WULFF DROPLETS AND THE METASTABLE RELAXATION OF KINETIC ISING MODELS

Roberto H. Schonmann and Senya B. Shlosman*

Abstract. We consider the kinetic Ising models (Glauber dynamics) corresponding to the infinite volume Ising model in dimension 2 with nearest neighbor ferromagnetic interaction and under a positive external magnetic field h. Minimal conditions on the flip rates are assumed, so that all the common choices are being considered. We study the relaxation towards equilibrium when the system is at an arbitrary subcritical temperature T and the evolution is started from a distribution which is stochastically lower than the $(-)$-phase. We show that as $h \searrow 0$ the relaxation time blows up as $\exp(\lambda_c(T)/h)$, with $\lambda_c(T) = w(T)^2/(12Tm^*(T))$. Here $m^*(T)$ is the spontaneous magnetization and $w(T)$ is the integrated surface tension of the Wulff body of unit volume. Moreover, for $0 < \lambda < \lambda_c$, the state of the process at time $\exp(\lambda/h)$ is shown to be close, when h is small, to the $(-)$-phase. The difference between this state and the $(-)$-phase can be described in terms of an asymptotic expansion in powers of the external field. This expansion can be interpreted as describing a set of $\mathcal{C}^\infty$ continuations in h of the family of Gibbs distributions with the negative magnetic fields into the region of positive fields.

Contents

1991 *Mathematics Subject Classification.* Primary 60K35, 82B27.

Key words and phrases. kinetic Ising model, stochastic Ising model, Glauber dynamics, metastability, relaxation, nucleation, droplet growth, Wulff shape, large deviations, asymptotic expansion, spectral gap estimates, block-rescaling, percolation techniques, coupling, attractiveness.

*The work of R.H.S was partially supported by the N.S.F. through grants DMS 9100725 and DMS 9400644 and that of S.B.S. by the N.S.F. through grant DMS 9208029 and by the Russian Fund for Fundamental Research through grant 930101470

1. Introduction

1.1. Preliminaries. This paper is a continuation of the paper [Sch 1], and contains substantial strengthening of the results of that paper in the case of dimension 2. We refer the reader to [Sch 1] for a discussion of the motivation and background of the problem. For introductions to metastability see, e.g., [GD] and [PL]. The precise results in the current paper can only be stated after enough notation is introduced and are therefore postponed to Section 1.4. We provide next an informal summary of our results.

Our concern in this paper is with the metastable behavior of the 2 dimensional Ising model, evolving with a reversible spin-flip dynamics, in the proximity of the phase-coexistence line. We study the system at an arbitrary subcritical temperature T and under a small positive external magnetic field h. The results proven all refer to limits in which $h \searrow 0$. These results fully confirm, in particular, a conjecture raised by Aizenman and Lebowitz in [AL]. This conjecture was that if started from a typical configuration of the $(-)$-phase, for times of order $\exp(\lambda/h)$ with λ below a critical value λ_c the system would be in a sort of metastable state, close to the $(-)$-phase (in spite of the presence of the positive external field). On the other hand, for a time of order $\exp(\lambda/h)$ with $\lambda > \lambda_c$ the system would have already relaxed and so would be close to the $(+)$-phase. In [Sch 1] a weaker version of this conjecture was proven, with the first scenario occurring for $\lambda < \lambda_1$ and the second for $\lambda > \lambda_2$, with these two constants λ_1 and λ_2 having been explicitly estimated, but both being non-optimal. Moreover the temperature was supposed to be substantially lower than the critical one and the initial distribution had to be concentrated on the configuration with all spins down. On the good side, the results in [Sch 1] are valid in arbitrary dimension.

Here we will only consider dimension 2, but strengthen the result in the following ways:

1) The constants λ_1 and λ_2 are shown to be identical, the common value being given by

$$\lambda_c = \lambda_c(T) = \frac{w(T)^2}{12\, T\, m^*(T)}. \tag{1.1}$$

Here $m^*(T)$ is the spontaneous magnetization and $w(T)$ is the integrated surface tension of the Wulff body of unit volume. Note that all quantities on the right-hand-side of (1.1) pertain to equilibrium statistical mechanics.

2) The (subcritical) temperature T can be arbitrarily close to the critical one.

3) The initial distribution is only required to be stochastically lower than the $(-)$-phase. In particular it can be a Gibbs distribution under any negative value of the external field (as it would if the system were allowed to first relax to equilibrium under a negative external field and then, suddenly, the field was switched to a small positive value).

4) We also show that in a certain technical fashion it makes sense to say that at times of order $\exp(\lambda/h)$, with $\lambda < \lambda_c$ rather than being in the $(-)$-phase the system is better described as being in a metastable state which is infinitesimally (in h) higher than the $(-)$-phase. The rigorous result is presented in the form of an asymptotic expansion in powers of h. The metastable states can then be seen as a family of $\mathcal{C}^\infty$ continuations into the region of positive external fields of the curve of the equilibrium states with negative external fields. (See also the discussion after the statement of the main result of this paper in the Section 1.4.) The result about the $\mathcal{C}^\infty$ continuations is not in conflict with the known fact, proven in [Isa], that at least at low enough temperature, there is no analytic continuation of the equilibrium states beyond the transition point.

Various comments regarding (1) above are in order. To our knowledge, this is the first rigorous relation established between the equilibrium Wulff shape and the time evolution of kinetic Ising models. (The first rigorous relation between the equilibrium surface tension and the time evolution of kinetic Ising models was, as far as we know, established in the fundamental paper [Mar], in the situation in which there is no external field.) It is important to point out that when we started the investigation which led to the current paper we could not see any evident reason for (1.1) to hold. This doubt was expressed, and to some extent discussed, in Section 1-iii of [Sch 1]. Since the doubt stemmed in part from the study of the metastable behavior of anisotropic Ising models in [KO], it is important to stress that regarding the problems treated in the current paper, our results and methods apply also to these models. In section 1.5 we will present a heuristic picture which predicts (1.1) and is based on considering the free-energy of individual droplets of the stable phase in the midst of the metastable phase and taking into account droplet growth at a fast enough speed. The aspect of the heuristics which originally seemed weak to us was the idea that the evolution of these droplets is governed by their equilibrium free-energy.

A recent detailed study of computer simulations of the metastable relaxation of two-dimensional kinetic Ising models in [RTMS], which was done independently of our work in this paper, also indicated the validity of (1.1).

It has become evident over the years that the metastable behavior of kinetic Ising models is very rich and that precise mathematical statements can be conjectured and sometimes proven in various different asymptotic regimes (see, e.g., Section 4 of [Sch 1] or a more complete discussion in [Sch2]). In a recent companion paper, [DS], results were obtained which are counterparts to some of those presented here, but in the case in which the external field is held fixed (positive and small) and the temperature is scaled to zero.

This paper is divided into three parts. In the remainder of the first one, to which this section belongs, we will be introducing notation and terminology, stating results, motivating these results heuristically and presenting some basic tools. In the second part we prove the results concerning the metastable regime, i.e., the behavior at times of order of $\exp(\lambda/h)$ with $0 < \lambda < \lambda_c$. In the third part we prove the results concerning the relaxation regime, i.e., the behavior at times of the order of $\exp(\lambda/h)$ with $\lambda > \lambda_c$.

Acknowledgements: Over the years during which we have worked on this project, we have enjoyed the benefit of several stimulating conversations with various colleagues. We are especially thankful to A. van Enter, R. Kotecký, F. Martinelli, E. Olivieri and E. Scoppola. Part of this work was done while R.H.S. was visiting Rome in the Fall of 1994, and he thankfully acknowledges the warm hospitality of the Physics Department of the University of Rome I and of the Mathematics Departments of the Universities of Rome II and III.

1.2. Notation and terminology. In this section we introduce a long sequence of definitions, notation and techniques. We tried to make everything as standard as possible, so that most readers will browse quickly through this section, finding few things which they are not familiar with. Most statements are made without proof, and we refer readers to the books [Geo] and [Lig], and other references therein, for explanation. Almost all the notation introduced below is identical to that in the papers [Sch 1] and [SS1].

The lattice : The cardinality of a set $\Lambda \subset \mathbb{Z}^2$ will be denoted by $|\Lambda|$. The expression $\Lambda \subset\subset \mathbb{Z}^2$ will mean that Λ is a finite subset of $\mathbb{Z}^2$. For each $x \in \mathbb{Z}^2$, we define the usual norms $\|x\|_p = (|x_1|^p + |x_2|^p)^{1/p}$, $p > 0$ finite, and $\|x\|_\infty = \max\{|x_1|, |x_2|\}$. The distance between two sets $\Lambda_1, \Lambda_2 \in \mathbb{Z}^2$ in each one of these norms will be denoted by

$$\mathrm{dist}_p(\Lambda_1, \Lambda_2) = \inf\{||x - y||_p : x \in \Lambda_1, y \in \Lambda_2\}.$$

In case $\Lambda_1 = \{x\}$, we also write $\mathrm{dist}_p(\Lambda_1, \Lambda_2) = \mathrm{dist}_p(x, \Lambda_2)$. The interior and exterior boundaries of a set $\Lambda \subset \mathbb{Z}^2$ will be denoted, respectively by

$$\partial_{\mathrm{int}}\Lambda = \{x \in \Lambda : \|x - y\|_1 = 1 \text{ for some } y \notin \Lambda\},$$

and

$$\partial_{\mathrm{ext}}\Lambda = \{x \notin \Lambda : \|x - y\|_1 = 1 \text{ for some } y \in \Lambda\}.$$

The p-norm diameter of a set $\Lambda \subset\subset \mathbb{Z}^2$ is defined by

$$\mathrm{diam}_p(\Lambda) = \max\{\mathrm{dist}_p(x, y) : x, y \in \Lambda\}.$$

Given a set $A \subset \mathbb{R}^2$, we will write $\Lambda(A) = A \cap \mathbb{Z}^2$. In case $A = [-l/2, l/2]^2$ is a $l \times l$ square centered at the origin, we simplify the notation to

$$\Lambda(l) = \Lambda(A) = \mathbb{Z}^2 \cap [-l/2, l/2]^2.$$

Given $A, B \subset \mathbb{R}^2$, $z \in \mathbb{R}^2$ and $c \in \mathbb{R}$, we define $A + B = \{x + y : x \in A, y \in B\}$, $A + z = A + \{z\}$, and $cA = \{cx : x \in A\}$.

The set of bonds, i.e., (unordered) pairs of nearest neighbors is defined as

$$\mathbb{B} = \{\{x, y\} : x, y \in \mathbb{Z}^d \text{ and } \|x - y\|_1 = 1\}.$$

Given a set $\Lambda \subset\subset \mathbb{Z}^2$ we define also

$$\mathbb{B}_\Lambda = \{\{x, y\} : x, y \in \Lambda \text{ and } \|x - y\|_1 = 1\},$$

$$\partial\mathbb{B}_\Lambda = \{\{x, y\} : x \in \Lambda, y \notin \Lambda \text{ and } \|x - y\|_1 = 1\}.$$

Notions from percolation : A chain is a sequence of distinct sites $x_1, \ldots, x_n$, with the property that for $i = 1, \ldots, n-1$, $||x_i - x_{i+1}||_1 = 1$. The sites x_1 and x_n are called the end-points of the chain $x_1, \ldots, x_n$, and n is its length. A (*)-chain, its end-points and its length are defined in the same way, but with $|| \cdot ||_1$ replaced by $|| \cdot ||_\infty$. Informally this means that while chains can only move along bonds of $\mathbb{Z}^2$, (*)-chains can also move along diagonals. A set of sites with the property that each two of them can be connected by a chain contained in the set is said to be a connected set. A chain or (*)-chain is said to connect two sets if it has one end-point in each set. A set of sites is said to be simply-connected in case it is connected and its complement is also a connected set. A circuit is a chain such that $||x_1 - x_n||_1 = 1$. Similarly a (*)-circuit is a (*)-chain such that $||x_1 - x_n||_\infty = 1$.

The configurations and observables : At each site in $\mathbb{Z}^2$ there is a spin which can take values -1 and $+1$. The configurations will therefore be elements of the set $\{-1, +1\}^{\mathbb{Z}^d} = \Omega$. Given $\sigma \in \Omega$, we write $\sigma(x)$ for the spin at the site $x \in \mathbb{Z}^2$. Two configurations are specially relevant, the one with all spins -1 and the one with all spins $+1$. We will use the simple notation $-$ and $+$ to denote them. The single spin space, $\{-1, +1\}$ is endowed with the discrete topology and Ω is endowed with the corresponding product topology. The following definition will be important when we introduce finite systems with boundary conditions later on; given $\Lambda \subset\subset \mathbb{Z}^2$ and a configuration $\eta \in \Omega$, we introduce

$$\Omega_{\Lambda,\eta} = \{\sigma \in \Omega : \sigma(x) = \eta(x) \text{ for all } x \notin \Lambda\}.$$

Real-valued functions with domain in Ω are called observables. For each observable f we use the notation $||f||_\infty = \sup_{\eta\in\Omega} |f(\eta)|$. Local observables are those which depend only on the values of finitely many spins, more precisely, $f : \Omega \to \mathbb{R}$ is a local observable if there exists a set $S \subset\subset \mathbb{Z}^2$ such that $f(\sigma) = f(\eta)$ whenever $\sigma(x) = \eta(x)$ for all $x \in S$. The smallest S with this property is called the support of f, denoted Supp(f). The topology introduced above on Ω, has the nice feature that it makes the set of local observables be dense in the set of all continuous observables.

In Ω the following partial order is introduced:

$$\eta \le \zeta \quad \text{if} \quad \eta(x) \le \zeta(x) \text{ for all } x \in \mathbb{Z}^2.$$

A particularly important role will be played in this paper by the non-decreasing local observables. Each local observable can be written as the difference between two non-decreasing ones.

A $(+)$-chain in a configuration σ is a chain of sites, $x_1, \ldots, x_n$, as defined above, with the property that for each $i = 1, \ldots, n$, $\sigma(x_i) = +1$. Given $\Lambda_1, \Lambda_2 \subset \mathbb{Z}^2$, we will use the notation $\{\Lambda_1 \stackrel{+}{\longleftrightarrow} \Lambda_2\}$ to denote the set of configurations in which there is a $(+)$-chain with one end-point in Λ_1 and one end-point in Λ_2. In case $\Lambda_1 = \{x\}$ we simplify this notation to $\{x \stackrel{+}{\longleftrightarrow} \Lambda_2\}$, and similarly for Λ_2. Given a configuration σ, we say that a site x is $(+)$-connected to a set $\Lambda \subset \mathbb{Z}^2$ in σ if $\sigma \in \{x \stackrel{+}{\longleftrightarrow} \Lambda\}$. The $(+)$-cluster of a set $\Lambda \subset \mathbb{Z}^2$ in the configuration σ is the set of sites which are $(+)$-connected to Λ in this configuration. Similar notions can be defined for $(-)$-connectedness, $(+^*)$-connectedness and $(-^*)$-connectedness. In particular the notation $\{\Lambda_1 \stackrel{-}{\longleftrightarrow} \Lambda_2\}$, $\{\Lambda_1 \stackrel{+*}{\longleftrightarrow} \Lambda_2\}$, and $\{\Lambda_1 \stackrel{-*}{\longleftrightarrow} \Lambda_2\}$ should have now self-explanatory meaning.

The probability measures : We endow Ω also with the Borel σ-algebra corresponding to the topology introduced above. In this fashion, each probability measure μ in this space can be identified by the corresponding expected values $\int f d\mu$ of all the local observables f. A sequence of probability measures, $(\mu_n)_{n=1,2,\ldots}$, is said to converge weakly to the probability measure ν in case

$$\lim_{n\to\infty} \int f d\mu_n = \int f d\nu \quad \text{for every continuous observable } f. \tag{1.2}$$

The family of probability measures on Ω will be partially ordered by the following relation: $\mu \leq \nu$ if

$$\int f d\mu \leq \int f d\nu \quad \text{for every continuous non-decreasing observable } f. \tag{1.3}$$

Because the local observables are dense in the set of continuous observables, we can restrict ourselves to the local ones in (1.2) and (1.3). Moreover, because every local observable is the difference between two non-decreasing ones, we can also restrict ourselves to those in (1.2).

The Gibbs measures : We will consider always the formal Hamiltonian.

$$H_h(\sigma) = -\frac{1}{2} \sum_{x,y \text{ n.n.}} \sigma(x)\sigma(y) - \frac{h}{2} \sum_x \sigma(x), \tag{1.4}$$

where $h \in \mathbb{R}$ is the external field and $\sigma \in \Omega$ is a generic configuration. In order to give precise definitions, we define, for each set $\Lambda \subset\subset \mathbb{Z}^2$ and each boundary condition $\eta \in \Omega$,

$$H_{\Lambda,\eta,h}(\sigma) = -\frac{1}{2} \sum_{\{x,y\} \in \mathbb{B}_\Lambda} \sigma(x)\sigma(y) - \frac{1}{2} \sum_{\substack{\{x,y\} \in \partial\mathbb{B}_\Lambda \\ y \notin \Lambda}} \sigma(x)\eta(y) - \frac{h}{2} \sum_{x \in \Lambda} \sigma(x). \tag{1.5}$$

In what follows the temperature T will often appear explicitly in the notation, for clarity. Later on in this paper we will usually be considering a situation in which the temperature is fixed, while we scale the external field h, and then the temperature will be omitted from the notation. Given $\Lambda \subset\subset \mathbb{Z}^2$, $\eta \in \Omega$, $E \subset \Omega$, $T > 0$ and $h \in \mathbb{R}$ we write

$$Z_{\Lambda,\eta,T,h}(E) = \sum_{\sigma \in \Omega_{\Lambda,\eta} \cap E} \exp(-\beta H_{\Lambda,\eta,T,h}(\sigma)),$$

where $\beta = 1/T$. We abbreviate $Z_{\Lambda,\eta,T,h} = Z_{\Lambda,\eta,T,h}(\Omega)$.

The Gibbs (probability) measure in Λ with boundary condition η under external field h and at temperature T is now defined on Ω as

$$\mu_{\Lambda,\eta,T,h}(\sigma) = \begin{cases} \dfrac{\exp(-\beta H_{\Lambda,\eta,h}(\sigma))}{Z_{\Lambda,\eta,T,h}}, & \text{if } \sigma \in \Omega_{\Lambda,\eta}, \\ 0, & \text{otherwise.} \end{cases}$$

The Gibbs measures satisfy the following monotonicity relations to which we will refer as the FKG-Holley inequalities.

If $\eta \le \zeta$ and $h_1 \le h_2$, then, for each $\Lambda \subset\subset \mathbb{Z}^2$ and $T > 0$, $\mu_{\Lambda,\eta,T,h_1} \le \mu_{\Lambda,\zeta,T,h_2}$.

A Gibbs measure for the infinite system on $\mathbb{Z}^2$ is defined as any probability measure μ which satisfies the DLR equations in the sense that for every $\Lambda \subset\subset \mathbb{Z}^2$ and μ-almost all $\eta \in \Omega$

$$\mu(\,\cdot\,|\Omega_{\Lambda,\eta}) = \mu_{\Lambda,\eta,T,h}(\,\cdot\,). \tag{1.6}$$

Alternatively and equivalently, Gibbs measures can be defined as elements of the closed convex hull of the set of weak limit points of sequences of the form $(\mu_{\Lambda_i,\eta_i,h})_{i=1,2,\ldots}$, where each Λ_i is finite and $\Lambda_i \to \mathbb{Z}^d$, as $i \to \infty$, in the sense that $\cup_{i=1}^{\infty} \cap_{j=i}^{\infty} \Lambda_j = \mathbb{Z}^d$.

For each value of T and h, $\mu_{\Lambda(l),-,T,h}$ (*resp.* $\mu_{\Lambda(l),+,T,h}$) converges weakly, as $l \to \infty$, to a probability measure that we will denote by $\mu_{-,T,h}$ (*resp.* $\mu_{+,T,h}$). If $h \neq 0$ it is known that $\mu_{-,T,h} = \mu_{+,T,h}$, which will then be denoted simply by $\mu_{T,h}$; it is also known that this is the only Gibbs measure for the infinite system in this case. If $h = 0$ the same is true if the temperature is larger than or equal to a critical value $T_c > 0$, and is false for $T < T_c$, in which case one says that there is phase coexistence. We use the following abbreviations and names:

$$\mu_{-,T,0} = \mu_{-,T} = \text{ the } (-)\text{-phase},$$

$$\mu_{+,T,0} = \mu_{+,T} = \text{ the } (+)\text{-phase}.$$

Another known fact is that for each fixed T

$$\mu_{T,h} \to \mu_{+,T} \text{ weakly, as } h \searrow 0, \tag{1.7}$$

and

$$\mu_{T,h} \to \mu_{-,T} \text{ weakly, as } h \nearrow 0. \tag{1.8}$$

For the expected value corresponding to a Gibbs measure $\mu_{...}$, in finite or infinite volume, we will use the notation

$$\langle f \rangle_{...} = \int f d\mu_{...},$$

where ... stands for arbitrary subscripts. The corresponding conditional expectation, given the event E will be denoted by $\langle f | E \rangle_{...}$. The spontaneous magnetization at temperature T is defined as

$$m^*(T) = \langle \sigma(0) \rangle_{+,T}.$$

(Here we are using a common and convenient form of abuse of notation: $\sigma(x)$ is being used to denote the observable which associates to each configuration the value of the spin at the site x in that configuration. This notation will also be used in other places.) It is known that $m^*(T) > 0$ if and only if $\mu_{-,T} \neq \mu_{+,T}$ and also that $\lim_{T \searrow 0} m^*(T) = 1$.

Surface tension and Wulff shape: The direction dependent 0-field surface tension is defined in the following way. First consider on $\mathbb{R}^2 \times \mathbb{R}^2$ the usual inner product $(x, y) = x_1 y_1 + x_2 y_2$. Let $\mathbb{S}^1 = \{x \in \mathbb{R}^2 : ||x||_2 = 1\}$, and for each vector $\mathbf{n} \in \mathbb{S}^1$, consider the following configuration, to be used as a boundary condition

$$\eta(\mathbf{n})(x) = \begin{cases} +1, & \text{if } (x, \mathbf{n}) \geq 0, \\ -1, & \text{if } (x, \mathbf{n}) < 0. \end{cases}$$

The surface tension in the direction perpendicular to $\mathbf{n}$ is given by

$$\tau_T(\mathbf{n}) = \lim_{l \to \infty} -\frac{1}{\beta ||y(l) - z(l)||_2} \log \frac{Z_{\Lambda(l),\eta(\mathbf{n}),T,0}}{Z_{\Lambda(l),+,T,0}},$$

where $y(l)$ and $z(l) = -y(l)$ are the points where the straight line $\{x \in \mathbb{R}^2 : (x, \mathbf{n}) = 0\}$ intersects the boundary of the square $\Lambda(l)$. It is known that for each $T < T_c$ the surface tension $\tau_T(\cdot)$ is a continuous strictly positive and finite function.

We shall use $\mathcal{D}$ to denote the set of all closed self-avoiding rectifiable curves $\gamma \subset \mathbb{R}^2$ that are a boundary of a bounded region, $\gamma = \partial V$, $V \subset \mathbb{R}^2$. Let us recall that a curve is called rectifiable if the supremum of the lengths of polygons, with edges connecting arbitrary collections of points chosen on the curve, in the order inherited from the curve, is finite (and equals then the length, $|\gamma|$, of the curve γ), and that a rectifiable curve has a tangent at almost every point. It is easy to verify that a curve γ that is the boundary of a convex bounded region belongs to $\mathcal{D}$. We can assign to each curve $\gamma \in \mathcal{D}$ the quantity

$$\mathcal{W}(\gamma) = \mathcal{W}_T(\gamma) = \int_\gamma \tau_T(\mathbf{n}_s) ds,$$

where s parametrizes the curve γ according to Euclidean length measured along this curve, and $\mathbf{n}_s$ is the unit outward normal vector to the curve at the point $s \in \gamma$ (i.e.

the vector orthogonal to the tangent in the considered point and oriented outward the region bounded by γ). The functional $\mathcal{W}_T$ will be called the Wulff functional associated to the zero-field direction-dependent surface tension $\tau_T(\cdot)$. Sometimes we will refer to it also as the integrated surface tension.

To every vector $\mathbf{n} \in \mathbb{S}^1$ and $\lambda > 0$ we assign the half-plane

$$L_{T,\mathbf{n},\lambda} = \left\{ x \in \mathbb{R}^2 : (x, \mathbf{n}) \leq \lambda \tau_T(\mathbf{n}) \right\}.$$

Let us consider the intersection

$$W_{T,\lambda} = \bigcap_{\mathbf{n} \in \mathbb{S}^1} L_{T,\mathbf{n},\lambda}. \tag{1.9}$$

These sets clearly satisfy the scaling relation $W_{T,\lambda} = \lambda W_{T,1}$. In particular they keep the same shape, as λ varies; this shape is called the Wulff shape. The Wulff body of volume 1 is defined as $W_T = W_{T,\lambda_0}$, where λ_0 is chosen so that its volume is indeed 1. W_T is clearly convex and thus its boundary $\partial W_T \in \mathcal{D}$. The following is therefore well defined,

$$w = w(T) = \mathcal{W}_T(\partial W_T).$$

For each $T < T_c$, the boundary of the Wulff body satisfies the following variational principle. For all $\gamma \in \mathcal{D}$ which are boundaries of regions of volume 1,

$$w(T) \leq \mathcal{W}_T(\gamma), \tag{1.10}$$

with equality only in case γ is a translation of ∂W_T.

The dynamics : We introduce now for the Ising model above, the type of time evolution which makes it into what is known as kinetic Ising model, or stochastic Ising model, or dynamic Ising model or Glauber dynamics. First we recall that a spin flip system is defined as a Markov process on the state space Ω, whose generator, L, acts on a generic local observable f as

$$(Lf)(\sigma) = \sum_{x \in \mathbb{Z}^d} c(x, \sigma)(f(\sigma^x) - f(\sigma)), \tag{1.11}$$

where σ^x is the configuration obtained from σ by flipping the spin at the site x, and $c(x, \sigma)$ is called the rate of flip of the spin at the site x when the system is in the state σ. In order for this generator to be well defined and indeed generate a unique Markov process, one has to assume that the rates $c(x, \sigma)$ satisfy certain regularity conditions. For our purposes here, we will actually restrict ourselves to the following conditions, which are more than enough to assure the existence and uniqueness of the process.

(H1) (*Translation invariance*) For every $x, y \in \mathbb{Z}^d$,

$$c(x, \sigma) = c(x + y, \theta_y \sigma),$$

where $\theta_y \sigma$ is the configuration obtained by shifting σ by y, i.e., $(\theta_y \sigma)(z) = \sigma(z - y)$.

(H2) (*Finite range*) There exists R such that

$$c(0,\eta) = c(0,\zeta) \quad \text{if} \quad \eta(x) = \zeta(x) \text{ whenever } \|x\|_\infty \le R.$$

The minimal such R is called the range of the interaction.

The connection between the rates of flip and the Hamiltonian (1.4) and the temperature $T = 1/\beta$ is established by imposing conditions which assure us that the Gibbs measures are not only invariant, but also reversible with respect to the dynamics. These conditions, called detailed balance, state that for each $x \in \mathbb{Z}^d$ and $\sigma \in \Omega$,

$$c(x,\sigma) = c(x,\sigma^x)\exp(-\beta\Delta_x H_h(\sigma)), \tag{1.12}$$

where

$$\Delta_x H_h(\sigma) = \sigma(x)\left(\sum_{y:\{x,y\}\in\mathbb{B}_\Lambda} \sigma(y) + h\right),$$

which formally equals $H_h(\sigma^x) - H_h(\sigma)$. We will usually make the dependence on h explicit, by writing $c_h(x,\sigma)$ for the rates. There are many examples of rates which satisfy the conditions of detailed balance (1.12) and also the other hypotheses, H(1) and H(2). The most common examples found in the literature are:

Example 1) *Metropolis Dynamics*

$$c_h(x,\sigma) = \exp(-\beta(\Delta_x H_h(\sigma))^+),$$

where $(a)^+ = \max\{a,0\}$ is the positive part of a.

Example 2) *Heat Bath Dynamics*

$$c_h(x,\sigma) = \frac{1}{1+\exp(\beta\Delta_x H_h(\sigma))}.$$

Example 3)

$$c_h(x,\sigma) = \exp\left(-\frac{\beta}{2}\Delta_x H_h(\sigma)\right).$$

Each one of these rates satisfies also the further conditions below which will be needed for the analysis in this paper to be possible.

(H3) (*Attractiveness and monotonicity in h*) If $\eta(x) \le \zeta(x)$ and $h_1 \le h_2$, then

$$\begin{aligned} c_{h_1}(x,\eta) &\le c_{h_2}(x,\zeta) \quad \text{if} \quad \eta(x) = \zeta(x) = -1, \\ c_{h_1}(x,\eta) &\ge c_{h_2}(x,\zeta) \quad \text{if} \quad \eta(x) = \zeta(x) = +1. \end{aligned}$$

(H4) (*Uniform boundedness of rates*) For each temperature T there is $h(T) > 0$ and $0 < c_{\min}(T) \le c_{\max}(T) < \infty$ such that for all $h \in (-h(T), h(T))$ and $\sigma \in \Omega$

$$c_{\min}(T) \le c_h(0,\sigma) \le c_{\max}(T).$$

Throughout this paper we will suppose that we have chosen and kept fixed a set of rates $c_h(x,\sigma)$ which satisfy the detailed balance conditions, (1.12) and all the hypotheses H(1) - H(4). This spin flip system will be denoted by $(\sigma^\eta_{h;t})_{t\geq 0}$, where η is the initial configuration. If this initial configuration is selected at random according to a probability measure ν, then the resulting process is denoted by $(\sigma^\nu_{h;t})_{t\geq 0}$. The probability measure on the space of trajectories of the process will be denoted by $\mathbb{P}$, and the corresponding expectation by $\mathbb{E}$. (Later, when we couple various related processes, we will also use the symbols $\mathbb{P}$ and $\mathbb{E}$ to denote probabilities and expectations in some larger probability spaces, but no confusion should arise from this.) The assumption of detailed balance, (1.12), assures that the Gibbs measures are invariant with respect to the stochastic Ising models. Moreover, from the assumption of attractiveness, H(3), one obtains the following convergence results

$$\sigma^-_{h;t} \to \mu_{-,T,h},$$

and

$$\sigma^+_{h;t} \to \mu_{+,T,h},$$

weakly, as $t\to\infty$.

We will want to consider, sometimes as a tool, and sometimes for its own sake, the counterpart of the stochastic Ising model that we are considering, on an arbitrary finite set $\Lambda \subset\subset \mathbb{Z}^2$, with some boundary condition $\xi\in\Omega$. This process, which will be denoted by $(\sigma^\eta_{\Lambda,\xi,h;t})_{t\geq 0}$, where $\eta\in\Omega_{\Lambda,\xi}$ is the initial configuration, is defined as the spin flip system with rates of flip given by

$$c_{\Lambda,\xi,h}(x,\sigma) = \begin{cases} c_h(x,\sigma) & \text{if } \sigma,\sigma^x \in \Omega_{\Lambda,h}, \\ 0 & \text{otherwise.}\end{cases}$$

When $\sigma,\sigma^x\in\Omega_{\Lambda,h,}$, (1.12) yields, for all $x\in\mathbb{Z}^d$,

$$\mu_{\Lambda,\xi,h}(\sigma)c_{\Lambda,\xi,h}(x,\sigma) = \mu_{\Lambda,\xi,h}(\sigma^x)c_{\Lambda,\xi,h}(x,\sigma^x), \tag{1.13}$$

which is the usual reversibility condition for finite state-space Markov processes. (Conversely, if one requires (1.13) to be satisfied for arbitrary $\Lambda\in\mathcal{F}$ and $\xi\in\Omega$, then one can deduce that (1.12) must hold.) It is clear from H(4) that $(\sigma^\eta_{\Lambda,\xi,h;t})$ is irreducible and hence from (1.13) it follows that, for any η,

$$\sigma^\eta_{\Lambda,\xi,h;t} \to \mu_{\Lambda,\xi,T,h},$$

weakly, as $t\to\infty$.

Graphical construction : In order to prove our claims in this paper, we will use a standard graphical construction which provides versions of the whole family of processes at a given temperature T, with arbitrary value of $h\in(-h(T),h(T))$, either on the infinite lattice $\mathbb{Z}^2$ or on any of its finite subsets, with arbitrary boundary conditions and starting from any initial configuration, all on the same probability space. This construction is the same one used in [Sch 1]. But the relevance of this construction will be even greater here than it was in that paper, since in part 3 of our paper we will use it to define the process on regions of space-time which are fairly general, and we will set up a rescaling procedure based on such objects.

The graphical construction that we use is a specific version of what is called basic coupling between spin flip processes: a coupling in which the spins flip together as much as possible, considering the constraint that they have to flip with certain rates. The construction is carried out by first associating to each site $x \in \mathbb{Z}^2$ two independent Poisson processes, each one with rate $c_{\max}(T)$. We will denote the successive arrival times (after time 0) of these Poisson processes $(\tau^+_{x,n})_{n=1,2,\dots}$ and $(\tau^-_{x,n})_{n=1,2,\dots}$. Assume that the Poisson processes associated to different sites are also mutually independent. We say that at each point in space-time of the form $(x, \tau^+_{x,n})$ there is an upward mark and that at each point of the form $(x, \tau^-_{x,n})$ there is a downward mark. Next we associate to each arrival time $\tau^*_{x,n}$, where $*$ stands for $+$ or $-$, a random variable $U^*_{x,n}$ with uniform distribution between 0 and 1. All these random variables are supposed to be independent among themselves and independent from the previously introduced Poisson processes. This finishes the construction of the probability space. The corresponding probability and expectation will be denoted, respectively, by $\mathbb{P}$ and $\mathbb{E}$. We have to say now how the various processes are constructed on this probability space. For finite Λ and arbitrary ξ, the process $(\sigma^\eta_{\Lambda,\xi,h;t})$ is constructed as follows. We know that almost surely the random times $\tau^*_{x,n}$, $x \in \Lambda$, $n = 1, 2, \dots$, $* = +, -$, are all distinct, and we update the state of the process at each time when there is a mark at some $x \in \Lambda$ according to the following rules. If the mark that we are considering is at the point $(x, \tau^*_{x,n})$, and the configuration immediately before time $\tau^*_{x,n}$ was σ, then

i) The spins not at x do not change.

ii) If $\sigma(x) = -1$ (*resp.* $\sigma(x) = +1$), then the spin at x can only flip if the mark is of upward type (*resp.* downward type).

iii) If the mark is upward and $\sigma(x) = -1$, or if the mark is downward and $\sigma(x) = +1$, then we flip the spin at x if and only if $c_{\Lambda,\xi,h}(x, \sigma) > U^*_{x,n} c_{\max}$.

One can readily see that the process constructed in this fashion has the correct rates of flip.

In principle, one would like to construct the processes on the infinite lattice $\mathbb{Z}^2$ in a similar fashion, with $c_h(x, \sigma)$ replacing $c_{\Lambda,\xi,h}(x, \sigma)$ in (iii). Some extra care has to be taken, because during any non-degenerate interval of time infinitely many marks occur. This is not a real problem, because of the assumption that the range of the interaction, R, is finite. Starting from a configuration η at time 0, we have to say how the spin at a generic site x at a time t is obtained. Using percolation arguments one can argue that on a set of probability 1 in the probability space where the marks were defined, for any fixed x and t, if we take any boundary condition ξ, then the sequence $(\sigma^\eta_{\Lambda(l),\xi,h;t}(x))_{l=1,2,\dots}$ will converge as $l \to \infty$ (i.e., will become constant for large l), to a limit which does not depend on ξ. This limit can then be taken to be the value of $\sigma^\eta_{h;t}(x)$, and it is clear that the process thus constructed has the correct flip rates and is therefore a version of the process $(\sigma^\eta_{h;t})$. The expected value of a function of that process will be denoted by $\langle\cdot\rangle^\eta_{h;t}$. A standard proof of the claim above about insensitivity to receding boundary conditions can be found in [Sch 1]. Those estimates presented in [Sch 1] show also that even if we let t grow with l, but keeping l/t large enough, then the spin at a fixed site x is almost insensitive up to time t to what happens outside of the box $\Lambda(l)$. We state this result in the form of a lemma for future reference. This lemma is a rigorous counterpart to the informal statement that because of the finite range of the interaction and of the uniform upper bound on the rates of flip, "the effects travel with a bounded speed".

Lemma 1.2.1. *For each temperature T, there exists a finite positive constant $C(T)$ such that if $l \geq C(T)t$, then there exist a positive function $C_1(x), x \in \mathbb{Z}^2$ and a positive constant C_2, such that for every site $x \in \mathbb{Z}^2$,*

$$\sup_{h\in(-h(T),h(T))} \sup_{\xi\in\Omega} \sup_{\eta\in\Omega} \mathbb{P}(\sigma^{\eta}_{h;t}(x) \neq \sigma^{\eta}_{\Lambda(l),\xi,h;t}(x)) \leq C_1(x)\exp\{-C_2 l\}.$$

Proof. See [Sch 1].

Because of the hypotheses (H3), of attractiveness and monotonicity in h, the coupling provided by the construction above preserves the order between the coupled marginal processes, in various cases. In this paper we will need the following facts. If $\eta \leq \zeta$, $\xi \leq \xi'$, $-h(T) < h_1 \leq h_2 < h(T)$ and $\Lambda \subset\subset \mathbb{Z}^2$ is arbitrary, then for all $t \geq 0$,

$$\sigma^{\eta}_{\Lambda,\xi,h_1;t} \leq \sigma^{\zeta}_{\Lambda,\xi',h_2;t}, \tag{1.14}$$

$$\sigma^{\eta}_{h_1;t} \leq \sigma^{\zeta}_{h_2;t}, \tag{1.15}$$

and

$$\sigma^{\eta}_{\Lambda,-,h_1;t} \leq \sigma^{\zeta}_{h_2;t}. \tag{1.16}$$

We will refer to these inequalities as basic-coupling inequalities. (Observe that the FKG-Holley inequalities for the models we are considering can be derived from (1.14).)

We will sometimes have to enlarge the probability space defined by the graphical construction to accommodate a random choice of the initial configuration, according to some distribution ν, performed independently of the subsequent time evolution of the process. In this case we will replace the initial configuration with ν in the notation for the process, but, in a slight abuse of notation, we will still use $\mathbb{P}$ and $\mathbb{E}$ for probabilities and expectations in this enlarged probability space.

A few more remarks on notation and conventions: We will use C, C_1, C_2,C',$C(T)$, etc... to denote positive finite constants, whose precise values are not relevant and may even change from appearance to appearance.

We will omit the temperature T in most of the notation, since it is fixed.

1.3. Some tools and further definitions.

Before we can state our main results, we need to introduce a few more concepts. This will by done in this section, intermingled with the presentation of some basic techniques.

A fundamental fact that we will use often is that if $T < T_c$ then there is a finite positive constant $C(T)$ such that for all $h \geq 0$, all $\Lambda \subset \mathbb{Z}^2$ and all $x, y \in \mathbb{Z}^2$,

$$\mu_{\Lambda,+,h}\left(x \overset{-*}{\longleftrightarrow} y\right) \leq \mu_{+,h}\left(x \overset{-*}{\longleftrightarrow} y\right) \leq \mu_{+}\left(x \overset{-*}{\longleftrightarrow} y\right) \leq \exp(-C(T)||x-y||_\infty). \tag{1.17}$$

The first two inequalities above are instances of the FKG-Holley inequalities (and hold, of course, for all T), and the third one is Theorem 1 in [CCS].

We review next some well known ways to exploit these inequalities in combination with the FKG-Holley inequalities and the Markov property of the Gibbs measures. Until we say otherwise, we will be supposing that f is an increasing local observable. We start with the observation that if ν_1 and ν_2 are two probability distributions on Ω, and the event $E \subset \Omega$ is such that $\nu_1(\,\cdot\,) \le \nu_2(\,\cdot\,|E^c)$ then

$$\begin{aligned}\int f d\nu_1 - \int f d\nu_2 &= \int f d\nu_1 - \nu_2(E) \int f d\nu_2(\,\cdot\,|E) - \nu_2(E^c) \int f d\nu_2(\,\cdot\,|E^c) \\ &= \int f d\nu_1 - \int f d\nu_2(\,\cdot\,|E^c) \\ &\quad + \nu_2(E) \left(- \int f d\nu_2(\,\cdot\,|E) + \int f d\nu_2(\,\cdot\,|E^c) \right) \\ &\le 2\ ||f||_\infty\ \nu_2(E). \end{aligned} \tag{1.18}$$

Suppose that $\Lambda \subset \Lambda' \subset\subset \mathbb{Z}^2$ and $\eta \in \Omega$. By partitioning the complement of the event $E = \left\{ \mathrm{Supp}(f) \stackrel{-}{\longleftrightarrow} \Lambda^c \right\}$ according to what the $(-)$-cluster of the set Λ^c is, and using the FKG-Holley inequalities and the Markov property of the Gibbs distributions, one obtains

$$\mu_{\Lambda',\eta,h}(\,\cdot\,|E^c) \ge \mu_{\Lambda,+,h}(\,\cdot\,). \tag{1.19}$$

Therefore, from the FKG-Holley inequalities and (1.18) we obtain for any $\eta \in \Omega$

$$0 \le \langle f \rangle_{\Lambda,+,h} - \langle f \rangle_{\Lambda',\eta,h} \le 2\ ||f||_\infty\ \mu_{\Lambda',\eta,h} \left(\mathrm{Supp}(f) \stackrel{-}{\longleftrightarrow} \Lambda^c \right). \tag{1.20}$$

If we let now $\Lambda' \to \mathbb{Z}^2$ and suppose that $\mu_{\Lambda',\eta,h} \to \mu$ weakly for some distribution μ, we obtain

$$0 \le \langle f \rangle_{\Lambda,+,h} - \int f d\mu \le 2\ ||f||_\infty\ \mu \left(\mathrm{Supp}(f) \stackrel{-}{\longleftrightarrow} \Lambda^c \right). \tag{1.21}$$

By taking $\eta = +$ and combining (1.21) with (1.17) we obtain, when $T < T_c$ and $h \ge 0$,

$$0 \le \langle f \rangle_{\Lambda,+,h} - \langle f \rangle_{+,h} \le C(f) \exp\Big(- C(T)\ \mathrm{dist}_\infty(\mathrm{Supp}(f), \Lambda^c) \Big). \tag{1.22}$$

Next we consider the correlation functions

$$\langle f; g \rangle_{\Lambda,+,h} = \langle fg \rangle_{\Lambda,+,h} - \langle f \rangle_{\Lambda,+,h} \langle g \rangle_{\Lambda,+,h},$$

where f and g are two local observables, not necessarily increasing. We claim that there is a finite positive numerical constant C such that for each T and h

$$|\langle f; g \rangle_{\Lambda,+,h}| \le C\ ||f||_\infty\ ||g||_\infty\ \mu_{\Lambda,+,h} \left(\mathrm{Supp}(f) \stackrel{-}{\longleftrightarrow} \mathrm{Supp}(g) \right). \tag{1.23}$$

There is no loss in generality in supposing that f and g are increasing and have $||f||_\infty \le 1$, $||g||_\infty \le 1$. Given $\eta \in \{-1,+1\}^{\mathrm{Supp}(g)}$, let E_η be the event that the configuration restricted to Supp(g) is identical to η, and let g_η be the value that g assumes on E_η. Note now that using first the FKG-Holley inequalities and then (1.20), we obtain

$$\begin{aligned}\langle fg\rangle_{\Lambda,+,h} &= \sum_{\eta\in\{-1,+1\}^{\mathrm{Supp}(g)}} \langle f|E_\eta\rangle_{\Lambda,+,h}\; g_\eta\; \mu_{\Lambda,+,h}(E_\eta)\\ &\le \langle f\rangle_{\Lambda\setminus\mathrm{Supp}(g),+,h}\;\langle g\rangle_{\Lambda,+,h}\\ &\le \Big[\langle f\rangle_{\Lambda,+,h} + 2\mu_{\Lambda,+,h}\Big(\mathrm{Supp}(f)\overset{-}{\longleftrightarrow}\mathrm{Supp}(g)\Big)\Big]\;\langle g\rangle_{\Lambda,+,h}.\end{aligned}$$

This immediately implies (1.23).

In case $T<T_c$ and $h\ge 0$, we can combine (1.23) with (1.17) to obtain

$$|\langle f;g\rangle_{\Lambda,+,h}| \le C(f,g)\exp\Big(-C(T)\ \mathrm{dist}_\infty(\mathrm{Supp}(f),\mathrm{Supp}(g))\Big), \tag{1.24}$$

an estimate that in particular is uniform in $h\ge 0$ and in Λ.

A consequence of the exponential decay of correlations in (1.24) is that when $T<T_c$ the function $\langle f\rangle_{+,h}$ of $h\ge 0$, which to $h=0$ associates $\langle f\rangle_+$ and to $h>0$ associates $\langle f\rangle_h$, is infinitely differentiable at $h=0$. Moreover, for $j=1,2,\ldots$, the following identity holds:

$$\left.\frac{d^j\langle f\rangle_{+,h}}{dh^j}\right|_{h=0_+} = \left(\frac{\beta}{2}\right)^j \sum_{x_1,\ldots,x_j\in\mathbb{Z}^2} \langle f;\sigma(x_1);\ldots;\sigma(x_j)\rangle_+. \tag{1.25}$$

In this expression, which will be justified below, the quantities that appear inside of the summation are called generalized Ursell functions, and are defined next. We start by defining the generalized Ursell functions for a Gibbs measure $\mu_{\Lambda,\eta,h}(\,\cdot\,)$, where Λ is a finite set. For this purpose we consider a generalization of the Hamiltonian, in which at each site x the external applied field may be different, and takes the value h_x. If $\underline{h}$ is the function that to each $x\in\mathbb{Z}^2$ associates h_x, then we will denote by $\mu_{\Lambda,\eta,\underline{h}}$ the corresponding Gibbs distribution in Λ with boundary condition η. With a local observable f given we define

$$\langle f;\sigma(x_1);\ldots;\sigma(x_j)\rangle_{\Lambda,\eta,h} = \left(\frac{2}{\beta}\right)^j \left.\frac{\partial^j\ \langle f\rangle_{\Lambda,\eta,\underline{h}}}{\partial h_{x_1}\ldots\partial h_{x_j}}\right|_{\underline{h}\equiv h},$$

where $\underline{h}\equiv h$ means that the function $\underline{h}$ is identically h. It is easy to see, by induction on j, that $\langle f;\sigma(x_1);\ldots;\sigma(x_j)\rangle_{\Lambda,\eta,h}$ is a linear combination of products of $\mu_{\Lambda,\eta,h}$-expected values of local observables, all of them with support contained in Supp$(f)\cup\{x_1,\ldots,x_j\}$. Convergence of the generalized Ursell functions as $\Lambda\to\mathbb{Z}^2$ follows from this, for arbitrary T, h and η, as long as $\mu_{\Lambda,\eta,h}$ converges weakly to some distribution. In case $T<T_c$, $h\ge 0$ and $\eta=+$, we can do better and obtain from (1.22) the bound

$$\begin{aligned}|\langle f;\sigma(x_1);\ldots;\sigma(x_j)\rangle_{\Lambda,+,h} &- \langle f;\sigma(x_1);\ldots;\sigma(x_j)\rangle_{+,h}|\\ &\le C_j(f)\exp\Big(-C(T)\ \mathrm{dist}_\infty(\mathrm{Supp}(f)\cup\{x_1,\ldots,x_j\},\Lambda^c)\Big).\end{aligned} \tag{1.26}$$

Observe that in particular this estimate is uniform in $h \geq 0$.

The proof of (1.25) can be found in Section 2 of [M-L]. In that paper the result was proven at low enough temperature, but this was so simply because the estimate (1.24), which is a consequence of (1.17) was not available then. Replacing Theorem 1 in [M-L] with (1.24), the proof in that paper applies up to T_c. The basic estimate used in [M-L], and to which we will return in section 2.4, states that the exponential decay of correlations (1.24) implies a similar exponential decay for the generalized Ursell functions, when the diameter of the set $\mathrm{Supp}(f) \cup \{x_1, ..., x_j\}$ becomes large:

$$|\langle f; \sigma(x_1); ...; \sigma(x_j)\rangle_{\Lambda,+,h}| \leq C_j(T, f) \exp\left(-C(T)\frac{\mathrm{diam}_\infty(\mathrm{Supp}(f) \cup \{x_1, ..., x_j\})}{j}\right). \tag{1.27}$$

Observe that in particular this estimate is uniform in $h \geq 0$ and in Λ. The proof that (1.27) follows from (1.24) in the special case in which f is of the form $f(\sigma) = \sigma(y_1) \cdots \sigma(y_m)$ for some set of sites $\{y_1, ..., y_m\}$ can be found in the Appendix B of [M-L]. The general case follows immediately from this one, since any local function f can be written as a linear combination of functions of this special form, with $\{y_1, ..., y_m\}$ running over all the subsets of $\mathrm{Supp}(f)$.

As explained in the proof of Theorem 4 in [M-L], (1.25) follows in a standard fashion from (1.27).

In connection to (1.25) it is worth mentioning that for $h > 0$ the function $\langle f\rangle_h$ is analytic, as follows from the Lee-Yang theorem. Moreover, in this case the Gibbs distributions are completely analytic in an appropriate sense (see [SS1] and references therein). The estimate (1.27) can be replaced for $h > 0$ by a stronger one:

$$|\langle f; \sigma(x_1); ...; \sigma(x_j)\rangle_{\Lambda,+,h}| \leq C'_j(T, h, f) \exp\left(-C(T, h)\ \mathrm{dist}_{\mathrm{tree}}(\mathrm{Supp}(f), x_1, ..., x_j)\right),$$

where $\mathrm{dist}_{\mathrm{tree}}(\Lambda_1, ..., \Lambda_j)$ is the length of the shortest tree in $\mathbb{R}^2$, connecting all the sets $\Lambda_1, ..., \Lambda_j \subset \mathbb{Z}^2$. However, the constants $C'_j(T, h, f)$ explode as $h \to 0$. It is also important to recall that, on the other hand, it has been proven in [Isa] that at low enough temperatures there is an essential singularity nevertheless at $h = 0$; this is expected to be so up to T_c, but no proof of that claim is available, as far as we know.

The identity (1.25) and the various related statements that we made above have, of course, analogues for $h \leq 0$. Those are the ones that will be relevant for us when we study metastability under small $h > 0$, since the "metastable state" should then be a "continuation" of the equilibrium states with $h \leq 0$.

1.4. Main results. Recall that we are considering a kinetic Ising model for the formal Hamiltonian (1.4) in dimension 2, which is supposed to satisfy conditions (H1), (H2), (H3) and (H4) of Section 1.2. Recall also that for $T < T_c$ we define $\lambda_c = \lambda_c(T) = w(T)^2/(12\, T\, m^*(T))$. The following theorem is our main result.

Theorem 1. *Suppose $T < T_c$. For every probability distribution $\nu \leq \mu_-$ the following happens.*

i) If $0 < \lambda < \lambda_c$, then for each $n \in \{1, 2, ...\}$ and for each local observable f,

$$\mathbb{E}\left(f\left(\sigma^{\nu}_{h;\exp(\lambda/h)}\right)\right) = \sum_{j=0}^{n-1} b_j(f)h^j + O(h^n)$$

for $h > 0$, where

$$b_j(f) = \frac{1}{j!}\frac{d^j\langle f\rangle_{-,h}}{dh^j}\bigg|_{h=0_-} = \frac{1}{j!}\left(\frac{\beta}{2}\right)^j \sum_{x_1,\dots,x_j\in\mathbb{Z}^2}\langle f;\sigma(x_1);...;\sigma(x_j)\rangle_-,$$

and $O(h^n)$ is a function of f and h which satisfies $\limsup_{h\searrow 0}|O(h^n)|/h^n < \infty$.

ii) If $\lambda > \lambda_c$, then for any finite positive C there is a finite positive C_1 such that for every local observable f,

$$\left|\mathbb{E}\left(f\left(\sigma^{\nu}_{h;\exp(\lambda/h)}\right)\right) - \langle f\rangle_h\right| \;\le\; C_1\,||f||_\infty\,\exp\left(-\frac{C}{h}\right),$$

for all $h > 0$.

From this theorem, the simple Proposition 1 in [Sch 1], and the fact that $\langle f\rangle_h \to \langle f\rangle_+$ as $h \searrow 0$ (see (1.7), or the paragraph which precedes (1.25)), the following corollary is obtained.

Corollary 1. *Suppose $T < T_c$. For every probability distribution $\nu \le \mu_-$ the following happens. If we let $h \searrow 0$ and $t \to \infty$ together, then for every local observable f*
i) $\mathbb{E}\left(f\left(\sigma^{\nu}_{h;t}\right)\right) \to \langle f\rangle_-$ *if* $\limsup h\log t < \lambda_c(T)$.
ii) $\mathbb{E}\left(f\left(\sigma^{\nu}_{h;t}\right)\right) \to \langle f\rangle_+$ *if* $\liminf h\log t > \lambda_c(T)$.

In other words, we are stating that the law of the random configuration $\sigma^{\nu}_{h;t}$ converges weakly to μ_- in case (i) and to μ_+ in case (ii). This corollary is already an important strengthening of Theorem 1 in [Sch 1] in the 2 dimensional case. The following aspects of that theorem are improved here: 1) There is a single constant λ_c separating the regimes (i) and (ii). 2) The temperature is now only required to be below T_c. 3) The initial distribution is much more general than in [Sch 1], where it was supposed to be concentrated on the configuration with all spins down. Note that, by the FKG-Holley inequalities, for each $h < 0$ the distribution μ_h satisfies the condition above on the initial distribution ν.

To illustrate and clarify the way in which Theorem 1(i) improves even further the statement in Corollary 1(i), let us take the local observable given by $f(\sigma) = \sigma(0)$ and $n = 2$. We have then, when $0 < \lambda < \lambda_c$

$$\mathbb{E}\left(\sigma^{\nu}_{h;\exp(\lambda/h)}(0)\right) = -m^* + \chi h + O(h^2),$$

when $h > 0$. Here

$$\chi = b_1(f) = \frac{d\langle\sigma(0)\rangle_{-,h}}{dh}\bigg|_{h=0_-} = \left(\frac{\beta}{2}\right)\sum_{x\in\mathbb{Z}^2}\langle\sigma(0);\sigma(x)\rangle_-,$$

is the susceptibility at $h = 0_-$. This means that when $h > 0$ is small the function $-m^* + \chi h$ is a better approximation to $\mathbb{E}\left(\sigma^{\nu}_{h;\exp(\lambda/h)}(0)\right)$ than the constant function identical to $-m^* = \langle f \rangle_-$. This function $-m^* + \chi h$ is the smooth linear continuation into the region $h \geq 0$ of the function which to $h < 0$ associates the equilibrium expectation $\langle f \rangle_h$. Similar interpretations can be given for larger values of n and arbitrary f.

Another way to express part of the content of the Theorem 1 is to observe that it claims that for any λ, $0 < \lambda < \lambda_c$ and any probability distribution $\nu \leq \mu_-$ the branch of states $\langle \cdot \rangle^{\nu}_{h;\exp(\lambda/h)}$ for $h > 0$ is a $\mathcal{C}^\infty$ continuation of the family $\langle \cdot \rangle_h$ for $h < 0$. That interpretation suggests that the phenomenon of metastability should be understood dynamically, in which case the physically meaningful smooth continuations through the critical point $h = 0$ become possible.

In the Physics literature (see, e.g., [BM]), one sometimes relates the metastable relaxation of a system to the presence of a "plateau" in the graph corresponding to the time evolution of a quantity of the type of $\mathbb{E}\left(f\left(\sigma^{\nu}_{h;t}\right)\right)$. Of course, strictly speaking there is no "plateau", and generically the slope of such a function is never 0. Still, from the experimental point of view a rough "plateau" can be seen and described as follows. In a relatively short time $\mathbb{E}\left(f\left(\sigma^{\nu}_{h;t}\right)\right)$ seems to converge to a value close to $\langle f \rangle_-$; after this, one sees an apparent flatness in the relaxation curve over a period of time which may be quite long compared with the time needed to first approach this value. But eventually the relaxation curve starts to deviate from this almost constant value and move towards the true asymptotic limit, close to $\langle f \rangle_+$. The experimentally almost flat portion of the relaxation curve is referred to as a "plateau". Theorem 1 can be seen to some extent as giving some precise meaning to such a "plateau", and we discuss now two ways in which this can be done. First note that if $0 < \lambda' < \lambda'' < \lambda_c$, then from Part (i) of the Theorem we have

$$\mathbb{E}\left(f\left(\sigma^{\nu}_{h;\exp(\lambda'/h)}\right)\right) - \mathbb{E}\left(f\left(\sigma^{\nu}_{h;\exp(\lambda''/h)}\right)\right) \to 0,$$

faster than any power of h. Observe that we are considering times which are of different order of magnitudes, when h is small, and still we are observing a nearly constant $\mathbb{E}\left(f\left(\sigma^{\nu}_{h;t}\right)\right)$. For a second way in which Theorem 1 can be seen as expressing the presence of a "plateau", we can think of plotting $\mathbb{E}\left(f\left(\sigma^{\nu}_{h;t}\right)\right)$ versus $\log(t)$, rather than versus t. This is somewhat the natural graph to consider, if one is interested in the order of magnitude of the relaxation time. If the $\log(t)$-axis is drawn in the proper scale, amounting to replacing it with $h\log(t)$, then, when h is small, Theorem 1 tells us that the graph should be close to that of a step function which jumps at the point λ_c, from the value $\langle f \rangle_-$ to the value $\langle f \rangle_+$.

Readers who are familiar with [Sch 1] and [Sch2] can expect that also Theorem 4 and Corollary 1 in [Sch 1], which refer to finite systems with $(-)$ boundary conditions and sizes which are scaled as $h \searrow 0$, have stronger versions along the lines of the current paper. This is indeed the case, but for brevity we will omit the statements of these theorems, which can easily be obtained and can be proved with the techniques introduced in this paper.

1.5 Heuristics. One of the appealing features of the results proven in this paper is that some of them can be correctly predicted based on a very simple and naive-looking heuristics. It is probably a challenge to a historian of science to trace back the origin and evolution of this non-rigorous approach to the problem, to give the proper credit to the people involved and to elucidate how the interplay among empirical observation of metastable systems, theoretical analysis and computer simulation led to the reasoning described in a simple fashion below. Here we will make no attempt to clarify the history of the subject. The interested reader will find a great deal of references to the earlier literature on the subject in the paper [RTMS]. It is worth stressing that certain parts of the heuristics were rediscovered more than once in different forms and contexts, so that giving proper credit is a very difficult task.

The reader may want to compare the heuristics present below with the one presented in [Sch1] and [Sch2]. The main difference is that here we are being more ambitious by basing the heuristics on the computation of the free-energy of a droplet with an arbitrary shape, rather than on the energy of a square droplet. At the time that those two former papers were written, it seemed to its author that there was no compelling reason to believe that the equilibrium free-energy of droplets would predict correctly the value of λ_c, i.e, the behavior of the relaxation time to the level of the correct rate of exponential growth with $1/h$.

The first ingredient of the heuristics is the idea of looking at an individual droplet of the stable phase (roughly the $(+)$-phase, since h is small) in a background given by the metastable phase (roughly the $(-)$-phase). Let S be the shape of that droplet, which a priori can be arbitrary. Say that l^2 is the volume (i.e., the number of sites) of the droplet, and let us find an expression for the free-energy of such a droplet. This free-energy may be seen as coming from two main contributions. There should be a bulk term, proportional to l^2. This term should be obtained by multiplying l^2 by the difference in free-energy per site between the $(+)$-phase and the $(-)$-phase in the presence of a small magnetic field $h > 0$. This difference in the free-energy per site of the two phases should come only from the term in the Hamiltonian which couples the spins to the external field and should therefore be given by $2m^*h/2 = m^*h$. The other relevant contribution to the free-energy of the droplet should come from its surface, where there is an interface between the $(+)$-phase and the $(-)$-phase. This contribution is proportional to the length of the interface, which is of the order of l. It should be multiplied by a constant w_{S} which depends on the shape of the droplet. This constant w_{S} represents the excess free-energy per unit of length integrated over the surface of the droplet when its scale is changed so that its volume becomes 1. Therefore, since the external field h is small, we can take for w_{S} the value $\mathcal{W}(\gamma)$, where γ is the boundary of the droplet, rescaled in this fashion. In particular, w_{S} is minimized when the droplet has the Wulff shape, and in this case $w_{\mathrm{S}} = w$. Adding the pieces, we obtain for the free-energy of the droplet the expression

$$\Phi_{\mathrm{S}}(l) = -m^*hl^2 + w_{\mathrm{S}}l.$$

The two terms in this expression become of the same order of magnitude, in case l is of the order of $1/h$. Therefore, for later convenience we write $l = b/h$, with a new variable $b \geq 0$. This yields

$$\Phi_{\mathrm{S}}(b/h) = \frac{\phi_{\mathrm{S}}(b)}{h},$$

where

$$\phi_{\mathrm{S}}(b) = -m^* b^2 + w_{\mathrm{S}} b.$$

This very simple function takes the value 0 at $b = 0$, grows with b on the interval $[0, B_c^{\mathrm{S}},]$, where $B_c^{\mathrm{S}} = B_c^{\mathrm{S}}(T) = \frac{w_{\mathrm{S}}}{2m^*}$, reaching its absolute maximum $\phi_{\mathrm{S}}(B_c^{\mathrm{S}}) = \frac{(w_{\mathrm{S}})^2}{4m^*} = A^{\mathrm{S}}(T) = A^{\mathrm{S}}$ at the end of this interval and decreases with b on the semi-infinite interval $[B_c^{\mathrm{S}}, \infty)$. It crosses the value 0 at the point $B_0^{\mathrm{S}}(T) = B_0^{\mathrm{S}} = \frac{w_{\mathrm{S}}}{m^*} = 2B_c$.

Metastability is then "understood" from the fact that systems in contact with a heat bath move towards lowering their free-energy, so that the presence of a free-energy barrier which needs to be overcome in order to create a large droplet of the stable phase with any shape keeps the system close to the metastable phase. Subcritical droplets are constantly being created by thermal fluctuations, in the metastable phase, but they tend to shrink, as dictated by the free-energy landscape. On the other hand, once a supercritical droplets is created due to a larger fluctuation, it will grow and drive the system to the stable phase, possibly colliding and coalescing in its growth with other supercritical droplets created elsewhere. As a function of h, the linear size of a critical droplet, B_c^{S}/h, blows up as $h \searrow 0$. One can then, in a somewhat circular, but heuristically-meaningful way, say that the macroscopic free-energy of droplets is indeed a relevant object of consideration. One can also hope then that sharp theorems could be conjectured and possibly proven regarding the asymptotic behavior of quantities of interest in the limit $h \searrow 0$.

Regarding the shape of the droplet, the height of this barrier is minimized by minimizing the value of the constant w_{S}, i.e., by considering Wulff-shaped droplets. This singles out the Wulff shape as the most relevant one in the heuristics above. We will simplify the notation by omitting the subscript S when talking about the Wulff shape. In particular,

$$B_c = \frac{w}{2m^*}, \quad A = \frac{w^2}{4m^*}. \tag{1.28}$$

Based on the expression above for the free-energy barrier, one predicts the rate of creation of supercritical droplets with center at a given place to be $\exp\left(\frac{-\beta A}{h}\right)$.

In what follows now we write d instead of 2, to make the role of the dimension clear in the geometric argument which comes next. We are concerned with an infinite system, and we are observing it through a local function f, which depends, say, on the spins in a finite set $\mathrm{Supp}(f)$. For us the system will have relaxed to equilibrium when $\mathrm{Supp}(f)$ is covered by a big droplet of the plus-phase, which appeared spontaneously somewhere and then grew, as discussed above. We want to estimate how long we have to wait for the probability of such an event to be large. If we suppose that the radius of supercritical droplets grows with a speed v, then we can see that the region in space-time where a droplet which covers $\mathrm{Supp}(f)$ at time t could have appeared is, roughly speaking, a cone with vertex in $\mathrm{Supp}(f)$

and which has as base the set of points which have time-coordinate 0 and are at most at distance tv from $\mathrm{Supp}(f)$. The volume of such a cone is of the order of $(vt)^d t$. The order of magnitude of the relaxation time, t_{rel}, before which the region $\mathrm{Supp}(f)$ is unlikely to have been covered by a large droplet and after which the region $\mathrm{Supp}(f)$ is likely to have been covered by such an object can now be obtained by solving the equation

$$(vt_{\mathrm{rel}})^d\, t_{\mathrm{rel}} \exp\left(-\frac{\beta A}{h}\right) = 1. \tag{1.29}$$

This gives us

$$t_{\mathrm{rel}} = v^{-d/(d+1)} \, \exp\left(\frac{\beta A}{(d+1)\, h}\right). \tag{1.30}$$

In order to use this relation to predict the way in which the relaxation time scales with h, one needs to figure out the way in which v scales with h. If we suppose, for instance, that v does not scale with h, or at least that if it goes to 0, as $h \searrow 0$, it does it so slowly that

$$\lim_{h \searrow 0} \, h^{d-1} \log v = 0, \tag{1.31}$$

then we can predict that

$$t_{\mathrm{rel}} = \exp\left(\frac{\beta A}{(d+1)\, h}\right) = \exp\left(\frac{\lambda_c}{h}\right),$$

where

$$\lambda_c = \frac{\beta A}{d+1} = \frac{\beta w^2}{(d+1)\, 4m^*} = \frac{\beta w^2}{12\, m^*},$$

in agreement with our result (1.1).

In Section 1-iii of [Sch1] and more explicitly in Section 4 of [Sch2] an argument was given in support of the conjecture that $v \sim Ch$ as $h \searrow 0$, a much stronger conjecture than (1.31). In the paper [RTMS] (see display (9) there) a different non-rigorous argument is described, in which the same conclusion is derived from an "Allen-Cahn approximation". In part 3 of this paper we will introduce a rescaling procedure and obtain results which can be seen as rigorous counterparts to (1.31). It is interesting to compare this feature of the regime of fixed T and $h \searrow 0$ with the case of fixed $h > 0$ and $T \searrow 0$, studied in [DS]. In that case the analogue of (1.31) is false, and consequently the v-term in (1.30) is of greater relevance than it is here.

2. Metastable regime

2.1. Preliminaries. In this part of the paper we will prove part (i) of Theorem 1. The first step will be to prove, in Sections 2.2 and 2.3, the following proposition.

Proposition 2.1.1. *Suppose that $T < T_c$ and $0 < \lambda < \lambda_c$. Then for each constant $a \in (0, 1/4)$, there is a positive finite constant C such that for each local observable f there is a positive finite constant $C(f)$ such that for $h > 0$*

$$\left|\mathbb{E}\left(f\left(\sigma^-_{h;\exp(\lambda/h)}\right)\right) - \langle f\rangle_{\Lambda(1/h^a),-,h}\right| \leq C(f)\exp(-C/h^a).$$

We will not try to optimize the constants C and $C(f)$ in this proposition. But we observe that from our proof, if the inequality displayed above is only required to hold for $h \leq h_0$ for some $h_0 > 0$ depending on f, then we can take $C(f) = C' ||f||_\infty |\mathrm{Supp}(f)|$, where C' does not depend on f.

Proposition 2.1.1 transforms our dynamical problem into an equilibrium one, in case the initial distribution is concentrated on the configuration with all spins down. In Section 2.4 we will study the behavior of $\langle f \rangle_{\Lambda(1/h^a),-,h}$ for small $h > 0$ and show that it gives rise to the asymptotic expansion claimed in part (i) of Theorem 1. Let us note here that if our goal were only to prove Corollary 1(i), with the initial distribution having all spins down, then Proposition 2.1.1 would have reduced our task to a very simple one. First, from the heuristic view-point, with $a > 0$ small, the box $\Lambda(1/h^a)$ is too small for any supercritical droplet to fit inside, so that one should expect to see the $(-)$-phase inside it. A rigorous argument to the effect that for $0 < a < 1$

$$\langle f \rangle_{\Lambda(1/h^a),-,h} \to \langle f \rangle_- \qquad \text{as } h \searrow 0 \tag{2.1}$$

can be obtained from the FKG-Holley inequalities in combination with part (a2) of Corollary 1 in [SS1]. But for $0 < a < 1/2$, which is our case, a simple and direct argument can also be given. It is clear that the following estimate, which can be seen as a uniform bound on a Radon-Nikodym derivative, holds. For all $\eta \in \Omega_{\Lambda(1/h^a),-}$,

$$\exp(-\beta |\Lambda(1/h^a)| h) \leq \frac{\mu_{\Lambda(1/h^a),-,h}(\eta)}{\mu_{\Lambda(1/h^a),-,0}(\eta)} \leq \exp(\beta |\Lambda(1/h^a)| h).$$

Since $|\Lambda(1/h^a)| h \leq h^{1-2a} \to 0$, as $h \searrow 0$, (2.1) follows from the weak convergence of $\mu_{\Lambda(1/h^a),-,0}$ to μ_- as $h \searrow 0$. In this argument we see directly that, with $a < 1/2$, the box $\Lambda(1/h^a)$ is too small for the external field $h > 0$, acting on each spin, to be able to win over the effect of the negative spins at the boundary.

The extension of Theorem 1(i) to arbitrary initial distributions $\nu \leq \mu_-$ will be obtained in Section 2.5. Interestingly enough a basic tool there will be a result obtained in [Mar], and extended to arbitrary subcritical temperatures in [CGMS], concerning the 2 dimensional Ising model evolving in the absence of an external field.

2.2. Bottlenecks for the dynamics. We start now the proof of Proposition 2.1.1. Several times in the proof of this proposition we will use arguments which are only true for small enough $h > 0$, but the constant $C(f)$ can be adjusted so that we do not have to require h to be small in the statement of the proposition.

In order to prove Proposition 2.1.1 there is no loss in generality in supposing that f is increasing and that it has $||f||_\infty \leq 1$. For the remainder of the proof of this proposition we will make these assumptions.

To simplify the notation we set

$$t_h = \exp(\lambda/h). \tag{2.2}$$

We turn first to the proof of the easy half of Proposition 2.1.1. We will show that for small $h > 0$

$$\mathbb{E}(f(\sigma^-_{h;t_h})) - \langle f \rangle_{\Lambda(1/h^a),-,h} \geq -\exp\Big(-\exp\Big(\frac{\lambda}{2h}\Big)\Big). \tag{2.3}$$

For this, observe first that from the basic-coupling inequalities we have

$$\mathbb{E}(f(\sigma^-_{h;t_h})) \geq \mathbb{E}(f(\sigma^-_{\Lambda(1/h^a),-,h;t_h})). \tag{2.4}$$

Let the process $(\sigma^-_{\Lambda(1/h^a),-,h;t})$ and the stationary process $(\sigma^{\mu_{\Lambda(1/h^a),-,h}}_{\Lambda(1/h^a),-,h;t})$ evolve on the probability space defined by the graphical construction, so that in particular once these processes hit each other they remain together forever. Note that, for some positive ϵ, the probability that these two processes will hit each other during any unit time interval is at least $\epsilon^{1/h^{2a}}$, regardless of their states at the begining of this time interval. Also, $f(\sigma^-_{\Lambda(1/h^a),-,h;t}) \leq f(\sigma^{\mu_{\Lambda(1/h^a),-,h}}_{\Lambda(1/h^a),-,h;t})$ with probability one. From these remarks it is clear that

$$0 \leq \langle f\rangle_{\Lambda(1/h^a),-,h} - \mathbb{E}(f(\sigma^-_{\Lambda(1/h^a),-,h;t_h})) \leq \left(1-\epsilon^{1/h^{2a}}\right)^{t_h-1} \leq \exp\left(-\exp\left(\frac{\lambda}{2h}\right)\right), \tag{2.5}$$

for small $h > 0$. The inequality (2.3) follows from (2.4) and (2.5).

The main task in this and the next sections is to prove the other half of Proposition 2.1.1, i.e., the inequality

$$\mathbb{E}(f(\sigma^-_{h;t_h})) - \langle f\rangle_{\Lambda(1/h^a),-,h} \leq C(f)\exp(-C/h^a). \tag{2.6}$$

We approach this problem borrowing some ideas from [Sch1]. As in that paper, set

$$\Lambda_h = \Lambda(\exp(\lambda_c/h)) \tag{2.7}$$

and observe that Lemma 1.2.1 (which is the same as Lemma 1 in [Sch1]) gives us the following stronger version of Lemma 2 in [Sch1]:

$$|\mathbb{E}(f(\sigma^-_{h;t_h})) - \mathbb{E}(f(\sigma^-_{\Lambda_h,-,h;t_h}))| \leq C_1(f)\exp(-C_2\exp(\lambda/h)), \tag{2.8}$$

where $C_1(f)$ and C_2 are positive and finite.

Next we will introduce a restricted set of configurations, in a way similar to [Sch1], and inspired there by [CCO] and by the heuristic idea of critical droplets. To make this idea precise one uses the standard notion of contours, on the dual lattice $\mathbb{Z}^2 + (1/2, 1/2)$, which separate spins -1 from $+1$. In the definition of these contours, we adopt here the splitting rules used in, e.g., [DKS] (see Section 3.1 there), which allow one to take the contours as self-avoiding curves, which are closed, when the boundary conditions are, e.g., $(-)$, as is our case. We will denote by $|\Gamma|$ the length of the contour Γ, by Int Γ the set of sites it surrounds, and by $V(\Gamma)$ the number of spins that it surrounds, which we call the volume of Γ. As usual, a contour is called an external contour if it is not enclosed by any other contour. If Γ is such a contour of the configuration σ, and the boundary conditions are $(-)$, then at certain sites x, attached to Γ, the values $\sigma(x)$ are uniquely defined by the presence of Γ. The set of such sites will be denoted by $\partial\Gamma$. We have

$$\partial\Gamma = \partial_-\Gamma \cup \partial_+\Gamma, \tag{2.9}$$

where $\sigma|_{\partial_\pm \Gamma} = \pm 1$. We will use the notation

$$\Omega_- = \bigcup_{l=1}^{\infty} \Omega_{\Lambda(l),-}.$$

Our restricted set of configurations is defined as

$$\mathcal{R} = \Big\{ \sigma \in \Omega_- : \text{each contour } \Gamma \text{ in } \sigma \text{ has } V(\Gamma) \le \Big(\frac{B_c}{h} \Big)^2 \Big\}, \tag{2.10}$$

where B_c is defined in (1.28).

We want to argue that up to a time as large as t_h the system evolving in the box Λ_h with $(-)$ boundary conditions and starting with all spins -1, will be unlikely to escape from $\mathcal{R}$, in which case the system indeed would look very much like the $(-)$ phase. In order to do this we introduce a modified dynamics evolving in $\Omega_{\Lambda_h,-}$, in which large droplets cannot, by definition, be formed and then we couple the unrestricted dynamics to this modified one, in a natural way. The modified dynamics is simply defined as the Markov process on $\Omega_{\Lambda_h,-}$ which evolves as the original stochastic Ising model in Λ_h, with $(-)$ boundary conditions, but for which all jumps out of $\mathcal{R}$ are suppressed. In other words, the rates, $\tilde{c}_{\Lambda_h,-,h}(x,\sigma)$, of the new process are identical to $c_{\Lambda_h,-,h}(x,\sigma)$ in case $\sigma, \sigma^x \in \Omega_{\Lambda_h,-} \cap \mathcal{R}$ and are 0 otherwise. We will denote this modified process, which is restricted to the state space $\mathcal{R}$, by $(\tilde{\sigma}^{\eta}_{\Lambda_h,-,h;t})_{t\ge 0}$,where $\eta \in \mathcal{R}$ is the initial configuration. It is well known, and very easy to prove, that such a modified process is also reversible and that since it is, in our case, irreducible, its unique invariant probability measure is $\tilde{\mu}_{\Lambda_h,-,h}$ given by

$$\tilde{\mu}_{\Lambda_h,-,h}(\,\cdot\,) = \mu_{\Lambda_h,-,h}(\,\cdot\,|\mathcal{R}).$$

This distribution is sometimes called a "restricted ensemble", and, informally speaking, represents the "metastable state".

For each initial configuration $\eta \in \mathcal{R}$, the process $(\tilde{\sigma}^{\eta}_{\Lambda_h,-,h;t})$ can be constructed on the same probability space corresponding to the graphical construction introduced in Section 1.2. For this purpose it is enough to suppress all jumps out of $\mathcal{R}$. In other words, Poisson marks which should cause such a jump are just ignored.

The important fact about this construction is that if we introduce

$$\tau = \inf\{t \ge 0 : \text{the process } (\tilde{\sigma}^{\tilde{\mu}_{\Lambda_h,-,h}}_{\Lambda_h,-,h;t}) \text{ has a suppressed jump at time } t\},$$

then

$$\sigma^-_{\Lambda_h,-,h;t} \le \tilde{\sigma}^{\tilde{\mu}_{\Lambda_h,-,h}}_{\Lambda_h,-,h;t} \quad \text{for all} \quad t < \tau. \tag{2.11}$$

(Readers who are familiar with the argumentation in [Sch1], have noted that while most of what we introduced above in connection to the restricted ensemble is similar to its counterparts in that paper, the notions are not strictly parallel to those there. The reason is that, while we are still pursuing the idea that the boundary of $\mathcal{R}$ is a bottleneck, the arguments used in [Sch1] to prove Lemma 5(i) there would give us an estimate that, while good enough to obtain Corollary 1 (i), would not be good enough to obtain an estimate as sharp as Proposition 2.1.1, which is needed for the proof of Theorem 1(i).)

We introduce now a family of sets whose union is the inner boundary of $\mathcal{R}$. For each site $x \in \mathbb{Z}^2$ define

$$\mathcal{F}_x^- = \left\{ \sigma \in \mathcal{R} : \sigma^x \notin \mathcal{R} \right\}.$$

Set also

$$\varphi = \sup_{x \in \Lambda_h} \tilde{\mu}_{\Lambda_h,-,h}(\mathcal{F}_x^-).$$

We will now consider a discrete time Markov chain embedded into the stationary process $(\tilde{\sigma}^{\tilde{\mu}_{\Lambda_h,-,h}}_{\Lambda_h,-,h;t})$. It is formed by the configurations of our process between the successive jumps. For this purpose order all the Poisson marks in the graphical construction which occur on Λ_h, according to the time they occur. Let $N(t)$ be the number of such marks from time 0 up to time t. Let $M_{x,k}$ be the event that the k-th such mark occurs at the site x, and $F_{x,k}$ be the event that immediately before this k-th mark the process $(\tilde{\sigma}^{\tilde{\mu}_{\Lambda_h,-,h}}_{\Lambda_h,-,h;t})$ is in $\mathcal{F}_x^-$. Note that $M_{x,k}$ and $F_{x,k}$ are independent events and that $\mathbb{P}(M_{x,k}) = 1/|\Lambda_h|$, while by stationarity $\mathbb{P}(F_{x,k}) = \tilde{\mu}_{\Lambda_h,-,h}(\mathcal{F}_x^-) \le \varphi$ for all x and k. Set also $K = \lfloor 2\,|\Lambda_h|\, c_{\max}(T)\, t_h \rfloor$ to obtain the estimate

$$\begin{aligned}\mathbb{P}(\tau \le t_h) &\le \mathbb{P}(N(t_h) > K) + \sum_{x \in \Lambda_h, k=1,\dots,K} \mathbb{P}(M_{x,k} \cap F_{x,k}) \\ &\le C_1 \exp(-C_2 \exp(\lambda/h)) + C_3 |\Lambda_h| t_h \varphi, \end{aligned} \tag{2.12}$$

where in the second inequality a standard large deviation estimate for Poisson random variables was used.

From (2.8), (2.11) and (2.12) it follows that

$$\mathbb{E}(f(\sigma^-_{h;t_h})) \le \int f\, d\tilde{\mu}_{\Lambda_h,-,h} + C_3 |\Lambda_h| t_h \varphi + C_4(f) \exp(-C_5 \exp(\lambda/h)). \tag{2.13}$$

All the quantities in the right-hand-side of (2.13) pertain to equilibrium statistical mechanics, so that we have reduced our dynamical problem of proving (2.6) to the equilibrium problems of proving the following two claims concerning the measure $\tilde{\mu}_{\Lambda_h,-,h}$:

i) for all $\epsilon > 0$ and $h > 0$ small enough

$$\varphi \le C_1 \exp\left(-\beta(1-\epsilon) \frac{A}{h} \right), \tag{2.14}$$

where A is defined by (1.28);

ii) for $h > 0$

$$\int f\, d\tilde{\mu}_{\Lambda_h,-,h} \le \langle f \rangle_{\Lambda(1/h^a),-,h} + C(f) \exp(-C_2/h^a). \tag{2.15}$$

It is interesting to note that the term $C_3|\Lambda_h|t_h\varphi$ in (2.13) has a direct connection with the heuristics in Section 1.4. The quantity φ plays the role of the rate of nucleation, while $|\Lambda_h|t_h$ is the space-time volume of a cylinder which plays a role similar to the space-time cone in the heuristics. The absence here of the velocity factor which appears in (1.29) is a consequence of our using an upper bound of order 1 for the velocity of propagation of effects, through (2.8). Once (2.14) is proven, it follows from the definitions (1.1), (1.28), (2.2) and (2.7) that $|\Lambda_h|t_h\varphi$ vanishes exponentially fast in $1/h$ as $h \searrow 0$.

2.3. The restricted ensemble. We start our study of the measure $\tilde{\mu}_{\Lambda_h,-,h}$ by observing that it is sufficient to study measures of the type $\tilde{\mu}_{\Lambda,-,h}$ on subsets Λ of $\mathbb{Z}^2$ which are much smaller than Λ_h. The definition of $\tilde{\mu}_{\Lambda,-,h}$ is analogous to that of $\tilde{\mu}_{\Lambda_h,-,h}$, with Λ replacing Λ_h. Suppose that for each sufficiently small value of $h > 0$ we have an event E_h which only depends on the values of the spins inside the box $x + \Lambda(1/h^3)$. Consider the larger box $x + \Lambda(2/h^3)$ and condition on what the set of exterior contours which surround at least one site in its complement, $(x + \Lambda(2/h^3))^c$, is. Let that set consist of contours $\{\Gamma_j\}$. Denote by $\Lambda(\{\Gamma_j\})$ the connected component of the complement, $(x+\Lambda(2/h^3)) \setminus \bigcup_j(\text{Int }\Gamma_j \cup \partial(\Gamma_j))$, which contains the set $x + \Lambda(1/h^3)$. Then one obtains, when h is small, that

$$\tilde{\mu}_{\Lambda_h,-,h}(E_h) = \sum_i \alpha_i \cdot \tilde{\mu}_{\Lambda^i,-,h}(E_h), \tag{2.16}$$

where the Λ^i-s denote different $\Lambda(\{\Gamma_j\})$-s, the index i runs over a finite set, $\alpha_i > 0$ and $\sum_i \alpha_i = 1$. The choices of the scales above and the need that h be small for (2.16) to hold, are clearly related to the fact that we are conditioning the Gibbs measure on the absence of any contour with volume larger than $(B_c/h)^2$. Therefore, with the choices above, we are sure that for small h the support of E_h will be disjoint from the set of sites surrounded by any contour which also surrounds any site in $(x + \Lambda(2/h^3))^c$.

The sets Λ^i which appear in (2.16) have an additional property, which will be important later: they are simply-connected.

For each value of $h > 0$ and each $x \in \Lambda_h$ the event $\mathcal{F}_x^-$ satisfies the condition above on E_h, so that to derive (2.14) it is enough to obtain a corresponding upper bound:

$$\sup_{x\in\Lambda_h} \sup_{\Lambda\in\mathcal{L}_{x,h}} \tilde{\mu}_{\Lambda,-,h}(\mathcal{F}_x^-) \le C_1 \exp\left(-\beta(1-\epsilon)\frac{A}{h}\right), \tag{2.17}$$

for small $h > 0$. Here $\mathcal{L}_{x,h}$ is the family of simply-connected sets Λ which satisfy $(x + \Lambda(1/h^3)) \cap \Lambda_h \subset \Lambda \subset x + \Lambda(2/h^3)$.

Similarly, the derivation of (2.15) is reduced to that of the following,

$$\sup_{\Lambda\in\mathcal{L}_{0,h}} \int f d\tilde{\mu}_{\Lambda,-,h} \le \langle f\rangle_{\Lambda(1/h^a),-,h} + C(f)\exp(-C_2/h^a), \tag{2.18}$$

for $h > 0$.

In order to derive (2.17) and (2.18) we will use the notion of the skeleton of a contour, as introduced in [DKS]. In what follows b is a fixed but arbitrary number in $(a, 1/4)$ and r is a fixed but also arbitrary number in $(0, b/2) \cap (0, a)$. A contour Γ will be said to be h-vertebrate if $V(\Gamma) > (1/h)^{2b}$, otherwise Γ will be said to be h-invertebrate. (Usually one says that Γ is large in the former case and small in the latter one, but this terminology would be confusing in the present paper, since "large" and "small" contours may also be used in connection with being supercritical or subcritical, with the threshold volume being the quantity $(B_c/h)^2$, which is much larger than $(1/h)^{2b}$.) Often we will omit the mention to h when referring to a vertebrate contour. Given now a vertebrate contour Γ one can associate to it, in an algorithmic way, a sequence of sites, $(x_1, ..., x_J)$ of the dual lattice $\mathbb{Z}^2 + (1/2, 1/2)$. We think of the sites $x_1, ..., x_J$ as the ordered vertices of a closed polygonal curve, with possible self-intersections (see fig. 5.3 on p.166 in [DKS]); we will denote this curve by γ in what follows and call it the skeleton of Γ. For the construction of γ, given Γ, the reader is referred to Chapter 5 of [DKS]; here we will limit ourselves to reviewing some of the basic properties that we can guarantee the skeleton to have:

(S.1) $x_i \in \Gamma$ for each i, moreover, the points x_i are consecutive on Γ (for one of the orientations of it).

(S.2) The length of each edge of γ is bounded between $C_1(1/h)^r$ and $C_2(1/h)^r$, where $0 < C_1 < C_2 < \infty$ are fixed appropriate constants.

(S.3) The Hausdorff distance between Γ and γ satisfies

$$\rho_H(\Gamma, \gamma) \leq (1/h)^r. \tag{2.19}$$

The length, $|\gamma|$ of a skeleton γ is defined as the sum of the Euclidean lengths of its edges. To each skeleton γ we associate its Wulff functional, $\mathcal{W}(\gamma)$, defined by summing over the edges of γ the product of the Euclidean length of each edge by the surface tension in the direction defined by the edge, i.e., $\tau_T(\mathbf{n})$, with $\mathbf{n}$ perpendicular to the edge. Observe that from the fact that the surface tension $\tau_T(\mathbf{n})$ is bounded away from 0 and ∞ uniformly in $\mathbf{n}$,

$$C_3\mathcal{W}(\gamma) \leq |\gamma| \leq C_4\mathcal{W}(\gamma). \tag{2.20}$$

From (S.2) and (2.20) it follows that the number $J(\gamma)$ of vertices in γ satisfies

$$C_5\mathcal{W}(\gamma)h^r \leq J(\gamma) \leq C_6\mathcal{W}(\gamma)h^r. \tag{2.21}$$

To each configuration $\sigma \in \Omega_-$ we can associate the collection $G = \{\Gamma_1, ...\Gamma_n\}$ of its external vertebrate contours. To this collection we can associate the collection $S = \{\gamma_1, ..., \gamma_n\}$ of their skeletons. The Wulff functional associated to the configuration σ is then defined as

$$\mathcal{W}(S) = \sum_{i=1}^{n} \mathcal{W}(\gamma_i),$$

with the convention that $\mathcal{W}(\emptyset) = 0$. Next we want to consider the volume surrounded by the external vertebrate contours $\Gamma_1, ..., \Gamma_n$ and say that it has to be close to the volume surrounded by the collection of skeletons $\gamma_1, ..., \gamma_n$. A difficulty lies in the fact that while the volume surrounded by the contours is easily defined as

$$V(G) = \sum_{i=1}^{n} V(\Gamma_i),$$

the fact that the skeletons can self-intersect and also intersect with each other makes the notion of the volume that they surround more delicate. Fortunately the notion of "phase volume", as defined in Section 2.10 of [DKS], solves this difficulty. This definition is as follows (a look at fig.2.5 on p.37 of [DKS] will probably lead the reader to guess correctly the definition). The set $\mathbb{R}^2 \setminus \cup \gamma_i$ splits up into a collection of connected components Q_α with exactly one unbounded component among them. A component Q_α will be called a minus-component if any path that connects its interior points with points of the unbounded component and intersects the curves from S in a finite number of points, intersects them in an odd number of points (counted with multiplicities). The phase volume of S, denoted by $\hat{V}(S)$, is defined as the joint volume of all the minus-components.

Motivated by (2.19), we want to show that $V(G)$ and $\hat{V}(S)$ have to be also relatively close to each other. If we remove from $\mathbb{R}^2$ all the points which are at a distance not larger than $(1/h)^r$ from $\cup\Gamma_i$, then the remaining set also splits up into connected components with exactly one unbounded component among them. It is easy to see that all the bounded components are subsets of minus-components in the splitting produced by S, while the unbounded component is a subset of the unbounded component in the splitting produced by S. It is also clear that the bounded components in this splitting are inside contours of G, while the unbounded component in this splitting is completely outside the contours of G. Hence

$$|V(G) - \hat{V}(S)| \le C_7 \left(\sum_{i=1}^{n} |\Gamma_i| \right) (1/h)^{2r}. \tag{2.22}$$

Similarly, by removing from $\mathbb{R}^2$ all the points which are at a distance not larger than $(1/h)^r$ from $\cup\gamma_i$, one can also derive

$$|V(G) - \hat{V}(S)| \le C_8 \left(\sum_{i=1}^{n} |\gamma_i| \right) (1/h)^{2r} \le C_9 \mathcal{W}(S)(1/h)^{2r}, \tag{2.23}$$

where in the second inequality use of (2.20) was made.

For convenience we introduce also another measure of the "volume" of a collection of skeletons, which is motivated by the procedure described in the explanation of why (2.22) and (2.23) hold. We define $\check{V}(S)$ as the number of sites inside the minus components Q_α, which are at a distance larger than $(1/h)^r$ from $\cup\gamma_i$. Clearly we have

$$\check{V}(S) \le \min\{V(G), \hat{V}(S)\}, \tag{2.24}$$

and

$$\check{V}(S) \ge \max\{V(G), \hat{V}(S)\} - C_{10}\mathcal{W}(S)(1/h)^{2r} \tag{2.25}$$

Given a finite set G of h-vertebrate contours we denote by $\mathcal{S}_G^{h,\mathrm{co}}$ the set of configurations which belong to Ω_- and which have as their collection of external

h-vertebrate contours the set G. We say that G is a compatible set of external h-vertebrate contours in case $\mathcal{S}_G^{h,\mathrm{co}}$ is not empty.

Similarly, given a finite set of skeletons S, we define $\mathcal{C}_S^h$ as the class of all G which are compatible sets of external h-vertebrate contours and which have as their set of skeletons the set S. We define also

$$\mathcal{S}_S^{h,\mathrm{sk}} = \bigcup_{G \in \mathcal{C}_S^h} \mathcal{S}_G^{h,\mathrm{co}}.$$

$\mathcal{S}_S^{h,\mathrm{sk}}$ is the set of configurations which belong to Ω_- and which have as their collection of skeletons corresponding to their external h-vertebrate contours the set S. We say that S is a compatible set of skeletons in case $\mathcal{S}_S^{h,\mathrm{sk}}$ is not empty.

Again similarly, given an interval of real numbers, I, we define $\mathcal{S}_I^{h,\mathcal{W}}$ as the set of configurations which belong to Ω_- and for which the collection S of skeletons corresponding to their external h-vertebrate contours satisfies $\mathcal{W}(S) \in I$.

One would like to say that the volume of a collection of skeletons is the sum of the volumes of the individual skeletons. One proper version of this is the following relation, which follows from the argumentation used to show (2.22) and (2.23). In this relation, and also later on, we use the simplified notation $\check{V}(\gamma_i)$ in place of the more cumbersome $\check{V}(\{\gamma_i\})$.

$$\text{If } S = \{\gamma_1, ..., \gamma_n\} \text{ is a compatible set of skeletons, then } \check{V}(S) = \sum_{i=1}^n \check{V}(\gamma_i). \tag{2.26}$$

A fundamental fact about the Wulff shape and the associated quantity w is the variational characterization (1.10). By scaling lengths we obtain easily from this and (2.24) that for any skeleton γ

$$w\sqrt{\check{V}(\gamma)} \le \mathcal{W}(\gamma). \tag{2.27}$$

This inequality will now be used to derive another one, which is of central relevance in this paper. This is the content of the next lemma.

Lemma 2.3.1. *For each configuration in $\mathcal{R}$, the associated set of skeletons S, corresponding to its set of vertebrate external contours satisfies*

$$\mathcal{W}(S) \ge 2m^* h \check{V}(S).$$

Proof. Say that the collection of external vertebrate contours for the configuration with which we are concerned is $G = \{\Gamma_1, ..., \Gamma_n\}$ and that Γ_i has skeleton γ_i, for $i = 1, ..., n$. From (2.24) and the definition (2.10) of $\mathcal{R}$ we have for each i

$$\sqrt{\check{V}(\gamma_i)} \le \frac{B_{\mathrm{c}}}{h}. \tag{2.28}$$

Multiplying the inequalities (2.27) (with $\gamma = \gamma_i$) and (2.28) by each other, and using the fact that $B_{\mathrm{c}} = w/(2m^*)$ we obtain

$$\mathcal{W}(\gamma_i) \ge 2m^* h \check{V}(\gamma_i).$$

The thesis now follows by adding over i, using (2.26). ■

In order to show (2.17), we need to know that the skeletons of configurations in $\mathcal{F}_x^-$ are associated with sufficiently large values of $\mathcal{W}(\cdot)$. This is the content of the next lemma.

Lemma 2.3.2. *Given $\epsilon > 0$ there is $h_0 > 0$ such that for all $0 < h \le h_0$ and all $x \in \mathbb{Z}^2$ the following holds. For each configuration in $\mathcal{F}_x^-$, the associated set of skeletons S, corresponding to its set of vertebrate external contours satisfies*

$$\mathcal{W}(S) \ge \frac{2A(1-\epsilon)}{h}.$$

Proof. Let $\sigma \in \mathcal{F}_x^-$ be the configuration with which we are concerned. By definition, $\sigma^x \in \mathcal{R}^c$, so the configuration σ^x has an external contour Γ with $V(\Gamma) > (B_c/h)^2$. On the other hand, $\sigma \in \mathcal{R}$, so the contour Γ of the configuration σ^x is attached to the site x. Let $sq(x)$ be the 3×3 square of the dual lattice, centered at x. Consider the set $\widetilde{G}(x,\sigma)$ of all dual bonds that are either in $sq(x)$ (there are 24 of them) or belong to a contour of σ, which contains some bonds of $sq(x)$. Let $G(x,\sigma) \subset \widetilde{G}(x,\sigma)$ be these bonds, which are "visible from infinity", i.e., are not screened away from it by other bonds in $\widetilde{G}(x,\sigma)$. Evidently, the set of bonds $G(x,\sigma)$ serves as a set of bonds of exterior contours of some new configuration $\eta = \eta(x,\sigma)$, which in fact has no interior contours. Let these contours be $G_1, G_2, ..., G_k$. Clearly, $k \le 6$, since each contour G_i passes through at least two boundary sites of $sq(x)$. Note that $\widetilde{G}(x,\sigma) = \widetilde{G}(x,\sigma^x)$, so the same is true also for $G(x,\sigma^x)$, $\eta(x,\sigma^x)$ and $G_1, G_2, ..., G_k$.

The idea of the proof is to observe that $\text{Int}\,\Gamma \subset \cup_i \text{Int}\, G_i$, so in particular $\sum_i V(G_i) \ge V(\Gamma)$. That implies that the contours G_i should be quite long. On the other hand, the set of dual bonds $G(x,\sigma) \setminus sq(x)$ belongs to exterior contours of σ, so they should be long as well.

The implementation of the above argument is as follows. Let $\widetilde{S}$ be the set of skeletons of the subset of all vertebrate contours among $G_1, G_2, ..., G_k$. Denote the corresponding subset of indices by ver $\subset \{1, ..., k\}$. We have then that $\widetilde{S} = \{\gamma_i, i \in \text{ver}\}$. Note that the volume of any invertebrate contour is $o(\frac{1}{h})$, so we have:

$$\sum_{i \in \text{ver}} V(G_i) \ge \left(\frac{B_c}{h}\right)^2 - o\left(\frac{1}{h}\right).$$

Hence the estimate (2.25) implies that also

$$\sum_{i \in \text{ver}} \check{V}(\gamma_i) \ge \left(\frac{B_c}{h}\right)^2 - o\left(\frac{1}{h}\right).$$

Using now (2.27) we have

$$\mathcal{W}(\widetilde{S}) \ge w \sum_{i \in \text{ver}} \sqrt{\check{V}(\gamma_i)} \ge w\sqrt{\left(\frac{B_c}{h}\right)^2 - o\left(\frac{1}{h}\right)} = \frac{2A}{h} - o\left(\frac{1}{h}\right), \tag{2.29}$$

where we are using an evident property of the square root and also the definitions (1.28).

The only remaining problem now is the relation between the skeletons S and $\widetilde{S}$. One would like to argue that in some sense $\widetilde{S} \subset S$, which would imply our claim. However, the last inclusion is almost always violated. The way out of this unlucky circumstance is the following. Note, that in fact the skeleton of a contour is not uniquely defined; it should just be a closed polygon satisfying the properties of the type (S.1)-(S.3) above. Once this is the case, any such polygon satisfies any statement above made about any skeleton. We are going to use this nonuniqueness in order to prepare a special family of the skeletons $\tilde{\tilde{S}}$ of the family $\{G_i, i \in \text{ver }\}$. This family is going to be constructed in such a way as to use as big pieces of the family S as possible. Namely, note that every maximal connected arc k_j of $G(x,\sigma) \setminus sq(x)$ is also an arc of some external contour of the configuration σ, and so it inherits a portion κ_j of the skeleton S. (Note that there are at most 6 such arcs.) Namely, define κ_j to be the maximal subpolygon of S with both endpoints in k_j. (It might happen that some κ_j are empty; for example, it will be the case when the external contour in question would be invertebrate.) It is immediate to see that the family of the (open) polygons $\cup\kappa_j$ can be made into a skeleton family of the contour family $\{G_i, i \in \text{ver }\}$ by adding at most six extra edges to it, which addition might require the prior removal of some ending edges from $\cup\kappa_j$ (also in the amount of at most six). This is the skeleton family $\tilde{\tilde{S}}$ sought. Now (2.29), being valid for every possible skeleton family of the contour family $\{G_i, i \in \text{ver }\}$, implies that

$$\mathcal{W}(\tilde{\tilde{S}}) \geq \frac{2A}{h} - o\left(\frac{1}{h}\right).$$

On the other hand,

$$\mathcal{W}(S) \geq \mathcal{W}(\tilde{\tilde{S}}) - o\left(\frac{1}{h}\right),$$

since every edge of $\tilde{\tilde{S}}$ except a finite number belongs to S. (Note that in fact the family S might be much longer than $\tilde{\tilde{S}}$, though it is immaterial for our argument). ■

The next lemma shows that vertebrate contours have a minimum cost.

Lemma 2.3.3. *There is a constant $h_0 > 0$ such that for $0 < h \leq h_0$ the following holds. For each configuration in Ω_-, the associated set of skeletons S, corresponding to its set of vertebrate external contours is either empty or satisfies*

$$\mathcal{W}(S) \geq \frac{w}{\sqrt{2}}\Big(\frac{1}{h}\Big)^b.$$

Proof. Suppose that S is not empty, so that there exists $\gamma \in S$. This means that γ is the skeleton of a vertebrate contour Γ. Suppose that the inequality that we should prove is false and use (2.25) to prove that then

$$\check{V}(\gamma) \geq V(\Gamma) - C\mathcal{W}(S)\Big(\frac{1}{h}\Big)^{2r} \geq \Big(\frac{1}{h}\Big)^{2b} - C\Big(\frac{1}{h}\Big)^{2r+b} \geq \frac{1}{2}\Big(\frac{1}{h}\Big)^{2b}.$$

In the last inequality we used the fact that $r < b/2$, and for this last inequality to be true h_0 has to be taken properly. Using (2.27) now, we obtain

$$\mathcal{W}(S) \geq \mathcal{W}(\gamma) \geq \frac{w}{\sqrt{2}}\Big(\frac{1}{h}\Big)^b.$$

■

The next two lemmas are of a somewhat technical nature. The first one of them will only be used in the proof of the second one, while that second one is a step towards proving (2.18) and it will also be used in the proof of the fundamental Lemma 2.3.6 below. Some readers may prefer to first read the proof of Lemma 2.3.6, which has an immediate heuristic appeal, and later return to Lemmas 2.3.4 and 2.3.5.

We will use the notation $\mathcal{P}(x,l) = \{x \overset{+*}{\longleftrightarrow} (x+\Lambda(l))^c\}$ for the event that x is (+,*)-connected to a site outside the box $x+\Lambda(l)$. Given a set G of compatible external h-vertebrate contours (possibly $G=\emptyset$), we define the following conditional Gibbs measure and corresponding expectation:

$$\mu^{G,h}_{\Lambda,-,h'} = \mu_{\Lambda,-,h'}(\,\cdot\,|\mathcal{S}^{h,\mathrm{co}}_G) \qquad \text{and} \qquad \langle f\rangle^{G,h}_{\Lambda,-,h'} = \int f\,d\mu^{G,h}_{\Lambda,-,h'}.$$

Lemma 2.3.4. *There are positive finite constants C_1 and C_2 such that for all $0<h'\le h\le 1$, all $x\in\mathbb{Z}^2$ and all finite $\Lambda\subset\mathbb{Z}^2$*

$$\mu^{\emptyset,h}_{\Lambda,-,h'}\left(\mathcal{P}\left(x,\frac{1}{2h^a}\right)\right) \le C_1\exp(-C_2/h^a).$$

Proof. We start with an observation akin to the one that originated (2.16), but with the box $x+\Lambda(3/h^{2b})$ replacing the box $x+\Lambda(2/h^3)$ used there. Conditioning on what the set of contours which surround at least one site in $(x+\Lambda(3/h^{2b}))^c$ is we obtain, when h is small,

$$\mu^{\emptyset,h}_{\Lambda,-,h'}\left(\mathcal{P}\left(x,\frac{1}{2h^a}\right)\right) = \sum_i \alpha_i\cdot\mu^{\emptyset,h}_{\Lambda^i,-,h'}\left(\mathcal{P}\left(x,\frac{1}{2h^a}\right)\right),$$

where the index i runs over a finite set, for each of its values Λ^i is a subset of $x+\Lambda(3/h^{2b})$ which contains $(x+\Lambda(1/h^{2b}))\cap\Lambda$ and $\alpha_i>0$; moreover $\sum_i\alpha_i=1$.

The event that the box Λ^i is free of h-vertebrate contours is a decreasing event, while $\mathcal{P}\left(x,\frac{1}{2h^a}\right)$ is an increasing event, so by the FKG-Holley inequalities we obtain

$$\mu^{\emptyset,h}_{\Lambda^i,-,h'}\left(\mathcal{P}\left(x,\frac{1}{2h^a}\right)\right) \le \mu_{\Lambda^i,-,h'}\left(\mathcal{P}\left(x,\frac{1}{2h^a}\right)\right).$$

Using now the fact that for each i, $|\Lambda^i|\le 9/h^{4b}$, that $b<1/4$ and that $0<h'\le h\le 1$, we have

$$\begin{aligned}
\mu_{\Lambda^i,-,h'}\left(\mathcal{P}\left(x,\frac{1}{2h^a}\right)\right) &= \frac{Z_{\Lambda^i,-,h'}(\mathcal{P}\left(x,\frac{1}{2h^a}\right))}{Z_{\Lambda^i,-,h'}}\\
&\le \frac{Z_{\Lambda^i,-,0}(\mathcal{P}\left(x,\frac{1}{2h^a}\right))}{Z_{\Lambda^i,-,0}}\exp(\beta|\Lambda^i|h)\\
&\le \mu_{\Lambda^i,-,0}\left(\mathcal{P}\left(x,\frac{1}{2h^a}\right)\right)\exp(C_1h^{1-4b})\\
&\le C_2\exp(-C_3/h^a),
\end{aligned}$$

were in the last inequality we are also using (1.17), with the role of +s and −s switched. This completes the proof of the lemma. ■

Lemma 2.3.5. *There are positive finite constants C_1 and C_2 such that for every local observable g there is $h_0 = h_0(g) > 0$ such that for all $0 < h' \le h \le h_0$ and all finite $\Lambda \subset \mathbb{Z}^2$*

$$|\langle g\rangle^{\emptyset,h}_{\Lambda,-,h'} - \langle g\rangle_{\Lambda\cap(\mathrm{Supp}(g)+\Lambda(1/h^a)),-,h'}| \le C_1||g||_\infty|Supp(g)|\exp(-C_2/h^a),$$

Proof. Without loss in generality we suppose that g is increasing and has $||g||_\infty \le 1$.

Set

$$E_1 = \bigcup_{x\in\mathrm{Supp}(g)} \mathcal{P}\left(x, \frac{1}{2h^a}\right)$$

Once more we use an argument similar to the one used to prove (2.16), this time we condition on what the (+,*)-cluster of the set $(\mathrm{Supp}(g) + \Lambda(1/h^a))^c$ is. In this fashion we obtain the following equality:

$$\langle g\rangle^{\emptyset,h}_{\Lambda,-,h'} = \left(\sum_i \alpha_i \cdot \langle g\rangle^{\emptyset,h}_{\Lambda^i,-,h'}\right)\mu^{\emptyset,h}_{\Lambda,-,h'}((E_1)^c) \ + \ \langle g|E_1\rangle^{\emptyset,h}_{\Lambda,-,h'}\ \mu^{\emptyset,h}_{\Lambda,-,h'}(E_1),$$

where the index i runs over a finite set, for each of its values Λ^i is a subset of $\Lambda\cap(\mathrm{Supp}(g)+\Lambda(1/h^a))$ which contains $\Lambda\cap\mathrm{Supp}(g)$ and $\alpha_i > 0$; moreover $\sum_i \alpha_i = 1$. Note that for small enough h_0 (depending on $\mathrm{Supp}(g)$) the fact that $a < b$ implies that no h-vertebrate contour can fit inside $\Lambda \cap (\mathrm{Supp}(g) + \Lambda(1/h^a))$, so that for each value of i we have $\langle\,\cdot\,\rangle^{\emptyset,h}_{\Lambda^i,-,h'} = \langle\,\cdot\,\rangle_{\Lambda^i,-,h'}$. The equality displayed above and Lemma 2.3.4 now lead to

$$\begin{aligned}\left|\langle g\rangle^{\emptyset,h}_{\Lambda,-,h'} - \sum_i \alpha_i \cdot \langle g\rangle_{\Lambda^i,-,h}\right| &\le 2\mu^{\emptyset,h}_{\Lambda,-,h'}(E_1)\\ &\le C_1|\mathrm{Supp}(g)|\exp(-C_2/h^a),\end{aligned} \tag{2.30}$$

for a proper choice of h_0.

Similarly, consider now the event

$$E_2 = \bigcup_{x\in\mathrm{Supp}(g)+\Lambda(1/h^a)} \mathcal{P}\left(x, \frac{1}{2h^a}\right),$$

and condition on what the (+,*)-cluster of $(\mathrm{Supp}(g) + \Lambda(2/h^a))^c$ is. The same reasoning above, gives us, for small enough h_0,

$$\begin{aligned}\left|\langle g\rangle^{\emptyset,h}_{\Lambda,-,h'} - \sum_i \alpha_j \cdot \langle g\rangle_{\Lambda^j,-,h}\right| &\le 2\mu^{\emptyset,h}_{\Lambda,-,h'}(E_2)\\ &\le C_1|\mathrm{Supp}(g)|\exp(-C_2/h^a),\end{aligned} \tag{2.31}$$

where the index j runs over a finite set, for each of its values Λ^j is a subset of $\Lambda \cap (\mathrm{Supp}(g) + \Lambda(2/h^a))$ which contains $\Lambda \cap (\mathrm{Supp}(g) + \Lambda(1/h^a))$ and $\alpha_j > 0$; moreover $\sum_j \alpha_j = 1$.

Thanks to the fact that g is being supposed to be increasing, we have for all i and all j as above

$$\langle g\rangle_{\Lambda^i,-,h} \leq \langle g\rangle_{\Lambda\cap(\mathrm{Supp}(g)+\Lambda(1/h^a)),-,h} \leq \langle g\rangle_{\Lambda^j,-,h}. \tag{2.32}$$

The lemma follows from (2.30), (2.31) and (2.32). ■

The next lemma is a fundamental step, in which an aspect of the heuristics about droplets and their surface and bulk free-energies is made into a rigorous results.

Lemma 2.3.6. *For any $p > 0$, given $\epsilon > 0$ there is a finite positive constant h_0 such that for any $0 < h \leq h_0$, any collection of skeletons S, and any simply-connected set $\Lambda \subset \mathbb{Z}^2$ which satisfies $|\Lambda| \leq 1/h^p$,*

$$\begin{aligned}\tilde{\mu}_{\Lambda,-,h}(\mathcal{S}_S^{h,\mathrm{sk}}) &\leq \exp\left(-\beta\Big((1-\epsilon)\mathcal{W}(S) - (1+\epsilon)hm^*\check{V}(S)\Big)\right)\\ &\leq \exp\left(-\beta(1-3\epsilon)\frac{\mathcal{W}(S)}{2}\right).\end{aligned}$$

Note. In this and some further statements the restriction $|\Lambda| \leq 1/h^p$ is not really necessary. Still, we are using it since it simplifies some arguments, and is enough for our purposes.

Proof. The second inequality is immediate from Lemma 2.3.1, so that we only have to prove the first one. It is enough to consider the case of nonempty skeleton S. We start with

$$\begin{aligned}\tilde{\mu}_{\Lambda,-,h}(\mathcal{S}_S^{h,\mathrm{sk}}) &\leq \sum_{G\in\mathcal{C}_S^h} \frac{Z_{\Lambda,-,h}(\mathcal{S}_G^{h,\mathrm{co}})}{Z_{\Lambda,-,h}(\mathcal{S}_\emptyset^{h,\mathrm{co}})}\\ &= \sum_{G\in\mathcal{C}_S^h} \frac{Z_{\Lambda,-,0}(\mathcal{S}_G^{h,\mathrm{co}})}{Z_{\Lambda,-,0}(\mathcal{S}_\emptyset^{h,\mathrm{co}})} \exp\left(\frac{\beta}{2}\int_0^h \sum_{x\in\Lambda}\left[\langle\sigma(x)\rangle^{G,h}_{\Lambda,-,h'} - \langle\sigma(x)\rangle^{\emptyset,h}_{\Lambda,-,h'}\right] dh'\right),\end{aligned} \tag{2.33}$$

where in the first step we used the fact that $\mathcal{R} \supset \mathcal{S}_\emptyset^{h,\mathrm{co}}$ when h is small, while in the second step we used the fact that for an arbitrary $E \subset \Omega_{\Lambda,-}$,

$$\frac{d}{dh}\log Z_{\Lambda,-,h}(E) = \frac{\beta}{2}\sum_{x\in\Lambda}\langle\sigma(x)|E\rangle_{\Lambda,-,h}.$$

Next we will show that given $\epsilon > 0$ it is possible to take $h_0 > 0$ small enough so that if $G \in \mathcal{C}_S^h$, $0 < h' \leq h \leq h_0$, and Λ is simply-connected and satisfies $|\Lambda| \leq 1/h^p$, then

$$\sum_{x\in\Lambda}\left[\langle\sigma(x)\rangle^{G,h}_{\Lambda,-,h'} - \langle\sigma(x)\rangle^{\emptyset,h}_{\Lambda,-,h'}\right] \leq 2(1+\epsilon)m^*\check{V}(S) + C\mathcal{W}(S)/h^{2a} + 1. \tag{2.34}$$

(The reader should not be confused by the fact that a seemingly does not enter the l.h.s. In fact, it enters, because the restriction $G \in \mathcal{C}_S^h$ depends on the related parameter r.)

For this purpose let $\partial_- G$ (*resp.* $\partial_+ G$) be the set of sites where each configuration in Ω_- with the set of external contours equal to G is doomed to be -1 (*resp.* $+1$). Let $\Lambda_{\rm ext}^G$ and $\Lambda_{\rm int}^G$ be the components of $\Lambda \backslash (\partial_- G \cup \partial_+ G)$ which are, respectively, external and internal to the contours in G. Let also $\check{\Lambda}_{\rm ext}^G$ and $\check{\Lambda}_{\rm int}^G$ be, respectively, the subsets of $\Lambda_{\rm ext}^G$ and $\Lambda_{\rm int}^G$ obtained by removing from these sets all sites which are at a distance not larger than $2/h^a$ from any point in any contour of G.

First we consider the sites x which are neither in $\check{\Lambda}_{\rm int}^G$ nor in $\check{\Lambda}_{\rm ext}^G$. For these we have, using (2.19) (2.20) and the fact that $r < a$, that when h_0 is small,

$$\sum_{x \in \Lambda \backslash (\check{\Lambda}_{\rm int}^G \cup \check{\Lambda}_{\rm ext}^G)} \left[\langle \sigma(x) \rangle_{\Lambda,-,h'}^{G,h} - \langle \sigma(x) \rangle_{\Lambda,-,h'}^{\emptyset,h} \right] \leq 2|\Lambda \backslash (\check{\Lambda}_{\rm int}^G \cup \check{\Lambda}_{\rm ext}^G)| \leq C\mathcal{W}(S)/h^{2a}. \tag{2.35}$$

Regarding now the sites $x \in \check{\Lambda}_{\rm ext}^G$, we observe that for these sites

$$\langle \sigma(x) \rangle_{\Lambda,-,h'}^{G,h} = \langle \sigma(x) \rangle_{\Lambda_{\rm ext}^G,-,h'}^{\emptyset,h}.$$

But for each such x we have $\Lambda_{\rm ext}^G \cap (x + \Lambda(1/h^a)) = \Lambda \cap (x + \Lambda(1/h^a))$, so that a double application of Lemma 2.3.5 gives us, for small h_0, that

$$|\langle \sigma(x) \rangle_{\Lambda_{\rm ext}^G,-,h'}^{\emptyset,h} - \langle \sigma(x) \rangle_{\Lambda,-,h'}^{\emptyset,h}| \leq C_1 \exp(-C_2/h^a).$$

Combining the last two displays, and using the fact that $|\check{\Lambda}_{\rm ext}^G| \leq |\Lambda| \leq 1/h^p$, we obtain for small enough h_0,

$$\sum_{x \in \check{\Lambda}_{\rm ext}^G} \left[\langle \sigma(x) \rangle_{\Lambda,-,h'}^{G,h} - \langle \sigma(x) \rangle_{\Lambda,-,h'}^{\emptyset,h} \right] \leq 1. \tag{2.36}$$

Finally, regarding now the sites $x \in \check{\Lambda}_{\rm int}^G$, we observe that for these sites

$$\langle \sigma(x) \rangle_{\Lambda,-,h'}^{G,h} = \langle \sigma(x) \rangle_{\Lambda_{\rm int}^G,+,h'}.$$

But since each such x is separated from the boundary of the set $\Lambda_{\rm int}^G$ by a minimal distance $1/h^a$, when h is small, we obtain from the FKG-Holley inequalities, (1.22) and (1.7)

$$\langle \sigma(x) \rangle_{\Lambda_{\rm int}^G,+,h'} \leq \langle \sigma(x) \rangle_{\Lambda_{\rm int}^G,+,h} \leq m(h) + C_1 \exp(-C_2/h^a) \leq m^*(1+\epsilon),$$

provided h_0 is chosen small enough. On the other hand, since Λ is simply-connected, for each $x \in \check{\Lambda}_{\rm int}^G$ we have $\Lambda \cap (x + \Lambda(1/h^a)) = x + \Lambda(1/h^a)$, so that Lemma 2.3.5 gives us, for small h_0,

$$|\langle \sigma(x) \rangle_{\Lambda,-,h'}^{\emptyset,h} - \langle \sigma(x) \rangle_{x+\Lambda(1/h^a),-,h'}| \leq C_1 \exp(-C_2/h^a).$$

By another application of the FKG-Holley inequalities and (1.22) (with $+1$ and -1 switched) we have

$$\langle \sigma(x) \rangle_{x+\Lambda(1/h^a),-,h'} \geq \langle \sigma(x) \rangle_{x+\Lambda(1/h^a),-,0} \geq -m^* - C_1 \exp(-C_2/h^a) \geq -m^*(1+\epsilon),$$

provided h_0 is chosen small enough. Combining the last four displays, and using the fact that $|\check{\Lambda}^G_{\text{int}}| \leq \check{V}(S)$, for small enough h_0 we obtain

$$\sum_{x \in \check{\Lambda}^G_{\text{int}}} \left[\langle \sigma(x) \rangle^{G,h}_{\Lambda,-,h'} - \langle \sigma(x) \rangle^{\emptyset,h}_{\Lambda,-,h'} \right] \leq 2(1+\epsilon) m^* \check{V}(S). \tag{2.37}$$

By adding (2.35), (2.36) and (2.37), we obtain (2.34).

Keeping in mind that our concern is with the r.h.s. of (2.33), we observe now that

$$\sum_{G \in \mathcal{C}^h_S} \frac{Z_{\Lambda,-,0}(\mathcal{S}^{h,\text{co}}_G)}{Z_{\Lambda,-,0}(\mathcal{S}^{h,\text{co}}_\emptyset)} = \frac{\mu_{\Lambda,-,0}(\mathcal{S}^{h,\text{sk}}_S)}{\mu_{\Lambda,-,0}(\mathcal{S}^{h,\text{co}}_\emptyset)}. \tag{2.38}$$

The numerator of this fraction is controlled in the fundamental Lemma 10.1 in [Pfi], where it is shown to satisfy

$$\mu_{\Lambda,-,0}(\mathcal{S}^{h,\text{sk}}_S) \leq \exp(-\beta \mathcal{W}(S) + CJ(S)),$$

where $J(S) = \sum J(\gamma_i)$, in case $S = \{\gamma_1, ..., \gamma_n\}$, with $J(\gamma)$ being the number of vertices of the skeleton γ. In reality, in the statement of Lemma 10.1 in [Pfi] the assumption that the temperature is low enough is made. But, as pointed out in [Iof2], this assumption is actually not needed for Pfister's elegant proof, based on a clever use of duality and Griffiths' inequalities, to work. Using (2.21) we can now write that given $\epsilon > 0$ there is $h_0 > 0$ such that for $0 < h \leq h_0$

$$\mu_{\Lambda,-,0}(\mathcal{S}^{h,\text{sk}}_S) \leq \exp(-\beta(1-\epsilon/2)\mathcal{W}(S)). \tag{2.39}$$

Turning now to the denominator in the r.h.s. of (2.38), observe that at the boundary of any external vertebrate contour of a configuration in Ω_- there is a $(+,^*)$-chain with l^∞-diameter at least $1/h^b$. If $\sigma \in (\mathcal{S}^{h,\text{co}}_\emptyset)^c$, then in σ there is such a $(+,^*)$-chain, and (1.17) (with $+1$ and -1 switched) can be used in combination with $|\Lambda| \leq 1/h^p$ to conclude that

$$\mu_{\Lambda,-,0}(\mathcal{S}^{h,\text{co}}_\emptyset) \geq \frac{1}{2}, \tag{2.40}$$

for small enough h.

The first inequality claimed in the statement of the lemma follows from (2.33), (2.34), (2.38), (2.39) and (2.40), since $Ch^{1-2a} < \epsilon/2$ if h is small.

∎

The next lemma takes care of some remaining entropy.

Lemma 2.3.7. *For any $p > 0$, given $\epsilon > 0$ there exists $h_0 > 0$ such that for all $u > 0$ and $D > 0$, there is a finite positive constant C such that for any $0 \le h \le h_0$, and any simply-connected set $\Lambda \subset \mathbb{Z}^2$ which satisfies $|\Lambda| \le 1/h^p$,*

$$\tilde{\mu}_{\Lambda,-,h}\left(\mathcal{S}^{h,\mathcal{W}}_{\left[\frac{D}{h^u},\infty\right)}\right) \le C \exp\left(-\beta(1-\epsilon)\frac{D}{2h^u}\right).$$

Proof. There is no loss in generality in supposing that $0 < \epsilon < 1$ and that $u > r$; the second of these claims being justified by Lemma 2.3.3 and the fact that $r < b$.

We start by estimating $\tilde{\mu}_{\Lambda,-,h}\left(\mathcal{S}^{h,\mathcal{W}}_{\left[\frac{D}{h^u}k,\frac{D}{h^u}(k+1)\right)}\right)$, for $k = 1, 2, \ldots$. To bound these quantities from above, using Lemma 2.3.6, all that we need is an upper bound on the number of choices of skeletons S which correspond to configurations in $\Omega_{\Lambda,-}$ and have $\mathcal{W}(S) \in \left[\frac{D}{h^u}k, \frac{D}{h^u}(k+1)\right)$. Using (2.21), the number $J(S)$ of vertices that S can have is bounded above by $N_k = C_1 D(k+1)/h^{u-r} \le C_2 Dk/h^{u-r}$. Let V be the set of distinct points which are possible vertices of S. The cardinality of V is bounded above by $4|\Lambda| \le 4/h^p$. We consider now ordered N_k-tuples of points in V and associate to each point in such an N_k-tuple one of the words "continue", "close", or "quit". To every such object we associate a collection of closed polygons in the following way. We start from the first point in the N_k-tuple and while we keep seeing the word "continue", we do the following. We join the first point to the second point, this one to the third point, and so on. When we first reach a point were we read "close", we connect it to the first point of the N_k-tuple, closing a first polygonal line, and then we jump to the point which in the N_k-tuple follows the point where we just read "close". While we do not see the word "quit", we proceed in this fashion, understanding that the word "close" means that we close the polygonal line which we are currently constructing, and that then we jump to the next point in the N_k-tuple and start the next polygonal line from there. The word "quit" is self-explanatory: we stop the procedure by closing the last polygon and disregard the remaining points of the N_k-tuple. The procedure that we described generates all the collections of skeletons with which we are concerned and plenty of additional garbage. At any rate, counting the number of options here gives us an upper bound on the quantity we are interested in. This upper bound on the number of choices of S is bounded above by

$$|V|^{N_k} 3^{N_k} \le \left(\frac{12}{h^p}\right)^{\frac{C_2 D}{h^{u-r}}k} \le \exp\left(\beta\frac{\epsilon D}{4h^u}k\right),$$

for small enough h. Combining this estimate with Lemma 2.3.6, gives us that for small h_0

$$\begin{aligned}\tilde{\mu}_{\Lambda,-,h}\left(\mathcal{S}^{h,\mathcal{W}}_{\left[\frac{D}{h^u},\infty\right)}\right) &= \sum_{k=1}^{\infty} \tilde{\mu}_{\Lambda,-,h}\left(\mathcal{S}^{h,\mathcal{W}}_{\left[\frac{D}{h^u}k,\frac{D}{h^u}(k+1)\right)}\right)\\ &\le \sum_{k=1}^{\infty} C\exp\left(-\beta\left(1-\frac{\epsilon}{2}\right)\frac{D}{2h^u}k\right)\exp\left(\beta\frac{\epsilon D}{4h^u}k\right) \le C'\exp\left(-\beta(1-\epsilon)\frac{D}{2h^u}\right),\end{aligned}$$

provided $\epsilon < 1$. ■

We are now close to completing the proof of Proposition 2.1.1. The inequality (2.17) is a direct consequence of Lemmas 2.3.2 and 2.3.7. For use in Section 3.2, we observe that the same argument gives for each $p > 0$ and $\epsilon > 0$ the existence of some finite C_1 so that

$$\sup_{\substack{\Lambda \text{ simply-connected} \\ |\Lambda| \le (1/h)^p}} \sup_{x \in \Lambda} \tilde{\mu}_{\Lambda,-,h}(\mathcal{F}_x^-) \le C_1 \exp\left(-\beta(1-\epsilon)\frac{A}{h}\right), \tag{2.41}$$

for small h.

Turning to (2.18), Lemma 2.3.5, with $h = h'$, gives us that when h is small, for all $\Lambda \in \mathcal{L}_{0,h}$

$$|\langle f \rangle_{\Lambda,-,h}^{\emptyset,h} - \langle f \rangle_{\Lambda(1/h^a)),-,h}| \le C(f)\exp(-C_1/h^a). \tag{2.42}$$

We will be done once we replace the conditional Gibbs distribution $\mu_{\Lambda,-,h}^{\emptyset,h}$, implicit in this expression, with the conditional Gibbs distribution $\tilde{\mu}_{\Lambda,-,h}$. To this end we combine Lemmas 2.3.3 and 2.3.7 to write

$$\tilde{\mu}_{\Lambda,-,h}\left((\mathcal{S}_\emptyset^{h,\mathrm{co}})^c\right) \le C_2 \exp(-C_3/h^b).$$

Combining this inequality with the equalities

$$\int f d\tilde{\mu}_{\Lambda,-,h} = \int_{\mathcal{S}_\emptyset^{h,\mathrm{co}}} f d\tilde{\mu}_{\Lambda,-,h} + \int_{\left(\mathcal{S}_\emptyset^{h,\mathrm{co}}\right)^c} f d\tilde{\mu}_{\Lambda,-,h},$$

and

$$\langle f \rangle_{\Lambda,-,h}^{\emptyset,h} = \frac{\int_{\mathcal{S}_\emptyset^{h,\mathrm{co}}} f d\tilde{\mu}_{\Lambda,-,h}}{\tilde{\mu}_{\Lambda,-,h}(\mathcal{S}_\emptyset^{h,\mathrm{co}})},$$

gives us

$$\left|\int f d\tilde{\mu}_{\Lambda,-,h} - \langle f \rangle_{\Lambda,-,h}^{\emptyset,h}\right| \le C_4(f)\exp(-C_3/h^b) \le C_4(f)\exp(-C_3/h^a). \tag{2.43}$$

Together (2.42) and (2.43) give us (2.18), and Proposition 2.1.1 is proved.

2.4. Asymptotic expansion.

Proposition 2.4.1. *Suppose the $T < T_c$. Then for each constant $a \in (0, 1/2)$ and local observable f, for $n = 1, 2, 3, \ldots$, the following expansion holds when $h > 0$:*

$$\langle f \rangle_{\Lambda(1/h^a),-,h} = \sum_{j=0}^{n-1} b_j(f) h^j + O(h^n),$$

where

$$b_j(f) = \frac{1}{j!}\frac{d^j \langle f \rangle_{-,h}}{dh^j}\bigg|_{h=0_-} = \frac{1}{j!}\left(\frac{\beta}{2}\right)^j \sum_{x_1,\ldots,x_j \in \mathbb{Z}^2} \langle f; \sigma(x_1); \ldots; \sigma(x_j)\rangle_-,$$

and $O(h^n)$ *is a function of* f *and* h *which satisfies* $\limsup_{h\searrow 0}|O(h^n)|/h^n<\infty$.

The existence of the derivatives $b_j(f)$ and their relations with the summations over generalized Ursell functions are contained in (1.25), modulo the interchange of the role of +s and −s. Below we will also be using various other relations from Section 1.3 modulo this symmetry; we will do it without further warning.

From (1.23) and (1.17), in combination with an idea already explained in Section 2.1 in connection with the derivation of (2.1) and also exploited at the end of the proof of Lemma 2.3.4 (and which amounts to an estimate on a Radon-Nikodym derivative), we have for each $a\in(0,1/2)$, $T<T_c$, $0\le h'\le h\le 1$ and local observables f and g,

$$
\begin{aligned}
|\langle f;g\rangle_{\Lambda(1/h^a),-,h'}| &\le C(f,g)\ \mu_{\Lambda(1/h^a),-,h'}\left(\mathrm{Supp}(f)\stackrel{+}{\longleftrightarrow}\mathrm{Supp}(g)\right)\\
&= C(f,g)\ \frac{Z_{\Lambda(1/h^a),-,h'}\left(\mathrm{Supp}(f)\stackrel{+}{\longleftrightarrow}\mathrm{Supp}(g)\right)}{Z_{\Lambda(1/h^a),-,h'}}\\
&\le C(f,g)\ \frac{Z_{\Lambda(1/h^a),-,0}\left(\mathrm{Supp}(f)\stackrel{+}{\longleftrightarrow}\mathrm{Supp}(g)\right)}{Z_{\Lambda(1/h^a),-,0}}\exp(\beta|\Lambda(1/h^a)|h')\\
&\le C(f,g)\ e^{\beta}\ \mu_{\Lambda(1/h^a),-,0}\left(\mathrm{Supp}(f)\stackrel{+}{\longleftrightarrow}\mathrm{Supp}(g)\right)\\
&\le C(f,g,T)\exp\Big(-C(T)\ \mathrm{dist}_\infty(\mathrm{Supp}(f),\mathrm{Supp}(g))\Big). \qquad (2.44)
\end{aligned}
$$

As with (1.27) in Section 1.3, the argument in Appendix B of [M-L] shows that from the exponential decay of correlations in (2.44) a similar exponential decay follows for the generalized Ursell functions. With a, T, h, h' and f as above,

$$
|\langle f;\sigma(x_1);...;\sigma(x_j)\rangle_{\Lambda(1/h^a),-,h'}|\le C_j(T,f)\exp\left(-C(T)\frac{\mathrm{diam}_\infty(\mathrm{Supp}(f)\cup\{x_1,...,x_j\})}{j}\right). \qquad (2.45)
$$

If we keep $0<h\le 1$ fixed and look at $\langle f\rangle_{\Lambda(1/h^a),-,h'}$ as a function of $h'\ge 0$, we can write the following Taylor expansion:

$$
\begin{aligned}
\langle f\rangle_{\Lambda(1/h^a),-,h} &= \sum_{j=0}^{n-1}\frac{1}{j!}\frac{d^j\langle f\rangle_{\Lambda(1/h^a),-,h'}}{d(h')^j}\bigg|_{h'=0}h^j+\frac{1}{n!}\frac{d^n\langle f\rangle_{\Lambda(1/h^a),-,h'}}{d(h')^n}\bigg|_{h'=h''(h)}h^n\\
&=\sum_{j=0}^{n-1}\left\{\frac{1}{j!}\left(\frac{\beta}{2}\right)^j\sum_{x_1,\dots,x_j\in\Lambda(1/h^a)}\langle f;\sigma(x_1);...;\sigma(x_j)\rangle_{\Lambda(1/h^a),-,0}\right\}h^j+\\
&\quad+\left\{\frac{1}{n!}\left(\frac{\beta}{2}\right)^n\sum_{x_1,\dots,x_n\in\Lambda(1/h^a)}\langle f;\sigma(x_1);...;\sigma(x_n)\rangle_{\Lambda(1/h^a),-,h''(h)}\right\}h^n, \qquad (2.46)
\end{aligned}
$$

where all that we know about $h''(h)$ is that $h''(h)\in[0,h]$. Thanks to the uniformity on h and h' of the bound (2.45), this is enough to conclude that the coefficient which multiplies h^n in the last term of (2.46) is bounded in absolute value by a constant which does not depend on h. In other words, this last term in (2.46) is indeed a $O(h^n)$.

To conclude the proof of Proposition 2.4.1 we must show that for $j = 0, 1, ..., n-1$, the coefficient which multiplies h^j converges fast enough to $b_j(f)$. We will show that this convergence occurs at a rate which is more than enough for this purpose. Indeed, using first (1.27) and then (1.26), we have for some finite positive C_1, C_2, C_3 and C_4,

$$\left| \sum_{x_1,...,x_j \in \Lambda(1/h^a)} \langle f; \sigma(x_1); ...; \sigma(x_j) \rangle_{\Lambda(1/h^a),-,0} - \sum_{x_1,...,x_j \in \mathbb{Z}^2} \langle f; \sigma(x_1); ...; \sigma(x_j) \rangle_- \right|$$

$$\leq \left| \sum_{x_1,...,x_j \in \Lambda(1/2h^a)} \langle f; \sigma(x_1); ...; \sigma(x_j) \rangle_{\Lambda(1/h^a),-,0} - \sum_{x_1,...,x_j \in \Lambda(1/2h^a)} \langle f; \sigma(x_1); ...; \sigma(x_j) \rangle_- \right|$$

$$+ C_1 \exp\left(-\frac{C_2}{h^a}\right)$$

$$\leq \sum_{x_1,...,x_j \in \Lambda(1/2h^a)} \left| \langle f; \sigma(x_1); ...; \sigma(x_j) \rangle_{\Lambda(1/h^a),-,0} - \langle f; \sigma(x_1); ...; \sigma(x_j) \rangle_- \right|$$

$$+ C_1 \exp\left(-\frac{C_2}{h^a}\right)$$

$$\leq C_3 \exp\left(-\frac{C_4}{h^a}\right).$$

This completes the proof of Proposition 2.4.1. ■

We have now finished the proof of Theorem 1(i) in the special case in which the initial distribution ν is concentrated on the configuration with all spins down.

2.5. More general initial distributions. In this section we will show that part (i) of Theorem 1 for an initial distribution $\nu \leq \mu_-$ can be derived from the same result for the particular case in which ν is concentrated on the configuration with all spins down.

With no loss in generality we will again suppose that f is increasing and has $||f||_\infty \leq 1$. In this case for all $t > 0$

$$\mathbb{E}(f(\sigma^\nu_{h;t})) \geq \mathbb{E}(f(\sigma^-_{h;t})).$$

The claim therefore will be proven once we show that for all λ and λ' which satisfy $0 < \lambda < \lambda'$, there are finite positive constants C_1 and C_2 such that for $h > 0$

$$\mathbb{E}(f(\sigma^\nu_{h;\exp(\lambda/h)})) \leq \mathbb{E}(f(\sigma^-_{h;\exp(\lambda'/h)})) + C_1 \exp(-C_2/h). \tag{2.47}$$

To prove this inequality we first note that for arbitrary $s, t \geq 0$,

$$\mathbb{E}(f(\sigma^\nu_{h;t})) \leq \mathbb{E}(f(\sigma^{\mu_-}_{h;t})) = \int \mathbb{P}(\sigma^{\mu_-}_{0;s} \in d\zeta) \mathbb{E}(f(\sigma^\zeta_{h;t})).$$

From Lemma 1.2.1 there exists finite positive C_3, C_4 and C_5 so that we have

$$\int \mathbb{P}(\sigma^{\mu-}_{0;s} \in d\zeta)\mathbb{E}(f(\sigma^{\zeta}_{h;t})) \le \int \mathbb{P}(\sigma^{-}_{0;s} \in d\zeta)\mathbb{E}(f(\sigma^{\zeta}_{h;t})) + \mathbb{P}\left(\sigma^{-}_{0;s}(x) \neq \sigma^{\mu-}_{0;s}(x) \text{ for some } x \in \Lambda(C_3 t)\right) \; + \; C_4 \exp(-C_5 t).$$

From the basic-coupling inequalities and the Markov property,

$$\int \mathbb{P}(\sigma^{-}_{0;s} \in d\zeta)\mathbb{E}(f(\sigma^{\zeta}_{h;t})) \le \int \mathbb{P}(\sigma^{-}_{h;s} \in d\zeta)\mathbb{E}(f(\sigma^{\zeta}_{h;t})) = \mathbb{E}(f(\sigma^{-}_{h;s+t})).$$

Combining the last three displays, and taking $t = s = \exp(\lambda/h)$ we obtain

$$\begin{aligned}
\mathbb{E}(f(\sigma^{\nu}_{h;\exp(\lambda/h)})) \le{}& \mathbb{E}(f(\sigma^{-}_{h;2\exp(\lambda/h)})) \\
&+ |\Lambda(C_3 \exp(\lambda/h))| \; \mathbb{P}\left(\sigma^{-}_{0;\exp(\lambda/h)}(0) \neq \sigma^{\mu-}_{0;\exp(\lambda/h)}(0)\right) \\
&+ C_4 \exp(-C_5 \exp(\lambda/h)) \\
\le{}& \mathbb{E}(f(\sigma^{-}_{h;\exp(\lambda'/h)})) \\
&+ C_6 \exp(2\lambda/h) \; \left(\mathbb{E}\left(\sigma^{+}_{0;\exp(\lambda/h)}(0)\right) - m^*\right) \\
&+ C_4 \exp(-C_5 \exp(\lambda/h)),
\end{aligned}$$

when h is small. In the second inequality above we used the basic-coupling inequalities twice (which imply in particular the monotonicity of $\mathbb{E}(f(\sigma^{-}_{h;t}))$ in t), and also the spin-reversal symmetry in case $h = 0$.

To complete the proof of (2.47) we have to show that $\mathbb{E}(\sigma^{+}_{0;u}(0)) \to m^*$ fast enough as $u \to \infty$. The following lemma, which states that this happens faster than any power of $1/u$, is clearly sufficient for our purpose.

Lemma 2.5.1. *Suppose $T < T_c$. Then for each $p > 0$ there is a positive finite constant C such that*

$$0 \le \mathbb{E}\left(\sigma^{+}_{0;u}(0)\right) - m^* \le C u^{-p}.$$

Proof. The lower bound is a standard application of a basic-coupling inequality. To prove the upper bound, we will first also use basic-coupling inequalities, in order to compare the infinite system with finite ones with $(+)$ boundary conditions. For these finite systems we will then use a result in [Mar], which was extended up to T_c in [CGMS].

The result of [Mar] and [CGMS] that we will use refers to the spectral gap of the generator a kinetic Ising model in a finite box. For each finite $\Lambda \subset \mathbb{Z}^2$, $\eta \in \Omega$ and h, the process $(\sigma^{\cdot}_{\Lambda,\eta,h;t})_{t \ge 0}$ is a finite-state-space reversible irreducible Markov process and its generator has its (discrete) spectrum contained in the interval $(-\infty, 0]$, with 0 being in the spectrum. The spectral gap, denoted by $\text{gap}(\Lambda, \eta, h)$, is then simply the absolute value of the largest non-zero number in the spectrum.

It is shown in [Mar] (Theorem 3.1) that for low enough T, given $\epsilon \in (0, 1/2)$, there exists a finite positive C so that

$$\text{gap}(\Lambda(l), +, 0) \ge \exp(-C l^{1/2+\epsilon}).$$

The common belief is that even up to T_c the lower bound on $\mathrm{gap}(\Lambda(l),+,0)$ in this inequality is far from optimal. No rigorous result in this direction is available, and for temperatures close to T_c only a weaker bound has been proven. That bound, which is implicitly derived in [CGMS] (see the introduction of that paper), states that for any $\epsilon > 0$

$$\mathrm{gap}(\Lambda(l),+,0) \geq \exp(-\epsilon l), \tag{2.48}$$

for large l. Fortunately for us, this estimate is suffices for our purpose here.

For any $l > 0$, we can write, using basic-coupling inequalities, a standard estimate for the relaxation to equilibrium of expected values of observables in terms of the spectral gap (see, e.g., inequality (59) in [Sch1]), and (1.22),

$$\begin{aligned} \mathbb{E}\left(\sigma^{+}_{0;u}\right) - m^* &\leq \left(\mathbb{E}\left(\sigma^{+}_{\Lambda(l),+,0;u}(0)\right) - \langle\sigma(0)\rangle_{\Lambda(l),+,0}\right) + \left(\langle\sigma(0)\rangle_{\Lambda(l),+,0} - m^*\right) \\ &\leq \frac{e^{-\mathrm{gap}(\Lambda(l),+,0)\,u}}{\mu_{\Lambda(l),+,0}(+)} + C_1\exp(-C_2 l). \end{aligned} \tag{2.49}$$

We use now (2.48) with $\epsilon = C_2/(2p)$, and we choose $l = (\log u)/(2\epsilon)$. Since $\mu_{\Lambda(l),+,0}(+) \geq \exp(-C_3 l^2)$, the first term in the right hand side of (2.49) is bounded above by $\exp(-\sqrt{u}/2)$, when u is large. This finishes the proof, since the second term in the right hand side of (2.49) is of the claimed form and this upper bound on the first term goes to 0 even faster.

■

3. Relaxation regime

3.1. Preliminaries. In this part of the paper we will prove part (ii) of Theorem 1. For this we suppose that $\lambda > \lambda_c$ is fixed and that $\nu \leq \mu_-$. Once more, there is no loss in generality in supposing that f is increasing and that it has $||f||_\infty \leq 1$. For the remainder of the proof we will make these assumptions.

Half of our goal is trivial, since the basic-coupling inequalities give for each $t \geq 0$

$$\mathbb{E}(f(\sigma^{\nu}_{h;t})) \leq \mathbb{E}(f(\sigma^{\mu_h}_{h;t})) = \langle f\rangle_h.$$

Our goal is therefore reduced to proving that for all $C > 0$ there exists a finite C_1 such that

$$\mathbb{E}(f(\sigma^{\nu}_{h;\exp(\lambda/h)})) \geq \langle f\rangle_h - C_1\exp(-C/h). \tag{3.1}$$

At this point there is also no loss in supposing that ν is concentrated on the configuration with all spins down, and therefore also that $\lambda_c < \lambda < 2\lambda_c$. In our argumentation this second assumption will simplify things.

In Section 3.2 we will introduce certain space-time structures, which we call inverted pyramids, and which will be used to obtain statements concerning droplet growth. To be able to use these results in order to prove (3.1), we will need to use such inverted pyramids as building blocks of a rescaling procedure, and will also need to obtain a mathematical counterpart to the notion of droplet creation at the correct rate; these two topics will be covered in Section 3.3. In both sections, 3.2

and 3.3, we will need to use some lemmas which can be seen as rigorous counterparts to the notion that the function $\phi(b) = wb - m^* b^2$ gives the free-energy of optimally-shaped droplets, and that equilibrium distributions can be studied based on this heuristics. To avoid distracting the reader with the technicalities behind these lemmas, they are only presented later in the paper, in Section 3.4. Some of the results in that section were already contained in the paper [SS1], but here we will need substantial strengthenings of them; moreover the techniques used here will be different from those in [SS1] and so provide an alternative to parts of that paper. Finally some estimates on the spectral gap of the generator of the dynamics of some kinetic Ising models on some finite sets will also be needed in Sections 3.2 and 3.3. Again, those will be postponed to the final section, 3.5, in this part of the paper, in order to avoid distracting the reader's attention from the main ideas in Sections 3.2 and 3.3.

The remainder of the paper is written having in mind a reader who will be following it in the order in which the sections are presented. With this in mind we tried to motivate and explain heuristically in Sections 3.2 and 3.3 the results which are used there but will only be proven later, in Sections 3.4 and 3.5. Readers who prefer following the lemmas and propositions in a strictly logical order, and who do not worry about the motivation behind each lemma can read the sections in the following order: 3.4, 3.5, 3.2, 3.3. Considering this possibility and also the length of the paper, we repeat some definitions from Sections 3.2 and 3.3 in Sections 3.4 and 3.5.

3.2. Inverted pyramids and droplet growth. In this section we will introduce two propositions which are counterparts to the statement: "If we start with a large enough Wulff-shaped droplet of the (+)-phase in the midst of the (−)-phase, then it is likely to grow with a linear speed larger than any negative exponential of $1/h$". In the first of these propositions "large enough" will be substantially larger than just "supercritical" (it will mean that the droplet has a negative free-energy); in the second of these propositions this aspect will be improved, at the cost of extra technicalities and extra work in the proof.

We will need to generalize the notion of a process $(\sigma^{\cdot}_{\Lambda,\eta,h;t})_{t\geq 0}$ evolving in a box Λ, with boundary condition η. We will have to consider the time evolution to occur in boxes which may change with time. Since we will construct all of our generalized processes using the graphical construction, their definition is very elementary. The space-time regions with which we will be concerned will all be of the following type. Let $t_0 < t_1 < t_2 < ... < t_{N+1}$ be an increasing finite sequence of times, and $\Lambda_0, \Lambda_1, ..., \Lambda_N$ be finite subsets of $\mathbb{Z}^2$. Our space-time region is

$$\Diamond = \bigcup_{i=0}^{N} [t_i, t_{i+1}] \times \Lambda_i.$$

We will refer to $[t_0, t_1] \times \Lambda_0$ as the bottom cylinder of $\Diamond$ and to $[t_N, t_{N+1}] \times \Lambda_N$ as the top cylinder of $\Diamond$. The base of $\Diamond$ is $t_0 \times \Lambda_0$, while the top of $\Diamond$ is $t_{N+1} \times \Lambda_N$.

While the boundary condition could be very general, in this paper we will only need to consider (−)-boundary conditions in space-time. We will need to start the evolution from an arbitrary time $s \in [t_0, t_{N+1})$, from a space configuration

$\eta \in \Omega_{\Lambda_j,-}$, where j is defined by $s \in [t_j, t_{j+1})$. The following notation will be used to denote the process:

$$(\sigma^{s,\eta}_{\Diamond,-,h;t})_{t\geq s}.$$

Its definition is as follows. We freeze all spins outside of $\Diamond$ as -1 and at time s set the configuration to η. We use then the graphical construction with its standard rules to update spins inside $\Diamond$, after time s.

The basic-coupling inequalities generalize in obvious ways. For instance, if $\eta \leq \zeta$, and $-h(T) < h_1 \leq h_2 < h(T)$, then for all $t_0 \leq s \leq t$,

$$\sigma^{s,\eta}_{\Diamond,-,h_1;t} \leq \sigma^{s,\zeta}_{\Diamond,-,h_2;t},$$

and, in case $t_0 = 0$,

$$\sigma^{0,\eta}_{\Diamond,-,h_1;t} \leq \sigma^{\zeta}_{h_2;t}.$$

We explain next the basic type (up to space-time translations) of space-time regions $\Diamond$ which we will deal with. For reasons which will become clear, we will refer to such regions as "inverted pyramids". For $i = 0, ..., N-1$, the set Λ_{i+1} will be obtained from Λ_i by adding one site to it. In particular we will have $\Lambda_0 \subset \Lambda_1 \subset ... \subset \Lambda_N$. These sets Λ_i will all be simply-connected subsets of $\mathbb{Z}^2$ and have "approximate Wulff shapes", in the following technical sense. Given a positive number l_0, we say that a subset Λ of $\mathbb{Z}^2$ is l_0-quasi-Wulff-shaped in case it is connected, simply connected and

$$\Lambda((l-l_0)W) \ \subset \ \Lambda \ \subset \ \Lambda((l+l_0)W),$$

for some number l. An l with this property will be said to be a linear-size-parameter of for the quasi-Wulff-shaped set Λ (of course, more than one such value of l can exist). In our case there will be an l_0 for which each Λ_i, $i = 0, ..., N$ will be l_0-quasi-Wulff-shaped. The absolute constant l_0 will be fixed throughout our work, and will be omitted from the notation. We will identify our inverted pyramids by four non-negative parameters: $b_1 < b_2$, h and δ. The first three of these parameters are related to the space dimensions of the inverted pyramid, by setting $\Lambda_0 = \Lambda(\frac{b_1}{h}W)$ and $\Lambda_N = \Lambda(\frac{b_2}{h}W)$. The parameters δ and again h are related to the time dimensions, by setting $t_{i+1} - t_i = \exp(\delta/h)$, for $i = 0, ..., N$.

It is not hard to see that if $l_0 > 2$ then for all b_1, b_2 and h as above there is a sequence,

$$\mathrm{Seq}(b_1, b_2; h) = (\Lambda_0, ..., \Lambda_N),$$

of boxes with the required properties, and with

$$N = N(b_1, b_2; h) \leq C\left(\frac{b_2}{h}\right)^2 \tag{3.2}$$

for some finite constant C. (Of course, more than one such sequence typically exists, and we suppose that some rule to choose one from among them is being used). Given also δ, the notation

$$\nabla = \nabla(b_1, b_2; h; \delta)$$

will be used for the inverted pyramid which we described above and which also has $t_0 = 0$.

The technical counterpart to the idea of the growth of a droplet will be contained in an event that we define next. With b_1, b_2, h and δ fixed (and omitted from the notation in several places, for simplicity), and also an initial configuration $\eta \in \Omega_{\Lambda(\frac{b_1}{h}W),-}$ given, we let G_η be the following event:

$$G_\eta = \left\{ \sigma^{0,\eta}_{\nabla,-,h;t_{N+1}} = \sigma^{t_N,+}_{\nabla,-,h;t_{N+1}} \right\}.$$

Observe that it is clear from the basic-coupling inequalities that, regardless of what η is, $\sigma^{0,\eta}_{\nabla,-,h;t_{N+1}} \leq \sigma^{t_N,+}_{\nabla,-,h;t_{N+1}}$, so that G_η only really requires that the complementary inequality holds.

Informally, G_+ can be seen as the event of a droplet of $+1$ spins of linear size proportional to b_1/h at time 0 growing to become a droplet of linear size proportional to b_2/h at time $t_N = N\exp(\delta/h) \leq C\,(b_2/h)^2 \exp(\delta/h)$.

Recall from Section 1.5 that heuristically the free-energy of a Wulff-shaped droplet $(b/h)W$ of the $(+)$-phase is given by $\phi(b)/h$, where $\phi(b) = wb - m^* b^2$, and that B_c is the value of b which maximizes this function, while $B_0 = 2B_c$ is the value of b above which this function becomes negative. In Section 1.5 this was used to predict heristically the behavior of the relaxation time as $h \searrow 0$. Similarly, the free-energy of such droplets can also be used to predict the typical aspect of a Gibbs distribution $\mu_{\Lambda(\frac{b}{h}W),-,h}(\,\cdot\,)$, when h is small, and this will be of great relevance in this paper. When $b < B_0$, one should expect this Gibbs distribution to resemble the $(-)$-phase, since droplets of the $(+)$-phase would all have positive free-energies. On the other hand, when $b > B_0$, one should expect this Gibbs distribution to resemble the $(+)$-phase, separated from the $(-)$ spins at the boundary by a large contour, since a single droplet of the $(+)$-phase of the size of the system itself would have the lowest possible free-energy. Rigorous results of this type were obtained in [SS1]. Unfortunately, for our purposes in the current paper we will need technically stronger results than those in [SS1]. As mentioned in Section 3.1, these technical results will only be presented and proven in Section 3.4. If the reader accepts the picture which we just presented and justified heuristically as reasonable, then, he or she should have no difficulty believing in the specific statements which appear in this and in the next section and are proven only in Section 3.4.

Next we state the first of the two main claims of this section.

Proposition 3.2.1. *Given $B_0 < b_1 < b_2$ and $\delta > 0$ there are positive finite constants C_1 and C_2 such that for $h > 0$*

$$\int d\mu_{\Lambda(\frac{b_1}{h}W),-,h}(\eta)\mathbb{P}(G_\eta) \geq 1 - C_1 \exp(-C_2/h). \tag{3.3}$$

In particular

$$\mathbb{P}(G_+) \geq 1 - C_1 \exp(-C_2/h). \tag{3.4}$$

Moreover the choice of C_2 does not depend on δ and b_2 and it can be taken arbitrarily large, provided b_1 is large enough.

Observe that from the heuristic picture described before the last proposition, we can see that in (3.3) we are starting from a droplet of the $(+)$-phase, and the statement is that it is likely to grow at a speed which is controlled in a useful way. In comparison, for $B_c < b_1 < B_0$, (3.3) should be false, since then in the Gibbs distribution $\mu_{\Lambda(\frac{b_1}{h}W),-,h}(\,\cdot\,)$ no droplet of the (+)-phase should be present. On the other hand, we should expect (3.4) to be true also in this case, since there we are starting from a droplet not just of the (+)-phase, but actually a solid droplet of (+) spins, with a supercritical size. This claim is contained in the next proposition. In this proposition we will use the following object.

Recall the definition (2.10) of $\mathcal{R}$ and set

$$\widehat{\mu}_{\Lambda,-,h}(\,\cdot\,) = \mu_{\Lambda,-,h}(\,\cdot\,|\mathcal{R}^c).$$

This is the Gibbs measure conditioned on the presence of a supercritical droplet. The following should be expected to happen, based on the heuristics. When $\Lambda = \Lambda(\frac{b}{h}W)$ and $b > B_0$ the conditioning has no major effect; but if $B_c < b < B_0$, then the conditioning produces a droplet of the (+)-phase, of roughly the size of the whole system, separated from the (−) spins at the boundary by a large contour. This is so since a single droplet of the (+)-phase of the size of the system itself would have the lowest possible free-energy compatible with the conditioning.

It is important to note that in the next proposition, in addition to having to modify (3.3) by introducing the conditioning on $\mathcal{R}^c$, also the way in which δ can be chosen is different. The reason for this difference will be clarified in the proof of the proposition.

Proposition 3.2.2. *Given $B_c < b_1 < b_2$ there are positive finite constants δ_0, C_1 and C_2 such that if $0 < \delta < \delta_0$, for $h > 0$*

$$\int d\widehat{\mu}_{\Lambda(\frac{b_1}{h}W),-,h}(\eta)\mathbb{P}(G_\eta) \geq 1 - C_1\exp(-C_2/h). \tag{3.5}$$

In particular

$$\mathbb{P}(G_+) \geq 1 - C_1\exp(-C_2/h). \tag{3.6}$$

In the remainder of this section we will prove Propositions 3.2.1 and 3.2.2. Besides leaving some technical lemmas which concern equilibrium distributions to Section 3.4, also some technical results which concern the kinetic Ising models run in certain finite boxes, including those of the type of $\Lambda\left(\frac{b}{h}W\right)$, will have their proofs postponed to Section 3.5. These results, when used in the current section, will be heuristically motivated, though.

Proof of Proposition 3.2.1 (modulo results in Sections 3.4 and 3.5). The second claim, (3.4), follows from the first one, (3.3), and the basic-coupling inequalities. Our task is to prove (3.3) and the claims about the value of the constant C_2. There is no loss in supposing that h is small, and we will assume that $0 < h \leq 1$.

For $i = 1, ..., N$, and arbitrary $\zeta \in \Omega_{\Lambda_{i-1},-}$ set

$$G^i_\zeta = \left\{\sigma^{t_{i-1},\zeta}_{\nabla,-,h;t_{i+1}} = \sigma^{t_i,+}_{\nabla,-,h;t_{i+1}}\right\}.$$

Note that

$$G_\eta \;\supset\; G^1_\eta \cap \left(\cap_{i=2}^N G^i_+\right). \tag{3.7}$$

Our goal now is to prove that for some positive finite C_1 and C_2, as in the statement of the proposition, for $i = 1, ..., N$,

$$\int d\mu_{\Lambda_{i-1},-,h}(\zeta)\mathbb{P}((G^i_\zeta)^c) \leq C_1 \exp(-C_2/h). \tag{3.8}$$

In particular this implies, as in the first paragraph in this proof, that

$$\mathbb{P}((G^i_+)^c) \leq C_1 \exp(-C_2/h). \tag{3.9}$$

By putting together (3.8) (used for $i = 1$), (3.9) (used for $i = 2, ..., N$), (3.7) and (3.2) we obtain (3.3).

The sets Λ_{i-1} and Λ_i differ only in that the latter has one extra site, say x, that the former does not have. For an arbitrary $\zeta \in \Omega_{\Lambda_{i-1},-}$, we will need to compare $\mu_{\Lambda_{i-1},-,h}(\zeta)$ with $\mu_{\Lambda_i,-,h}(\zeta)$. For this purpose we introduce the notation $\mathcal{S}^-_x = \{\sigma : \sigma(x) = -1\}$ for the event that the spin at the site x is negative, and let

$$\alpha = \inf_{|h|\leq 1} \inf_{\xi\in\Omega} \mu_{\{0\},\xi,h}(\mathcal{S}^-_0) \tag{3.10}$$

be the largest lower bound on the probability of having in equilibrium a spin -1 at the origin, given any information about the other spins, for values of h in the arbitrarily chosen neighborhood $[-1, +1]$ of the origin. Clearly $\alpha > 0$ for each temperature $T > 0$. With this notation,

$$\mu_{\Lambda_i,-,h}(\zeta) = \mu_{\Lambda_i,-,h}(\mathcal{S}^-_x)\ \mu_{\Lambda_i,-,h}(\zeta|\mathcal{S}^-_x) \geq \alpha\ \mu_{\Lambda_{i-1},-,h}(\zeta),$$

which can be seen as a uniform estimate on a Radon-Nikodym derivative:

$$\sup_{\zeta\in\Omega_{\Lambda_{i-1},-}} \frac{d\mu_{\Lambda_{i-1},-,h}}{d\mu_{\Lambda_i,-,h}}(\zeta) \leq \frac{1}{\alpha}. \tag{3.11}$$

Using the stationarity of the processes $(\sigma^{\mu_{\Lambda,-,h}}_{\Lambda,-,h})_{t\geq 0}$, and (3.11) we obtain

$$\begin{aligned}
\int d\mu_{\Lambda_{i-1},-,h}(\zeta)\mathbb{P}\left((G^i_\zeta)^c\right) &= \int d\mu_{\Lambda_{i-1},-,h}(\zeta)\mathbb{P}\left(\sigma^{t_{i-1},\zeta}_{\nabla,-,h;t_{i+1}} \neq \sigma^{t_i,+}_{\nabla,-,h;t_{i+1}}\right) \\
&= \int d\mu_{\Lambda_{i-1},-,h}(\zeta)\mathbb{P}\left(\sigma^{t_i,\zeta}_{\nabla,-,h;t_{i+1}} \neq \sigma^{t_i,+}_{\nabla,-,h;t_{i+1}}\right) \\
&\leq \frac{1}{\alpha}\int d\mu_{\Lambda_i,-,h}(\zeta)\mathbb{P}\left(\sigma^{t_i,\zeta}_{\nabla,-,h;t_{i+1}} \neq \sigma^{t_i,+}_{\nabla,-,h;t_{i+1}}\right) \\
&= \frac{1}{\alpha}\int d\mu_{\Lambda_i,-,h}(\zeta)\mathbb{P}\left(\sigma^{\zeta}_{\Lambda_i,-,h;\exp(\delta/h)} \neq \sigma^{+}_{\Lambda_i,-,h;\exp(\delta/h)}\right).
\end{aligned} \tag{3.12}$$

This may have seen at first sight as a minor and trivial maneuver, but it is actually a central step in our approach towards controlling droplet growth. We have just transformed our problem pertaining to "growth" into a problem pertaining to "rapid loss of memory" or, in other words, "rapid convergence to equilibrium", since (3.12) will provide us with the aimed (3.8), once we show that

$$\int d\mu_{\Lambda_i,-,h}(\zeta)\mathbb{P}\left(\sigma^{\zeta}_{\Lambda_i,-,h;\exp(\delta/h)} \neq \sigma^{+}_{\Lambda_i,-,h;\exp(\delta/h)}\right) \leq C_1 \exp(-C_2/h). \qquad (3.13)$$

Proving (3.13) seems like a standard problem, due to the vast current literature on this type of issue: we have a reversible Markov process $(\sigma^{+}_{\Lambda_i,-,h;t})_{t\geq 0}$, and we want to show that it reaches equilibrium in a time of the order of $\exp(\delta/h)$. There are nevertheless still major hurdles to overcome. The standard approach to such a problem starts with the derivation of a lower bound on the spectral gap of the generator of the process. The result that we are after would then follow if the time with which we are concerned were much larger that the inverse of the lower bound on the spectral gap. Such an approach is nevertheless unfeasible in our case, due to the fact that here the spectral gap is of the order of $\exp(-\beta A/h)$, so that its inverse is much larger than $\exp(\delta/h)$, when δ is small. We will not give the full proof of this claim on the value of the spectral gap, since it will be of no use for us (some readers may want to take it as an exercise, with the hint that it can be solved using techniques in this paper), and will limit ourselves to explaining the nature of the difficulty at the heuristic level. This difficulty lies precisely in the sort of metastability studied in this paper. We are considering a Glauber dynamics in the box Λ_i, which is almost the same as a set $\Lambda(\frac{b}{h}W)$ with some $b \in [b_1, b_2]$. Since $B_0 < b_1 \leq b$, in equilibrium we should have a large droplet of the (+)-phase, covering the box Λ_i almost entirely. If we look at the process started from all spins down, $(\sigma^{-}_{\Lambda_i,-,h;t})_{t\geq 0}$, then to reach equilibrium this big droplet has to be formed, and the system has to go through the bottleneck presented by the situation with a critical droplet. Hence the free-energy barrier to be overcome has height A/h. The system should then reach equilibrium in a time of order $\exp(\beta A/h)$. (Because the linear size of the box is of the order of $1/h$, droplet growth is of no relevance for the estimate of the order of magnitude of the relaxation time inside the box.) Starting with all spins down should maximize the relaxation time, since equilibrium is basically the (+)-phase, and the inverse of this relaxation time should hence give the order of magnitude of the spectral gap.

The same heuristics above which pointed out the problem with using the spectral gap for our purpose of proving (3.13) indicates also why we should believe that this inequality holds nevertheless. The difficulty pointed out concerns the long time needed to relax towards equilibrium if the process is started with all spins down. We are, on the other hand, concerned with the case in which we start with all spins up, much closer to equilibrium. One can talk of an heuristic picture with a double well structure. The configuration with all spins down is in the higher (metastable) well, separated from the other one by the free-energy barrier, but the configuration with all spins up is inside the deeper (stable) well, and our problem concerns only relaxation inside of this well. The problem still remains at this point of how to exploit this heuristic picture and prove (3.13). The solution will be to use the basic-coupling inequalities in order to compare our process with some modified ones, for which the spectral gaps can be proven to be large enough for our purposes. In doing so we were inspired by arguments in Section 5 of [Mar].

One of the two comparison processes that we will use is $(\sigma^{\cdot}_{\Lambda_i,+,h;t})_{t\geq 0}$, in which (+)-boundary conditions are used. The other one will be denoted by $\left(\sigma^{\cdot}_{\Lambda_i\setminus\Lambda^{\mathrm{core}},(+,-),h;t}\right)_{t\geq 0}$, and has the following meaning. The box in which it is run is the annulus $\Lambda_i\setminus\Lambda^{\mathrm{core}}$, where $\Lambda^{\mathrm{core}} = \Lambda(\frac{B_c+B_0}{2h}W)$, and the boundary condition denoted by $(+,-)$ refers to freezing the spins up inside the core Λ^{core} and down outside Λ_i. In Section 3.5, Propositions 3.5.1 and 3.5.2, we will show that for any $\delta > 0$ the generators of these two processes satisfy

$$\inf_{i=0,\ldots,N} \ \mathrm{gap}(\Lambda_i,+,h) \ \geq \ \exp\left(-\frac{\delta}{2h}\right), \tag{3.14}$$

$$\inf_{i=0,\ldots,N} \ \mathrm{gap}\left(\Lambda_i\setminus\Lambda^{\mathrm{core}},(+,-),h\right) \ \geq \ \exp\left(-\frac{\delta}{2h}\right), \tag{3.15}$$

for small enough h. The intuitive reason behind these relatively large spectral gaps, is that the extra (+)'s introduced as boundary conditions eliminate metastability in the time evolution of these two processes. In equilibrium these systems are again basically in the (+)-phase, and if we start them with all spins down, then there is no need to nucleate a critical droplet in order to relax to equilibrium. In the first case the (+)-phase drifts inwards from the (+)-boundary towards the center. In the second case a supercritical (+)-droplet is frozen by hand in the center of the box, so that the relaxation is a "down-hill" movement on the free-energy landscape; the (+)-phase should drift outwards from the center towards the outer boundary of the box.

Our task now is to show how (3.14) and (3.15) can be used to derive (3.13). We partition Λ_i into two sets:

$$\Lambda^{\mathrm{in}} = \Lambda\left(\frac{b_1+B_0}{2h}W\right),$$
$$\Lambda_i^{\mathrm{out}} = \Lambda_i\setminus\Lambda^{\mathrm{in}}.$$

Using the basic-coupling inequalities, we have

$$\begin{aligned}
&\int d\mu_{\Lambda_i,-,h}(\zeta)\mathbb{P}\left(\sigma^{\zeta}_{\Lambda_i,-,h;\exp(\delta/h)} \neq \sigma^{+}_{\Lambda_i,-,h;\exp(\delta/h)}\right)\\
&\leq \int d\mu_{\Lambda_i,-,h}(\zeta)\sum_{y\in\Lambda_i}\left\{\mathbb{P}\left(\sigma^{+}_{\Lambda_i,-,h;\exp(\delta/h)}(y)=+1\right)-\mathbb{P}\left(\sigma^{\zeta}_{\Lambda_i,-,h;\exp(\delta/h)}(y)=+1\right)\right\}\\
&=\sum_{y\in\Lambda_i}\left\{\mathbb{P}\left(\sigma^{+}_{\Lambda_i,-,h;\exp(\delta/h)}(y)=+1\right)-\mu_{\Lambda_i,-,h}(\{\sigma:\sigma(y)=+1\})\right\}\\
&\leq\sum_{y\in\Lambda^{\mathrm{in}}}\left\{\mathbb{P}\left(\sigma^{+}_{\Lambda_i,+,h;\exp(\delta/h)}(y)=+1\right)-\mu_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}\right)\right\}\\
&+\sum_{y\in\Lambda_i^{\mathrm{out}}}\left\{\mathbb{P}\left(\sigma^{+}_{\Lambda_i\setminus\Lambda^{\mathrm{core}},(+,-),h;\exp(\delta/h)}(y)=+1\right)-\mu_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}\right)\right\}.
\end{aligned}$$

At this point we can use (3.14) and (3.15) in a standard way. For instance, from inequality (59) in [Sch1],

$$\left|\mathbb{P}\left(\sigma^{+}_{\Lambda_i,+,h;\exp(\delta/h)}(y)=+1\right)-\mu_{\Lambda_i,+,h}\left(\{\sigma:\sigma(y)=+1\}\right)\right|$$
$$\leq \frac{e^{-\exp(\delta/h)\mathrm{gap}(\Lambda_i,+,h)}}{\mu_{\Lambda_i,+,h}(+)}\leq \exp\left(C\left(\frac{b_2}{h}\right)^2\right)\exp\{-e^{\frac{\delta}{2h}}\}$$

and

$$\left|\mathbb{P}\left(\sigma^{+}_{\Lambda_i\setminus\Lambda^{\mathrm{core}},(+,-),h;\exp(\delta/h)}(y)=+1\right)-\mu_{\Lambda_i\setminus\Lambda^{\mathrm{core}},(+,-),h}\left(\{\sigma:\sigma(y)=+1\}\right)\right|$$
$$\leq \frac{e^{-\exp(\delta/h)\mathrm{gap}(\Lambda_i\setminus\Lambda^{\mathrm{core}},(+,-),h)}}{\mu_{\Lambda_i\setminus\Lambda^{\mathrm{core}},(+,-),h}(+)}\leq \exp\left(C\left(\frac{b_2}{h}\right)^2\right)\exp\{-e^{\frac{\delta}{2h}}\}.$$

Combining the last three displayed inequalities, we obtain

$$\begin{aligned}
&\int d\mu_{\Lambda_i,-,h}(\zeta)\mathbb{P}\left(\sigma^{\zeta}_{\Lambda_i,-,h;\exp(\delta/h)}\neq\sigma^{+}_{\Lambda_i,-,h;\exp(\delta/h)}\right)\\
&\quad\leq\sum_{y\in\Lambda^{\mathrm{in}}}\left\{\mu_{\Lambda_i,+,h}\left(\{\sigma:\sigma(y)=+1\}\right)-\mu_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}\right)\right\}\\
&\quad+\sum_{y\in\Lambda_i^{\mathrm{out}}}\left\{\mu_{\Lambda_i\setminus\Lambda^{\mathrm{core}},(+,-),h}\left(\{\sigma:\sigma(y)=+1\}\right)-\mu_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}\right)\right\}\\
&\quad+C'\left(\frac{b_2}{h}\right)^2\exp\left(C\left(\frac{b_2}{h}\right)^2\right)\ \exp\{-e^{\frac{\delta}{2h}}\}.
\end{aligned}\tag{3.16}$$

Our remaining problem concerns only equilibrium. We will use a result from Section 3.4, but observe that this result is intuitively natural, based on the heuristics presented immediately before the statement of Proposition 3.2.1. Set

$$\mathcal{B}_1=\left\{\Lambda^{\mathrm{in}}\overset{-*}{\longleftrightarrow}\left(\Lambda\left(\frac{b_1}{h}W\right)\right)^c\right\},\tag{3.17}$$

$$\mathcal{B}_2=\left\{\Lambda^{\mathrm{core}}\overset{-*}{\longleftrightarrow}\left(\Lambda^{\mathrm{in}}\right)^c\right\},\tag{3.18}$$

From Lemma 3.4.8 we know that for every $b_1>B_0$ there exists a $C_2=C_2(b_1)>0$, $C_2\to\infty$ as $b_1\to\infty$ and a finite C_1 such that for $j=1,2$,

$$\sup_{i=0,\ldots,N}\mu_{\Lambda_i,-,h}(\mathcal{B}_j)\leq C_1\exp(-C_2/h).\tag{3.19}$$

From (1.20), for each $y\in\Lambda^{\mathrm{in}}$,

$$\mu_{\Lambda_i,+,h}\left(\{\sigma:\sigma(y)=+1\}\right)-\mu_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}\right)\leq 2\mu_{\Lambda_i,-,h}\left(\mathcal{B}_1\right).\tag{3.20}$$

Similarly, for each $y\in\Lambda_i^{\mathrm{out}}$,

$$\mu_{\Lambda_i\setminus\Lambda^{\mathrm{core}},(+,-),h}\left(\{\sigma:\sigma(y)=+1\}\right)-\mu_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}\right)\leq 2\mu_{\Lambda_i,-,h}\left(\mathcal{B}_2\right).\tag{3.21}$$

The desired inequality (3.3) and the claims about the choice of the constant C_2 which appears there follow from combining (3.16), (3.20), (3.21) and (3.19)

■

Proof of Proposition 3.2.2 (modulo results in Sections 3.4 and 3.5). We will explain how the proof of Proposition 3.2.1 can be adapted to prove this proposition. There are several extra complications, since $\widehat{\mu}_{\Lambda_i,-,h}$ is not an invariant distribution for the process $(\sigma^{\cdot}_{\Lambda_i,-,h})_{t\geq 0}$, and in particular this is the reason for which we will have to choose δ_0 small enough. The idea is to look at this distribution instead as a "metastable state" for this process, and to use techniques from Part 2 of the present paper in this connection.

Using the graphical construction we can define the following processes restricted to $\mathcal{R}^c$. For arbitrary $s \in [t_0, t_{N+1})$, and $\eta \in \Omega_{\Lambda_j,-} \cap \mathcal{R}^c$, where j is defined by $s \in [t_j, t_{j+1})$, the process

$$(\widehat{\sigma}^{s,\eta}_{\nabla,-,h;t})_{t\geq s}.$$

is obtained in the following simple way. We freeze all spins outside of ∇ as -1 and at time s set the configuration to η. We use then the graphical construction with its standard rules modified by suppressing jumps which would bring the system to $\mathcal{R}$, to update spins inside ∇ after time s.

Bottlenecks for this dynamics are the sets

$$\mathcal{F}_x^+ = \{\sigma : \sigma^x \in \mathcal{F}_x^-\}.$$

Define also

$$\mathcal{F}_x = \mathcal{F}_x^- \cup \mathcal{F}_x^+.$$

For each x the event $\mathcal{F}_x$ depends only on the spins at sites other than x. If $|h| \leq 1$, we have from the definition (3.10) of α, that for $i = 0, ..., N$ and $x \in \Lambda_i$,

$$\mu_{\Lambda_i,-,h}(\mathcal{F}_x^-) \geq \alpha\ \mu_{\Lambda_i,-,h}(\mathcal{F}_x). \tag{3.22}$$

Using this inequality in combination with the bottleneck estimate (2.41) we have that for arbitrary $\epsilon > 0$,

$$\begin{aligned}\widehat{\mu}_{\Lambda_i,-,h}(\mathcal{F}_x^+) &= \frac{\mu_{\Lambda_i,-,h}(\mathcal{F}_x^+)}{\mu_{\Lambda_i,-,h}(\mathcal{R}^c)} \leq \frac{\mu_{\Lambda_i,-,h}(\mathcal{F}_x)}{\mu_{\Lambda_i,-,h}(\mathcal{R}^c)} \leq \frac{\frac{1}{\alpha}\mu_{\Lambda_i,-,h}(\mathcal{F}_x^-)}{\mu_{\Lambda_i,-,h}(\mathcal{R}^c)} \\ &= \frac{1}{\alpha}\tilde{\mu}_{\Lambda_i,-,h}(\mathcal{F}_x^-)\frac{\mu_{\Lambda_i,-,h}(\mathcal{R})}{\mu_{\Lambda_i,-,h}(\mathcal{R}^c)} \leq \frac{1}{\alpha}\frac{C\exp\left(-\beta(1-\epsilon)\frac{A}{h}\right)}{\mu_{\Lambda_i,-,h}(\mathcal{R}^c)},\end{aligned}$$

for small h, independent of i. Since $\Lambda\left(\frac{b_1}{h}W\right) \subset \Lambda_i$ and $b_1 > B_c$, Lemma 3.4.4 gives us for some $\delta_0 > 0$,

$$\mu_{\Lambda_i,-,h}(\mathcal{R}^c) \geq \exp\left(-\beta\frac{A}{h} + \frac{2\delta_0}{h}\right),$$

for small h. So, if our choice of ϵ above is made properly, we obtain

$$\widehat{\varphi} = \sup_{i=0,\ldots,N}\ \sup_{x\in\Lambda_i}\ \widehat{\mu}_{\Lambda_i,-,h}(\mathcal{F}_x^+)\ \leq\ C_1\exp\left(-\frac{\delta_0}{h}\right), \tag{3.23}$$

for some finite C_1.

Define now

$$\tau_i^\zeta = \inf\{t \geq t_i : \text{the process } (\widehat{\sigma}^{t_i,\zeta}_{\nabla,-,h;t}) \text{ has a suppressed jump at time } t\}.$$

Clearly

$$\widehat{\sigma}^{t_i,\zeta}_{\nabla,-,h;t} = \sigma^{t_i,\zeta}_{\nabla,-,h;t} \quad \text{for} \quad t_i \leq t < \tau_i^\zeta. \tag{3.24}$$

From the same argument used to prove (2.12), we obtain now, using (3.23), for $i = 0, ..., N$, and $\delta < \delta_0$,

$$\begin{aligned} \int d\widehat{\mu}_{\Lambda_i,-,h}(\zeta)\mathbb{P}(\tau_i^\zeta \leq t_{i+1}) &\leq C_2 \exp\left(-C_3 e^{\delta/h}\right) + C_4 \left|\Lambda\left(\frac{b_2}{h}W\right)\right| e^{\delta/h}\,\widehat{\varphi} \\ &\leq C_5 \exp(-C_6/h), \end{aligned} \tag{3.25}$$

for small h. (We remind the reader that $t_{i+1} = t_i + e^{\delta/h}$.)

Before we can proceed with the adaptation of the proof of Proposition 3.2.1 to prove Proposition 3.2.2, we need to derive an analogue to (3.11). We will show that for small enough h, for $i = 1, ..., N$,

$$\sup_{\zeta \in \Omega_{\Lambda_{i-1},-}} \frac{d\widehat{\mu}_{\Lambda_{i-1},-,h}}{d\widehat{\mu}_{\Lambda_i,-,h}}(\zeta) \leq \frac{2}{\alpha}. \tag{3.26}$$

To this end, as in the argumentation for (3.11), we will use the notation $x = \Lambda_i \backslash \Lambda_{i-1}$, and $\mathcal{S}_x^- = \{\sigma : \sigma(x) = -1\}$. First note now that by partitioning $(\mathcal{F}_x)^c \cap \mathcal{R}^c$ according to what the configuration in Λ_{i-1} is and denoting by $\{E_j\}$ the resulting parts, we have

$$\begin{aligned} \widehat{\mu}_{\Lambda_i,-,h}(\mathcal{S}_x^- | (\mathcal{F}_x)^c) &= \sum_j \widehat{\mu}_{\Lambda_i,-,h}(E_j | (\mathcal{F}_x)^c)\, \widehat{\mu}_{\Lambda_i,-,h}(\mathcal{S}_x^- | E_j) \\ &= \sum_j \widehat{\mu}_{\Lambda_i,-,h}(E_j | (\mathcal{F}_x)^c)\, \mu_{\Lambda_i,-,h}(\mathcal{S}_x^- | E_j) \\ &\geq \sum_j \widehat{\mu}_{\Lambda_i,-,h}(E_j | (\mathcal{F}_x)^c)\, \alpha = \alpha. \end{aligned}$$

Therefore, using (3.23),

$$\begin{aligned} \widehat{\mu}_{\Lambda_i,-,h}(\mathcal{S}_x^-) &\geq \widehat{\mu}_{\Lambda_i,-,h}((\mathcal{F}_x)^c)\ \widehat{\mu}_{\Lambda_i,-,h}(\mathcal{S}_x^- | (\mathcal{F}_x)^c) \\ &\geq \left(1 - \widehat{\mu}_{\Lambda_i,-,h}(\mathcal{F}_x^+)\right) \widehat{\mu}_{\Lambda_i,-,h}(\mathcal{S}_x^- | (\mathcal{F}_x)^c) \geq \frac{\alpha}{2}, \end{aligned}$$

for small enough h, uniformly in i. Since $\zeta(x) = -1$, we have

$$\begin{aligned} \widehat{\mu}_{\Lambda_i,-,h}(\zeta) &= \widehat{\mu}_{\Lambda_i,-,h}(\mathcal{S}_x^-)\ \widehat{\mu}_{\Lambda_i,-,h}(\zeta | \mathcal{S}_x^-) \\ &= \widehat{\mu}_{\Lambda_i,-,h}(\mathcal{S}_x^-)\ \widehat{\mu}_{\Lambda_{i-1},-,h}(\zeta) \geq \frac{\alpha}{2}\ \widehat{\mu}_{\Lambda_{i-1},-,h}(\zeta), \end{aligned}$$

completing the proof of (3.26).

We are now ready to explain how the proof of Proposition 3.2.1 can be modified to prove Proposition 3.2.2.

In place of (3.8), we have to prove the analogous statement:

$$\int d\widehat{\mu}_{\Lambda_{i-1},-,h}(\zeta)\mathbb{P}((G^i_\zeta)^c) \leq C_1 \exp(-C_2/h). \tag{3.27}$$

For this we use (3.24), (3.25), the stationarity of $(\widehat{\sigma}^{t_{i-1},\widehat{\mu}_{\Lambda_{i-1},-,h}}_{\nabla,-,h;t})_{t_{i-1}\leq t\leq t_i}$, and (3.26) to obtain the following replacement of (3.12)

$$\begin{aligned}
&\int d\widehat{\mu}_{\Lambda_{i-1},-,h}(\zeta)\mathbb{P}\left((G^i_\zeta)^c\right) = \int d\widehat{\mu}_{\Lambda_{i-1},-,h}(\zeta)\mathbb{P}\left(\sigma^{t_{i-1},\zeta}_{\nabla,-,h;t_{i+1}} \neq \sigma^{t_i,+}_{\nabla,-,h;t_{i+1}}\right)\\
&\leq \int d\widehat{\mu}_{\Lambda_{i-1},-,h}(\zeta)\mathbb{P}\left(\sigma^{t_i,\zeta}_{\nabla,-,h;t_{i+1}} \neq \sigma^{t_i,+}_{\nabla,-,h;t_{i+1}}\right) + \int d\widehat{\mu}_{\Lambda_{i-1},-,h}(\zeta)\mathbb{P}(\tau^\zeta_{i-1}\leq t_i)\\
&\leq \frac{2}{\alpha}\int d\widehat{\mu}_{\Lambda_i,-,h}(\zeta)\mathbb{P}\left(\sigma^{t_i,\zeta}_{\nabla,-,h;t_{i+1}} \neq \sigma^{t_i,+}_{\nabla,-,h;t_{i+1}}\right) + C_5\exp(-C_6/h).
\end{aligned} \tag{3.28}$$

We can show that (3.28) leads to (3.27) by adapting the steps used to prove that (3.12) leads to (3.8). The following are the changes in the argument.

This time we take

$$\Lambda^{\text{core}} = \Lambda\left(\frac{2B_c+b_1}{3h}W\right),$$
$$\Lambda^{\text{in}} = \Lambda\left(\frac{B_c+2b_1}{3h}W\right),$$

and

$$\Lambda^{\text{out}}_i = \Lambda_i\backslash\Lambda^{\text{in}}.$$

Similarly to the derivations of (3.16) and (3.28), we can use the basic-coupling inequalities, the spectral gap estimates in Propositions 3.5.1 and 3.5.2, (3.24), (3.25), and the stationarity of $(\widehat{\sigma}^{t_i,\widehat{\mu}_{\Lambda_i,-,h}}_{\nabla,-,h;t})_{t_i\leq t\leq t_{i+1}}$ to derive

$$
\begin{aligned}
&\int d\widehat{\mu}_{\Lambda_i,-,h}(\zeta)\mathbb{P}\left(\sigma^{t_i,\zeta}_{\nabla,-,h;t_{i+1}} \neq \sigma^{t_i,+}_{\nabla,-,h;t_{i+1}}\right)\\
&\leq \int d\widehat{\mu}_{\Lambda_i,-,h}(\zeta)\sum_{y\in\Lambda_i}\left\{\mathbb{P}\left(\sigma^{t_i,+}_{\nabla,-,h;t_{i+1}}(y)=+1\right)-\mathbb{P}\left(\sigma^{t_i,\zeta}_{\nabla,-,h;t_{i+1}}(y)=+1\right)\right\}\\
&\leq \sum_{y\in\Lambda^{\mathrm{in}}}\left\{\mathbb{P}\left(\sigma^{+}_{\Lambda_i,+,h;\exp(\delta/h)}(y)=+1\right)-\widehat{\mu}_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}\right)\right\}\\
&+\sum_{y\in\Lambda_i^{\mathrm{out}}}\left\{\mathbb{P}\left(\sigma^{+}_{\Lambda_i\setminus\Lambda^{\mathrm{core}},(+,-),h;\exp(\delta/h)}(y)=+1\right)-\widehat{\mu}_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}\right)\right\}\\
&+|\Lambda_i|\int d\widehat{\mu}_{\Lambda_i,-,h}(\zeta)\mathbb{P}(\tau_i^{\zeta}\leq t_{i+1})\\
&\leq \sum_{y\in\Lambda^{\mathrm{in}}}\left\{\mu_{\Lambda_i,+,h}\left(\{\sigma:\sigma(y)=+1\}\right)-\widehat{\mu}_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}\right)\right\}\\
&+\sum_{y\in\Lambda_i^{\mathrm{out}}}\left\{\mu_{\Lambda_i\setminus\Lambda^{\mathrm{core}},(+,-),h}\left(\{\sigma:\sigma(y)=+1\}\right)-\widehat{\mu}_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}\right)\right\}\\
&+C_7\left(\frac{b_2}{h}\right)^2\left\{C_5\exp(-C_6/h)\;+\;\exp\left(C_8\left(\frac{b_2}{h}\right)^2\right)\exp\{-e^{\frac{\delta}{2h}}\}\right\}. \qquad (3.29)
\end{aligned}
$$

The definitions of $\mathcal{B}_1$ and $\mathcal{B}_2$ are the same as before (see (3.17) and (3.18)), but with the modified choices above of Λ^{core} and Λ^{in}. From Lemma 3.4.8 we know that for $j=1,2$,

$$\sup_{i=0,\dots,N}\widehat{\mu}_{\Lambda_i,-,h}(\mathcal{B}_j)\leq C_9\exp(-C_{10}/h), \qquad (3.30)$$

Due to the conditioning in the definition of $\widehat{\mu}_{\Lambda_i,-,h}$, the derivation of the analogues of (3.20) and (3.21) are somewhat more delicate. For $y\in\Lambda^{\mathrm{in}}$, we let $\{E_j\}$ denote the partition of $(\mathcal{B}_1)^c$ according to what the $(-,*)$-cluster of $\left(\Lambda\left(\frac{b_1}{h}W\right)\right)^c$ is. We obtain the following:

$$
\begin{aligned}
\widehat{\mu}_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}|(\mathcal{B}_1)^c\right)&=\sum_j\alpha_j\;\widehat{\mu}_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}|E_j\right)\\
&=\sum_j\alpha_j\;\mu_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}|E_j\right)\\
&\geq \mu_{\Lambda_i,+,h}\left(\{\sigma:\sigma(y)=+1\}\right)
\end{aligned}
$$

where in the second equality we used the fact that for each j, $E_j\subset\mathcal{R}^c$, and in the final inequality we used the same standard argument which gives rise to (1.19) and the fact that $\sum_j\alpha_j=1$. From (1.18) it follows then that

$$\mu_{\Lambda_i,+,h}\left(\{\sigma:\sigma(y)=+1\}\right)-\widehat{\mu}_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}\right)\leq 2\widehat{\mu}_{\Lambda_i,-,h}\left(\mathcal{B}_1\right). \qquad (3.31)$$

Similarly we can derive, for each $y\in\Lambda_i^{\mathrm{out}}$,

$$\mu_{\Lambda_i\setminus\Lambda^{\mathrm{core}},(+,-),h}\left(\{\sigma:\sigma(y)=+1\}\right)-\widehat{\mu}_{\Lambda_i,-,h}\left(\{\sigma:\sigma(y)=+1\}\right)\leq 2\widehat{\mu}_{\Lambda_i,-,h}\left(\mathcal{B}_2\right). \qquad (3.32)$$

Our goal, (3.27), follows from combining (3.28), (3.29), (3.30), (3.31) and (3.32). ■

3.3. Rescaling and droplet creation. The inverted pyramid $\nabla = \nabla(b_1, b_2; h; \delta)$ and the event G_+ were conceived having in mind their use in a rescaling procedure. To each point $k = (k_1, k_2, k_3)$ of the rescaled space time $\mathbb{Z}^2 \times \mathbb{Z}_+$ we associate the following translate of ∇:

$$\nabla_k = \left(\frac{b'}{h}k_1, \frac{b'}{h}k_2, t_N k_3\right) + \nabla,$$

where as before N is 1 less than the number of elements in the sequence $\mathrm{Seq}(b_1, b_2; h)$, $t_N = N \exp(\delta/h)$ and $b' > 0$ is a new parameter. We suppose that b_1, b_2 and b' are such that

$$\left(\Lambda\left(\frac{b_1}{h}W\right) + \left(\frac{b'}{h}, 0\right)\right) \bigcap \Lambda\left(\frac{b_1}{h}W\right) = \emptyset, \tag{3.33}$$

and

$$\Lambda\left(\frac{b_1}{h}W\right) + \left(\frac{b'}{h}, 0\right) \subset \Lambda\left(\frac{b_2}{h}W\right). \tag{3.34}$$

It will be important that this can be done for arbitrarily large values of b_1. Indeed, with b_1 given we can choose b' for (3.33) to hold, and then choose b_2 for (3.34) to hold.

Of course, we will sometimes confuse inverted pyramids with their indices in the terminology being introduced below, and no inconvenience should arise from this.

We say that two inverted pyramids ∇_k and $\nabla_{k'}$ are neighbors in case $|k_3 - k_3'| = 1$ and $(k_1, k_2) - (k_1', k_2') \in \{(0,0), (0,1), (1,0), (0,-1), (-1,0)\}$.

Note that $\nabla = \nabla_{(0,0,0)}$ has exactly 5 neighbors, all at rescaled time 1. The bottom cylinders of these 5 inverted pyramids will be pairwise disjoint and each will be contained in the top cylinder of the inverted pyramid $\nabla_{(0,0,0)}$.

A rescaled-space-time oriented chain will be a sequence $(k^{(1)}, ..., k^{(n)})$ of elements of $\mathbb{Z}^2 \times \mathbb{Z}^+$ such that $k^{(i)}$ and $k^{(i+1)}$ are neighbors and $k_3^{(i+1)} = k_3^{(i)} + 1$, for $i = 1, ..., n-1$. The start of the chain is $k^{(1)}$, and its end is $k^{(n)}$.

We will say that the inverted pyramid ∇_k is open if the corresponding event G_+ happens for it. This event will be denoted by $G_{+,k}$. More formally, $G_{+,k}$ is the set of realizations of the graphical construction which would be in G_+ after space and time were translated by the amount $-(\frac{b'}{h}k_1, \frac{b'}{h}k_2, t_N k_3)$.

The events $G_{+,k}$ have the very nice property that they are well suited for being concatenated. The simplest version of this idea is the following. Suppose that $(k^{(1)}, ..., k^{(n)})$ is a rescaled-space-time oriented chain. We will be concerned with the process $(\sigma^{\cdot}_{\cup_i \nabla_{k^{(i)}}, -, h; t})$. To fix some notation, say that the bottom cylinder of the bottom inverted pyramid, $\nabla_{k^{(1)}}$, is the set $\Lambda^{\mathrm{bot}} \times [s_0, s_1]$ and that the top cylinder of the top inverted pyramid, $\nabla_{k^{(n)}}$, is the set $\Lambda^{\mathrm{top}} \times [s_3, s_4]$. Standard applications of the basic-coupling inequalities yield

$$\cap_{i=1,\ldots,n} G_{+,k^{(i)}} \subset \left\{\sigma^{s_0,+}_{\cup_i \nabla_{k^{(i)}}, -, h; s_4} = \sigma^{s_3,+}_{\cup_i \nabla_{k^{(i)}}, -, h; s_4}\right\}.$$

Pictorially this amounts to a droplet of the $(+)$-phase flowing through the open tube $\cup_i \nabla_{k^{(i)}}$. This relation is not yet strong enough for our purposes, because it requires a solid blob of $(+)$ spins at the bottom, and we will not have such a solid droplet of (+)'s. Nevertheless the following stronger version of this relation can be derived via the same standard use of the basic-coupling inequalities. For an arbitrary $\eta \in \Omega_{\Lambda^{\mathrm{bot}},-}$

$$\left\{\sigma^{s_0,\eta}_{\nabla_{k(1)},-,h;s_1} = \sigma^{s_0,+}_{\nabla_{k(1)},-,h;s_1}\right\} \bigcap \cap_{i=1,\ldots,n} G_{+,k^{(i)}} \subset \left\{\sigma^{s_0,\eta}_{\cup_i \nabla_{k(i)},-,h;s_4} = \sigma^{s_3,+}_{\cup_i \nabla_{k(i)},-,h;s_4}\right\}. \tag{3.35}$$

We will now construct a space-time structure motivated by the cone which appeared in the heuristics in Section 1.5. Our goal is to look at the system at time $\exp(\lambda/h)$, with $\lambda_c \leq \lambda \leq 2\lambda_c$, and prove (3.1). Set

$$M = \lfloor \exp(\lambda/h)/t_N \rfloor.$$

Moving backwards in time we choose now a set $\mathcal{C}_M$ of inverted pyramids ∇_k. From the inverted pyramids with rescaled time coordinate $M-1$ we take only $\nabla_{(0,0,M-1)}$. Inductively, once we have selected the inverted pyramids at a certain rescaled time $m > 0$, we take at rescaled time $m-1$ the inverted pyramids which are neighbors to at least one inverted pyramid already included in our set. The procedure stops at rescaled time 0. Note that if we consider the indices of the selected inverted pyramids, what we have done is precisely to construct the discrete rescaled analog of the space-time cone, as in the heuristics. An alternative definition of $\mathcal{C}_M$ is that it is the set of inverted pyramids from which we can start a rescaled space-time oriented chain which ends at $\nabla_{(0,0,M-1)}$.

The cardinality of the set $\mathcal{C}_M$ clearly satisfies

$$C_1 M^3 \leq |\mathcal{C}_M| \leq C_2 M^3.$$

On the other hand, the bounds on λ and the fact that, due to (3.2),

$$t_N \leq C \left(\frac{b_2}{h}\right)^2 \exp\left(\frac{\delta}{h}\right),$$

give us, for small h,

$$C' \left(\frac{h}{b_2}\right)^2 \exp\left(\frac{\lambda-\delta}{h}\right) \leq M \leq \exp\left(\frac{2\lambda_c}{h}\right).$$

Therefore

$$C_3 \left(\frac{h}{b_2}\right)^6 \exp\left(\frac{3(\lambda-\delta)}{h}\right) \leq |\mathcal{C}_M| \leq C_4 \exp\left(\frac{6\lambda_c}{h}\right).$$

Let us choose δ as

$$\delta = \frac{\lambda - \lambda_c}{2}, \tag{3.36}$$

which means that $\lambda - \delta = \lambda_c + \delta$. Then, since $\lambda_c = \beta A/3$, we have

$$C_3 \left(\frac{h}{b_2}\right)^6 \exp\left(\frac{3\delta}{h}\right) \exp\left(\frac{\beta A}{h}\right) \leq |\mathcal{C}_M| \leq C_4 \exp\left(\frac{6\lambda_c}{h}\right). \tag{3.37}$$

The lower bound in (3.37) is central to our analysis, and we will return to it later, when we discuss droplet creation. The upper bound in (3.37) is of technical relevance in connection to droplet growth, because we will want to have all the events $G_{+,k}$, $k \in \mathcal{C}_M$ happening. At this point recall that our goal is to prove (3.1), in which an arbitrarily large constant C is involved. It is clear from Proposition 3.2.1 and the upper bound in (3.37) that given a value for the constant C in (3.1) we can choose b_1 so large that for some finite C_1

$$\mathbb{P}(\cup_{k\in\mathcal{C}_M}(G_{+,k})^c) \leq C_1 \exp(-C/h), \tag{3.38}$$

for all $h > 0$.

Define the space-time region

$$\Delta = \left(\bigcup_{k\in\mathcal{C}_M} \nabla_k \right) \bigcup \Box,$$

where

$$\Box = \Lambda\left(\frac{b_2}{h}W\right) \times \left[Mt_N, \exp\left(\frac{\lambda}{h}\right)\right].$$

We can think of this space-time region Δ as playing the role of the cone in the heuristics. Informally speaking, (3.38) assures that once a supercritical droplet of size $(b_1/h)W$ is born at the bottom of some ∇_k, $k \in \mathcal{C}_M$, then with large probability it will reach the top of the uppermost inverted pyramid, $\nabla_{(0,0,M-1)}$, of $\mathcal{C}_M$. If it survives also through the top cylinder $\Box$ of Δ, it will reach the time $\exp(\lambda/h)$.

A further condition on b_2 is required to assure us that the cylinder at the top of Δ is wide enough so that the aimed conclusion (3.1) will be proven with an arbitrarily large constant C. Technically this condition is the following. Given the constant C in (3.1), we will need to take b_2 large enough for there to exist a finite C_1 so that

$$\langle f\rangle_{\Lambda(\frac{b_2}{h}W),-,h} \geq \langle f\rangle_h - C_1\exp(-C/h), \tag{3.39}$$

for all $h > 0$ small enough. That such a choice is possible is heuristically reasonable, since in the double-well picture of the Gibbs distribution $\mu_{\Lambda(\frac{b_2}{h}W),-,h}$ with b_2 large, the mass is concentrated in the well corresponding to the (+)-phase, which can be made very deep by choosing b_2 appropriately large. From the rigorous viewpoint, aside from the claim that C can be arbitrarily large, a proof of (3.39) can be found in [SS1] (see the proof of Theorem 1.b.3 there). A complete proof of (3.39) is obtained by combining Lemma 3.4.8 with (1.20) and the FKG-Holley inequalities.

Observe that our choices are made in the following order. Suppose λ and C are given. First (3.36) gives us the value of δ. Then (3.38) gives us the value of b_1. Afterwards, (3.33) gives us b', and finally (3.34) and (3.39) give us b_2.

So far we have developed mathematically rigorous counterparts to the notion that if a supercritical droplet is created close to the bottom of any of the inverted pyramids ∇_k, $k \in \mathcal{C}_M$ it is likely to grow and bring the (+)-phase to the neighborhood of the origin at time $\exp(\lambda/h)$. What we still need is to make mathematical sense of the creation of such a supercritical droplet as occurring at a rate predicted by the heuristics. The lower bound in (3.37) tells us that if we could say that close to the bottom of each one of these inverted pyramids, and independently of what happens close to the bottom of the other inverted pyramids, there is probability of the order of $\exp(-\beta\frac{A}{h})$ of creating such a droplet, then we would be done. This would be akin to saying that the rate of creation of supercritical droplets is $\exp(-\beta\frac{A}{h})$, as we expect.

Motivated by (3.35) we will say that supercritical droplet creation occurs in the bottom of the inverted pyramid ∇_k, which has $\Lambda^{\text{bot}} \times [s_0, s_1]$ as its bottom cylinder, in case the following event happens:

$$F_k \;=\; \left\{\sigma^{s_0,-}_{\nabla_k,-,h;s_1} = \sigma^{s_0,+}_{\nabla_k,-,h;s_1}\right\}.$$

The events F_k, $k \in \mathcal{C}_M$ are clearly mutually independent, since they are determined by the graphical construction marks in disjoint regions of space-time. The final part of this section will be concerned with proving that for each k for small $h > 0$

$$\mathbb{P}(F_k) = \mathbb{P}(F_0) \geq \frac{1}{2}\exp\left(-\frac{\beta A}{h}\right). \tag{3.40}$$

If for the moment we suppose that this is known, we can complete the proof of (3.1) as follows. Consider the event

$$E = (\cap_{k\in\mathcal{C}_M} G_{+,k}) \bigcap (\cup_{k\in\mathcal{C}_M} F_k).$$

From (3.38), the lower bound in (3.37), and (3.40) we have for small h

$$\mathbb{P}(E^c) \leq C_1\exp(-C/h) + \left(1 - \frac{1}{2}\exp\left(-\frac{\beta A}{h}\right)\right)^{|\mathcal{C}_M|} \leq 2C_1\exp(-C/h).$$

Now, using basic-coupling inequalities, (3.35) and (3.39) we obtain

$$\begin{aligned}
\mathbb{E}\left(f\left(\sigma^{\nu}_{h;\exp(\lambda/h)}\right)\right) &\geq \mathbb{E}\left(f\left(\sigma^{0,-}_{\Delta,-,h;\exp(\lambda/h)}\right)\right)\\
&\geq \mathbb{E}\left(f\left(\sigma^{0,-}_{\Delta,-,h;\exp(\lambda/h)}\right)\;;\;E\right) - \mathbb{P}(E^c)\\
&= \mathbb{E}\left(f\left(\sigma^{Mt_N,+}_{\Box,-,h;\exp(\lambda/h)}\right)\;;\;E\right) - \mathbb{P}(E^c)\\
&\geq \mathbb{E}\left(f\left(\sigma^{Mt_N,+}_{\Box,-,h;\exp(\lambda/h)}\right)\right) - 2\mathbb{P}(E^c)\\
&\geq \langle f\rangle_{\Lambda(\frac{b_2}{h}W),-,h} - 4C_1\exp(-C/h)\\
&\geq \langle f\rangle_h - C_1'\exp(-C/h).
\end{aligned}$$

All that remains to be done in this section is to show (3.40). For this purpose we will insert another inverted pyramid inside the bottom cylinder of each one of our inverted pyramids ∇_k, $k \in \mathcal{C}_M$. To distinguish the new inverted pyramids from the ones that we have been discussing so far (parametrized by b_1, b_2, δ and h), we will call the old ones "growth inverted pyramids" and the ones which we are introducing now "creation inverted pyramids". The creation inverted pyramid which is inserted inside the bottom cylinder of $\triangledown = \triangledown_0$ is described next. It will be of the form

$$\triangledown_{\mathrm{cr}} = \bigcup_{i=0}^{N_{\mathrm{cr}}} [u_i, u_{i+1}] \times \Lambda_i^{\mathrm{cr}},$$

with the following features. As with the growth inverted pyramids, for $i = 0, ..., N_{\mathrm{cr}} - 1$, the set $\Lambda_{i+1}^{\mathrm{cr}}$ will be obtained from Λ_i^{cr} by adding one site to it. In particular we will have $\Lambda_0^{\mathrm{cr}} \subset \Lambda_1^{\mathrm{cr}} \subset ... \subset \Lambda_N^{\mathrm{cr}}$. These sets Λ_i^{cr} will all be l_0-quasi-Wulff-shaped. We will take $\Lambda_0^{\mathrm{cr}} = \Lambda\left(\frac{b_0}{h}W\right)$, where $b_0 \in (B_c, b_1)$ is a new parameter. At the other end, $\Lambda_{N_{\mathrm{cr}}}^{\mathrm{cr}} = \Lambda\left(\frac{b_1}{h}W\right)$, where b_1 is the same one used for the growth inverted pyramids. The bottom cylinder of $\triangledown_{\mathrm{cr}}$ will have height $u_1 - u_0 = \exp\left(\frac{\delta}{2h}\right)$, while all its other cylinders will have height $u_{i+1} - u_i = \exp\left(\frac{\delta'}{h}\right)$, for $i = 1, ..., N_{\mathrm{cr}}$, where $\delta' < \delta$ is also a new parameter.

With the parameters b_1, δ, b_0 and δ' fixed and satisfying the conditions above, it is clear that for small enough $h > 0$, there is an inverted pyramid as described above which fits inside the bottom cylinder of $\triangledown$, and has its top $\Lambda_{N_{\mathrm{cr}}}^{\mathrm{cr}} \times u_{N_{\mathrm{cr}}+1}$ coinciding with the top of the bottom cylinder of $\triangledown$, i.e., such that $u_{N_{\mathrm{cr}}+1} = \exp(\delta/h)$. This is the inverted pyramid that we will denote by $\triangledown_{\mathrm{cr}}$.

The basic-coupling inequalities imply that

$$\left\{\sigma^{u_0,-}_{\triangledown_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}} = \sigma^{u_{N_{\mathrm{cr}}},+}_{\triangledown_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}}\right\} \subset \left\{\sigma^{0,-}_{\triangledown,-,h;\exp(\delta/h)} = \sigma^{0,+}_{\triangledown,-,h;\exp(\delta/h)}\right\}.$$

Therefore the next lemma implies (3.40).

Lemma 3.2. *Given $b_1 > B_c$ and $\delta > 0$ there are $b_0 \in (B_c, b_1)$ and $\delta' > 0$ such that for small $h > 0$*

$$\mathbb{P}\left(\sigma^{u_0,-}_{\triangledown_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}} = \sigma^{u_{N_{\mathrm{cr}}},+}_{\triangledown_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}}\right) \geq \frac{1}{2}\exp\left(-\frac{\beta A}{h}\right).$$

Proof (modulo results in Sections 3.4 and 3.5). Before starting the rigorous proof we will motivate it heuristically. Consider first the bottom cylinder of $\triangledown_{\mathrm{cr}}$. Intuitively, when b_0 is close to B_c, we can see the system in the box $\Lambda\left(\frac{b_0}{h}W\right)$ as having a double well structure with the deeper well corresponding to the $(-)$-phase, and the higher well corresponding to the presence of a supercritical droplet of the $(+)$-phase. The barrier between these wells is given by the configurations with a critical droplet. Note that if we were in equilibrium at time u_1 inside the box $\Lambda\left(\frac{b_0}{h}W\right)$ with $(-)$ boundary conditions, then the quantity $\frac{1}{2}\exp\left(-\frac{\beta A}{h}\right)$ would indeed be a lower bound on the probability of being in the higher well and hence having a supercritical droplet.

We would therefore like to say that inside this bottom cylinder the system started at time u_0 with all spins down should reach equilibrium at the top of this cylinder, i.e., at time $u_1 = u_0 + \exp\left(\frac{\delta}{2h}\right)$. In other words, we would like to say that the relaxation time for the process $\left(\sigma^{\cdot}_{\Lambda\left(\frac{b_0}{h}W\right),-,h;t}\right)$ is shorter than $\exp\left(\frac{\delta}{2h}\right)$. For this to be true it should be enough to take b_0 close enough to B_c, so that from the higher well there is a very small barrier to overcome to reach equilibrium. This free-energy barrier can be made smaller than $\frac{\delta}{2h}$, so the available time should be enough to equilibrate the system.

The heuristics in the last paragraph can be made rigorous by considering the spectral gap of the generator of the process. This will be done in Section 3.5. From Proposition 3.5.3 we have that if b_0 is chosen close enough to B_c, then

$$\operatorname{gap}\left(\Lambda\left(\frac{b_0}{h}W\right), -, h\right) \geq \exp\left(-\frac{\delta}{4h}\right), \tag{3.41}$$

for small enough h.

We start now the rigorous proof of the lemma. First we break things down according to what happens at time $u_1 = u_0 + \exp\left(\frac{\delta}{2h}\right)$.

$$\mathbb{P}\left(\sigma^{u_0,-}_{\nabla_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}} = \sigma^{u_{N_{\mathrm{cr}}},+}_{\nabla_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}}\right) = \\ \sum_{\zeta} \mathbb{P}\left(\sigma^{u_0,-}_{\nabla_{\mathrm{cr}},-,h;u_1} = \zeta\right) \mathbb{P}\left(\sigma^{u_1,\zeta}_{\nabla_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}} = \sigma^{u_{N_{\mathrm{cr}}},+}_{\nabla_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}}\right).$$

Combining (3.41) with the standard inequality (59) in [Sch1], we have

$$\Bigg| \sum_{\zeta} \mathbb{P}\left(\sigma^{u_0,-}_{\nabla_{\mathrm{cr}},-,h;u_1} = \zeta\right) \mathbb{P}\left(\sigma^{u_1,\zeta}_{\nabla_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}} = \sigma^{u_{N_{\mathrm{cr}}},+}_{\nabla_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}}\right) - \\ \sum_{\zeta} \mu_{\Lambda(\frac{b_0}{h}W),-,h}(\zeta)\, \mathbb{P}\left(\sigma^{u_1,\zeta}_{\nabla_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}} = \sigma^{u_{N_{\mathrm{cr}}},+}_{\nabla_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}}\right) \Bigg| \\ \leq \frac{e^{-\exp(\frac{\delta}{2h})\operatorname{gap}(\Lambda(\frac{b_0}{h}W),-,h)}}{\mu_{\Lambda(\frac{b_0}{h}W),-,h}(-)} \leq \exp\left(C\left(\frac{b_0}{h}\right)^2\right) \exp\{-e^{\frac{\delta}{4h}}\}.$$

But using Lemma 3.4.4 and Proposition 3.2.2, we can take δ' small enough so that

$$\sum_{\zeta} \mu_{\Lambda(\frac{b_0}{h}W),-,h}(\zeta)\, \mathbb{P}\left(\sigma^{u_1,\zeta}_{\nabla_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}} = \sigma^{u_{N_{\mathrm{cr}}},+}_{\nabla_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}}\right) \\ \geq \mu_{\Lambda(\frac{b_0}{h}W),-,h}(\mathcal{R}^c) \sum_{\zeta\in\mathcal{R}^c} \widehat{\mu}_{\Lambda(\frac{b_0}{h}),-,h}(\zeta)\, \mathbb{P}\left(\sigma^{u_1,\zeta}_{\nabla_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}} = \sigma^{u_{N_{\mathrm{cr}}},+}_{\nabla_{\mathrm{cr}},-,h;u_{N_{\mathrm{cr}}+1}}\right) \\ \geq \left(\exp\left(-\frac{\beta A}{h}\right)\right) \cdot \frac{3}{4},$$

for small h.

The three displays above combined give us the lemma. ■

At this point the proof of part (ii) of Theorem 1 has been reduced to proving the claims in Sections 3.4 and 3.5.

3.4. Double well structure of equilibrium distributions. In this section we will study some of the features of the Gibbs distributions on finite simply-connected sets with $(-)$-boundary conditions. We will extend and strengthen results contained in Theorem 1 of [SS1]; the purpose being the use of these stronger results in several other sections of the current paper.

While the main results that we will derive in this section and use in other ones could be derived using the same approach as in [SS1], we will nevertheless introduce an alternative method for proving them, based on results and techniques from Section 2.3 of this paper. Basically, in [SS1] we started from results on large deviations of the average spin (i.e., the magnetization) inside of a box, under no external field. We could see the external field then as tilting the distribution. Such an approach is appealing and even natural, but since in the current paper our interest lies primarily on contours and not on the magnetization, it seemed even more natural to search for a direct approach to the problems, in which the magnetization need not to be mentioned.

Having in mind that we are interested in the contours of configurations in Ω_-, it is natural to regard $\mathcal{W}$ and $\hat{V}$, defined in Section 2.3, as random variables. Given a configuration $\eta \in \Omega_-$, the associated values of these random variables are, respectively, $\mathcal{W}(S)$ and $\hat{V}(S)$, where S is the collection of skeletons corresponding to the external contours of η.

Basically we will show that in a sense the function

$$\phi(b) = wb - m^* b^2 \qquad b \geq 0,$$

plays the role of a large-deviation rate function for the random variable $\hat{V}$. The title of this section derives from the shape of this function. The reader will realize that in the lemmas below the results are stated and proven with a certain amount of uniformity over the allowed sets Λ; this uniformity is needed in some of our applications of the lemmas, since for instance in Section 3.2 we need uniformity over all the sets which are bases of cylinders of inverted pyramids.

Recall that B_c is the value of b which maximizes ϕ, with $\phi(B_c) = A$, while $B_0 = 2B_c$ is the value of b above which this function becomes negative. Recall also that given a finite set of vertebrate contours G we denote by $\mathcal{S}_G^{h,\mathrm{co}}$ the set of configurations which belong to Ω_- and which have as their collection of external vertebrate contours the set G. In some of the lemmas below, we will use as a "reference" the set of configurations with no external vertebrate contours, $\mathcal{S}_\emptyset^{h,\mathrm{co}}$. Our next lemma shows that this is essentially the same as using the set $\mathcal{R}$, defined by (2.10), as a "reference".

We will use the notation $o_h(1)$ to represent some function of $h > 0$, satisfying the property $\lim_{h \searrow 0} o_h(1) = 0$.

Lemma 3.4.1. *For any $p > 0$ there is a function $o_h(1)$ such that for any simply-connected set $\Lambda \subset \mathbb{Z}^2$ which satisfies $|\Lambda| \leq 1/h^p$,*

$$1 \leq \frac{Z_{\Lambda,-,h}(\mathcal{R})}{Z_{\Lambda,-,h}\left(\mathcal{S}_\emptyset^{h,\mathrm{co}}\right)} \leq 1 + o_h(1),$$

Proof. From Lemmas 2.3.3 and 2.3.7 we have, for small enough h depending on p and some finite C also depending on p,

$$\tilde{\mu}_{\Lambda,-,h}\left(\left(\mathcal{S}_{\emptyset}^{h,\mathrm{co}}\right)^{c}\right) \leq C \exp\left(-\beta \frac{w}{4\sqrt{2}h^{\tilde{b}}}\right).$$

(Here we use $\widetilde{b}$ to denote the parameter entering the definition of the vertebrate contour.) The thesis follows immediately from the last estimate. ■

Lemma 3.4.2. *For any $p > 0$ there is $h_0 > 0$ and a function $o_h(1)$ such that given $D_0 > 0$ there is a finite constant C so that for any $D > D_0$, any $b > 0$, any $0 < h \leq h_0$, and any simply-connected set $\Lambda \subset \mathbb{Z}^2$ which satisfies $|\Lambda| \leq 1/h^p$,*

$$\frac{Z_{\Lambda,-,h}\left(\mathcal{W} \geq \frac{D}{h}\ ,\ \hat{V} \leq \left(\frac{b}{h}\right)^2\right)}{Z_{\Lambda,-,h}\left(\mathcal{S}_{\emptyset}^{h,\mathrm{co}}\right)} \leq C \exp\left(-\frac{\beta}{h}(D - m^* b^2)(1 + o_h(1))\right).$$

Proof. The proof of Lemma 2.3.6 and (2.24) show that we can choose h_0 and C so as to have for each collection of skeletons S such that $\hat{V}(S) \leq \left(\frac{b}{h}\right)^2$

$$\sum_{G \in \mathcal{C}_S^h} \frac{Z_{\Lambda,-,h}(\mathcal{S}_G^{h,\mathrm{co}})}{Z_{\Lambda,-,h}(\mathcal{S}_{\emptyset}^{h,\mathrm{co}})} \leq \exp\left(-\frac{\beta}{h}(\mathcal{W}(S) - m^* b^2)(1 + o_h(1))\right).$$

Summing then over S with $\mathcal{W}(S) \geq \frac{D}{h}$, using the entropy estimate in the proof of Lemma 2.3.7, we obtain the desired conclusion. ■

In what follows, with $b > 0$ and $0 < \rho < 1$ given, $E_{b,\rho}^h$ will denote the event that there is an external contour which surrounds $\frac{b(1-\rho)}{h} W$ and is contained in $\frac{b(1+\rho)}{h} W$, and that moreover this is the only external vertebrate contour.

Lemma 3.4.3. *For any $p > 0$ and $\gamma < 1$ there are finite positive constants h_0 and C_1 and a function $o_h(1)$ such that for any $0 < h \leq h_0$, any $b > 0$, any $\rho = \rho(h) > h^\gamma$ and any simply-connected set $\Lambda \subset \mathbb{Z}^2$ which satisfies $|\Lambda| \leq 1/h^p$ and contains $\Lambda\left(\frac{b(1+\rho)}{h} W\right)$,*

$$\frac{Z_{\Lambda,-,h}\left(E_{b,\rho}^h\right)}{Z_{\Lambda,-,h}\left(\mathcal{S}_{\emptyset}^{h,\mathrm{co}}\right)} \geq C_1 \exp\left(-\frac{\beta}{h}\phi(b)(1 + o_h(1))\right).$$

In particular, ρ can be a positive constant.

Proof. Without loss of generality we can suppose that $\rho = h^\gamma$ with $\gamma < 1$.

Partition $E_{b,\rho}^h$ according to what the vertebrate external contours are, as

$$E_{b,\rho}^h = \bigcup_{G \in \mathcal{G}_{b,\rho}^h} \mathcal{S}_G^{h,\mathrm{co}}.$$

By the definition of $E^h_{b,\rho}$, each $G \in \mathcal{G}^h_{b,\rho}$ is a singleton. Exactly as in (2.33), we have

$$\frac{Z_{\Lambda,-,h}\left(E^h_{b,\rho}\right)}{Z_{\Lambda,-,h}\left(\mathcal{S}^{h,\mathrm{co}}_{\emptyset}\right)} = \sum_{G \in \mathcal{G}^h_{b,\rho}} \frac{Z_{\Lambda,-,h}(\mathcal{S}^{h,\mathrm{co}}_{G})}{Z_{\Lambda,-,h}(\mathcal{S}^{h,\mathrm{co}}_{\emptyset})}$$
$$= \sum_{G \in \mathcal{G}^h_{b,\rho}} \frac{Z_{\Lambda,-,0}(\mathcal{S}^{h,\mathrm{co}}_{G})}{Z_{\Lambda,-,0}(\mathcal{S}^{h,\mathrm{co}}_{\emptyset})} \exp\left(\frac{\beta}{2}\int_0^h \sum_{x\in\Lambda}\left[\langle\sigma(x)\rangle^{G,h}_{\Lambda,-,h'} - \langle\sigma(x)\rangle^{\emptyset,h}_{\Lambda,-,h'}\right] dh'\right).$$

Arguments analogous to the ones which led to (2.34) provide also the following bound, in the opposite direction. For small enough h_0, and some finite positive C_2, if $G \in \mathcal{G}^h_{b,\rho}$ and $0 < h' \le h \le h_0$, then

$$\sum_{x\in\Lambda}\left[\langle\sigma(x)\rangle^{G,h}_{\Lambda,-,h'} - \langle\sigma(x)\rangle^{\emptyset,h}_{\Lambda,-,h'}\right] \ge 2m^*(1+o_h(1))\left(\frac{b}{h}\right)^2.$$

Therefore we have

$$\frac{Z_{\Lambda,-,h}\left(E^h_{b,\rho}\right)}{Z_{\Lambda,-,h}\left(\mathcal{S}^{h,\mathrm{co}}_{\emptyset}\right)} \ge C_1 \exp\left(\frac{\beta}{h}\left(m^* b^2(1+o_h(1))\right)\right) \sum_{G \in \mathcal{G}^h_{b,\rho}} \frac{Z_{\Lambda,-,0}(\mathcal{S}^{h,\mathrm{co}}_{G})}{Z_{\Lambda,-,0}(\mathcal{S}^{h,\mathrm{co}}_{\emptyset})}$$
$$\ge C_1 \exp\left(\frac{\beta}{h}\left(m^* b^2(1+o_h(1))\right)\right) \mu_{\Lambda,-,0}(E^h_{b,\rho}).$$

The problem is now reduced to a 0-field problem. We will use the notation introduced in the proof of Lemma 2.3.6. For each $G \in \mathcal{G}^h_{b,\rho}$ let F_G be the event that the spins in $\partial_- G$ are all negative and those in $\partial_+ G$ are all positive. With this notation we can write

$$\mu_{\Lambda,-,0}(E^h_{b,\rho}) = \sum_{G \in \mathcal{G}^h_{b,\rho}} \mu_{\Lambda,-,0}(F_G)\mu_{\Lambda^G_{\mathrm{ext}},-,0}(\mathcal{S}^{h,\mathrm{co}}_{\emptyset})$$
$$\ge \mu_{\Lambda,-,0}(\mathcal{S}^{h,\mathrm{co}}_{\emptyset}) \sum_{G \in \mathcal{G}^h_{b,\rho}} \mu_{\Lambda,-,0}(F_G) \ge \frac{1}{2}\sum_{G \in \mathcal{G}^h_{b,\rho}} \mu_{\Lambda,-,0}(F_G),$$

where in the second step we used the FKG-Holley inequalities, and in the last step we used (2.40).

To complete the proof, we note that from Lemma 5.2 in [Iof1], or from the techniques in Section 7 of [SS2], we have

$$\sum_{G \in \mathcal{G}^h_{b,\rho}} \mu_{\Lambda,-,0}(F_G) \ge \exp\left(-\frac{\beta}{h} w b (1+o_h(1))\right).$$

■

Lemma 3.4.4. *For any $p > 0$ and any $b > B_c$ there are finite positive constants h_0 and ϵ such that for any $0 < h \le h_0$ and any simply-connected set $\Lambda \subset \mathbb{Z}^2$ which satisfies $|\Lambda| \le 1/h^p$ and contains $\Lambda\left(\frac{b}{h}W\right)$,*

$$\mu_{\Lambda,-,h}\left(\mathcal{R}^c\right) \ge \exp\left(-\frac{\beta}{h}A(1-\epsilon)\right).$$

Proof. From Lemmas 3.4.1 and 3.4.3 we have, for small enough h_0, ϵ, small constant ρ and some positive finite C_1 and C_2,

$$\begin{aligned}\frac{Z_{\Lambda,-,h}\left(\mathcal{R}^c\right)}{Z_{\Lambda,-,h}\left(\mathcal{R}\right)} &\ge \frac{Z_{\Lambda,-,h}\left(E^h_{b(1-\rho),\rho}\right)}{2\ Z_{\Lambda,-,h}\left(\mathcal{S}^{h,\mathrm{co}}_{\emptyset}\right)}\\ &\ge C_1\exp\left(-\frac{\beta}{h}\phi(b(1-\rho))(1+o_h(1))\right)\\ &\ge 2\exp\left(-\frac{\beta}{h}A(1-\epsilon)\right).\end{aligned}$$

The conclusion is now immediate. ∎

Lemma 3.4.5. *For any $p > 0$, there are finite positive constants h_0, C_1 and C_2 and a function $o_h(1)$ such that given also $0 < b_1 < b_2$ there is a finite constant C_1' so that for any $0 < h \le h_0$, any $b \in [b_1, b_2]$, any $\kappa \in (0,1)$, any integer k and any simply-connected set $\Lambda \subset \mathbb{Z}^2$ which satisfies $|\Lambda| \le 1/h^p$ and contains $\Lambda\left(\frac{b}{h}(1-\kappa^2/4)W\right)$,*

$$\begin{aligned}&C_1\exp\left(-\frac{\beta}{h}\Big(\min_{b(1-\kappa)<\widetilde{b}<b}\phi(\widetilde{b})(1+o_h(1))\Big)\right)\\ &\qquad\le\frac{Z_{\Lambda,-,h}\left(\left(\frac{b}{h}(1-\kappa)\right)^2\le\hat{V}\le\left(\frac{b}{h}\right)^2\right)}{Z_{\Lambda,-,h}\left(\mathcal{S}^{h,\mathrm{co}}_{\emptyset}\right)}\\ &\qquad\le kC_1'\exp\left(-\frac{\beta}{h}\Big(\min_{b(1-\kappa)<\widetilde{b}<b}\phi(\widetilde{b})(1+o_h(1))-C_2\frac{\kappa}{k}b\Big)\right).\end{aligned}$$

Proof. For the first inequality, simply note that if $\epsilon > 0$ is small, $\rho = \epsilon/2$, then for small h for all $\widetilde{b} > 0$,

$$E^h_{\widetilde{b}(1-\rho),\rho} \subset \left\{\left(\frac{\widetilde{b}}{h}(1-\epsilon)\right)^2 \le \hat{V} \le \left(\frac{\widetilde{b}}{h}\right)^2\right\},$$

and also $\Lambda\left(\frac{b(1-\rho)(1+\rho)}{h}W\right) \subset \Lambda$. Hence the claim follows from Lemma 3.4.3.

For the second inequality, note that from the variational result for families of curves presented in Section 2.9 of [DKS] we know that if $\hat{V} \ge \left(\frac{b}{h}\right)^2$ then $\mathcal{W} \ge w\frac{b}{h}$.

We can therefore apply Lemma 3.4.2 to each of the k events $\left(\frac{b}{h}(1-i\frac{\kappa}{k})\right)^2 \leq \hat{V} \leq \left(\frac{b}{h}(1-(i-1)\frac{\kappa}{k})\right)^2$, $i=1,2,...,k$, to conclude that

$$\frac{Z_{\Lambda,-,h}\left(\left(\frac{b}{h}(1-i\frac{\kappa}{k})\right)^2 \leq \hat{V} \leq \left(\frac{b}{h}(1-(i-1)\frac{\kappa}{k})\right)^2\right)}{Z_{\Lambda,-,h}\left(\mathcal{S}_{\emptyset}^{h,\mathrm{co}}\right)}$$

$$\leq \frac{Z_{\Lambda,-,h}\left(\mathcal{W} \geq w\frac{b}{h}(1-i\frac{\kappa}{k})\ ,\ \hat{V} \leq \left(\frac{b}{h}(1-(i-1)\frac{\kappa}{k})\right)^2\right)}{Z_{\Lambda,-,h}\left(\mathcal{S}_{\emptyset}^{h,\mathrm{co}}\right)}$$

$$\leq C_1' \exp\left(-\frac{\beta}{h}\left(\phi(b(1-(i-1)\frac{\kappa}{k})(1+o_h(1)) - C_2\frac{\kappa}{k}b\right)\right).$$

■

The previous lemma basically contains the promised characterization of $\phi(\cdot)$ as a large-deviation rate function for the random variable $\hat{V}$. Informally it tells us that under appropriate conditions on Λ,

$$\frac{Z_{\Lambda,-,h}\left(\hat{V} \approx \left(\frac{b}{h}\right)^2\right)}{Z_{\Lambda,-,h}\left(\mathcal{S}_{\emptyset}^{h,\mathrm{co}}\right)} \sim \exp\left(-\frac{\beta}{h}\phi(b)\right).$$

There is a technical difficulty in applying Lemma 3.4.5, nevertheless, and this is the motivation for the next lemma. The issue is that, as stated above, the lemma cannot be use to estimate $Z_{\Lambda,-,h}\left(\hat{V} \in (I/h)^2\right)/Z_{\Lambda,-,h}\left(\mathcal{S}_{\emptyset}^{h,\mathrm{co}}\right)$, if I is, e.g., an interval of the form $[0,b]$, for some $b>0$, since $C_1' = C_1'(b_1,b_2)$ can explode as $b_1 \to 0$.

Lemma 3.4.6. *For any $p>0$, there is $h_0>0$ and a function $o_h(1)$ such that for any $0<h\leq h_0$, any $0<b<B_0$, and any simply-connected set $\Lambda \subset \mathbb{Z}^2$ which satisfies $|\Lambda| \leq 1/h^p$,*

$$1 \leq \frac{Z_{\Lambda,-,h}\left(\hat{V} \leq \left(\frac{b}{h}\right)^2\right)}{Z_{\Lambda,-,h}\left(\mathcal{S}_{\emptyset}^{h,\mathrm{co}}\right)} \leq 1 + o_h(1).$$

Proof. Only the upper bound has to be explained. If $b<B_c$ the claim follows from Lemma 3.4.1. The general case is reduced to this one by taking $0<b' \leq \min\{B_c, b\}$ and using Lemma 3.4.5 to control the contribution from the interval $[b',b]$. ■

Let $\bar{E}_{b,\rho}^h$ be the event that there is an external contour which up to a translation surrounds $\frac{b(1-\rho)}{h}W$ and is contained in $\frac{b(1+\rho)}{h}W$.

Lemma 3.4.7. *For any $p>0$ and $\rho>0$ there are finite positive constants ϵ_0 and C_2 such that given also $0<b_1<b_2$ there are finite positive constants h_0 and C_1*

such that for any $0 < h \leq h_0$, any $b \in [b_1, b_2]$, any $\epsilon \leq \epsilon_0$ and any simply-connected set $\Lambda \subset \mathbb{Z}^2$ which satisfies $|\Lambda| \leq 1/h^p$ and contains $\Lambda\left(\frac{b}{h}(1-\epsilon^2/4)W\right)$,

$$\mu_{\Lambda,-,h}\left((\bar{E}^h_{b,\rho})^c \,\middle|\, \left(\frac{b}{h}(1-\epsilon)\right)^2 \leq \hat{V} \leq \left(\frac{b}{h}\right)^2\right) \leq C_1 \exp\left(-\frac{C_2 b}{h}\right).$$

Proof. We will use the stability result for the Wulff functional (in the case of families of curves) contained in Theorem 2.9 of [DKS].

Consider a configuration in

$$(\bar{E}^h_{b,\rho})^c \cap \left\{\left(\frac{b}{h}(1-\epsilon)\right)^2 \leq \hat{V} \leq \left(\frac{b}{h}\right)^2\right\}.$$

Clearly there is a positive finite C (depending only on the temperature) such that, if h is small, for any such configuration there is no external contour whose skeleton, even after being translated by any amount, is at a Hausdorff distance less than $C\frac{b\rho}{h}$ from the boundary of $\frac{b}{h}W$.

To use Theorem 2.9 in [DKS], we scale lengths down by dividing them by $\frac{b(1-\epsilon)}{h}$. The rescaled collection of skeletons corresponding to the external contours has a phase volume bounded below by 1. On the other hand, no translate of any of the rescaled skeletons is at a Hausdorff distance less than $C\frac{\rho}{1-\epsilon}$ from the boundary of $\frac{1}{1-\epsilon}W$.

The quoted theorem in [DKS] tells us then that the Wulff functional associated to the collection of scaled skeletons is bounded below by an expression of the type $w(1+G(\rho,\epsilon))$, where $\lim_{\epsilon\searrow 0} G(\rho,\epsilon) = G(\rho,0) > 0$. Scaling lengths back to their original value, we have obtained the following lower bound for the Wulff functional of any configuration in the event with which we are concerned:

$$\mathcal{W} \geq w\frac{b}{h}(1-\epsilon)(1+G(\rho,\epsilon)).$$

If ϵ_0 is chosen small enough so that $(1-\epsilon_0)(1+G(\rho,\epsilon_0)) > 1$, and then h_0 is chosen small enough, based on Lemma 3.4.2, we obtain for all $\epsilon \leq \epsilon_0$

$$\begin{aligned}
&\frac{Z_{\Lambda,-,h}\left((\bar{E}^h_{b,\rho})^c \cap \left\{\left(\frac{b}{h}(1-\epsilon)\right)^2 \leq \hat{V} \leq \left(\frac{b}{h}\right)^2\right\}\right)}{Z_{\Lambda,-,h}\left(\mathcal{S}^{h,\mathrm{co}}_{\emptyset}\right)} \\
&\leq \frac{Z_{\Lambda,-,h}\left(\mathcal{W} \geq w\frac{b}{h}(1-\epsilon)(1+G(\rho,\epsilon))\ ,\ \hat{V} \leq \left(\frac{b}{h}\right)^2\right)}{Z_{\Lambda,-,h}\left(\mathcal{S}^{h,\mathrm{co}}_{\emptyset}\right)} \\
&\leq C_1 \exp\left(-\frac{\beta}{h}\left(wb(1-\epsilon)(1+G(\rho,\epsilon))(1+o_h(1)) - m^* b^2(1+o_h(1))\right)\right) \\
&\leq C_1 \exp\left(-\frac{\beta}{h}\phi(b)(1+o_h(1)) - \frac{C_2' b}{h}\right),
\end{aligned}$$

where $C_2' = \beta w((1-\epsilon)G(\rho,\epsilon)-\epsilon)/2$.

The result now follows from the comparison between this estimate and the first inequality in Lemma 3.4.5. In this fashion, with both h_0 and ϵ_0 small enough, we can take $C_2 = \beta w G(\rho, 0)/3$. ∎

As motivation for the next lemma, we recall that by controlling how deeply the $(-,*)$ cluster of the boundary penetrates in a set Λ, we can obtain estimates similar to (1.20) for the expected value of observables. Define

$$\mathcal{B}^{-*}_{h,b}(\epsilon) = \bigcup_{b'\in(\epsilon b, b]} \left\{ \Lambda\left(\frac{b'-\epsilon b}{h}W\right) \overset{-*}{\longleftrightarrow} \left(\Lambda\left(\frac{b'}{h}W\right)\right)^c \right\}.$$

Recall that a set $\Lambda \subset \mathbb{Z}^2$ is said to be l_0-quasi-Wulff-shaped with linear parameter l in case it is simply-connected and

$$\Lambda((l-l_0)W) \;\subset\; \Lambda \;\subset\; \Lambda((l+l_0)W).$$

Recall also that $\widehat{\mu}_{\Lambda,-,h}(\,\cdot\,) = \mu_{\Lambda,-,h}(\,\cdot \mid \mathcal{R}^c)$.

Lemma 3.4.8. *For any $l_0 > 0$ and any $\epsilon > 0$ there is $C_2 > 0$ such that given also $B_c < b_1 \le b_2$ there are positive finite constants h_0 and C_1 such that if $0 < h \le h_0$ then for any l_0-quasi-Wulff-shaped set Λ which has linear size parameter b/h with $b_1 \le b \le b_2$,*

$$\widehat{\mu}_{\Lambda,-,h}(\mathcal{B}^{-*}_{h,b}(\epsilon)) \le C_1 \exp\left(-\frac{C_2(b_1-B_c)}{h}\right). \tag{3.42}$$

In case $b_1 > B_0$ the same holds also for the measure $\mu_{\Lambda,-,h}$:

$$\mu_{\Lambda,-,h}(\mathcal{B}^{-*}_{h,b}(\epsilon)) \le C_1 \exp\left(-\frac{C_2(b_1-B_0)}{h}\right). \tag{3.43}$$

Proof. We start by introducing more terminology. The event that a certain contour Γ is present is equivalent to the statement that the spins in a certain set of sites $S_1(\Gamma)$ all have the same sign and the spins in a certain other set of sites $S_2(\Gamma)$ all have the opposite sign. Exactly one of the two sets, $S_1(\Gamma)$ or $S_2(\Gamma)$ is surrounded by the contour, while the other is completely outsite of the contour. The one which is surrounded by Γ will be denoted by $\partial_{\text{int}}\Gamma$, while the one which is outside of Γ will be denoted by $\partial_{\text{ext}}\Gamma$.

Let $\mathcal{D}^+_{h,b}(\epsilon)$ be the event that for some contour Γ which surrounds $\frac{b(1-\epsilon/2)}{h}W$ all spins in $\partial_{\text{int}}\Gamma$ are $+1$.

Our first goal is to prove that

$$\widehat{\mu}_{\Lambda,-,h}\left(\left(\mathcal{D}^+_{h,b}(\epsilon)\right)^c\right) \le C_1 \exp\left(-\frac{C_2(b_1-B_c)}{h}\right). \tag{3.44}$$

In order to do it let us consider the event $\hat{V} \ge \left(\frac{b'}{h}\right)^2$, which for $b' < B_c$ and h small enough is bigger than $\mathcal{R}^c$, that is $\mathcal{R}^c \subset \left\{\hat{V} \ge \left(\frac{b'}{h}\right)^2\right\}$. If b' is just slightly smaller than B_c, then the bigger event is a good approximation to the smaller one. So instead of considering the conditional distribution $\widehat{\mu}_{\Lambda,-,h}$, we begin with the distribution

$$\bar{\mu}_{\Lambda,-,h}(\ \cdot\ ) = \mu_{\Lambda,-,h}(\ \cdot \mid \hat{V} \geq (b'/h)^2).$$

The choice of b' is immaterial. The only thing we need is that $b' < B_c$ and that it is close enough to B_c, so that $\phi(b') > \phi(b_1)$. Under such a choice this value b' would not even appear in our estimates.

Elementary geometric considerations show that for small enough ρ, dependent on ϵ but not on b and h,

$$\Omega_{\Lambda,-} \cap \left(\mathcal{D}^+_{h,b}(\epsilon)\right)^c \subset \left(\bar{E}^h_{b,\rho}\right)^c.$$

And obviously every configuration in $\Omega_{\Lambda,-}$ has $\hat{V} \leq \left(\frac{b}{h}\right)^2 + \frac{C}{h}$, for some finite constant C, which depends on b. So we can use Lemma 3.4.7 to conclude that we can choose an appropriate b'' (larger than but close enough to b), and small $\epsilon' > 0$, ϵ'' and $h_0 > 0$ so as to have

$$\begin{aligned}
&\bar{\mu}_{\Lambda,-,h}\left(\left(\mathcal{D}^+_{h,b}(\epsilon)\right)^c \middle| \hat{V} \geq \left(\frac{b}{h}(1-\epsilon')\right)^2\right) \\
&\quad = \mu_{\Lambda,-,h}\left(\left(\mathcal{D}^+_{h,b}(\epsilon)\right)^c \middle| \hat{V} \geq \left(\frac{b}{h}(1-\epsilon')\right)^2\right) \\
&\quad = \mu_{\Lambda,-,h}\left(\left(\mathcal{D}^+_{h,b}(\epsilon)\right)^c \middle| \hat{V} \geq \left(\frac{b''}{h}(1-\epsilon'')\right)^2\right) \\
&\quad = \mu_{\Lambda,-,h}\left(\left(\mathcal{D}^+_{h,b}(\epsilon)\right)^c \middle| \left(\frac{b''}{h}(1-\epsilon'')\right)^2 \leq \hat{V} \leq \left(\frac{b''}{h}\right)^2\right) \\
&\quad \leq \mu_{\Lambda,-,h}\left(\left(\bar{E}^h_{b,\rho}\right)^c \middle| \left(\frac{b''}{h}(1-\epsilon'')\right)^2 \leq \hat{V} \leq \left(\frac{b''}{h}\right)^2\right) \\
&\quad \leq \mu_{\Lambda,-,h}\left(\left(\bar{E}^h_{b'',\rho/2}\right)^c \middle| \left(\frac{b''}{h}(1-\epsilon'')\right)^2 \leq \hat{V} \leq \left(\frac{b''}{h}\right)^2\right) \\
&\quad \leq C_3 \exp\left(-\frac{C_4 b}{h}\right),
\end{aligned} \tag{3.45}$$

for some finite positive C_3 and C_4, which depend on ρ, and hence on ϵ, but with C_4 not depending on b, b_1 and b_2.

Lemma 3.4.5 implies that given $\epsilon''' > 0$, after possibly readjusting the value of h_0, we will also have

$$\begin{aligned}
\bar{\mu}_{\Lambda,-,h}\left(\hat{V} \leq \left(\frac{b}{h}(1-\epsilon''')\right)^2\right) &\leq C_5 \exp\left(-\frac{\beta}{h}\frac{\phi(b(1-\epsilon''')) - \phi(b)}{2}\right) \\
&\leq C_5 \exp\left(-\frac{C\epsilon'''(b_1 - B_c)}{h}\right).
\end{aligned} \tag{3.46}$$

The last step above is based on elementary calculus, and the resulting $C > 0$ is a constant which depends only on the temperature.

From (3.45) and (3.46) (with $\epsilon' = \epsilon'''$) we have

$$\bar{\mu}_{\Lambda,-,h}\left(\left(\mathcal{D}^{+}_{h,b}(\epsilon)\right)^{c}\right) \le C_6 \exp\left(-\frac{C_7(b_1 - B_c)}{h}\right). \tag{3.47}$$

To go back from $\bar{\mu}_{\Lambda,-,h}$ to $\widehat{\mu}_{\Lambda,-,h}$, observe that if we choose an appropriate ϵ''', from (3.46) and an application of Lemma 3.4.7 similar to the one above we obtain

$$\bar{\mu}_{\Lambda,-,h}(\mathcal{R}) \le C_8 \exp\left(-\frac{C_9 b_1}{h}\right).$$

Because $\mathcal{R}^c \subset \left\{\hat{V} \ge \left(\frac{b'}{h}\right)^2\right\}$, we have

$$\widehat{\mu}_{\Lambda,-,h}\left(\left(\mathcal{D}^{+}_{h,b}(\epsilon)\right)^{c}\right) = \widehat{\mu}_{\Lambda,-,h}\left(\left(\mathcal{D}^{+}_{h,b}(\epsilon)\right)^{c} \;\middle|\; \hat{V} \ge \left(\frac{b'}{h}\right)^2\right)$$
$$= \bar{\mu}_{\Lambda,-,h}\left(\left(\mathcal{D}^{+}_{h,b}(\epsilon)\right)^{c} \;\middle|\; \mathcal{R}^c\right) \le \frac{\bar{\mu}_{\Lambda,-,h}\left(\left(\mathcal{D}^{+}_{h,b}(\epsilon)\right)^{c}\right)}{\bar{\mu}_{\Lambda,-,h}(\mathcal{R}^c)}.$$

From the last three displays we obtain (3.44)

Let now $\{E_j\}$ partition $\mathcal{D}^{+}_{h,b}(\epsilon)$ according to what the outermost contour Γ in its definition is. In this fashion we obtain the following:

$$\begin{aligned}
\widehat{\mu}_{\Lambda,-,h}\left(\mathcal{B}^{-*}_{h,b}(\epsilon) \;\middle|\; \mathcal{D}^{+}_{h,b}(\epsilon)\right) &= \sum_j \alpha_j \; \widehat{\mu}_{\Lambda,-,h}\left(\mathcal{B}^{-*}_{h,b}(\epsilon) \;\middle|\; E_j\right) \\
&= \sum_j \alpha_j \; \mu_{\Lambda,-,h}\left(\mathcal{B}^{-*}_{h,b}(\epsilon) \;\middle|\; E_j\right) \\
&\le \mu_{\Lambda,+,h}\left(\mathcal{B}^{-*}_{h,b}(\epsilon/2)\right) \\
&\le C_{10}\left(\frac{b_2}{h}\right)^4 \exp\left(-\frac{C_{11} b}{h}\right) \\
&\le C_{12} \exp\left(-\frac{C_{13} b_1}{h}\right),
\end{aligned} \tag{3.48}$$

where in the second equality we used the fact that for each j, $E_j \subset \mathcal{R}^c$, in the first inequality we used the the Markov property of Gibbs distributions, the FKG-Holley inequalities and the fact that $\sum_j \alpha_j = 1$, and in the second inequality we used (1.17).

The first claim that we wanted to prove, (3.42), follows from (3.44) and (3.48). Regarding the second claim, (3.43), which refers to the case in which $b_1 > B_0$, note that then Lemmas 3.4.5 (lower bound), 3.4.6 and 3.4.7 imply

$$\mu_{\Lambda,-,h}(\mathcal{R}) \le C_{14} \exp\left(\frac{\beta}{h}\frac{\phi(b)}{2}\right) + C_1 \exp\left(-\frac{C_2 b}{h}\right)$$
$$\le C_{14} \exp\left(\frac{\beta}{h}\frac{\phi(b_1)}{2}\right) + C_1 \exp\left(-\frac{C_2 b_1}{h}\right) \le C_{15} \exp\left(-\frac{C_{16}(b_1 - B_0)}{h}\right).$$

This shows that our second claim follows from the first one, already proven, since

$$\mu_{\Lambda,-,h}(\ \cdot\ ) \leq \widehat{\mu}_{\Lambda,-,h}(\ \cdot\ ) + \mu_{\Lambda,-,h}(\mathcal{R}).$$

■

3.5. Spectral gap estimates. In this section we will prove three propositions which provide lower bounds on the spectral gap of the generator of the kinetic Ising models on some finite sets. A basic technique to be used comes from the fundamental paper [Mar], where for the first time (to our knowledge) a rigorous mathematical relation was established between relaxation times of kinetic Ising models and the equilibrium surface tension. This was done in a setting in which there is no external applied field, and the system was taken in a square box of size $l \times l$ with free boundary conditions and at low temperature (in that paper only very low temperatures where considered, but the main results were later extended up to T_c in [CGMS]). The "time to jump between the (+)-phase and the (−)-phase" was shown in that paper to behave as $\exp(\beta\bar{\tau}l)$, as $l \to \infty$, where $\bar{\tau} = \tau((1,0))$ is the surface tension in a coordinate direction.

For each finite $\Lambda \subset \mathbb{Z}^2$, $\eta \in \Omega$ and h, the process $(\sigma^{\cdot}_{\Lambda,\eta,h;t})_{t\geq 0}$ is a finite-state-space reversible irreducible Markov process and its generator has its (discrete) spectrum contained in the interval $(-\infty, 0]$, with 0 being in the spectrum. The spectral gap, denoted by $\text{gap}(\Lambda,\eta,h)$, is then simply the absolute value of the largest non-null number in the spectrum.

We will prove the three propositions below, the first two of them being needed in Section 3.2 and the third one being needed in Section 3.3. The heuristics behind these three propositions was explained in those sections.

For a recap of the terminology used in these lemmas, see the begining of Section 3.4 and the paragraph in that section which preceds Lemma 3.4.8.

Proposition 3.5.1. *For any $l_0 > 0$, any $\epsilon > 0$ and any $\bar{b} > 0$ there is $h_0 > 0$ such that if $0 < h \leq h_0$ then for any l_0-quasi-Wulff-shaped set Λ which has linear size parameter b/h with $b \leq \bar{b}$,*

$$gap(\Lambda, +, h) \ \geq \ \exp\left(-\frac{\epsilon}{h}\right).$$

Proposition 3.5.2. *For any $l_0 > 0$, any $\epsilon > 0$ and any $B_c < b_{\text{core}} < \bar{b}$ there is $h_0 > 0$ such that if $0 < h \leq h_0$ then for any l_0-quasi-Wulff-shaped set Λ which has linear size parameter b/h with $b_{\text{core}} < b \leq \bar{b}$,*

$$gap(\Lambda \backslash \Lambda^{\text{core}}, (+,-), h) \ \geq \ \exp\left(-\frac{\epsilon}{h}\right),$$

where $\Lambda^{\text{core}} = \Lambda\left(\frac{b_{\text{core}}}{h}W\right)$, and the boundary condition $(+,-)$ refers to freezing the spins up inside the core Λ^{core} and down outside Λ.

Proposition 3.5.3. *For any $\epsilon > 0$ there are $b > B_c$ and $h_0 > 0$ such that if $0 < h \leq h_0$ then*

$$gap\left(\Lambda\left(\frac{b}{h}W\right), -, h\right) \geq \exp\left(-\frac{\epsilon}{h}\right).$$

We will explain how certain results and techniques in [Mar] can be used to prove the three propositions above. The specific problem in [Mar] which is close to ours is the subject of Section 3 in that paper. This problem concerns the spectral gap for the process with no external field, in a square box of size $l \times l$ with (+)-boundary conditions. It is shown that given $\epsilon \in (0, 1/2)$, at low enough temperature there is $C > 0$ so that

$$\text{gap}(\Lambda(L), +, 0) \geq \exp(-CL^{1/2+\epsilon}). \tag{3.49}$$

Heuristically speaking, in this problem (contrary to the case of free-boundary conditions, which is the main concern in [Mar]), there is no free-energy barrier to overcome for the system either starting with all spins down or all spins up to reach equilibrium. One simply expects the (+)-phase to drift inwards from the boundary. (This indicates that the result (3.49) is far from optimal, and that the corresponding gap should be much larger than this lower bound.) This situation is similar to our problems, as stated in the propositions above. In the first two propositions we are dealing with situations without free-energy barriers, and in the third one the free-energy barrier can be made as small as needed, by adjusting the value of b.

The technique used in [Mar] to prove (3.49) consists in comparing the spectral gap for the kinetic Ising model with the one for a block-dynamics. The estimate on the spectral gap for the block-dynamics is reduced to equilibrium-statistical-mechanics problems, which are then solved.

By a block-dynamics the following is meant. Suppose that $\{\Lambda^0, ..., \Lambda^J\}$ is a finite collection of finite subsets of $\mathbb{Z}^2$ and that $\Lambda \subset \cup_{j=0,...,J}\Lambda^j$ is another finite subset of $\mathbb{Z}^2$. The block-dynamics in Λ with blocks $\{\Lambda^0, ..., \Lambda^J\}$ and with boundary condition $\eta \in \Omega_{\Lambda,-}$ will be denoted by

$$(\sigma^{\cdot}_{\Lambda,\{\Lambda^0,...,\Lambda^J\},\eta,h;t})_{t\geq 0}.$$

It is defined by updating each block $\Lambda^j \cap \Lambda$ at rate 1, independently of the other blocks, and at each update of $\Lambda^j \cap \Lambda$ replacing the configuration inside this block with a configuration chosen at random according to the Gibbs distribution $\mu_{\Lambda^j\cap\Lambda,\sigma,h}$, where σ is the current configuration. The corresponding generator is given by

$$(\mathcal{L}f)(\sigma) = \sum_{j=0,...,J} \mu_{\Lambda^j\cap\Lambda,\sigma,h}(\sigma')(f(\sigma') - f(\sigma)).$$

To state an inequality which compares the spectral gap $\text{gap}(\Lambda, \eta, h)$ with the spectral gap $\text{gap}(\Lambda, \{\Lambda^0, ..., \Lambda^J\}, \eta, h)$ of the block dynamics, we need to introduce some notation. Set

$$L_j = \max_{k\in\mathbb{Z}} |\{(x_1, x_2) \in \Lambda^j : x_2 = k\}|,$$

$$L = \max_{j=0,...,J} L_j,$$

and

$$V = \max_{j=0,\dots,J} |\Lambda^j|.$$

Theorem 2.1 in [Mar] provides us with the following bound:

$$\mathrm{gap}(\Lambda,\eta,h) \;\geq\; \frac{C_1\exp(-C_2 L)}{V}\,\mathrm{gap}(\Lambda,\{\Lambda^0,\dots,\Lambda^J\},\eta,h), \tag{3.50}$$

where C_1 and C_2 are finite positive constants which depend only on the temperature. The proof of this result in [Mar] is restricted to the case in which the blocks are of a certain type, adapted to the needs in that paper, but this restriction is clearly not relevant in the proof. Somewhat more importantly, the rates of the kinetic Ising models in [Mar] are not as general as ours, with only a special case of rates, which satisfy detailed-balance and our assumptions (H1) – (H4) being considered. This is not a problem, and indeed, once (3.50) is established for one choice of rates satisfying detailed-balance and (H4), it holds for all such rates, thanks to the fact that the spectral gap is bounded below and above by, respectively, the same equilibrium quantity multiplyed by $c_{\min}(T)$ and $c_{\max}(T)$ (for this see, e.g., equation (61) in [Sch1]).

In order to prove Propositions 3.5.1 – 3.5.3, our blocks will be Wulff-annuli, defined as follows. Given $\rho > 0$ set

$$A^0_\rho = 2\rho W$$
$$A^j_\rho = (j+2)\rho W \setminus j\rho W \quad \text{for} \quad j = 0,1,\dots$$

and for $h > 0$, and $j = 0,1,\dots$

$$\Lambda^j = \Lambda\left(\frac{1}{h}A^j_\rho\right),$$

where the value of ρ is chosen in a fashion that we describe next, and which depends on the value of ϵ and $\bar b$ in the propositions that we are proving (in the case of Proposition 3.5.3 we can choose some arbitrary $\bar b > B_c$). Note that for each $\rho > 0$, $j = 0,1,\dots$, and $r \in \mathbb{R}$, the set $A^j_\rho \cap \{(x_1,x_2) \in \mathbb{R}^2 : x_2 = r\}$ if not empty consists of either an interval or the union of two intervals, and that its Lebegue measure satisfies

$$\max_{r\in\mathbb{R}} \max_{j=0,\dots,\lfloor \bar b/\rho\rfloor} |A^j_\rho \cap \{(x_1,x_2)\in\mathbb{R}^2 : x_2 = r\}| \to 0 \quad \text{as} \quad \rho \to 0.$$

Therefore we can choose ρ small enough and take $J = \lfloor \bar b/\rho \rfloor$, so that the corresponding collection of blocks, $\{\Lambda^0,\dots,\Lambda^J\}$, satisfies for each $h > 0$,

$$L \leq \frac{\epsilon}{2C_2 h}, \tag{3.51}$$

where C_2 comes from (3.50).

Since V grows only as a power of $1/h$, from (3.50) and (3.51) we see that all that remains is to show that with the choices above and the pertinent Λ and η,

$\mathrm{gap}(\Lambda, \{\Lambda^1, ..., \Lambda^J\}, \eta, h)$ can be assured to be large enough. In the case of Proposition 3.5.1 and 3.5.2 this amounts to showing that this quantity can be bounded below by a positive constant which does not depend on h. In the case of Proposition 3.5.3, we need to show that by taking b sufficiently close to B_c this will also be the case. (The way we set things up above, some blocks Λ^j may not intersect the set Λ in some situations. This, of course, is not important, since such blocks have no effect on the dynamics. The setup above was chosen for notational convenience.)

To study $\mathrm{gap}(\Lambda, \{\Lambda^0, ..., \Lambda^J\}, \eta, h)$ one can couple the processes $(\sigma^-_{\Lambda,\{\Lambda^0,...,\Lambda^J\},\eta,h;t})_{t\geq 0}$ and $(\sigma^+_{\Lambda,\{\Lambda^0,...,\Lambda^J\},\eta,h;t})_{t\geq 0}$ in such a way that the first marginal never lies above the second one, and after they hit each other they coalesce and remain together. This can be done, for instance, via a graphical construction, in which to each block Λ^j we associate a rate 1 exponential Poisson process. The occurrence times of this Poisson process then determine the moments of update of $\Lambda^j \cap \Lambda$ in both marginal processes, and the updates are coupled in a way that preserves the order, which is possible due to the FKG-Holley inequalities. In what follows we will use $\mathbb{P}$ to denote the probabilities associated to this coupling.

The goal is now to show that there are positive finite constants t_0, and h_0 so that for all $0 < h \leq h_0$ and all Λ with which we are concerned, and corresponding η,

$$\mathbb{P}\left(\sigma^-_{\Lambda,\{\Lambda^0,...,\Lambda^J\},\eta,h;t_0} \neq \sigma^+_{\Lambda,\{\Lambda^0,...,\Lambda^J\},\eta,h;t_0}\right) \leq \frac{1}{2}. \tag{3.52}$$

From this inequality it then follows that for all $t > 0$,

$$\mathbb{P}\left(\sigma^-_{\Lambda,\{\Lambda^0,...,\Lambda^J\},\eta,h;t} \neq \sigma^+_{\Lambda,\{\Lambda^0,...,\Lambda^J\},\eta,h;t}\right) \leq \left(\frac{1}{2}\right)^{\lfloor t/t_0 \rfloor} \leq C \exp\left(-\frac{\log(1/2)}{t_0}\, t\right).$$

In particular then we have

$$\mathrm{gap}(\Lambda, \{\Lambda^0, ..., \Lambda^J\}, \eta, h) \geq \frac{\log(1/2)}{t_0},$$

as needed to complete the proofs.

Concerning the proof of (3.52), we start by observing that we have only a finite number of blocks. Therefore, if t_0 is chosen large enough it will be likely that before time t_0 a sequence of updates will have occurred in the coupled processes in which the blocks were updated in a particular, predetermined order. In the case of Propositions 3.5.1 and 3.5.3 the good order is the decreasing one, from J down to 0, while in the case of Proposition 3.5.2 it is the increasing order, from 0 up to J. The point is that at the end of a sequence of updates produced in the good order it is very likely, when h is small, that the two marginal processes will have hit each other. To show this, we will use an equilibrium estimate on how much the boundary conditions can influence the Gibbs distributions inside the annular blocks. This is the contend of the following lemma. After stating it and before going into the proof of it we will explain the idea of how it has to be used in order to prove (3.52). In the lemma we will use the notation

$$\mathcal{P}^-_{h,b}(\epsilon) = \left\{\Lambda\left(\frac{b(1-\epsilon)}{h}W\right) \stackrel{-}{\longleftrightarrow} \left(\Lambda\left(\frac{b(1+\epsilon)}{h}W\right)\right)^c\right\},$$

and

$$\mathcal{P}^+_{h,b}(\epsilon) = \left\{ \Lambda\left(\frac{b(1-\epsilon)}{h} W\right) \stackrel{+}{\longleftrightarrow} \left(\Lambda\left(\frac{b(1+\epsilon)}{h} W\right)\right)^c \right\}.$$

Lemma 3.5.1. *For any $l_0 > 0$, any $\epsilon > 0$ and any $0 < b_1 < b_2$ there are finite positive constants C_1 and C_2 such that*

(a) *If $\phi(b_1) < \phi(b_2)$ then for all Λ_1 and Λ_2 which are l_0-quasi-Wulff shaped with respective linear-size-parameters $\frac{b_1}{h}$ and $\frac{b_2}{h}$,*

$$\mu_{\Lambda_2 \setminus \Lambda_1, (+,-), h}(\mathcal{P}^+_{h,b_1}(\epsilon)) \leq C_1 \exp(-C_2/h),$$

for all $h > 0$. And if also $x \in \left(\Lambda\left(\frac{b_1+\epsilon}{h} W\right)\right)^c$ then

$$|\langle \sigma(x) \rangle_{\Lambda_2 \setminus \Lambda_1, (+,-), h} - \langle \sigma(x) \rangle_{\Lambda_2 \setminus \Lambda_1, (-,-), h}| \leq C_1 \exp(-C_2/h),$$

for all $h > 0$.

(b) *If $\phi(b_1) > \phi(b_2)$ then for all Λ_1 and Λ_2 which are l_0-quasi-Wulff shaped with respective linear-size-parameters $\frac{b_1}{h}$ and $\frac{b_2}{h}$,*

$$\mu_{\Lambda_2 \setminus \Lambda_1, (+,-), h}(\mathcal{P}^-_{h,b_2}(\epsilon)) \leq C_1 \exp(-C_2/h),$$

for all $h > 0$. And if also $x \in \Lambda\left(\frac{b_2-\epsilon}{h} W\right)$ then

$$|\langle \sigma(x) \rangle_{\Lambda_2 \setminus \Lambda_1, (+,-), h} - \langle \sigma(x) \rangle_{\Lambda_2 \setminus \Lambda_1, (+,+), h}| \leq C_1 \exp(-C_2/h),$$

for all $h > 0$.

(c) *For all Λ_1 and Λ_2 which are l_0-quasi-Wulff shaped with respective linear-size-parameters $\frac{b_1}{h}$ and $\frac{b_2}{h}$,*

$$\mu_{\Lambda_2 \setminus \Lambda_1, (-,+), h}(\mathcal{P}^-_{h,b_1}(\epsilon)) \leq C_1 \exp(-C_2/h),$$

for all $h > 0$. And if also $x \in \left(\Lambda\left(\frac{b_1+\epsilon}{h} W\right)\right)^c$ then

$$|\langle \sigma(x) \rangle_{\Lambda_2 \setminus \Lambda_1, (-,+), h} - \langle \sigma(x) \rangle_{\Lambda_2 \setminus \Lambda_1, (+,+), h}| \leq C_1 \exp(-C_2/h),$$

for all $h > 0$.

Proof of (3.52). Here we explain the idea of deriving the estimate (3.52) from the above lemma. We will do it for the case of the Proposition 3.5.1, for which the statement (c) of the lemma is used. The main observation is very simple. As the reader remembers, we are concerned with the event that the updates of the blocks happen in decreasing order, that is first the block Λ^J is updated, then the block Λ^{J-1} is, and so on. Note, that the boundary condition on the outer boundary of Λ^J is $(+)$. As we learn from the lemma, after the update of the block Λ^J, with overwhelming probability it becomes almost completely filled with $(+)$-phase, no matter what the boundary condition on the inner boundary of Λ^J is. In particular, its middle line — which is the outer boundary of the block Λ^{J-1} — is in the $(+)$-phase. So the argument can be repeated. For the complete argument the reader is referred to [Mar], Theorem 3.1.

In the final remark before going into the proof of the Lemma 3.5.1 let us explain the relations between different parts of the lemma and Propositions 3.5.1 – 3.5.3: part (a) has to be used for the Proposition 3.5.3, part (b) - for the Proposition 3.5.2, and part (c) - for the Proposition 3.5.1. The choice of b in Proposition 3.5.3 has to be done in such a way that $\phi(b) > \phi(b-\rho)$ with ρ chosen before (3.51).

Proof of Lemma 3.5.1.

In each one of the three parts of the lemma, the second claim, about the expected value of the spin at a site x, follows from the first claim in the same part of the lemma, by arguments analogous to (1.20).

Regarding the first claim in each part of the lemma, we could in principle develop a machinery similar to the one developed in Section 3.4, to deal with the Gibbs distribution inside Wulff annuli. Some technical complications would arise from the fact that such sets are not simply-connected. It turns out, nevertheless, that for our purposes we can avoid this lengthy approach, and rather use the results for simply-connected sets (more specifically, for quasi-Wulff-shaped sets) of Section 3.4, combined with some tricks involving conditioning. To stress that our approach is to some extent natural, we observe that the hypothesis in our lemma refer the values of $\phi(b_1)$ and $\phi(b_2)$, and hence somehow we should use our knowledge about quasi-Wulff-shaped boxes to prove these results on the annuli. And to make the conditioning below appear less of a trick, observe that, say in part (a), we are interested in (+)-boundary conditions in the center, something that is akin to having a droplet of the (+)-phase placed there; the conditioning places such a droplet in the center.

We turn now to the proof of the first claim in part (a) of the lemma.

Recall the definitions in the first paragraph of the proof of Lemma 3.4.8. Using that notation, we will say that Γ is a negative-outside (same as positive-inside) contour of a configuration η in case η is identically -1 on $\partial_{\text{ext}}\Gamma$ and identically $+1$ on $\partial_{\text{int}}\Gamma$. External contours of any configuration in Ω_- are negative-outside.

Let E be the event that there is a contour which surrounds Λ_1 and $E' \subset E$ be the event that such a contour is contained in $\frac{b_1(1+\epsilon/2)}{h}W$ and is negative-outside. We are going to argue that if E happens, then with very high probability E' happens as well. So we need an estimate on $\mu_{\Lambda_2,-,h}(E)$ from below, and an estimate on $\mu_{\Lambda_2,-,h}(E\cap(E')^c)$ from above. From Lemma 3.4.3 we know that the first probability is at least of the order of $\exp\left(-\frac{\beta}{h}\phi(b_1)\right)$. To obtain the second estimate, let us introduce the number $b' > b_1$, which is close enough to b_1, so that $\delta = 1-\frac{b_1}{b'}$ is small enough. The choice of δ will be made later. One sees immediately that

$$
\begin{aligned}
&\mu_{\Lambda_2,-,h}\left((E')^c \ \cap\ E\right)\\
&\quad=\mu_{\Lambda_2,-,h}\left((E')^c \cap E \cap \left[\left(\frac{b_1}{h}\right)^2 \le \hat{V} \le \left(\frac{b_2}{h}\right)^2\right]\right)\\
&\quad\le\mu_{\Lambda_2,-,h}\left((E')^c \cap E \,\middle|\, \left(\frac{b'}{h}(1-\delta)\right)^2 \le \hat{V} \le \left(\frac{b'}{h}\right)^2\right)\times\\
&\qquad\times\,\mu_{\Lambda_2,-,h}\left(\left(\frac{b'}{h}(1-\delta)\right)^2 \le \hat{V} \le \left(\frac{b'}{h}\right)^2\right)\\
&\quad+\mu_{\Lambda_2,-,h}\left(\left(\frac{b'}{h}\right)^2 \le \hat{V} \le \left(\frac{b_2}{h}\right)^2\right).
\end{aligned}
$$

Note now, that if both events $(E')^c \cap E$ and $\hat{V} \le \left(\frac{b'}{h}\right)^2$ happen, and δ is small enough, then we can claim that the following three properties hold:
there is an exterior contour surrounding Λ_1;
this contour can not be shifted so as to fit inside $\frac{b'(1+\epsilon/4)}{h}W$;
there are no other exterior contours which can surround Λ_1 even after being shifted.
In other words, we have the inclusion $(E')^c \cap E \cap \left(\hat{V} \le \left(\frac{b'}{h}\right)^2\right) \subset (\bar{E}^h_{b',\epsilon/4})^c$. So from the hypothesis that $\phi(b_1) < \phi(b_2)$ and Lemmas 3.4.3, 3.4.5 and 3.4.7 we have

$$\mu_{\Lambda_2,-,h}\left((E')^c \mid E\right) \ \le\ C_1 \exp(-C_2/h), \tag{3.53}$$

for some positive finite constants C_1 and C_2.

Set

$$F = \left\{\Lambda\left(\frac{b_1(1+\epsilon/2)}{h}W\right) \overset{+}{\longleftrightarrow} \left(\Lambda\left(\frac{b_1(1+\epsilon)}{h}W\right)\right)^c\right\},$$

Partitioning E' according to what the innermost negative-outside contour around Λ_1 is and using the FKG-Holley inequalities we obtain

$$\mu_{\Lambda_2,-,h}\left(F \mid E'\right) \le \mu_{\Lambda_2,-,h}\left(F\right).$$

In order to have $\phi(b_1) < \phi(b_2)$ we must have $b_2 < B_0$. Clearly also, for some $\epsilon' > 0$ which depends on ϵ and b_1, $F \cap \Omega_{\Lambda_2,-} \ \subset\ \left\{\mathcal{W} \ge \frac{\epsilon'}{h}\right\}$. Therefore, if we choose a conveniently small $\epsilon'' > 0$ and use Lemmas 3.4.2, and 3.4.5 we obtain

$$
\begin{aligned}
\mu_{\Lambda_2,-,h}\left(F\right) &\le \mu_{\Lambda_2,-,h}\left(\mathcal{W} \ge \frac{\epsilon'}{h}\ ,\ \hat{V} \le \left(\frac{\epsilon''}{h}\right)^2\right) + \mu_{\Lambda_2,-,h}\left(\hat{V} \ge \left(\frac{\epsilon''}{h}\right)^2\right)\\
&\le C_1 \exp(-C_2/h).
\end{aligned}
$$

Combining the various inequalities displayed above, we have

$$
\begin{aligned}
\mu_{\Lambda_2,-,h}\left(F \mid E\right) &\le \mu_{\Lambda_2,-,h}\left((E')^c \mid E\right) + \mu_{\Lambda_2,-,h}\left(F \mid E'\right)\\
&\le C_1 \exp(-C_2/h).
\end{aligned} \tag{3.54}
$$

To proceed, we have to introduce a new notion. We will call a (*)-circuit $\gamma = x_1, \ldots, x_n$ a c-circuit iff there exists a configuration $\sigma \in \Omega_-$ with one contour, Γ, such that $\partial_+\Gamma = \{x_1, \ldots, x_n\}$. (Note that our definition depends on the choice of the splitting rules, used in connection with transforming of contours into closed curves.) The interior $\mathrm{Int}\,\gamma$ of a c-circuit γ is by definition the interior of the corresponding contour Γ. Let γ_1, γ_2 be two c-circuits; we define the intersection c-circuits δ_k in the following way. Let $I = \mathrm{Int}\,\gamma_1 \cap \mathrm{Int}\,\gamma_1$ and I_k be connected components of I in the sense of the above mentioned splitting rules. Since I is simply-connected, so are the I_k-s. Let δ_k be c-circuits, such that $\mathrm{Int}\,\delta_k = I_k$. For immediate use we need the following property of c-circuits: let γ_1 and γ_2 be two c-circuits, δ be one of the intersection c-circuits, and a configuration σ be given, such that both γ_1 and γ_2 are $(+*)$-circuits of it. Then so is the circuit δ. To see this, let us introduce the corresponding contours Γ_1, Γ_2 and Δ, and let first a site x of the c-circuit δ be at the distance $\frac{1}{2}$ from some bond b of Δ. But this bond b evidently belongs to at least one of the contours Γ_1, Γ_2, while x is inside this contour, hence $\sigma(x) = +1$. In the remaining case the distance $\mathrm{dist}(x, \Delta) = \frac{\sqrt{2}}{2}$, and in such a situation there are two adjacent bonds b_1, b_2 of Δ, with $\mathrm{dist}(x, b_1) = \mathrm{dist}(x, b_2) = \frac{\sqrt{2}}{2}$, and two nearest neighbors x_1, x_2 of x, belonging to δ, such that $\mathrm{dist}(x_i, b_i) = \frac{1}{2}, i = 1, 2$. Consider another nearest neighbor y of the sites x_1, x_2. It stays outside Δ, hence it also stays outside at least one of the contours Γ_1, Γ_2, while all three sites x, x_1, x_2 are inside both of them. Hence both bonds b_1, b_2 belong to that contour, which again implies that $\sigma(x) = +1$.

Let E'' be the event that some c-circuit γ which surrounds Λ_1 and is contained in $\frac{b_1(1+\epsilon/2)}{h}W$ is a $(+*)$-circuit. Let $\{E''_j\}$ be the partition of E'' according to what the innermost such circuit γ is. (The preceding paragraph ensures the existence of such an innermost $(+)$-c-circuit.) Using the fact that for each j we have $\Omega_{\Lambda_2,-} \cap E''_j \subset E$, the FKG-Holley inequalities and (3.53) we obtain

$$
\begin{aligned}
\mu_{\Lambda_2,-,h}\left(F \mid E\right) &\geq \sum_j \mu_{\Lambda_2,-,h}\left(F \mid E''_j \cap E\right)\ \mu_{\Lambda_2,-,h}\left(E''_j \mid E\right) \\
&= \sum_j \mu_{\Lambda_2,-,h}\left(F \mid E''_j\right)\ \mu_{\Lambda_2,-,h}\left(E''_j \mid E\right) \\
&\geq \mu_{\Lambda_2\setminus\Lambda_1,(+,-),h}(F)\ \sum_j \mu_{\Lambda_2,-,h}\left(E''_j \mid E\right) \\
&= \mu_{\Lambda_2\setminus\Lambda_1,(+,-),h}(F)\ \mu_{\Lambda_2,-,h}\left(E'' \mid E\right) \\
&\geq \mu_{\Lambda_2\setminus\Lambda_1,(+,-),h}(F)\ \mu_{\Lambda_2,-,h}\left(E' \mid E\right) \\
&\geq \frac{1}{2}\ \mu_{\Lambda_2\setminus\Lambda_1,(+,-),h}(F),
\end{aligned}
\tag{3.55}
$$

for small h.

Comparing (3.54) with (3.55) we obtain

$$
\mu_{\Lambda_2\setminus\Lambda_1,(+,-),h}(\mathcal{P}^+_{h,b_1}(\epsilon)) \leq \mu_{\Lambda_2\setminus\Lambda_1,(+,-),h}(F) \leq 2C_1\exp(-C_2/h).
$$

This completes the proof of part (a) of the lemma, and we turn to to the proof of the first claim in part (b) of the lemma.

Let this time E be the event that there is a contour which surrounds $\frac{b_1(1-\epsilon)}{h}W$ and E' be the event that there is a contour which surrounds $\frac{b_2(1-\epsilon/2)}{h}W$. With no loss of generality, we will suppose that ϵ is small enough so that $E' \subset E$.

From the hypothesis that $\phi(b_1) > \phi(b_2)$ and Lemmas 3.4.5 and 3.4.7 we have

$$\mu_{\Lambda_2,-,h}\left((E')^c \mid E\right) \ \leq \ C_1 \exp(-C_2/h), \tag{3.56}$$

for some positive finite constants C_1 and C_2.

Therefore also

$$\mu_{\Lambda_2,-,h}\left(\mathcal{P}^-_{h,b_2}(\epsilon) \,\middle|\, E\right) \ \leq \ C_1 \exp(-C_2/h). \tag{3.57}$$

Let this time E'' be the event that there is a (+)-circuit which surrounds $\frac{b_1(1-\epsilon)}{h}W$. and is contained in $\frac{b_1(1-\epsilon/2)}{h}W$. Let $\{E''_j\}$ partition E'' according to what the innermost such (+)-circuit is.

By partitioning $(\mathcal{P}^-_{h,b_2}(\epsilon))^c$ according to what the (−)-cluster of $(\Lambda_2)^c$ is and using the FKG-Holley inequalities combined with (1.17) and the graph-theoretic duality between connectivity and (*)-connectivity, in addition to (3.57), we obtain

$$\begin{aligned} \mu_{\Lambda_2,-,h}\left((E'')^c \mid E\right) &\leq \mu_{\Lambda_2,-,h}\left(\mathcal{P}^-_{h,b_2}(\epsilon) \,\middle|\, E\right) + \mu_{\Lambda_2,-,h}\left((E'')^c \,\middle|\, \left(\mathcal{P}^-_{h,b_2}(\epsilon)\right)^c\right) \\ &\leq C_1 \exp(-C_2/h). \end{aligned} \tag{3.58}$$

Using the fact that for each j we have $\Omega_{\Lambda_2,-} \cap E''_j \ \subset \ E$, the FKG-Holley inequalities and (3.58) we obtain

$$\begin{aligned} \mu_{\Lambda_2,-,h}\left(\mathcal{P}^-_{h,b_2}(\epsilon) \,\middle|\, E\right) &\geq \sum_j \mu_{\Lambda_2,-,h}\left(\mathcal{P}^-_{h,b_2}(\epsilon) \,\middle|\, E''_j \cap E\right) \ \mu_{\Lambda_2,-,h}\left(E''_j \mid E\right) \\ &= \sum_j \mu_{\Lambda_2,-,h}\left(\mathcal{P}^-_{h,b_2}(\epsilon) \,\middle|\, E''_j\right) \ \mu_{\Lambda_2,-,h}\left(E''_j \mid E\right) \\ &\geq \mu_{\Lambda_2\setminus\Lambda_1,(+,-),h}\left(\mathcal{P}^-_{h,b_2}(\epsilon)\right) \ \sum_j \mu_{\Lambda_2,-,h}\left(E''_j \mid E\right) \\ &= \mu_{\Lambda_2\setminus\Lambda_1,(+,-),h}\left(\mathcal{P}^-_{h,b_2}(\epsilon)\right) \ \mu_{\Lambda_2,-,h}\left(E'' \mid E\right) \\ &\geq \frac{1}{2} \ \mu_{\Lambda_2\setminus\Lambda_1,(+,-),h}\left(\mathcal{P}^-_{h,b_2}(\epsilon)\right), \end{aligned} \tag{3.59}$$

for small h.

Comparing (3.57) with (3.59) we obtain

$$\mu_{\Lambda_2\setminus\Lambda_1,(+,-),h}\left(\mathcal{P}^+_{h,b_1}(\epsilon)\right) \leq 2C_1 \exp(-C_2/h).$$

This completes the proof of part (b) of the lemma, and we turn to to the proof of the first claim in part (c) of the lemma.

Because in Section 3.4 we did not study systems with (+) boundary conditions, we will use a somewhat artificial approach to part (c), by reducing it to part (a), studied above. (In doing so we will proceed as someone who first heats up cold water in order to freeze it later, because he knows how to freeze hot water, but has never frozen cold water). First note that by reversing all signs and then using the FKG-Holley inequalities, we obtain, for any $h' \geq -h$,

$$\mu_{\Lambda_2\setminus\Lambda_1,(-,+),h}\left(\mathcal{P}^-_{h,b_1}(\epsilon)\right) = \mu_{\Lambda_2\setminus\Lambda_1,(+,-),-h}\left(\mathcal{P}^+_{h,b_1}(\epsilon)\right) \leq \mu_{\Lambda_2\setminus\Lambda_1,(+,-),h'}\left(\mathcal{P}^+_{h,b_1}(\epsilon)\right).$$

Suppose that $h' > 0$ and set

$$b_1' = \frac{h'}{h}b_1 \qquad \text{and} \qquad b_2' = \frac{h'}{h}b_2.$$

If h' is small enough we have $0 < b_1' < b_2' < B_c$ and hence $\phi(b_1') < \phi(b_2')$. But Λ_1 and Λ_2 are l_0-quasi-Wulff shaped with respective linear-size-parameters $\frac{b_1}{h} = \frac{b_1'}{h'}$ and $\frac{b_2}{h} = \frac{b_2'}{h'}$. So we can quote part (a) of the lemma to conclude the proof of part (c).

■

References

[AL] M. Aizenman and J. L. Lebowitz, *Metastability effects in bootstrap percolation*, J. Phys. A **21** (1988), 3801–3813.

[BM] K. Binder and H. Müller-Krumbhaar, *Investigation of metastable states and nucleation in the kinetic Ising model*, Physical Review B **9** (1974), 2328–2353.

[CCO] D. Capocaccia, M. Cassandro and E.Olivieri, *A study of metastability in the Ising model*, Commun. Math. Phys. **39** (1974), 185–205.

[CCS] J. Chayes, L. Chayes and R. H. Schonmann, *Exponential decay of connectivities in the two-dimensional Ising model*, J. Stat. Phys. **49** (1987), 433–445.

[CGMS] F. Cesi, G. Guadagni, F. Martinelli, and R. H. Schonmann On the 2D dynamical Ising model in the phase coexistence region near the critical point, J. Stat. Phys. **85** (1996), 55–102.

[DS] P. Dehghanpour and R. H. Schonmann, *Metropolis dynamics relaxation via nucleation and growth*, preprint (1996).

[DKS] R. L. Dobrushin, R. Kotecký and S. B. Shlosman, *Wulff construction: a global shape from local interaction*, AMS translations series, Providence (Rhode Island), 1992.

[Geo] H.-O. Georgii, *Gibbs measures and phase transitions*, Walter de Gruyter, Berlin, New York, 1988.

[GD] J. D. Gunton and M. Droz, *Introduction to the theory of metastable and unstable states*, Lecture Notes in Physics, vol. 183, Springer-Verlag, Berlin, Heidelberg, New York, 1983.

[Iof1] D. Ioffe, *Large deviations for the 2D Ising model: a lower bound without cluster expansions*, J. Stat. Phys. **74** (1994), 411–432.

[Iof2] D. Ioffe, *Exact large deviation bounds up to T_c for the Ising model in two dimensions*, Prob. Th. rel. Fields **102** (1995), 313–330.

[Isa] S. N. Isakov, *Nonanalytic features of the first order phase transition in the Ising model*, Commun. Math. Phys. **95** (1984), 427–443.

[KO] R.Kotecký and E.Olivieri, *Droplet dynamics for asymmetric Ising model*, J. Stat. Phys. **70** (1993), 1121–1148.

[Lig] T. Liggett, *Interacting Particle Systems*, Springer-Verlag, Berlin, Heidelberg, New York, 1985.

[Mar] F. Martinelli, *On the two-dimensional dynamical Ising model in the phase coexistence region*, J. Stat. Phys. **76** (1994), 1179–1246.

[M-L] A.Martin-Löf, *Mixing properties, differentiability of the free energy and the central limit theorem for a pure phase in the Ising model at low temperature*, Commun. Math. Phys. **32** (1973), 75–92.

[PL] O. Penrose and J. L. Lebowitz, *Towards a rigorous molecular theory of metastability*, in *Fluctuation Phenomena* (second edition), E. W. Montroll, J. L. Lebowitz, editors, North–Holland Physics Publishing (1987).

[Pfi] C.E.Pfister, *Large deviations and phase separation in the two-dimensional Ising model*, Helv. Phys. Acta **64** (1991), 953–1054.

[RTMS] P. A. Rikvold, H. Tomita, S. Miyashita, and S. W. Sides, *Metastable lifetimes in a kinetic Ising model: dependence on field and system size*, Phys. Rev. E **49** (1994), 5080–5090.

[Sch1] R. H. Schonmann, *Slow droplet-driven relaxation of stochastic Ising models in the vicinity of the phase coexistence region*, Commun. Math. Phys. **161** (1994), 1–49.

[Sch2] R. H. Schonmann, *Theorems and conjectures on the droplet driven relaxation of stochastic Ising models*, in *Probability theory of spatial disorder and phase transition*, G. Grimmett, ed., Kluwer Publ. Co (1994), 265–301.

[SS1] R. H. Schonmann and S. B. Shlosman, *Complete analyticity for 2D Ising completed*, Commun. Math. Phys. **170** (1995), 453–482.

[SS2] R. H. Schonmann and S. B. Shlosman, *Constrained variational problem with applications to the Ising model*, J. Stat. Phys. **83** (1996), 867–905.

R.H.Schonmann
Mathematics Department
University of California at Los Angeles
Los Angeles, CA 90024, U.S.A.

S.B.Shlosman
Mathematics Department
University of California at Irvine
Irvine, CA 92697, U.S.A.

and

Institute for the Information Transmission Problems
Russian Academy of Sciences
Moscow, Russia