

## ON THE WULFF CRYSTAL IN THE ISING MODEL

BY RAPHAËL CERF<sup>1</sup> AND ÁGOSTON PISZTORA<sup>2</sup>

*University Paris Sud and Carnegie Mellon University*

We study the phase separation phenomenon in the Ising model in dimensions  $d \geq 3$ . To this end we work in a large box with plus boundary conditions and we condition the system to have an excess amount of negative spins so that the empirical magnetization is smaller than the spontaneous magnetization  $m^*$ . We confirm the prediction of the phenomenological theory by proving that with high probability a single droplet of the minus phase emerges surrounded by the plus phase. Moreover, the rescaled droplet is asymptotically close to a definite deterministic shape, the Wulff crystal, which minimizes the surface free energy. In the course of the proof we establish a surface order large deviation principle for the magnetization. Our results are valid for temperatures  $T$  below a limit of slab-thresholds  $\widehat{T}_c$  conjectured to agree with the critical point  $T_c$ . Moreover,  $T$  should be such that there exist only two extremal translation invariant Gibbs states at that temperature, a property which can fail for at most countably many values and which is conjectured to be true for every  $T$ . The proofs are based on the Fortuin–Kasteleyn representation of the Ising model along with coarse-graining techniques. To handle the emerging macroscopic objects we employ tools from geometric measure theory which provide an adequate framework for the large deviation analysis. Finally, we propose a heuristic picture that for subcritical temperatures close enough to  $T_c$ , the dominant minus spin cluster of the Wulff droplet *permeates* the entire box and has a strictly positive local density everywhere.

**1. Introduction and statement of results.** In order to develop heuristics we begin with an informal discussion of phase separation and phase coexistence phenomena, the main subject of this paper, with a simplistic description of a physical system familiar from everyday experience. The system we propose to consider is the mixture of two substances: oil and water. For our purpose it will be enough to recall that water and oil tend to repel each other, and yet a certain (even if only a very small) amount of oil can be dissolved in water, and vice versa. (A mixture of phenol and water would be a more realistic choice since they are more miscible, however the potential reader is hardly familiar with that mixture. A mixture of salt and water is also a good example but the phenomenon we want to describe is more complex in that case.)

As we know, solubility is not unlimited. Depending on the temperature  $T$ , there exist threshold densities  $d_c^{o/w}(T)$  and  $d_c^{w/o}(T)$  (both increasing in  $T$ ) corresponding to *saturated* solutions of oil in water and water in oil, respec-

---

Received April 1999; revised February 2000.

<sup>1</sup>Supported in part by Center for Nonlinear Analysis, Carnegie Mellon University.

<sup>2</sup>Supported in part by NSF Grant DMS-97-04901.

AMS 1991 *subject classifications*. 60K35, 82B20, 60F10.

*Key words and phrases*. Phase separation, Wulff crystal, Ising model, large deviations, FK percolation.

tively. These two types of saturated solutions are called the pure phases “o/w” and “w/o” and they correspond to a perfect balance between energy and entropy.

In order to see phase separation in this system all we have to do is to take an almost saturated solution of oil in water with density  $d$  at temperature  $T$  and let the system cool down to a temperature  $T'$  such that  $d > d_c^{o/w}(T')$ . The excess amount of oil precipitates and macroscopic droplets emerge. The inside of the droplets does not contain pure oil, rather the droplets are regions where we see the phase w/o. They swim around in the phase o/w (assuming there is no gravitation, otherwise the phase w/o gathers at the top.) What is the law which governs the behavior of the coexisting phases?

The classical phenomenological theory asserts that there is a positive “surface free energy”  $\mathcal{S}$  associated with the macroscopic phase boundaries and the system will settle down in a state minimizing this energy. Assuming isotropy,  $\mathcal{S}$  is proportional to the surface area. Hence, in perfect equilibrium, there should be one single droplet of the phase w/o with a spherical shape floating in the phase o/w, since, by the isoperimetric inequality, this is the energetically most favorable configuration.

A mathematical challenge is to confirm this theory starting from a microscopic description of the system, to explain the existence and breakdown of solubility, the occurrence of phase separation on a macroscopic scale, and to verify the prediction about the existence of a single droplet with a specific shape.

Let us try to set up a simple model for our system. A convenient choice is a lattice model: each site of the lattice is occupied either by a water particle or an oil particle which we indicate by  $+1$  or  $-1$ . The interaction between particles is repulsive and occurs when the substances are in immediate contact. Hence a repulsive nearest neighbor interaction is a sensible choice. If we want to focus only on the dominant repulsive interaction between different molecules, we could simplify the model by making the two substances “symmetric” by assuming that their self-interactions are of equal magnitude, or equivalently, equal to zero. Thus the total energy of a configuration should be simply the sum of all nearest neighbor pairs with different sign. Recall that in our experiment the density of oil is fixed, therefore we have a constraint on the possible configurations: the proportion of  $+1$ -s and  $-1$ -s has to be fixed. The reader familiar with the Ising model has already observed that our simple model is equivalent to the Ising model with plus boundary conditions (guaranteeing the water dominance) *conditioned* on the event that the magnetization is fixed and is smaller than the spontaneous magnetization (corresponding to the saturation density  $d_c^{o/w}$ ) at the given temperature. Note that this event is extremely unlikely when the system is large. In fact, the study of this restricted system leads to questions about large deviations in the (unrestricted) Ising model.

There is an additional difficulty due to the use of the lattice: the system is anisotropic; thus we have to deal with a surface energy depending on the direction which leads to a nonspherical droplet. It is called the *Wulff droplet*

or *Wulff crystal* after Wulff [69] who first studied and solved the corresponding isoperimetric problem.

The goal of the present paper is to continue the study of the questions stated above in the context of the Ising model (which is, as we have just seen, a fairly natural choice) with the specific aim of developing methods which work independently from dimension and temperature. Before we start with the presentation of our results we give a brief summary of the previous work on this subject.

The first efforts were devoted to the study of large deviations of the empirical magnetization, that is, the average value of the spins in a large box. A *volume order* large deviation principle (LDP) has been established for the Ising model by various authors: Comets [18], Ellis [30], Föllmer and Orey [34], Olla [56]. The corresponding rate function has been found to vanish in  $[-m^*, m^*]$  where  $m^*$  denotes the spontaneous magnetization. In fact, it was suspected that the correct order of decay is exponential to *surface order*. Indeed, Schonmann [62] found a proof of this conjecture, valid for any dimensions and low enough temperatures and Chayes, Chayes and Schonmann [17] extended the result for the supercritical  $\beta > \beta_c$  regime in the two-dimensional case. Föllmer and Ort [35] investigated this phenomenon on the level of empirical measures. Finally, inspired by the work of Kesten and Zhang [47] on related questions in percolation, Pisztora [59] established surface order upper bounds for the remaining dimensions  $d \geq 3$  above the slab-threshold  $\hat{\beta}_c$ , introduced in the same work, which is conjectured to agree with the critical point  $\beta_c$ . In that work a coarse-graining scheme has been developed for supercritical Fortuin–Kasteleyn percolation (or random cluster model) in conjunction with a stochastic domination argument (generalized and improved in [50]) which allows controlling the renormalized process, and so, the original one.

The monograph of Dobrushin, Kotecký and Shlosman [27] opened the way to the rigorous study of the phase separation phenomenon, creating thereby an immense interest and activity which has lasted up to the present time. Their analysis, which provided the first mathematical proof of phase separation, had been performed in the context of the Ising model. The main tool of their work is the cluster expansion, which, on the one hand allowed the derivation of results much finer than necessary to verify the Wulff construction, on the other hand restricted the validity of the results to two dimensions and low temperatures. It was a challenge to improve on those results. Pfister [57] simplified the proof through duality arguments. Alexander, Chayes and Chayes [5] have proved the Wulff construction in the entire supercritical phase of two-dimensional Bernoulli percolation. Alexander [4] subsequently refined the probabilistic estimates. By using Pfister's approach and certain coarse-graining techniques from [59], Ioffe [44, 45] extended the basic large deviation principle for the magnetization up to the critical temperature. Finally, Ioffe and Schonmann [46] extended the results of [27] up to  $T_c$ . The results around the Wulff construction for the two-dimensional Ising model are now fairly precise. They go far beyond large deviations statements. Dobrushin and Hryniv [26, 43] managed to describe the Gaussian fluctuations around the interface.

This whole area of research is now very active (see [16, 58, 63, 64, 65]). Nevertheless these works are confined to dimension two. The main reason is that the probabilistic estimates for the presence of an interface separating the two phases rely on the skeleton coarse-graining technique. Through the skeleton, one can approximate a polygonal line drawn on the lattice by a coarser one and then use a combinatorial bound for the number of polygonal lines. Unfortunately, it seems hard to find a higher-dimensional analogue of this elegant and efficient technique.

A next challenge was to analyze phase separation for short range models in higher dimensions. A first step towards this goal was accomplished by Cerf [14]: large deviations statements are proven for the *cluster shapes* in three-dimensional Bernoulli percolation, from which a picture of the Wulff crystal of this model emerges. From the very beginning, the problem is embedded in a continuous setting, and adequate tools from geometric measure theory (à la Caccioppoli and De Giorgi) are used with the crucial benefit that the two-dimensional combinatorial argument associated with skeletons can be circumvented by a compactness argument. The key ingredients to get the required probabilistic estimates are the coarse-graining results of Pisztora [59], specialized to the case of Bernoulli percolation.

The current paper is a natural successor to [14] and [59]. We proceed by using the FK representation of the Ising model. Equipped with the renormalization technology of [59], we follow the same strategy as in [14] to derive a full large deviation principle for the *shape of the minus phase* with respect to the unconstrained Ising measure with plus boundary conditions. A novel issue is that thereby we need a very accurate control of the dependence between events occurring in distant regions; the corresponding decoupling results, Proposition 3.1 and Lemma 3.2, might be of independent interest. The aforementioned LDP then allows deriving the desired results on the large deviations of the (empirical) magnetization (Theorem 1.1) and on the existence of the Wulff crystal (Theorem 1.2). (These results were announced at the statistical mechanics conference in Paris on January 27, 1999.)

In an independent line of development, fundamental ideas were already present in a series of works in the context of the Ising model with Kac potentials. Alberti, Bellettini, Cassandro, Presutti [3] and Bellettini, Cassandro, Presutti [7] developed the general philosophy of embedding the problem in a continuous setting in order to use the BV framework (the equivalent functional formulation of the sets of finite perimeter). For that aim, they introduced the local averaging of the magnetization to study surface order large deviations. The subsequent works of Benoist, Bodineau, Buttà and Presutti [8] and Benoist, Bodineau and Presutti [9] contain useful techniques for handling the renormalized picture. In particular, the idea of transforming adequately the configuration in order to get estimates on certain probabilities, which is an important element of the interface lemma, appears there as well. However, these works lose the microscopic structure of the model: they perform a mean-field limit where the range of interaction tends to infinity and everything becomes isotropic. While preparing the manuscript of this article, we received

the preprint [11]. Building upon [8, 9, 14, 59], Bodineau has proved results similar to ours for sufficiently low temperatures. A discussion of similarities and differences between [11] and the present paper is postponed to after the description of the strategy of our proofs.

We next state our results and outline the strategy of the proofs.

*Range of validity of the results.* We consider Fortuin–Kasteleyn (FK) percolation (also known as the random cluster model) in dimensions  $d \geq 3$  in the regime  $q \geq 1$ ,  $p > \hat{p}_c$ ,  $\theta^f(p) = \theta^w(p)$ . Here  $\hat{p}_c$  is the slab percolation threshold introduced in [59] and  $\theta^f(p)$ ,  $\theta^w(p)$  are the densities of the infinite open cluster for the infinite volume FK measures  $\Phi_\infty^f, \Phi_\infty^w$  corresponding to free and wired boundary conditions. The equality  $\theta^f(p) = \theta^w(p)$  implies that there exists a unique infinite volume FK measure  $\Phi_\infty$  on the cubic lattice  $\mathbb{L}^d$  corresponding to the parameters  $(p, q)$ , by results of Lebowitz and Martin-Löf [49, 48] and Grimmett [41]. It is also known that the condition  $\theta^f(p) = \theta^w(p)$  holds for values of  $p$  close to 1 and might only be violated at countably many values. It is conjectured that  $\theta^f(p) = \theta^w(p)$  for every  $p \neq p_c$  and that  $\hat{p}_c$  coincides with the critical point  $p_c$ . (This is the case at least for Bernoulli percolation ( $q = 1$ ) by the result of Grimmett and Marstrand [42] or [40].) For the Ising model, we choose  $q = 2$  and the inverse temperature  $\beta = 1/T$  is related to  $p$  via the relation  $p = 1 - \exp(-\beta)$ . The spontaneous magnetization is denoted by  $m^*(\beta)$  and it is well known that  $m^*(\beta) = \theta(p)$ . We set  $\hat{\beta}_c = 1/\hat{T}_c = -\log(1 - \hat{p}_c)$  and  $\mathcal{W}(d) = \{-\log(1 - p) : p \text{ such that } \theta^f(p) = \theta^w(p)\}$ . The set  $\mathcal{W}(d)$  is the domain of inverse temperatures where there exist only two extremal translation invariant Gibbs states. Our results for the Ising model hold in the region  $d \geq 3$ ,  $\beta > \hat{\beta}_c$ ,  $\beta \in \mathcal{W}(d)$ .

*Results.* We first extract from the FK percolation model a direction-dependent surface tension  $\tau$ . For a unit vector  $\nu$ , let  $A$  be a unit hypersquare orthogonal to  $\nu$ , let  $\text{cyl } A$  be the cylinder  $A + \mathbb{R}\nu$ , then  $\tau(\nu)$  is equal to the limit

$$\lim_{n \rightarrow \infty} -\frac{1}{n^{d-1}} \log \Phi_\infty \left( \begin{array}{l} \text{inside } n \text{ cyl } A \text{ there exists a finite set of closed edges} \\ E \text{ cutting } n \text{ cyl } A \text{ in at least two unbounded com-} \\ \text{ponents and the edges of } E \text{ at distance less than } 2d \\ \text{from the boundary of } n \text{ cyl } A \text{ are at distance less} \\ \text{than } 2d \text{ from } nA. \end{array} \right)$$

This function  $\tau$  satisfies the weak simplex inequality, is continuous, positive and invariant under the isometries which leave  $\mathbb{Z}^d$  invariant. The Wulff crystal  $\mathcal{W}_\tau$  of  $\tau$  is

$$\mathcal{W}_\tau = \{x \in \mathbb{R}^d \mid x \cdot \nu \leq \tau(\nu) \text{ for all unit vectors}\}.$$

The crystal  $\mathcal{W}_\tau$  is convex, closed, bounded and contains the origin in its interior. We define the surface energy  $\mathcal{S}(\mathcal{W}_\tau)$  of the Wulff crystal by

$$\mathcal{S}(\mathcal{W}_\tau) = \sup \left\{ \int_{\mathcal{W}_\tau} \text{div} f(x) dx : f \in C_0^1(\mathbb{R}^d, \mathcal{W}_\tau) \right\},$$

where  $C_0^1(\mathbb{R}^d, \mathscr{W}_\tau)$  is the set of the compactly supported  $C^1$  vector fields taking values in the Wulff crystal  $\mathscr{W}_\tau$  and  $\text{div}$  is the usual divergence operator; cf. (6) for an explanation of this formula.

Our first theorem gives the logarithmic asymptotics of the probability of having a defect of magnetization in the cubic box  $\Lambda(n)$  of diameter  $n$  with plus boundary conditions. For  $m \in [-m^*, m^*]$ , we define the rescaled Wulff crystal

$$\mathscr{W}(m) = \left( \frac{m^* - m}{2m^* \text{vol } \mathscr{W}_\tau} \right)^{1/d} \mathscr{W}_\tau.$$

Let  $m$  be such that  $\mathscr{W}(m)$  fits completely into the unit box. It is easy to see that this is the case if

$$1 - 2(\text{diam } \mathscr{W}_\tau)^{-d} \text{vol } \mathscr{W}_\tau < m/m^* < 1$$

and we call such  $m$  *admissible*.

We denote by  $\mu_{\Lambda(n)}^+$  the Ising Gibbs measure in the box  $\Lambda(n)$  at temperature  $1/\beta$  with plus boundary conditions.

**THEOREM 1.1** (Logarithmic asymptotics of the magnetization). *Let  $d \geq 3$ ,  $\beta > \hat{\beta}_c$ ,  $\beta \in \mathscr{U}(d)$  and  $m$  be admissible. Then*

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mu_{\Lambda(n)}^+ \left[ \frac{1}{n^d} \sum_{x \in \Lambda(n)} \sigma(x) \leq m \right] \\ = - \left( \frac{m^* - m}{2m^* \text{vol } \mathscr{W}_\tau} \right)^{(d-1)/d} \mathscr{I}(\mathscr{W}_\tau) = -\mathscr{I}(\mathscr{W}(m)). \end{aligned}$$

Here we choose to work with finite boxes rather than with the infinite volume measure since usually a finite volume result is more useful in applications. As a benefit, on the technical side, we gain compactness after rescaling. The drawback is that we have to pay attention to the boundary; it is necessary to handle separately the case of an interface sitting on the boundary, for the probabilistic estimates as well as for the geometric approximations.

Our second theorem gives further information on the mechanism creating a defect of magnetization in  $\Lambda(n)$ . Alternatively, it describes the equilibrium of the system when forced to have an excess amount of negative spins. As indicated earlier, a single Wulff droplet of the minus phase emerges (with a local magnetization close to  $-m^*$ ) which is surrounded by the plus phase and contains all the excess amount of negative spins. A convenient way to localize this droplet is to look at local averages of the magnetization over a (small) intermediate scale.

For  $x \in \mathbb{R}^d$  and  $r > 0$  we define the box  $\Lambda(x, r)$  by

$$\Lambda(x, r) = \{y = (y_1, \dots, y_d) \in \mathbb{R}^d \mid -r/2 < y_i - x_i \leq r/2, 1 \leq i \leq d\}.$$

The  $L^1$  distance between two Borel sets  $E, F$  is the volume of their symmetric difference; that is,

$$\text{dist}_{L^1}(E, F) = \text{vol}(E \Delta F).$$

The mass center  $\text{mc}(E)$  of a (bounded) Borel set  $E$  is

$$\text{mc}(E) = \frac{1}{\text{vol } E} \int_E x \, dx.$$

We make the convention that  $\text{mc}(\emptyset) = 0$ . Let  $f(n)$  be a fixed function from  $\mathbb{N}$  to  $\mathbb{N}$  such that both  $n/f(n)^{d-1}$  and  $f(n)/\log n$  tend to  $\infty$  as  $n \rightarrow \infty$ . The *locally averaged magnetization*  $\sigma_n$  is the map from the closed unit cube  $\Gamma = [-1/2, 1/2]^d$  to  $[-1, 1]$  defined by

$$\forall x \in \Gamma \quad \sigma_n(x) = \frac{1}{f(n)^d} \sum_{y \in \Lambda(nx, f(n)) \cap \Lambda(n)} \sigma(y).$$

We partition  $\Gamma$  into the random sets  $\Gamma_n^-, \Gamma_n^0$  and  $\Gamma_n^+$  according to whether the value of the local magnetization is smaller, equal to or larger than zero.

**THEOREM 1.2** (Typical configuration of the locally averaged magnetization). *Let  $d \geq 3$ ,  $\beta > \hat{\beta}_c$ ,  $\beta \in \mathcal{U}(d)$ . For  $\delta > 0$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mu_{\Lambda(n)}^+ \left[ \int_{\Gamma_n^-} |\sigma_n(x) + m^*| \, dx + \text{vol } \Gamma_n^0 + \int_{\Gamma_n^+} |\sigma_n(x) - m^*| \, dx > \delta \right] = -\infty.$$

*Let  $m$  be admissible. There exist constants  $b = b(d, \beta, m, \delta)$ ,  $c = c(d, \beta, m, \delta) > 0$  such that*

$$\begin{aligned} \mu_{\Lambda(n)}^+ \left[ \text{dist}_{L^1}(\Gamma_n^-, \text{mc}(\Gamma_n^-) + \mathcal{W}(m)) < \delta \mid \frac{1}{n^d} \sum_{x \in \Lambda(n)} \sigma(x) \leq m \right] \\ \geq 1 - b \exp(-cn^{d-1}). \end{aligned}$$

We finish the presentation of our results with some comments and speculations.

**REMARKS ON THE MICROSCOPIC STRUCTURE OF THE WULFF DROPLET.** It is well known [27, 46] that in two dimensions the Wulff droplet can be identified with a (random) region surrounded by a *minus spin cluster* (a set surrounds itself). Its external boundary is therefore a large contour separating plus and minus spins and its shape (when rescaled) follows closely the Wulff shape in the sense of the Hausdorff metric. It has already been realized in [27] that in higher dimensions this picture might be false; long but very thin “hairs” (or spikes) might be attached to the droplet without significantly increasing its surface energy. A single hair is depicted in [27], Figure 1.3.

Without addressing the question of hairs here, we argue that for  $d \geq 3$  the situation is even more cumbersome. We expect that the crossing of the threshold  $\beta_p$ , indicating the onset of percolation of minority spins, implies a drastic change in the physical appearance of the Wulff droplet as seen in the spin model on the microscopic level. In the low-temperature regime  $\beta > \beta_p$ ,

we expect that the situation is as predicted by [27]: there is a big minus spin cluster surrounding the Wulff droplet with attached tiny (logarithmic) hairs. In particular, there is a well-defined *sharp interface* between the two phases.

We recall that in dimensions  $d \geq 3$ ,  $\beta_p(d)$  is conjectured to be strictly larger than  $\beta_c$ , meaning that minus spins (as well as plus spins) percolate in the plus phase for  $\beta_c < T < \beta_p$ , but not when  $\beta > \beta_p$ . It is known [1] that this is the case for dimensions large enough. In the case of  $d = 2$  this phenomenon does not occur; that is,  $\beta_p = \beta_c$ , cf. [17].

In the regime  $(\beta_c, \beta_p)$ , however, we expect that the dominant minus spin cluster  $S^-$  of the Wulff droplet will percolate all the way to the boundary of the box. In fact, it permeates the entire box with a certain positive density in the complement of the Wulff region. More precisely, there exist two big spin clusters in the box,  $S^+$  and  $S^-$ , which are both omnipresent in the entire box. Moreover there exist densities  $0 < \rho_1 < \rho_2 < 1$  such that the (local) density of  $S^-$  within the Wulff region and the density of  $S^+$  in the complement are both concentrated around  $\rho_2$  while the density of  $S^-$  outside and  $S^+$  inside the Wulff region is close to  $\rho_1$ . This means that in this regime the phase boundaries cannot be described directly with contours. In fact, it seems that there is no unique phase boundary on the microscopic level; a phenomenon which we call *interface fuzziness*. It has to be contrasted with the picture observed at low temperatures where a sharp interface is present.

At the same time the corresponding object at the FK percolation level does not show this strange phenomenon: in the entire regime  $\beta > \beta_c$ , the Wulff region can be identified with a region associated with a single big FK cluster  $C^-$  with the right shape. The phenomenon described above occurring at the spin level is simply a consequence of the percolation of negatively colored small FK clusters starting at small FK clusters sitting next to the Wulff cluster  $C^-$ .

In our view, this phenomenon is one of the numerous examples where the FK picture captures certain features of the physics of the Ising model in a more efficient and transparent way than the original spin picture does. Finally we remark that in this paper we make no attempt to prove the phenomenon described above; that remains a challenging open problem.

*The strategy of the proofs.* One of our main tools is the FK representation to relate the events occurring at the Ising level to events occurring at the FK level.

It is natural to ask about the significance of this representation in our (or, in fact, in any) approach. To prove a large deviation *upper* bound it is, in general, necessary to show that there is enough independence in the system. This will be achieved in this type of model by some kind of decoupling results. It is here that the FK representation provides a decisive advantage: in FK percolation we do have asymptotic independence from (imposed) boundary conditions, which is definitely not the case in the spin model. (This corresponds to the uniqueness of infinite volume FK measures versus the multiple Gibbs states in the spin model.) The aforementioned decoupling property has been utilized in [59] in the form of the stochastic domination inequality ([59], Proposition 4.1)

on the level of the renormalized process which is of course used in this work as well. Another even more explicit instance of the decoupling property is formulated in Proposition 3.1 and Lemma 3.2 of the present paper: they assert a high level of independence of boundary conditions. They are crucial ingredients for recovering the exact surface energy. Finally, we remark that, at least for low enough temperatures, there might be ways to prove decoupling properties within the spin framework but at moderate temperatures we don't know how to achieve that.

In the present paper we try to provide simple and efficient proofs of Theorems 1.1 and 1.2; hence we do not care to describe accurately the picture at the FK level. A defect of magnetization in a cubic box with positive boundary conditions is caused with overwhelming probability by the presence of one or several large open FK clusters which have been colored negatively ("large" means here of diameter exceeding  $f(n)$ ). Thus we keep track of the region  $M^-$  consisting of the points of  $n\Gamma$  whose  $f(n)$  neighborhood intersects only large clusters colored negatively. We first show that

$$(1) \quad \frac{1}{n^d} \sum_{x \in \Lambda(n)} \sigma(x) = m^*(1 - 2 \text{vol}(M^-/n)) + \delta(n),$$

where the (random) correction term  $\delta(n)$  goes to 0 in probability faster than exponential to surface order. Therefore everything boils down to getting probabilistic estimates on the law of the random set  $M^-/n$  as  $n \rightarrow \infty$ . Obviously the law of  $M^-/n$  converges to the Dirac mass at the empty set: the typical state of the system is the pure phase  $\mu^+$ . We are interested in the large deviations event  $\{\text{vol}(M^-/n) \geq (m^* - m)/(2m^*)\}$ . We first obtain a lower bound: for  $\lambda, x$  such that  $0 < \lambda \text{diam } \mathcal{W}_\tau \leq 1$  and  $x + \lambda \mathcal{W}_\tau \subset \Gamma$ , for any  $\delta > 0$ ,

$$(2) \quad \liminf_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ [\text{dist}_{L^1}(M^-/n, x + \lambda \mathcal{W}_\tau) < \delta] \geq -\lambda^{(d-1)} \mathcal{J}(\mathcal{W}_\tau),$$

where  $\mathbb{P}_{\Lambda(n)}^+$  is the measure on edge-spin configurations realizing the coupling between the FK measure  $\Phi_{\Lambda(n)}^w$  and the Ising Gibbs measure  $\mu_{\Lambda(n)}^+$ . The difficult part is to get the upper bound. In order to control the combinatorial explosion of all possible discrete shapes for  $M^-/n$ , we embed  $M^-/n$  in a continuous space and we use a compactness argument. This is the place where geometric measure theory enters the game. Notice that the use of this theory has been suggested already in [27] (see Remark 1.11.2c, page 14.). The general philosophy of embedding the problem in a continuous space appeared for the first time in the context of the Ising model with Kac potentials [3, 7]. The appropriate geometric setting, namely, the BV-framework, was then successfully combined with large deviations techniques [8, 9]; the Van der Waals surface tension, however, could only be recovered in the Lebowitz–Penrose mean-field limit where the range of interactions tends to infinity and the model becomes isotropic.

We consider the random rescaled region  $M^-/n$  as an element of the space  $\mathcal{B}(\Gamma)$  of the Borel subsets of  $\Gamma$  endowed with the metric  $\text{dist}_{L^1}$ . We define the

surface energy  $\mathcal{J}(E)$  of a Borel set  $E$  by [see (6) for another form]

$$(3) \quad \mathcal{J}(E) = \sup \left\{ \int_E \operatorname{div} f(x) \, dx : f \in C_0^1(\mathbb{R}^d, \mathcal{H}_\tau) \right\}.$$

Notice that the surface energy might be infinite. First, we prove that the law of  $M^-/n$  concentrates fast near sets having a finite surface energy: there exists a positive constant  $c = c(d, p, q)$  such that

$$(4) \quad \forall \lambda, \delta > 0, \quad \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ [\operatorname{dist}_{L^1}(M^-/n, \mathcal{J}^{-1}([0, \lambda])) \geq \delta] \leq -c\lambda.$$

Second, we estimate the probability that  $M^-/n$  is close to a fixed set having finite surface energy: for  $E$  in  $\mathcal{B}(\Gamma)$  such that  $\mathcal{J}(E) < \infty$  and for  $\varepsilon > 0$ , there exists  $\delta = \delta(E, \varepsilon) > 0$  such that,

$$(5) \quad \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ [\operatorname{dist}_{L^1}(M^-/n, E) < \delta] \leq -(1 - \varepsilon)\mathcal{J}(E).$$

The essential tools for the bound (5) are the coarse-graining results of [59] to get the probabilistic estimates and the Vitali covering theorem for Hausdorff measures to get appropriate geometric approximations.

The surface energy  $\mathcal{J}(E)$  is infinite unless  $E$  is a set of finite perimeter in the sense of Caccioppoli and De Giorgi. The Gauss–Green theorem shows that, in case the boundary  $\partial E$  of  $E$  is smooth, then

$$(6) \quad \mathcal{J}(E) = \int_{\partial E} \sup_{y \in \mathcal{H}_\tau} (y \cdot \nu_E(x)) \mathcal{H}^{d-1}(dx) = \int_{\partial E} \tau(\nu_E(x)) \mathcal{H}^{d-1}(dx),$$

where  $\nu_E(x)$  is the exterior normal vector to  $E$  at  $x$  and  $\mathcal{H}^{d-1}$  is the  $(d - 1)$ -dimensional Hausdorff measure in  $\mathbb{R}^d$ . The theory of the sets of finite perimeter was historically invented by Caccioppoli [12, 13] and subsequently developed by De Giorgi [19, 20, 21, 22]. The goal of Caccioppoli was to build a general theory of integration for differential forms and to extend the classical Gauss–Green theorem to sets whose boundary is not  $\mathcal{C}^1$ . Independently, De Giorgi was seeking to generalize some isoperimetric problems, starting with the Gauss–Green theorem. This framework is extremely convenient for dealing with variational problems. Indeed the surface energy  $\mathcal{J}$  is lower semicontinuous (l.s.c.) with respect to the  $L^1$  convergence. In fact, the surface energy defined in (3) is the largest l.s.c. extension of the expression (6) to the class of all Borel sets. Moreover the level sets  $\{E \in \mathcal{B}(\Gamma) : \mathcal{J}(E) \leq \lambda\}$ ,  $\lambda \geq 0$ , are compact. This crucial compactness property, in conjunction with the local estimate (5) and the exponential tightness given by (4), yield the upper bound

$$(7) \quad \forall v > 0 \quad \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ [\operatorname{vol}(M^-/n) \geq v] \leq - \inf_{\operatorname{vol} E \geq v} \mathcal{J}(E).$$

It turned out that the class of the sets of locally finite perimeter (also called Caccioppoli sets) is isomorphic to the currents of codimension one of the general geometric measure theory developed by Federer [33]. Taylor [66, 67, 68] proved the Wulff isoperimetric theorem with the theory of currents. This theorem (originally due to Wulff [69], followed by Dinghas [25]) states that the Wulff crystal  $\mathscr{W}_\tau$  is the only solution to the variational problem

$$\text{minimize } \mathscr{J}(E) \text{ under the constraint } \text{vol } E \geq \text{vol } \mathscr{W}_\tau.$$

The Wulff isoperimetric theorem has been reworked and slightly generalized in the framework of the Caccioppoli sets by Fonseca [36] and Fonseca and Müller [37]. Theorem 1.1 is an easy consequence of (1), (2), (7) and the Wulff isoperimetric theorem. For Theorem 1.2, one needs in addition some slightly refined coarse-graining estimates from [59].

Finally, we address the question of the relation between [11] and the present paper. There is a lot in common in these works; namely, the large deviations framework, the geometric setting (functions of bounded variation and sets of finite perimeter are equivalent), and the coarse-graining results of [59]. The basic difference between the two approaches is the following. In [11] almost everything is described in the spin language. In order to recover the exact surface tension factor (along the lines of [14]) the author resorts to the FK representation and verifies that the spin and the percolation definitions of the surface tension agree. It is here that the results have to be confined to low enough temperatures. (Of course this is a crucial step in the proof since this is the very place where the surface tension will be linked to the rate function of the desired LDP). In our approach, the relevant quantities of the spin model are translated into (FK) percolation terms from the very beginning and we work then only at the FK level. The decoupling property for FK measures allows us to push all the results until the limit of the slab-percolation thresholds. There is no analogous result in [11].

**2. Preliminaries.** In this section we introduce the notation first and we give some basic definitions. In the second part, we recall some useful properties of FK (or random clusters measures and we give a short description of the Ising model and its FK representation.

*Notation.* The cardinality of a set  $A$  is denoted by  $|A|$ . The symmetric difference between two sets  $A_1, A_2$  is denoted by  $A_1 \Delta A_2$ . If  $\mathscr{A}$  is a family of sets we write  $\text{cup } \mathscr{A}$  for  $\bigcup_{A \in \mathscr{A}} A$ . For  $r \in \mathbb{R}$ ,  $[r]$  denotes the integer part of  $r$  and  $\lceil r \rceil$  stands for the smallest integer larger or equal to  $r$ .

*Metric.* We denote by  $d_p$  the metric associated with the  $p$ -norm, that is,  $d_p(x, y) = |x - y|_p$  for any  $x, y$  in  $\mathbb{R}^d$ . We will only use the 1, 2 and  $\infty$  norms. The  $d_p$  distance between two subsets  $E_1$  and  $E_2$  of  $\mathbb{R}^d$  is  $d_p(E_1, E_2) = \inf\{|x_1 - x_2|_p: x_1 \in E_1, x_2 \in E_2\}$ . The  $r$ -neighborhood of  $E \subseteq \mathbb{R}^d$  with respect to the  $d_p$  metric is the set  $\mathscr{V}_p(E, r) = \{x \in \mathbb{R}^d: d_p(x, E) < r\}$ . The  $d_p$  diameter of a subset  $E$  of  $\mathbb{R}^d$  is  $\text{diam}_p E = \sup\{|x - y|_p: x, y \in E\}$ . We will usually work with the Euclidean distance  $d_2$  on the continuous space  $\mathbb{R}^d$  and with the

distance  $d_1$  or  $d_\infty$  on the discrete lattice  $\mathbb{Z}^d$ . By default, when we speak of the diameter of a set without any specification, we mean the  $d_\infty$  diameter.

*Geometric objects.* We denote by  $\Gamma$  the closed unit cube  $\Gamma = [-1/2, 1/2]^d$ . Let  $x = (x_1, \dots, x_d)$  be a point of  $\mathbb{R}^d$  and let  $r$  be positive. The closed ball of center  $x$  and Euclidean radius  $r$  is denoted by  $B(x, r)$ . The sphere of center  $x$  and radius  $r$  is  $\partial B(x, r)$ . When  $x = 0$  we usually drop  $x$  from the notation; for instance  $B(r) = B(0, r)$ . This convention applies to all of the objects described below. The unit sphere of  $\mathbb{R}^d$  is denoted by  $S^{d-1}$ . The projective sphere  $PS^{d-1}$  is obtained by identifying opposite points on  $S^{d-1}$ . Let  $w$  be a unit vector. We set

$$\text{hyp}(x, w) = \{y \in \mathbb{R}^d \mid (y - x) \cdot w = 0\}.$$

The specific hyperplane containing the origin and perpendicular to the  $d$ th axis is denoted by  $\mathbb{D}^d$ . For  $r_1, r_2$  in  $\mathbb{R} \cup \{-\infty, +\infty\}$ , we define

$$\text{slab}(x, w, r_1, r_2) = \{y \in \mathbb{R}^d \mid r_1 \leq (y - x) \cdot w \leq r_2\}.$$

The half spaces  $\text{slab}(x, w, 0, \infty)$  and  $\text{slab}(x, w, -\infty, 0)$  are denoted by  $\mathbb{H}_+(x, w)$  and  $\mathbb{H}_-(x, w)$ , respectively. We set

$$B_-(x, r, w) = B(x, r) \cap \mathbb{H}_-(x, w), \quad B_+(x, r, w) = B(x, r) \cap \mathbb{H}_+(x, w).$$

The ‘‘upper’’ half space  $\{x \in \mathbb{Z}^d \mid x_d \geq 0\}$  is denoted by  $\mathbb{H}^d$ . By  $\text{disc}(x, r, w)$  we denote the closed disc centered at  $x$  of radius  $r$  and normal vector  $w$ . A box is a set of the form

$$\Lambda(x, r) = \{y = (y_1, \dots, y_d) \in \mathbb{R}^d \mid -r_i/2 < y_i - x_i \leq r_i/2, i = 1, \dots, d\},$$

where  $x, r \in \mathbb{R}^d$ . Clearly,  $x$  is the center and  $r$  determines the side lengths of the box. If  $r_i = t$  for each  $i = 1, \dots, d$ , where  $t \in \mathbb{R}$ , then we write simply  $\Lambda(x, t)$ . Notice that  $\Lambda(x, t)$  has diameter  $t$  and is neither open nor closed. If  $d_\infty(x, y) \geq t$  then  $\Lambda(x, t)$  and  $\Lambda(y, t)$  are disjoint. Let  $A$  be a subset of  $\mathbb{R}^d$  of linear dimension  $d - 1$ ; that is,  $A$  spans a hyperplane of  $\mathbb{R}^d$ , which we denote by  $\text{hyp } A$ . We call such a set a hyperset. By  $\text{nor } A$  we denote one of the two unit vectors orthogonal to  $\text{hyp } A$ , or equivalently, the element of  $PS^{d-1}$  orthogonal to  $\text{hyp } A$ . The cylinder of basis  $A$  is the set

$$\text{cyl } A = \{x + t \text{ nor } A \mid t \in \mathbb{R}, x \in A\}.$$

We set also  $\text{cyl}(A, r) = \{x + t \text{ nor } A : |t| \leq r; x \in A\} = \text{cyl } A \cap \text{slab}(x, \text{nor } A, -r, r)$ . A hyperrectangle is a hyperset which, up to an orthonormal change of coordinates, is a  $(d - 1)$ -dimensional box.

*Topology and measures.* Let  $E$  be a subset of  $\mathbb{R}^d$ . We denote its interior by  $\text{int } E$ , its closure by  $\text{clo } E$ , its boundary by  $\partial E$ . Whenever  $A$  is a hyperset of  $\mathbb{R}^d$ , that is,  $A$  spans a hyperplane of  $\mathbb{R}^d$ , we use the induced  $(d - 1)$ -dimensional topology of  $\text{hyp } A$  to define  $\partial A$ ,  $\text{int } A$ ,  $\text{clo } A$ . The collection of all the Borel subsets of a set  $E$  of  $\mathbb{R}^d$  is denoted by  $\mathcal{B}(E)$ . The volume  $\text{vol } E$  of a Borel set  $E$  is simply its Lebesgue measure. A Borel set is said to be negligible if its volume is zero. We define a (pseudo) metric  $\text{dist}_{L^1}$  on  $\mathcal{B}(\mathbb{R}^d)$  by

$$\forall E, F \in \mathcal{B}(\mathbb{R}^d) \quad \text{dist}_{L^1}(E, F) = \text{vol}(E \Delta F).$$

When dealing with topological questions on the space  $\mathcal{B}(\mathbb{R}^d)$ , we consider the equivalence classes of the Borel sets modulo negligible sets. Notice that  $\mathcal{B}(\Gamma)$  is a closed subset of  $\mathcal{B}(\mathbb{R}^d)$ . We denote by  $\mathcal{H}^k$  the standard  $k$ -dimensional Hausdorff measure, for  $k = 1, 2, \dots, d$  (see Section A.1 for the definition).

*The lattice  $\mathbb{L}^d$ .* We turn  $\mathbb{Z}^d$  into a graph with vertex set  $\mathbb{Z}^d$  and edge set

$$\mathbb{E}^d = \{\{x, y\} \mid x \in \mathbb{Z}^d, y \in \mathbb{Z}^d, d_1(x, y) = 1\}.$$

This graph is called the *d-dimensional cubic lattice* and is denoted by  $\mathbb{L}^d$ . We often think of this graph as embedded in  $\mathbb{R}^d$ , the edges  $\{x, y\}$  being straight line segments  $[x, y]$  between nearest neighbors. If  $x$  and  $y$  are nearest neighbors, we denote this relation by  $x \sim y$ .

Let  $D$  be a subset of  $\mathbb{R}^d$ . An edge  $\{x, y\}$  of  $\mathbb{E}^d$  is said to be included in  $D$  if both sites  $x, y$  belong to  $D$ . We denote by  $\mathbb{E}^d(D)$  the set of the edges of  $\mathbb{E}^d$  included in  $D$ . For  $D$  a subset of  $\mathbb{Z}^d$ , the graph  $(D, \mathbb{E}^d(D))$  will be often identified with its vertex set  $D$ . For  $E$  a subset of  $\mathbb{E}^d$ , a formula like  $E \subset \mathbb{E}^d(D)$  will be abbreviated into  $E \subset D$ .

To simplify notation, we will sometimes identify subsets of  $\mathbb{R}^d$  with their traces on the lattice; that is, we identify  $A \subseteq \mathbb{R}^d$  with  $A \cap \mathbb{Z}^d$ . For example,  $\Lambda(n)$  denotes a box both in the continuum and in the lattice. On the other hand, if  $A \subseteq \mathbb{Z}^d$ , we define an associated thickened region, regarded as a continuum object in  $\mathbb{R}^d$ , by

$$\text{cube } A = \bigcup_{x \in A} \Lambda(x, 1).$$

*Discrete topology.* Let  $A$  be a subset of  $\mathbb{Z}^d$ . We define its

edge boundary:  $\partial^{\text{edge}} A = \{\{x, y\} \in \mathbb{E}^d \mid x \in A, y \in A^c\};$

inner vertex boundary:  $\partial^{\text{in}} A = \{x \in A \mid \exists y \in A^c \text{ such that } y \sim x\};$

outer vertex boundary:  $\partial^{\text{out}} A = \{x \in A^c \mid \exists y \in A \text{ such that } y \sim x\}.$

These definitions are extended to the subsets of  $\mathbb{R}^d$  by setting, for any  $E \subset \mathbb{R}^d$ ,  $\partial^* E = \partial^*(\mathbb{Z}^d \cap E)$ , where  $*$  stands for *edge, in* or *out*.

A path  $\gamma$  in  $(\mathbb{Z}^d, \mathbb{E}^d)$  (or in any graph) is an alternating sequence  $x_0, e_0, x_1, e_1, \dots, e_{n-1}, x_n, \dots$  of distinct vertices  $x_i$  and edges  $e_i$ , where  $e_i$  is the edge between  $x_i$  and  $x_{i+1}$ . The path is said to connect every pair of its vertices. If the path terminates at some vertex  $x_n$  it is said to have length  $n$ ; otherwise it is infinite. Two paths are disjoint if they have no edges in common. The set  $A$  is said to be connected if the graph  $(A, \mathbb{E}^d(A))$  is connected. Let  $A, B, D$  be subsets of  $\mathbb{R}^d$ . A set of edges  $E \subseteq \mathbb{E}^d$  is said to *separate*  $A$  and  $B$  in  $D$  if there is no path in the graph  $(\mathbb{Z}^d \cap D, \mathbb{E}^d(D) \setminus E)$  connecting  $A$  and  $B$ . The set  $E$  *separates  $\infty$  in  $D$*  if the graph  $(\mathbb{Z}^d \cap D, \mathbb{E}^d(D) \setminus E)$  has at least two infinite components.

*The lattice  $\mathbb{L}^{d, \infty}$ .* We introduce another graph structure on  $\mathbb{Z}^d$ . First we define the edge set  $\mathbb{E}^{d, \infty} = \{\{x, y\} \mid d_\infty(x, y) = 1\}$ . The lattice  $\mathbb{L}^{d, \infty}$  is defined to be the graph  $(\mathbb{Z}^d, \mathbb{E}^{d, \infty})$ . The relevance of this lattice stems from the fact that the *exterior* boundary of any connected finite set  $A$  in  $\mathbb{Z}^d$  is itself connected

when regarded as a subgraph of  $\mathbb{L}^{d,\infty}$  (but not of  $\mathbb{L}^d$ ). To be more precise, let us define the residual components of  $A$  as the connected components of the graph  $(A^c, \mathbb{E}^d(A^c))$ . A subset  $A$  of  $\mathbb{Z}^d$  is  $\mathbb{L}^{d,\infty}$ -connected if the graph  $(A, \mathbb{E}^{d,\infty}(A))$  is connected, where  $\mathbb{E}^{d,\infty}(A)$  is the set of the edges of  $\mathbb{E}^{d,\infty}$  whose both endpoints belong to  $A$ . Note that connectedness in the usual ( $\mathbb{L}^d$ ) sense implies  $\mathbb{L}^{d,\infty}$ -connectedness. Let  $A$  be a  $\mathbb{L}^{d,\infty}$ -connected subset of  $\mathbb{Z}^d$ . If  $R$  is a residual component of  $A$ , then its inner and outer vertex boundaries are also  $\mathbb{L}^{d,\infty}$ -connected (cf. [23]). Suppose in addition that  $A$  is finite. Then exactly one of these residual components, say  $R_1$ , is infinite. The boundary between  $R_1$  and  $A$  is called the exterior (edge, inner vertex, outer vertex) boundary of  $A$ . For future reference, we prove next a little geometrical lemma.

LEMMA 2.1. *For any finite  $\mathbb{L}^{d,\infty}$ -connected subset  $A$  of  $\mathbb{Z}^d$ , for  $r \geq 4$ ,*

$$\text{vol } \mathcal{V}_\infty(A, r) \leq 4^{d+1} r^{d-1} (|A| \vee r).$$

PROOF. If  $\text{diam } A \leq r$ , then  $\text{vol } \mathcal{V}_\infty(A, r) \leq (3r)^d$ . Suppose now that  $r < \text{diam } A < \infty$ . Let  $\{x_1, \dots, x_l\}$  be a collection of vertices of  $A$  of maximal cardinality such that

$$\forall i, j \in \{1 \dots l\}, i \neq j, \quad \Lambda(x_i, r) \cap \Lambda(x_j, r) = \emptyset.$$

The maximality of the collection implies that  $A \subset \Lambda(x_1, 2r) \cup \dots \cup \Lambda(x_l, 2r)$ . Because  $l$  is necessarily larger than or equal to two and  $A$  is  $\mathbb{L}^{d,\infty}$ -connected, for each  $i$  in  $\{1, \dots, l\}$ ,

$$|A \cap \Lambda(x_i, r)| \geq d_\infty(x_i, \partial^{\text{in}} \Lambda(x_i, r)) \geq r/2 - 1$$

so that  $|A| \geq l(r/2 - 1)$ . Since  $\mathcal{V}_\infty(A, r)$  is included in  $l$  boxes of diameter  $4r$ , we obtain

$$\text{vol } \mathcal{V}_\infty(A, r) \leq l(4r)^d \leq (4r)^d (r/2 - 1)^{-1} |A| \leq 4^{d+1} r^{d-1} |A|. \quad \square$$

*Ising Gibbs measure.* Let  $\Lambda \subset \mathbb{Z}^d$  be a box. A *spin configuration* in  $\Lambda$  is an element  $\sigma$  of  $\{-, +\}^\Lambda$ . We denote by  $\sigma(x)$  the spin at site  $x$  in the configuration  $\sigma$ . We define the Hamiltonian or energy  $H_\Lambda(\sigma)$  of a spin configuration  $\sigma$  in  $\Lambda$  by

$$H_\Lambda(\sigma) = -\frac{1}{2} \sum_{\{x, y\} \in \mathbb{E}^d(\Lambda)} \sigma(x)\sigma(y).$$

Let  $\beta > 0$ . The Ising Gibbs measure  $\mu_{\Lambda, \beta}^+$  in  $\Lambda$  with  $+$  boundary conditions at inverse temperature  $\beta$  is the probability measure on  $\{-, +\}^\Lambda$  defined by  $\mu_{\Lambda, \beta}^+(\sigma) = 0$  if  $\sigma(x) = -$  for some  $x \in \partial^{\text{in}} \Lambda$  and

$$\mu_{\Lambda, \beta}^+(\sigma) = (Z_{\Lambda, \beta}^+)^{-1} \exp(-\beta H_\Lambda(\sigma))$$

otherwise, where  $Z_{\Lambda, \beta}^+$  is a normalizing factor called the partition function. The following is known. The limit  $\lim_{n \rightarrow \infty} \mu_{\Lambda(n), \beta}^+(\sigma(0))$  exists for any  $\beta > 0$ . We denote it by  $m^*$ . For  $d \geq 2$ , there exists a critical value  $\beta_c = \beta_c(d)$  in  $(0, \infty)$  such that  $m^* > 0$  for  $\beta > \beta_c$  and  $m^* = 0$  for  $\beta < \beta_c$ .

*FK percolation.*

*Edge configurations.* The basic probability space for edge processes is given by  $\Omega = \{0, 1\}^{\mathbb{E}^d}$ ; its elements are called *edge configurations in  $\mathbb{Z}^d$* . The natural projections are given by  $\text{pr}_e: \omega \in \Omega \mapsto \omega(e) \in \{0, 1\}$ , where  $e \in \mathbb{E}^d$ . An edge  $e$  is called open in the configuration  $\omega$  if  $\text{pr}_e(\omega) = 1$ , and closed otherwise.

For  $E \subseteq \mathbb{E}^d$  with  $E \neq \emptyset$ , we write  $\Omega(E)$  for the set  $\{0, 1\}^E$ ; its elements are called *configurations in  $E$* . Note that there is a one-to-one correspondence between cylinder sets and configurations, which is given by  $\eta \in \Omega(E) \mapsto \{\eta\} := \{\omega \in \Omega \mid \omega(e) = \eta(e) \text{ for every } e \in E\}$ . We will use the following convention: the set  $\Omega$  is regarded as a cylinder (set) corresponding to the “empty configuration” (with the choice  $E = \emptyset$ .) We will sometimes identify cylinders with the corresponding configuration. For  $A \subseteq \mathbb{Z}^d$ , let  $\Omega_A$  stand for the set of the configurations in  $A$ :  $\{0, 1\}^{\mathbb{E}^d(A)}$  and  $\Omega^A$  for the set of the configurations *outside*  $A$ :  $\{0, 1\}^{\mathbb{E}^d \setminus \mathbb{E}^d(A)}$ . In general, for  $A \subseteq B \subseteq \mathbb{Z}^d$ , we set  $\Omega_B^A = \{0, 1\}^{\mathbb{E}^d(B) \setminus \mathbb{E}^d(A)}$ . Given  $\omega \in \Omega$  and  $E \subseteq \mathbb{E}^d$ , we denote by  $\omega(E)$  the restriction of  $\omega$  to  $\Omega(E)$ . Analogously,  $\omega_B^A$  stands for the restriction of  $\omega$  to the set  $\mathbb{E}^d(B) \setminus \mathbb{E}^d(A)$ .

Given  $\eta \in \Omega$ , we denote by  $\mathcal{C}(\eta)$  the set of the edges of  $\mathbb{E}^d$  which are open in the configuration  $\eta$ . The connected components of the graph  $(\mathbb{Z}^d, \mathcal{C}(\eta))$  are called  $\eta$ -clusters. The path  $\gamma = (x_1, e_1, x_2, \dots)$  is said to be  $\eta$ -open if all the edges  $e_i$  belong to  $\mathcal{C}(\eta)$ . We write  $\{A \leftrightarrow B\}$  for the event that there exists an open path joining some site in  $A$  with some site in  $B$ . Similarly, we denote by  $\{A \leftrightarrow \infty\}$  the event that there exists  $x \in A$  lying in an infinite component.

If  $V \subseteq \mathbb{Z}^d$  and  $E$  consists of all the edges between vertices in  $V$ , the graph  $G = (V, E) \subseteq \mathbb{L}^d$  is called the maximal subgraph of  $\mathbb{L}^d$  on the vertices  $V$ . Let  $\omega$  be an edge configuration in  $\mathbb{Z}^d$  (or in a subgraph of  $\mathbb{L}^d$ ). We can look at the *open clusters in  $V$*  or alternatively the open  $V$ -clusters. These clusters are simply the connected components of the random graph  $(V, \mathcal{C}(\omega(E)))$ , where  $\omega(E)$  is the restriction of  $\omega$  to  $E$ .

Given  $E \subseteq \mathbb{E}^d$ , we write  $\mathcal{F}(E)$  for the  $\sigma$ -field generated by the finite-dimensional cylinders associated with configurations in  $\Omega(E)$ . Similarly, for  $A \subseteq B \subseteq \mathbb{Z}^d$ , we use the notation  $\mathcal{F}_B^A$  for the  $\sigma$ -field generated by finite-dimensional cylinders associated with configurations in  $\Omega_B^A$ . If  $A = \emptyset$  or  $B = \mathbb{Z}^d$ , then we omit them from the notation.

*Stochastic domination.* There is a partial order  $\leq$  in  $\Omega$  given by  $\omega \leq \omega'$  iff  $\omega(e) \leq \omega'(e)$  for every  $e \in \mathbb{E}^d$ . A function  $f: \Omega \rightarrow \mathbb{R}$  is called *increasing* if  $f(\omega) \leq f(\omega')$  whenever  $\omega \leq \omega'$ . An event is called increasing if its characteristic function is increasing. Let  $\mathcal{F}$  be a  $\sigma$ -field of subsets of  $\Omega$ . For a pair of probability measures  $\mu$  and  $\nu$  on  $(\Omega, \mathcal{F})$ , we say that  $\mu$  (*stochastically*) *dominates*  $\nu$  if for any  $\mathcal{F}$ -measurable increasing function  $f$  the expectations satisfy  $\mu(f) \geq \nu(f)$ . If, in addition, for each  $\mathcal{F}$ -measurable cylinder  $Z$  with  $\mu(Z) \wedge \nu(Z) > 0$ , we have  $\mu(f \mid Z) \geq \nu(f \mid Z)$ , then we say that  $\mu$  *strongly dominates*  $\nu$ , and we denote this relation by  $\mu \geq \nu$ .

*FK measures.* Let  $V \subseteq \mathbb{Z}^d$  be finite and  $E = \mathbb{E}^d(V)$ . We first introduce (*partially wired*) *boundary conditions* as follows. Consider a partition  $\pi$  of the

set  $\partial^{\text{in}}V$ , say  $\pi = \{B_1, \dots, B_n\}$ . (The sets  $B_i$  are disjoint nonempty subsets of  $\partial^{\text{in}}V$  with  $\bigcup_{i=1, \dots, n} B_i = \partial^{\text{in}}V$ .) We say that  $x, y \in \partial^{\text{in}}V$  are  $\pi$ -wired, if  $x, y \in B_i$  for an  $i \in \{1, \dots, n\}$ . Fix a configuration  $\eta \in \Omega_V$ . We want to count the  $\eta$ -clusters in  $V$  in such a way that  $\pi$ -wired sites are considered to be connected. This can be done in the following formal way. We introduce an equivalence relation on  $V$ :  $x$  and  $y$  are said to be  $\pi \cdot \eta$ -wired if they are both joined by  $\eta$ -open paths to (or identical with) sites  $x', y' \in \partial^{\text{in}}V$  which are themselves  $\pi$ -wired. The new equivalence classes are called  $\pi \cdot \eta$ -clusters, or  $\eta$ -clusters in  $V$  with respect to the boundary condition  $\pi$ . The number of clusters with respect to the boundary condition  $\pi$  (i.e., the number of  $\pi \cdot \eta$ -clusters) is denoted by  $\text{cl}^\pi(\eta)$ . (Note that  $\text{cl}^\pi$  is simply a random variable.)

For fixed  $p \in [0, 1]$  and  $q \geq 1$ , the FK measure with parameters  $(p, q)$  and boundary conditions  $\pi$  is a probability measure on the  $\sigma$ -field  $\mathcal{F}_V$ , defined by the formula

$$(8) \quad \forall \eta \in \Omega_V \quad \Phi_V^{\pi, p, q}[\{\eta\}] = \frac{1}{Z_V^{\pi, p, q}} \left( \prod_{e \in E} p^{\eta(e)}(1-p)^{1-\eta(e)} \right) q^{\text{cl}^\pi(\eta)},$$

where  $Z_V^{\pi, p, q}$  is the appropriate normalization factor. Since  $\mathcal{F}_V$  is an atomic  $\sigma$ -field with atoms  $\{\eta\}$ ,  $\eta \in \Omega_V$ , (8) determines a unique measure on  $\mathcal{F}_V$ . Note that every cylinder has nonzero probability. There are two extremal b.c.s: the free boundary condition corresponds to the partition  $f$  defined to have exactly  $|\partial^{\text{in}}V|$  classes, and the wired b.c. corresponds to the partition  $w$  with only one class. The set of all such measures called FK (or random cluster) measures corresponding to different b.c.s will be denoted by  $\mathcal{FK}(p, q, V)$ , and we write  $c\mathcal{FK}(p, q, V)$  for its convex hull. The stochastic process  $(\text{pr}_e)_{e \in \mathbb{E}^d(V)}: \Omega \rightarrow \Omega_V$  given on the probability space  $(\Omega, \mathcal{F}, \Phi_V^{\pi, p, q})$  is called FK percolation with boundary conditions  $\pi$ .

We will list some useful properties of FK measures. A property of crucial importance is that for  $q \geq 1$ , every  $\Phi \in \mathcal{FK}(p, q, V)$  is strong FKG. This means that for every  $\mathcal{F}_V$ -measurable cylinder  $Z$ , and for all  $\mathcal{F}_V$ -measurable increasing functions  $f, g$ , we have

$$(9) \quad \Phi[fg \mid Z] \geq \Phi[f \mid Z]\Phi[g \mid Z].$$

In some cases it is possible to compare FK measures with different b.c.s. There is a partial order on the set of partitions of  $\partial^{\text{in}}V$ . We say that  $\pi$  dominates  $\pi'$ ,  $\pi \geq \pi'$ , if  $x, y$   $\pi'$ -wired implies that they are  $\pi$ -wired. We then have  $\Phi_V^{\pi', p, q} \leq \Phi_V^{\pi, p, q}$ . This implies immediately that for each  $\Phi \in \mathcal{FK}(p, q, V)$ ,  $\Phi_V^{f, p, q} \leq \Phi \leq \Phi_V^{w, p, q}$ . Next we discuss properties of conditional FK measures. For given  $U \subseteq V$  and  $\omega \in \Omega$ , we define a partition  $W_V^U(\omega)$  of  $\partial^{\text{in}}U$  by declaring  $x, y \in \partial^{\text{in}}U$  to be  $W_V^U(\omega)$ -wired if they are joined by an  $\omega_V^U$ -open path. Fix a partition  $\pi$  of  $\partial^{\text{in}}V$ . We define a new partition of  $\partial^{\text{in}}U$ , denoted by  $\pi \cdot W_V^U(\omega)$ , by considering  $x, y \in \partial^{\text{in}}U$  to be  $\pi \cdot W_V^U(\omega)$ -wired if they are both joined by  $\omega_V^U$ -open paths to (or identical with) sites  $x', y'$ , which are themselves  $\pi$ -wired.

Then, for every  $\mathcal{F}_U$ -measurable function  $f$ ,

$$(10) \quad \Phi_V^{\pi, p, q}[f | \mathcal{F}_V^U](\omega) = \Phi_U^{\pi, W_V^U(\omega), p, q}[f], \quad \Phi_V^{\pi, p, q} \text{ a.s.}$$

Note that (10) can be interpreted as a kind of Markov property. A direct consequence of this formula is that the restriction of every FK measure  $\Phi$  in  $\mathcal{FK}(p, q, V)$  to  $\mathcal{F}_U$  is contained in the convex hull  $c\mathcal{FK}(p, q, U)$ .

*FK Ising coupling.* We describe the fundamental coupling between the FK measures and the Ising Gibbs measure (see [29, 38, 55] for more details). Let  $\Lambda \subset \mathbb{Z}^d$  be a box. An *edge-spin configuration* in  $\Lambda$  is an element  $(\omega, \sigma)$  of  $\{0, 1\}^{\mathbb{E}^d(\Lambda)} \times \{-, +\}^\Lambda$ . Let  $p$  belong to  $(0, 1)$ . Let  $\mathbb{P}_\Lambda^+$  be the probability measure on the space of edge-spin configurations in  $\Lambda$  obtained through the following procedure.

The edges in  $\Lambda$  are opened with probability  $p$  and closed with probability  $1 - p$ . The spin value of the sites in  $\partial^{\text{in}}\Lambda$  is set to  $+$ . The spin value of the sites in  $\Lambda \setminus \partial^{\text{in}}\Lambda$  is drawn randomly with the uniform distribution on  $\{-, +\}$ . The previous operations are performed independently. Finally the measure is conditioned on the event that there is no open edge in  $\Lambda$  between two sites with different spin values.

The support of  $\mathbb{P}_\Lambda^+$  consists of the edge-spin configurations  $(\omega, \sigma)$  in  $\Lambda$  such that  $\sigma(x) = +$  for any  $x \in \partial^{\text{in}}\Lambda$ , and all the sites belonging to one  $\omega$ -cluster  $C$  have the same spin value, which we denote by  $\sigma(C)$ . The first marginal of  $\mathbb{P}_\Lambda^+$  on  $\{0, 1\}^{\mathbb{E}^d(\Lambda)}$  is the FK measure  $\Phi_\Lambda^{w, p, 2}$ , its second marginal on  $\{-, +\}^\Lambda$  is the Ising Gibbs measure  $\mu_{\Lambda, \beta}^+$  where  $\beta = -\log(1 - p)$ . Therefore, to draw a spin configuration in  $\Lambda$  according to  $\mu_{\Lambda, \beta}^+$  we can proceed as follows. First, we draw an edge configuration in  $\Lambda$  according to  $\Phi_\Lambda^{w, p, 2}$ . Second, we color each open cluster independently, with  $+$  for clusters intersecting  $\partial^{\text{in}}\Lambda$  and with the uniform distribution on  $\{-, +\}$  for the other clusters.

*Coarse graining of FK processes.*

*The blocks and the rescaled lattice.* Let  $K$  be a fixed positive integer. We divide  $\mathbb{Z}^d$  into small boxes called *blocks* of size  $K$  in the following way. For  $\underline{x} \in \mathbb{Z}^d$ , we define the block indexed by  $\underline{x}$  as  $B(\underline{x}) = \Lambda(K\underline{x}, K)$ . Note that the blocks partition  $\mathbb{R}^d$  (or  $\mathbb{Z}^d$ ). Let  $A$  be a region in  $\mathbb{R}^d$  (or  $\mathbb{Z}^d$ ). Depending on the context, we define the rescaled region  $\underline{A}$  as either  $\underline{A} = \{\underline{x} | B(\underline{x}) \subseteq A\}$  or  $\underline{A} = \{\underline{x} | B(\underline{x}) \cap A \neq \emptyset\}$ . In general, we use underlining in the notation to emphasize that we are dealing with rescaled objects. For instance, we denote by  $\underline{\Lambda}(n) = \{\underline{x} | B(\underline{x}) \subseteq \Lambda(n)\}$  the box  $\Lambda(n)$  rescaled by a factor  $K$ . Clearly, the rescaled lattice inherits the structure of  $\mathbb{Z}^d$ . In particular, we can equip it with the graph structures corresponding to  $\mathbb{L}^d$  or  $\mathbb{L}^{d, \infty}$ .

With a block we often associate events which can be observed in the block or in a certain neighborhood of the block. Let  $\alpha$  be a positive integer, called the *event-block size*. For  $\underline{x} \in \mathbb{Z}^d$ , we introduce a larger block  $B'(\underline{x})$  around  $K\underline{x}$ , called the *event-block*, by setting

$$(11) \quad B'(\underline{x}) = \bigcup_{\underline{z}: d_\infty(\underline{z}, \underline{x}) < \alpha} B(\underline{z}).$$

When rescaling a finite object, like a large box  $\Lambda$ , some care is needed in the definition of the blocks sitting close to the boundary if a partition of  $\Lambda$  is desired. For  $\underline{x} \in \partial^{\text{in}} \underline{\Lambda}$ , we modify the definition of the block associated with  $\underline{x}$  as follows. First we set

$$\mathcal{H}(\underline{x}) = \{\underline{x}\} \cup \{\underline{z} \in \mathbb{Z}^d \mid d_1(\underline{z}, \underline{x}) = 1, \Lambda(K\underline{z}, K) \cap \Lambda \neq \emptyset, \Lambda(K\underline{z}, K) \cap \Lambda^c \neq \emptyset\}.$$

The block  $B(\underline{x})$  is then defined as the smallest box containing the set  $\Lambda \cap (\bigcup_{\underline{z} \in \mathcal{H}(\underline{x})} \Lambda(K\underline{z}, K))$ ; see Figure 1. Note that  $B(\underline{x})$  is a box with (in general unequal) side-lengths between  $K$  and  $2K$ , and the blocks  $B(\underline{x})$ ,  $\underline{x} \in \underline{\Lambda}$ , partition  $\Lambda$ . For  $\underline{x} \in \underline{\Lambda}$ , the event-blocks are still defined with formula (11) but using the enlarged blocks.

*Block events.* Let  $\Lambda$  be a box with side lengths between  $n$  and  $2n$ . An open cluster within  $\Lambda$  is called “crossing for  $\Lambda$ ” if it intersects each of the  $2d$  faces of  $\partial^{\text{in}} \Lambda$ . Let  $g$  be an increasing function from  $\mathbb{N}$  to  $\mathbb{R}^+$  such that  $g(n) \leq n$  for all  $n$  and let  $\delta > 0$ . We consider the following events:

$$\begin{aligned} U(\Lambda) &= \{\text{there exists a unique open crossing cluster } C^* \text{ in } \Lambda\}; \\ R(\Lambda, g) &= U(\Lambda) \cap \{\exists! \text{ open cluster with diameter } \geq g(n)\}; \\ (12) \quad O(\Lambda, g) &= R(\Lambda, g) \cap \{C^* \text{ intersects every sub-box} \\ &\quad \text{of } \Lambda \text{ of diameter } \geq g(n)\}; \\ V(\Lambda, \delta) &= U(\Lambda) \cap \{(\theta - \delta)|\Lambda| < |C^*| < (\theta + \delta)|\Lambda|\}; \\ T(\Lambda, g, \delta) &= O(\Lambda, g) \cap \{\text{for any box } \Lambda' \subset \Lambda \text{ of diameter } \geq g(n), \\ &\quad (\theta - \delta)|\Lambda'| < |C^* \cap \Lambda'| < (\theta + \delta)|\Lambda'|\}. \end{aligned}$$

Theorem 3.1 in [59] implies that for  $d \geq 3$ ,  $q \geq 1$ ,  $p > \hat{p}_c$ ,  $\theta^f(p) = \theta^w(p)$ ,  $\delta > 0$ , there exist positive constants  $b = b(p, q, d\delta)$ ,  $c = c(p, q, d\delta)$  and  $\kappa = \kappa(p, q, d\delta)$ , such that for each  $n \geq 1$ , each box with side-lengths between  $n$  and  $2n$ , and each measure  $\Phi \in c\mathcal{F}\mathcal{N}(p, q, \Lambda)$ ,

$$(13) \quad \Phi[U(\Lambda)^c] \leq \Phi[V(\Lambda, \delta)^c] \leq b \exp(-cn).$$

Moreover, if  $\kappa \log n \leq g(n) \leq n$  for all  $n$  in  $\mathbb{N}$ ,

$$(14) \quad \Phi[R(\Lambda, g)^c] \leq \Phi[O(\Lambda, g)^c] \leq \Phi[T(\Lambda, g, \delta)^c] \leq b \exp(-cg(n)).$$

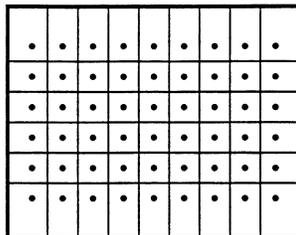


FIG. 1. A rescaled box.

Notice that we have introduced a new type of event named  $T(\Lambda, g, \delta)$ . The corresponding estimate follows from Theorems 3.1 and 1.2 in [59].

*Block variables.* In the course of the proofs we will often use coarse graining in  $\Lambda(n)$  by looking at a block process  $(X(\underline{x}))_{\underline{x} \in \Lambda(n)}$ , indicating the non-occurrence of one of the typical events listed in (12) in the corresponding event-block. By controlling the coarse-grained process  $X$  we can extract useful information about the underlying FK process; in fact that is our main tool for analyzing the microscopic behavior of the model. The definition of the events and estimates (13) and (14) guarantee that the block process satisfies the following properties:

(15) The variable  $X(\underline{x})$  depends only on the edges in  $B'(\underline{x})$ ;  

$$\max_{\Phi \in \mathcal{CFN}(p, q, B'(\underline{x}))} \Phi[X(\underline{x}) = 1] \leq \varepsilon.$$

These two properties imply furthermore

(16) 
$$\max_{\Phi \in \mathcal{CFN}(p, q, \Lambda(n))} \Phi[X(\underline{x}) = 1 \mid \sigma(X(\underline{z}), d_\infty(\underline{x}, \underline{z}) \geq 2\alpha - 1)] \leq \varepsilon.$$

The properties (15), (16) alone imply some simple estimates for the block process. For future reference we formulate them in the subsequent lemmas.

LEMMA 2.2. *Consider a 0–1 valued random field  $(X_z)_{z \in \Lambda(m)}$  with the property that there exist a positive integer  $D$  and  $\varepsilon \in [0, 1]$  such that for each  $z \in \Lambda(m)$ ,*

(17) 
$$P[X_z = 1 \mid \sigma(X_y; d_\infty(z, y) \geq D)] \leq \varepsilon.$$

Then, for every  $\delta \in (\varepsilon, 1]$ ,

$$P\left[\frac{1}{m^d} \sum_{z \in \Lambda(m)} X_z \geq \delta\right] \leq D^d \exp\left(-\Lambda_\varepsilon^*(\delta) \left\lfloor \frac{m}{D} \right\rfloor^d\right),$$

where  $\Lambda_\varepsilon^*(\delta) = \delta \log(\delta/\varepsilon) + (1 - \delta) \log[(1 - \delta)/(1 - \varepsilon)]$  is the Legendre transform of the logarithmic moment generating function of a Bernoulli variable with parameter  $\varepsilon$ . (We remark that if  $\varepsilon < \delta \leq 1/2$ , then  $\Lambda_\varepsilon^*(\delta) \geq \delta \log(\delta/\varepsilon) - \log 2$ .)

PROOF. We introduce an equivalence relation on  $\Lambda(m)$ :  $z \approx y$  iff  $D$  divides each component of  $z - y$ . The corresponding equivalence classes  $V_1, \dots, V_{D^d}$  partition  $\Lambda(m)$ . Condition (17) guarantees that each of the fields  $(X_z)_{z \in V_i}$ ,  $i = 1, \dots, D^d$ , is stochastically dominated by i.i.d. Bernoulli variables with parameter  $\varepsilon$ . By applying optimized exponential Chebyshev estimates (cf. the section on Cramér’s Theorem in [24]), and using  $|V_i| \geq \lfloor m/D \rfloor^d$  for  $i = 1, \dots, D^d$ , we arrive at the claim.  $\square$

Applying the previous lemma to a block process satisfying (16), we obtain that for every  $\Phi \in \mathcal{FK}(p, q, \Lambda(n))$  and  $\delta > \varepsilon$ ,

$$(18) \quad \Phi \left[ \frac{1}{|\underline{\Lambda}(n)|} \sum_{\underline{x} \in \underline{\Lambda}(n)} X(\underline{x}) \geq \delta \right] \leq (2\alpha)^d \exp \left( -\Lambda_\varepsilon^*(\delta) \frac{1}{(2\alpha)^d} \left\lfloor \frac{n}{K} \right\rfloor^d \right).$$

Note that this estimate is valid for each fixed  $n, \varepsilon, \alpha, K$  (with  $\alpha K < n$ ). In particular  $\varepsilon, \alpha, K$  may depend on  $n$ .

The block process can be viewed as a (dependent) site percolation process where a site  $\underline{x}$  is occupied iff  $X(\underline{x}) = 1$ . The occupied  $\infty$ -cluster of the site  $\underline{x}$ , that is, the  $\infty$ -connected component of the occupied sites containing  $\underline{x}$ , is then denoted by  $\underline{C}(\underline{x})$ .

LEMMA 2.3. *Assume (15) holds. There exists a dimension-dependent constant  $b(d) > 0$  such that, for any  $\Phi \in \mathcal{FK}(p, q, \Lambda(n))$ , any  $s, t > 0$ ,*

$$(19) \quad \begin{aligned} & \Phi[|\{\underline{x} \in \underline{\Lambda}(n) \mid |\underline{C}(\underline{x})| \geq t\}| \geq s] \\ & \leq \sum_{j \geq s} 2 \exp j(2d \log(n/K)/t + \log b + (2\alpha)^{-d} \log \varepsilon). \end{aligned}$$

For the proof, which is based on a standard counting (Peierls) argument, we refer to the proof of Lemma 7.9 in [14].

**3. Decay of boundary effects and decoupling.** The main results of this section are Proposition 3.1, which gives sufficient control of boundary effects and leads directly to Lemma 3.2, which allows decoupling rare events in distant regions. This lemma is one of the key results which allows extending the large deviations results for percolation of [14] to FK percolation. We begin with the statement of the main result and a useful consequence of it. The proof will be given at the end of the section after a series of preparatory lemmas.

Let  $\Theta \subseteq \mathbb{R}^d$  be a box building, that is, the union of finitely many  $d$ -dimensional boxes with nonempty interior. Fix a monotone increasing function  $\phi: \mathbb{N} \rightarrow \mathbb{N}$  satisfying  $\lim_{n \rightarrow \infty} \phi(n) = \infty$  and  $\lim_{n \rightarrow \infty} \phi(n)/n = 0$ . We will consider the  $\phi(n)$ -interior of the building  $n\Theta$  which is defined as

$$\text{int}(n\Theta, \phi(n)) = \{x \in n\Theta \mid d_\infty(x, \partial^{\text{in}} n\Theta) > \phi(n)\}.$$

PROPOSITION 3.1. *Assume  $d \geq 3, q \geq 1, p > \hat{p}_c$  with  $\theta^f(p) = \theta^w(p)$ . Let  $S_n$  be a sequence of events such that  $S_n$  depends only on the edges in  $\text{int}(n\Theta, \phi(n))$  and for each  $n \in \mathbb{N}$ , let  $\pi(n)$  be a partially wired b.c. on  $n\Theta$ . Then*

$$(20) \quad \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_{n\Theta}^{\pi(n), p, q}[S_n] = \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_{n\Theta}^{f, p, q}[S_n].$$

*The same equality is valid when  $\limsup$  is replaced by  $\liminf$ .*

REMARK. (i) Note that free b.c.s on the right-hand side could be replaced by wired b.c.s as well. A direct consequence of (20) is

$$\begin{aligned}
 (21) \quad \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_{n\Theta}^{f, p, q}[S_n] &= \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \sup_{\Phi \in \mathcal{C}\mathcal{F}\mathcal{K}(p, q, n\Theta)} \Phi[S_n] \\
 &= \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \inf_{\Phi \in \mathcal{C}\mathcal{F}\mathcal{K}(p, q, n\Theta)} \Phi[S_n].
 \end{aligned}$$

(ii) We will work in the box  $\Lambda(n)$  with wired boundary conditions; hence we will need the following slight generalization of Proposition 3.1, whose proof is similar. Assume  $d \geq 3$ ,  $q \geq 1$ ,  $p > \hat{p}_c$  with  $\theta^f(p) = \theta^w(p)$ . Let  $\Theta$  be a box building included in  $\Lambda(1)$ . We define the  $\phi(n)$ -interior of the building  $n\Theta$  relative to  $\Lambda(n)$  by

$$\text{int}_{\Lambda(n)}(n\Theta, \phi(n)) = \{x \in n\Theta \mid d_\infty(x, \partial^{\text{in}} n\Theta \setminus \partial^{\text{in}} \Lambda(n)) > \phi(n)\}.$$

For  $n \in \mathbb{N}$ , let  $S_n$  be an event depending only on the edges in  $\text{int}_{\Lambda(n)}(n\Theta, \phi(n))$  and let  $\pi(n)$  be a partially wired b.c. on  $n\Theta$  such that the sites in  $\partial^{\text{in}} n\Theta \cap \partial^{\text{in}} \Lambda(n)$  are wired together. We have

$$(22) \quad \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_{n\Theta}^{\pi(n), p, q}[S_n] = \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_{\Lambda(n)}^{w, p, q}[S_n].$$

LEMMA 3.2 (Decoupling lemma). *Let  $d \geq 3$ ,  $q \geq 1$ ,  $p > \hat{p}_c$  with  $\theta^f(p) = \theta^w(p)$ . Let  $D_i, i \in I$ , be a finite collection of disjoint compact subsets of the unit cube  $\Gamma = [-1/2, 1/2]^d$ . Assume that these sets have nonempty connected interiors. For  $i \in I$  let  $S_n^i$  be a sequence of events such that  $S_n^i$  depends on the edges in  $nD_i \cap \Lambda(n)$ . Then*

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_{\Lambda(n)}^w \left[ \bigcap_{i \in I} S_n^i \right] \leq \sum_{i \in I} \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_{\Lambda(n)}^w [S_n^i].$$

REMARK. The same result is valid for  $\liminf$  with the opposite inequality. Again,  $w$  could be replaced by  $f$ . The statements are valid also for the infinite volume measure.

PROOF. It is sufficient to prove the statement for two sets  $D_1$  and  $D_2$  only. We suppose that  $D_1$  and  $D_2$  are closed subsets of the interior of  $\Gamma$ . Let  $\Theta_1$  be a box building such that  $D_1 \subset \text{int } \Theta_1$ ,  $D_2 \cap \text{clo } \Theta_1 = \emptyset$ . Let  $\phi: \mathbb{N} \rightarrow \mathbb{N}$  be an increasing function such that  $\lim_{n \rightarrow \infty} \phi(n) = \infty$  and  $\lim_{n \rightarrow \infty} \phi(n)/n = 0$ . For  $n$  large enough, the event  $S_n^1$  depends only on the edges in  $\text{int}(n\Theta_1, \phi(n))$ , while  $S_n^2$  is in the  $\sigma$ -field  $\mathcal{F}_{\Lambda(n)}^{n\Theta_1}$ . Therefore

$$\begin{aligned}
 (23) \quad \Phi_{n\Gamma}^w [S_n^1 \cap S_n^2] &= \Phi_{n\Gamma}^w [1_{S_n^2} \Phi_{n\Gamma}^w [S_n^1 \mid \mathcal{F}_{\Lambda(n)}^{n\Theta_1}]] \\
 &\leq \left( \max_{\Phi \in \mathcal{C}\mathcal{F}\mathcal{K}(n\Theta_1)} \Phi[S_n^1] \right) \Phi_{n\Gamma}^w [S_n^2].
 \end{aligned}$$

Hence taking the lim sups and using (21) we obtain the claim of Lemma 3.2. For sets intersecting the boundary of  $\Gamma$ , the argument is analogous and it relies on (22).  $\square$

We next present the lemmas involved in the proof of Proposition 3.1. The first lemma is a simple observation which, nonetheless, plays a crucial role in the derivation of the results mentioned above.

LEMMA 3.3 (Monotone perturbation of boundary conditions in FK percolation). *Let  $q \geq 1$  and  $\pi' \leq \pi$  be two (partially wired) b.c.s on the region  $V \subseteq \mathbb{Z}^d$ . We denote by  $|\pi|$  the number of equivalence classes in  $\pi$  (note  $|\pi| \leq |\pi'|$ ). Assume  $|\pi'| - |\pi| \leq C$  for a certain constant  $C$ . Then for any event  $S$  depending only on the edges in  $V$ ,*

$$q^{-C} \leq \frac{\Phi_V^{\pi'}[S]}{\Phi_V^\pi[S]} \leq q^C.$$

PROOF. Note that  $|\pi'| - |\pi| \leq C$  together with  $\pi \succeq \pi'$  imply that for each  $\omega$ ,

$$0 \leq \text{cl}^{\pi'}(\omega) - \text{cl}^\pi(\omega) \leq C.$$

For any event  $A$  we define the partition sum  $Z^\pi[A]$  by

$$Z^\pi[A] = \sum_{\omega \in A} \left( \prod_{e \in E} p^{\omega(e)} (1-p)^{1-\omega(e)} \right) q^{\text{cl}^\pi(\omega)}.$$

We have the inequalities  $Z^\pi[A] \leq Z^{\pi'}[A] \leq q^C Z^\pi[A]$ . The first inequality is obvious and the second follows from

$$Z^{\pi'}[A] = \sum_{\omega \in A} \left( \prod_{e \in E} p^{\omega(e)} (1-p)^{1-\omega(e)} \right) q^{\text{cl}^\pi(\omega)} q^{\text{cl}^{\pi'}(\omega) - \text{cl}^\pi(\omega)} \leq q^C Z^\pi[A].$$

The inequalities of the lemma can be proved as follows:

$$\begin{aligned} \Phi_V^\pi[S] &= \frac{Z^\pi[S]}{Z^\pi} \leq \frac{Z^{\pi'}[S]}{q^{-C} Z^{\pi'}} = q^C \Phi_V^{\pi'}[S], \\ \Phi_V^{\pi'}[S] &= \frac{Z^{\pi'}[S]}{Z^{\pi'}} \leq \frac{q^C Z^\pi[S]}{Z^\pi} = q^C \Phi_V^\pi[S]. \end{aligned} \quad \square$$

The other ingredient of the proof of Proposition 3.1 is the control of (random) boundary conditions on  $\text{int}(n\Theta, \phi(n))$  induced by the measures  $\Phi_{n\Theta}^{\pi(n)}$ . Denote by  $A(n)$  the “boundary region”  $n\Theta \setminus \text{int}(n\Theta, \phi(n))$ . We will show that the average number of  $A(n)$ -clusters per site is concentrated around  $\kappa^+$ , which is the half-space analogue of the expected number of clusters per site  $\kappa$ . We begin the lengthy but rather simple proof with the introduction of *half-clusters*. Given an edge configuration in  $\mathbb{Z}^d$ , the half-clusters are simply the  $\mathbb{H}^d$ -clusters, that is, the open clusters of the configuration restricted to  $\mathbb{H}^d$ . For  $A \subseteq \mathbb{D}^d = \{x \in \mathbb{Z}^d \mid x_d = 0\}$  we denote by  $K'_A$  the number of half-clusters intersecting  $A$ . Finally, we set  $\kappa'_n(p, q) = \Phi_\infty^{p, q}[K'_{D(n)}/n^{d-1}]$ , where  $D(n)$  is the hypersquare  $\mathbb{D}^d \cap \Lambda(n)$ .

LEMMA 3.4. Assume  $d \geq 3, q \geq 1, 0 \leq p \leq 1, \theta^f(p) = \theta^w(p)$ . The limit  $\lim_{n \rightarrow \infty} \kappa'_n = \kappa^+$  exists and is equal to  $\inf_n \kappa'_n$ .

REMARK. If  $0 < p < 1$ , the limit  $\kappa^+$  is easily seen to be confined to  $(0,1)$ .

PROOF OF LEMMA 3.4. The proof is based on the following subadditive property: if  $A, B \subseteq \mathbb{D}^d$  and  $\omega$  is a configuration, then

$$K'_{A \cup B}(\omega) \leq K'_A(\omega) + K'_B(\omega)$$

and the same is true for the expectations. The claim of the lemma follows from a well-known  $d$ -dimensional generalization of the subadditive inequality. For the reader's convenience we sketch the argument. It is enough to show that for each fixed  $m, \limsup_{n \rightarrow \infty} \kappa'_n \leq \kappa'_m$ . Let  $n \geq m$  and divide the hypersquare  $D(n) \subseteq \mathbb{D}^d$  into disjoint hypersquares congruent to  $D(m)$ . There are  $\lfloor n/m \rfloor^{d-1}$  boxes which fit into  $D(n)$  and it remains an uncovered region of size not exceeding  $2(d-1)mn^{d-2}$ . By subadditivity and translation invariance of  $\Phi_\infty^{p,q}$ ,

$$\begin{aligned} \Phi_\infty[K'_n/n^{d-1}] &\leq \lfloor n/m \rfloor^{d-1} (m/n)^{d-1} \Phi_\infty[K'_m/m^{d-1}] \\ &\quad + \frac{1}{n^{d-1}} 2(d-1)mn^{d-2}. \end{aligned}$$

Taking  $\limsup_{n \rightarrow \infty}$  we arrive at the claim.  $\square$

We next introduce the number  $K_n$  of  $\Lambda(n) \cap \mathbb{H}^d$ -clusters intersecting the hypersquare  $D(n - \sqrt{n})$  and set  $\kappa_n = \Phi_\infty^{p,q}[K_n/n^{d-1}]$ .

LEMMA 3.5. Assume  $d \geq 3, q \geq 1, p > \hat{p}_c, \theta^f(p) = \theta^w(p)$ . Then  $\lim_{n \rightarrow \infty} \kappa_n = \kappa^+$ .

PROOF. We define the event

$$(24) \quad R_n = \{ \exists! \Lambda(n) \cap \mathbb{H}^d\text{-cluster } C^* \text{ with diameter } \geq \sqrt{n} \text{ and } C^* \cap D(n - \sqrt{n}) \neq \emptyset \}.$$

On  $R_n$ , if two sites in  $D(n - \sqrt{n})$  are connected in  $\mathbb{H}^d$  then they are already connected in  $\Lambda(n) \cap \mathbb{H}^d$ . Hence, on  $R_n$ ,

$$|K_n - K'_n| \leq |D(n) \setminus D(n - \sqrt{n})| \leq 2(d-1)n^{d-2}\sqrt{n}.$$

By Theorem 3.1 in [59] we know that  $\lim_{n \rightarrow \infty} \Phi_\infty^{p,q}[R_n^c] = 0$ . Thus

$$\begin{aligned} \limsup_{n \rightarrow \infty} |\kappa_n - \kappa'_n| &\leq \limsup_{n \rightarrow \infty} \Phi_\infty^{p,q}[n^{1-d}|K_n - K'_n|; R_n] \\ &\quad + \limsup_{n \rightarrow \infty} \Phi_\infty^{p,q}[R_n^c] = 0. \end{aligned} \quad \square$$

LEMMA 3.6. *Assume  $d \geq 3, q \geq 1, p > \hat{p}_c$  with  $\theta^f(p) = \theta^w(p)$ . The quantity  $K_n/n^{d-1}$  concentrates around  $\kappa^+$  as  $n$  goes to  $\infty$ , that is, for each  $\varepsilon > 0$ ,*

$$\lim_{n \rightarrow \infty} \max_{\Phi \in c\mathcal{F}\mathcal{N}(p, q, \Lambda(n))} \Phi \left[ \left| K_n/n^{d-1} - \kappa^+ \right| > \varepsilon \right] = 0.$$

PROOF. Let  $\varepsilon > 0$ . Since  $K_n$  is decreasing, we need only to show

$$(25) \quad \lim_{n \rightarrow \infty} \Phi_{\Lambda(n)}^w \left[ K_n/n^{d-1} < \kappa^+ - 5\varepsilon \right] = 0,$$

$$(26) \quad \lim_{n \rightarrow \infty} \Phi_{\Lambda(n)}^f \left[ K_n/n^{d-1} > \kappa^+ + 5\varepsilon \right] = 0.$$

By Theorem 3.1 in [59] and Lemma 3.5, there exists  $N = N(\varepsilon)$  such that

$$(27) \quad \max_{\Phi \in c\mathcal{F}\mathcal{N}(p, q, \Lambda(N) \cap \mathbb{H}^d)} \Phi \left[ R_N^c \right] \leq \varepsilon/2,$$

$$(28) \quad \kappa^+ - \varepsilon/2 \leq \kappa_N = \Phi_\infty \left[ K_N/N^{d-1} \right] \leq \kappa^+ + \varepsilon/2$$

[the event  $R_N$  was defined in (24)]. We require in addition that

$$(29) \quad 2(d-1)/\sqrt{N} < \varepsilon.$$

Note that  $K_n$  is also a local variable. Since according to our assumptions  $\Phi_\infty^f = \Phi_\infty^w (= \Phi_\infty)$ , then the measures  $\Phi_{\Lambda(m)}^*, * = f$  or  $w$ , converge weakly toward  $\Phi_\infty$  as  $m \rightarrow \infty$ , and there exists  $M_0 = M_0(\varepsilon, N)$  such that for every  $m \geq M_0$ ,

$$(30) \quad \begin{aligned} \Phi_{\Lambda(m)}^f \left[ K_N/N^{d-1} \right] - \varepsilon/2 &\leq \Phi_\infty \left[ K_N/N^{d-1} \right] \\ &\leq \Phi_{\Lambda(m)}^w \left[ K_N/N^{d-1} \right] + \varepsilon/2. \end{aligned}$$

Therefore, by (28) and (30), for  $M \geq M_0(\varepsilon, N)$ ,

$$(31) \quad \kappa^+ - \varepsilon \leq \Phi_{\Lambda(M)}^w \left[ K_N/N^{d-1} \right] \leq \Phi_{\Lambda(M)}^f \left[ K_N/N^{d-1} \right] \leq \kappa^+ + \varepsilon.$$

Let  $M \geq M_0(\varepsilon, N)$  be such that  $M/N$  is an odd integer greater than or equal to 3. Thus for these values  $N, M$  the conditions (27) and (31) hold true. Let  $n > M^2$  and  $l = \lfloor (n - \sqrt{n})/N \rfloor$ . We use the blocks  $B(\underline{i}) = \Lambda(\underline{i}N, N)$ ,  $\underline{i} \in D(l)$ , to rescale  $D(n - \sqrt{n})$ . For given  $\underline{i}$ , we introduce the corresponding “middle square”  $D(\underline{i}) = \underline{i}N + D(N - \sqrt{N})$  and we denote by  $K(\underline{i})$  the number of clusters in  $B(\underline{i}) \cap \mathbb{H}^d$  intersecting  $D(\underline{i})$ . We call the block  $B(\underline{i})$  regular if the event described in (24) occurs in  $B(\underline{i}) \cap \mathbb{H}^d$  [instead of  $\Lambda(n)$ ]; in this case we set  $X(\underline{i}) = 0$  and  $X(\underline{i}) = 1$  otherwise. Condition (27) guarantees that the block process  $(X(\underline{i}))_{\underline{i} \in D(l)}$  is stochastically dominated by i.i.d. Bernoulli variables with parameter  $\varepsilon/2$ .

Let  $T_n$  denote the event that the proportion of irregular blocks  $B(\underline{i})$ ,  $\underline{i} \in D(l)$ , exceeds  $\varepsilon$ . By Cramér’s theorem there exists  $c = c(\varepsilon) > 0$  such that

$$(32) \quad \Phi_{\Lambda(n)}^w \left[ T_n \right] \leq \exp(-c|D(l)|) = \exp \left( -c \lfloor (n - \sqrt{n})/N \rfloor^{d-1} \right).$$

We next observe that  $K_n \geq \sum_{\underline{i} \in D(l)} K(\underline{i})(1 - X(\underline{i})) - |D(l)|$ , since the number of clusters in  $\mathbb{H}^d \cap \Lambda(n)$  intersecting  $D(\underline{i})$ , when  $\underline{i}$  is regular, can be determined by looking at the configuration in  $\mathbb{H}^d \cap B(\underline{i})$ . Moreover, in each such block there exists at most one half-cluster which can intersect another block. Hence the maximal overcount is bounded by the number of regular blocks (which is bounded by  $|D(l)|$ .) On  $T_n^c$  we can use the following estimate:

$$(33) \quad K_n \geq \sum_{\underline{i} \in D(l)} K(\underline{i}) - |D(l)| - \varepsilon |D(l)| N^{d-1} \geq \sum_{\underline{i} \in D(l)} K(\underline{i}) - 2\varepsilon n^{d-1}.$$

The term  $\varepsilon |D(l)| N^{d-1}$  is a bound on the number of sites in  $D(n - \sqrt{n})$  not covered by regular blocks, which itself is a crude upper bound on the (extra) overcount. The second inequality follows from (29).

It will be useful to subdivide all blocks in the collection  $B(\underline{i}), \underline{i} \in D(l)$  into further classes containing blocks which are sufficiently separated in space. First we set  $B'(\underline{i}) = \Lambda(\underline{i}N, M)$ . Recall that  $M/N$  is an odd integer greater than or equal to 3. Thus the blocks  $B(\underline{k}), \underline{k} \in D(M/N)$ , partition  $B'(\underline{0})$ . For  $\underline{k} \in D(M/N)$ , we look at the collection of (indices of) blocks

$$J(\underline{k}) = \left\{ \underline{j} \in D(l) \mid \underline{j} = \underline{k} + \underline{i}(M/N) \text{ for some } \underline{i} \in \mathbb{D}^d \right\}.$$

Note that  $J(\underline{k}), \underline{k} \in D(M/N)$  is a partition of  $D(l)$  and the  $d$ -dimensional blocks  $B'(\underline{j}), \underline{j} \in J(\underline{k})$  are all disjoint and contained in  $\Lambda(n)$  (since  $\sqrt{n} > M$ .) Returning to the proof of (25),

$$(34) \quad \begin{aligned} & \limsup_{n \rightarrow \infty} \Phi_{\Lambda(n)}^w \left[ K_n / n^{d-1} < \kappa^+ - 5\varepsilon \right] \\ & \leq \limsup_{n \rightarrow \infty} \Phi_{\Lambda(n)}^w [T_n] \\ & \quad + \limsup_{n \rightarrow \infty} \Phi_{\Lambda(n)}^w \left[ \frac{1}{n^{d-1}} \sum_{\underline{k} \in D(M/N)} \sum_{\underline{j} \in J(\underline{k})} K(\underline{j}) < \kappa^+ - 3\varepsilon; T_n^c \right] \\ & \leq \limsup_{n \rightarrow \infty} \Phi_{\Lambda(n)}^w \left[ \exists \underline{k} \in D(M/N) \frac{|D(M/N)|}{n^{d-1}} \sum_{\underline{j} \in J(\underline{k})} K(\underline{j}) < \kappa^+ - 3\varepsilon \right] \\ & \leq \limsup_{n \rightarrow \infty} |D(M/N)| \\ & \quad \times \max_{\underline{k} \in D(M/N)} \Phi_{\Lambda(n)}^w \left[ \frac{1}{|J(\underline{k})|} \sum_{\underline{j} \in J(\underline{k})} K(\underline{j}) / N^{d-1} < \kappa^+ - 2\varepsilon \right]. \end{aligned}$$

(Note that  $|D(M/N)| / n^{d-1} \sim 1 / (|J(\underline{k})| N^{d-1}) \sim (M/nN)^{d-1}$  as  $n \rightarrow \infty$ .) In order to estimate the probabilities above we use the decoupling event

$$F(\underline{k}) = \bigcap_{\underline{j} \in J(\underline{k})} \{ \text{each edge connecting two sites in } \partial^{\text{in}} B'(\underline{j}) \text{ is open} \}.$$

Under the conditional measure  $\Phi_{\Lambda(n)}^w[\cdot | F(\underline{k})]$ , the variables  $K(\underline{j}) / N^{d-1}, \underline{j} \in J(\underline{k})$ , become i.i.d., and by (31) we know that their expected values are at least  $\kappa^+ - \varepsilon$ . Hence, for any  $\underline{k} \in D(M/N)$ , by the FKG inequality and Cramér's

theorem we have

$$\begin{aligned}
 & \Phi_{\Lambda(n)}^w \left[ \frac{1}{|J(\underline{k})|} \sum_{j \in J(\underline{k})} K(\underline{j})/N^{d-1} < \kappa^+ - 2\varepsilon \right] \\
 (35) \quad & \leq \Phi_{\Lambda(n)}^w \left[ \frac{1}{|J(\underline{k})|} \sum_{j \in J(\underline{k})} K(\underline{j})/N^{d-1} < \kappa^+ - 2\varepsilon \mid F(\underline{k}) \right] \\
 & \leq \exp(-c' |J(\underline{k})|)
 \end{aligned}$$

for some  $c' = c'(\varepsilon) > 0$ . Note finally that  $|J(\underline{k})| \geq (n/M)^{d-1}/2$  for  $n$  large enough, which together with (34) and (35) implies (25).

The proof of (26) is very similar; in fact it is easier since subadditivity can be used instead of the regularity argument. The block construction is the same as before except that we set  $l = \lceil (n - \sqrt{n})/N \rceil$  to have a complete cover of  $D(n - \sqrt{n})$  and we choose  $N = N(\varepsilon)$  and  $M = M(\varepsilon, N)$  such that (27) and (31) are satisfied and  $M/N$  is an odd integer  $\geq 3$ . In this case,

$$K_n \leq |D(n) \setminus D(n - \sqrt{n})| + \left| \bigcup_{i \in D(l)} (B(i) \cap \mathbb{D}^d) \setminus D(i) \right| + \sum_{i \in D(l)} K(i).$$

By (29), both  $|D(n) \setminus D(n - \sqrt{n})|$  and  $|\bigcup_{i \in D(l)} (B(i) \cap \mathbb{D}^d) \setminus D(i)|$  are bounded by  $\varepsilon n^{d-1}$ . Hence,

$$\begin{aligned}
 & \limsup_{n \rightarrow \infty} \Phi_{\Lambda(n)}^f \left[ K_n/n^{d-1} > \kappa^+ + 5\varepsilon \right] \\
 & \leq \limsup_{n \rightarrow \infty} \Phi_{\Lambda(n)}^f \left[ n^{1-d} \sum_{i \in D(l)} K(i) > \kappa^+ + 3\varepsilon \right] \\
 & \leq \limsup_{n \rightarrow \infty} \Phi_{\Lambda(n)}^f \left[ \exists \underline{k} \in D(M/N) \frac{|D(M/N)|}{n^{d-1}} \sum_{j \in J(\underline{k})} K(\underline{j}) > \kappa^+ + 3\varepsilon \right] \\
 & \leq |D(M/N)| \limsup_{n \rightarrow \infty} \max_{\underline{k} \in D(M/N)} \Phi_{\Lambda(n)}^f \left[ \frac{1}{|J(\underline{k})|} \sum_{j \in J(\underline{k})} K(\underline{j})/N^{d-1} > \kappa^+ + 2\varepsilon \right].
 \end{aligned}$$

To estimate the probabilities above, we use the decoupling event

$$F(\underline{k}) = \bigcap_{j \in J(\underline{k})} \{\text{each edge in } \mathcal{J}^{\text{edge}} B'(j) \text{ is closed}\}.$$

Note that under the conditional measure  $\Phi_{\Lambda(n)}^f[\cdot \mid F(\underline{k})]$ , the variables  $K(\underline{j})/N^{d-1}$ ,  $j \in J(\underline{k})$ , are i.i.d. with expected value smaller than  $\kappa^+ + \varepsilon$ . The proof can be finished as in the previous case by employing the FKG inequality and Cramér’s theorem.  $\square$

**PROOF OF PROPOSITION 3.1.** To preserve transparency, in particular to simplify notation, we will give the proof for the choice  $\Theta = \Lambda(1)$ ; the generalization to arbitrary buildings is straightforward. Note that for  $\Theta = \Lambda(1)$ ,  $n\Theta = \Lambda(n)$ .

We begin with the description of the idea behind the proof. Let us first recall Lemma 3.3 which says that whenever two b.c.s  $\pi, \pi'$  given on a region are not too different, more precisely when  $\pi' \leq \pi$  and the difference of the number of equivalence classes is bounded by some number  $C$ , then the corresponding probabilities of any event  $S$  given in that region are comparable in the following sense:

$$q^{-C} \leq \Phi^{\pi'}[S]/\Phi^\pi[S] \leq q^C.$$

We will use a certain monotone coupling of the two measures  $\Phi_{\Lambda(n)}^f$  and  $\Phi_{\Lambda(n)}^\pi$  where  $\pi$  is a partially wired boundary condition on  $\Lambda(n)$ . Our first goal is to show that for any  $\varepsilon > 0$ , with exceedingly high probability (i.e., up to LDs of higher than surface order) we are able to find a random (centered) box  $B$  in  $\Lambda(n)$  [only a little smaller than  $\Lambda(n)$  itself] such that the b.c.s  $W_B^f$  and  $W_B^\pi$  on  $B$ , induced by the configuration in  $\Lambda(n) \setminus B$ , satisfy  $0 \leq |W_B^f| - |W_B^\pi| \leq \varepsilon n^{d-1}$ . It is then easy to show, by using Lemma 3.3, that

$$q^{-\varepsilon n^{d-1}} \leq \Phi_{\Lambda(n)}^\pi[S]/\Phi_{\Lambda(n)}^f[S] \leq q^{\varepsilon n^{d-1}}$$

for any event  $S$  which is measurable in a region not too close to the boundary of  $\Lambda(n)$ . Since  $\varepsilon$  is arbitrarily small, the result follows easily.

We begin the proof with the description of the before-mentioned monotone coupling  $\mathbb{P}_{\Lambda(n)}$  of the measures  $\Phi_{\Lambda(n)}^f$  and  $\Phi_{\Lambda(n)}^\pi$  governing two layers of configurations  $\omega = (\omega^f, \omega^\pi)$  with  $\omega^f \leq \omega^\pi$ . We first choose an arbitrary *inward spiral* ordering of all the bonds,  $b_1, b_2, \dots$  in the box  $\Lambda(n)$  beginning with some edge on the boundary [such an edge links two sites in  $\partial^{\text{in}}\Lambda(n)$ ] and we assign i.i.d. variables  $X_i$  to the bonds  $b_i$  in  $\Lambda(n)$  which are uniformly distributed on  $[0,1]$ . By an inward spiral ordering we simply mean that for  $m = 1, \dots, n$ , each bond between sites in  $\partial^{\text{in}}\Lambda(m)$  has a smaller index than every bond linking  $\partial^{\text{in}}\Lambda(m)$  to  $\partial^{\text{in}}\Lambda(m-1)$  which themselves have smaller indices than bonds in  $\partial^{\text{in}}\Lambda(m-1)$ , etc. The coupling will be constructed in an algorithmic way as follows: one takes the first bond  $b_1$  and declares the corresponding bond on the  $*$ -layer to be open ( $Y_1^* = 1$ ) if  $X_1 \leq \Phi_{\Lambda(n)}^*[b_1 \text{ is open}]$ , otherwise  $b_1$  is closed ( $Y_1^* = 0$ ). Note that the monotonicity of the coupling is guaranteed simply by the FKG property. The second bond on the  $*$ -layer will be open or closed according to whether  $X_2 \leq \Phi_{\Lambda(n)}^*[b_2 \text{ is open} \mid \text{the status of } b_1 \text{ is given by } Y_1^*]$ . Again, the strong FKG inequality (and the relation  $Y_1^w \geq Y_1^f$ ) guarantees the monotonicity of the coupling. In general, the  $k$ th bond on the  $*$ -layer is open iff  $X_k \leq \Phi_{\Lambda(n)}^*[b_k \text{ is open} \mid \text{the status of } b_1, \dots, b_{k-1} \text{ is given by } Y_1^*, \dots, Y_{k-1}^*]$ . One proceeds in this way until all the bonds have been assigned their status. One important property of this coupling is that by construction, for each  $k$  and each configuration  $\eta = (\eta^f, \eta^w) \in \{0, 1\}^{\{b_1, \dots, b_k\}} \times \{0, 1\}^{\{b_1, \dots, b_k\}}$ , the conditional measure  $\mathbb{P}_{\Lambda(n)}[\cdot \mid \eta]$  restricted to the  $*$ -layer agrees with  $\Phi_{\Lambda(n)}^*[\cdot \mid \eta^*]$ . In particular, if  $\eta$  is a (double) configuration defined on  $\Lambda(n) \setminus \Lambda(m)$  with  $m = 1, \dots, n-1$ , the same statement is true due to the particular choice of the ordering of the edges.

Recall that  $\phi: \mathbb{N} \rightarrow \mathbb{N}^+$  is a fixed monotone increasing function such that  $\lim_{n \rightarrow \infty} \phi(n) = \infty$  and  $\lim_{n \rightarrow \infty} \phi(n)/n = 0$ . Set  $L = L(n) = \lfloor \phi(n)^{1/(d+1)} \rfloor$  if

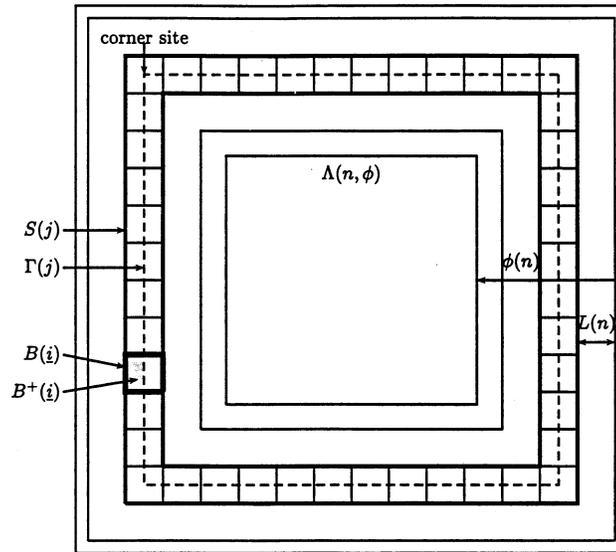


FIG. 2. The blocks and the shells.

this number is odd and let  $L = \lfloor \phi(n)^{1/(d+1)} \rfloor - 1$  otherwise. We assume that  $n$  is large enough such that  $L \geq 3$ . We divide  $\mathbb{Z}^d$  into blocks of side length  $L$  in the usual way: for  $\underline{i} \in \mathbb{Z}^d$  we set  $B(\underline{i}) = \Lambda(L\underline{i}, L)$ . Set  $\Lambda(n, \phi) = \text{int}(\Lambda(n), \phi(n))$  and let  $\underline{\Lambda}(n, \phi)$  be the set of indices  $\underline{i}$  such that  $\Lambda(n, \phi) \cap B(\underline{i}) \neq \emptyset$ . Note that  $\underline{\Lambda}(n, \phi)$  is itself a box of the form  $\Lambda(j'_0)$  for some  $j'_0 \in \mathbb{N}$ . We set  $j_0 = j'_0$  if this is an odd number and  $j_0 = j'_0 + 1$  otherwise. For  $k \geq 1$ , the set  $\underline{S}(k) = \Lambda(j_0 + 2k) \setminus \Lambda(j_0 + 2(k-1))$  is called the  $k$ th rescaled shell around  $\Lambda(n, \phi)$  and  $S(k) = \bigcup_{\underline{i} \in \underline{S}(k)} B(\underline{i})$  is the  $k$ th shell around  $\Lambda(n, \phi)$ . Note that  $S(k) \subseteq \Lambda(n) \setminus \Lambda(n, \phi)$  for  $k = 1, 2, \dots, s$ , where  $s = s(n) = \lfloor \phi(n)/(2L(n)) \rfloor - 1$ . We divide the  $k$ th rescaled shell into *corner* and *noncorner* sites (blocks) according to whether their  $d_1$  distance to  $\underline{\Lambda}(j_0 + 2(k-1))$  is equal to or strictly larger than one. In the subsequent arguments the middle hypersquares of these blocks will play an important role. They are contained in the inner vertex boundary of the box

$$\Gamma(k) = \Lambda((j_0 + 2k)L).$$

Note that  $\partial^{\text{in}}\Gamma(k)$  intersects every noncorner block of the  $k$ th shell in its middle hypersquare. We define for  $k = 1, \dots, s$ ,  $\underline{i} \in S(k)$ ,  $\underline{i}$  noncorner,

$$B^+(\underline{i}) = B(\underline{i}) \cap \left( \partial^{\text{in}}\Gamma(k) \cup (\Lambda(n) \setminus \Gamma(k)) \right),$$

$$D(\underline{i}) = \left\{ x \in B(\underline{i}) \cap \partial^{\text{in}}\Gamma(k) \mid d_\infty(x, \partial^{\text{in}}B(\underline{i})) \geq \sqrt{L} \right\},$$

$$K(\underline{i}) = \text{number of } B^+(\underline{i})\text{-clusters intersecting } D(\underline{i}).$$

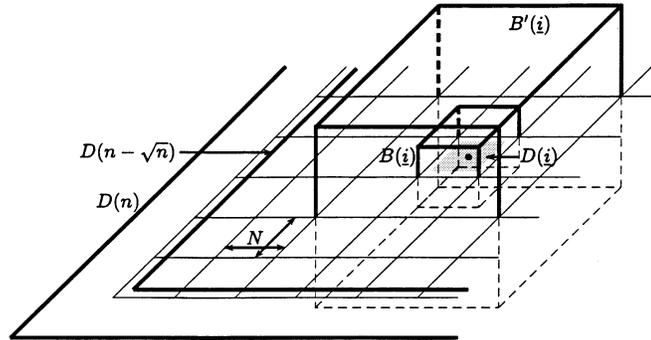


FIG. 3. The blocks  $B(\underline{i})$ ,  $B'(\underline{i})$  and the middle plane  $D(\underline{i})$ .

Finally, for  $\varepsilon > 0$ , we define the block event

$$R_{\underline{i}}(\varepsilon) = \left\{ \exists! B^+(\underline{i})\text{-cluster } C^* \text{ with diameter } \geq \sqrt{L} \right. \\ \left. \text{intersecting } D(\underline{i}), |K(\underline{i})/L^{d-1} - \kappa^+| < \varepsilon \right\}.$$

A noncorner block  $\underline{i}$  is called *regular* if  $R_{\underline{i}}(\varepsilon)$  occurs in *both* layers. We denote the corresponding block process by  $X_{\underline{i}}$ . Hence  $X_{\underline{i}} = 0$  if  $\underline{i}$  is regular and  $X_{\underline{i}} = 1$  otherwise. For  $k = 1, \dots, s$ , we set

$$G_k(\varepsilon) = \left\{ \text{the proportion of regular noncorner blocks} \right. \\ \left. \text{in the } k\text{th shell is larger than } 1 - 2\varepsilon \right\}.$$

The uniform estimates (14), Lemma 3.6 and the specific coupling guarantee that for every  $n \geq n_0(\varepsilon, d, p, q)$ , the process  $X_{\underline{i}}$ , indexed by noncorner blocks in the  $k$ th shell, is stochastically dominated by i.i.d. Bernoulli variables with parameter  $\varepsilon$ . Hence, there exists  $c = c(\varepsilon) > 0$  such that for every  $n \geq n_0$ ,

$$\begin{aligned} & \mathbb{P}_{\Lambda(n)} \left[ \bigcap_{k=1, \dots, s} G_k^c \right] \\ (36) \quad & \leq \mathbb{P}_{\Lambda(n)} \left[ \left( \text{the proportion of regular blocks in } \bigcup_{k=1, \dots, s} \underline{S}(k) \right) \leq 1 - 2\varepsilon \right] \\ & \leq \exp \left( -c' \left| \bigcup_{k=1, \dots, s} \underline{S}(k) \right| \right) \leq \exp(-c' n^{d-1} L(n)), \end{aligned}$$

where we used Cramér’s theorem. In words: we know that up to negligible events, we can find a shell in the  $\phi(n)$ -boundary of  $\Lambda(n)$  where most of the blocks are regular. As we will show, this implies that the induced boundary conditions in each layer are close to each other. For  $* = f$  or  $\pi$ , we denote by  $W_{\Gamma(k)}^*(\omega^*)$  the induced b.c.s  $* \cdot W_{\Lambda(n)}^{\Gamma(k)}(\omega^*)$  (see Section 2 for the notation).

Note that these are random b.c.s on  $\Gamma(k)$  which arise as a combination of the b.c.'s on  $\Lambda(n)$  and of the configuration occurring in the annulus

$$A(k) = \Lambda(n) \setminus \Gamma(k).$$

We claim that there exists  $n_0 = n_0(\varepsilon, d, \phi)$  such that for every  $n \geq n_0$  and  $k = 1, \dots, s$ , we have on the event  $G_k$ ,

$$(37) \quad 0 \leq |W_{\Gamma(k)}^f| - |W_{\Gamma(k)}^\pi| \leq 9\varepsilon 2d n^{d-1}.$$

Note that the first inequality is a simple consequence of the monotonicity of the coupling. To prove the other bound we first observe that if  $B(\underline{i})$  is a regular noncorner block in the  $k$ th shell then for  $* = f$  and  $* = \pi$ ,  $K(\underline{i})(\omega^*)$  is equal to the number of  $A(k)$ -clusters on the  $*$ -layer which intersect  $D(\underline{i})$ . The next step is to derive an appropriate upper bound on the number  $|W_{\Gamma(k)}^f(\omega)|$  of  $A(k)$ -clusters intersecting  $\partial^{\text{in}}\Gamma(k)$  in the  $f$  layer, when  $\omega \in G_k$ . We estimate it (using subadditivity) by the sum of the numbers of  $A(k)$ -clusters intersecting  $\partial^{\text{in}}\Gamma(k)$ :

- in corner blocks,  $= N_1$ ;
- in irregular non-corner blocks,  $= N_2$ ;
- in regular noncorner blocks,  $B(\underline{i})$  “outside”  $D(\underline{i})$ ,  
that is, in  $(\partial^{\text{in}}\Gamma(k) \cap B(\underline{i})) \setminus D(\underline{i})$ ,  $= N_3$ ;
- in  $D(\underline{i})$  when  $B(\underline{i})$  is a regular noncorner block,  $= N_4$ .

The corresponding estimates are as follows:

$$\begin{aligned} N_1 &\leq \varepsilon n^{d-1}; \text{ for } n \text{ large enough depending on } \varepsilon, d, \phi; \\ N_2 &\leq 2d2\varepsilon(n/L(n))^{d-1}L(n)^{d-1} = 4d\varepsilon n^{d-1}; \text{ on } G_k; \\ N_3 &\leq 2d(n/L(n))^{d-1} \times L(n)^{d-2} \sqrt{L(n)} < \varepsilon n^{d-1} \\ &\quad \text{for } n \text{ large enough depending on } \varepsilon, d, \phi; \\ N_4 &\leq \sum_{\substack{\underline{i} \in \underline{S}(k), B(\underline{i}) \text{ is} \\ \text{regular, noncorner}}} K(\underline{i}) \leq L(n)^{d-1}(\kappa^+ + \varepsilon)2d(n/L(n))^{d-1} \\ &= 2d(\kappa^+ + \varepsilon)n^{d-1}. \end{aligned}$$

This implies

$$(38) \quad |W_{\Gamma(k)}^f| \leq N_1 + N_2 + N_3 + N_4 \leq 2d(\kappa^+ + 4\varepsilon)n^{d-1}.$$

To get a lower bound on  $|W_{\Gamma(k)}^\pi(\omega)|$  when  $\omega \in G_k$  we first observe that in a regular noncorner block  $B(\underline{i})$ ,  $\underline{i} \in \underline{S}(k)$  there are  $K(\underline{i}) - 1 \geq (\kappa^+ - \varepsilon)L^{d-1} - 1$   $A(k)$ -clusters with diameter strictly smaller than  $\sqrt{L}$  intersecting  $D(\underline{i})$ . Note that these clusters cannot intersect any other block  $B(\underline{j})$ ,  $\underline{j} \in \underline{S}(k) \setminus \{\underline{i}\}$ . Thus the total number of  $A(k)$ -clusters intersecting  $\partial^{\text{in}}\Gamma(k)$  is certainly not smaller than the sum of the numbers of  $A(k)$ -clusters with diameter strictly smaller

than  $\sqrt{L}$  intersecting  $D(\underline{i})$  for some  $\underline{i} \in \underline{S}(k)$ ,  $B(\underline{i})$  noncorner and regular. The number of such blocks is bounded from below by

$$\left(\frac{n - 2\phi(n)}{L(n)}\right)^{d-1} 2d(1 - 2\varepsilon) > (1 - 3\varepsilon)2d(n/L(n))^{d-1},$$

where  $n$  is large enough (depending on  $\varepsilon, d, \phi$  only). Hence for  $\omega \in G_k$  and  $n$  large,

$$|W_{\Gamma(k)}^\pi(\omega)| \geq (1 - 3\varepsilon)2d(n/L)^{d-1}((\kappa^+ - \varepsilon)L^{d-1} - 1) \geq 2d(\kappa^+ - 5\varepsilon)n^{d-1}.$$

This, together with (38) gives (37).

The remainder of the proof of Proposition 3.1 is based on Lemma 3.3. For an integer  $m < n$  we set

$$J_m = J_m(\varepsilon) = \left\{ \omega \mid |W_{\Lambda(m)}^f| - |W_{\Lambda(m)}^\pi| < \varepsilon n^{d-1} \right\}.$$

By (36) and (37) we know that

$$(39) \quad \limsup_{n \rightarrow \infty} n^{1-d} \log \mathbb{P}_{\Lambda(n)} \left[ \left( \bigcup_{n-\phi(n) < m < n} J_m \right)^c \right] = -\infty.$$

Let  $S_n$  be an event which depends only on the edges in  $\Lambda(n, \phi)$ . In the coupled model denote by  $S_n^*$ ,  $*$  =  $f$  or  $\pi$ , the event that  $S_n$  occurs on the  $*$ -layer. Let  $\eta = (\eta^f, \eta^\pi)$  be a configuration in  $\Lambda(n) \setminus \Lambda(m)$  and assume  $\eta \in J_m(\varepsilon)$ . Then

$$\mathbb{P}_{\Lambda(n)}[S_n^* | \eta] = \Phi_{\Lambda(n)}^*[S_n | \eta^*] = \Phi_{\Lambda(m)}^{W_{\Lambda(m)}^*(\eta^*)}[S_n].$$

Recall that the first equality is a consequence of the specific choice of the coupling and the second one follows from (10). By Lemma 3.3 we have

$$(40) \quad q^{-\varepsilon n^{d-1}} \mathbb{P}_{\Lambda(n)}[S_n^\pi | \eta] \leq \mathbb{P}_{\Lambda(n)}[S_n^f | \eta] \leq q^{\varepsilon n^{d-1}} \mathbb{P}_{\Lambda(n)}[S_n^\pi | \eta].$$

Hence

$$\begin{aligned} \mathbb{P}_{\Lambda(n)}[S_n^f \cap J_m] &= \sum_{\substack{\eta: \text{config. on } \Lambda(n) \setminus \Lambda(m) \\ \{\eta\} \subseteq J_m}} \mathbb{P}_{\Lambda(n)}[S_n^f | \eta] \mathbb{P}_{\Lambda(n)}[\eta] \\ (41) \quad &\leq \sum_{\{\eta\} \subseteq J_m} q^{\varepsilon n^{d-1}} \mathbb{P}_{\Lambda(n)}[S_n^\pi | \eta] \mathbb{P}_{\Lambda(n)}[\eta] \\ &= q^{\varepsilon n^{d-1}} \mathbb{P}_{\Lambda(n)}[S_n^\pi \cap J_m] \leq q^{\varepsilon n^{d-1}} \mathbb{P}_{\Lambda(n)}[S_n^\pi]. \end{aligned}$$

Therefore, by using (39) and (41),

$$\begin{aligned} \limsup_{n \rightarrow \infty} n^{1-d} \log \Phi_{\Lambda(n)}^f[S_n] &= \limsup_{n \rightarrow \infty} n^{1-d} \log \mathbb{P}_{\Lambda(n)} \left[ S_n^f \cap \left( \bigcup_{n-\phi(n) < m < n} J_m \right) \right] \\ &\leq \limsup_{n \rightarrow \infty} n^{1-d} \log \left( \phi(n) \max_{n-\phi(n) < m < n} \mathbb{P}_{\Lambda(n)}[S_n^f \cap J_m] \right) \\ &\leq \limsup_{n \rightarrow \infty} n^{1-d} \log q^{\varepsilon n^{d-1}} \mathbb{P}_{\Lambda(n)}[S_n^{\pi(n)}] \\ &\leq \varepsilon \log q + \limsup_{n \rightarrow \infty} n^{1-d} \log \Phi_{\Lambda(n)}^{\pi(n)}[S_n]. \end{aligned}$$

Since  $\varepsilon > 0$  can be chosen arbitrarily, we have

$$(42) \quad \limsup_{n \rightarrow \infty} n^{1-d} \log \Phi_{\Lambda(n)}^f[S_n] \leq \limsup_{n \rightarrow \infty} n^{1-d} \log \Phi_{\Lambda(n)}^{\pi(n)}[S_n].$$

By interchanging the roles of  $f$  and  $\pi(n)$  and using the first inequality in (40) we derive similarly the opposite inequality which gives the claim (20). The derivation of the inequality involving  $\liminf$  is analogous.  $\square$

**4. Surface tension.** We will have to work with enlargements of continuous subsets of  $\mathbb{R}^d$  so that they have a significant trace on the discrete lattice  $\mathbb{Z}^d$ . We fix a real number  $\zeta > 2d$  and we enlarge a subset  $A$  of  $\mathbb{R}^d$  by considering its  $\zeta$ -neighborhood  $\mathcal{V}_2(A, \zeta)$ . A minimal requirement for choosing  $\zeta$  is that, whenever  $A$  is an arcwise connected subset of  $\mathbb{R}^d$ , the graph  $(\mathbb{Z}^d \cap \mathcal{V}_2(A, \zeta), \mathbb{E}^d(\mathcal{V}_2(A, \zeta)))$  is also connected. Some of the constants appearing in the statements and the proofs depend on  $\zeta$ . However the direction-dependent surface tension and the probabilistic estimates are independent of the particular choice of  $\zeta > 2d$ .

We next identify the surface tension of the model, whose existence is guaranteed by subadditivity and the FKG inequality. We work here in the following regime:  $d \geq 3$ ,  $q \geq 1$ ,  $p > \hat{p}_c$ ,  $\theta^f(p) = \theta^w(p)$ . We study the surface tension of the infinite volume FK measure  $\Phi_\infty$ .

Let  $A$  be a closed hyperrectangle and let  $s$  be positive or infinite. Recall that  $\text{cyl } \partial A = \{x+t \text{ nor } A \mid x \in \partial A, t \in \mathbb{R}\}$ . We denote by  $W(\partial A, s, \zeta)$  the event that there exists a finite set of closed edges  $E \subseteq \mathcal{V}_2(\text{hyp } A, s)$  such that:

1.  $E$  separates  $\infty$  in  $\text{cyl } A$ .
2. The edges in  $E \cap \mathcal{V}_2(\text{cyl } \partial A, \zeta)$  are closed to  $\text{hyp } A$  so that they are contained in  $\mathcal{V}_2(\text{hyp } A, \zeta)$ .

Loosely speaking, the second condition means that the “boundary” of  $E$  is “pinned down” at  $\partial A$  within a distance  $\zeta$ . Note that the event  $W(\partial A, s, \zeta)$  is decreasing and it depends only on edges inside  $\text{cyl } A \cap \mathcal{V}_2(\text{hyp } A, s)$ .

PROPOSITION 4.1. *Let  $A$  be a hyperrectangle. Let  $\phi(n)$  be a function from  $\mathbb{N}$  to  $\mathbb{R}^+ \cup \{\infty\}$  such that  $\lim_{n \rightarrow \infty} \phi(n) = \infty$ . The limit*

$$\lim_{n \rightarrow \infty} -\frac{1}{\mathcal{H}^{d-1}(nA)} \log \Phi_\infty[W(\partial nA, \phi(n), \zeta)]$$

*exists  $\in [0, \infty]$  and depends only on  $\text{nor } A$ . We denote it by  $\tau(\text{nor } A)$  and call it the surface tension in the direction  $\text{nor } A$ .*

For the proof, see Section A.2.

REMARK. It turns out that our definition in the case of integer  $q$  agrees with the classical one for spin systems. We are planning to give a proof in a subsequent paper [15].

*Properties.* We derive next some basic properties of the surface tension. In the context of lattice spin systems, where the definition of surface tension is significantly different, the corresponding properties have been derived in [53]. Although the techniques are not original (apart perhaps from Proposition 4.2), for the sake of completeness, we include the proofs in Section A.2.

The surface tension  $\tau$  inherits automatically some symmetry properties from the model. For instance, if  $f$  is a linear isometry of  $\mathbb{R}^d$  such that  $f(0) = 0$  and  $f(\mathbb{Z}^d) = \mathbb{Z}^d$  then  $\tau \circ f = \tau$ . Besides, the surface tension  $\tau$  satisfies another important inequality called the weak triangle inequality. For details and results concerning this kind of inequality, see [28, 53, 54].

PROPOSITION 4.2 (Weak triangle inequality). *Let  $(ABC)$  be a nondegenerate triangle in  $\mathbb{R}^d$  and let  $v_A, v_B, v_C$  be the exterior normal unit vectors to the sides  $[BC], [AC], [AB]$  in the plane spanned by  $A, B, C$ . Then*

$$\mathcal{H}^1([BC])\tau(v_A) \leq \mathcal{H}^1([AC])\tau(v_B) + \mathcal{H}^1([AB])\tau(v_C).$$

PROOF. We consider first the case where  $BA \cdot BC \geq 0$  and  $CA \cdot CB \geq 0$ . Let  $\varepsilon, h$  be positive with  $\varepsilon \leq 1 \leq h$ . Let  $(e_1, \dots, e_d)$  be an orthonormal basis of  $\mathbb{R}^d$  such that  $e_1, e_2$  belong to the two-dimensional space spanned by  $A, B, C$ . Let  $K$  be the compact convex set defined by

$$K = \left\{ x + \sum_{3 \leq i \leq d} u_i e_i \mid x \in (ABC), (u_3, \dots, u_d) \in [0, h]^{d-2} \right\}.$$

The boundary of  $K$  consists of the three hyperrectangles  $R_A, R_B, R_C$  defined by

$$R_A = \left\{ x + \sum_{3 \leq i \leq d} u_i e_i \mid x \in [BC], (u_3, \dots, u_d) \in [0, h]^{d-2} \right\},$$

$$R_B = \left\{ x + \sum_{3 \leq i \leq d} u_i e_i \mid x \in [AC], (u_3, \dots, u_d) \in [0, h]^{d-2} \right\},$$

$$R_C = \left\{ x + \sum_{3 \leq i \leq d} u_i e_i \mid x \in [AB], (u_3, \dots, u_d) \in [0, h]^{d-2} \right\}$$

and the set

$$T = \bigcup_{3 \leq j \leq d} \left\{ x + \sum_{3 \leq i \leq d} u_i e_i \mid x \in (ABC), u_j \in \{0, h\}, \right. \\ \left. (u_3, \dots, u_{j-1}, u_{j+1}, \dots, u_d) \in [0, h]^{d-3} \right\}.$$

Notice that the set  $T$  is connected in dimension  $d \geq 4$  and consists of two disjoint triangles in dimension  $d = 3$ . The intersection of the hyperrectangles  $R_A$  and  $R_B$  is a  $d - 2$ -dimensional rectangle and it is denoted by  $R_{A, B}$ . Similar notation is used for the other intersections. Let  $E_0$  be the set of the edges included in

$$\left( \text{cyl}(\text{hyp } nR_A \cap \mathcal{V}_2(\partial nR_A, 4\varepsilon n)) \cap \mathcal{V}_2(\text{hyp } nR_A, \zeta) \right) \\ \cup \mathcal{V}_2(nR_{B,C}, 2\zeta) \cup \mathcal{V}_2(nR_{A,C}, 2\zeta) \cup \mathcal{V}_2(nR_{A, B}, 2\zeta) \cup \mathcal{V}_2(nT, 2\zeta).$$

There exists a constant  $c = c(d, \zeta)$  such that

$$(43) \quad |E_0| \leq c \left( \varepsilon n^{d-1} h^{d-2} + (nh)^{d-2} + 2(d-2)n^{d-1} h^{d-3} \right).$$

Let  $R^\varepsilon$  be a hyperrectangle in  $\text{hyp } R_A$  such that  $R_A \subset R^\varepsilon \subset \mathcal{V}_2(R_A, 4\varepsilon)$  and  $d_2(\partial R^\varepsilon, R_A) > 2\varepsilon$ . For  $n$  large enough, so that  $\varepsilon n > \zeta$ , if the events

$$W(\partial nR_C, \varepsilon n, \zeta), \quad W(\partial nR_B, \varepsilon n, \zeta), \quad \{\text{all the edges of } E_0 \text{ are closed}\}$$

occur simultaneously, then the event  $W(\partial nR^\varepsilon, \infty, \zeta)$  occurs as well; by the assumptions  $BA \cdot BC \geq 0$  and  $CA \cdot CB \geq 0$ , the set  $\mathcal{V}_2(nR_C \cup nR_B, \varepsilon n)$  is included in  $\mathcal{V}_2(\text{cyl } nR_A, \varepsilon n)$  and does not intersect  $\mathcal{V}_2(\text{cyl } \partial nR^\varepsilon, \zeta)$ , so that the separating sets will be correctly localized. By the FKG inequality, this inclusion implies

$$(44) \quad (1-p)^{|E_0|} \Phi_\infty[W(\partial nR_C, \varepsilon n, \zeta)] \Phi_\infty[W(\partial nR_B, \varepsilon n, \zeta)] \\ \leq \Phi_\infty[W(\partial nR^\varepsilon, \infty, \zeta)].$$

The inequalities (43), (44) and Proposition 4.1 yield

$$(45) \quad \mathcal{H}^{d-1}(R^\varepsilon) \tau(v_A) \leq \mathcal{H}^{d-1}(R_C) \tau(v_C) + \mathcal{H}^{d-1}(R_B) \tau(v_B) \\ - c(\varepsilon h^{d-2} + 2(d-2)h^{d-3}) \log(1-p).$$

We observe

$$\mathcal{H}^{d-1}(R_B) = h^{d-2} \mathcal{H}^1([AC]), \\ \mathcal{H}^{d-1}(R_C) = h^{d-2} \mathcal{H}^1([AB]), \quad h^{d-2} \mathcal{H}^1([BC]) \leq \mathcal{H}^{d-1}(R^\varepsilon).$$

By substituting this into the inequality (45) and dividing by  $h^{d-2}$ ,

$$\mathcal{H}^1([BC]) \tau(v_A) \leq \mathcal{H}^1([AC]) \tau(v_B) + \mathcal{H}^1([AB]) \tau(v_C) \\ - c(\varepsilon + 2(d-2)/h) \log(1-p).$$

By letting  $h$  go to  $\infty$  and  $\varepsilon$  go to 0, we obtain the weak triangle inequality for the triangle  $(ABC)$ . Let now  $A, B, C$  be three points such that  $BA \cdot BC < 0$ ,  $CA \cdot CB \geq 0$ . Let  $D$  be the orthogonal projection of  $B$  on  $[AC]$ . Then  $BC \cdot BD > 0$ ,  $DB \cdot DA = 0$ ,  $BA \cdot BD > 0$ . We apply the weak triangle inequality to the triangles  $(BCD)$  and  $(BDA)$ :

$$\begin{aligned} \mathcal{H}^1([BC])\tau(v_A) &\leq \mathcal{H}^1([BD])\tau(v_{BD}) + \mathcal{H}^1([DC])\tau(v_B), \\ \mathcal{H}^1([BD])\tau(v_{BD}) &\leq \mathcal{H}^1([AB])\tau(v_C) + \mathcal{H}^1([AD])\tau(v_B), \end{aligned}$$

where  $v_{BD}$  is a unit vector orthogonal to  $[BD]$ . Combining the two inequalities, we get the weak triangle inequality for the triangle  $(ABC)$ . The case  $BA \cdot BC \geq 0$ ,  $CA \cdot CB < 0$  is similar.  $\square$

The weak triangle inequality implies a lot of nice properties for the surface tension.

**COROLLARY 4.3.** *The homogeneous extension  $\tau_0$  of  $\tau$  to  $\mathbb{R}^d$  defined by  $\tau_0(0) = 0$  and*

$$\forall w \in \mathbb{R}^d \setminus \{0\}, \quad \tau_0(w) = |w|_2 \tau(w/|w|_2)$$

*is a convex function.*

The convexity of  $\tau_0$  is in fact equivalent to the weak triangle inequality. The next corollary is a consequence of [28], Theorem 3.1: the weak triangle inequality automatically implies the weak simplex inequality.

**COROLLARY 4.4 (Weak simplex inequality).** *Let  $A_0, \dots, A_d$  be  $d + 1$  points in general position in  $\mathbb{R}^d$ . For  $i$  in  $\{0 \cdots d\}$  let  $\Delta(i)$  be the hypersimplex defined by the points  $\{A_0, \dots, A_d\} \setminus \{A_i\}$ . Up to the sign, there exists a unique family of unit vectors  $v_0, \dots, v_d$  such that for  $i$  in  $\{0 \cdots d\}$ , the vector  $v_i$  is orthogonal to the vector space spanned by the hypersimplex  $\Delta(i)$ . Then*

$$\mathcal{H}^{d-1}(\Delta(0))\tau(v_0) \leq \mathcal{H}^{d-1}(\Delta(1))\tau(v_1) + \dots + \mathcal{H}^{d-1}(\Delta(d))\tau(v_d).$$

**PROPOSITION 4.5.** *The surface tension  $\tau : S^{d-1} \mapsto \mathbb{R}^+$  is bounded, continuous and does not vanish.*

For the proof, see Section A.2.

The previous properties of  $\tau$  can equivalently be described through its Wulff crystal,

$$\mathcal{W}_\tau = \{x \in \mathbb{R}^d \mid x \cdot w \leq \tau(w) \text{ for all } w \text{ in } S^{d-1}\}.$$

**COROLLARY 4.6.** *The Wulff crystal  $\mathcal{W}_\tau$  associated with  $\tau$  is bounded, closed, convex and contains 0 in its interior. If  $f$  is a linear isometry of  $\mathbb{R}^d$  such that  $f(0) = 0$  and  $f(\mathbb{Z}^d) = \mathbb{Z}^d$  then  $f(\mathcal{W}_\tau) = \mathcal{W}_\tau$ . The surface tension  $\tau$  is the support function of its Wulff crystal, that is,*

$$\forall \nu \in S^{d-1}, \quad \tau(\nu) = \sup\{x \cdot \nu \mid x \in \mathcal{W}_\tau\}.$$

These properties are equivalent to the symmetry properties of  $\tau$  and Corollary 4.3, Proposition 4.5. The function  $\tau$  is the support function of  $\mathscr{H}_\tau$  because  $\tau_0$  is convex and coincides with its bipolar; see, for instance, [60], Corollary 13.2.1, [36], Proposition 3.5, or [28], Theorem 2.1, Corollary 3.6.

*Separating sets.* With the help of the surface tension, we next estimate the probability of the occurrence of a separating set of closed edges near an hyperplane. Let  $A$  be a hyperset in  $\mathbb{R}^d$  and let  $r$  be positive or infinite. We denote by  $S(A, r)$  the event that there exists a finite set of closed edges in  $\text{cyl } A \cap \mathscr{H}_2(\text{hyp } A, r)$  which separates  $\infty$  in  $\text{cyl } A$ ; that is,

$$S(A, r) = \{ \exists E \subset \text{cyl } A \cap \mathscr{H}_2(\text{hyp } A, r), \\ |E| < \infty, \forall e \in E, \omega(e) = 0, E \text{ separates } \infty \text{ in } \text{cyl } A \}.$$

From now on, we work with a fixed value of  $\zeta$  larger than  $2d$  and we drop  $\zeta$  in the notation  $W(\partial A, s, \zeta)$ , thus writing simply  $W(\partial A, s)$ . The proofs of Lemmas 4.7, 4.8, 4.9 and Corollary 4.10 are to be found in Section A.2.

LEMMA 4.7. *Let  $O$  be an open hyperset in  $\mathbb{R}^d$  and let  $\phi(n)$  be a function from  $\mathbb{N}$  to  $\mathbb{R}^+ \cup \{\infty\}$  such that  $\lim_{n \rightarrow \infty} \phi(n) = \infty$ . We have*

$$\liminf_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_\infty [S(nO, \phi(n))] \geq -\mathscr{H}^{d-1}(O)\tau(\text{nor } O).$$

LEMMA 4.8. *There exists a positive constant  $c = c(p, d, \zeta)$  such that, for any hyperrectangle  $A$  of  $\mathbb{R}^d$ , for any positive  $r$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_\infty [S(nA, nr)] \leq -\tau(\text{nor } A)\mathscr{H}^{d-1}(A) + cr\mathscr{H}^{d-2}(\partial A).$$

LEMMA 4.9. *There exists a positive constant  $c = c(p, d, \zeta)$  such that for any open hyperset  $O$  in  $\mathbb{R}^d$ , for any finite family  $(A_i, i \in I)$  of disjoint hyperrectangles included in  $O$ , for any positive  $r$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_\infty [S(nO, nr)] \\ \leq -\tau(\text{nor } O) \sum_{i \in I} \mathscr{H}^{d-1}(A_i) + cr \sum_{i \in I} \mathscr{H}^{d-2}(\partial A_i).$$

COROLLARY 4.10. *There exists a positive constant  $c = c(p, d, \zeta)$  such that, for any  $x$  in  $\mathbb{R}^d$ , any positive  $\rho, \eta$  with  $\eta \leq \rho$ , any  $w$  in  $S^{d-1}$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_\infty [S(n \text{ disc}(x, \rho, w), n\eta)] \leq -\alpha_{d-1}\rho^{d-1}\tau(w) + c\eta\rho^{d-2}.$$

**5. The surface energy.** We work here within the region  $d \geq 3, q \geq 1, p > \hat{p}_c, \theta^f(p) = \theta^w(p)$ . With the help of the surface tension  $\tau$  defined in Proposition 4.1, or equivalently its Wulff crystal  $\mathscr{H}_\tau$ , we build a surface energy

functional defined on the collection of the Borel sets of  $\mathbb{R}^d$ . The surface energy  $\mathcal{I}(A)$  of a Borel set  $A$  is defined as

$$\mathcal{I}(A) = \sup \left\{ \int_A \operatorname{div} f(x) \, dx \mid f \in C_0^1(\mathbb{R}^d, \mathcal{W}_\tau) \right\},$$

where  $C_0^1(\mathbb{R}^d, \mathcal{W}_\tau)$  is the set of the compactly supported  $C^1$  vector fields taking values in the Wulff crystal  $\mathcal{W}_\tau$  and  $\operatorname{div}$  is the usual divergence operator. The definition readily implies that the map  $\mathcal{I}: (\mathcal{B}(\mathbb{R}^d), \operatorname{dist}_{L^1}) \rightarrow [0, \infty]$  is lower semicontinuous. We denote by  $\mathcal{P}(A)$  the isotropic perimeter of a set  $A$ , introduced by Caccioppoli [12, 13]. Some references and basic results on the theory of sets of finite perimeter are recalled in Section A.1. In case  $A$  is a set of finite perimeter, we denote by  $\partial^* A$  its *reduced boundary*. Let  $\tau_{\min}$  and  $\tau_{\max}$  be the infimum and supremum of  $\tau$  over  $S^{d-1}$ . Then  $B(\tau_{\min}) \subseteq \mathcal{W}_\tau \subseteq B(\tau_{\max})$  so that for every  $A \in \mathcal{B}(\mathbb{R}^d)$ ,

$$(46) \quad \tau_{\min} \mathcal{P}(A) \leq \mathcal{I}(A) \leq \tau_{\max} \mathcal{P}(A).$$

By Proposition 4.5, we have  $0 < \tau_{\min} \leq \tau_{\max} < \infty$ . Thus a set  $A$  has finite surface energy if and only if it has finite perimeter. In this case

$$\mathcal{I}(A) = \int_{\partial^* A} \tau(\nu_A(x)) \mathcal{H}^{d-1}(dx)$$

(see Proposition 6.5 in [14] for a detailed proof). Of crucial importance is the following compactness result.

*Compactness of the level sets of  $\mathcal{I}$  in  $\Gamma$ .* The sets  $\{E \in \mathcal{B}(\Gamma) \mid \mathcal{I}(E) \leq \lambda\}$ ,  $\lambda > 0$ , are compact.

Since  $\mathcal{I}$  is lower semicontinuous, its level sets are closed. By inequality (46), the level set of  $\mathcal{I}$  in  $\Gamma$  associated to  $\lambda$  is included in  $\{E \in \mathcal{B}(\Gamma) : \mathcal{P}(E) \leq \lambda/\tau_{\min}\}$ , which is compact for the metric  $\operatorname{dist}_{L^1}$  (see Section A.1).

We state next the geometric approximation results needed to prove the lower bound (2) and the local upper bound (5).

The lower bound (2) relies on the possibility of approximating the Wulff crystal by a polyhedral set. A Borel subset of  $\mathbb{R}^d$  is polyhedral if its boundary is included in the union of a finite number of hyperplanes.

LEMMA 5.1. *For any  $\varepsilon > 0$ , there exists a polyhedral convex set  $E$  such that  $\mathcal{W}_\tau \subset E \subset \mathcal{V}_2(\mathcal{W}_\tau, \varepsilon)$  and  $\mathcal{I}(E) < \mathcal{I}(\mathcal{W}_\tau) + \varepsilon$ .*

REMARK. This result is stronger than the general polyhedral approximation result for Caccioppoli sets (see [14], Proposition 6.9). Indeed, we are able to approximate the Wulff crystal from the outside and with respect to the Hausdorff metric, thanks to its convexity.

PROOF OF LEMMA 5.1. We first recall some results from the theory of convex sets (for details, see [61]). Let  $A$  be a convex compact set with nonempty interior. A point  $x$  of  $\partial A$  is called “smooth” if the supporting hyperplane of  $A$  at  $x$  is unique. By a classical result due to Reidemeister (see [61],

Theorem 2.2.4), the boundary of a convex compact set admits  $\mathcal{H}^{d-1}$ -almost everywhere a unique supporting hyperplane. In particular,  $\mathcal{H}^{d-1}(\partial A \setminus \partial^* A) = 0$ . The area measure  $S_{d-1}(A, \cdot)$  associated to  $A$  is the Borel measure on the sphere  $S^{d-1}$  defined by

$$\forall E \in \mathcal{B}(S^{d-1}), \quad S_{d-1}(A, E) = \mathcal{H}^{d-1}(\{x \in \partial A \mid \nu_A(x) \in E\}).$$

We can express the surface energy  $\mathcal{J}(A)$  of  $A$  with the help of  $S_{d-1}(A, \cdot)$ :

$$\mathcal{J}(A) = \int_{S^{d-1}} \tau(\nu) S_{d-1}(A, d\nu).$$

Moreover, if  $(A_n)_{n \in \mathbb{N}}$  is a sequence of convex compact sets converging for the Hausdorff metric to a convex compact set  $A$  having nonempty interior, then the sequence of the area measures  $S_{d-1}(A_n, \cdot)_{n \in \mathbb{N}}$  converges weakly to  $S_{d-1}(A, \cdot)$  (see, e.g., the proof of Theorem 4.2.5 in [61]).

Let  $(\nu_n)_{n \in \mathbb{N}}$  be a dense subset of  $S^{d-1}$  and let for  $n$  in  $\mathbb{N}$ ,

$$A_n = \bigcap_{1 \leq i \leq n} \{x \in \mathbb{R}^d \mid x \cdot \nu_i \leq \tau(\nu_i)\}.$$

Since  $\tau$  is bounded (by Proposition 4.5), then for  $n$  large enough, the set  $A_n$  is convex polyhedral and compact. The sequence  $(A_n)_{n \in \mathbb{N}}$  is decreasing and converges to  $\mathcal{W}_\tau$ . Therefore  $S_{d-1}(A_n, \cdot)_{n \in \mathbb{N}}$  converges weakly to  $S_{d-1}(\mathcal{W}_\tau, \cdot)$ . Since the surface tension  $\tau$  is continuous (Proposition 4.5), the surface energy  $\mathcal{J}(A_n)$  converges toward  $\mathcal{J}(\mathcal{W}_\tau)$ .  $\square$

The proof of the local upper bound (5) relies on the following covering lemma.

LEMMA 5.2. *Let  $E$  be a Borel subset of  $\Gamma$  having finite perimeter. For  $\varepsilon, \delta > 0$  there exists a finite collection of disjoint balls  $B(x_i, r_i), i \in I$ , such that for any  $i \in I$ :*

- (i) *Either  $x_i \in \partial^* E \cap \text{int } \Gamma$  and  $B(x_i, r_i) \subseteq \text{int } \Gamma$ ,*
- (ii) *or  $x_i \in \partial^* E \cap \partial \Gamma$ ,  $B_-(x_i, r_i, \nu_E(x_i)) \subseteq \Gamma$ .*

*In both cases  $\text{dist}_{L^1}(E \cap B(x_i, r_i), B_-(x_i, r_i, \nu_E(x_i))) \leq \delta r_i^d$ .*

*Moreover,*

$$\left| \mathcal{J}(E) - \sum_{i \in I} \alpha_{d-1} r_i^{d-1} \tau(\nu_E(x_i)) \right| \leq \varepsilon.$$

PROOF. Let  $\varepsilon, \delta$  be positive with  $\varepsilon < 1/2$ . By definition of the measure theoretic normal, for any  $x \in \partial^* E$  there exists a positive  $r_1(x, \delta)$  such that for any  $r < r_1(x, \delta)$ ,

$$\text{dist}_{L^1}(E \cap B(x, r), B_-(x, r, \nu_E(x))) \leq \delta r^d.$$

Since  $E \subset \Gamma$ , then  $\partial^* E \subset \text{int } \Gamma \cup \partial^* \Gamma$ . If  $x$  belongs to  $\partial^* E \cap \partial^* \Gamma$ , then  $\nu_E(x) = \nu_\Gamma(x)$ . The map  $x \in \partial^* E \mapsto \nu_E(x)$  is  $\|\nabla \chi_E\|$  measurable and the map

$\tau: S^{d-1} \rightarrow \mathbb{R}^+$  is continuous and bounded, thus the map  $x \in \partial^* E \mapsto \tau(\nu_E(x))$  is  $\|\nabla \chi_E\|$  measurable and bounded. By the Besicovitch differentiation theorem (see Section A.1), for  $\mathcal{H}^{d-1}$  almost every  $x$  in  $\partial^* E$ ,

$$(47) \quad \lim_{r \rightarrow 0} (\alpha_{d-1} r^{d-1})^{-1} \mathcal{H}^{d-1}(\partial^* E \cap B(x, r)) = 1,$$

$$(48) \quad \lim_{r \rightarrow 0} (\alpha_{d-1} r^{d-1})^{-1} \int_{\partial^* E \cap B(x, r)} \tau(\nu_E(y)) \mathcal{H}^{d-1}(dy) = \tau(\nu_E(x)).$$

Let  $\partial^{**} E$  be the set of the points of  $\partial^* E$  where (47) and (48) hold simultaneously. Clearly  $\mathcal{H}^{d-1}(\partial^* E \setminus \partial^{**} E) = 0$ . For any  $x \in \partial^{**} E$ , there exists a positive  $r_2(x, \varepsilon)$  such that for any  $r < r_2(x, \varepsilon)$ ,

$$\begin{aligned} \left| \mathcal{H}^{d-1}(\partial^* E \cap B(x, r)) - \alpha_{d-1} r^{d-1} \right| &\leq \varepsilon \alpha_{d-1} r^{d-1}, \\ \left| \int_{\partial^* E \cap B(x, r)} \tau(\nu_E(y)) \mathcal{H}^{d-1}(dy) - \alpha_{d-1} r^{d-1} \tau(\nu_E(x)) \right| &\leq \varepsilon \alpha_{d-1} r^{d-1}. \end{aligned}$$

The family of the balls  $B(x, r)$ ,  $x \in \partial^{**} E$  and  $0 < r < \min(r_1(x, \delta), r_2(x, \varepsilon), d_\infty(x, \partial\Gamma \setminus \partial^* \Gamma))$  is a Vitali class for  $\partial^{**} E$  (where  $\partial\Gamma \setminus \partial^* \Gamma$  is the union of the boundaries of the faces of  $\Gamma$ ). By the Vitali covering theorem for  $\mathcal{H}^{d-1}$ , we can select a countable collection of disjoint balls  $B(x_i, r_i)$ ,  $i \in I$ , belonging to this family such that either  $\mathcal{H}^{d-1}(\partial^{**} E \setminus \cup_{i \in I} B(x_i, r_i)) = 0$  or  $\sum_{i \in I} r_i^{d-1} = \infty$ . By our assumption,  $\mathcal{P}(E) = \mathcal{H}^{d-1}(\partial^* E)$  is finite. For each  $i \in I$ ,  $r_i$  is smaller than  $r_2(x_i, \varepsilon)$ , thus

$$(1 - \varepsilon) \sum_{i \in I} \alpha_{d-1} r_i^{d-1} \leq \mathcal{H}^{d-1}(\partial^{**} E) < \infty.$$

Therefore, the first case must occur and we may select a finite subset  $J \subseteq I$  such that

$$\mathcal{H}^{d-1} \left( \partial^{**} E \setminus \bigcup_{i \in J} B(x_i, r_i) \right) < \varepsilon \mathcal{H}^{d-1}(\partial^{**} E).$$

We claim that the collection of balls indexed by  $J$  enjoys the desired properties. Indeed, there is only the last condition to be checked:

$$\begin{aligned} &\left| \mathcal{S}(E) - \sum_{i \in I} \alpha_{d-1} r_i^{d-1} \tau(\nu_E(x_i)) \right| \\ &\leq \int_{\partial^{**} E \setminus \cup_{i \in J} B(x_i, r_i)} \tau(\nu_E(y)) \mathcal{H}^{d-1}(dy) \\ &\quad + \sum_{i \in J} \left| \int_{\partial^{**} E \cap B(x_i, r_i)} \tau(\nu_E(y)) \mathcal{H}^{d-1}(dy) - \alpha_{d-1} r_i^{d-1} \tau(\nu_E(x_i)) \right| \\ &\leq \varepsilon \mathcal{H}^{d-1}(\partial^* E) (\tau_{\max} + 2) \\ &\leq \varepsilon \mathcal{S}(E) (\tau_{\max} + 2) / \tau_{\min}, \end{aligned}$$

where we have used inequality (46) in the last step.  $\square$

**6. The interface lemma.** This section is devoted to the proof of a lemma which is crucial for linking the surface tension to the desired large deviation upper bounds. The interface lemma gives a probabilistic estimate for the local presence of a collection of open clusters creating a small flat interface near a middle hyperplane of a ball. “Near” is understood with respect to the  $L^1$  topology. The estimate is uniform with respect to the location, the size and the direction of the interface. Moreover, this lemma suggests an alternative way for defining the surface tension, which is physically more natural.

*Notation.* Let  $B(x, r)$  be a ball in  $\Gamma = [-1/2, 1/2]^d$ , let  $w$  belong to  $S^{d-1}$ ,  $n$  to  $\mathbb{N}$  and let  $r, \delta$  be positive. Recall that for a collection  $\mathcal{C}$  of sets we denote by  $\text{cup } \mathcal{C}$  the union of all the sets belonging to the collection  $\mathcal{C}$ . If  $A \subseteq \mathbb{Z}^d$  we denote its “thickened” version, regarded as a continuous object of  $\mathbb{R}^d$ , by cube  $A = \bigcup_{x \in A} \Lambda(x, 1)$ . Recall that the open  $B(nx, nr)$ -clusters are the open clusters in the configuration restricted to the ball  $B(nx, nr)$ . Let  $\text{Sep}(n, x, r, w, \delta)$  be the event: there exists a collection  $\mathcal{C}$  of open  $B(nx, nr)$ -clusters such that

$$(49) \quad \text{dist}_{L^1}(\text{cube cup } \mathcal{C}, B_-(nx, nr, w)) \leq \delta r^d n^d.$$

We will work in the unit cube  $\Gamma$ , hence we have to examine the possibility of having an interface intersecting  $\partial\Gamma$ . Let  $B(x, r)$  be a ball such that  $x$  belongs to  $\partial^*\Gamma$  and  $B_-(x, r, \nu_\Gamma(x))$  is included in  $\Gamma$ . Let  $n \in \mathbb{N}$  and  $\delta > 0$ . Let  $\text{Sep}^{\text{bd}}(n, x, r, \delta)$  be the event: there exists a collection  $\mathcal{C}$  of open  $B_-(nx, nr, \nu_\Gamma(x))$ -clusters such that none of them intersects  $\partial^{\text{in}}\Lambda(n)$  and the inequality (49) is satisfied with  $w$  replaced by  $\nu_\Gamma(x)$ .

**LEMMA 6.1 (Interface lemma).** *Let  $d \geq 3$ ,  $q \geq 1$ ,  $p > \hat{p}_c$ ,  $\theta^f(p) = \theta^w(p)$ . There exists a constant  $c = c(p, q, d, \zeta)$  such that for every  $r$  in  $(0, 1)$ , every ball  $B(x, r) \subseteq \Gamma$ , every unit vector  $w$  in  $S^{d-1}$  and every  $\delta$  in  $(0, \theta/2)$ ,*

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \left( \max_{\Phi \in \mathcal{C}\mathcal{F}\mathcal{K}(p, q, \Lambda(n))} \Phi[\text{Sep}(n, x, r, w, \delta)] \right) \\ & \leq -\alpha_{d-1} r^{d-1} \tau(w) (1 - c\delta^{1/2}). \end{aligned}$$

**REMARK.** (i) Note that the event  $\text{Sep}(n, x, r, w, \delta)$  depends only on the status of the edges inside  $B(nx, nr)$  and it can be seen as an event guaranteeing the existence of a certain set of closed edges inside  $B(nx, nr)$ , hence it is decreasing.

(ii) For  $\eta$  sufficiently small (depending on  $\delta$ ), we have  $S(n \text{ disc}(x, r, w), n\eta) \subset \text{Sep}(n, x, r, w, \delta)$ . Proposition 3.1 and Lemma 4.7 together imply the corresponding lower bound

$$\liminf_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \left( \max_{\Phi \in \mathcal{C}\mathcal{F}\mathcal{K}(p, q, \Lambda(n))} \Phi[\text{Sep}(n, x, r, w, \delta)] \right) \geq -\alpha_{d-1} r^{d-1} \tau(w).$$

(iii) The event  $\text{Sep}^{\text{bd}}(n, x, r, \delta)$  is also decreasing. By the monotonicity of FK measures with respect to boundary conditions,

$$\Phi_{\Lambda(n)}^w[\text{Sep}^{\text{bd}}(n, x, r, \delta)] \leq \Phi_\infty[\text{Sep}^{\text{bd}}(n, x, r, \delta)] \leq \Phi_\infty[\text{Sep}(n, x, r, \nu_\Gamma(x), \delta)]$$

so that the interface lemma yields also an upper bound for  $\Phi_{\Lambda(n)}^w[\text{Sep}^{\text{bd}}(n, x, r, \delta)]$ . This upper bound turns out to be of the correct order. Indeed, the  $L^1$  constraint is weak and the interface might be created at a positive distance from the boundary: for  $\eta$  small enough (depending on  $\delta$ ),

$$S(n \text{ disc}(x - \eta\nu_\Gamma(x), r, \nu_\Gamma(x)), n\eta) \subset \text{Sep}^{\text{bd}}(n, x, r, \delta)$$

and we get the lower bound with the help of Proposition 3.1 and Lemma 4.7 as in Remark (ii) above. However, the situation could be radically different with another kind of boundary condition.

(iv) In the different context of the Ising model with Kac potentials, a procedure has been developed to get a probabilistic estimate for the local presence of an interface, whose spirit is similar to the technique of our proof (see the paragraph “minimal section” in [8]). In this context, a coarse-graining procedure specific to the Kac model is employed. The coarse-grained configuration is described by attributing a label  $-1, 0, +1$  to each mesoscopic box. The labels  $-1, +1$  correspond to boxes which have relaxed to the minus and plus phases, the label  $0$  to a box in an indeterminate status. The authors of [8] localize the interface between the minus and the plus phase and cover it with a collection of parallelepipeds (as we do with balls). Inside a parallelepiped, they map the configuration of labels on a configuration where there is no sequence of cubes of the same phase which crosses the parallelepiped, by modifying the labels in two strips, where the number of bad cubes is minimal (the bad cubes are the cubes having label  $0$  or cubes of the wrong phase). An essential difference is, in our view, that the aforementioned work stays at the mesoscopic level of the boxes and does not go in depth to handle the microscopic structure of the model, so that the estimates are not precise when the range of interactions is finite.

Lemma 6.1 and Remark (ii) suggest the following alternative definition of surface tension.

**COROLLARY 6.2** (Alternative definition of surface tension). *Let  $d \geq 3, q \geq 1, p > \hat{p}_c, \theta^f(p) = \theta^w(p)$ . Let  $\delta(n): \mathbb{N} \rightarrow (0, 1)$  be such that  $\lim_{n \rightarrow \infty} \delta(n) = 0, \lim_{n \rightarrow \infty} n\delta(n) = \infty$ . For  $x \in \mathbb{R}^d$  and  $w \in S^{d-1}$ , the limit*

$$\lim_{n \rightarrow \infty} -(\alpha_{d-1}n^{d-1})^{-1} \log \Phi_\infty[\text{Sep}(n, x, 1, w, \delta(n))]$$

*exists and depends only on  $w$ . It is equal to the surface tension  $\tau(w)$ .*

**REMARK.** The above limit does not depend on  $x$  because our model is translation invariant. Yet in a more general model, the surface tension might

depend simultaneously on the direction and a space variable; hence it might be more natural to define it as the above limit.

PROOF OF COROLLARY 6.2. Since the event  $\text{Sep}(n, x, 1, w, \delta)$  is nondecreasing with respect to  $\delta$ , then, for any  $\delta > 0$  and  $n$  large enough so that  $\delta(n) < \delta$ , we have  $\text{Sep}(n, x, 1, w, \delta(n)) \subset \text{Sep}(n, x, 1, w, \delta)$ , whence by Lemma 6.1,

$$\limsup_{n \rightarrow \infty} (\alpha_{d-1} n^{d-1})^{-1} \log \Phi_\infty[\text{Sep}(n, x, 1, w, \delta(n))] \leq -\tau(w).$$

Next, we have

$$S\left(n \text{ disc}(x, 1, w), n\delta(n)/(2\alpha_{d-1})\right) \subset \text{Sep}(n, x, 1, w, \delta(n)).$$

Since  $\lim_{n \rightarrow \infty} n\delta(n) = \infty$ , Lemma 4.7 yields

$$\liminf_{n \rightarrow \infty} (\alpha_{d-1} n^{d-1})^{-1} \log \Phi_\infty[\text{Sep}(n, x, 1, w, \delta(n))] \geq -\tau(w). \quad \square$$

An essential ingredient of the proof is Lemma 6.3 which might be of independent interest. This lemma relates the probabilities of events which can be transformed into each other by closing (or opening) certain configuration-dependent edges “by force”, a procedure which we refer to as perturbation or surgery. It can be used for estimating the probability of an event  $S$  about which we know that by “changing it a little bit” we get (a subset of) another event whose probability is known or can be controlled. Of course, the change (surgery) will cost some penalty; the corresponding factors can be seen in (50).

Let  $E$  be a finite set of edges in a graph and let  $S$  be an event defined on these edges. Let  $\psi: S \rightarrow \Omega_E$  be a map changing certain open edges into closed ones. Assume that the number of changed edges is uniformly bounded by  $N(\psi)$ . Note that the change in the number of clusters is nonnegative since we destroy connections, and it is bounded by  $N(\psi)$ . Let  $d(\psi) = \max_{\psi(S)} |\psi^{-1}|$  be the “degree of degeneracy” (noninjectiveness) of the map  $\psi$ .

LEMMA 6.3 (Monotone perturbation of an event in FK percolation). *Let  $q > 0$ ,  $p \in (0, 1)$  and  $\Phi \in \mathcal{FH}(p, q, E)$ . Then*

$$(50) \quad \Phi[S] \leq d(\psi) \left(1 \vee \frac{1}{q}\right)^{N(\psi)} \left(1 \vee \frac{p}{1-p}\right)^{N(\psi)} \Phi[\psi(S)].$$

*Let  $S_j$ ,  $j \in J$ , be a finite partition of  $S$  and for  $j \in J$ , let  $\psi_j = \psi|_{S_j}$  be the restriction of  $\psi$  to  $S_j$ . Assume that the number of changed edges in the mapping  $\psi_j$  is bounded by  $N(\psi_j)$ . Then*

$$(51) \quad \Phi[S] \leq |J| \left(\max_{j \in J} d(\psi_j)\right) \left(\left(1 \vee \frac{1}{q}\right) \left(1 \vee \frac{p}{1-p}\right)\right)^{\max_{j \in J} N(\psi_j)} \Phi[\psi(S)].$$

REMARK. Analogous estimates can be derived for opening edges (instead of closing) as well as for nonmonotonic perturbations.

PROOF OF LEMMA 6.3. For  $\omega \in \Omega_E$  and  $\Phi \in \mathcal{F}\mathcal{H}(p, q, E)$ , we denote by  $\text{cl}(\omega)$  the number of clusters in  $\omega$  counted with respect to the partially wired boundary conditions associated with  $\Phi$ . Then

$$\begin{aligned} \Phi[S] &= Z^{-1} \sum_{\omega \in S} q^{\text{cl}(\omega)} \left( \prod_{e \in E} p^{\omega(e)} (1-p)^{1-\omega(e)} \right) \\ &= Z^{-1} \sum_{\omega' \in \psi(S)} \sum_{\omega \in \psi^{-1}(\omega')} q^{\text{cl}(\omega')} \left( \prod_{e \in E} p^{\omega'(e)} (1-p)^{1-\omega'(e)} \right) \left( \frac{1}{q} \right)^{\text{cl}(\omega') - \text{cl}(\omega)} \\ &\quad \times \left( \prod_{e \in E} \left( \frac{p}{1-p} \right)^{\omega(e) - \omega'(e)} \right) \\ &\leq Z^{-1} \sum_{\omega' \in \psi(S)} \sum_{\omega \in \psi^{-1}(\omega')} q^{\text{cl}(\omega')} \left( \prod_{e \in E} p^{\omega'(e)} (1-p)^{1-\omega'(e)} \right) \left( 1 \vee \frac{1}{q} \right)^{\text{cl}(\omega') - \text{cl}(\omega)} \\ &\quad \times \prod_{e \in E} \left( 1 \vee \frac{p}{1-p} \right)^{\omega(e) - \omega'(e)} \\ &\leq \sum_{\omega' \in \psi(S)} \sum_{\omega \in \psi^{-1}(\omega')} \Phi[\omega'] \left( 1 \vee \frac{1}{q} \right)^N \left( 1 \vee \frac{p}{1-p} \right)^N \\ &= \left( 1 \vee \frac{1}{q} \right)^N \left( 1 \vee \frac{p}{1-p} \right)^N \sum_{\omega' \in \psi(S)} \Phi[\omega'] |\psi^{-1}(\omega')| \\ &\leq \left( \max_{\psi(S)} |\psi^{-1}| \right) \left( 1 \vee \frac{1}{q} \right)^N \left( 1 \vee \frac{p}{1-p} \right)^N \Phi[\psi(S)]. \end{aligned}$$

Inequality (51) is a direct consequence of (50).  $\square$

PROOF OF THE INTERFACE LEMMA. The basic idea of the proof is as follows. We will show that whenever the event ‘‘Sep’’ occurs, it is always possible to perform a surgery in the spirit of Lemma 6.3 by closing not too many edges so that in the modified configuration a separating set of closed edges appears near the middle hyperplane  $\text{hyp}(nx, w)$  of the ball  $B(nx, nr)$ . Recall that the occurrence of such a separating set of closed edges is directly related to our definition of surface tension. By controlling the penalty factors in (51) we will be able to give an estimate for  $\Phi[\text{Sep}]$  in terms of the surface tension  $\tau(w)$ .

Throughout the proof, we fix  $x, r, w, \delta$  as in the statement of the interface lemma and we note  $\text{Sep}(n) = \text{Sep}(n, x, r, w, \delta)$ . By Proposition 3.1, the proof can be carried out for the measure  $\Phi_{\Lambda(n)}^w$ . Let  $f(n)$  be a function from  $\mathbb{N}$  to  $\mathbb{N}$  such that  $n/f(n)^{d-1}$  goes to  $\infty$  as  $n$  goes to  $\infty$  and  $f(n) \geq \kappa \log n$  for all  $n$ , where  $\kappa$  is a sufficiently large constant, so that the estimate (14) holds. We work in the box  $\Lambda(n)$  rescaled by a factor  $f(n)$  with event-block size 1. For  $\underline{x}$  in  $\underline{\Lambda}(n)$ , let  $X(\underline{x})$  be the indicator function of  $R(B'(\underline{x}), f(n) - 1)^c \cup V(B(\underline{x}), \theta/2)^c$ . Suppose that the event  $\text{Sep}(n, x, r, w, \delta)$  occurs and let

$\mathcal{C}$  be a collection of open  $B(nx, nr)$ -clusters realizing it. Let  $\rho, \eta$  be such that  $0 < \eta < \rho < r, 0 < 2\eta < \sqrt{r^2 - \rho^2}$  ( $\rho$  will be chosen later to be close to  $r$ ). Set  $Z = \text{cyl}(n \text{ disc}(x, \rho, w))$ . We suppose that  $n$  is large enough so that every asymptotic inequality coming up in the proof is fulfilled (like  $4\eta n > \zeta > f(n)^{d-1}/n$ , where  $\zeta > 2d$  is an arbitrary but fixed constant which we used to define the surface tension). It will be useful to define the following subsets of  $B(nx, nr)$  (see Figure 4):

$$\begin{aligned} D &= D(n, x, w, \eta, \zeta, \rho) = Z \cap \text{slab}(nx, w, -n\eta - \zeta, n\eta + \zeta), \\ D^+ &= Z \cap \text{slab}(nx, w, 1, n\eta + \zeta), \\ D^- &= Z \cap \text{slab}(nx, w, -n\eta - \zeta, 0), \\ \partial^+ D &= Z \cap \text{slab}(nx + n\eta w, w, -\zeta, \zeta), \\ \partial^- D &= Z \cap \text{slab}(nx - n\eta w, w, -\zeta, \zeta), \\ \partial^- D^+ &= Z \cap \text{slab}(nx, w, 1, 1 + \zeta), \\ \partial^+ D^- &= Z \cap \text{slab}(nx, w, -\zeta, 0). \end{aligned}$$

Let  $C$  be an open  $D$ -cluster joining the sets  $\partial^+ D$  and  $\partial^- D$ . There are two possibilities: either  $C$  is contained in some cluster of  $\mathcal{C}$  or it is contained in its complement  $D \setminus \text{cup } \mathcal{C}$ . In the first case there exists an open  $D^+$ -cluster  $C'$  connecting  $\partial^- D^+$  and  $\partial^+ D$  such that  $C' \subset C$  and therefore,

$$C' \subset (\text{cube cup } \mathcal{C}) \setminus B_-(nx, nr, w).$$

In the second case there exists an open  $D^-$ -cluster  $C'$  connecting  $\partial^- D$  and  $\partial^+ D^-$  such that  $C' \subset C$  and therefore,

$$C' \subset B_-(nx, nr, w) \setminus \text{cube cup } \mathcal{C}.$$

Let  $F$  be the union of all such clusters  $C'$ . Then, on  $\text{Sep}(n)$ ,

$$(52) \quad |F| \leq \text{dist}_{L^1}(\text{cube cup } \mathcal{C}, B_-(nx, nr, w)) \leq \delta r^d n^d.$$

We define

$$\begin{aligned} \underline{F} &= \{ \underline{x} \in \underline{\Lambda}(n) \mid B(\underline{x}) \cap F \neq \emptyset, \\ &\quad B'(\underline{x}) \subset \text{slab}(nx, w, 1, \eta n + \zeta) \cup \text{slab}(nx, w, -\eta n - \zeta, 0) \}. \end{aligned}$$

Let  $\underline{x}$  be a regular block in  $\underline{F}$ , that is,  $X(\underline{x}) = 0$ , and such that  $B'(\underline{x}) \subset Z$ . This implies in particular that  $B'(\underline{x}) \subset D^- \cup D^+$ . Since  $F$  intersects  $B(\underline{x})$  and the diameter of any open cluster contained in  $F$  is larger than  $f(n)$ , the box  $B'(\underline{x})$  contains an open path of diameter larger than  $f(n) - 1$  included in  $F$ . The occurrence of the event  $R(B'(\underline{x}), f(n) - 1)$  implies that this open path is contained in the crossing cluster  $C^*(B'(\underline{x}))$ , which in turn implies that the latter cluster is contained in  $F$ . Similarly, the crossing cluster  $C^*(B(\underline{x}))$  associated with the event  $U(B(\underline{x}))$  has diameter  $f(n) - 1$  and it is thus contained in  $C^*(B'(\underline{x}))$ . Since the event  $V(B(\underline{x}), \theta/2)$  occurs as well, then

$$|F \cap B(\underline{x})| \geq |C^*(B(\underline{x}))| \geq \theta f(n)^d / 2.$$

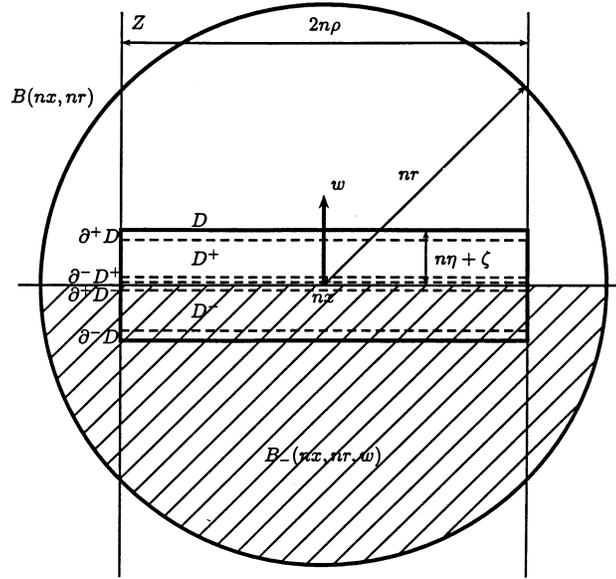


FIG. 4. The sets in  $B(nx, nr)$ .

Summing over  $\underline{x}$  yields

$$\begin{aligned} |F| &\geq \sum_{\underline{x} \in \underline{F}, X(\underline{x})=0, B'(\underline{x}) \subset Z} |F \cap B(\underline{x})| \\ &\geq (\theta/2)f(n)^d |\{\underline{x} \in \underline{F} \mid X(\underline{x}) = 0, B'(\underline{x}) \subset Z\}|. \end{aligned}$$

In addition, there exists a positive constant  $\gamma = \gamma(d)$  depending only upon the dimension such that

$$|\{\underline{x} \in \underline{F} \mid B'(\underline{x}) \cap \partial Z \neq \emptyset\}| < \gamma \eta n^{d-1} r^{d-2} / f(n).$$

Combining the two previous inequalities with (52), on  $\text{Sep}(n)$ , we can bound the number of the regular blocks in  $\underline{F}$ ,

$$|\{\underline{x} \in \underline{F} \mid X(\underline{x}) = 0\}| < \frac{2\delta}{\theta} \left( \frac{rn}{f(n)} \right)^d + \gamma \eta r^{d-2} \frac{n^{d-1}}{f(n)}.$$

On the other hand, by the estimate (14), the block process  $X(\underline{x})$  satisfies (15) with  $\varepsilon = b \exp(-cf(n))$ . Applying (18) with  $\delta = an^{d-1}(|\underline{\Lambda}(n)|f(n))^{-1}$ , where  $a > 0$ , we obtain

$$\begin{aligned} \Phi_{\Lambda(n)}^w \left[ \sum_{\underline{x} \in \underline{F}} X(\underline{x}) \geq a \frac{n^{d-1}}{f(n)} \right] &\leq \Phi_{\Lambda(n)}^w \left[ \frac{1}{|\underline{\Lambda}(n)|} \sum_{\underline{x} \in \underline{\Lambda}(n)} X(\underline{x}) \geq \frac{an^{d-1}}{|\underline{\Lambda}(n)|f(n)} \right] \\ &\leq 2^d \exp \left( -2^{-d} \Lambda_\varepsilon^* \left( \frac{an^{d-1}}{|\underline{\Lambda}(n)|f(n)} \right) \left[ \frac{n}{f(n)} \right]^d \right) \\ &\leq b \exp(-can^{d-1}), \end{aligned}$$

where  $b, c$  are two positive constants. By setting

$$\text{Sep}^*(n) = \text{Sep}(n) \cap \left\{ |\underline{F}| \leq (a + \gamma\eta r^{d-2}) \frac{n^{d-1}}{f(n)} + \frac{2\delta}{\theta} \left( \frac{rn}{f(n)} \right)^d \right\}$$

and using the previous estimates, we see that

$$(53) \quad \Phi_{\Lambda(n)}^w[\text{Sep}(n)] \leq b \exp(-can^{d-1}) + \Phi_{\Lambda(n)}^w[\text{Sep}^*(n)].$$

For  $i$  in  $\mathbb{Z}$ , we define the slablike sets

$$\begin{aligned} \underline{H}(i) &= \{ \underline{y} \in \underline{\Lambda}(n) \mid B(\underline{y}) \cap D \neq \emptyset, \\ &\quad \text{if}(n) \leq w \cdot (f(n)\underline{y} - nx) < (i + 1)f(n) \}. \end{aligned}$$

The sets  $\underline{H}(i)$ ,  $i \in \mathbb{Z}$ , are pairwise disjoint. Hence for any subset  $I$  of  $\mathbb{Z}$ , the sum  $\sum_{i \in I} |\underline{F} \cap \underline{H}(i)|$  is less than  $|\underline{F}|$  and there exists  $i$  in  $I$  such that  $|\underline{F} \cap \underline{H}(i)| \leq |\underline{F}|/|I|$ . Applying the preceding remark to the sets  $(-\eta n/f(n) + 3, -3) \cap \mathbb{Z}$  and  $(3, \eta n/f(n) - 3) \cap \mathbb{Z}$ , whose cardinalities are larger than  $\eta n/f(n) - 7 > \eta n/(2f(n))$ , we find that there exist two random indices  $I_-$  and  $I_+$  in  $\mathbb{Z}$  such that  $-\eta n/f(n) + 3 < I_- < -3$ ,  $3 < I_+ < \eta n/f(n) - 3$  and both  $\underline{F} \cap \underline{H}(I_-)$ ,  $\underline{F} \cap \underline{H}(I_+)$  have a cardinality less than  $2f(n)(\eta n)^{-1}|\underline{F}|$ . We choose the indices  $I_-$  and  $I_+$  with the smallest possible absolute value. We decompose the event  $\text{Sep}^*(n)$  according to the values of  $I_-, I_+$  and the sets  $\underline{F} \cap \underline{H}(I_-)$ ,  $\underline{F} \cap \underline{H}(I_+)$ . For

$$-\eta n/f(n) + 3 < i_- < -3, \quad 3 < i_+ < \eta n/f(n) - 3,$$

$$T_{\pm} \subseteq \underline{H}(i_{\pm}) \text{ with } |T_{\pm}| \leq \frac{2f(n)}{\eta n} |\underline{F}|,$$

we define the event

$$G(n, i_{\pm}, T_{\pm}) = \text{Sep}^*(n) \cap \{ I_{\pm} = i_{\pm}, \underline{F} \cap \underline{H}(I_{\pm}) = T_{\pm} \}.$$

Notice that the events  $G(n, i_{\pm}, T_{\pm})$  corresponding to different values of  $i_-, i_+, T_-, T_+$  are disjoint. Hence these events partition  $\text{Sep}^*(n)$ . Using Lemma 6.3, we now derive an estimate for  $\Phi_{\Lambda(n)}^w[\text{Sep}^*(n)]$ . Let  $\psi: \text{Sep}^*(n) \rightarrow \Omega_{\Lambda(n)}$  be the map defined on  $G(n, i_{\pm}, T_{\pm})$  by

$$\psi(\omega)[e] = \begin{cases} 0, & \text{if } e \in \bigcup_{\underline{x} \in T_+ \cup T_-} \partial^{\text{edge}} B(\underline{x}), \\ \omega(e), & \text{otherwise.} \end{cases}$$

Assume  $\omega \in G(n, i_{\pm}, T_{\pm})$ . By the definition of  $\underline{F}$ , each open path in  $D$  connecting  $\partial^+ D$  and  $\partial^- D$  intersects either  $\bigcup_{\underline{x} \in T_+} B(\underline{x})$  or  $\bigcup_{\underline{x} \in T_-} B(\underline{x})$  depending on whether it belongs to  $\mathcal{C}$  or not. Therefore, when we close all the edges in  $\bigcup_{\underline{x} \in T_+ \cup T_-} \partial^{\text{edge}} B(\underline{x})$ , we destroy every open connection in  $D$  between  $\partial^+ D$  and  $\partial^- D$ . Hence the map  $\psi$  has values in the set of configurations  $S(n \text{ disc}(x, \rho, w), n\eta)$ ; see Section 4 before Lemma 4.7 for its definition. To apply (51), we have to estimate the penalty factors. We first observe that there exists a dimension-dependent constant  $c' = c'(d)$  such that for all  $i \in \mathbb{Z}$ ,

$$|\underline{H}(i)| \leq c' \left( \frac{nr}{f(n)} \right)^{d-1}.$$

The first penalty factor in (51),  $|J|$ , is easily seen to be bounded by

$$\begin{aligned}
 |J| &\leq \left(\frac{n\eta}{f(n)}\right)^2 \exp\left(\left(\max_{i_+} |H(i_+)| + \max_{i_-} |H(i_-)|\right) \log 2\right) \\
 &\leq \left(\frac{\eta n}{f(n)}\right)^2 \exp\left(2c' \left(\frac{nr}{f(n)}\right)^{d-1}\right).
 \end{aligned}
 \tag{54}$$

To get a bound on the maximal possible number  $N$  of edges to close in the surgery, we observe that

$$\begin{aligned}
 &\left(\max_{x \in T_+ \cup T_-} |\partial^{\text{edge}} B(x)|\right) \times |T_+ \cup T_-| \\
 &\leq 2d^2 f(n)^{d-1} 2 \frac{2f(n)}{\eta n} \left( (a + \gamma \eta r^{d-2}) \frac{n^{d-1}}{f(n)} + \frac{2\delta}{\theta} \left(\frac{rn}{f(n)}\right)^d \right) \\
 &= \frac{8d^2}{\eta} \left( (a + \gamma \eta r^{d-2}) \frac{f(n)^{d-1}}{n} + \frac{2\delta r^d}{\theta} \right) n^{d-1} = N.
 \end{aligned}
 \tag{55}$$

The degree of degeneracy  $d(\psi_j)$  in our case is simply bounded by  $2^N$  [since on  $G(n, i_{\pm}, T_{\pm})$  the location of all the changes is known]. Note that all these estimates are uniform in  $i_+, i_-, T_+, T_-$ . Therefore, by (51),

$$\begin{aligned}
 \Phi_{\Lambda(n)}^w[\text{Sep}^*(n)] &\leq |J| \left(\frac{2p}{1-p}\right)^N \Phi_{\Lambda(n)}^w[\psi(\text{Sep}^*(n))] \\
 &\leq |J| \left(\frac{2p}{1-p}\right)^N \Phi_{\Lambda(n)}^w[S(n \text{ disc}(x, \rho, w), n\eta)].
 \end{aligned}$$

Coming back to (53), we get

$$\begin{aligned}
 \Phi_{\Lambda(n)}^w[\text{Sep}(n, x, r, w, \delta)] &\leq b \exp(-can^{d-1}) \\
 &\quad + |J| \left(\frac{2p}{1-p}\right)^N \Phi_{\Lambda(n)}^w[S(n \text{ disc}(x, \rho, w), n\eta)].
 \end{aligned}$$

Letting  $n$  go to  $\infty$  and using (54), (55) and Corollary 4.10, we get

$$\begin{aligned}
 &\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_{\Lambda(n)}^w[\text{Sep}(n, x, r, w, \delta)] \\
 &\leq -\min \left\{ ca, \alpha_{d-1} \rho^{d-1} \tau(w) - c\eta \rho^{d-2} + \frac{16d^2 r^d}{\theta \eta} \delta \log \frac{2p}{1-p} \right\},
 \end{aligned}$$

where  $c = c(p, d, \zeta)$ . We choose now  $\eta = \sqrt{\delta}r/3$ ,  $\rho = r\sqrt{1-\delta}$  and we let  $n$  go to  $\infty$  in the preceding inequality. Because  $\tau$  is bounded away from 0 (Proposition 4.5), there exists a constant  $c'' = c''(p, q, d, \zeta)$  such that

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_{\Lambda(n)}^w[\text{Sep}(n, x, r, w, \delta)] \leq -\alpha_{d-1} r^{d-1} \tau(w) (1 - c'' \delta^{1/2}).$$

This inequality holds for every ball  $B(x, r) \subseteq \Gamma$ ,  $\delta$  in  $(0, \theta/2)$  and  $w$  in  $S^{d-1}$ . □

**7. Proofs of the main theorems.**

7.1. *Asymptotics of the magnetization.* We consider an edge–spin configuration  $(\omega, \sigma)$  in  $\Lambda(n)$  drawn under the FK–Ising coupling measure  $\mathbb{P}_{\Lambda(n)}^+$ . We analyze the behavior of the magnetization of the Ising model in the box  $\Lambda(n)$  via the FK representation. The magnetization  $n^{-d} \sum_{x \in \Lambda(n)} \sigma(x)$  of the spin configuration  $\sigma$  is related to the open clusters of the FK configuration  $\omega$  through the formula

$$\frac{1}{n^d} \sum_{x \in \Lambda(n)} \sigma(x) = \frac{1}{n^d} \sum_{C \in \mathcal{C}} \sigma(C)|C|,$$

where  $\mathcal{C}$  is the collection of the open clusters in  $\Lambda(n)$  and  $\sigma(C)$  is the color associated with the cluster  $C$ . We first show that the relevant information is carried by the large (macroscopic) clusters of  $\mathcal{C}$ . Throughout the proof, we use a fixed function  $f: \mathbb{N} \rightarrow \mathbb{N}$  such that both  $n/f(n)^{d-1}$  and  $f(n)/\log n$  tend to  $\infty$  as  $n \rightarrow \infty$ . A cluster in  $\Lambda(n)$  is called *small* if its diameter is strictly smaller than  $f(n)$  and *large* otherwise. The next lemma controls the coloring of the small clusters.

LEMMA 7.1. *For any  $\delta > 0$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ \left[ \left| \frac{1}{n^d} \sum_{\substack{C \in \mathcal{C} \\ C \text{ small}}} \sigma(C)|C| \right| \geq \delta \right] = -\infty.$$

PROOF. For later use, we will prove the statement with respect to a measure  $\mathbb{P}_{\Lambda(n)}$  describing a coupling between an arbitrary FK measure  $\Phi \in c\mathcal{FK}(\Lambda(n))$  and a coloring measure with the property that the colors of the clusters not touching the boundary are i.i.d.  $\pm 1$  with probability 1/2 each. (Clearly,  $\mathbb{P}_{\Lambda(n)}^+$  is such a measure). An analogous statement has already been proved in the proof of Theorem 1.1 [59], hence we will merely sketch the argument. Set

$$\mathcal{C}' = \{C \in \mathcal{C} \mid C \text{ is small and } C \cap \Lambda(n - 2f(n)) \neq \emptyset\}.$$

Let  $n$  be so big that  $2dn^{d-1}2f(n) < \delta n^d$ . Then

$$\left\{ \left| \sum_{\substack{C \in \mathcal{C} \\ C \text{ small}}} \sigma(C)|C| \right| \geq 2\delta n^d \right\} \subseteq \left\{ \left| \sum_{C \in \mathcal{C}'} \sigma(C)|C| \right| \geq \delta n^d \right\}.$$

There exists  $\alpha = \alpha(p, d) > 0$  and  $c_i = c_i(p, d) > 0$ ,  $i = 1, 2$ , such that

$$\mathbb{P}_{\Lambda(n)}[|\mathcal{C}'| < \alpha n^d] \leq c_1 \exp(-c_2 n^d).$$

Note that already the number of the point clusters  $C \in \mathcal{C}'$ ,  $|C| = 1$ , is proportional to  $n^d$ , up to volume order large deviations: to prove this, one can use, for instance, the fact that the FK measure is stochastically dominated by the Bernoulli percolation. We next condition on the edge configuration  $\omega$  observed

on the FK level and assume that  $\omega \in \{|\mathcal{C}'| \geq \alpha n^d\}$ . The coloring variables of the clusters of  $\mathcal{C}'$  are i.i.d.  $\pm 1$  valued variables with zero mean (since the clusters of  $\mathcal{C}'$  cannot touch the boundary of  $\Lambda(n)$ ). Let  $X$  denote a  $\pm 1$  valued variable with zero mean. By Lemma 5.3 in [59],

$$\begin{aligned} & \mathbb{P}_{\Lambda(n)} \left[ \left| \sum_{C \in \mathcal{C}'} \sigma(C) |C| \right| \geq \delta n^d \mid \omega \right] \\ & \leq \mathbb{P}_{\Lambda(n)} \left[ \left| \frac{1}{|\mathcal{C}'|} \sum_{C \in \mathcal{C}'} \sigma(C) |C| \right| \geq \delta/\alpha \mid \omega \right] \\ & \leq 2 \exp(-|\mathcal{C}'| \Lambda_{f(n)X}^*(\delta/\alpha)) \leq 2 \exp\left(-\frac{c_3 \delta^2 n^d}{\alpha f(n)^2}\right), \end{aligned}$$

where we have used the fact that the Legendre-transform  $\Lambda_X^*(x)$  of the log-moment generating function of  $X$  satisfies  $\Lambda_X^*(x) \geq c_3 x^2$  for an appropriate  $c_3 > 0$ . Since the previous estimate is uniform on the set  $\{|\mathcal{C}'| \geq \alpha n^d\}$ , the claim follows. Let us finally remark that the estimate is not optimal; the correct order of decay is  $\sim \exp(-cn^d)$  and it could be proved with slightly more effort.  $\square$

We next treat the large clusters. For  $* = -$  or  $+$ , let  $\mathcal{C}^*$  be the collection of the large clusters colored with  $*$  and  $\mathcal{M}^* = \text{cup } \mathcal{C}^*$ . We divide the continuous box  $n\Gamma$  into three disjoint random sets,

$$\begin{aligned} M^- &= \{x \in n\Gamma \mid d_\infty(x, \mathcal{M}^-) \leq f(n), d_\infty(x, \mathcal{M}^+) > f(n)\}, \\ M^+ &= \{x \in n\Gamma \mid d_\infty(x, \mathcal{M}^-) > f(n), d_\infty(x, \mathcal{M}^+) \leq f(n)\} \end{aligned}$$

and  $M^0 = n\Gamma \setminus (M^- \cup M^+)$ . Note that the region  $M^-$  consists of the points of  $n\Gamma$  whose  $f(n)$  neighborhood intersects only large clusters with negative color, and at least one such cluster.

LEMMA 7.2. For any  $\delta > 0$ ,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ \left[ \left| |\mathcal{M}^-| - \theta \text{vol } M^- \right| + \text{vol } M^0 \right. \\ \left. + \left| |\mathcal{M}^+| - \theta \text{vol } M^+ \right| > \delta n^d \right] = -\infty. \end{aligned}$$

PROOF. We work with the box  $\Lambda(n)$  rescaled by a factor  $f(n)$  with event-block size 2. For  $\underline{x} \in \underline{\Lambda}(n)$ , the block variable  $X(\underline{x})$  is the indicator function of the event  $T(B'(\underline{x}), f(n) - 1, \delta/2)^c$ . Suppose that  $X(\underline{x}) = 0$  for some  $\underline{x} \in \underline{\Lambda}(n)$ ; that is, the block is regular. Then there exists a unique large cluster  $C$  intersecting the block  $B(\underline{x})$  and the block  $B(\underline{x})$  is included in  $M^{\sigma(C)}$ , while  $\mathcal{M}^{\sigma(C)} \cap B(\underline{x}) = C \cap B(\underline{x})$  and  $\mathcal{M}^{-\sigma(C)} \cap B(\underline{x}) = \emptyset$ . Moreover, for  $n$  large enough,

$$\left| |\mathcal{M}^{\sigma(C)} \cap B(\underline{x})| - \theta |B(\underline{x})| \right| = \left| |C \cap B(\underline{x})| - \theta |B(\underline{x})| \right| \leq (\delta/2) f(n)^d.$$

Therefore,

$$||\mathcal{M}^-| - \theta \text{vol}M^-| + \text{vol}M^0 + ||\mathcal{M}^+| - \theta \text{vol}M^+| \leq \delta n^d/2 + f(n)^d \sum_{\underline{x} \in \Lambda(n)} X(\underline{x}).$$

By the estimate (14), the block process  $X(\underline{x})$  satisfies (15) with  $\varepsilon = b \exp(-cf(n))$ . Applying (18), we obtain

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ [f(n)^d \sum_{\underline{x} \in \Lambda(n)} X(\underline{x}) > \delta n^d/2] = -\infty$$

and the proof is complete.  $\square$

Lemmas 7.1, 7.2 yield the following result.

COROLLARY 7.3. *For  $\delta > 0$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ \left[ \left| \frac{1}{n^d} \sum_{x \in \Lambda(n)} \sigma(x) - \theta(1 - 2 \text{vol}(M^-/n)) \right| > \delta \right] = -\infty.$$

Thus the magnetization  $n^{-d} \sum_{x \in \Lambda(n)} \sigma(x)$  and  $\theta(1 - 2 \text{vol}(M^-/n))$  are exponentially contiguous which implies (1). To derive the large deviation statement for the magnetization we need only to consider the random variable  $\text{vol}(M^-/n)$ . We prove successively the lower bound (2) and the upper bound (5).

PROOF OF THE LOWER BOUND. We first prove (2).

PROPOSITION 7.4. *Let  $\lambda > 0$  be such that  $\lambda \text{diam } \mathcal{W}_\tau < 1$  and let  $x$  be such that  $x + \lambda \mathcal{W}_\tau$  is included in  $\text{int } \Gamma$ . For any  $\delta > 0$ ,*

$$\liminf_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ [\text{dist}_{L^1}(M^-/n, x + \lambda \mathcal{W}_\tau) < \delta] \geq -\lambda^{(d-1)} \mathcal{I}(\mathcal{W}_\tau).$$

PROOF. By Lemma 5.1, for any  $\delta, \delta' > 0$ , there exists a polyhedral set  $E$  such that

$$\begin{aligned} \text{dist}_{L^1}(E, x + \lambda \mathcal{W}_\tau) &< \delta/3, & x + \lambda \mathcal{W}_\tau &\subset E \subset \mathcal{V}_2(x + \lambda \mathcal{W}_\tau, \delta), \\ \mathcal{I}(E) &< \lambda^{(d-1)} \mathcal{I}(\mathcal{W}_\tau) + \delta'. \end{aligned}$$

In particular, since  $x + \lambda \mathcal{W}_\tau$  is included in  $\text{int } \Gamma$ , then for  $\delta$  sufficiently small,  $E$  is also included in the interior of  $\Gamma$ . Let  $F_i, i \in I$ , be the relative interiors (w.r.t. the  $(d-1)$ -dimensional topology) of the faces of  $E$ . These are polyhedral open hypersets of  $\mathbb{R}^d$ . Let  $U$  be an open connected subset of  $E$  such that  $\text{dist}_{L^1}(U, E) < \delta/3$  and  $\partial U \cap \partial E = \emptyset$ . Let  $V$  be an open connected subset of  $\Gamma \setminus E$  such that  $\text{dist}_{L^1}(V, \Gamma \setminus E) < \delta/3$  and  $\partial V \cap \partial(\Gamma \setminus E) = \emptyset$ . Pick  $\varepsilon > 0$  such that  $4\varepsilon < d_\infty(U, \partial E) \wedge d_\infty(V, \partial(\Gamma \setminus E))$ . We cover the compact set  $\text{clo}(U \cup V)$  by a finite collection of boxes  $\Lambda(x_j, \varepsilon), j \in J$ , included in  $\text{int } \Gamma$  and centered

in  $U \cup V$ . Let  $n$  be large enough so that  $f(n) < \varepsilon n$ . Let  $\zeta > 2d$  be fixed and consider the following events:

$$W(n) = \bigcap_{i \in I} S(nF_i, f(n)) \cap \{ \text{all the edges in } \mathcal{V}_2(\text{cyl } n\partial F_i, \zeta) \\ \cap \mathcal{V}_2(\text{hyp } nF_i, f(n)) \text{ are closed} \},$$

$$R(n, \varepsilon) = \bigcap_{j \in J} O(n\Lambda(x_j, 2\varepsilon), f(n)).$$

In words, we have surfaces of closed edges near the faces of  $E$ , and typical configurations inside the regions  $U$  and  $V$ . Whenever  $W(n)$  and  $R(n, \varepsilon)$  occur, there exist two distinct large clusters  $C_1$  and  $C_2$  such that

$$nU \subseteq \mathcal{V}_\infty(C_1, f(n)) \subseteq \mathcal{V}_\infty(nE, f(n)),$$

$$nV \subseteq \mathcal{V}_\infty(C_2, f(n)) \subseteq \mathcal{V}_\infty(n(\Gamma \setminus E), f(n))$$

and no other large cluster intersects the region  $n\mathcal{V}_\infty(U, \varepsilon) \cup n\mathcal{V}_\infty(V, \varepsilon)$ . If  $C_1$  is colored with  $-$  and  $C_2$  with  $+$  then the region  $M^-$  will satisfy  $nU \subseteq M^- \subseteq n(\Gamma \setminus V)$ . Therefore,

$$(1/4)\mathbb{P}_{\Lambda(n)}^+[W(n) \cap R(n, \varepsilon)] \\ = \mathbb{P}_{\Lambda(n)}^+[W(n) \cap R(n, \varepsilon) \cap \{ \sigma(C_1) = -, \sigma(C_2) = + \}] \\ \leq \mathbb{P}_{\Lambda(n)}^+[\text{dist}_{L^1}(M^-/n, x + \lambda\mathcal{W}_\tau) < \delta].$$

The event  $R(n, \varepsilon)$  depends only on the edges inside  $n\mathcal{V}_\infty(U \cup V, 2\varepsilon)$ , while the event  $W(n)$  depends on the edges in  $n\mathcal{V}_\infty(\partial E, \varepsilon)$ . By the estimate (14), for  $j$  in  $J$ , the probability of each event  $O(n\Lambda(x_j, 2\varepsilon), f(n))$  goes to 1, uniformly over the b.c.s on  $n\Lambda(x_j, 2\varepsilon)$ . Therefore,  $\lim_{n \rightarrow \infty} \mathbb{P}_{\Lambda(n)}^+[R(n, \varepsilon) | W(n)] = 1$ . Moreover, by the FKG inequality and Lemma 4.7,

$$\liminf_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+[W(n)] \geq - \sum_{i \in I} \mathcal{H}^{d-1}(F_i) \tau(\text{nor } F_i) \\ = -\mathcal{J}(E) \geq -\lambda^{(d-1)} \mathcal{J}(\mathcal{W}_\tau) - \delta'.$$

Thus, for any  $\delta, \delta' > 0$  sufficiently small,

$$\liminf_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+[\text{dist}_{L^1}(M^-/n, x + \lambda\mathcal{W}_\tau) < \delta] \geq -\lambda^{(d-1)} \mathcal{J}(\mathcal{W}_\tau) - \delta'.$$

The result of the proposition follows by letting  $\delta'$  go to 0.  $\square$

We now complete the proof of the lower bound for Theorem 1.1. Let  $m$  as in Theorem 1.1 and let  $\delta > 0$  such that  $1 - 2(\text{diam } \mathcal{W}_\tau)^{-d} \text{vol } \mathcal{W}_\tau < (m - 2\delta)/m^*$ .

By applying successively Corollary 7.3 and Proposition 7.4, we get

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ \left[ \frac{1}{n^d} \sum_{x \in \Lambda(n)} \sigma(x) \leq m \right] \\ & \geq \liminf_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ \left[ \text{vol}(M^-/n) \geq \frac{m^* + \delta - m}{2m^*} \right] \\ & \geq \liminf_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ \left[ \text{dist}_{L^1} \left( M^-/n, \left( \frac{m^* + 2\delta - m}{2m^* \text{vol } \mathcal{W}_\tau} \right)^{1/d} \mathcal{W}_\tau \right) < \frac{\delta}{m^*} \right] \\ & \geq - \left( \frac{m^* + 2\delta - m}{2m^* \text{vol } \mathcal{W}_\tau} \right)^{(d-1)/d} \mathcal{I}(\mathcal{W}_\tau). \end{aligned}$$

Letting  $\delta$  go to 0, we obtain the correct lower bound for Theorem 1.1. This lower bound is certainly not correct for  $m$  such that  $1 - 2(\text{diam } \mathcal{W}_\tau)^{-d} \text{vol } \mathcal{W}_\tau > m/m^*$  (see [63] for the corresponding question in  $d = 2$ ).

**PROOF OF THE UPPER BOUND.** We build two auxiliary sets *fifa*  $M^-$  and *aglu*  $M^-$ . Both *fifa*  $M^-$  and *aglu*  $M^-$  are exponentially contiguous to  $M^-$ . The set *fifa*  $M^-$  helps to prove the exponential tightness (4), while *aglu*  $M^-$  is used to prove the local estimate (5).

*The set fifa  $M^-$ .* This set is a coarser version of  $M^-$  obtained through the succession of a filling and a fattening operation, hence the name “fifa.” We work with the box  $\Lambda(n)$  rescaled by a factor  $k$  with event-block size 1 ( $k$  is a fixed integer strictly larger than  $2d$  and large enough for some of the subsequent estimates to hold). The block variable  $X(\underline{x})$  is the indicator function of the event  $O(B'(\underline{x}), k - 1)^c$ . We introduce a coarse-grained image of the set  $\mathcal{M}^-$  as follows:

$$\underline{\mathcal{M}}^- = \{ \underline{x} \in \underline{\Lambda}(n) \mid B(\underline{x}) \cap \mathcal{M}^- \neq \emptyset \}.$$

This set, in general, contains lots of large and small holes. To get rid of the small ones we first fill out these holes by the operation “fill” which we now describe. We look at the residual components of  $\underline{\mathcal{M}}^-$ , more precisely the  $\mathbb{L}^{d, \infty}$ -connected components of  $\mathbb{Z}^d \setminus \underline{\mathcal{M}}^-$ . We define

$$\begin{aligned} \text{fill } \underline{\mathcal{M}}^- &= \underline{\mathcal{M}}^- \cup \{ \underline{R} \mid \underline{R} \text{ is a finite residual component of } \underline{\mathcal{M}}^-, \\ & \quad |\partial^{\text{edge}} \underline{R}| \leq f(n)/k \}. \end{aligned}$$

Notice that any  $\mathbb{L}^{d, \infty}$ -connected component of  $\partial^{\text{in}} \text{fill } \underline{\mathcal{M}}^-$  has cardinality strictly larger than  $f(n)/(2dk)$ . This is clearly true, by construction, for the parts of  $\partial^{\text{in}} \text{fill } \underline{\mathcal{M}}^-$  associated with finite residual components (holes). Let  $\underline{A}$  be an  $\mathbb{L}^d$ -connected component of  $\underline{\mathcal{M}}^-$ . Since we deal with large clusters,  $\text{diam } \underline{A} \geq f(n)/k - 1$  and  $|\partial^{\text{ext, edge}} \underline{A}| \geq \text{diam } \underline{A} + 2 > f(n)/k$ . This in turn implies that  $|\partial^{\text{ext, in}} \underline{A}| > f(n)/(2dk)$ . Note finally that every  $\mathbb{L}^{d, \infty}$ -connected piece of  $\partial^{\text{in}} \text{fill } \underline{\mathcal{M}}^-$  consists of such boundary pieces, which implies the claim.

We next claim that if  $x$  belongs to  $\partial^{\text{in}}\text{fill } \mathcal{M}^- \subseteq \partial^{\text{in}}\mathcal{M}^-$  then  $X(x) = 1$ . To see this we first observe that  $x \in \partial^{\text{in}}\underline{\Lambda}(n) \cap \partial^{\text{in}}\mathcal{M}^-$  implies that the event  $O(B'(x), k - 1)$  does not happen, otherwise the crossing cluster of  $B'(x)$  would be colored “+” and the block of  $x$  would not intersect  $\mathcal{M}^-$ . If, on the other hand,  $x \in \partial^{\text{in}}\mathcal{M}^- \setminus \partial^{\text{in}}\underline{\Lambda}(n)$  then it has a nearest neighbor  $y$  in  $\underline{\Lambda}(n)$  such that  $B(y) \cap \mathcal{M}^- = \emptyset$ , hence the event  $O(B'(x), k - 1)$  does not happen either.

Finally, we define the fattened region

$$\text{fifa } M^- = \bigcup_{x \in \text{fill } \mathcal{M}^-} B(x).$$

LEMMA 7.5. *For  $k$  large enough, depending on  $p$  and  $d$ , there exist  $b, c > 0$  such that*

$$\forall s > 0 \quad \mathbb{P}_{\Lambda(n)}^+[\mathcal{P}(\text{fifa } M^-) > s] \leq b \exp(-cs).$$

PROOF. We have the following bound on the perimeter of  $\text{fifa } M^-$ :

$$\mathcal{P}(\text{fifa } M^-) \leq 2dk^{d-1}|\partial^{\text{in}}\text{fill } \mathcal{M}^-| \leq 2dk^{d-1}|\{x \in \underline{\Lambda}(n) \mid |C(x)| \geq f(n)/2dk\}|.$$

By the estimate (14), the block process  $X(x)$  satisfies (15) with  $\varepsilon = b \exp(-ck)$ . The claim of the lemma follows from Lemma 2.3 applied with  $t = f(n)/(2dk)$ .  $\square$

LEMMA 7.6. *For  $n$  large enough,*

$$\text{dist}_{L^1}(M^-, \text{fifa } M^-) \leq \text{vol}M^0 + 5^d f(n)^{d-1} \mathcal{P}(\text{fifa } M^-).$$

PROOF. We estimate the distance between  $M^-$  and  $\text{fifa } M^-$ . In the filling operation performed to build  $\text{fifa } M^-$ , all the holes which are filled on the rescaled lattice have a diameter not exceeding  $f(n)/((2d - 2)k)$ . Therefore, for  $n$  large enough, so that  $10k < f(n)$ ,

$$(56) \quad \text{fifa } M^- \subseteq \{x \in n\Gamma \mid d_\infty(x, \mathcal{M}^-) \leq f(n)\},$$

whence  $\text{fifa } M^- \setminus M^- \subseteq M^0$ . Moreover (notice that  $\mathcal{M}^- \subset \text{fifa } M^-$ ),

$$(57) \quad M^- \setminus \text{fifa } M^- \subseteq \{x \in n\Gamma \mid d_\infty(x, \partial \text{fifa } M^-) \leq f(n)\}.$$

Notice that either  $\mathcal{M}^- = \emptyset$  or  $\text{diam } \text{fifa } M^- \geq f(n)$ . If  $\mathcal{M}^- = \emptyset$ , then  $M^- = \text{fifa } M^- = \emptyset$ . Suppose that  $\text{diam } \text{fifa } M^- \geq f(n)$ . Let  $A_1, \dots, A_l$  be the  $\mathbb{L}^{d, \infty}$  components of  $\partial^{\text{in}} \text{fifa } M^-$ . Each of these components has cardinality larger than  $f(n)k^{d-2}/(2d) > f(n)$  (recall that  $k > 2d$ ). By Lemma 2.1,

$$\begin{aligned} \text{vol}\{x \in n\Gamma \mid d_\infty(x, \partial \text{fifa } M^-) \leq f(n)\} &\leq \sum_{1 \leq i \leq l} \text{vol } \mathcal{V}_\infty(A_i, f(n) + 1) \\ &\leq \sum_{1 \leq i \leq l} 4^{d+1}(f(n) + 1)^{d-1}|A_i| \\ &\leq 5^d f(n)^{d-1} \mathcal{P}(\text{fifa } M^-), \end{aligned}$$

where the last inequality holds for  $n$  large enough.  $\square$

We now prove the exponential tightness (4).

LEMMA 7.7. *There exists a positive constant  $c = c(p, d)$  such that for any  $\lambda, \delta > 0$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ [\text{dist}_{L^1}(M^-/n, \mathcal{J}^{-1}([0, \lambda])) \geq \delta] \leq -c\lambda.$$

PROOF. Let  $\lambda, \delta > 0$  be given. Using Lemma 7.6 and inequality (46),

$$\begin{aligned} & \mathbb{P}_{\Lambda(n)}^+ [\text{dist}_{L^1}(M^-/n, \mathcal{J}^{-1}([0, \lambda])) \geq \delta] \\ & \leq \mathbb{P}_{\Lambda(n)}^+ [\text{dist}_{L^1}(n^{-1} \text{fifa } M^-, n^{-1} M^-) \geq \delta/2] \\ & \quad + \mathbb{P}_{\Lambda(n)}^+ [\text{dist}_{L^1}(n^{-1} \text{fifa } M^-, \mathcal{J}^{-1}([0, \lambda])) \geq \delta/2] \\ & \leq \mathbb{P}_{\Lambda(n)}^+ [\text{vol } M^0 + 5^d f(n)^{d-1} \mathcal{P}(\text{fifa } M^-) \geq n^d \delta/2] \\ & \quad + \mathbb{P}_{\Lambda(n)}^+ [\mathcal{J}(\text{fifa } M^-) \geq \lambda n^{d-1}] \\ & \leq \mathbb{P}_{\Lambda(n)}^+ \left[ \text{vol } M^0 \geq \frac{\delta n^d}{4} \right] + \mathbb{P}_{\Lambda(n)}^+ \left[ \mathcal{P}(\text{fifa } M^-) \geq \frac{\delta n^d}{4 \cdot 5^d f(n)^{d-1}} \right] \\ & \quad + \mathbb{P}_{\Lambda(n)}^+ \left[ \mathcal{P}(\text{fifa } M^-) \geq \frac{\lambda n^{d-1}}{\tau_{\max}} \right]. \end{aligned}$$

By setting  $c = 1/\tau_{\max}$ , Lemmas 7.2 and 7.5 imply the desired conclusion.  $\square$

*The set aglu  $M^-$ .* We begin with the definition of a certain enlargement  $A$  of  $\mathcal{C}^-$ . The set  $A$  consists of the sites belonging either to  $\mathcal{M}^-$  or to small clusters not touching the boundary of  $\Lambda(n)$  and whose distance to  $\mathcal{M}^-$  is bounded by  $3f(n)$ ; that is,

$$A = \mathcal{M}^- \cup \text{cup}\{C \text{ is a small cluster} \mid d_\infty(C, \mathcal{M}^-) \leq 3f(n), C \cap \partial^{\text{in}} \Lambda(n) = \emptyset\}.$$

The set aglu  $M^-$  (the word ‘‘aglu’’ stands for agglutination) is defined as the union of all the unit continuous boxes  $\Lambda(x, 1) \subseteq \mathbb{R}^d$  centered at a vertex  $x$  in  $A$ ; that is, aglu  $M^- = \text{cube}A$ . We first show that  $M^-$  and aglu  $M^-$  are exponentially contiguous.

LEMMA 7.8. *For any  $\delta > 0$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ [\text{dist}_{L^1}(M^-, \text{aglu } M^-) > \delta n^d] = -\infty.$$

PROOF. We work with the box  $\Lambda(n)$  rescaled by a factor  $f(n)$  and with event-block size 4. The block variable  $X(\underline{x})$  is the indicator function of the event  $R(B'(\underline{x}), f(n) - 1)^c$ . We introduce an intermediate block  $B''(\underline{x})$  ‘‘around’’  $\underline{x}$  corresponding to the event-block size 2: hence, if  $d_\infty(\underline{x}, \partial^{\text{in}} \underline{\Lambda}(n)) \geq 3$ , then  $B''(\underline{x}) = \Lambda(f(n)\underline{x}, 5f(n))$ . Let  $\underline{x} \in \underline{\Lambda}(n)$  be such that  $d_\infty(\underline{x}, \partial^{\text{in}} \underline{\Lambda}(n)) \geq 3$  and suppose that  $B(\underline{x}) \cap \mathcal{M}^- \neq \emptyset$  and  $X(\underline{x}) = 0$ . Then there is a unique crossing

cluster in  $B'(\underline{x})$ , and all the other clusters intersecting the intermediate block  $B''(\underline{x})$  have a diameter less than  $f(n) - 1$ . Therefore  $B''(\underline{x}) \subseteq \text{aglu} M^-$ . Since

$$M^- \subseteq \{x \in n\Gamma \mid d_\infty(x, \mathcal{M}^-) \leq f(n)\} \subseteq \bigcup_{\underline{x}: B(\underline{x}) \cap \mathcal{M}^- \neq \emptyset} B''(\underline{x}),$$

we have

$$\begin{aligned} M^- \setminus \text{aglu} M^- &\subseteq \{x \in \Lambda(n) \mid d_\infty(x, \partial n\Gamma) \leq 6f(n)\} \\ &\cup \left( \bigcup_{\underline{x}: B(\underline{x}) \cap \mathcal{M}^- \neq \emptyset} B''(\underline{x}) \setminus \bigcup_{\substack{\underline{x}: B(\underline{x}) \cap \mathcal{M}^- \neq \emptyset \\ X(\underline{x})=0}} B''(\underline{x}) \right) \\ &\subseteq \{x \in \Lambda(n) \mid d_\infty(x, \partial n\Gamma) \leq 6f(n)\} \cup \bigcup_{\underline{x} \in \Lambda(n), X(\underline{x})=1} B''(\underline{x}) \end{aligned}$$

so that

$$\text{vol}(M^- \setminus \text{aglu} M^-) \leq 12df(n)n^{d-1} + 5^d f(n)^d \sum_{\underline{x} \in \Lambda(n)} X(\underline{x}).$$

On the other hand, using (56) and (57),

$$\begin{aligned} \text{aglu} M^- \setminus M^- &\subseteq (\text{aglu} M^- \setminus \text{fifa} M^-) \cup (\text{fifa} M^- \setminus M^-) \\ &\subseteq \{x \in \Lambda(n) \mid d_\infty(x, \partial \text{fifa} M^-) \leq 4f(n)\} \cup M^0. \end{aligned}$$

Using Lemma 2.1 and proceeding as in Lemma 7.6, we obtain

$$\begin{aligned} \text{dist}_{L^1}(M^-, \text{aglu} M^-) &\leq 12df(n)n^{d-1} + 5^d f(n)^d \sum_{\underline{x} \in \Lambda(n)} X(\underline{x}) \\ &\quad + \text{vol} M^0 + 5^d (4f(n))^{d-1} \mathcal{P}(\text{fifa} M^-). \end{aligned}$$

By the estimate (14), the block process  $X(\underline{x})$  satisfies (15) with  $\varepsilon = b \exp(-cf(n))$ . Applying (18) and Lemma 7.5, we get the exponential contiguity between  $M^-$  and  $\text{aglu} M^-$ .  $\square$

We now prove the upper bound (5) for a set having finite surface energy.

LEMMA 7.9. *Let  $E$  be a Borel set of  $\Gamma$  such that  $\mathcal{J}(E) < \infty$ . For every  $\varepsilon > 0$ , there exists  $\delta = \delta(E, \varepsilon) > 0$  such that*

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ [\text{dist}_{L^1}(M^-/n, E) < \delta] \leq -(1 - \varepsilon)\mathcal{J}(E).$$

PROOF. By the triangle inequality for  $\text{dist}_{L^1}$  and Lemma 7.8, we need only to prove the above statement for the set  $\text{aglu} M^-$ . Let  $E$  be a Borel set of  $\Gamma$  such that  $0 < \mathcal{J}(E) < \infty$ . For  $\varepsilon > 0$ , set  $\varepsilon' = \varepsilon(1 + 1/\mathcal{J}(E))^{-1}$ . Pick  $\delta_0 \in (0, \theta/2)$  such that  $c\sqrt{\delta_0} < \varepsilon'$  where  $c = c(p, q = 2, d, \zeta)$  is the constant appearing in the interface Lemma 6.1. Let  $B(x_i, r_i)$ ,  $i \in I$ , be a finite

collection of disjoint balls associated with  $E, \varepsilon'$  and  $\delta_0/3$ , as given in the covering Lemma 5.2. Let  $\delta > 0$  be smaller than  $\min\{\delta_0 r_i^d/3 | i \in I\}$ . Suppose that  $\text{dist}_{L^1}(\text{aglu } M^-/n, E) < \delta$ .

For  $i$  in  $I$  let  $\mathcal{C}(i)$  be the collection of the open clusters of the configuration restricted to  $\text{aglu } M^- \cap B(nx_i, nr_i)$ . Note that these clusters are open clusters of the configuration restricted to  $B(nx_i, nr_i)$ . We have

$$\begin{aligned}
 & \text{dist}_{L^1}(\text{cube cup } \mathcal{C}(i), B(nx_i, nr_i)) \\
 & \leq \text{dist}_{L^1}(\text{cube cup } \mathcal{C}(i), \text{aglu } M^- \cap B(nx_i, nr_i)) \\
 (58) \quad & + \text{dist}_{L^1}(\text{aglu } M^- \cap B(nx_i, nr_i), nE \cap B(nx_i, nr_i)) \\
 & + \text{dist}_{L^1}(nE \cap B(nx_i, nr_i), B_-(nx_i, nr_i, \nu_E(x_i))) \\
 & \leq c'(d)n^{d-1} + \delta n^d + \delta_0(nr_i)^d/3 \leq \delta_0(nr_i)^d,
 \end{aligned}$$

where  $c'(d)$  is an appropriate constant depending only on the dimension. If  $x_i$  belongs to  $\text{int } \Gamma$  then the collection  $\mathcal{C}(i)$  realizes the event  $\text{Sep}(n, x_i, r_i, \nu_E(x_i), \delta_0)$ . If  $x_i$  belongs to  $\partial^* \Gamma$  then the collection  $\mathcal{C}(i)$  realizes the event  $\text{Sep}^{\text{bd}}(n, x_i, r_i, \delta_0)$ . Indeed, none of the clusters inside  $\text{aglu } M^-$  intersects  $\partial^{\text{in}} \Lambda(n)$ : the large clusters in  $\text{aglu } M^-$  are colored with minus, and the small clusters attached to  $\mathcal{C}^-$  to build  $\text{aglu } M^-$  do not intersect  $\partial^{\text{in}} \Lambda(n)$ . We conclude that

$$\begin{aligned}
 \{\text{dist}_{L^1}(M^-/n, E) < \delta\} \subseteq & \bigcap_{\substack{i \in I \\ x_i \in \text{int } \Gamma}} \text{Sep}(n, x_i, r_i, \nu_E(x_i), \delta_0) \\
 & \cap \bigcap_{\substack{i \in I \\ x_i \in \partial^* \Gamma}} \text{Sep}^{\text{bd}}(n, x_i, r_i, \delta_0).
 \end{aligned}$$

Note that the sets  $\Gamma \cap B(x_i, r_i)$  are compact and disjoint. The decoupling Lemma 3.2 and the interface Lemma 6.1 together imply

$$\begin{aligned}
 & \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ \left[ \text{dist}_{L^1}(\text{aglu } M^-/n, E) < \delta \right] \\
 & \leq - \sum_{i \in I} \alpha_{d-1} r_i^{d-1} \tau(\nu_E(x_i))(1 - c\sqrt{\delta_0}) \\
 & \leq -\mathcal{J}(E)(1 - \varepsilon') + \varepsilon' = -\mathcal{J}(E)(1 - \varepsilon). \quad \square
 \end{aligned}$$

We complete now the proof of the upper bound for Theorem 1.1.

Let  $v \in [0, 1]$  be fixed and  $\lambda > 0$ . We define the sets  $S(v) = \{E \in \mathcal{B}(\Gamma) | \text{vol } E \geq v\}$  and  $S(v, \lambda) = \{E \in S(v) | \mathcal{J}(E) \leq \lambda\}$ . The sets  $S(v, \lambda), \lambda < \infty$ , are compact. Let  $\varepsilon > 0$ . For each  $E$  in  $S(v - \varepsilon, \lambda)$  we choose  $\delta(E, \varepsilon) > 0$  according to Lemma 7.9. The family

$$\{\{A \in \mathcal{B}(\Gamma) | \text{dist}_{L^1}(A, E) < \delta(E, \varepsilon)\}, E \in S(v - \varepsilon, \lambda)\}$$

is an open covering of  $S(v - \varepsilon, \lambda)$  from which we can extract a finite subcover associated with a finite collection of sets  $E_i, i \in I$ . Since  $S(v - \varepsilon, \lambda)$  is compact, we can choose  $\delta \in (0, \varepsilon)$  such that the  $\delta$ -neighborhood of  $S(v - \varepsilon, \lambda)$  is still contained in the subcover, that is,

$$\begin{aligned} & \{A \in \mathcal{B}(\Gamma) \mid \text{dist}_{L^1}(A, S(v - \varepsilon, \lambda)) < \delta\} \\ & \subset \bigcup_{i \in I} \{A \in \mathcal{B}(\Gamma) \mid \text{dist}_{L^1}(A, E_i) < \delta(E_i, \varepsilon)\}. \end{aligned}$$

Since  $\text{dist}_{L^1}(S(v), S(v - \varepsilon)^c) \geq \varepsilon > \delta$ ,

$$\begin{aligned} \mathbb{P}_{\Lambda(n)}^+[\text{vol } M^-/n \geq v] & \leq \mathbb{P}_{\Lambda(n)}^+[\text{dist}_{L^1}(M^-/n, S(v - \varepsilon, \lambda)) < \delta] \\ & \quad + \mathbb{P}_{\Lambda(n)}^+[M^-/n \in S(v), \text{dist}_{L^1}(M^-/n, S(v - \varepsilon, \lambda)) \geq \delta] \\ & \leq \sum_{i \in I} \mathbb{P}_{\Lambda(n)}^+[\text{dist}_{L^1}(M^-/n, E_i) < \delta(E_i, \varepsilon)] \\ & \quad + \mathbb{P}_{\Lambda(n)}^+[\text{dist}_{L^1}(M^-/n, \mathcal{I}^{-1}([0, \lambda])) \geq \delta]. \end{aligned}$$

Passing to the lim sup, the choice of the  $\delta(E_i, \varepsilon)$ -s and Lemma 7.7 yield

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+[\text{vol}(M^-/n) \geq v] & \leq -\left(c\lambda \wedge (1 - \varepsilon) \min_{i \in I} \mathcal{I}(E_i)\right) \\ & \leq -\left(c\lambda \wedge (1 - \varepsilon) \inf_{S(v-\varepsilon, \lambda)} \mathcal{I}\right). \end{aligned}$$

Since the sets  $S(v - \varepsilon, \lambda)$  are compact and  $\mathcal{I}$  is l.s.c.,

$$\lim_{\varepsilon \rightarrow 0} \inf_{S(v-\varepsilon, \lambda)} \mathcal{I} = \inf_{S(v, \lambda)} \mathcal{I} \geq \inf_{S(v)} \mathcal{I}.$$

By letting  $\varepsilon \rightarrow 0, \lambda \rightarrow \infty$  and applying the Wulff isoperimetric theorem [36], we obtain (7),

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+[\text{vol}(M^-/n) \geq v] \\ \leq -\inf_{S(v)} \mathcal{I} = -(v/\text{vol } \mathcal{W}_\tau)^{(d-1)/d} \mathcal{I}(\mathcal{W}_\tau). \end{aligned}$$

Let  $m$  belong to  $[-m^*, +m^*]$ . Let  $\delta > 0$  such that  $m + \delta < m^*$ . By Corollary 7.3,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ \left[ \frac{1}{n^d} \sum_{x \in \Lambda(n)} \sigma(x) \leq m \right] \\ \leq \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ \left[ \text{vol}(M^-/n) \geq \frac{m^* - \delta - m}{2m^*} \right] \\ \leq -\left(\frac{m^* - \delta - m}{2m^* \text{vol } \mathcal{W}_\tau}\right)^{(d-1)/d} \mathcal{I}(\mathcal{W}_\tau). \end{aligned}$$

Letting  $\delta$  go to 0, we obtain the correct upper bound for Theorem 1.1. Notice that this upper bound is valid for any  $m$  in  $[-m^*, +m^*]$ . However, we have the corresponding lower bound only for  $1 - 2(\text{diam } \mathcal{W}_\tau)^{-d} \text{vol } \mathcal{W}_\tau < m/m^* < 1$ . For the remaining values of  $m$  in  $(-m^*, m^*)$ , the correct value should be

$$\inf\{\mathcal{J}(E) \mid E \in \mathcal{B}(\Gamma), \text{vol } E \geq (m^* - m)/(2m^*)\}.$$

See [63] for the corresponding question in dimension two.

7.2. *The locally averaged magnetization.* We express  $\sigma_n$  with the help of the FK representation,

$$\forall x \in \Gamma \quad \sigma_n(x) = \frac{1}{f(n)^d} \sum_{C \in \mathcal{C}} \sigma(C) |C \cap \Lambda(nx, f(n))|.$$

We separate the contribution of the small and the large clusters by setting, for  $x \in \Gamma$ ,

$$\sigma_n^{\text{small}}(x) = \frac{1}{f(n)^d} \sum_{\substack{C \in \mathcal{C} \\ C \text{ small}}} \sigma(C) |C \cap \Lambda(nx, f(n))|$$

and  $\sigma_n^{\text{large}}(x)$  as in the previous case except that the sum is running over the large clusters.

LEMMA 7.10. *For any  $\delta > 0$ ,*

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ \left[ \int_{\Gamma} |\sigma_n^{\text{small}}(x)| dx > \delta \right] = -\infty.$$

PROOF. Let  $\delta \in (0, 1/2)$  and  $n$  so big that  $2dn^{d-1}2f(n) < \delta n^d$ . Set  $\Lambda(n)' = \Lambda(n - 2f(n)) \cap \mathbb{Z}^d$ . We denote by  $B(z)$  the open unit cube centered at  $z \in \mathbb{Z}^d$ . Note that  $\sigma_n^{\text{small}}(x)$  is constant in  $B(z)$  with common value  $\sigma_n^{\text{small}}(z)$ . Thus

$$\begin{aligned} \mathbb{P}_{\Lambda(n)}^+ \left[ \int_{\Lambda} |\sigma_n^{\text{small}}(x)| dx > 3\delta \right] &\leq \mathbb{P}_{\Lambda(n)}^+ \left[ \sum_{z \in \Lambda(n)'} \int_{B(z)} |\sigma_n^{\text{small}}(x/n)| dx > 2\delta n^d \right] \\ &= \mathbb{P}_{\Lambda(n)}^+ \left[ \sum_{z \in \Lambda(n)'} |\sigma_n^{\text{small}}(z/n)| > 2\delta n^d \right]. \end{aligned}$$

Let  $z \in \Lambda(n)'$ . If  $C$  is a small cluster intersecting  $\Lambda(z, f(n))$  then  $C \cap \delta^{\text{in}} \Lambda(n) = \emptyset$ ; hence  $\sigma(C)$  is  $\pm 1$  with probability  $1/2$  each and is independent of everything. Arguing as in the proof of Lemma 7.1 (no need for  $\mathcal{C}'$  here) we have

$$\mathbb{P}_{\Lambda(n)}^+ \left[ |\sigma_n^{\text{small}}(z)| > \delta \right] \leq 2 \exp\left(-\frac{c_3 \delta^2 f(n)^{d-2}}{\alpha}\right).$$

We call  $z$  regular if  $|\sigma_n^{\text{small}}(z)| \leq \delta$ . We set  $X_z = 0$  if  $z$  is regular and  $X_z = 1$  otherwise. There exists  $c_4 = c_4(p, d) > 0$  such that

$$(59) \quad \mathbb{P}_{\Lambda(n)}^+ \left[ X_z = 1 \mid \sigma(X_y; d_\infty(z, y) \geq 2f(n)) \right] \leq 2 \exp(-c_4 \delta^2 f(n)^{d-2}).$$

Then, for  $n$  large enough,

$$\begin{aligned} & \mathbb{P}_{\Lambda(n)}^+ \left[ \sum_{z \in \Lambda(n)'} |\sigma_n^{\text{small}}(z)| > 2\delta n^d \right] \\ &= \mathbb{P}_{\Lambda(n)}^+ \left[ \sum_{\substack{z \in \Lambda(n)' \\ X_z=0}} |\sigma_n^{\text{small}}(z)| + \sum_{\substack{z \in \Lambda(n)' \\ X_z=1}} |\sigma_n^{\text{small}}(z)| > 2\delta n^d \right] \\ &\leq \mathbb{P}_{\Lambda(n)}^+ \left[ \sum_{z \in \Lambda(n)'} X_z > \delta n^d \right]. \end{aligned}$$

In order to apply Lemma 2.2, by using (59) we first estimate  $\Lambda_\varepsilon^*(\delta)$ ,

$$\begin{aligned} \Lambda_{\varepsilon(n)}^*(\delta) &\geq \delta \log(\delta/\varepsilon(n)) - \log 2 \geq \delta \log((\delta/2) \exp(c_4 \delta^2 f(n)^{d-2})) - \log 2 \\ &\geq c_4 \delta^3 f(n)^{d-2} / 2 - 2 \log 2. \end{aligned}$$

Thus, for  $n$  large,

$$\mathbb{P}_{\Lambda(n)}^+ \left[ \sum_{z \in \Lambda(n)'} X_z > \delta n^d \right] \leq (2f(n))^d \exp\left(-\frac{cn^d}{2f(n)^2}\right)$$

for a certain constant  $c > 0$  and the claim follows.  $\square$

LEMMA 7.11. For  $\delta > 0$ ,

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ \left[ \int_{M^-/n} |\sigma_n^{\text{large}}(x) + m^*| dx \right. \\ \left. + \int_{M^+/n} |\sigma_n^{\text{large}}(x) - m^*| dx > \delta \right] = -\infty. \end{aligned}$$

PROOF. We work with the box  $\Lambda(n)$  rescaled by a factor  $f(n)$  and block size 2. The block variable  $X(\underline{x})$  is the indicator function of the event  $T(B'(\underline{x}), f(n) - 1, \delta/2)^c$ . Suppose that  $X(\underline{x}) = 0$  and let  $x$  be such that  $y = xn \in B(\underline{x})$ . Then there exists exactly one large open cluster  $C$  such that  $d_\infty(y, C) \leq f(n)$ . This cluster  $C$  satisfies in addition,

$$|\|C \cap \Lambda(y, f(n))\| - \theta f(n)^d| < (\delta/2)f(n)^d.$$

Therefore,  $y$  belongs to  $M^- \cup M^+$  and  $|\sigma_n^{\text{large}}(x) - \theta\sigma(C)| < \delta/2$ . We split the region of integration into the blocks  $B(\underline{x})/n, \underline{x} \in \underline{\Lambda}(n)$  and we use the previous

inequality:

$$\begin{aligned} & \int_{M^-/n} |\sigma_n^{\text{large}}(x) + m^*| dx + \int_{M^+/n} |\sigma_n^{\text{large}}(x) - m^*| dx \\ & \leq 2^d (f(n)/n)^d \sum_{\underline{x} \in \underline{\Lambda}(n)} X(\underline{x}) + \delta/2. \end{aligned}$$

By the estimate (14), the block process  $X(\underline{x})$  satisfies (15) with  $\varepsilon = b \exp(-cf(n))$ . The result follows from (18).  $\square$

We next compare the two random partitions  $(M^-/n, M^0/n, M^+/n)$  and  $(\Gamma_n^-, \Gamma_n^0, \Gamma_n^+)$ .

COROLLARY 7.12. *For any  $\delta > 0$ ,*

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ \left[ \text{dist}_{L^1}(M^-/n, \Gamma_n^-) + \text{vol } \Gamma_n^0 \right. \\ \left. + \text{dist}_{L^1}(M^+/n, \Gamma_n^+) > \delta \right] = -\infty. \end{aligned}$$

PROOF. With the help of some set algebra, we have

$$\begin{aligned} & \text{dist}_{L^1}(M^-/n, \Gamma_n^-) + \text{dist}_{L^1}(M^0/n, \Gamma_n^0) + \text{dist}_{L^1}(M^+/n, \Gamma_n^+) \\ & \leq 2 \text{vol}(M^-/n \cap (\Gamma_n^0 \cup \Gamma_n^+)) + 2 \text{vol}(M^0/n) + 2 \text{vol}(M^+/n \cap (\Gamma_n^0 \cup \Gamma_n^-)) \\ & \leq \frac{2}{m^*} \left( \int_{M^-/n} |\sigma_n(x) + m^*| dx + \text{vol}(M^0/n) + \int_{M^+/n} |\sigma_n(x) - m^*| dx \right) \end{aligned}$$

The claim of the corollary is a consequence of Lemmas 7.2, 7.10, 7.11.  $\square$

We now complete the proof of Theorem 1.2. Lemmas 7.10, 7.11 and Corollary 7.12 imply the first claim of Theorem 1.2. Let  $m$  be such that  $1 - 2(\text{diam } \mathscr{W}_\tau)^{-d} \text{vol } \mathscr{W}_\tau < m/m^* < 1$ . Let  $\delta > 0$ . By Theorem 1.1,

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mu_{\Lambda(n)}^+ \left[ \text{dist}_{L^1}(\Gamma_n^-, \text{mc}(\Gamma_n^-) + \mathscr{W}(m)) \geq \delta \mid \frac{1}{n^d} \sum_{x \in \Lambda(n)} \sigma(x) \leq m \right] \\ & \leq \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ \left[ \text{dist}_{L^1}(\Gamma_n^-, \text{mc}(\Gamma_n^-) + \mathscr{W}(m)) \geq \delta, \right. \\ & \qquad \qquad \qquad \left. \frac{1}{n^d} \sum_{x \in \Lambda(n)} \sigma(x) \leq m \right] + \mathcal{I}(\mathscr{W}(m)). \end{aligned}$$

Next, by Corollaries 7.3, 7.12 and Lemma 7.7, for any  $\lambda > 0, \varepsilon > 0$ ,

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ \left[ \text{dist}_{L^1}(\Gamma_n^-, \text{mc}(\Gamma_n^-) + \mathcal{W}(m)) \geq \delta, \frac{1}{n^d} \sum_{x \in \Lambda(n)} \sigma(x) \leq m \right] \\ & \leq -c\lambda \wedge \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \\ & \quad \times \log \mathbb{P}_{\Lambda(n)}^+ \left[ \begin{array}{l} \text{dist}_{L^1}(\Gamma_n^-, \text{mc}(\Gamma_n^-) + \mathcal{W}(m)) \geq \delta, \text{dist}_{L^1}(\Gamma_n^-, M^-/n) < \varepsilon, \\ \text{dist}_{L^1}(M^-/n, \mathcal{J}^{-1}([0, \lambda])) < \varepsilon, \text{vol}(M^-/n) \geq \text{vol} \mathcal{W}(m) - \varepsilon \end{array} \right]. \end{aligned}$$

Let  $\eta$  such that  $0 < 2\eta < \text{vol}(\mathcal{W}(m))$  and let us denote by  $\mathcal{E}(\delta, \lambda, \eta)$  the subset of  $\mathcal{B}(\Gamma)$  defined by

$$\begin{aligned} \mathcal{E}(\delta, \lambda, \eta) = \{ E \in \mathcal{B}(\Gamma) \mid & \text{dist}_{L^1}(E, \text{mc}(E) + \mathcal{W}(m)) \geq \delta - \eta, \\ & \text{vol } E \geq \text{vol}(\mathcal{W}(m)) - \eta, \mathcal{J}(E) \leq \lambda \}. \end{aligned}$$

The map  $E \in \mathcal{B}(\Gamma) \mapsto \text{mc}(E) \in \Gamma$  is continuous at each set having positive volume. Therefore for  $\varepsilon$  small enough, the inequalities involved in the above event imply the further inequality  $\text{dist}_{L^1}(M^-/n, \mathcal{E}(\delta, \lambda, \eta)) < \varepsilon$ . Moreover the set  $\mathcal{E}(\delta, \lambda, \eta)$  is compact. Proceeding as in the proof of the upper bound of Theorem 1.1, with the help of Lemmas 7.7, 7.9, we get the bound

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mathbb{P}_{\Lambda(n)}^+ [\text{dist}_{L^1}(M^-/n, \mathcal{E}(\delta, \lambda, \eta)) < \varepsilon] \\ & \leq - \left( c\lambda \wedge \inf_{E \in \mathcal{E}(\delta, \lambda, \eta)} \mathcal{J}(E) \right). \end{aligned}$$

Combining the previous inequalities, sending  $\lambda$  to  $\infty$  and  $\eta$  to 0, we arrive at

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \mu_{\Lambda(n)}^+ \left[ \text{dist}_{L^1}(\Gamma_n^-, \text{mc}(\Gamma_n^-) + \mathcal{W}(m)) \geq \delta \mid \frac{1}{n^d} \sum_{x \in \Lambda(n)} \sigma(x) \leq m \right] \\ & \leq - \inf_{E \in \mathcal{E}(\delta, \infty, 0)} \mathcal{J}(E) + \mathcal{J}(\mathcal{W}(m)). \end{aligned}$$

Since  $\mathcal{J}$  restricted to  $\Gamma$  has compact level sets it attains its minimum over the set  $\mathcal{E}(\delta, \infty, 0)$ . However this set contains only sets of volume larger than or equal to  $\text{vol} \mathcal{W}(m)$  but no translate of  $\mathcal{W}(m)$ . The uniqueness statement in the Wulff isoperimetric Theorem [37] implies that

$$\inf_{E \in \mathcal{E}(\delta, \infty, 0)} \mathcal{J}(E) > \mathcal{J}(\mathcal{W}(m)),$$

concluding thereby the proof of the second claim of Theorem 1.2.  $\square$

APPENDIX

**A.1. Results from geometric measure theory.** We recall first that for any subset  $A$  of  $\mathbb{R}^d$ , denoting by  $\alpha_k$  the volume of the unit ball of  $\mathbb{R}^k$ , the  $k$ -dimensional Hausdorff measure  $\mathcal{H}^k$  is defined by

$$\mathcal{H}^k(A) = \sup_{\delta > 0} \inf \left\{ \alpha_k 2^{-k} \sum_{i \in I} (\text{diam}_2 E_i)^k \mid A \subset \bigcup_{i \in I} E_i, \sup_{i \in I} \text{diam}_2 E_i \leq \delta \right\}.$$

Let  $E$  be a Borel subset of  $\mathbb{R}^d$ . A collection of sets  $\mathcal{U}$  is called a Vitali class for  $E$  if for each  $x \in E$  and  $\delta > 0$ , there exists a set  $U \in \mathcal{U}$  containing  $x$  such that  $0 < \text{diam}_2 U < \delta$ .

The *Vitali covering Theorem for  $\mathcal{H}^{d-1}$*  ([32], Theorem 1.10). Let  $E$  be an  $\mathcal{H}^{d-1}$ -measurable subset of  $\mathbb{R}^d$  and let  $\mathcal{U}$  be a Vitali class of closed sets for  $E$ . Then we may select a (countable) disjoint sequence  $(U_i)_{i \in I}$  from  $\mathcal{U}$  such that either  $\sum_{i \in I} (\text{diam } U_i)^{d-1} = \infty$  or  $\mathcal{H}^{d-1}(E \setminus \bigcup_{i \in I} U_i) = 0$ . If  $\mathcal{H}^{d-1}(E) < \infty$  then, given  $\varepsilon > 0$ , we may also require that  $\mathcal{H}^{d-1}(E) \leq \alpha_{d-1} 2^{-d+1} \sum_{i \in I} (\text{diam}_2 U_i)^{d-1} + \varepsilon$ .

For a general version concerning Radon measures, see [51], Theorem 2.8.

We recall next some facts concerning the class of the sets of finite perimeter, introduced initially by Caccioppoli [12, 13] and subsequently developed by De Giorgi [19, 20, 21, 22] (see also [31, 33, 39, 52, 70]). The perimeter of a Borel set  $E$  of  $\mathbb{R}^d$  is defined as

$$\mathcal{P}(E) = \sup \left\{ \int_E \text{div } f(x) \, dx : f \in C_0^\infty(\mathbb{R}^d, B(1)) \right\},$$

where  $C_0^\infty(\mathbb{R}^d, B(1))$  is the set of the compactly supported  $C^\infty$  vector functions from  $\mathbb{R}^d$  to the unit ball  $B(1)$  and  $\text{div}$  is the usual divergence operator. The set  $E$  is of finite perimeter if  $\mathcal{P}(E)$  is finite. A set  $E$  is a Caccioppoli set if it is locally of finite perimeter. In this paper, we deal with bounded sets, hence we need only to consider sets of finite perimeter. A set  $E$  has finite perimeter if and only if its characteristic function  $\chi_E$  is a function of bounded variation. The distributional derivative  $\nabla \chi_E$  of  $\chi_E$  is then a vector Radon measure and  $\mathcal{P}(E) = \|\nabla \chi_E\|(\mathbb{R}^d)$ , where  $\|\nabla \chi_E\|$  is the total variation measure of  $\nabla \chi_E$ . The perimeter  $\mathcal{P}$  is l.s.c. on the space  $(\mathcal{B}(\mathbb{R}^d), \text{dist}_{L^1})$ .

*Compactness property of sets of finite perimeter in  $\Gamma$ .* For every bounded domain  $U$  and every  $\lambda > 0$ , the set  $\{E \in \mathcal{B}(U) : \mathcal{P}(E) \leq \lambda\}$  is compact for the metric  $\text{dist}_{L^1}$ .

This result is stated in this precise form in [22], Teorema 2.4, or [20], Teorema I. It is also an immediate consequence of the compactness theorem stated in [52], Chapter 2, page 70. Modern presentations are formulated through functions of bounded variations: if  $O$  is an open bounded domain with sufficiently regular boundary (say  $C^1$ ), then a set of functions in  $L^1(O)$  uniformly bounded in BV-norm is relatively compact in  $L^1(O)$  (see any of the following references: [31], Section 5.2.3, [39], Theorem 1.19, [70], Corollary 5.3.4). To deduce the compactness result on sets of finite perimeter, we choose an open

bounded domain  $O$  with regular boundary containing  $U$  in its interior. We embed  $\mathcal{B}(U)$  in  $L^1(O)$  by associating to a Borel set  $E$  its characteristic function  $\chi_E$  and we simply remark that the set  $\{\chi_E \mid E \in \mathcal{B}(U)\}$  is a closed subset of  $L^1(O)$ .

Let  $E$  be a set of finite perimeter. Its reduced boundary  $\partial^*E$  consists of the points  $x$  such that:

1.  $\|\nabla\chi_E\|(B(x, r)) > 0$  for any  $r > 0$ .
2. If  $\nu_r(x) = -\nabla\chi_E(B(x, r))/\|\nabla\chi_E\|(B(x, r))$  then, as  $r$  goes to 0,  $\nu_r(x)$  converges toward a limit  $\nu_E(x)$  such that  $|\nu_E(x)|_2 = 1$ .

The reduced boundary  $\partial^*E$  is countably  $(d - 1)$ -rectifiable, that is  $\partial^*E \subset N \cup \bigcup_{i \in \mathbb{N}} M_i$  where  $\mathcal{H}^{d-1}(N) = 0$  and each  $M_i$  is a  $d - 1$ -dimensional embedded  $C^1$  submanifold of  $\mathbb{R}^d$ . For a point  $x$  belonging to  $\partial^*E$ , the vector  $\nu_E(x)$  is called the generalized exterior normal to  $E$  at  $x$ . A unit vector  $\nu$  is called the measure theoretic exterior normal to  $E$  at  $x$  if

$$\lim_{r \rightarrow 0} r^{-d} \text{vol}(B_-(x, r, \nu) \setminus E) = 0, \quad \lim_{r \rightarrow 0} r^{-d} \text{vol}(B_+(x, r, \nu) \cap E) = 0.$$

At each point  $x$  of the reduced boundary  $\partial^*E$  of  $E$ , the generalized exterior normal  $\nu_E(x)$  is also the measure theoretic exterior normal to  $E$  at  $x$ . The map  $x \in \partial^*E \mapsto \nu_E(x) \in S^{d-1}$  is  $\|\nabla\chi_E\|$  measurable. For any Borel set  $A$  of  $\mathbb{R}^d$ ,

$$\|\nabla\chi_E\|(A) = \mathcal{H}^{d-1}(A \cap \partial^*E), \quad \nabla\chi_E(A) = \int_{A \cap \partial^*E} -\nu_E(x) \mathcal{H}^{d-1}(dx).$$

Let  $f: \partial^*E \mapsto \mathbb{R}$  be a  $\|\nabla\chi_E\|$  measurable bounded function. By the Besicovitch derivation Theorem [6, 10] applied to the measure  $\|\nabla\chi_E\|$ , for  $\mathcal{H}^{d-1}$  almost all  $x$  in  $\partial^*E$ ,

$$\lim_{r \rightarrow 0} (\alpha_{d-1} r^{d-1})^{-1} \int_{B(x, r) \cap \partial^*E} f(y) \mathcal{H}^{d-1}(dy) = f(x).$$

For any vector function  $f$  in  $C_0^1(\mathbb{R}^d, \mathbb{R}^d)$ , any Caccioppoli set  $E$ , by the generalized Gauss–Green Theorem,

$$\int_E \text{div } f(x) dx = \int_{\partial^*E} f(x) \cdot \nu_E(x) \mathcal{H}^{d-1}(dx).$$

**A.2. Proofs of the results on surface tension.** For the sake of completeness, we include here the proofs on basic results on surface tension. They follow from minor adaptations of arguments in [14] and [53].

PROOF OF PROPOSITION 4.1. This result is proved with the help of the same subadditivity argument used in [2], Proposition 2.4. The only additional problem is that we work with curves whose position with respect to the discrete lattice  $\mathbb{Z}^d$  is arbitrary. Let  $w$  be a unit vector of  $\mathbb{R}^d$  and let  $A, A'$  be two hyperrectangles such that  $\text{nor } A = \text{nor } A' = w$ . Let  $\phi(n), \phi'(n)$  be two functions from  $\mathbb{N}$  to  $\mathbb{R}^+ \cup \{\infty\}$  such that  $\lim_{n \rightarrow \infty} \phi(n) = \infty, \lim_{n \rightarrow \infty} \phi'(n) = \infty$ . Let  $\zeta, \zeta'$  be two real numbers larger than  $2d$ . Let  $n, m$  in  $\mathbb{N}$  be such that  $n \text{ diam}_2 A > m \text{ diam}_2 A' > \max(\zeta, \zeta')$ . Because we deal with hyperrectangles,

certainly there exists a collection of sets  $(T(i), i \in I)$  such that each set  $T(i)$  is a translate of  $mA'$  intersecting the set

$$D(m, n) = \{x \in nA: d_2(x, n\partial A) > 2m \operatorname{diam}_2 A'\};$$

the sets  $(T(i), i \in I)$  have pairwise disjoint interiors; their union  $\bigcup_{i \in I} T(i)$  contains the set  $D(m, n)$ . Since  $A$  is a hyperrectangle, then

$$(60) \quad \begin{aligned} & \mathcal{H}^{d-1}(nA) - 2m(\operatorname{diam}_2 A')\mathcal{H}^{d-2}(n\partial A) \\ & \leq \mathcal{H}^{d-1}(D(m, n)) \leq |I|\mathcal{H}^{d-1}(mA') \leq \mathcal{H}^{d-1}(nA). \end{aligned}$$

For each  $i$  in  $I$ , let  $t(i)$  be a vector in  $\mathbb{R}^d$  such that  $|t(i)|_\infty \leq 1$  and  $t(i) + T(i)$  is the image of  $mA'$  by an integer translation (a translation that leaves  $\mathbb{Z}^d$  globally invariant). Let  $T'(i) = t(i) + T(i)$ . Suppose that all the events  $W(\partial T'(i), \phi(n), \zeta')$ ,  $i \in I$ , occur, and let  $E(i)$ ,  $i \in I$ , be finite sets of closed edges realizing these events. Let  $c(d, \zeta')$  be a positive constant and let  $E_0$  be the set of edges included in

$$\begin{aligned} & (\operatorname{cyl}(nA \setminus D(m, n)) \cap \mathcal{V}_2(\operatorname{hyp} nA, \zeta)) \\ & \cup \bigcup_{i \in I} (\mathcal{V}_2(\operatorname{cyl} \partial T'(i), c(d, \zeta')) \cap \mathcal{V}_2(\operatorname{hyp} nA, c(d, \zeta'))). \end{aligned}$$

Let  $E = E_0 \cup \bigcup_{i \in I} E(i)$ . Clearly  $E$  is finite. The constant  $c(d, \zeta')$  can be chosen large enough (depending only on  $d, \zeta'$ ) to guarantee that the edges of  $E$  separate  $\infty$  inside  $\operatorname{cyl} nA$ . Then the set of edges  $E$  realizes the event  $W(\partial nA, \phi(n), \zeta)$ . An attempt of proof is done in [14], Proposition 5.2. Therefore

$$\{\omega \mid \forall e \in E_0 \omega(e) = 0\} \cap \left( \bigcap_{i \in I} W(\partial T'(i), \phi(n), \zeta') \right) \subseteq W(\partial nA, \phi(n), \zeta).$$

Since all these events are decreasing, by the FKG inequality,

$$(61) \quad \Phi_\infty[W(\partial nA, \phi(n), \zeta)] \geq (1-p)^{|E_0|} \prod_{i \in I} \Phi_\infty[W(\partial T'(i), \phi(n), \zeta')].$$

Since the model is invariant under the integer translations, for any  $i$  in  $I$ ,

$$(62) \quad \Phi_\infty[W(\partial T'(i), \phi(n), \zeta')] = \Phi_\infty[W(\partial mA', \phi(n), \zeta')].$$

Because  $\phi(n)$  goes to  $\infty$  as  $n$  goes to  $\infty$ ,

$$\lim_{n \rightarrow \infty} \Phi_\infty[W(\partial mA', \phi(n), \zeta')] = \Phi_\infty[W(\partial mA', \infty, \zeta')],$$

whence, for  $n$  sufficiently large,

$$(63) \quad \Phi_\infty[W(\partial mA', \phi(n), \zeta')] \geq (1/2)\Phi_\infty[W(\partial mA', \infty, \zeta')].$$

For such integers  $n$ , combining (61), (62), (63) and passing to the logarithm,

$$(64) \quad \begin{aligned} \log \Phi_\infty[W(\partial nA, \phi(n), \zeta)] & \geq |I| \log \Phi_\infty[W(\partial mA', \phi'(m), \zeta')] \\ & \quad + |E_0| \log(1-p) - |I| \log 2. \end{aligned}$$

There exists a further constant  $c(d, \zeta, \zeta', A, A')$  such that

$$(65) \quad |E_0| \leq c(d, \zeta, \zeta', A, A')(n^{d-2}m + n^{d-1}/m + 1).$$

Using the inequalities (60), (64) and (65), we obtain

$$\begin{aligned} & \mathcal{H}^{d-1}(nA)^{-1} \log \Phi_\infty[W(\partial nA, \phi(n), \zeta)] \\ & \geq \mathcal{H}^{d-1}(mA')^{-1} \log \Phi_\infty[W(\partial mA', \phi'(m), \zeta')] \\ & \quad + c(d, \zeta, \zeta', A, A') \mathcal{H}^{d-1}(A)^{-1} (m/n + 1/m + 1/n^{d-1}) \log(1-p) \\ & \quad - \mathcal{H}^{d-1}(mA')^{-1} \log 2. \end{aligned}$$

Sending successively  $n$  to  $\infty$  and then  $m$  to  $\infty$  yields

$$\begin{aligned} & \liminf_{n \rightarrow \infty} \mathcal{H}^{d-1}(nA)^{-1} \log \Phi_\infty[W(\partial nA, \phi(n), \zeta)] \\ & \geq \limsup_{m \rightarrow \infty} \mathcal{H}^{d-1}(mA')^{-1} \log \Phi_\infty[W(\partial mA', \phi'(m), \zeta')], \end{aligned}$$

which implies the result of the proposition.  $\square$

PROOF OF PROPOSITION 4.5. Let  $w \in S^{d-1}$  and let  $A$  be a hyperrectangle orthogonal to  $w$  such that  $\mathcal{H}^{d-1}(A) = 1$ . Let  $E(n)$  be the set of the edges included in  $\text{cyl } nA \cap \mathcal{V}_2(\text{hyp } nA, 2d)$ . Then  $|E(n)| \leq c(d)n^{d-1}$  and

$$\Phi_\infty[W(\partial nA, \infty, 2d)] \geq \Phi_\infty[\text{the edges of } E(n) \text{ are closed}] \geq (1-p)^{|E(n)|}.$$

Passing to the limit, we get  $\tau(w) \leq -c(d) \log(1-p)$ . Since  $\tau_0$  is homogeneous, convex (by Corollary 4.3) and bounded on  $S^{d-1}$ , it is finite everywhere. By a standard result of convex analysis [60], Corollary 10.1.1, it follows that  $\tau_0$  is continuous, as well as  $\tau$ .

Recall that we work here in the region  $d \geq 3, q \geq 1, p > \hat{p}_c, \theta^f(p) = \theta^w(p)$ . We first show that  $\tau(1, 0, \dots, 0)$  is positive. Let  $\delta$  be positive and let  $F_1, \dots, F_{2d}$  be the  $2d$  faces of the cubic box  $\Lambda(1-2\delta)$ . Let  $E_0$  be the set of edges included in  $\mathcal{V}_2(n\partial F_1 \cup \dots \cup n\partial F_{2d}, 2\zeta)$ . There exists a constant  $c(d, \zeta)$  such that  $|E_0| \leq c(d, \zeta)n^{d-2}$ . If the events  $W(n\partial F_i, \delta n, \zeta), 1 \leq i \leq 2d$ , {all the edges of  $E_0$  are closed}, occur simultaneously, then there exists a set of closed edges inside the box  $\Lambda(n)$  which separates the box  $n\Lambda(1-4\delta)$  from  $\partial^{\text{in}}\Lambda(n)$ . By the FKG inequality,

$$(1-p)^{|E_0|} \prod_{1 \leq i \leq 2d} \Phi_\infty[W(n\partial F_i, \delta n, \zeta)] \leq \Phi_\infty[\{n\Lambda(1-4\delta) \leftrightarrow \partial^{\text{in}}\Lambda(n)\}^c].$$

The bound on  $|E_0|$ , Proposition 4.1 and the symmetry of  $\tau$  imply that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \left( (1-p)^{|E_0|} \prod_{1 \leq i \leq 2d} \Phi_\infty[W(n\partial F_i, \delta n, \zeta)] \right) \\ & = -2d(1-2\delta)^{d-1} \tau(1, 0, \dots, 0). \end{aligned}$$

Next, the event  $\{n\Lambda(1 - 4\delta) \leftrightarrow \partial^{\text{in}}\Lambda(n)\}$  depends only on the status of the edges inside  $\Lambda(n)$ . By conditioning on the configuration outside  $\Lambda(n)$ , we get

$$\begin{aligned} & \Phi_\infty[\{n\Lambda(1 - 4\delta) \leftrightarrow \partial^{\text{in}}\Lambda(n)\}^c] \\ &= \Phi_\infty[\Phi_\infty[\{n\Lambda(1 - 4\delta) \leftrightarrow \partial^{\text{in}}\Lambda(n)\}^c | \mathcal{F}^{\Lambda(n)}]] \\ &\leq \sup_{\Phi \in c\mathcal{FK}(\Lambda(n))} \Phi[\{n\Lambda(1 - 4\delta) \leftrightarrow \partial^{\text{in}}\Lambda(n)\}^c]. \end{aligned}$$

Yet the event  $\{n\Lambda(1 - 4\delta) \leftrightarrow \partial^{\text{in}}\Lambda(n)\}^c$  implies that any cluster in  $\Lambda(n)$  intersecting  $\partial^{\text{in}}\Lambda(n)$  is included in  $\Lambda(n) \setminus n\Lambda(1 - 4\delta)$  and has therefore cardinality less than  $n^d(1 - (1 - 4\delta)^d)$ . We choose  $\delta$  small enough so that  $1 - (1 - 4\delta)^d \leq \theta/2$ . By the result of Pisztora ([59], Theorem 1.2) we have then

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \sup_{\Phi \in c\mathcal{FK}(\Lambda(n))} \Phi[\{\Lambda(n(1 - 4\delta)) \leftrightarrow \partial^{\text{in}}\Lambda(n)\}^c] < 0$$

so that  $\tau(1, 0, \dots, 0)$  is positive.

Suppose that  $\tau(\nu) = 0$  for some  $\nu$  in  $S^{d-1}$ . Let  $f_1, \dots, f_{d-1}$  be linear isometries of  $\mathbb{R}^d$  such that  $f_1(0) = \dots = f_{d-1}(0) = 0$ ,  $f_1(\mathbb{Z}^d) = \dots = f_{d-1}(\mathbb{Z}^d) = \mathbb{Z}^d$  and  $(f_1(\nu), \dots, f_{d-1}(\nu), \nu)$  is an orthonormal basis of  $\mathbb{R}^d$ . Then  $\tau(f_1(\nu)) = \dots = \tau(f_{d-1}(\nu)) = 0$ . Applying the weak simplex inequality to a pyramid having for basis a hypersimplex orthogonal to  $(1, 0, \dots, 0)$  and whose  $d$  other faces are orthogonal to  $f_1(\nu), \dots, f_{d-1}(\nu), \nu$ , we obtain that  $\tau(1, 0, \dots, 0) = 0$ , a contradiction. Thus the surface tension  $\tau$  does not vanish on  $S^{d-1}$ . Since  $\tau$  is continuous on  $S^{d-1}$ , it is bounded away from 0 on  $S^{d-1}$ .  $\square$

PROOF OF LEMMA 4.7. Let  $(A_i, i \in I)$  be a finite family of hyperrectangles in hyp  $O$  having disjoint relative interiors and covering  $O$ . Let  $c = c(d, \zeta)$  be a large constant. Let  $E_0$  be the set of edges included in the union

$$\bigcup_{i \in I} \mathcal{V}_2(\text{cyl } \partial nA_i, c) \cap \mathcal{V}_2(\text{hyp } nO, c).$$

There exists a further constant  $c' = c'(d, \zeta, c)$  such that  $|E_0| \leq c'|I|n^{d-2}$ . If all the events  $W(\partial nA_i, \phi(n))$ ,  $i \in I$ , occur and all the edges of  $E_0$  are closed, then  $S(nO, \phi(n))$  occurs as well, provided the constant  $c$  is large enough. By the FKG inequality,

$$\Phi_\infty[S(nO, \phi(n))] \geq (1 - p)^{|E_0|} \prod_{i \in I} \Phi_\infty[W(\partial nA_i, \phi(n))],$$

whence, by Proposition 4.1 and the bound on  $|E_0|$ ,

$$\liminf_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_\infty[S(nO, \phi(n))] \geq -\tau(\text{nor } O) \sum_{i \in I} \mathcal{H}^{d-1}(A_i).$$

By taking the supremum of the right-hand side over all possible coverings of  $O$ , we obtain the claim of the lemma.  $\square$

PROOF OF LEMMA 4.8. Let  $\varepsilon > 0$ . Let  $A^\varepsilon$  be a hyperrectangle such that

$$\mathcal{V}_\infty(A, 2\varepsilon) \cap \text{hyp } A \subset A^\varepsilon \subset \mathcal{V}_\infty(A, 3\varepsilon) \cap \text{hyp } A.$$

Let  $E_0$  be the set of the edges included in

$$(\mathcal{V}_2(\text{cyl } n\partial A, \zeta) \cap \mathcal{V}_2(\text{hyp } nA, nr)) \cup (\text{cyl}(nA^\varepsilon \setminus nA) \cap \mathcal{V}_2(\text{hyp } nA, \zeta)).$$

Suppose that the event  $S(nA, nr)$  occurs, and let  $E_S$  be a set of closed edges realizing it. Suppose also that all the edges of  $E_0$  are closed. Then the set of closed edges  $E_0 \cup E_S$  realizes the event  $W(\partial nA^\varepsilon, nr)$ . Therefore

$$\{\omega \mid \forall e \in E_0 \ \omega(e) = 0\} \cap S(nA, nr) \subset W(\partial nA^\varepsilon, nr).$$

Since all these events are decreasing, by the FKG inequality,

$$\Phi_\infty[S(nA, nr)](1-p)^{|E_0|} \leq \Phi_\infty[W(\partial nA^\varepsilon, nr)].$$

There exists a constant  $c = c(d, \zeta)$  such that

$$|E_0| \leq cn^{d-1}(r\mathcal{H}^{d-2}(\partial A) + \mathcal{H}^{d-1}(A^\varepsilon \setminus A)),$$

whence, passing to the logarithm,

$$\begin{aligned} \log \Phi_\infty[S(nA, nr)] &\leq \log \Phi_\infty[W(\partial nA^\varepsilon, nr)] - cn^{d-1}(r\mathcal{H}^{d-2}(\partial A) \\ &\quad + \mathcal{H}^{d-1}(A^\varepsilon \setminus A)) \log(1-p). \end{aligned}$$

Letting  $n$  go to  $\infty$ , applying Proposition 4.1 and sending  $\varepsilon$  to 0, we get the desired inequality.  $\square$

PROOF OF LEMMA 4.9. The very definition of the event  $S(nO, nr)$  implies that

$$S(nO, nr) \subset \bigcap_{i \in I} S(nA_i, nr).$$

For  $i$  in  $I$ , the event  $S(nA_i, nr)$  depends only on the status of the edges inside  $n \text{cyl}(A_i, r)$ . Since the hyperrectangles  $(A_i, i \in I)$  are pairwise disjoint and compact, so are the sets  $(\text{cyl}(A_i, r), i \in I)$ . Thus the decoupling Lemma 3.2

is in force and

$$\limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_\infty[S(nO, nr)] \leq \sum_{i \in I} \limsup_{n \rightarrow \infty} \frac{1}{n^{d-1}} \log \Phi_\infty[S(nA_i, nr)].$$

The conclusion follows from Lemma 4.8.  $\square$

**PROOF OF COROLLARY 4.10.** We apply Lemma 4.9 with  $O = \text{disc}(x, \rho, w)$ ,  $r = \eta$ . There exists a constant  $c' = c'(d)$  such that, for any  $\varepsilon$  positive, there exists a finite family  $(A_i, i \in I)$  of disjoint hyperrectangles included in  $O$  such that

$$\sum_{i \in I} \mathcal{H}^{d-1}(A_i) \geq \alpha_{d-1} \rho^{d-1} - \varepsilon, \quad \sum_{i \in I} \mathcal{H}^{d-2}(\partial A_i) \leq c' \rho^{d-2}.$$

The result follows by taking the infimum over all possible families in the inequality stated in Lemma 4.9.  $\square$

**Acknowledgments.** We thank D. Ioffe for pointing out to us that the specification of the coupling in the proof of the decoupling lemma was missing in the initial version. We are grateful to R. B. Griffiths and M. Widom for giving us advice on the physical aspects of phase separation. R. Cerf thanks the Center for Nonlinear Analysis at Carnegie Mellon University for its hospitality.

## REFERENCES

- [1] AIZENMAN, M., BRICMONT, J. and LEBOWITZ, J. L. (1987). Percolation of the minority spins in high-dimensional Ising models. *J. Statist. Phys.* **49** 859–865.
- [2] AIZENMAN, M., CHAYES, J. T., CHAYES, L., FRÖHLICH, J. and RUSSO, L. (1983). On a sharp transition from area law to perimeter law in a system of random surfaces. *Comm. Math. Phys.* **92** 19–69.
- [3] ALBERTI, G., BELLETTINI, G., CASSANDRO, M. and PRESUTTI, E. (1996). Surface tension in Ising systems with Kac potentials. *J. Statist. Phys.* **82** 743–796.
- [4] ALEXANDER, K. S. (1992). Stability of the Wulff minimum and fluctuations in shape for large finite clusters in two-dimensional percolation. *Probab. Theory Related Fields* **91** 507–532.
- [5] ALEXANDER, K. S., CHAYES, J. T. and CHAYES, L. (1990). The Wulff construction and asymptotics of the finite cluster distribution for two-dimensional Bernoulli percolation. *Comm. Math. Phys.* **131** 1–50.
- [6] ASSOUAD, P. and QUENTIN DE GROMARD, T. (1998). Sur la dérivation des mesures dans  $\mathbb{R}^n$ . Note.
- [7] BELLETTINI, G., CASSANDRO, M. and PRESUTTI, E. (1996). Constrained minima of nonlocal free energy functionals. *J. Statist. Phys.* **84** 1337–1349.
- [8] BENOIS, O., BODINEAU, T., BUTTÀ, P. and PRESUTTI, E. (1997). On the validity of van der Waals theory of surface tension. *Markov Process. Related Fields* **3** 175–198.
- [9] BENOIS, O., BODINEAU, T. and PRESUTTI, E. (1998). Large deviations in the van der Waals limit. *Stochastic Process. Appl.* **75** 89–104.
- [10] BESICOVITCH, A. S. (1946). A general form of the covering principle and relative differentiation of additive functions. *Proc. Cambridge Philos. Soc.* **41** 103–110; II. *Proc. Cambridge Philos. Soc.* **42** 1–10.

- [11] BODINEAU, T. (1999). The Wulff construction in three and more dimensions. *Comm. Math. Phys.* **207** 197–229.
- [12] CACCIOPPOLI, R. (1952). Misura e integrazione sugli insiemi dimensionalmente orientati I, II. *Rend. Accad. Naz. Lincei Cl. Sci. Fis. Mat. Nat.* (8) **XII** 3–11, 137–146.
- [13] CACCIOPPOLI, R. (1952). Misura e integrazione sulle varietà parametriche I, II, III. *Rend. Accad. Naz. Lincei Cl. Sci. Fis. Mat. Nat.* (8) **XII** 219–227, 365–373, 629–634.
- [14] CERF, R. (1998). Large deviations for three-dimensional supercritical percolation. *Astérisque* **267** (2000).
- [15] CERF, R. and PISZTORA, Á. (2000). Phase coexistence in Ising, Potts and percolation models. Preprint.
- [16] CESI, F., GUADAGNI, G., MARTINELLI, F. and SCHONMANN, R. (1996). On the 2D stochastic Ising model in the phase coexistence region near the critical point. *J. Statist. Phys.* **85** 55–102.
- [17] CHAYES, J. T., CHAYES, L. and SCHONMANN, R. H. (1987). Exponential decay of connectivities in the two-dimensional Ising model. *J. Statist. Phys.* **49** 433–445.
- [18] COMETS, F. (1986). Grandes déviations pour des champs de Gibbs sur  $\mathbb{Z}^d$ . *C. R. Acad. Sci. Sér. I* **303** 511–513.
- [19] DE GIORGI, E. (1954). Su una teoria generale della misura  $(r - 1)$ -dimensionale in uno spazio ad  $r$  dimensioni. *Ann. Mat. Pura Appl.* **36** 191–213.
- [20] DE GIORGI, E. (1955). Nuovi teoremi relativi alle misure  $(r - 1)$ -dimensionali in uno spazio ad  $r$  dimensioni. *Ricerche Mat.* **4** 95–113.
- [21] DE GIORGI, E. (1958). Sulla proprietà isoperimetrica dell'ipersfera, nella classe degli insiemi aventi frontiera orientata di misura finita. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Nat.* (8) **5** 33–44.
- [22] DE GIORGI, E., COLOMBINI, F. and PICCININI, L. C. (1972). Frontiere orientate di misura minima e questioni collegate. Scuola Normale Superiore di Pisa.
- [23] DEUSCHEL, J.-D. and PISZTORA, Á. (1996). Surface order large deviations for high-density percolation. *Probab. Theory Related Fields* **104** 467–482.
- [24] DEUSCHEL, J.-D. and STROOCK, D. W. (1989). *Large Deviations*. Academic Press, New York.
- [25] DINGHAS, A. (1944). Über einen geometrischen Satz von Wulff für die Gleichgewichtsform von Kristallen. *Z. Kristallogr.* **105** 304–314.
- [26] DOBRUSHIN, R. K. and HRYNIV, O. (1997). Fluctuations of the phase boundary in the 2D Ising ferromagnet. *Comm. Math. Phys.* **189** 395–445.
- [27] DOBRUSHIN, R. L., KOTECKÝ, R. and SHLOSMAN, S. B. (1992). Wulff construction: a global shape from local interaction. *Amer. Math. Soc. Transl. Ser. 2*.
- [28] DOBRUSHIN, R. L. and SHLOSMAN, S. B. (1992). Thermodynamic inequalities for the surface tension and the geometry of the Wulff construction. In *Ideas and Methods in Quantum and Statistical Physics* (S. Albeverio, ed.) 461–483. Cambridge Univ. Press.
- [29] EDWARDS, R. G. and SOKAL, A. D. (1988). Generalization of the Fortuin–Kasteleyn–Swendsen–Wang representation and Monte Carlo algorithm. *Phys. Rev. D* **38** 2009–2012.
- [30] ELLIS, R. S. (1986). *Entropy, Large Deviations and Statistical Mechanics*. Springer, New York.
- [31] EVANS, L. C. and GARIÉPY, R. F. (1992). *Measure theory and fine properties of functions*. CRC Press, Boca Raton, FL.
- [32] FALCONER, K. J. (1985). *The Geometry of Fractal Sets*. Cambridge Univ. Press.
- [33] FEDERER, H. (1969). *Geometric Measure Theory*. Springer, New York.
- [34] FÖLLMER, H. and OREY, S. (1988). Large deviations for the empirical field of a Gibbs measure. *Ann. Probab.* **16** 961–977.
- [35] FÖLLMER, H. and ORT, M. (1988). Large deviations and surface entropy for Markov fields. *Astérisque* 157–158, 173–190.
- [36] FONSECA, I. (1991). The Wulff theorem revisited. *Proc. Roy. Soc. London Ser. A* **432** 125–145.
- [37] FONSECA, I. and MÜLLER, S. (1991). A uniqueness proof for the Wulff theorem. *Proc. Roy. Soc. Edinburgh Ser. A* **119** 125–136.
- [38] FORTUIN, C. M. and KASTELEYN, P. W. (1972). On the random cluster model. I. *Physica* **57** 536–564.
- [39] GIUSTI, E. (1984). *Minimal surfaces and functions of bounded variation*. Birkhäuser, Boston.

- [40] GRIMMETT, G. R. (1989). *Percolation*. Springer, New York.
- [41] GRIMMETT, G. R. (1995). The stochastic random-cluster process and the uniqueness of random-cluster measures. *Ann. Probab.* **23** 1461–1510.
- [42] GRIMMETT, G. R. and MARSTRAND, J. M. (1990). The supercritical phase of percolation is well behaved. *Proc. Roy. Soc. Lond. Ser. A* **430** 439–457.
- [43] HRYNIV, O. (1998). On local behaviour of the phase separation line in the 2D Ising model. *Probab. Theory Related Fields* **110** 91–107.
- [44] IOFFE, D. (1993). Large deviations for the 2D Ising model: a lower bound without cluster expansions. *J. Statist. Phys.* **74** 411–432.
- [45] IOFFE, D. (1995). Exact large deviation bounds up to  $T_c$  for the Ising model in two dimensions. *Probab. Theory Related Fields* **102** 313–330.
- [46] IOFFE, D. and SCHONMANN, R. (1998). Dobrushin–Kotecký–Shlosman Theorem up to the critical temperature. *Comm. Math. Phys.* **199** 117–167.
- [47] KESTEN, H. and ZHANG, Y. (1990). The probability of a large finite cluster in supercritical Bernoulli percolation. *Ann. Probab.* **18** 537–555.
- [48] LEBOWITZ, J. L. (1977). Coexistence of phases in Ising ferromagnets. *J. Statist. Phys.* **16** 463–476.
- [49] LEBOWITZ, J. L. and MARTIN-LÖF, A. (1972). On the uniqueness of the equilibrium state for Ising spin systems. *Comm. Math. Phys.* **25** 276–282.
- [50] LIGGETT, T. M., SCHONMANN, R. H. and STACEY, A. M. (1997). Domination by product measures. *Ann. Probab.* **25** 71–95.
- [51] MATTILA, P. (1995). *Geometry of Sets and Measures in Euclidean Spaces. Fractals and Rectifiability*. Cambridge Univ. Press.
- [52] MASSARI, U. and MIRANDA, M. (1994). *Minimal Surfaces of Codimension One*. North-Holland, Amsterdam.
- [53] MESSENGER, A., MIRACLE-SOLÉ, S. and RUIZ, J. (1992). Convexity properties of the surface tension and equilibrium crystals. *J. Statist. Phys.* **67** 449–469.
- [54] MIRACLE-SOLÉ, S. (1995). Surface tension, step free energy, and facets in the equilibrium crystal. *J. Statist. Phys.* **79** 183–214.
- [55] NEWMAN, C. M. (1993). Disordered Ising systems and random cluster representations. In *Probability and Phase Transition* (G. Grimmett, ed.) 247–260. Kluwer, Dordrecht.
- [56] OLLA, S. (1988). Large deviations for Gibbs random fields. *Probab. Theory Related Fields* **77** 395–409.
- [57] PFISTER, C. E. (1991). Large deviations and phase separation in the two-dimensional Ising model. *Helv. Phys. Acta* **64** 953–1054.
- [58] PFISTER, C. E. and VELENIK, Y. (1997). Large deviations and continuum limit in the 2D Ising model. *Probab. Theory Related Fields* **109** 435–506.
- [59] PISZTORA, Á. (1996). Surface order large deviations for Ising, Potts and percolation models. *Probab. Theory Related Fields* **104** 427–466.
- [60] ROCKAFELLAR, R. T. (1970). *Convex Analysis*. Princeton Univ. Press.
- [61] SCHNEIDER, R. (1993). Convex bodies: the Brunn–Minkowski Theory. *Encyclopedia of Mathematics and Its Applications* **44**. Cambridge Univ. Press.
- [62] SCHONMANN, R. H. (1987). Second order large deviation estimates for ferromagnetic systems in the phase coexistence region. *Comm. Math. Phys.* **112** 409–422.
- [63] SCHONMANN, R. and SHLOSMAN, S. B. (1996). Constrained variational problem with applications to the Ising model. *J. Statist. Phys.* **83** 867–905.
- [64] SCHONMANN, R. and SHLOSMAN, S. B. (1996). Complete analyticity for the 2D Ising model completed. *Comm. Math. Phys.* **170** 453–482.
- [65] SCHONMANN, R. and SHLOSMAN, S. B. (1998). Wulff droplets and the metastable relaxation of kinetic Ising models. *Comm. Math. Phys.* **194** 389–462.
- [66] TAYLOR, J. E. (1974). Existence and structure of solutions to a class of nonelliptic variational problems. *Sympos. Math.* **14** 499–508.
- [67] TAYLOR, J. E. (1975). Unique structure of solutions to a class of nonelliptic variational problems. *Proc. Symp. Pure Math.* **27** 419–427.
- [68] TAYLOR, J. E. (1978). Crystalline variational problems. *Bull. Amer. Math. Soc.* **84** 568–588.

- [69] WULFF, G. (1901). Zur Frage der Geschwindigkeit des Wachstums und der Auflösung der Kristallflächen. *Z. Kristallogr.* **34** 449–530.
- [70] ZIEMER, W. P. (1989). *Weakly differentiable Functions. Sobolev Spaces and Functions of Bounded Variation*. Springer, New York.

CNRS, UNIVERSITÉ PARIS SUD  
MATHÉMATIQUE, BÂTIMENT 425  
91405 ORSAY CEDEX  
FRANCE

DEPARTMENT OF MATHEMATICAL SCIENCES  
CARNEGIE MELLON UNIVERSITY  
PITTSBURGH, PENNSYLVANIA 15213-3890