}$$

By performing a union bound we see that

$$\widehat{\mathbf{P}}^u\left(B_u^{\complement}\right) \le 2h^{-3d}\exp\left(-\frac{\rho_h^{-\frac{1}{4}}}{8C_d}\right), \tag{6.56}$$

and the proof is complete. □

In the proof of Lemma 6.5 we make use of the following estimate whose proof is postponed.

Lemma 6.11. *For $d \ge 3$ and for every $\kappa = 1, 2, \dots$ we have*

$$\sup_{B\subset\mathbb{Z}^d:\,|B|=\kappa}\sum_{(x,y)\in B^2} G(x,y) \le c(d)\kappa^{1+\frac{2}{d}}, \tag{6.57}$$

where $c(d) = 2^{d+4}\mathsf{c}_d$ (c_d is the constant appearing in (2.44)).

Proof of Lemma 6.5. Recall (6.24) and (6.25). We use

$$\mathbf{P}^{\widehat{\phi}}_{N_1}\left(A_\kappa^{\complement}\right) = \mathbf{P}^{\widehat{\phi}}_{N_1}\left(A^{(1)^{\complement}}\right) + \mathbf{P}^{\widehat{\phi}}_{N_1}\left(A_\kappa^{(2)^{\complement}} \cap A^{(1)}\right), \tag{6.58}$$

and we estimate the two terms on the right-hand side, uniformly over $\widehat{\phi} \in B_u$.

For the first term we start by observing that since $\widehat{\phi} \in B_u$

$$\mathbf{P}^{\widehat{\phi}}_{N_1}\left(\left|H^{\phi}_{\mathbb{C}_j}(x) - u\right| > \rho_h^{\frac{1}{8}}\right) \le \mathbf{P}^{\widehat{\phi}}_{N_1}\left(\left|H^{\phi}_{\mathbb{C}_j}(x) - \mathbf{E}^{\widehat{\phi}}_{N_1}\left[H^{\phi}_{\mathbb{C}_j}(x)\right]\right| > \frac{\rho_h^{\frac{1}{8}}}{2}\right), \tag{6.59}$$

and we apply Lemma 6.9 with $N = 2N_0$ and $M = N_0/2$ to obtain that for every $x \in \mathbb{C}_j$

$$\mathrm{var}_{\mathbf{P}^{\widehat{\phi}}_{N_1}}\left(H^{\phi}_{\mathbb{C}_j}(x)\right) \le C_d\left(\frac{N_0}{2}\right)^{-\frac{d-2}{2}} \le 2^{\frac{d-2}{2}}C_d\rho_h^{\frac{1}{2}} =: C'_d\rho_h^{\frac{1}{2}}, \tag{6.60}$$

and therefore for such x's

$$\mathbf{P}^{\widehat{\phi}}_{N_1}\left(\left|H^{\phi}_{\mathbb{C}_j}(x) - u\right| > \rho_h^{\frac{1}{8}}\right) \le 2\exp\left(-\frac{\rho_h^{-\frac{1}{4}}}{8C'_d}\right). \tag{6.61}$$

Therefore, by a union bound, we obtain

$$\sup_{\widehat{\phi}\in B_u}\mathbf{P}^{\widehat{\phi}}_{N_1}\left(A^{(1)^{\complement}}\right) \le N_1^d 2\exp\left(-\frac{\rho_h^{-\frac{1}{4}}}{8C'_d}\right), \tag{6.62}$$

where $N_1^d = h^{-3d}$ and therefore this term is $O(h^\alpha)$ for any $\alpha > 0$. So, in view of (6.26), we can safely focus on the second term in the right-hand side of (6.58).

For such a term, we first observe that

$$\mathbf{P}^{\widehat{\phi}}_{N_1}\left(A^{(2)^{\complement}}_{\kappa}\cap A^{(1)}\right) \,=\, \mathbf{P}^{\widehat{\phi}}_{N_1}\left(\cup_{j\in\mathcal{J}}\left\{\sum_{x\in\mathbb{B}_j}\delta_x\,>\,\kappa\right\}\cap A^{(1)}\right)$$
$$\leq \sum_{j\in\mathcal{J}}\mathbf{P}^{\widehat{\phi}}_{N_1}\left(\left\{\sum_{x\in\mathbb{B}_j}\delta_x\,>\,\kappa\right\}\cap\left\{\max_{x\in\mathbb{B}_j}\left|H^{\phi}_{\mathbb{C}_j}(x)-u\right|\,\leq\,\rho_h^{\frac{1}{8}}\right\}\right). \tag{6.63}$$

We now condition on the σ-algebra generated by $(\phi_x)_{x\in\cup_{j\in\mathcal{J}_w}\partial\mathbb{C}_j}$ by the Markov property we obtain that

$$\mathbf{P}^{\widehat{\phi}}_{N_1}\left(A^{(2)^{\complement}}_{\kappa}\cap A^{(1)}\right)\,\leq\,\sum_{j\in\mathcal{J}}\sup{}'\,\mathbf{P}^{\widehat{\phi}}_{\mathbb{C}_j}\left(\sum_{x\in\mathbb{B}_j}\delta_x\,>\,\kappa\right), \tag{6.64}$$

where $\sup'$ stands for the supremum over $\widehat{\phi}$ such that $\max_{x\in\mathbb{B}_j}|H^{\widehat{\phi}}_{\mathbb{C}_j}(x)-u|\,\leq\,\rho_h^{\frac{1}{8}}$. To bound the $\sup'$ term, we proceed as follows:

$$\mathbf{P}^{\widehat{\phi}}_{\mathbb{C}_j}\left(\sum_{x\in\mathbb{B}_j}\delta_x\,>\,\kappa\right)\,\leq\sum_{\substack{x_1,\ldots,x_\kappa\in\mathbb{B}_j:\\ x_i\neq x_{i'}\text{ for }i\neq i'}}\mathbf{P}^{\widehat{\phi}}_{\mathbb{C}_j}\left(\frac{1}{\kappa}\sum_{i=1}^{\kappa}|\phi_{x_i}|\,\leq 1\right), \tag{6.65}$$

and

$$\mathbf{P}^{\widehat{\phi}}_{\mathbb{C}_j}\left(\frac{1}{\kappa}\sum_{i=1}^{\kappa}|\phi_{x_i}|\,\leq 1\right)\,\leq\,\mathbf{P}^{\widehat{\phi}}_{\mathbb{C}_j}\left(\frac{1}{\kappa}\sum_{i=1}^{\kappa}\phi_{x_i}\,\leq 1\right)\,\leq\,\mathbf{P}^{0}_{\mathbb{C}_j}\left(\frac{1}{\kappa}\sum_{i=1}^{\kappa}\phi_{x_i}\,\leq\,1-\frac{1}{\kappa}\sum_{i=1}^{\kappa}H^{\widehat{\phi}}_{\mathbb{C}_j}(x)\right)$$
$$\leq\,\mathbf{P}^{0}_{\mathbb{C}_j}\left(\frac{1}{\kappa}\sum_{i=1}^{\kappa}\phi_{x_i}\,\leq\,1-u+\rho_h^{\frac{1}{8}}\right)\,\leq\,\mathbf{P}^{0}_{\mathbb{C}_j}\left(\frac{1}{\kappa}\sum_{i=1}^{\kappa}\phi_{x_i}\,\geq\,u-2\right), \tag{6.66}$$

where in the last step we used the symmetry and the choice of subtracting 2 is arbitrary (any number larger than 1 would do, and we have to choose h sufficiently small). It is now a matter of evaluating the variance of $\frac{1}{\kappa}\sum_{i=1}^{\kappa}\phi_{x_i}$ and we perform this estimate uniformly over the location of $x_1,\ldots,x_\kappa$. In fact, we apply Lemma 6.11 to obtain that

$$\mathrm{var}_{\mathbf{P}^0_{\mathbb{C}_j}}\left(\frac{1}{\kappa}\sum_{i=1}^{\kappa}\phi_{x_i}\right)\,\leq\,c(d)\kappa^{-1+\frac{2}{d}}\,\leq\,c(d)\kappa^{-1/3}, \tag{6.67}$$

and therefore (6.3) yields

$$\mathbf{P}^{0}_{\mathbb{C}_j}\left(\frac{1}{\kappa}\sum_{i=1}^{\kappa}\phi_{x_i}\,\geq\,u-2\right)\,=\,\mathbf{P}\left(\mathcal{N}\geq\kappa^{1/6}c(d)^{-1/2}(u-2)\right)\,\leq\,h^{\frac{\kappa^{1/3}}{2c(d)}}, \tag{6.68}$$

where $\mathcal{N}$ is a standard Gaussian random variable. With this estimate we now go back to (6.65) and we obtain

$$\mathbf{P}^{\widehat{\phi}}_{\mathbb{C}_j}\left(\sum_{x\in\mathbb{B}_j}\delta_x\,>\,\kappa\right)\,\leq\,N_0^{\kappa}h^{\frac{\kappa^{1/3}}{2c(d)}}\,\leq\,h^{\frac{\kappa^{1/3}}{3c(d)}}, \tag{6.69}$$

for h sufficiently small. In turn, this estimate yields the control (see (6.64))

$$\mathbf{P}^{\widehat{\phi}}_{N_1}\left(A^{(2)\complement}_{\kappa}\cap A^{(1)}\right)\,\le\,h^{-3d}h^{\frac{\kappa^{1/3}}{3c(d)}}\,\le\,h^{\kappa^{1/4}}\,, \tag{6.70}$$

for κ such that $9d\,c(d)\le\kappa^{\frac13}-3c(d)\kappa^{\frac14}$. Let us choose $\kappa^{\frac13}\ge 6c(d)\,\kappa^{\frac14}$, that is $\kappa\ge(6c(d))^{12}$, and, under this assumption, $9d\,c(d)\le\kappa^{\frac13}-3c(d)\kappa^{\frac14}$ is satisfied if $\kappa^{\frac13}\ge 18d\,c(d)$. Taking into account these two lower bounds on κ we see that is suffices to choose $\kappa\ge d^3 6^{12}c(d)^{12}$. The proof of Proposition 6.3 is therefore complete. □

Proof of Lemma 6.11. First of all we introduce $d_\square(x,y):=\max_{i=1,\dots,d}|x_i-y_i|\le|x-y|$ and we start from the direct consequence of (2.44): for every $x,y\in\mathbb{Z}^d$

$$G(x,y)\,\le\,\frac{\mathsf{c}_d}{(1+d_\square(x,y))^{d-2}}\,=:\,g_\square\left(d_\square(x,y)\right)\,. \tag{6.71}$$

Thanks to (6.71), it suffices to prove the statement for $\sum_{(x,y)\in B^2}g_\square(x-y)$. We then observe that for any $B\subset\mathbb{Z}^d$ with $|B|=j$ and every $z\notin B$ we have

$$\sum_{(x,y)\in(B\cup\{z\})^2}g_\square(x-y)-\sum_{(x,y)\in B^2}g_\square(x-y)\,\le\,2\sum_{i=1}^{j}g_\square(x_i)+g_\square(0)\,, \tag{6.72}$$

where $x_1,x_2,\dots$ yields a *fully* packed configuration of points around the origin. By this we mean that if $B(n)=\{-n,\dots,n\}^d$ and $A(n)=B(n)\setminus B(n-1)$, then $x_1,\dots,x_{|B(1)|}$ is an arbitrary numbering of the points in $A(1)$, $x_{|B(1)|+1},\dots,x_{|B(2)|}$ is an arbitrary numbering of the points in $A(2)$, and so on. Of course z disappears because it has been translated to the origin. We have

$$\begin{aligned}2\sum_{i=1}^{j}g_\square(x_i)\,&\le\,2c_d\sum_{\substack{m=1,2,\dots\\ m\le\frac{j^{1/d}}{2}+\frac12}}\frac{|A(m)|}{(1+m)^{d-2}}\,\le\,2^{d+1}c_d\sum_{\substack{m=1,2,\dots\\ m\le j^{1/d}}}(m+1)\,\le\\ &\le\,2^dc_d\left(j^{1/d}+2\right)^2\,\le\,2^{d+1}c_d\left(j^{2/d}+4\right)\,\le\,5\cdot2^{d+1}c_dj^{2/d}\,,\end{aligned} \tag{6.73}$$

where in the first step we have simply made $g_\square(\cdot)$ explicit, used the fact that is constant on annuli and that with j points we cannot go beyond filling $\frac{j^{1/d}}{2}+\frac12(\le j^{1/d})$ annuli. In the second step instead we used $A(n)\le 2d(2m+1)^{d-1}\le d2^d(m+1)^{d-1}$. Therefore for any B with $|B|=j$ and every $z\notin B$ we have

$$\sum_{(x,y)\in(B\cup\{z\})^2}g_\square(x-y)-\sum_{(x,y)\in B^2}g_\square(x-y)\,\le\,5\cdot2^{d+1}c_dj^{2/d}+c_d\,, \tag{6.74}$$

so that for any B with $|B|=\kappa$

$$\begin{aligned}\sum_{(x,y)\in B^2}g_\square(x-y)\,&\le\,\kappa g_\square(0)+5\cdot2^{d+1}c_d\sum_{j=1}^{\kappa-1}j^{2/d}\\ &\le\,\kappa c_d+5\cdot2^{d+1}c_d\kappa^{1+\frac2d}\,\le\,2^{d+4}c_d\kappa^{1+\frac2d}\,,\end{aligned} \tag{6.75}$$

and the proof is complete. □

7. The two dimensional case

This section is dedicated to the proof of Theorem 2.3. The first step of the proof in Section 7.1 is to establish a finite volume criterion similar to (4.11) . Then in Sections 7.2 and 7.3 we will use replica coupling arguments in the spirit of Lemma 6.6 to have a bound on the free energy of a system with finite volume.

7.1. A finite volume criterion and replica coupling in dimension two.

We want to have a criterion similar to (4.11) in two dimension but the problem we have to face is that there is no infinite volume limit for the two dimensional lattice free field. A way out is to consider the massive free field (2.48) and then find a way to compare the free energy with the original one.

We let $\mathbf{P}^{u,m}$ be the law the infinite volume limit of the massive free field with mean u and mass m (we write $\widehat{\mathbf{P}}^{u,m}$ when the variable is denoted by $\widehat{\phi}$). We define also the measure in finite volume with boundary conditions:

$$\mathbf{P}_N^{\widehat{\phi},u,m}(\mathrm{d}\phi) = \frac{1}{\mathcal{Z}_N^{\widehat{\phi},u,m}} \left(\prod_{\substack{x,y\in\Lambda_N \\ x\sim y}} \exp\left(-\frac{(\phi_x-\phi_y)^2}{4}\right)\right) \left(\prod_{x\in\Lambda_N\setminus\partial\Lambda_N} e^{-\frac{m^2}{2}(\phi_x-u)^2}\mathrm{d}\phi_x\right). \tag{7.1}$$

where

$$\mathcal{Z}_N^{\widehat{\phi}} := \int_{\mathbb{R}^{\Lambda_N\setminus\partial\Lambda_N}} \left(\prod_{\substack{x,y\in\Lambda_N \\ x\sim y}} \exp\left(-\frac{(\phi_x-\phi_y)^2}{4}\right)\right) \left(\prod_{x\Lambda_N\setminus\partial\Lambda_N} e^{-\frac{m^2}{2}(\phi_x-u)^2}\mathrm{d}\phi_x\right). \tag{7.2}$$

The particular case where $\widehat{\phi}\equiv u$ is denoted by $\mathbf{P}_N^{u,m}$. We set

$$Z_{N,h}^{\beta,\omega,\widehat{\phi},u,m} := \mathbf{E}_N^{\widehat{\phi},u,m}\left[\exp\left(\sum_{x\in\widetilde{\Lambda}_N}(\beta\omega_x-\lambda(\beta)+h)\delta_x\right)\right], \tag{7.3}$$

and denote by $Z_{N,h}^{\beta,\omega,u,m}$ the partition function corresponding to constant boundary conditions u. Similarly to Proposition 4.2, we can prove:

Proposition 7.1. *For any m and u*

$$\lim_{N\to\infty}\frac{1}{N^2}\mathbb{E}\left[\log Z_{N,h}^{\beta,\omega,u,m}\right] = \lim_{N\to\infty}\frac{1}{N^2}\mathbb{E}\widehat{\mathbf{E}}^{u,m}\left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi},u,m}\right] = \mathrm{F}(\beta,h,m,u). \tag{7.4}$$

and furthermore for any value of N

$$\frac{1}{N^d}\mathbb{E}\widehat{\mathbf{E}}^{u,m}\left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi},u,m}\right] \le \mathrm{F}(\beta,h,m,u). \tag{7.5}$$

Note that, unlike the massless ($m=0$) case, there is now a dependence on u (and on m). Now, for this criterion to be useful, we need to be able to compare $\mathrm{F}(\beta,h,m,u)$ with $\mathrm{F}(\beta,h)$. The idea for this is the following: the derivative of the massive free field measure with respect to the non massive one has the following expression

$$\frac{\mathrm{d}\mathbf{P}_N^{u,m}}{\mathrm{d}\mathbf{P}_N^{u}} = \frac{1}{W_N^m}\exp\left(\sum_{x\in\Lambda_N\setminus\partial\Lambda_N} -\frac{m^2}{2}(\phi_x-u)^2\right). \tag{7.6}$$

where

$$W_N^m = \mathbf{E}_N^u \left[-\frac{m^2}{2} \sum_{x \in \Lambda_N \setminus \partial \Lambda_N} (\phi_x - u)^2 \right] = \mathbf{E}_N \left[\exp\left(-\frac{m^2}{2} \sum_{x \in \Lambda_N \setminus \partial \Lambda_N} \phi_x^2 \right) \right] \tag{7.7}$$

Lemma 7.2. *We have*

$$\lim_{N\to\infty} \frac{1}{N^2} \log W_N^m = -\frac{1}{2} \int_{[0,1]^2} \log \left(1 + \frac{m^2}{4[\sin^2(\pi x/2) + \sin^2(\pi y/2)]} \right) \mathrm{d}x\mathrm{d}y = -f(m). \tag{7.8}$$

Around zero we have the following equivalence

$$f(m) \overset{m\searrow 0}{\sim} c_W m^2 |\log m| \,, \qquad \text{with } c_W := \frac{1}{4\pi}. \tag{7.9}$$

Furthermore we have

$$\mathrm{F}(\beta, h, m, u) \le \mathrm{F}(\beta, h) + f(m). \tag{7.10}$$

Proof. Let us start to prove (7.10) is deduced from (7.8). Because the exponential in (7.6) is always smaller than 1 we have

$$Z_{N,h}^{\beta,\omega,u,m} \le \frac{1}{W_N^m} Z_{N,h}^{\beta,\omega,u}. \tag{7.11}$$

The result is obtained by taking log, dividing by N^d and passing to the limit. The functions u_i, $i = 1, \dots, N-1$ defined

$$u_i(k) := \sqrt{\frac{2}{N}} \sin(ik\pi/N) \tag{7.12}$$

forms an orthogonal base of eigenfunctions of the one dimensional Laplacian with Dirichlet boundary conditions on $[0, N] \cap \mathbb{Z}$. We let Let $0 > -\lambda_1 > -\lambda_2 > \cdots > -\lambda_{(N-1)}$ denote the associated eigenvalues where

$$\lambda_i = 2(1 - \cos(i\pi/N)) \,. \tag{7.13}$$

We set

$$v_{i,j}(x_1, x_2) := u_i(x_1) u_j(x_2), \tag{7.14}$$

and

$$\alpha_{i,j}(\phi) := \sum_{x \in \Lambda_N \setminus \partial \Lambda_N} v_{i,j}(x) \phi_x. \tag{7.15}$$

From Parseval's formula we have

$$\sum_{x \in \Lambda_N \setminus \partial \Lambda_N} \phi_x^2 = \sum_{i,j=1}^{N-1} \alpha_{i,j}(\phi)^2. \tag{7.16}$$

Now note that when ϕ has law $\mathbf{P}_N$, the $\alpha_{i,j}(\phi)$ are independent Gaussian variables. Their variance is equal to $(\lambda_i + \lambda_j)^{-1}$. Hence

$$W_N^m := \prod_{i,j=1}^{N-1} \mathbb{E} \left[\exp\left(-\frac{m^2}{2(\lambda_i + \lambda_j)} \mathcal{N}^2 \right) \right] = \left(\prod_{i,j=1}^{N-1} \sqrt{1 + \frac{m^2}{(\lambda_i + \lambda_j)}} \right)^{-1}. \tag{7.17}$$

where $\mathcal{N}$ is a standard Gaussian. It is then standard (it is a Riemann sum) to check that

$$\frac{1}{N^2} \log W_N^m := -\frac{1}{2N^2} \sum_{i,j=1}^{N-1} \log \left(1 + \frac{m^2}{(\lambda_i + \lambda_j)} \right), \tag{7.18}$$

converges to

$$-\frac{1}{2}\int_{[0,1]^2}\log\left(1+\frac{m^2}{4[\sin^2(x\pi/2)+\sin^2(y\pi/2)]}\right)\mathrm{d}x\mathrm{d}y\,. \tag{7.19}$$

For the leading order asymptotic behavior one can restrict the domain of integration to positive x and y such that $x^2+y^2\le\varepsilon^2$, ε arbitrarily small. Passing to polar coordinates the estimate becomes rather straightforward. □

As a consequence of Lemma 7.2 and Proposition 7.1 we obtain the following finite volume criterion:

Corollary 7.3. *For every m, N and u one has*

$$\mathrm{F}(\beta,h)\ge\frac{1}{N^2}\mathbb{E}\widehat{\mathbf{E}}^{u,m}\left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi},u,m}\right]-f(m). \tag{7.20}$$

To estimate the quantity $\mathbb{E}\widehat{\mathbf{E}}^{u,m}\left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi},u,m}\right]$ we will use (as for Lemma 6.6) the following bound derived from replica-coupling. For the proof see Appendix A where the slightly more involved proof of Lemma 6.6 is given in detail.

Lemma 7.4. *We have for all $u\in\mathbb{R}$ for all $m>0$ and all boundary conditions $\widehat{\phi}$*

$$\log Z_{N,h}^{\beta,\omega,\widehat{\phi},u,m}\ge\log\mathbf{E}_N^{\widehat{\phi},u,m}\left[e^{h\sum_{x\in\widetilde{\Lambda}_N}\delta_x}\right]-\log\left\langle\exp\left(2\beta^2\sum_{x\in\widetilde{\Lambda}_N}\delta_x^{(1)}\delta_x^{(2)}\right)\right\rangle^{\otimes 2}\,, \tag{7.21}$$

where

$$\langle\,\cdot\,\rangle=\langle\,\cdot\,\rangle_{N,h,\widehat{\phi},u,m}:=\frac{\mathbf{E}_N^{\widehat{\phi},u,m}\left[\,\cdot\,\exp\left(h\sum_{x\in\widetilde{\Lambda}_N}\delta_x\right)\right]}{\mathbf{E}_N^{\widehat{\phi},u,m}\left[\exp\left(h\sum_{x\in\widetilde{\Lambda}_N}\delta_x\right)\right]}\,. \tag{7.22}$$

7.2. **A first, rough bound on the critical point (warm up argument).** In order to use (7.20) in the most efficient way, for a given β we must tune the value of m and u and N in to obtain the best possible bound.

We start by stating and proving a weaker version of Theorem 2.3 obtained by choosing $u=0$. The proof uses some of the steps that will be used for Theorem 2.3, but not all and it is considerably simpler. So it can be viewed as a warm up.

Proposition 7.5. *When $d=2$ and ω is Gaussian, for every $\beta_0>0$ there exists a constant $c>0$ such that for all $\beta<\beta_0$*

$$h_c(\beta)\;\le\;\frac{c\beta^2}{\sqrt{|\log\beta|}}\,. \tag{7.23}$$

Proof. It suffices to prove the result for β small. So let us choose $\beta>0$ and let us set $N=1/\beta$: with slight abuse we will assume $N\in\mathbb{N}$ without taking integer part. We introduce a mass m and exploit Corollary 7.3, but with $u=0$. We aim at showing that there exists $c>0$ such that if $h>c\frac{\beta^2}{\sqrt{|\log\beta|}}$ there exists m_0 such that for $m<m_0$

$$\frac{1}{N^2}\mathbb{E}\widehat{\mathbf{E}}^{0,m}\left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi},0,m}\right]-f(m)\;>\;0\,. \tag{7.24}$$

When applying Lemma 7.4 we have then to deal with two terms:

- For the first one we obtain a lower bound just by computing and using Jensen's inequality:

$$\frac{1}{N^2}\widehat{\mathbf{E}}^{0,m}\left[\log \mathbb{E}Z_{N,h}^{\beta,\omega,\widehat{\phi},0,m}\right] = \frac{1}{N^2}\widehat{\mathbf{E}}^{0,m}\log Z_{N,h}^{\widehat{\phi},0,m} \geq \frac{1}{N^2}\widehat{\mathbf{E}}^{0,m}\mathbf{E}^{\widehat{\phi},m,0}\left[h\sum_{x\in\widetilde{\Lambda}_N}\delta_x\right] \tag{7.25}$$
$$= h\mathbf{E}^{0,m}\left[\delta_0\right].$$

- For the second term we need an upper bound and we observe that

$$\frac{1}{N^2}\widehat{\mathbf{E}}^{0,m}\log\left\langle \exp\left(2\beta^2\sum_{x\in\widetilde{\Lambda}_N}\delta_x^{(1)}\delta_x^{(2)}\right)\right\rangle_{N,0,m,h;\widehat{\phi}}^{\otimes 2}$$
$$\leq \frac{\beta^2(e^2-1)}{N^2}\widehat{\mathbf{E}}^{0,m}\left\langle\sum_{x\in\widetilde{\Lambda}_N}\delta_x^{(1)}\delta_x^{(2)}\right\rangle_{N,0,m,h;\widehat{\phi}}^{\otimes 2} = \frac{\beta^2(e^2-1)}{N^2}\widehat{\mathbf{E}}^{0,m}\sum_{x\in\widetilde{\Lambda}_N}\langle\delta_x\rangle_{N,0,m,h;\widehat{\phi}}^2$$
$$\leq \frac{7\beta^2}{N^2}\widehat{\mathbf{E}}^{0,m}\sum_{x\in\widetilde{\Lambda}_N}\langle\delta_x\rangle_{N,0,m,0;\widehat{\phi}}^2, \tag{7.26}$$

where in the first inequality we have used the fact that $\beta^2\sum_{x\in\widetilde{\Lambda}_N}\delta_x^{(1)}\delta_x^{(2)}\leq 1$ and Lemma 6.8 and in the second we have used the fact that $h=o(\beta^2)$ therefore the Radon-Nykodym density (7.22) can be made arbitrarily close to 1 when β becomes small.

By separating the contribution of the boundary in the second term (i.e. the rightmost term in (7.26)), we realize that it suffices to show that

$$h\mathbf{E}^{0,m}\left[\delta_0\right] - \frac{7\beta^2}{N^2}\widehat{\mathbf{E}}^{0,m}\sum_{x\in\mathring{\Lambda}_N}\langle\delta_x\rangle_{N,0,m,0;\widehat{\phi}}^2 - \frac{7d\beta^2}{N}\mathbf{E}^{0,m}\left[\delta_0\right] - 2c_W m^2|\log m| > 0, \tag{7.27}$$

where the boundary term is irrelevant because $\frac{7d\beta^2}{N}=O(\beta^3)$, hence it is dominated by h for β small. We are playing on choosing β small, hence N large, but one should think that we have chosen β, possibly small, and then we choose m as small as we wish or need.

In particular

$$\langle\delta_x\rangle_{N,0,m,0;\widehat{\phi}} \leq \langle\delta_x\rangle_{N,0,m,0;0} = P\left(|\mathcal{N}|\leq\frac{1}{\gamma_x}\right) \leq \sqrt{\frac{2}{\pi}}\frac{1}{\gamma_x}, \tag{7.28}$$

where γ_x^2 is the variance of ϕ_x under the massive field with Dirichlet boundary conditions. And hence

$$\widehat{\mathbf{E}}^{0,m}\langle\delta_x\rangle_{N,0,m,0;\widehat{\phi}}^2 \leq \sqrt{\frac{2}{\pi}}\frac{1}{\gamma_x}\widehat{\mathbf{E}}^{0,m}\langle\delta_x\rangle_{N,0,m,0;\widehat{\phi}} = \sqrt{\frac{2}{\pi}}\frac{1}{\gamma_x}\mathbf{E}^{0,m}[\delta_0] \tag{7.29}$$

By choosing m sufficiently small we can say that γ_x^2 is bounded below by the variance in the $m=0$ case times a positive number smaller than one, that is by (2.46)

$$\gamma_x^2 \geq \frac{1}{5\pi}\log d_N(x), \tag{7.30}$$

at least if $d_N(x) := \mathrm{dist}(x, \partial\Lambda_N)$ is larger than a fixed positive number d_0. It is at this stage more practical to lose track of the constants and choose a $c' > 0$ such that $\gamma_x^2 \geq c' \log(d_N(x)+1)$ for every $x \in \mathring{\Lambda}_N$. Therefore

$$\sqrt{\frac{2}{\pi}}\frac{7\beta^2}{N^2}\mathbf{E}^{0,m}\left[\delta_0\right]\sum_{x\in\mathring{\Lambda}_N}\frac{1}{\gamma_x} \leq 7\sqrt{\frac{2}{c'\pi}}\left(\frac{1}{N^2}\sum_{x\in\mathring{\Lambda}_N}\frac{1}{\sqrt{\log(d_N(x)+1)}}\right)\beta^2\mathbf{E}^{0,m}\left[\delta_0\right] \leq c''\frac{\beta^2}{\sqrt{|\log\beta|}}\mathbf{E}^{0,m}\left[\delta_0\right], \quad (7.31)$$

where c'' is a positive constant that one can easily express in terms of c'.

Going back to (7.27) we therefore see that it suffices to prove that

$$\left(c - c'' - 7d\beta\sqrt{|\log\beta|}\right)\frac{\beta^2}{\sqrt{|\log\beta|}}\mathbf{E}^{0,m}\left[\delta_0\right] - 2c_W m^2|\log m| > 0\,, \quad (7.32)$$

and it is clearly necessary to choose $c > c''$, which we do, hence $c - c'' - 7d\beta\sqrt{|\log\beta|} > 0$ for β suitably small. It is now clear that if $\mathbf{E}^{0,m}\left[\delta_0\right] \gg m^2|\log m|$ we are done. But

$$\mathbf{E}^{0,m}\left[\delta_0\right] = P\left(|\mathcal{N}| \leq \frac{1}{\sigma_m}\right) \quad (7.33)$$

where σ_m^2 is the variance of the infinite volume massive field, which satisfies

$$\sigma_m^2 = \int_0^\infty e^{-m^2 t}P^0\left(X(t)=0\right)\mathrm{d}t \overset{m\searrow 0}{\sim} \frac{1}{2\pi}|\log m|\,, \quad (7.34)$$

where the asymptotic equivalence is a direct consequence of the Local Central Limit Theorem [37, Th. 2.5.6]: $P^0(X(t)=0) \sim (4\pi t)^{-1}$ as t tends to infinity (recall the speed factor 4 with respect to [37]). We therefore see $\mathbf{E}^{0,m}\left[\delta_0\right]$ is bounded below by $1/\sqrt{|\log m|}$ (times a positive constant) and the proof is complete. □

7.3. **Proof of Theorem 2.3.** For the proof we choose an arbitrary (small) $\beta > 0$ and $N = 1/\beta$: again, with slight abuse we will assume $N \in \mathbb{N}$ without taking integer part. We will then make estimates by introducing a massive field: the mass m will be taken to go to zero and the height of the field u will be a function of m, see (7.39) below, that tends to infinity as $m \searrow 0$. Our estimates correspond to taking $m \searrow 0$ first and then $\beta \searrow 0$. Let us focus first on choosing the mean height $u = u_m$ of the massive field that we intend to exploit.

Setting the parameters of the massive field (mass and height). Let us start by recalling the behavior of the the variance $\sigma^2 = \sigma_m^2$ of the infinite volume massive field (7.34). We now assume that u_m is such that $\lim_{m\to 0} u_m/\sigma_m^2 = \sqrt{8\pi} =: C$. A precise choice of u_m is made in (7.39) below but for now we need neither this expression nor the precise value of the positive constant C. To make formulas lighter we write $\sigma = \sigma_m$ and $u = u_m$. We then compute

$$\begin{aligned}\mathbf{E}^{u,m}[\delta_0] &= \int_{-2}^0 \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-u+1)^2}{2\sigma^2}\right)\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(u-1)^2}{2\sigma^2}\right)\int_{-2}^0\exp\left(\frac{x(u-1)}{\sigma^2} - \frac{x^2}{2\sigma^2}\right)\mathrm{d}x\,,\end{aligned} \quad (7.35)$$

and since $\lim_{m\downarrow 0}(u-1)/\sigma^2 = C$ we readily see that the integral in the last line converges to

$$\int_{-2}^{0} \exp(xC)\mathrm{d}x = (1-e^{-2C})/C.$$

On the other hand we see that

$$\exp\left(-\frac{(u-1)^2}{2\sigma^2}\right) \overset{m\searrow 0}{\sim} \exp\left(-\frac{u^2}{2\sigma^2}\right)\exp(C)\,, \tag{7.36}$$

so that we obtain

$$\mathbf{E}^{u,m}[\delta_0] \overset{m\searrow 0}{\sim} \exp\left(-\frac{u^2}{2\sigma^2}\right)\frac{\sinh(C)}{\sqrt{\frac{\pi}{2}}C}\frac{1}{\sigma} \;\sim\; \exp\left(-\frac{u^2}{2\sigma^2}\right)\frac{\sinh(C)}{C}\frac{2}{\sqrt{|\log m|}}\,. \tag{7.37}$$

We now choose $u = u_m$ such that

$$\frac{1}{\sqrt{|\log m|}}\exp\left(-\frac{u^2}{2\sigma^2}\right) \;=\; C'm^2|\log m|^2\,, \qquad \text{for } C' = \frac{C}{2\sinh(C)}\,. \tag{7.38}$$

Let us point out that, because of (7.34), the choices (7.38) and $u/\sigma^2 \sim C$ force $C = \sqrt{8\pi}$.

Remark 7.6. *The choice* (7.38) *is linked to the fact that in the replica argument this term correspond to the energy gain and it needs to beat the loss due to the presence of the term* $f(m) \sim c_W m^2|\log m|$ *in the free energy lower bound* (7.20)*.*

Here is a slightly more explicit expression for u

$$\frac{u^2}{2\sigma^2} \;=\; 2|\log m| - \frac{5}{2}\log|\log m| - \log C'\,. \tag{7.39}$$

For the ease of the reader we collect here the asymptotic behaviors ($m \searrow 0$)

$$u \;\sim\; \frac{\sqrt{2}}{\sqrt{\pi}}|\log m|\,, \qquad \sigma \sim \frac{1}{\sqrt{2\pi}}\sqrt{|\log m|}\,, \qquad \frac{u}{\sigma^2} \;\sim\; \sqrt{8\pi}\,. \tag{7.40}$$

Replica coupling estimates. Recall that $N = 1/\beta$. We aim at showing that if $h = \beta^b$, any $b < 3$, there exists a β_0 such that for $\beta < \beta_0$ there exists m_0 such that for $m < m_0$

$$\frac{1}{N^2}\mathbb{E}\widehat{\mathbf{E}}^{u,m}\left[\log Z_{N,h}^{\beta,\omega,\widehat{\phi},u,m}\right] - f(m) \;>\; 0\,. \tag{7.41}$$

Using Lemma 7.4 we see that it suffices to show that

$$\frac{1}{N^2}\widehat{\mathbf{E}}^{u,m}\left[\log \mathbb{E}Z_{N,h}^{\beta,\omega,\widehat{\phi},u,m}\right] - \frac{1}{N^2}\widehat{\mathbf{E}}^{u,m}\log\left\langle \exp\left(2\beta^2\sum_{x\in\widetilde{\Lambda}_N}\delta_x^{(1)}\delta_x^{(2)}\right)\right\rangle^{\otimes 2} - f(m) > 0\,. \tag{7.42}$$

From the choice of u (cf. (7.37) and (7.38)), once the choice of β (hence of N and h) is made, for m sufficiently small we have

$$\frac{1}{N^2}\widehat{\mathbf{E}}^{u,m}\left[\log \mathbb{E}Z_{N,h}^{\beta,\omega,\widehat{\phi},u,m}\right] - f(m) \;\ge\; h\,\mathbf{E}^{u,m}\left[\delta_0\right] - f(m) \;\ge\; \frac{h}{2}m^2|\log m|^2\,, \tag{7.43}$$

where the first inequality is Jensen's. Therefore we are left with estimating

$$\frac{1}{N^2}\widehat{\mathbf{E}}^{u,m}\log\left\langle \exp\left(2\beta^2\sum_{x\in\widetilde{\Lambda}_N}\delta_x^{(1)}\delta_x^{(2)}\right)\right\rangle^{\otimes 2}_{N,u,m,h;\widehat{\phi}}\,, \tag{7.44}$$

from above. The first point to remark is that, thanks to the choice of N, the argument of the exponential is bounded by 2. This way at the cost of loosing a multiplicative constant, the exponential and log essentially cancels. More precisely, by Lemma 6.8, the expression in (7.44) is bounded by e^2-1 times

$$\frac{2\beta^2}{N^2}\widehat{\mathbf{E}}^{u,m}\left\langle \sum_{x\in\widetilde{\Lambda}_N}\delta_x^{(1)}\delta_x^{(2)}\right\rangle^{\otimes 2}_{N,u,m,h;\widehat{\phi}} = \frac{2\beta^2}{N^2}\sum_{x\in\widetilde{\Lambda}_N}\widehat{\mathbf{E}}^{u,m}\left(\langle\delta_x\rangle^2_{N,u,m,h;\widehat{\phi}}\right). \tag{7.45}$$

With our choice of $h=o(\beta^2)$ we see that that the Radon-Nikodyn derivative in (7.22) is bounded above and below uniformly in β small and we can even replace $\langle\cdot\rangle_{N,u,m,h;\widehat{\phi}}$ with the original measure at the expense of a multiplicative constant that can be chosen arbitrarily close to 1. Hence

$$\frac{1}{N^2}\widehat{\mathbf{E}}^{u,m}\log\left\langle \exp\left(2\beta^2\sum_{x\in\widetilde{\Lambda}_N}\delta_x^{(1)}\delta_x^{(2)}\right)\right\rangle^{\otimes 2}_{N,u,m,h;\widehat{\phi}} \geq \frac{2(e^2-1)\beta^2}{N^2}\sum_{x\in\widetilde{\Lambda}_N}\widehat{\mathbf{E}}^{u,m}\left(\mathbf{E}_N^{u,m,\widehat{\phi}}\delta_x\right)^2, \tag{7.46}$$

where of course $\beta^2/N^2=\beta^4$.

Lemma 7.7. *For any $\varepsilon>0$, there exists a constant $C_\varepsilon>0$ such that for all N, for all $m\geq m_0(N)$ and all $x\in\widetilde{\Lambda}_N$*

$$\widehat{\mathbf{E}}^{u_m,m}\left(\mathbf{E}_N^{u_m,m,\widehat{\phi}}\delta_x\right)^2 \leq C_\varepsilon m^2|\log m|^2\left(d_N(x)+1\right)^{-1+\varepsilon}, \tag{7.47}$$

where here $d_N(x)$ is the distance of x to $\partial\Lambda_N$.

In view of (7.43) and (7.46), Lemma 7.7 directly implies the result we are after. In fact for $b=3-2\varepsilon$ we have

$$\begin{aligned}\frac{1}{N^2}\widehat{\mathbf{E}}^{u,m}\left[\log\mathbb{E}Z^{\beta,\omega,\widehat{\phi},u,m}_{N,h}\right] - \frac{1}{N^2}\widehat{\mathbf{E}}^{u,m}\log\left\langle\exp\left(2\beta^2\sum_{x\in\widetilde{\Lambda}_N}\delta_x^{(1)}\delta_x^{(2)}\right)\right\rangle^{\otimes 2}_{N,u,m,h;\widehat{\phi}} - f(m)\\ \geq \frac{h}{2}m^2|\log m|^2 - 2(e^2-1)\frac{\beta^2}{N^2}\sum_{x\in\widetilde{\Lambda}_N}\widehat{\mathbf{E}}^{u,m}\left(\mathbf{E}_N^{u,m,\widehat{\phi}}\delta_x\right)^2\\ > \left(\frac{1}{2}\beta^{3-2\varepsilon}-C_{\varepsilon,d}\beta^{3-\varepsilon}\right)m^2|\log m|^2,\end{aligned} \tag{7.48}$$

where $C_{\varepsilon,d}$ is a positive constant, where in the last step we have used the bound

$$\sum_{x\in\widetilde{\Lambda}_N}\frac{1}{\left(d_N(x)+1\right)^{1-\varepsilon}} \leq 4N\sum_{n=1}^{N}\frac{1}{n^{1-\varepsilon}} \leq \frac{4}{\varepsilon}N^{1+\varepsilon} = \frac{4}{\varepsilon}\beta^{-1-\varepsilon}. \tag{7.49}$$

Since for β sufficiently small the right-hand side in (7.48) is positive, we are done. □

Proof of Lemma 7.7. First of all remark that the result is trivial for $d_N(x)=O(1)$ because by (7.37) and (7.38)

$$\widehat{\mathbf{E}}^{u_m,m}\left(\mathbf{E}_N^{u_m,m,\widehat{\phi}}\delta_x\right)^2 \leq \widehat{\mathbf{E}}^{u_m,m}\left(\mathbf{E}_N^{u_m,m,\widehat{\phi}}\delta_x\right) = \mathbf{E}^{u_m,m}\delta_x = O(m^2|\log m|^2). \tag{7.50}$$

Hence, it suffices to prove (7.47) for x such that $d_N(x)$ is larger than a constant that may depend on ε. At this point we note that the variable ϕ_x can be written the sum of two independent Gaussian variables

$$\varphi_x := \mathbf{E}_N^{u,m,\widehat{\phi}}\phi_x \qquad \text{and} \qquad \psi_x := \phi_x - \varphi_x\,. \tag{7.51}$$

The variance of ψ_x tends, as $m \searrow 0$, to $G_{\Lambda_N}(x,x)$. As from (2.46) we have

$$G_{\Lambda_N}(x,x) \ge (1-(\varepsilon/2))\frac{1}{2\pi}\log d_N(x)$$

for x away from the boundary (how far depends just on ε and not on N). Hence for m sufficiently small we have

$$\eta^2 := \mathrm{Var}(\psi_x) \ge (1-\varepsilon)\frac{1}{2\pi}\log(x)\,.$$

Since we are performing the estimates by sending first m to zero we are effectively performing our estimates in the regime

$$\sigma^2 - \eta^2 \gg \eta^2 \gg 1\,. \tag{7.52}$$

Recall that $g_{\sigma^2}(\cdot)$ is the density of a centered Gaussian variable with variance σ^2. One has

$$\widehat{\mathbf{E}}^{u,m}\left(\mathbf{E}_N^{u,m,\widehat{\phi}}\delta_x\right)^2 = \int_{-\infty}^{\infty} g_{\sigma^2-\eta^2}(s)\left(\int_{u-1-s}^{u+1-s} g_{\eta^2}(t)\mathrm{d}t\right)^2 \mathrm{d}s\,, \tag{7.53}$$

and to this expression we can directly apply the next lemma.

Lemma 7.8. *Recall that $\sigma = \sigma_m$. For every $\varepsilon \in (0,1)$ there exists $\eta_0 > 0$ such that for every $\eta > \eta_0$ there exists $m_0 > 0$ such that for every $m \in (0, m_0)$ we have $\sigma > \eta$ and*

$$\int_{-\infty}^{\infty} g_{\sigma^2-\eta^2}(s)\left(\int_{u-1-s}^{u+1-s} g_{\eta^2}(t)\mathrm{d}t\right)^2 \mathrm{d}s \le 2C'm^2|\log m|^2\exp\left(-2\pi(1-\varepsilon)\eta^2\right)\,. \tag{7.54}$$

It is now just a matter of observing that

$$\exp\left(-2\pi(1-\varepsilon)\eta^2\right) \le d_N(x)^{-(1-\varepsilon)^2}\,, \tag{7.55}$$

and we are done. □

7.4. **Proof of Lemma 7.8.** In (7.54) we consider separately the case of s larger or smaller than $u-\eta$: note that $u-\eta \sim u$ in the limit that suffices to consider to establish the result, i.e. $\sigma \gg \eta \gg 1$, even if, at this stage, we cannot replace u with $u-\eta$. The choice of $u-\eta$ is arbitrary in the sense that $u - c\eta$, $c \ge 1/\sqrt{2}$, would do.

We start with considering $s \ge u-\eta$ and we have

$$\int_{u-\eta}^{\infty} g_{\sigma^2-\eta^2}(s)\left(\int_{u-1-s}^{u+1-s} g_{\eta^2}(t)\mathrm{d}t\right)^2 \mathrm{d}s \le \int_{u-\eta}^{\infty} g_{\sigma^2-\eta^2}(s)\mathrm{d}s = P\left(\mathcal{N} \ge \frac{u-\eta}{\sqrt{\sigma^2-\eta^2}}\right). \tag{7.56}$$

Since $P(Z \ge t) \le \frac{1}{\sqrt{2\pi}t}\exp(-t^2/2)$ for every $t > 0$ we can continue the chain of inequality in (7.56) by

$$P\left(\mathcal{N} \ge \frac{u-\eta}{\sqrt{\sigma^2-\eta^2}}\right) \le \sqrt{\frac{2}{\pi}}\frac{\sigma}{u}\exp\left(-\frac{(u-\eta)^2}{2(\sigma^2-\eta^2)}\right)\,, \tag{7.57}$$

for m such that $u \geq 2\eta$. By recalling that $\lim_m u/\sigma^2$ is a positive constant we see that

$$\frac{(u-\eta)^2}{2(\sigma^2-\eta^2)} \;=\; \frac{u^2}{2\sigma^2} - \frac{\eta u}{2\sigma^2} + \frac{u^2\eta^2}{2\sigma^4} + O\left(\frac{\eta^4}{\sigma^4}\right). \tag{7.58}$$

More precisely, since $\lim_m u/\sigma^2 = \sqrt{8\pi}$, we see that for every $q < 1$, but we choose $q \in (1/2, 1)$, we have that

$$-\frac{\eta u}{2\sigma^2} + \frac{u^2\eta^2}{2\sigma^4} \;\geq\; 4\pi q\eta^2\,, \tag{7.59}$$

for η sufficiently large and m sufficiently small. Therefore, by choosing if needed m even smaller so that $\sigma/u \sim 1/(2\sqrt{|\log m|})$ is smaller than $1/\sqrt{|\log m|}$ and that the $O[(\eta/\sigma)^4]$ term in (7.58) can be absorbed by replacing q with a smaller value still larger than $\frac{1}{2}$, we have

$$\begin{aligned}\int_{u-\eta}^{\infty} g_{\sigma^2-\eta^2}(s)\left(\int_{u-1-s}^{u+1-s} g_{\eta^2}(t)\mathrm{d}t\right)^2 \mathrm{d}s \;&\leq\; \frac{1}{\sqrt{|\log m|}}\exp\left(-\frac{u^2}{2\sigma^2}\right)\exp\left(-4\pi q\eta^2\right)\\ &=\; C'm^2|\log m|^2\exp\left(-4\pi q\eta^2\right),\end{aligned} \tag{7.60}$$

where in the last step we have used (7.38). In view of what we need to establish, that is (7.54), we can move to look for an upper bound for the case $s \leq u-\eta$.

For $s \leq u-\eta$, we use instead

$$\begin{aligned}\int_{u-1-s}^{u+1-s} g_{\eta^2}(t)\mathrm{d}t \;&=\; P\left(\mathcal{N} \in \frac{u-s-1}{\eta} + \left[0, \frac{2}{\eta}\right]\right) \;\leq\; P\left(\mathcal{N} \geq \frac{u-s-1}{\eta}\right)\\ &\leq \frac{1}{\sqrt{2\pi}}\frac{\eta}{u-s-1}\exp\left(-\frac{1}{2}\left(\frac{u-s-1}{\eta}\right)^2\right) \;\leq\; \frac{\eta}{u-s}\exp\left(-\frac{1}{2}\left(\frac{u-s}{\eta_\varepsilon}\right)^2\right),\end{aligned} \tag{7.61}$$

where $\eta_\varepsilon := \eta/(1-\varepsilon)$, $\varepsilon \in (0, 1/20)$, and we have used the bounds on the distribution of $\mathcal{N}$ recalled just after (7.56) (we are choosing m and $1/\eta$ sufficiently small). Hence we have

$$\int_{-\infty}^{u-\eta} g_{\sigma^2-\eta^2}(s)\left(\int_{u-1-s}^{u+1-s} g_{\eta^2}(t)\mathrm{d}t\right)^2 \mathrm{d}s \;\leq\; \frac{1}{(2\pi)^{3/2}}\frac{\eta^2}{\sigma_1}\int_{-\infty}^{u-\eta}\frac{\exp\left(-\frac{s^2}{2\sigma_1^2} - \frac{(u-s)^2}{\eta_\varepsilon^2}\right)}{(u-s)^2}\mathrm{d}s\,, \tag{7.62}$$

where we have introduced

$$\sigma_1 := \sqrt{\sigma^2-\eta^2}.$$

Recall that we look for a result in a regime in which $\sigma_1 \sim \sigma$. We now reconstruct the square in the term in the exponential and we obtain that the right-hand side of (7.62) equals

$$\frac{1}{(2\pi)^{3/2}}\frac{\eta^2}{\sigma_1}\exp\left(-\frac{u^2}{\eta_\varepsilon^2+2\sigma_1^2}\right)\int_{-\infty}^{u-\eta}\frac{1}{(u-s)^2}\exp\left(-\frac{\eta_\varepsilon^2+2\sigma_1^2}{2\eta_\varepsilon^2\sigma_1^2}\left(s - \frac{2u\sigma_1^2}{\eta_\varepsilon^2+2\sigma_1^2}\right)^2\right)\mathrm{d}s\,, \tag{7.63}$$

Let us introduce

$$a_{m,\eta_\varepsilon} \;:=\; u - \frac{2u\sigma_1^2}{\eta_\varepsilon^2+2\sigma_1^2} \overset{m\searrow 0}{\sim} \sqrt{2\pi}\eta_\varepsilon^2\,, \tag{7.64}$$

so that the integral term in (7.63) can be rewritten as

$$\int_{-\infty}^{a_{m,\eta_\varepsilon}-\eta}\frac{1}{(a_{m,\eta_\varepsilon}-s)^2}\exp\left(-\frac{\eta_\varepsilon^2+2\sigma_1^2}{2\eta_\varepsilon^2\sigma_1^2}s^2\right)\mathrm{d}s\,, \tag{7.65}$$

and since $\frac{\eta_\varepsilon^2+2\sigma_1^2}{2\eta_\varepsilon^2\sigma_1^2} \ge \frac{1}{\eta_\varepsilon^2}$ we see that we can bound the expression in (7.65) by

$$\frac{1}{\eta_\varepsilon}\int_{-\infty}^{\frac{a_{m,\eta_\varepsilon}}{\eta_\varepsilon}-1+\varepsilon} \frac{1}{\left(\frac{a_{m,\eta_\varepsilon}}{\eta_\varepsilon}-s\right)^2}\exp\left(-s^2\right)\mathrm{d}s \;\sim\; \sqrt{\pi}\frac{\eta_\varepsilon}{(a_{m,\eta_\varepsilon})^2} \;\sim\; \frac{1}{2\sqrt{\pi}\,\eta_\varepsilon^3}\,, \tag{7.66}$$

where the asymptotic limit is as $m \searrow 0$ and then $\eta \to \infty$. Therefore, by (7.64)–(7.66) and the choice of $\varepsilon < 1/20$, we see that the integral term in (7.63) is smaller than $1/(4\eta^3)$ for suitably small m and $1/\eta$. Therefore, going back to (7.62) and (7.63) we see that

$$\int_{-\infty}^{u-\eta} g_{\sigma^2-\eta^2}(s)\left(\int_{u-1-s}^{u+1-s} g_{\eta^2}(t)\mathrm{d}t\right)^2 \mathrm{d}s \;\le\; \frac{1}{4(2\pi)^{3/2}}\frac{1}{\eta\sigma_1}\exp\left(-\frac{u^2}{\eta_\varepsilon^2+2\sigma_1^2}\right)\,, \tag{7.67}$$

and in turn, with $c_\varepsilon := 2-(1-\varepsilon^{-2} \ge 1-3\varepsilon$ ($\varepsilon < 1/20$), we have

$$\begin{aligned}\exp\left(-\frac{u^2}{\eta_\varepsilon^2+2\sigma_1^2}\right) \;&=\; \exp\left(-\frac{u^2}{2\sigma^2}\right)\exp\left(-\frac{c_\varepsilon\eta^2u^2}{2\sigma^2\left(2\sigma^2-c_\varepsilon\eta^2\right))}\right)\\ &=\; C'm^2|\log m|^{5/2}\exp\left(-\frac{c_\varepsilon\eta^2u^2}{2\sigma^2(2\sigma^2-c_\varepsilon\eta^2)}\right) \le C'm^2|\log m|^{5/2}e^{-(1-3\varepsilon)\eta^2\left(\frac{u}{2\sigma^2}\right)^2}\,,\end{aligned} \tag{7.68}$$

where we have used (7.38). Finally, since $(\frac{u}{2\sigma^2})^2 > 2\pi(1-\varepsilon)$ for m small and recalling also (7.34), going back to (7.67) we obtain

$$\int_{-\infty}^{u-\eta} g_{\sigma^2-\eta^2}(s)\left(\int_{u-1-s}^{u+1-s} g_{\eta^2}(t)\mathrm{d}t\right)^2 \mathrm{d}s \;\le\; C'm^2|\log m|^2\frac{\exp(-2\pi(1-4\varepsilon)\eta^2)}{\eta}\,, \tag{7.69}$$

which, together with (7.60), yields (7.54) and the proof of Lemma 7.8 is complete. □

Remark 7.9. *The warm up argument of Section 7.2 does not yield interesting information in the case of the co-membrane, simply because the probability of visiting the lower half plane is* $1/2$ *for a centered field and the quadratic term in the replica computation is too large. But the arguments of Section 7.3 have a chance to be generalized because we introduce a shift in the field that makes the probability of visiting the lower half plane small. And they do generalize, giving the analog of Theorem 2.3 for the co-membrane model: let us quickly see why. The estimate* (7.35) *becomes*

$$\mathbf{E}^{u,m}[\Delta_0] \;=\; \int_{-\infty}^{0}\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-u)^2}{2\sigma^2}}\mathrm{d}x \;=\; \mathbf{P}\left(\sigma\mathcal{N} > u\right) \sim \frac{1}{\sqrt{2\pi}}\frac{\sigma}{u}e^{-\frac{u^2}{2\sigma^2}}\,, \tag{7.70}$$

and $u/\sigma \sim 2\sqrt{|\log m|}$ *and, apart for the value of* C' *(in* (7.38)*), we have the analog of* (7.37)*. In the remainder of the proof in reality we estimate the probabilities either by replacing* $[-1,1]$ *with* $\mathbb{R}$*, this is the case in* (7.56)*, or with* $(-\infty,1]$*, see* (7.61)*. Therefore the proof can be adapted to the co-membrane set-up.*

Remark 7.10. *To complete the discussion in Section 2.3 we observe that a lower bound on the non disordered free energy* $\mathrm{F}(h)$ *for* $d=2$ *can be obtained by first localizing the* ϕ *field, by introducing a mass, so that we can apply the approach in Section 2.3 and then by optimizing the choice of* m *as a function of* h*. More precisely, by Corollary 7.3 (for* $\beta = 0$*) and by applying the same argument as for the lower bound in Section 2.3 we obtain*

$$\mathrm{F}(h) \;\ge\; h\sqrt{\frac{2}{\pi}}\frac{1}{\sigma_m} - f(m)\,. \tag{7.71}$$

The two terms in the right-hand side are then estimated for m small by (7.34) *and Lemma 7.2 to obtain that for every positive $c < 1$ (that can be chosen arbitrarily close to* 1*) we have*

$$\mathrm{F}(h) \,\geq\, \frac{2}{c}\frac{h}{\sqrt{|\log m|}} - 2\pi c\, m^2 |\log m|\,, \tag{7.72}$$

for m smaller than a constant that depends on the choice of c. It is now sufficient to choose m equal to h to a power larger than $1/2$, for example $m = h^{3/4}$ to obtain that there exists $C_2 > 0$ such that

$$\mathrm{F}(h) \,\geq\, C_2 \frac{h}{\sqrt{|\log h|}}\,, \tag{7.73}$$

which should be compared to (2.15)*.*

Appendix A. Replica coupling: proof of Lemma 6.6

The argument follows closely the main argument in [39]. We do not detail the proof of Lemma 7.4 which is extremely similar (and simpler). Let us fix $\widehat{\phi}$. Given an event $A \subset \mathbb{R}^{\mathbb{Z}^d}$ (in the specific application A is measurable with respect to $\{\phi_x : x \in \mathring{\Lambda}_N\}$, but at this stage we just require $\mathbf{P}_N^{\widehat{\phi}}(A) > 0$), we write

$$\begin{aligned} F_N(\beta, h; A) \,&:=\, \frac{1}{N^d}\mathbb{E}\left[\log \mathbf{E}_N^{\widehat{\phi}}\left[\exp\left(\sum_{x\in\Lambda}\left(\beta\omega_x - \frac{\beta^2}{2} + h\right)\delta_x\right);\, A\right]\right] \\ &=\, \frac{1}{N^d}\log \mathbf{E}_N^{\widehat{\phi}}\left[\exp\left(h\sum_{x\in\Lambda}\delta_x\right)\right] + R_{N,h}(\beta; A)\,, \end{aligned} \tag{A.1}$$

where $\Lambda \subset \Lambda_N$ (for this proof it suffices to consider $\Lambda = \Lambda_{N_1,N_0}$, but this specific choice is irrelevant at this stage) and

$$R_{N,h}(\beta; A) \,:=\, \frac{1}{N^d}\mathbb{E}\left[\log\left\langle \exp\left(\sum_{x\in\Lambda}\left(\beta\omega_x - \frac{\beta^2}{2}\right)\delta_x\right);\, A\right\rangle_{N,h;\widehat{\phi}}\right]. \tag{A.2}$$

Of course

$$\langle\cdot\rangle_{N,h;\widehat{\phi}} \,:=\, \frac{\mathbf{E}_N^{\widehat{\phi}}\left[\,\cdot\, \exp\left(h\sum_{x\in\Lambda}\delta_x\right)\right]}{\mathbf{E}_N^{\widehat{\phi}}\left[\exp\left(h\sum_{x\in\Lambda}\delta_x\right)\right]}\,. \tag{A.3}$$

By (Gaussian) integration by parts – the basic formula being $E[\mathcal{N}F(\mathcal{N})] = E[F'(\mathcal{N})]$ which holds for $F \in C^1$ with a suitable growth condition at $\pm\infty$ – we obtain

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}t}\left(-R_{N,h}(\sqrt{t}\beta; A)\right) \,= \\ &\qquad \frac{\beta^2}{2N^d}\sum_{x\in\Lambda}\mathbb{E}\left[\left(\frac{\left\langle \delta_x \exp\left(\sum_{x\in\Lambda}\left(\sqrt{t}\beta\omega_x - t\frac{\beta^2}{2}\right)\delta_x\right);\, A\right\rangle_{N,h;\widehat{\phi}}}{\left\langle \exp\left(\sum_{x\in\Lambda}\left(\sqrt{t}\beta\omega_x - t\frac{\beta^2}{2}\right)\delta_x\right);\, A\right\rangle_{N,h;\widehat{\phi}}}\right)^2\right]. \end{aligned} \tag{A.4}$$

At this point we introduce

$$\psi_{N,h}(t, \lambda, \beta; A) \,:=\, \frac{1}{2N^d}\mathbb{E}\log\left\langle \exp\left(H_N\right);\, A^2\right\rangle_{N,h;\widehat{\phi}}^{\otimes 2}\,, \tag{A.5}$$

where

$$H_N := \sum_{x\in\Lambda}\left(\sqrt{t}\beta\omega_x - t\frac{\beta^2}{2}\right)\left(\delta_x^{(1)}+\delta_x^{(2)}\right) + \lambda\beta^2\sum_{x\in\Lambda}\delta_x^{(1)}\delta_x^{(2)}\,. \tag{A.6}$$

In particular

$$\psi_{N,h}(0,\lambda,\beta;A) \;=\; \frac{1}{2N^d}\mathbb{E}\log\left\langle \exp\left(\lambda\beta^2\sum_{x\in\Lambda}\delta_x^{(1)}\delta_x^{(2)}\right);\, A^2\right\rangle_{N,h;\widehat{\phi}}^{\otimes 2}\,, \tag{A.7}$$

and

$$\psi_{N,h}(t,0,\beta;A) \;=\; R_{N,h}(\sqrt{t}\beta;A)\,. \tag{A.8}$$

Again by integration by parts we obtain

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\psi_{N,h,u}(t,\lambda,\beta;A) \;\le\; \frac{\beta^2}{2N^d}\mathbb{E}\widehat{\mathbf{E}}^u\left[\frac{\sum_{x\in\Lambda}\delta_x^{(1)}\delta_x^{(2)}\left\langle \exp\left(H_N\right);\, A^2\right\rangle_{N,h;\widehat{\phi}}^{\otimes 2}}{\left\langle \exp\left(H_N\right);\, A^2\right\rangle_{N,h;\widehat{\phi}}^{\otimes 2}}\right]\\ = \frac{\mathrm{d}}{\mathrm{d}\lambda}\psi_{N,h}(t,\lambda,\beta;A)\,,\end{aligned} \tag{A.9}$$

where the inequality comes from neglecting the (negative) term coming from the derivative of the denominator. We therefore see that $(\mathrm{d}/\mathrm{d}s)\psi_{N,h}(t-s,\lambda+s,\beta;A)\ge 0$ for $s\in[0,t]$ and therefore

$$\psi_{N,h}(t,\lambda,\beta;A) \;\le\; \psi_{N,h}(0,\lambda+t,\beta;A)\,. \tag{A.10}$$

Now we go back to (A.4) and we remark that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(-R_{N,h}(\sqrt{t}\beta;A)\right) \;=\; \frac{\mathrm{d}}{\mathrm{d}\lambda}\psi_{N,h}(t,\lambda,\beta;A)\Big|_{\lambda=0}\,, \tag{A.11}$$

and for $t\in[0,1]$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\lambda}\psi_{N,h}(t,\lambda,\beta;A)\Big|_{\lambda=0} &\le \frac{\psi_{N,h}(t,2-t,\beta;A)-R_{N,h}(\sqrt{t}\beta;A)}{2-t}\\ &\le \psi_{N,h}(0,2,\beta;A)-R_{N,h}(\sqrt{t}\beta;A)\,,\end{aligned} \tag{A.12}$$

where the first bound follows by convexity of $\psi_{N,h}(t,\cdot,\beta;A)$ and (A.8), while for the second we use $2-t\ge 1$, non negativity of the numerator and (A.10). By (A.4) and by integrating the differential inequality obtained by combining (A.11) and (A.12) one obtains

$$\frac{1}{N^d}\log\langle \mathbf{1}_A\rangle_{N,h,\widehat{\phi}} \ge R_{N,h}(\beta;A) \ge \frac{1}{N^d}\log\langle \mathbf{1}_A\rangle_{N,h,\widehat{\phi}} - (e-1)\,\psi_{N,h}(0,2,\beta;A)\,. \tag{A.13}$$

Therefore, since

$$\langle \mathbf{1}_A\rangle_{N,h,\widehat{\phi}} \;=\; \frac{\mathbf{E}_N^{\widehat{\phi}}\left[\exp\left(h\sum_{x\in\Lambda}\delta_x\right);\, A\right]}{\mathbf{E}_N^{\widehat{\phi}}\left[\exp\left(h\sum_{x\in\Lambda}\delta_x\right)\right]}\,, \tag{A.14}$$

by putting (A.1) and (A.13) together we obtain

$$F_N(\beta,h;A) \quad\ge\quad \frac{1}{N^d}\log\mathbf{E}_N^{\widehat{\phi}}\left[\exp\left(h\sum_{x\in\Lambda}\delta_x\right);\, A\right] \;-\; (e-1)\,\psi_{N,h}(0,2,\beta;A)\,. \tag{A.15}$$

The expressions in the statement of Lemma 6.6 are retrieved from (A.15) by setting $\Lambda=\Lambda_{N_1,N_0}$, $A=A_\kappa$ and by replacing $e-1$ by the larger value 2. □

Appendix B. Proof of Proposition 2.4

We give the proof in four steps.

Step 1: upper bound on the contact density. The fractional moment method yields also a quantitative upper bound on the contact fraction: if for $c>0$ we introduce the event $B_{N,c}=\{\sum_{x\in\widetilde{\Lambda}_N}\delta_x\geq cN^d\}$, as in (3.6) and (3.8), but this time with $\alpha=2h/\beta$ (compare with (3.5) in the Gaussian case)

$$\left(\mathbb{E}\left[\sqrt{Z^{\beta,\omega}_{N,h}\left(B_{N,ChN^d}\right)}\right]\right)^2\leq\widetilde{\mathbb{E}}\left[Z^{\beta,\omega}_{N,h}\left(B_{N,ChN^d}\right)\right]\exp\left(4\frac{h^2}{\beta^2}N^d\right)$$
$$\leq Z_{N,-h}\left(B_{N,ChN^d}\right)\exp\left(4\frac{h^2}{\beta^2}N^d\right)\leq\exp\left(-\left(C-\frac{4}{\beta^2}\right)h^2N^d\right),\quad\text{(B.1)}$$

so that if we choose $C=6/\beta^2$, by Markov inequality, we have

$$\mathbb{P}\left(Z^{\beta,\omega}_{N,h}\left(B_{N,ChN^d}\right)\geq\exp\left(-2\frac{h^2}{\beta^2}N^d\right)\right)\leq\exp\left(-\frac{h^2}{\beta^2}N^d\right).\quad\text{(B.2)}$$

We now focus on $\mathbf{P}^{\beta,\omega}_{N,h}(B_{N,ChN^d})=Z^{\beta,\omega}_{N,h}\left(B_{N,ChN^d}\right)/Z^{\beta,\omega}_{N,h}$ which we are going to bound from above simply by one for ω in the event whose probability is estimated in (B.2) and otherwise we use, as in Section 2.2 the *entropic repulsion* estimate [38] $\inf_\omega Z^{\beta,\omega}_{N,h}\geq\exp(-r(N))$ with $r(N)=o(N^d)$. This yields

$$\mathbb{E}\mathbf{P}^{\beta,\omega}_{N,h}\left(B_{N,6hN^d/\beta^2}\right)\leq\exp\left(-\frac{h^2}{\beta^2}N^d\right)+\exp\left(r(N)-2\frac{h^2}{\beta^2}N^d\right).\quad\text{(B.3)}$$

The punchline of step 1 is that for every $h>0$

$$\lim_{N\to\infty}\mathbb{E}\mathbf{P}^{\beta,\omega}_{N,h}\left(B_{N,6hN^d/\beta^2}\right)=0.\quad\text{(B.4)}$$

Step 2: neighbor averages below a threshold for too many sites implies high contact fraction. Set $\overline{\phi}_x=(1/2d)\sum_{y\sim x}\phi_y$ and consider the event that on the even sites there is a density of at least $\varepsilon/2$ of the $\overline{\phi}$ variables that in absolute value are smaller than $\sqrt{1/(4d)\log(1/h)}$:

$$F_{N,\varepsilon}=\left\{\phi:\sum_{x\in\mathring{\Lambda}_N:\,x\text{ even}}\mathbf{1}_{\left(-\sqrt{\frac{1}{4d}\log(1/h)},\sqrt{\frac{1}{4d}\log(1/h)}\right)}\left(\overline{\phi}_x\right)\geq\frac{\varepsilon}{4}N^d\right\}.\quad\text{(B.5)}$$

We aim at showing that there exists h_0 such that for $h\in(0,h_0)$

$$\lim_N\mathbb{E}\mathbf{P}^{\beta,\omega}_{N,h}\left(F_{N,\varepsilon}\right)=0.\quad\text{(B.6)}$$

By (B.4), (B.6) is implied by

$$\lim_{N\to\infty}\mathbb{E}\mathbf{P}^{\beta,\omega}_{N,h}\left(B^{\complement}_{N,6hN^d/\beta^2}\cap F_{N,\varepsilon}\right)=0,\quad\text{(B.7)}$$

and, by writing once again the probability as ratio of partition functions, by using the lower bound on the denominator given by the entropic repulsion estimate and by performing the $\mathbb{P}$ expectation, we see that

$$\mathbb{E}\mathbf{P}^{\beta,\omega}_{N,h}\left(B^{\complement}_{N,6hN^d/\beta^2}\cap F_{N,\varepsilon}\right)\leq\exp(-r(n)+hN^d)\mathbf{P}_N\left(B^{\complement}_{N,6hN^d/\beta^2}\cap F_{N,\varepsilon}\right),\quad\text{(B.8)}$$

so that we are done if we show that for a $c \in (0,1)$ and an $h_0 > 0$ we have that for $h \in (0, h_0)$

$$\mathbf{P}_N \left(B^{\complement}_{N,6hN^d/\beta^2} \cap F_{N,\varepsilon} \right) \le \exp\left(-h^c N^d \right) . \tag{B.9}$$

For this use the fact that the event $B_{N,6hN^d/\beta^2}$ contains the event that the $6hN^d/\beta^2$ (or more) contacts are all on the even sites. By conditioning on the odd sites and by using the Markov property – remark that $F_{N,\varepsilon}$ is measurable with respect to the σ-algebra of the odd variables – one realizes that the random variables δ_x, x even, are independent Bernoulli variables of parameter

$$p_x := P\left(\frac{1}{\sqrt{2d}} \mathcal{N} + \overline{\phi}_x \in [-1,1] \right) . \tag{B.10}$$

If $|\overline{\phi}_x| \le \sqrt{1/(4d)\log(1/h)}$ by the standard Gaussian tail estimate – one can use (6.3) even if what we claim here is substantially rougher – we see that for h sufficiently small

$$p_x \ge \exp\left(-\frac{1}{2} \log\left(\frac{1}{h} \right) \right) = h^{\frac{1}{2}} =: p = p(h) . \tag{B.11}$$

So, once $\phi \in F_{N,\varepsilon}$ is chosen and hence the set (of at least $\varepsilon N^d/4$ sites x) on which $p_x \ge p$ is determined, we are simply left with a Large Deviation upper bound on a binomial random variable $B(n,p)$: it is a well known fact, direct consequence of the exponential form of the Markov inequality, that the probability that for $\Delta \in (0,1)$

$$P\left(B(n,p) \le p\Delta n \right) \le \exp(-n f(p,\Delta)) \tag{B.12}$$

where

$$f(p,\Delta) := \Delta p \log \Delta + (1 - \Delta p) \log((1-\Delta p)/(1-p)).$$

If $\Delta = 1/2$

$$f(p,1/2) \overset{p\to 0}{\sim} \frac{1}{2}(1 - \log 2)p,$$

so that $f(p,1/2) \ge p/10$ for p sufficiently small. Therefore for $p(h) = h^{1/2}$ this implies in a rather direct way that

$$\mathbf{P}_N \left(B^{\complement}_{N,\frac{\varepsilon}{8}h^{\frac{1}{2}}N^d} \cap F_{N,\varepsilon} \right) \le \exp\left(-h^{\frac{1}{2}} \frac{\varepsilon}{40} N^d \right) , \tag{B.13}$$

and this implies (B.9) for any $c \in (1/2, 1)$ and h sufficiently small.

Step 3: the Gaussian Hamiltonian cannot be too large under the pinning model. If we set

$$H_N(\phi) := \sum_{\substack{(x,y)\in(\Lambda)^2\setminus(\partial\Lambda)^2 \\ x\sim y}} \frac{(\phi_x - \phi_y)^2}{2} =: (\phi, A_N\phi)_N , \tag{B.14}$$

where A_N is a positive definite symmetric $(N-1)^d \times (N-1)^d$ and $(\cdot,\cdot)_N$ is the scalar product on $\mathbb{R}^{N-1}$. In fact A_N is just discrete Laplacian operator with zero Dirichlet boundary condition which is the generator of the simple random walk killed upon hitting $\partial\Lambda_N$. We have

$$\begin{aligned} \mathbb{E}\mathbf{P}^{\beta,\omega}_{N,h}\left(H_N(\phi) > CN^d \right) &\le e^{r_N} \mathbb{E}\mathbf{E}_N \left[e^{\sum_{x\in\tilde{\Lambda}_N}\left(\beta\omega_x - \frac{\beta^2}{2} + h\right)\delta_x} ; H_N(\phi) > CN^d \right] \\ &\le e^{r_N + hN^d} \mathbf{P}_N\left(H_N(\phi) > CN^d \right) , \end{aligned} \tag{B.15}$$

where in the first step we have applied once again the entropic repulsion bound ($r_N = o(N^d)$) and in the second step we have performed the expectation with respect to the disorder and then we have bounded in the obvious way the pinning part. We are left with estimating the remaining probability, which is a Gaussian computation: since for $\lambda < 1/2$

$$\mathbf{E}_N\left[\exp(\lambda H_N(\phi))\right] = \int_{\mathbb{R}^{\mathring{\Lambda}_N}} \sqrt{\frac{\det\left(A_N\right)}{(2\pi)^{(N-1)^d}}} \exp\left(-\frac{(1-2\lambda)}{2}\left(\phi, A_N\phi\right)_N\right) \prod_{x\in\mathring{\Lambda}_N} \mathrm{d}\phi_x$$
$$= (1-2\lambda)^{-\frac{1}{2}(N-1)^d}, \quad \text{(B.16)}$$

where we have used $\det(cA_N) = c^{(N-1)^d}\det(A_N)$, $c > 0$. By applying the Markov inequality with $\lambda = 3/8$ we obtain

$$\mathbf{P}_N\left(H_N(\phi) > CN^d\right) \le 2^{(N-1)^d}\exp(-(3/8)CN^d.$$

Therefore, by recalling (B.15), we have

$$\mathbb{E}\mathbf{P}_{N,h}^{\beta,\omega}\left(H_N(\phi) > CN^d\right) \le \exp\left(-\frac{C}{4}N^d\right), \quad \text{(B.17)}$$

for any $C > 8\log 2$ (for example $C = 6$), h small and N sufficiently large.

Step 4: conclusion. It is now a matter of putting together (B.6) and (B.17). Let us first observe that for h sufficiently small

$$\lim_{N\to\infty} \mathbb{E}\mathbf{P}_{N,h}^{\beta,\omega}\left(F_{N,\frac{1}{2},\varepsilon}^{\complement} \cap \left\{\phi : \sum_{\substack{x\in\mathring{\Lambda}_N \\ x \text{ even}}} \mathbf{1}_{\left\{\exists y,\ y\sim x \text{ and } |\phi_y|\le\sqrt{\frac{1}{8d}\log(1/h)}\right\}} \ge \frac{\varepsilon N^d}{2}\right\}\right) = 0. \quad \text{(B.18)}$$

This is because on the event whose probability is computed in (B.18) there will be at least $\varepsilon N^d/4$ even sites x on which $|\overline{\phi}_x| > \sqrt{\frac{1}{4d}\log(1/h)}$ and at least for one of the neighboring sites y we have $|\phi_y| \le \sqrt{\frac{1}{8d}\log(1/h)}$, while instead there is another neighbor y' of x for which $|\phi_{y'}| > \sqrt{\frac{1}{4d}\log(1/h)}$. Therefore it is not difficult to see that this implies $(\phi_y - \phi_x)^2 + (\phi_{y'} - \phi_x)^2 \ge \frac{1}{100d}\log(1/h)$, and in turn $H_N(\phi) \ge \log(1/h)\varepsilon N^d/(400d)$. By choosing h sufficiently small we see that the event under analysis becomes a subset of $\{H_N(\phi) > 6N^d\}$ and, by (B.17), we see that (B.18) tends to zero. It remains then to repeat the argument (that is even simpler) to show that also

$$\lim_{N\to\infty} \mathbb{E}\mathbf{P}_{N,h}^{\beta,\omega}\left(F_{N,\frac{1}{2},\varepsilon}^{\complement} \cap \left\{\phi : \sum_{\substack{x\in\mathring{\Lambda}_N \\ x \text{ even}}} \mathbf{1}_{\left\{|\phi_x|\le\sqrt{\frac{1}{8d}\log(1/h)}\right\}} \ge \frac{\varepsilon N^d}{2}\right\}\right) = 0. \quad \text{(B.19)}$$

If we combine (B.18) and (B.19), by recalling (B.6) we conclude. □

References

[1] M. Aizenman and J. Wehr, *Rounding effects of quenched randomness on first-order phase transitions*, Comm. Math. Phys. **130** (1990), 489-528.

[2] K. S. Alexander, *The effect of disorder on polymer depinning transitions*, Commun. Math. Phys. **279** (2008), 117-146.

[3] K. S. Alexander and N. Zygouras, *Quenched and annealed critical points in polymer pinning models*, Comm. Math. Phys. **291** (2009), 659-689.

[4] K. S. Alexander and N. Zygouras, *Path properties of the disordered pinning model in the delocalized regime*, Ann. Appl. Probab. **24** (2014), 599-615.

[5] Q. Berger, *Pinning model in random correlated environment: appearance of an infinite disorder regime*, J. Stat. Phys. **155** (2014), 544-570.

[6] Q. Berger and H. Lacoin, *Sharp critical behavior for pinning models in a random correlated environment*, Stochastic Process. Appl. **122** (2012), 1397-1436.

[7] Q. Berger and H. Lacoin, *Pinning on a defect line: characterization of marginal disorder relevance and sharp asymptotics for the critical point shift*, J. Inst. Math. Jussieu (published online).

[8] E. Bolthausen, *Random copolymers*, in *Correlated random systems: Five different methods*, CIRM Jean-Morlet Chair: Spring 2013, V. Gayrard et N. Kistler Eds., Springer Lecture Notes in Mathematics, 2015.

[9] E. Bolthausen and D. Brydges, *Localization and decay of correlations for a pinned lattice free field in dimension two*, in: State of the Art in Probability and Statistics, Festschrift for Willem R. van Zwet, IMS Lecture Notes Vol. 36 (2001), 134-149.

[10] E. Bolthausen, J.-D. Deuschel and O. Zeitouni, *Absence of a wetting transition for a pinned harmonic crystal in dimensions three and larger*, J. Math. Phys. **41** (2000), 1211-1223.

[11] E. Bolthausen and Y. Velenik, *Critical behavior of the massless free field at the depinning transition*, Comm. Math. Phys. **223** (2001), 161-203.

[12] A. Bovier, *Statistical mechanics of disordered systems. A mathematical perspective*, Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2006.

[13] J. Bricmont and A. Kupiainen, *Phase transition in the 3d random field Ising model*, Comm. Math. Phys. **116** (1988), 539-572.

[14] P. Caputo and Y. Velenik, *A note on wetting transition for gradient field*, Stoch. Proc. Appl. **87** (2000), 107-113.

[15] F. Caravenna, G. Giacomin and F. L. Toninelli *Copolymers at selective interfaces: settled issues and open problems*, in Probability in Complex Physical Systems, volume 11 of Springer Proceedings in Mathematics, pages 289-311. Springer Berlin Heidelberg, 2012.

[16] F. Caravenna, R. Sun and N. Zygouras, *The continuum disordered pinning model*, Probab. Theory Related Fields, **164** (2016), 17-59.

[17] D. Cheliotis and F. den Hollander, *Variational characterization of the critical curve for pinning of random polymers* Ann. Probab. **41** (2013) 1767-1805.

[18] A. Chiarini, A. Cipriani and R. S. Hazra, *A note on the extremal process of the supercritical Gaussian Free Field*, Elec. Comm. Probab. **20** (2015) 74.

[19] L. Coquille and P. Milos, *A note on the discrete Gaussian free field with disordered pinning on* $\mathbb{Z}^d$*,* $d \geq 2$, Stoch. Proc. and Appl. **123** (2013) 3542-3559.

[20] L. Coquille and P. Milos, *A second note on the discrete Gaussian free field with disordered pinning on* $\mathbb{Z}^d$*,* $d \geq 2$, (preprint) arXiv:1303.6770.

[21] B. Derrida, G. Giacomin, H. Lacoin and F. L. Toninelli, *Fractional moment bounds and disorder relevance for pinning models*, Commun. Math. Phys. **287** (2009) 867–887.

[22] B. Derrida and M. Retaux, *The depinning transition in presence of disorder: a toy model*, J. Stat. Phys. **156** (2014), 268-290.

[23] B. Derrida, V. Hakim and J. Vannimenus, *Effect of disorder on two-dimensional wetting*, J. Statist. Phys. **66** (1992), 1189–1213.

[24] F. Dunlop, J. Magnen, V. Rivasseau and P. Roche, *Pinning of an interface by a weak potential*, J. Stat. Phys. **66** (1992), 71-98.

[25] S. Finch, *Mathematical Constants,* Encyclopedia of Mathematics and its Applications, Cambridge University Press, 2003.

[26] M. E. Fisher, *Walks, walls, wetting, and melting*, J. Statist. Phys. **34** (1984), 667-729.

[27] G. Giacomin, *Random polymer models*, Imperial College Press, World Scientific (2007).

[28] G. Giacomin, *Disorder and critical phenomena through basic probability models*, École d'été de probablités de Saint-Flour XL-2010, Lecture Notes in Mathematics **2025**, Springer, 2011.

[29] G. Giacomin, H. Lacoin and F. L. Toninelli, *Hierarchical pinning models, quadratic maps and quenched disorder*, Probab. Theory Relat. Fields **147** (2010), 185-216.

[30] G. Giacomin, H. Lacoin and F. L. Toninelli, *Marginal relevance of disorder for pinning models*, Commun. Pure Appl. Math. **63** (2010) 233-265.
[31] G. Giacomin and F. L. Toninelli, *Smoothing effect of quenched disorder on polymer depinning transitions*, Commun. Math. Phys. **266** (2006), 1–16.
[32] F. Guerra and F. L. Toninelli, *Quadratic replica coupling in the Sherrington-Kirkpatrick mean field spin glass model*, J. Math. Phys. **43** (2002), 3704-3716.
[33] A. B. Harris, *Effect of random defects on the critical behaviour of Ising models* J. Phys. C **7** (1974), 16711692.
[34] D. Ioffe and Y. Velenik, *A note on the decay of correlations under δ-pinning*, Probab. Theory Relat. Fields **116** (2000), 379-389.
[35] H. Lacoin, *The martingale approach to disorder irrelevance for pinning models*, Electron. Commun. Probab. **15** (2010), 418-427.
[36] H. Lacoin, *Pinning and disorder for the Gaussian free field II: the two dimensional case*, (preprint) arXiv:1512.05240 [math-ph].
[37] G. F. Lawler and V. Limic, *Random Walk: A Modern Introduction*, Cambridge Studies in Advanced Mathematics **123**, 2010.
[38] J. L. Lebowitz and C. Maes, *The effect of an external field on an interface, entropic repulsion*, J. Statist. Phys. **46** (1987), 39-49.
[39] F. L. Toninelli, *A replica-coupling approach to disordered pinning models*, Commun. Math. Phys. **280** (2008), 389-401.
[40] F. L. Toninelli, *Disordered pinning models and copolymers: beyond annealed bounds*, Ann. Appl. Probab. **18** (2008), 1569-1587.
[41] Y. Velenik, *Localization and delocalization of random interfaces*, Probab. Surv. **3** (2006), 112-169.

Université Paris Diderot, Sorbonne Paris Cité, Laboratoire de Probabilités et Modèles Aléatoires, UMR 7599, F- 75205 Paris, France

IMPA, Institudo de Matemàtica Pura e Aplicada, Estrada Dona Castorina 110 Rio de Janeiro, CEP-22460-320, Brasil.