Commun. Math Phys 161, 1–49 (1994)

Communications in Mathematical Physics

Slow Droplet-Driven Relaxation of Stochastic Ising Models in the Vicinity of the Phase Coexistence Region

Roberto H. Schonmann★
Mathematics Department, UCLA, Los Angeles, CA 90024, U S.A

Received: 31 December 1992/in revised form: 18 June 1993

Abstract: We consider the stochastic Ising models (Glauber dynamics) corresponding to the infinite volume basic Ising model in arbitrary dimension $d \geqq 2$ with nearest neighbor interaction and under a positive external magnetic field h. Under minimal assumptions on the rates of flip (so that all the common choices are included), we obtain results which state that when the system is at low temperature T, the relaxation time when the evolution is started with all the spins down blows up, when $h \searrow 0$, as $\exp(\lambda(T)/h^{d-1})$ (the precise results are lower and upper bounds of this form). Moreover, after a time which does not scale with h and before a time which also grows as an exponential of a multiple of $1/h^{d-1}$ as $h \searrow 0$, the law of the state of the process stays, when h is small, close to the minus-phase of the same Ising model without an external field. These results may be considered as a partial vindication of a conjecture raised by Aizenman and Lebowitz in connection to the metastable behavior of these stochastic Ising models.

Contents

★ Partially supported by NSF, under grant DMS 91-00725

1. Introduction

In this paper we address the old question of understanding the relaxation patterns of stochastic Ising models in the vicinity of the phase transition region, i.e., at low temperature and under small but non-null external field h, when initially they are far from equilibrium (for instance close to the equilibrium state with opposite value of the external field). We will consider the basic models on $\mathbb{Z}^d$, with formal Hamiltonian

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

where $\sigma(x) = \pm 1$ is the spin at the site $x \in \mathbb{Z}^d$, and the first sum runs over pairs of sites which are nearest neighbors in $\mathbb{Z}^d$, each pair counted only once. The time evolution is introduced as a spin flip Markov process which is reversible with respect to the corresponding Gibbs measures at temperature T. The flip rates will be supposed to satisfy certain regularity conditions, but those will be very mild, so that essentially all common choices for these rates will be covered by our results.

These systems that we are considering are probably the favorite model systems for investigators addressing the issue of relaxation to equilibrium of systems close to a first order phase transition. The literature on the subject is vast, because the problem is of interest to researchers in such diverse areas as metallurgy, chemistry, physics and probability. One of the conspicuous features of relaxation phenomena close to discontinuous transitions is the presence of metastable behavior, in which the system seems, for a long time, to have reached equilibrium, but in a state which is actually far from the true equilibrium state, and is close to what the equilibrium would be for values of the parameters at the other side of the transition region. A considerable number of review papers and monographs has been written on the subjects of metastability and relaxation close to transition regions. The reader may consult for instance, [GD, GSS and Koc] for accounts which emphasize non-rigorous results. A good review of rigorous investigations on the problem of metastability is [PL].

Many papers have been written on simulations of the stochastic Ising models in the regime which concerns us. The reader will find a large number of references in the reviews quoted above and a constant stream of papers on the subject in more recent issues of journals in statistical mechanics. We want to point out only two aspects of the relaxation pattern which were put in evidence by such simulations and that are of great relevance in this paper. The first one is the notion of "plateaus" in relaxation curves, as emphasized by Binder and Müller-Krubhaar, in their classic paper [BM]. Suppose that one runs the dynamics under a small positive external field, starting from all spins down (and uses periodic or free boundary conditions). One is interested in the time evolution of a local observable, say, the value of the spin at the origin. An average is taken over a large number of independent repetitions of the same evolution from time 0 up to a certain time. Under these conditions there is manifestation of metastable behavior in the form of a "plateau" in the relaxation curve that is obtained: in a relatively short time the average value of the spin at the origin seems to converge to a value close to the opposite of the spontaneous magnetization, after this, one sees an apparent flatness in the relaxation curve over a stretch 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 constant value and move upwards, towards the true asymptotic limit, close to the spontaneous magnetization. The experimentally-almost-flat portion of the relaxation curve is referred to as a plateau. Of course, for given values of the parameters T and h the relaxation curve is strictly monotone increasing, and there is no clear cut definition of what the plateau is. On the other hand, repeating the numerical experiment with smaller and smaller values of h (at the same temperature T) one sees that the flatness becomes more and more evident, in the sense that the first portion of the relaxation curve, which is observed while the system is moving towards its "metastable state" becomes essentially independent of h, while the length of the apparent plateau increases (see Figs. 6, 7 and 8 in [BM]).

Here are some of the reasons the problem captured the attention of mathematicians: can one still make and prove precise mathematical statements which rephrase the idea that a plateau seems to be approached for a very long time but that eventually the system will move away from that plateau? Can one estimate how long the plateau is when h is very small? More generally, what can be said about the asymptotic behavior of the system in the limit $h \searrow 0$, in which we approach the phase transition region?

The second aspect of the relaxation pattern which can be seen in the simulations and which is another reason for much of the interest in the problem is the particularly relevant role played by the behavior of individual droplets of spins $+1$ (possibly with holes where the spins are $-$) in the sea of spins -1, during the evolution. In the "metastable state" one sees such droplets appearing spontaneously throughout the system, but shrinking and disappearing before they become large, in a sort of equilibrium which resembles the minus-phase. Eventually one of these droplets grows to a larger size, apparently by chance, and then it keeps growing and eventually "covers" the whole system, which is then in the true equilibrium phase. While this droplet is growing, it sometimes happens that other large droplets appear somewhere else and also grow, so that the system is, in this case, driven to equilibrium when such droplets coalesce and "cover" the system. This phenomenon, which is also observed in real experiments (see the reviews quoted above), is known as "nucleation and growth." Many theoretical and numerical studies have focused on these aspects of the evolution and on simplified, single-droplet, or independent-droplets, models. It is a common saying that one can "understand" the behavior of the individual droplets on purely "energetic," or rather "free-energetic" terms, as a problem of escaping from a potential well. A very heuristic form of this reasoning will be reviewed later in this paper, and indeed served to orientate our approach towards proving rigorous results.

Recently, in collaboration with E.J. Neves, the author introduced, in [NS1, NS2 and Sch2], an approach which gave a precise mathematical meaning to the notion of critical droplets and metastability for the same stochastic Ising models considered in the present paper, but in the different regime in which the external field h is held small but fixed, and the temperature T is scaled to 0. For a review of this project, in the stage it was in mid 1990, see [Sch1], where also references are given to papers which motivated the approach and other related papers, including those by Martinelli, Olivieri and Scoppola on rapid convergence to equilibrium and

Swensen–Wang dynamics. More recently further results on these lines appeared in the work of Kotecký and Olivieri, [KO1, KO2, KO3], who considered the same type of time evolution, but for different Hamiltonians, obtaining interesting differences between the correct patterns of relaxation and some "common wisdom," at least in this regime. At the time that this paper was being written up, Scoppola was finishing [Sco], in which she presents a general approach to problems of this type, based on the separation of the relevant time scales.

The approach addressed in the previous paragraph, to which we refer as "the limit of very low temperatures," is helping to clarify the way in which droplets behave, and how this affects the evolution of the systems. It also points out the presence of metastability effects which, in an asymptotic sense, are sharply defined, in a regime which is not quite the same as the one usually considered in the numerical and real experiments and in most of the theoretical non-rigorous study of the problem, but which is, nevertheless, still close to that regime. The relaxation patterns of stochastic Ising models are now relatively well understood, at a mathematically rigorous level, in this limit of very low temperatures.

In the present paper we address the more challenging and, from the point of view of physics, also more interesting regime in which the temperature is kept fixed and the external field is scaled to 0 (from the positive side, say). The problem is considerably harder in this regime, and we were able to say far less than in the other case, so far. Also in this regime the consideration of individual droplets turned out to be a key notion in the analysis. The results essentially vindicate a conjecture raised by Aizenman and Lebowitz in [AL], that gives a precise meaning to the following two statements:

i) The relaxation time grows in this regime as an exponential of $1/h^{d-1}$.
ii) Moreover, before a time which grows also as an exponential of $1/h^{d-1}$ the system stays in a metastable situation, in which locally it is close to the minus-phase, i.e., the equilibrium measure under no external field obtained as a limit of equilibrium measures under a vanishing negative external field.

Another way to phrase it is by saying that in the relaxation curve of any local observable there is indeed, when h is small, a part which is almost a flat plateau, where the expected value of the observable is close to the equilibrium value in the minus-phase. This plateau has a length at least of the order of an exponential of $1/h^{d-1}$, and this estimate is, to some extent, optimal, since the system also relaxes to equilibrium in a time of the order of another exponential of $1/h^{d-1}$. These quantitative features of the relaxation curves were observed in simulations by Stauffer, published in [Sta].

Before we can state the theorems, we need to introduce precise definitions. This is done in the next subsection, where we also review basic techniques and results, mostly without proofs.

The main result in this paper was announced in [Sch3].

1-i. The Models and Basic Techniques. In this subsection 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 subsection, finding few things with which they are not familiar. Most statements are made without proof, and we refer readers to the books [Lig1] and [Rue], and other references therein, if they need explanation.

The lattice. We will consider models on the lattices $\mathbb{Z}^d$, where d is the space dimensionality. Because the dimension d will in general be arbitrary but fixed, we will omit it in most of the notation. The cardinality of a set $\Gamma \subset \mathbb{Z}^d$ will be denoted by $|\Gamma|$. The family of finite subset of $\mathbb{Z}^d$ will be denoted by $\mathscr{F}$. For each $x \in \mathbb{Z}^d$, we define the usual norms $\|x\|_p = (|x_1|^p + \cdots + |x_d|^p)^{1/p}$, $p > 0$ finite, and $\|x\|_\infty = \max\{|x_1|, \ldots, |x_d|\}$. The interior and exterior boundaries of a set $\Gamma \subset \mathbb{Z}^d$ will be denoted, respectively by

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

and

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

For integer i, we introduce the notation

$$V_i = \{x \in \mathbb{Z}^d\colon \|x\|_\infty \leqq i\} ,$$

for the box centered at the origin which has side-length $2i + 1$. But because usually the side-length of such a box is of particular importance for us, we will mostly be using the alternative notation

$$\Lambda(l) = \text{largest } V_i \text{ which has side-length not larger than } l .$$

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

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

Given a set $\Gamma \in \mathscr{F}$ we define also

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

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

A chain is a sequence of distinct sites $x_1, \ldots, x_n$, with the property that for $i = 1, \ldots, n - 1$, $\{x_i, x_{i+1}\} \in \mathbb{B}$. The sites x_1 and x_n are called the end-points of the chain $x_1, \ldots, x_n$. 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.

The configurations and observables. At each site in $\mathbb{Z}^d$ 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 $x \in \mathbb{Z}^d$. Two configurations are specially relevant: $-\underline{1}$ and $+\underline{1}$, which are, respectively, the ones with all spins -1 and $+1$. When these configurations appear as a subscript or superscript, we will usually abbreviate them by, respectively, $-$ and $+$. 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 $\Gamma \in \mathscr{F}$ and a configuration $\eta \in \Omega$, we introduce

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

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, $f\colon \Omega \to \mathbb{R}$ is a local observable if there exists a set $S \in \mathscr{F}$ 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. Clearly, if f is a local observable, then $\|f\|_\infty < \infty$. 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:

$$\eta \leqq \zeta \text{ if } \eta(x) \leqq \zeta(x) \text{ for all } x \in \mathbb{Z}^d .$$

A particularly important role will be played in this paper by the non-decreasing local observables. Clearly every local observable is of bounded variation, and, as such, can be written as the difference between two non-decreasing ones.

A $-$ chain in a configuration σ, or simply a σ-chain, is a chain of sites, $x_1, \ldots, x_n$, as defined above, with the property that for each $i = 1, \ldots, n$, $\sigma(x_i) = -1$. The $-$ clusters in a configuration σ are the connected components of the set of sites where the spin is -1 in the configuration σ. A $-$ cluster is called infinite if it contains infinitely many sites.

Contours. Contours are important tools in describing and counting configurations with certain properties. We adopt the following definition. A contour, γ, is a set of bonds which separates $\mathbb{Z}^d$ into a finite connected set, $\Theta(\gamma)$, and an infinite set, in the sense that:

i) If $x_1, \ldots, x_n$ is a chain which has one end-point in $\Theta(\gamma)$ and the other one outside this set, then there is $i \in \{1, \ldots, n-1\}$ such that $\{x_i, x_{i+1}\} \in \gamma$, i.e., this chain must "cross γ."
ii) If γ' is strictly contained in γ, then the property above fails if we replace γ by γ'. In other words, if we remove any bond from γ, we do not separate anymore the lattice $\mathbb{Z}^d$ into two sets.

The set of sites $\Theta(\gamma)$ is called the interior of the contour γ, and its complement is called the exterior of γ. We also say that the spins in $\Theta(\gamma)$ are surrounded by γ. The number of elements (bonds) in γ is called the size, or surface, of γ and denoted by $|\gamma|$. The number of sites surrounded by γ, $|\Theta(\gamma)|$, is called the volume of γ. The interior and exterior boundaries of γ are defined, respectively, by

$$\partial_{\text{int}}\gamma := \{x \in \Theta(\gamma): \{x, y\} \in \gamma \text{ for some } y \notin \Theta(\gamma)\} ,$$

$$\partial_{\text{ext}}\gamma := \{x \notin \Theta(\gamma): \{x, y\} \in \gamma \text{ for some } y \in \Theta(\gamma)\} .$$

A contour γ is said to border on a set $\Gamma \subset \mathbb{Z}^d$ if Γ intersects the interior or exterior boundary of γ. In this case we also say that Γ is adjacent to γ. A contour γ_1 is said to surround another contour γ_2 in case $\Theta(\gamma_2) \subset \Theta(\gamma_1)$. We say that a contour γ is inside a set $\Gamma \in \mathscr{F}$ in case $\Theta(\gamma) \subset \Gamma$.

Given a configuration $\sigma \in \Omega$, we say that the contour γ is present in σ if, in the configuration σ, the spins in $\partial_{\text{ext}}\gamma$ all have the same sign and the spins in $\partial_{\text{int}}\gamma$ all have the opposite of this sign. An outer contour present in a configuration is a contour which is present in this configuration and is not surrounded by any other contour present in this configuration. A contour γ which is present in a configuration σ is said to be a positive (resp. negative) contour in this configuration if the spins in $\partial_{\text{int}}\gamma$ are all positive (resp. negative) in the configuration σ. A family of distinct contours is said to be compatible if there is at least one configuration in which all these contours are present.

Given a set $\Gamma \in \mathscr{F}$, every configuration in $\Omega_{\Gamma,-}$ can be identified by the collection of contours present in this configuration. The same is true for configurations in $\Omega_{\Gamma,+}$. We will use the notation $\Omega(\gamma_1, \ldots, \gamma_n)$ to denote the event that the contours $\gamma_1, \ldots, \gamma_n$ are all present as outer contours.

One can visualize better the contours by means of the following construction. Consider $\mathbb{Z}^d$ embedded in $\mathbb{R}^d$, and for each $x \in \mathbb{Z}^d$, set $Q(x) := \{y \in \mathbb{R}^d : |x_i - y_i| \leqq 1/2 \text{ for } i = 1, \ldots, d\}$. To each bond $\{x, y\}$, we associate what is called the face between x and y, and is defined as $F_{\{x,y\}} := Q(x) \cap Q(y)$. Given a contour γ, one introduces $\hat{\Theta}(\gamma) := \bigcup_{x \in \Theta(\gamma)} Q(x)$ and $\hat{\gamma} := \bigcup_{\{x,y\} \in \gamma} F_{\{x,y\}}$. $\hat{\Theta}(\gamma)$ is a solid of volume $|\Theta(\gamma)|$ and $\hat{\gamma}$ is its boundary, whose surface is $|\gamma|$. An immediate consequence of this construction is the following isoperimetric inequality, valid for every contour γ.

$$|\Theta(\gamma)| \leqq \left(\frac{|\gamma|}{2d}\right)^{d/(d-1)} . \tag{2}$$

More generally, if $\gamma_1, \ldots, \gamma_n$ is a family of contours, then

$$\sum_{i=1,\ldots,n} |\Theta(\gamma_i)| \leqq \left(\frac{\sum_{i=1,\ldots,n} |\gamma_i|}{2d}\right)^{d/(d-1)} . \tag{3}$$

The first of these inequalities can be derived from Theorem 1.1 in [Tay], in the following way. Define a surface tension which is 1 for planes perpendicular to each one of the coordinate directions and is $+\infty$ otherwise. One can easily see that the corresponding Wulff shape is a cube, and hence (2) follows from the theorem to which we referred above. In the case $d = 2$ we can also provide a simpler argument: let a and b be the sides of the smallest rectangle which circumscribes the contour γ. Then the following two inequalities are easily checked:

$$|\Theta(\gamma)| \leqq ab \quad \text{and} \quad |\gamma| \geqq 2(a+b) .$$

These inequalities imply (2). The inequality (3) follows easily from the same arguments used to prove (2) and the observation that by translating the sets $\hat{\Theta}(\gamma)$ we can make them coalesce into a single solid of volume $\sum_{i=1,\ldots,n} |\Theta(\gamma_i)|$ and surface not larger than $\sum_{i=1,\ldots,n} |\gamma_i|$.

Finally we recall the exponential upper bounds on the number of choices for families of compatible contours which have to satisfy simultaneously two types of constraints: all the contours must intersect a certain set of bonds S, which is fixed and has cardinality k, and the sum of the sizes of the contours in the family has to be a certain number l. Let $M(S, l)$ be the number of different choices of families of contours with these properties, then there exists $b < \infty$, which depends only on the dimension, such that

$$M(S, l) \leqq b^{l+k} . \tag{4}$$

One way to prove (4) is to observe that the set of bonds can be thought of as the set of vertices of a graph, in which two bonds v_1 and v_2 are connected by an edge if and only if the corresponding faces, F_{v_1} and F_{v_2}, have a non-empty intersection. Observe that the graph that we introduce in this fashion is infinite but of bounded degree. Say that a set of vertices is connected in case for every pair of vertices in this set there is a sequence of vertices also in the set, starting from one of these two vertices and finishing at the other, with the property that successive vertices in this sequence are connected by an edge of the graph. Contours have to be connected sets in this sense, and so the proof of (4) is reduced to the proof that for an arbitrary graph with bounded degree, the number of ways in which we can choose a set V, with l vertices, which is the union of connected sets, each one of which has a non-empty intersection with a given fixed set of vertices, S, of cardinality k, is bounded above by b^{l+k}, for some finite b, which may depend on the graph. A proof of this inequality, in the case in which the set S is a singleton, so that the set of vertices which is obtained has actually to be connected, can be found for instance in [RL], Proposition 2 in Sect. 4. For the general case, let $s_1, \ldots, s_n$ be the elements of S, ordered in some arbitrary fashion. Divide each set of vertices V, with the desired properties, into the family of its maximal connected components, and associate to each of these components the smallest vertex in S contained in it. Now associate to V the sequence of numbers $l_1, \ldots, l_k$, where l_i is 0 if the vertex s_i is not associated to any of the connected components, and otherwise is the size of the connected component to which s_i is associated. Recall that the number of ways in which we can pick k non-negative integer numbers $l_1, \ldots, l_k$, constrained by $\sum_{i=1, \ldots, k} l_i = l$ is

$$\binom{l+k-1}{l} \leqq 2^{l+k} .$$

The result now follows easily from the case $k = 1$, at the cost of doubling the value of b.

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

$$\lim_{n\to\infty} \int f d\mu_n = \int f d\nu \text{ for every continuous observable } f . \tag{5}$$

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

$$\int f d\mu \leqq \int f d\nu \text{ for every continuous non-decreasing observable } f . \tag{6}$$

Because the local observables are dense in the set of continuous observables, we can restrict ourselves to the local ones in (5) and (6). Moreover, because every local observable is the difference between two non-decreasing ones, we can also restrict ourselves to those in (5); we will make heavy use of this observation.

The Gibbs measures. We will consider always the formal Hamiltonian (1). In order to give precise definitions, we define, for each set $\Gamma \in \mathscr{F}$ and each boundary condition $\eta \in \Omega$,

$$H_{\Gamma,\eta,h}(\sigma) = -\frac{1}{2} \sum_{\{x,y\}\in \mathbb{B}_\Gamma} \sigma(x)\sigma(y) - \frac{1}{2} \sum_{\substack{\{x,y\}\in \partial\mathbb{B}_\Gamma \\ y\notin\Gamma}} \sigma(x)\eta(y) - \frac{h}{2} \sum_{x\in\Gamma} \sigma(x) , \tag{7}$$

where $h \in \mathbb{R}$ is the external field and $\sigma \in \Omega$ is a generic configuration. The Gibbs (probability) measure in Γ with boundary condition η under external field h and at temperature $T = 1/\beta$ is now defined on Ω as

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

Observe that we omit in the notation reference to the temperature T, because it will be usually fixed. The following property is a consequence of the fact that the Hamiltonian only involves interactions between nearest neighbors: given $\Gamma \in \mathscr{F}$, if $\eta(x) = \zeta(x)$ for every $x \in \partial_{\text{ext}}\Gamma$, then

$$\int f d\mu_{\Gamma,\eta,h} = \int f d\mu_{\Gamma,\zeta,h} , \tag{8}$$

for every local observable f whose support is contained in Γ. The next property is known as the DLR equations: given $\Gamma \subset \Gamma' \in \mathscr{F}$ and a pair of configurations η and η' which are identical off Γ', we have

$$\mu_{\Gamma',\eta',h}(\cdot \mid \Omega_{\Gamma,\eta}) = \mu_{\Gamma,\eta,h}(\cdot) . \tag{9}$$

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

If $\eta \leqq \zeta$ and $h_1 \leqq h_2$, then, for each $\Gamma \in \mathscr{F}$, $\mu_{\Gamma,\eta,h_1} \leqq \mu_{\Gamma,\zeta,h_2}$.

A Gibbs measure for the infinite system on $\mathbb{Z}^d$ is defined now as any probability measure, μ, which satisfies the DLR equations in the sense that for every $\Gamma \in \mathscr{F}$ and μ-almost all $\eta \in \Omega$,

$$\mu(\cdot \mid \Omega_{\Gamma,\eta}) = \mu_{\Gamma,\eta,h}(\cdot) . \tag{10}$$

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 $(\mu_{\Gamma_i,\eta_i,h})_{i=1,2,}$, where each Γ_i is finite and $\Gamma_i \to \mathbb{Z}^d$, as $i \to \infty$, in the sense that $\bigcup_{i=1}^{\infty} \bigcap_{j=i}^{\infty} \Gamma_j = \mathbb{Z}^d$. Together, (8) and the DLR equations, (9) and (10), imply the Markov property for the Gibbs measures; for instance, if μ is a Gibbs measure for the infinite system under external field h, then for arbitrary $\Gamma \in \mathscr{F}$ and μ-almost all $\eta, \zeta \in \Omega$ such that $\eta(x) = \zeta(x)$ for every $x \in \partial_{\text{ext}}\Gamma$,

$$\int f d\mu(\cdot \mid \Omega_{\Gamma,\eta}) = \int f d\mu(\cdot \mid \Omega_{\Gamma,\zeta}) ,$$

for every local observable f whose support is contained in Γ.

The Holley-FKG inequalities can be used to prove that for each value of T and h, $\mu_{\Lambda(l),-,h}$ (resp. $\mu_{\Lambda(l),+,h}$) converges weakly, as $l \to \infty$, to a probability measure that we will denote by $\mu_{-,h}$ (resp. $\mu_{+,h}$). If $h \neq 0$, or $d = 1$, it is also known that

$$\mu_{-,h} = \mu_{+,h} =: \mu_h , \tag{11}$$

while if $d \geqq 2$ and $h = 0$ the same is true if the temperature is larger than a critical value $T_c > 0$, which depends on the dimension, and is false for $T < T_c$. Moreover, for the values of T and h for which (11) holds, any weak limit of any sequence of the form $(\mu_{\Gamma_i,\eta_i,h})_{i=1,2,}$, where $\mathscr{F} \ni \Gamma^i \to \mathbb{Z}^d$, is identical to μ_h. Therefore we conclude that whenever (11) holds there is a unique Gibbs measure for the infinite system. When (11) fails, there is more than one Gibbs measure for the infinite system, and we say that there is phase coexistence. We use the following abbreviations and names:

$$\mu_{-,0} := \mu_- = \text{the minus phase} ,$$

$$\mu_{+,0} := \mu_+ = \text{the plus phase} .$$

Another known fact is that for fixed T

$$\mu_h \to \mu_+ \text{ weakly, as } h \searrow 0 ,$$

and

$$\mu_h \to \mu_- \text{ weakly, as } h \nearrow 0 .$$

For the expected value corresponding to a Gibbs measure $\mu\ldots$, in finite or infinite volume, we will use the notation

$$\langle f \rangle \quad := \int f d\mu\ldots ,$$

where ... stands for arbitrary subscripts. The spontaneous magnetization at temperature T is defined as

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

(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 $\mu_- \neq \mu_+$, and also that $\lim_{T \searrow 0} m^*(T) = 1$.

The dynamics. We introduce now for the Ising model above, the time evolution known as stochastic 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

$$(Lf)(\sigma) = \sum_{x \in \mathbb{Z}^d} c(x, \sigma)(f(\sigma^x) - f(\sigma)) , \tag{12}$$

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

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

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

The connection between the rates of flip and the Hamiltonian (1) and the temperature $T = 1/\beta$ is established by imposing conditions which assure 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 $\sigma \in \Omega$,

$$c(x, \sigma) = c(x, \sigma^x) \exp(-\beta \Delta_x H_h(\sigma)) , \tag{13}$$

where

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

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 (13) 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)

$$c_h(x, \sigma) = \exp(-(\beta/2)\Delta_x H_h(\sigma)).$$

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 $\eta(x) \leqq \zeta(x)$ and $h_1 \leqq h_2$, then

$$c_{h_1}(x, \eta) \leqq c_{h_2}(x, \zeta) \quad \text{if } \eta(x) = \zeta(x) = -1,$$

$$c_{h_1}(x, \eta) \geqq c_{h_2}(x, \zeta) \quad \text{if } \eta(x) = \zeta(x) = +1.$$

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

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

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, (13) and all the hypotheses H(1)–H(4). This spin flip system will be denoted by $(\sigma^{\eta}_{h;t})_{t \geqq 0}$, 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 $(\sigma^{\nu}_{h;t})_{t \geqq 0}$. 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, (13), 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

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

and

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

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 $\Gamma \in \mathscr{F}$, with some boundary condition $\xi \in \Omega$. This process, which will be denoted by $(\sigma^{\eta}_{\Gamma,\xi,h;t})_{t \geqq 0}$, where $\eta \in \Omega_{\Gamma,\xi}$ is the initial configuration, is defined as the spin flip system with rates of flip given by

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

When $\sigma, \sigma^x \in \Omega_{\Gamma,h}$, (13) yields, for all $x \in \mathbb{Z}^d$,

$$\mu_{\Gamma,\xi,h}(\sigma) c_{\Gamma,\xi,h}(x,\sigma) = \mu_{\Gamma,\xi,h}(\sigma^x) c_{\Gamma,\xi,h}(x,\sigma^x) , \tag{14}$$

which is the usual reversibility condition for finite state-space Markov processes. (Conversely, if one requires (14) to be satisfied for arbitrary $\Gamma \in \mathscr{F}$ and $\xi \in \Omega$, then one can deduce that (13) must hold.) It is clear from H(4) that $(\sigma^{\eta}_{\Gamma,\xi,h;t})$ is irreducible and hence from (14) it follows that, for any η,

$$\sigma^{\eta}_{\Gamma,\xi,h;t} \to \mu_{\Gamma,\xi,h} ,$$

weakly, as $t \to \infty$.

Graphical construction. In order to prove a few lemmas in the form needed in this paper, we introduce next a 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 $\mathbb{Z}^d$ or on any of its finite subsets, with arbitrary boundary conditions and starting from any initial configuration, all on the same probability space. The type of construction and the proofs that we present next are not new, but we could not find in the literature the precise type of results that we needed, including the necessary uniformity of the estimates in $h \in (-h(T), h(T))$, and for this reason we provide a self-contained exposition. The construction below 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 $x \in \mathbb{Z}^d$ 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 $(\tau^+_{x,n})_{n=1,2,\ldots}$ and $(\tau^-_{x,n})_{n=1,2,\ldots}$. 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 Γ and arbitrary ξ, the process $(\sigma^{\eta}_{\Gamma,\xi,h;t})$ is constructed as follows. We know that almost surely the random times $\tau^*_{x,n}$, $x \in \Gamma$, $n = 1, 2, \ldots$, $* = +, -$, are all distinct, and we update the state of the process at each time when there is a mark at some $x \in \Gamma$ 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 $c_{\Gamma,\xi,h}(x,\sigma) > U^*_{x,n} c_{\max}$.

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 $\mathbb{Z}^d$ in a similar fashion, with $c_h(x,\sigma)$ replacing $c_{\Gamma,\xi,h}(x,\sigma)$ 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. We will argue that on a set of probability 1 in the space where the marks were defined, for any fixed x and t, if we take any boundary condition ξ, then the sequence $(\sigma^\eta_{\Lambda(l),\xi,h;t}(x))_{l=1,2,\ldots}$ 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 version of the process $(\sigma^\eta_{h;t})$ constructed in this fashion has the correct flip rates. To prove the claim above about insensitivity to receding boundary conditions, we introduce the events $E(x,t,l)$ that there exists a sequence of points in space-time $(x_0, 0), (x_1, t_1), \ldots, (x_n, t_n)$ with the properties that $0 < t_1 < \cdots < t_n < t$, $x_0 \notin \Lambda(l)$, $x_n = x$, $\|x_i - x_{i-1}\| \leqq R$ for $i = 1, \ldots, n$, and that at each point (x_i, t_i), $i = 1, \ldots, n$, there is a mark. It is easy to see that out of the event $E(x,t,l)$, $\sigma^\eta_{\Lambda(l),\xi,h;t}(x)$ does not depend on ξ. Because $E(x,t,l) \subset E(x,u,l)$, when $t \leqq u$, our claim is reduced to the statement that for each x and integer t, $E(x,t,l)$ happens for only finitely many values of l, $\mathbb{P}$-almost surely. We have to show now that the probability of $E(x,t,l)$ vanishes fast enough as $l \to \infty$, so that we can apply the Borel–Cantelli Lemma. In order to do it we observe that for a given n, the sites $x_0, \ldots, x_{n-1}$ in the definition of $E(x,t,l)$ cannot be chosen in more than $((2R+1)^d)^n$ ways. Also, for large l, n cannot be less than $l/(3R)$. Therefore,

$$\mathbb{P}(E(x,t,l)) \leqq \sum_{n \geqq l/(3R)} (2R+1)^{dn} \mathbb{P}(Z \geqq n),$$

where Z is a Poisson random variable, with mean $tc_{\max} =: r$. We will use now the following elementary inequality, valid for $n \geqq r$,

$$\mathbb{P}(Z \geqq n) = e^{-r} r^n \sum_{k \geqq n} \frac{r^{k-n}}{k!} \leqq e^{-r} r^n \sum_{k \geqq 0} \frac{n^{k-n}}{k!}$$
$$= (r/n)^n e^{n-r} \leqq \exp(-n(\log(n/r) - 1)).$$

Combining the last two estimates we obtain for large l (depending only on x)

$$\mathbb{P}(E(x,t,l)) \leqq \sum_{n \geqq l/(3R)} (2R+1)^{dn} \exp(-n(\log(l/(3Rtc_{\max})) - 1)),$$

which goes to 0 faster than any exponential of l.

Actually, the estimate above shows 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 of the informal statement that because of the finite range of the interaction, and the uniform upper bound on the rates of flip, "the effects travel with a bounded speed."

Lemma 1. *For each dimension d and temperature T, there exists a finite positive constant $C(d, T)$ such that if we let $l \to \infty$ and $t \to \infty$ together, keeping $l \geqq C(d, T)t$, then for every site $x \in \mathbb{Z}$,*

$$\sup_{h \in (-h(T), h(T))} \sup_{\xi \in \Omega} \sup_{\eta \in \Omega} \mathbb{P}(\sigma^{\eta}_{h;t}(x) \neq \sigma^{\eta}_{\Lambda(l), \xi, h; t}(x)) \to 0 ,$$

exponentially fast in l.

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 (particular cases of) the following facts. If $\eta \leqq \zeta$, $\xi \leqq \xi'$, $-h(T) < h_1 \leqq h_2 < h(T)$ and $\Gamma \in \mathscr{F}$ is arbitrary, then for all $t \geqq 0$,

$$\sigma^{\eta}_{\Gamma, \xi, h_1; t} \leqq \sigma^{\zeta}_{\Gamma, \xi', h_2; t} , \tag{15}$$

$$\sigma^{\eta}_{h_1; t} \leqq \sigma^{\zeta}_{h_2; t} , \tag{16}$$

and

$$\sigma^{\eta}_{\Gamma, -, h_1; t} \leqq \sigma^{\zeta}_{h_2; t} . \tag{17}$$

We will refer to these inequalities as basic-coupling inequalities. (Observe that the Holley-FKG inequalities for the models we are considering can be derived from (15).)

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

We will use the notation $\beta' = \beta - \log b$, where b is the constant (dependent on d) which appeared in the counting inequality (4). Several times we will encounter the fraction β'/β, and we observe that it satisfies

$$\beta'/\beta = 1 - T \log b \nearrow 1 \quad \text{as } T \searrow 0 . \tag{18}$$

Given $\Gamma \in \mathscr{F}$, $\xi \in \Omega$, $h \in \mathbb{R}$ and $E \subset \Omega$, we write

$$Z_{\Gamma, \xi, h}(E) := \sum_{\sigma \in \Omega_{\Gamma, \xi} \cap E} \exp(-\beta H_{\Gamma, \xi, h}(\sigma)) . \tag{19}$$

1-ii. Main Result. The following theorem is our main result.

Theorem 1. *For each dimension $d \geqq 2$ there is $T_0 > 0$ such that for every temperature $T \in (0, T_0)$ the following happens. There are constants $0 < \lambda_1(T) \leqq \lambda_2(T) < \infty$ such that if we let $h \searrow 0$ and $t \to \infty$ together, then for every local observable f,*

i) $\mathbb{E}(f(\sigma^{-}_{h;t})) \to \int f d\mu_-$ *if* $\limsup h^{d-1} \log t < \lambda_1(T)$.
ii) $\mathbb{E}(f(\sigma^{-}_{h;t})) \to \int f d\mu_+$ *if* $\liminf h^{d-1} \log t > \lambda_2(T)$.

We can take $\lambda_1(T) = (2^d(d-1)^{d-1}/(d+1))(\beta'/\beta)^d \beta$, and $\lambda_2(T) = (2^d d^{d-1})(1 + \delta(T))\beta$, where $\delta(T)$ is a positive-valued function which vanishes as $T \searrow 0$.

In other words, we are stating that the law of the random configuration $\sigma^{-}_{h;t}$ converges weakly to μ_- in case (i) and to μ_+ in case (ii).

Theorem 1, apart from the explicit estimates on $\lambda_1(T)$ and $\lambda_2(T)$, was conjectured by Aizenman and Lebowitz in [AL], where they proved a similar result for certain deterministic cellular automata evolving from initial random configurations selected according to translation invariant product measures. Actually they conjectured the stronger result, which states that also $\lambda_1(T) = \lambda_2(T) =: \lambda_c(T)$. This is a natural further conjecture, but we believe that it will be extremely difficult to prove it, because it is not even clear what the common value of $\lambda_1(T)$ and $\lambda_2(T)$ should be, as we will explain in the next subsection.

The statements (i) and (ii) in Theorem 1 should be compared with the following slightly weaker, but more explicit, pair of statements.

i′) $\mathbb{E}(f(\sigma^-_{h;t})) \to \int f d\mu_-$ if $t = \exp(\lambda/h^{d-1})$, $0 < \lambda < \lambda_1(T)$ and $h \searrow 0$.
ii′) $\mathbb{E}(f(\sigma^-_{h;t})) \to \int f d\mu_+$ if $t = \exp(\lambda/h^{d-1})$, $\lambda_2(T) < \lambda < \infty$ and $h \searrow 0$.

Statements (i′) and (ii′) are clearly implied by (i) and (ii), respectively. On the other hand, using monotonicity in t of $\mathbb{E}(f(\sigma^-_{h;t}))$, when f is monotone (and writing a generic observable f as a sum of two monotone parts), one can easily see that if we assume that $0 < \liminf h^{d-1} \log t \leqq \limsup h^{d-1} \log t < \infty$, then, actually, (i′) and (ii′) imply (i) and (ii), respectively. In other words, the cases covered, for instance, by (i) and not by (i′) are those in which $t \to \infty$ slower than any exponential of $1/h^{d-1}$. In this regard, (i) actually says that the system relaxes to the "metastable situation" in a time which has to be large compared to the unit of time, but which does not scale with h. Similarly, (ii), contrary to (ii′), covers cases in which $t \to \infty$ faster than any exponential of $1/h^{d-1}$. In reality the content of (i) and (ii) which is missing in (i′) and (ii′) is contained in the part of Theorem 1 which is easy to prove. We state this part and prove it next.

Proposition 1. *For each dimension d and for every temperature T, if we let h* $\searrow 0$ *and* $t \to \infty$ *together, then for every non-decreasing local observable f*

i″) $\liminf \mathbb{E}(f(\sigma^-_{h;t})) \geqq \int f d\mu_-$.
ii″) $\limsup \mathbb{E}(f(\sigma^-_{h;t})) \leqq \int f d\mu_+$.

Proof. Using the basic-coupling inequality (16), write

$$\mathbb{E}(f(\sigma^-_{h;t})) \geqq \mathbb{E}(f(\sigma^-_{0;t})) .$$

Now, as $t \to \infty$, the right-hand side converges to $\int f d\mu_-$, and hence (i″) holds.

To prove (ii″) observe that also from the same basic-coupling inequality, if $h' > 0$ is kept fixed, then, when $h \leqq h'$, we have

$$\mathbb{E}(f(\sigma^-_{h;t})) \leqq \mathbb{E}(f(\sigma^-_{h';t})) .$$

Letting $h \searrow 0$ and $t \to \infty$, in any fashion, gives

$$\limsup \mathbb{E}(f(\sigma^-_{h;t})) \leqq \lim_{t\to\infty} \mathbb{E}(f(\sigma^-_{h';t})) = \int f d\mu_{h'} .$$

The proof is completed by observing that $h' > 0$ is arbitrary and recalling that μ_+ is the weak limit of $\mu_{h'}$ as $h' \searrow 0$. q.e.d.

A simple-but-nice immediate consequence of Proposition 1 is the following result, which, in contrast to Theorem 1, says that for temperatures for which there is no phase transition, the relaxation to equilibrium occurs in a time of order 1 (no scaling with h).

Proposition 2. *For each dimension d and for every temperature T for which* $\mu_- = \mu_+ =: \mu_0$, *if we let* $h \searrow 0$ *and* $t \to \infty$ *together, then for every local observable f,*

$$\lim \mathbb{E}(f(\sigma^-_{h;t})) = \int f d\mu_0 .$$

We make now some remarks on the explicit form of $\lambda_1(T)$ and the semiexplicit form of $\lambda_2(T)$ provided in the statement of Theorem 1. First observe that, thanks to (18), as $T \searrow 0$, $\lambda_1(T)$ is asymptotic to $(2^d(d-1)^{d-1}/(d+1))\beta$. We will argue in the next subsection that this asymptotics should actually be optimal, in that for the conjectured constant $\lambda_c(T)$ the same asymptotics should hold. On the other hand, we have as $T \searrow 0$, $\lambda_2(T)$ asymptotic to $(2^d d^{d-1})\beta$, which we believe to be off the correct behavior by a factor $[(d/(d-1))^{d-1}][d+1]$. We will see later that the origin of each one of the factors singled out inside each one of the pairs of square brackets is different.

1-iii. Heuristics. We present now the heuristics behind Theorem 1. This heuristic reasoning comes in two parts, the first one of which is very well known, while the second one seems to have escaped most of the attention.

First part. We want to consider the behavior of an individual droplet of spins $+1$ in a background of spins -1. When the temperature is low, it is reasonable, on energetic grounds, to consider simply a cube full of spins $+1$, the other spins being all -1, as such a droplet. If the side-length of the cube is l, then the energy of such a configuration, with respect to the energy of the configuration with all spins -1 is given by

$$e_h(l) := 2dl^{d-1} - l^d h .$$

As a function of l, considered now as a continuous quantity, $e_h(l)$ grows from 0 to its maximum

$$E_{\max} = 2^d(d-1)^{d-1}/h^{d-1} ,$$

when l varies from 0 to $l_c = 2(d-1)/h$. For $l > l_c$, $e_h(l)$ decreases; it crosses the value 0 when $l = 2d/h$, and goes to $-\infty$ when $l \to \infty$.

If we assume that the droplet evolves in such a way as to lower the energy of the system, then we are led to the conclusion that droplets with side-length smaller than l_c tend to shrink and that droplets with side-length larger than l_c tend to grow and cover the whole system. Also by analogy with other phenomena related to passage over potential barriers, one would expect that the time needed for a droplet to pop up spontaneously, due to a thermal fluctuation, *in a given place* is of the order of

$$\exp(\beta E_{\max}) ,$$

which grows exponentially with $1/h^{d-1}$.

Second part. From the discussion above one could naively predict for the system a relaxation time of the order of $\exp(\beta E_{\max})$. Actually, this is only reasonable if the whole system is not much larger than the size of a critical droplet, so that the time for such a droplet to first appear should indeed be of that order and, moreover, when such a droplet appears, it will cover the whole system in a comparably negligible time. For instance, this seems to be a good prediction if the linear size of the system scales as B/h with a large fixed B. (In this regard, see Corollary 1.) But we are concerned with a larger (infinite) system, and we are observing it through a local function f, which depends, say, on the spins in a finite set S. For us the system will have relaxed to equilibrium when S 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 fixed speed v, then we can see that the region in space-time where a droplet which covers S at time t could have appeared is, roughly speaking, a cone with vertex in S and which has as base the set of points which have time-coordinate 0 and are at most at distance tv from S. The volume of such a cone is of the order of $(vt)^d t$. Now, from the discussion in the first part of the heuristics, one can infer that "the rate with which supercritical droplets appear by thermal fluctuations" at a given location should be of the order of $\exp(-\beta E_{\max})$. The order of magnitude of the relaxation time, t_{rel}, before which the region S is unlikely to have been covered by a large droplet and after which the region S is likely to have been covered by such an object can now be obtained by solving the equation

$$(vt_{\text{rel}})^d t_{\text{rel}} \exp(-\beta E_{\max}) = 1 .$$

This gives us

$$t_{\text{rel}} = v^{-d/(d+1)} \exp(\beta E_{\max}/(d+1)) .$$

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 that at least it goes to 0, as $h \searrow 0$, so slowly that

$$\lim_{h \searrow 0} h^{d-1} \log v = 0 , \tag{20}$$

then we can predict that

$$t_{\text{rel}} \sim \exp(\beta E_{\max}/(d+1)) = \exp\left(\frac{\beta 2^d (d-1)^{d-1}}{(d+1) h^{d-1}}\right) . \tag{21}$$

We will explain now why it seems reasonable to suppose that (20) is true. v should be the asymptotic speed of the droplet, when it becomes very large (much larger than the critical size), and in this regime we can neglect the curvature of the surface of the droplet and regard the growth of its radius as resulting from the movement of its boundary as that of a (mesoscopically) flat interface, caused by the fact that h is positive. Thinking of the surface as a flat interface and keeping in mind that h is small, we can, in first approximation, assume that on one side of the interface we have the minus-phase and on the other side the plus-phase, which are symmetric, and that protuberances of each phase into the other at the interface are essentially similar. The movement of the interface is then caused simply by the larger rate of flip of spins in the upward direction, caused by the fact that $h > 0$, when we compare two situations which are related by spin reversal at all sites. By checking in examples of rates, or from the type of argument that we will use later to derive (22), one can see that this difference in the rates of flip caused by the external field h is of the order of h. From this one obtains $v \sim h$ as $h \searrow 0$, which implies (20).

From (21) one sees that the relaxation time, even for the infinite system, should grow exponentially with h^{1-d}, and what the rate of this exponential growth should approximately be when T is close to 0. The fact that in part (i) of Theorem 1 we have $\lambda_1(T)$ which is asymptotic as $T \searrow 0$ to the value of $\lambda_c(T)$ predicted in (21) is a pleasant feature of the method used to prove this side of the theorem. On the other hand, in part (ii) of Theorem 1 we are missing the factor $1/(d+1)$, in $\lambda_2(T)$ because we are not able to control rigorously the growth of the supercritical droplets and make complete sense out of (20). The other factor by which $\lambda_2(T)$ differs from $\lambda_c(T)$ even as $T \to 0$ is there for other technical reasons.

A major question, which seems to be controversial even from a heuristic standpoint, is the prediction of the correct value of $\lambda_c(T)$, for each T (small enough, if necessary), and not just its asymptotic behavior as $T \to 0$. A certain type of "common wisdom" says that one should repeat the computation above but with the cubes replaced by solids which have the Wulff shape corresponding to the surface tension at temperature T. This idea has, nevertheless, been challenged by the results obtained in the limit of very low temperature by Kotecký and Olivieri in [KO2 and KO3] (results announced in [KO1]). (After discussions with these two colleagues, it seems to me that there is no compelling evidence that in the limit considered in the present paper Wulff shapes should be more likely to come into play in this problem than in the limit of very low temperatures.) In connection to this discussion, one may want to refer to the fact that investigations have been carried out on simulations and analytic (non-rigorous) studies of supercritical droplet growth (see for instance [DS] and references therein). Nevertheless such investigations refer to the growth of droplets which are very supercritical and should develop an asymptotic shape related to the different asymptotic speed of growth in different directions. The asymptotic shape is not given by the equilibrium Wulff construction, but by a similar construction based on the speed of growth as a function of the direction. In any case this asymptotic shape obtained when a droplet is moving downhill, "with the drift," does not clarify the controversy about the first droplets which appear and are likely to grow (a completely different, large-deviations type problem, related to moving uphill, "against the drift").

1-iv. Byproducts. In our way to prove Theorem 1, we will prove two results which are interesting for their own sake. The first one can be considered as a static (or equilibrium) version of the metastable behavior shown by the Ising model, in this case when the boundary condition competes with the external field.

Theorem 2. *For each dimension $d \geqq 2$ there is $T_0 > 0$ such that for every temperature $T \in (0, T_0)$ the following happens. There are constants $0 < B_1(T) \leqq B_2(T) < \infty$ such that if we let $h \searrow 0$, then for every local observable f*

i) $\langle f \rangle_{\Lambda(B/h), -, h} \to \int f d\mu_-$ *if* $B < B_1(T)$,
ii) $\langle f \rangle_{\Lambda(B/h), -, h} \to \int f d\mu_+$ *if* $B > B_2(T)$.

We can take $B_1(T) = 2d(\beta'/\beta)$, and $B_2(T) = 2d(1 + \delta(T))$, where $\delta(T)$ is a positive-valued function which vanishes as $T \searrow 0$.

While we do not know, at a rigorous level, that $B_1(T) = B_2(T)$, the theorem above states that at least in the limit $T \searrow 0$ these quantities converge both to $2d$. The value $2d$ for the common limit can be easily understood, since when $B < 2d$ the ground state inside $\Lambda(B/h)$ with $-$ boundary conditions is the configuration identically -1, while when $B > 2d$, it is the configuration identically $+1$.

Part (i) of Theorem 2 will follow from the arguments that will be developed in Sect. 2 in order to prove part (i) of Theorem 1; its proof will be presented at the end of that section. In contrast, part (ii) of Theorem 2 will actually be used to prove part (ii) of Theorem 1. Our proof of part (ii) of Theorem 2 will be based on Theorem 3 below, which will be proved in Subsect. 3-i. In order to state this theorem, we need to introduce the following definition. We will denote by $\mathscr{B}$ the set of configurations in Ω in which the box $\Lambda(d/h)$ intersects an infinite cluster of spins -1.

Theorem 3. *For each* $B > 2d$, *there exists* $T(B) > 0$ *so that for all* $T \in (0, T(B))$

$$\lim_{h \searrow 0} \mu_{\Lambda(B/h), -, h}(\mathscr{B}) = 0 .$$

Part (ii) of Theorem 2 follows from Theorem 3 in a standard fashion. Observe that for each configuration in $\Omega_{\Lambda(B/h), -} \backslash \mathscr{B}$ there is an outer contour present which surrounds $\Lambda(d/h)$. Conditioning on what this contour is and using the Markov property of the Gibbs measures and the FKG-Holley inequality, one obtains for each non-decreasing observable f,

$$\liminf_{h \searrow 0} \langle f \rangle_{\Lambda(B/h), -, h} \geqq \int f d\mu_+ ,$$

because the spins in the interior boundary of the outer contour are all $+1$. The complementary inequality,

$$\limsup_{h \searrow 0} \langle f \rangle_{\Lambda(B/h), -, h} \leqq \int f d\mu_+ ,$$

follows from the same simple argument used to prove part (ii″) of Proposition 1, based on the FKG-Holley inequalities.

Theorem 3 is technically the most difficult part of the present paper. It is a strengthening of the main result obtained by Martirosyan in [Mar]. That result was stated there as follows: for each dimension d, for low enough temperature T, there is a finite constant $B(T)$ such that the event that in the annulus between the boxes $\Lambda(B/h)$ and $\Lambda(B/(6h))$ there is an outer contour has $\mu_{\Lambda(B/h), -, h}$-probability which converges to 1 as $h \searrow 0$. The best possible choice of $B(T)$ is discussed in Sect. 3.12 of [Mar] (but beware that in the fourth displayed relation in that section, $c_2(v, \beta)$ should actually be the inverse of this quantity, and also that what is called h in [Mar] is βh in the present paper) and it is clear that the techniques in that paper do not allow one to take $B(T)$ arbitrarily close to $2d$ (provided we are willing to take a low enough temperature) as in Theorem 3 above. The fact that the box which is surrounded by a contour has side-length, respectively, d/h and $B/(6h)$ in the theorems that we are comparing is actually not serious; with some minor extra work, we can strengthen the statement of Theorem 3, by replacing in the definition of $\mathscr{B}$ the box $\Lambda(d/h)$ by the box $\Lambda(B(1-\varepsilon))/h$, where $\varepsilon > 0$ is arbitrary. It seems that also in [Mar] the factor 1/6 which multiplies B/h could be replaced by any fixed factor less than 1. In spite of the fact that we provide a stronger statement, our proof is actually simpler than the one in [Mar], interestingly enough, because we exploit a result (Lemma 6) which was primarily considered in connection to the proof of part (i) of Theorem 1.

1-v. Finite Systems. From the perspective of physics, the motivation behind statements such as those in Theorem 1, which refer to infinite systems, is actually the idea that such systems are idealizations of very large, but finite, systems. By very large, here, it should be understood that the system is much larger than any relevant space scale in the problem. From the heuristics above, it should be clear that, in the present case then, a very large system should be one with linear size of the order of $\exp(D/h^{d-1})$, where D is a large enough constant (how large depends on T also). In such a case the relevant space-time cones introduced in the heuristics will be fully contained inside the system.

Suppose now that we have a smaller system, the process with $-$ boundary conditions evolving in the box $\Lambda(l)$ of side l, say, where l scales in some fashion with h. If l stays constant as $h \searrow 0$, or even if it grows too slowly as $h \searrow 0$, then even in equilibrium the effect of the boundary conditions do not vanish as $h \searrow 0$; this is the type of problem discussed in the previous subsection, where we pointed out that for the equilibrium measure to be close to the plus-phase, l has to grow at least as fast as B/h, where B is a large enough constant. From the heuristics above this is also clear, since otherwise critical droplets may be larger than the whole system. On the other hand, if we suppose that l grows fast enough to avoid this problem, then we obtain the theorem below, which generalizes Theorem 1.

Theorem 4. *For each dimension $d \geqq 2$ there is $T_0 > 0$ such that for every temperature $T \in (0, T_0)$ and every constant $D \geqq 0$ the following happens. There are constants $0 < \lambda_1(T, D) \leqq \lambda_2(T, D) < \infty$ and $B(T) < \infty$ such that if we let $h \searrow 0$, $t \to \infty$ and $l \to \infty$ together in such a fashion that $\liminf hl > B(T)$ and $\lim h^{d-1} \log l = D$, then for every local observable f*

i) $\mathbb{E}(f(\sigma^{-}_{\Lambda(l), -, h; t})) \to \int f d\mu_-$ *if* $\limsup h^{d-1} \log t < \lambda_1(T, D)$,
ii) $\mathbb{E}(f(\sigma^{-}_{\Lambda(l), -, h; t})) \to \int f d\mu_+$ *if* $\liminf h^{d-1} \log t > \lambda_2(T, D)$.

We can take

$$\lambda_1(T, D) = \max\{2^d(d-1)^{d-1}(\beta'/\beta)^d\beta - dD, (2^d(d-1)^{d-1}/(d+1))(\beta'/\beta)^d\beta\},$$

$\lambda_2(T, D) = (2^d d^{d-1})(1 + \delta_2(T))\beta$, *and* $B(T) = 2d(1 + \delta_3(T))$, *where for* $i = 2, 3$, $\delta_i(T)$ *are positive-valued functions which vanish as* $T \searrow 0$.

Part (ii) of this theorem can be proven in exactly the same way in which part (ii) of Theorem 1 is proven. Also the adaptation of the proof of part (i) of Theorem 1 to prove part (i) of Theorem 4 is not difficult, and we will restrict ourselves to sketching it at the end of Sect. 2.

Because the statements in Theorem 4 are somewhat involved, we single out next, as a corollary, the particular case in which $D = 0$ (for instance, l may grow as $(1/h)^a$, where $a > 1$, or as $(2.001)d/h$ if the temperature is low enough). This case is conceptually simpler than the case of infinite systems, covered by Theorem 1, in that the notion of "growth of droplets" is irrelevant here. In particular there is no factor $1/(d+1)$ in the exponential rate of growth of the relaxation time with $1/h^{d-1}$. The asymptotic behavior of $\lambda_2(T, 0)$ as $T \searrow 0$, unfortunately, is still not what one predicts from the heuristics, nevertheless it is interesting to see that, when the dimension becomes large, the ratio between $\lambda_1(T, 0)$ and $\lambda_2(T, 0)$ stays bounded.

Corollary 1. *For each dimension $d \geqq 2$ there is $T_0 > 0$ such that for every temperature $T \in (0, T_0)$ the following happens. There are constants $0 < \lambda_1(T, 0) \leqq \lambda_2(T, 0) < \infty$ and $B(T) < \infty$ such that if we let $h \searrow 0$, $t \to \infty$ and $l \to \infty$ together in* *such a fashion that* $\liminf hl > B(T)$ *and* $\lim h^{d-1} \log l = 0$, *then for every local observable* f,

i) $\mathbb{E}(f(\sigma^{-}_{\Lambda(l),-,h;t})) \to \int f d\mu_-$ *if* $\limsup h^{d-1} \log t < \lambda_1(T,0)$,
ii) $\mathbb{E}(f(\sigma^{-}_{\Lambda(l),-,h;t})) \to \int f d\mu_+$ *if* $\liminf h^{d-1} \log t > \lambda_2(T,0)$.

We can take $\lambda_1(T,0) = 2^d(d-1)^{d-1}(\beta'/\beta)^d \beta$, $\lambda_2(T,0) = (2^d d^{d-1})(1+\delta_2(T))\beta$, *and* $B(T) = 2d(1+\delta_3(T))$, *where for* $i = 2, 3$, $\delta_i(T)$ *are positive-valued functions which vanish as* $T \searrow 0$.

1-vi. A Simple Argument for the Presence of a Plateau. The presence of a plateau in the relaxation curves of local observables, when we are close to the phase transition region, can actually be understood in a very simple way. Because of the bounded speed of propagation of effects, up to an arbitrary time t, which may be very large, the spins in the support S of a local observable f depend only weakly on what happens up to time t outside of the box $\Lambda(l)$, if l is taken large enough, depending on t. But if we are now free to take h as small as we want, we may make it very unlikely that the system could "feel the presence of h inside the space-time cylinder $\Lambda(l) \times [0, t]$." The conclusion is that the system behaves locally, up to time t almost as if it were submitted to no external field. In other words, up to time t the relaxation curve for f must be close to the relaxation curve that we would have under no external field. If we are starting from $-\underline{1}$ at time 0, then the relaxation curve for f under no external field approaches asymptotically, as $t \to \infty$, a straight line, corresponding to the equilibrium value of f in the minus-phase. Because t can be made large, by making h small, we can see as much of a plateau as we want in the relaxation curve of f under external field h.

The basic coupling, in the form that we constructed it, provides a method for giving a precise meaning to the argument described above, and for estimating the lower bound on the "length of the plateau" that can be obtained in this fashion. Before we state and prove the exact result, it is worth observing that implicit in the discussion above is the assumption that as $h \searrow 0$, the rates $c_h(x, \sigma)$ converge to $c_0(x, \sigma)$. This can be checked easily for each one of the explicit examples that we introduced before. We would have added this sort of continuity to the list of hypotheses that we are assuming, were it not for the fact that it can actually be deduced from some of the other hypotheses already introduced. Moreover, as we will see below, under the hypotheses that we are already assuming, the speed of this convergence of the rates is also constrained in a fashion that does not depend on the specific choice of the rates.

Proposition 3. *For each dimension* $d \geqq 2$ *and every temperature* $T \in (0, T_c)$ *the following happens. If we let* $h \searrow 0$ *and* $t \to \infty$ *together in such a way that* $\limsup ht^{d+1} = 0$, *then for every local observable* f

$$\mathbb{E}(f(\sigma^{-}_{h;t})) \to \int f d\mu_- \,.$$

Proof. First we analyze how the rates $c_h(x, \sigma)$ behave when h is close to 0. Due to translation invariance we set $x = 0$ and define

$$g(h) := \sup_{\sigma} |c_h(0, \sigma) - c_0(0, \sigma)| \,.$$

Given a configuration σ such that $\sigma(0) = +1$, we use the notation $a(h) := c_h(0, \sigma)$, and $b(h) := c_h(0, \sigma^0)$. The detailed balance condition (13) states that

$$\frac{b(h)}{a(h)} = \exp\left(\beta \sum_{y:\ \|x-y\|_1 = 1} \sigma(y) + \beta h\right) = \frac{b(0)}{a(0)} \exp(\beta h) .$$

Hence

$$\log(b(h)/b(0)) + \log(a(0)/a(h)) = \beta h .$$

From the hypotheses of monotonicity in h, (H3), $a(h)$ decreases with h, while $b(h)$ increases with h, therefore the two logarithms above are positive, and hence must vanish as $h \searrow 0$, implying that $b(h) \to b(0)$ and $a(h) \to a(0)$. Moreover

$$\beta h/2 \leqq \max\{\log(b(h)/b(0)), \log(a(0)/a(h))\} \leqq \beta h .$$

Hence for small h (depending on σ and β),

$$\beta h/4 \leqq \max\{(b(h) - b(0))/b(0)), (a(0) - a(h))/a(h))\} \leqq 2\beta h .$$

Using now the hypotheses of boundedness of the rates, (H4), we can conclude that there are positive finite constants $C_1(\beta, \sigma)$ and $C_2(\beta, \sigma)$ such that

$$C_1(\beta, \sigma)h \leqq \max\{|b(h) - b(0)|, |a(0) - a(h)|\} \leqq C_2(\beta, \sigma)h .$$

But because of the finite range hypotheses, (H2), we may consider only a finite number of possible choices for σ, and still would be covering completely arbitrary choices. This allows us to conclude that there are positive finite constants $C_3(\beta)$ and $C_4(\beta)$ for which

$$C_3(\beta)h \leqq g(h) \leqq C_4(\beta)h , \tag{22}$$

when h is small enough.

To complete the proof of the proposition we will use Lemma 1. Using the notation there, take $l = C(d, T)t$. Choose an arbitrary boundary condition ξ, denote the support of f by S, and write

$$\begin{aligned} |\mathbb{E}(f(\sigma^-_{h;t})) - \textstyle\int f d\mu_-| \leqq{} & |\mathbb{E}(f(\sigma^-_{0;t})) - \textstyle\int f d\mu_-| \\ & + \|f(\zeta)\|_\infty \sum_{x \in S} [\mathbb{P}(\sigma^-_{h;t}(x) \neq \sigma^-_{\Lambda(l),\xi,h;t}(x)) \\ & + \mathbb{P}(\sigma^-_{0;t}(x) \neq \sigma^-_{\Lambda(l),\xi,0;t}(x)) \\ & + \mathbb{P}(\sigma^-_{\Lambda(l),\xi,h;t}(x) \neq \sigma^-_{\Lambda(l),\xi,0;t}(x))] . \end{aligned} \tag{23}$$

To control the last term above, we observe that, for each $x \in S$, the event whose probability appears there can only occur if there is a mark inside the cylinder $\Lambda(l) \times [0, t]$ for which the corresponding $U^*_{x,n}$ lies between $c_h(x, \sigma)/c_{\max}$ and $c_0(x, \sigma)/c_{\max}$, for some configuration σ. But for each mark this occurs with probability bounded above by $g(h)/c_{\max}$. Since the number of marks in the cylinder that we are considering is a Poisson random variable with mean $|\Lambda(l)|tc_{\max}$, it follows from (22) and the hypotheses $\limsup ht^{d+1} = 0$ that, for each $x \in S$,

$$\mathbb{P}(\sigma^-_{\Lambda(l),\xi,h;t}(x) \neq \sigma^-_{\Lambda(l),\xi,0;t}(x)) \to 0 . \tag{24}$$

The proposition follows now from (23), (24), Lemma 1 and the fact that as $t \to \infty$, $\sigma^-_{0;t} \to \mu_-$, weakly. q.e.d.

The argument above is very soft and works at all temperatures below T_c, but unfortunately it only gives a lower bound for the "length of the plateau" of the order $1/h^{1/(d+1)}$. This proposition allows us to focus on what the main content of Theorem 1 really is: an estimate of how long the plateau is, or, in other words, for how long a time the system stays far from equilibrium, displaying a metastable type of behavior close to the wrong phase.

2. Lower Bound on the Relaxation Time

Our main goal in this section is to prove part (i) of Theorem 1. Throughout this section we will set $\lambda_1(T) = ((2^d(d-1)^{d-1})/(d+1))(\beta'/\beta)^d\beta$. Recall that because we already proved Proposition 1, part (i), all that remains is to prove that for every $\lambda < \lambda_1(T)$, if we define $t_h := \exp(\lambda/h^{d-1})$, then for every non-increasing local observable f,

$$\limsup_{h \searrow 0} \mathbb{E}(f(\sigma^-_{h;\,t_h})) \leqq \langle f \rangle_- \,. \tag{25}$$

The first step towards proving (25) is the observation that, because "effects propagate with a maximum speed which does not scale with h," up to time t_h the configuration in the support of f is not much affected by anything that happened outside the box $\Lambda_h := \Lambda(\exp(\lambda_1(T)/h^{d-1}))$. In particular we can keep the spins outside this region frozen, assuming any value that we may choose. The precise statement, in the form we will use it, with $-$ boundary conditions, is contained in the lemma below, which follows from Lemma 1.

Lemma 2. *If $\lambda < \lambda_1(T)$ then for any local observable f,*

$$\lim_{h \searrow 0} |\mathbb{E}(f(\sigma^-_{h;\,t_h})) - \mathbb{E}(f(\sigma^-_{\Lambda_h,\,-,\,h;\,t_h}))| = 0 \,.$$

Motivated by the heuristic discussion related to the behavior of droplets, we introduce now a set of configurations which are free of large droplets in a certain sense. Because we want to consider systems at temperatures which are low, but fixed and different from 0, we should not stick to the naive notion of droplets as cubes full of spins $+1$, but rather consider the contours which separate spins -1 from $+1$. Motivated by the classic paper [CCO], on which we will comment later, we denote by $\mathscr{R}_c$ the set of configurations in which no single contour surrounds a number of sites larger than c^d, i.e.,

$$\mathscr{R}_c := \{\sigma \in \Omega\colon \text{if } \gamma \text{ is a contour which is present in } \sigma, \text{ then } |\Theta(\gamma)| \leqq c^d\} \,.$$

The choices of c will be made later, in convenient ways, different for different uses, but usually of the form $c = A/h$, for some constant A.

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 have escaped from $\mathscr{R}_c$, where now $c = 2(d-1)(\beta'/\beta)/h$. In order to do this we introduce for each $\Gamma \in \mathscr{F}$ a modified dynamics evolving in $\Omega_{\Gamma,-}$, 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 $\Omega_{\Gamma,-}$ which evolves as the original stochastic Ising model in Γ, with $-$ boundary conditions, but for which all jumps out of $\mathscr{R}_c$ are suppressed. In other words, the rates, $\tilde{c}^c_{\Gamma,-,h}(x, \sigma)$, of the new process are identical to $c_{\Gamma,-,h}(x,\sigma)$ in case $\sigma^x \in \mathscr{R}_c$ and are 0 otherwise. We will denote this modified process, restricted to the state space $\Omega_{\Gamma,-} \cap \mathscr{R}_c$, by $(\tilde{\sigma}^{c,\eta}_{\Gamma,-,h;t})_{t \geqq 0}$, where $\eta \in \Omega_{\Gamma,-} \cap \mathscr{R}_c$ 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 $\tilde{\mu}^c_{\Gamma,-,h}$ given by

$$\tilde{\mu}^c_{\Gamma,-,h}(\cdot) := \mu_{\Gamma,-,h}(\cdot \mid \mathscr{R}_c) .$$

Now we observe that for each $\Gamma \in \mathscr{F}$ we can couple the process $(\sigma^-_{\Gamma,-,h;t})$ to the stationary process $(\tilde{\sigma}^{c,\tilde{\mu}^c_{\Gamma,-,h}}_{\Gamma,-,h;t})$ in such a way that the former will not escape from $\mathscr{R}_c$ before the latter visits the boundary relative to Γ of this set, defined by

$$\partial_\Gamma \mathscr{R}_c := \{\eta \in \mathscr{R}_c \cap \Omega_{\Gamma,-} : \eta^x \notin \mathscr{R}_c \text{ for some } x \in \Gamma\} .$$

This coupling can be constructed as follows. First we couple $(\sigma^-_{\Gamma,-,h;t})$ to $(\sigma^{\tilde{\mu}^c_{\Gamma,-,h}}_{\Gamma,-,h;t})$ using the basic coupling, so that for all $t \geqq 0$ we have

$$\sigma^-_{\Gamma,-,h;t} \leqq \sigma^{\tilde{\mu}^c_{\Gamma,-,h}}_{\Gamma,-,h;t} .$$

Now we enlarge the space on which this construction was made to accommodate $(\tilde{\sigma}^{c,\tilde{\mu}^c_{\Gamma,-,h}}_{\Gamma,-,h;t})$ which evolves together with $(\sigma^{\tilde{\mu}^c_{\Gamma,-,h}}_{\Gamma,-,h;t})$ up to the moment when this one escapes from $\mathscr{R}_c$; at this moment $(\tilde{\sigma}^{c,\tilde{\mu}^c_{\Gamma,-,h}}_{\Gamma,-,h;t})$ does not move, and afterwards it evolves independently of the other two processes, $(\sigma^-_{\Gamma,-,h;t})$ and $(\sigma^{\tilde{\mu}^c_{\Gamma,-,h}}_{\Gamma,-,h;t})$. The important fact about this coupling is that if we introduce

$$\tau^c_{\Gamma,h} := \inf\{t : \tilde{\sigma}^{c,\tilde{\mu}^c_{\Gamma,-,h}}_{\Gamma,-,h;t} \in \partial_\Gamma \mathscr{R}_c\} ,$$

then, because only one spin flips at a time in the type of dynamics that we are considering,

$$\sigma^-_{\Gamma,-,h;t} \leqq \tilde{\sigma}^{c,\tilde{\mu}^c_{\Gamma,-,h}}_{\Gamma,-,h;t} \quad \text{for all } t < \tau^c_{\Gamma,h} . \tag{26}$$

Lemma 4 below implies that in a sense and for certain choices of c, $\partial_\Gamma \mathscr{R}_c$ plays the role of a bottleneck. In combination with (26) above, it can be used to show that $(\sigma^-_{\Gamma,-,h;t})$ stays a long time in $\mathscr{R}_c$. That lemma follows from Peierls type of estimates, somewhat modified by the presence of the external field, which is compensated by the fact that we are suppressing large contours; this is the content of the preparatory Lemma 3.

Recall that we use the notation $\Omega(\gamma_1, \ldots, \gamma_k)$ to denote the event that $\gamma_1, \ldots, \gamma_k$ are present as outer contours.

Lemma 3. *If $h \geqq 0$, then for arbitrary $\Gamma \in \mathscr{F}$ and $c > 0$,*

$$\tilde{\mu}^c_{\Gamma,-,h}(\Omega(\gamma_1, \ldots, \gamma_k)) \leqq \exp\left((-\beta + \beta c h/(2d)) \sum_{i=1}^{k} |\gamma_i|\right) .$$

Proof. If there is any incompatibility among the γ_i, $i = 1, \ldots, k$, or if any of the regions they surround has more than c^d sites the bound stated in the lemma is trivially satisfied, therefore we suppose that they are compatible and that for each $i = 1, \ldots, k$,

$$|\Theta(\gamma_i)| \leqq c^d . \tag{27}$$

Combining these inequalities with the isoperimetric inequality (2), by taking an appropriate geometric mean of the corresponding bounds, we obain the further bound

$$|\Theta(\gamma_i)| \leqq |\gamma_i| c/(2d) . \tag{28}$$

Let $\Omega'(\gamma_1, \ldots, \gamma_k)$ be the set of configurations which can be obtained from configurations in $\Omega(\gamma_1, \ldots, \gamma_k)$ by flipping all the spins surrounded by one of the contours γ_i, $i = 1, \ldots, k$. Exploiting the fact that such a transformation is a one-to-one mapping and using an obvious estimate for the variation in energy resulting from this transformation we obtain, using the notation (19),

$$\begin{aligned}\tilde{\mu}^c_{\Gamma,-,h}(\Omega(\gamma_1, \ldots, \gamma_k)) &= \frac{Z_{\Gamma,-,h}(\mathcal{R}_c \cap \Omega(\gamma_1, \ldots, \gamma_k))}{Z_{\Gamma,-,h}(\mathcal{R}_c)} \\ &\leqq \frac{Z_{\Gamma,-,h}(\mathcal{R}_c \cap \Omega(\gamma_1, \ldots, \gamma_k))}{Z_{\Gamma,-,h}(\mathcal{R}_c \cap \Omega'(\gamma_1, \ldots, \gamma_k))} \\ &\leqq \exp\left(-\beta \sum_{i=1}^{k} |\gamma_i| + \beta h \sum_{i=1}^{k} |\Theta(\gamma_i)|\right). \end{aligned} \tag{29}$$

The proof of the lemma is finished by inserting (28) into (29). q.e.d.

Recall that $\beta' = \beta - \log b$, where b was introduced with the inequality (4).

Lemma 4. *Suppose that $\beta > \log b$ and $c = A/h$, where $A < 2d\beta'/\beta$ and $0 < h < 1$, then for arbitrary $\Gamma \in \mathcal{F}$,*

$$\tilde{\mu}^c_{\Gamma,-,h}(\partial_\Gamma \mathcal{R}_c) \leqq C(T, A)|\Gamma| \exp((-2d\beta' A^{d-1} + \beta A^d)/h^{d-1}).$$

Proof. Each configuration in $\partial_\Gamma \mathcal{R}_c$ has the property that there is at least one site $x \in \Gamma$ which is exterior but adjacent to a collection of outer contours $\gamma_1, \ldots, \gamma_k$, where k is 1, or 2, or $\ldots$, or $2d$, such that the flipping of the spin at x (from -1 to $+1$) would make these contours coalesce into a new outer contour γ, with $|\Theta(\gamma)| > c^d$. Because $\Theta(\gamma) = (\bigcup_i \Theta(\gamma_i)) \cup \{x\}$ we must have

$$\sum_{i=1}^{k} |\Theta(\gamma_i)| \geqq c^d - 1. \tag{30}$$

We sum now over all x and $\{\gamma_i\}$, making a partition according to the value of $\sum |\gamma_i| =: l$, and using Lemma 3 and the counting bound (4); in this fashion we obtain

$$\tilde{\mu}^c_{\Gamma,-,h}(\partial_\Gamma \mathcal{R}_c) \leqq C_1 |\Gamma| \sum_{l \geqq l_0} \exp((-\beta' + \beta ch/(2d))l),$$

where l_0 is the minimum value that $\sum |\gamma_i|$ can assume. An estimate for l_0 can be obtained from (28) (valid here from the same reason it was valid in the proof of Lemma 3) and (30):

$$\sum |\gamma_i| \geqq (2d/c)(c^d - 1) = 2dc^{d-1} - 2d/c.$$

Using now the facts that $c = A/h$, $A < 2d\beta'/\beta$ and $h < 1$ we obtain the bound stated in the lemma. q.e.d.

Observe that in order to exploit Lemma 4 the optimal choice is

$$A = 2(d-1)\beta'/\beta, \tag{31}$$

which leads to the bound

$$\tilde{\mu}^c_{\Gamma,-,h}(\partial_\Gamma \mathcal{R}_c) \leqq C(T, A)|\Gamma| \exp(-2^d(d-1)^{d-1}(\beta'/\beta)^d \beta/h^{d-1}). \tag{32}$$

Observe that, thanks to (18), the choice (31) and the bound (32), are very close, at low temperatures, to the values suggested by the heuristic reasoning presented before. Also the maximum allowed value of A in Lemma 4 makes the maximum value for c there close to $2d/h$, which is the threshold for the side-length of a cubic droplet of spins $+1$ to the droplet to have a negative energy, with respect to the configuration $-\underline{1}$.

We will use the notation $\langle f \rangle^c_{\Gamma,-,h}$ to denote the expectation of the observable f with respect to the measure $\tilde{\mu}^c_{\Gamma,-,h}$.

Lemma 5. *Suppose that* $\beta > \log b$, $A = 2(d-1)\beta'/\beta$, *and* $c = A/h$, *then*
i)

$$\mathbb{P}(\tau^c_{\Lambda_h,h} < t_h) \to 0 \quad \text{as } h \searrow 0 \,.$$

ii) *For every non-decreasing local observable* f

$$\limsup_{h \searrow 0} \mathbb{E}(f(\sigma^-_{\Lambda_h,-,h;t_h})) \leqq \limsup_{h \searrow 0} \langle f \rangle^{A/h}_{\Lambda_h,-,h} \,.$$

Proof. Part (ii) follows from part (i), (26) and the stationarity of the process $(\tilde{\sigma}^{c,\tilde{\mu}^c_{\Lambda_h,-,h}}_{\Lambda_h,-,h;t})$.

We turn now to the proof of part (i). For this we also exploit the stationarity of the process $(\tilde{\sigma}^{c,\tilde{\mu}^c_{\Lambda_h,-,h}}_{\Lambda_h,-,h;t})$. If time were discrete ($=0,1,2,\ldots$), then we could bound $\mathbb{P}(\tau^c_{\Lambda_h,h} < t_h)$ from above by $(t_h+1)\tilde{\mu}^c_{\Lambda_h,-,h}(\partial_{\Lambda_h}\mathscr{R}_c)$, which vanishes as $h \searrow 0$, thanks to (32) and to the facts that $|\Lambda_h| \leqq \exp(d\lambda_1(T)/h^{d-1})$ and $t_h = \exp(\lambda/h^{d-1})$, with $\lambda < \lambda_1(T) = (2^d(d-1)^{d-1}/(d+1))(\beta'/\beta)^d\beta$. To adapt such an estimate to continuous time, we consider first times which are multiples of $\Delta := h^{d+1}$. The probability of the event $E(h)$, that for some time $s \in [0, t_h)$ which is of the form $s = j\Delta$ for some $j = 0,1,2,\ldots$, we have $\tilde{\sigma}^{c,\tilde{\mu}^c_{\Lambda_h,-,h}}_{\Lambda_h,-,h;s} \in \partial_{\Lambda_h}\mathscr{R}_c$, is bounded above by $(h^{-d-1}t_h+1)\tilde{\mu}^c_{\Lambda_h,-,h}(\partial_{\Lambda_h}\mathscr{R}_c)$, which still vanishes as $h \searrow 0$. We have to show now that it is unlikely that $\tau^c_{\Lambda_h,h} < t_h$ and $E(h)$ fails. Because of the strong Markov property, it is sufficient that we show that for an arbitrary configuration $\eta \in \partial_{\Lambda_h}\mathscr{R}_c$ it is unlikely that the process $(\tilde{\sigma}^{c,\eta}_{\Lambda_h,-,h;s})_{s \geqq 0}$ leaves the set $\partial_{\Lambda_h}\mathscr{R}_c$ before time $s = \Delta$. To show this, we use again the fact that, since $\eta \in \partial_{\Lambda_h}\mathscr{R}_c$, there is somewhere in Λ_h a site x which is exterior but adjacent to a collection of outer contours, $\gamma_1,\ldots,\gamma_k$, present in η, where k is 1, or 2, or $\ldots$, or $2d$, such that the flipping of the spin at x (from -1 to $+1$) would make these contours coalesce into a new outer contour γ, with $|\Theta(\gamma)| > c^d$. (If there is more than one such arrangement in the configuration η, choose one of them, according to some arbitrary rule.) The system, started from η will not leave $\partial_{\Lambda_h}\mathscr{R}_c$ before at least one spin in $\bigcup \Theta(\gamma^i)$ flips. Fom the fact that $|\bigcup \Theta(\gamma^i)| \leqq 2dc^d = Ch^{-d}$, and the upper bound on the rates of flip provided by H(4), it is clear that the desired result holds. This completes the proof of the lemma. q.e.d.

The proof of (25) will be completed by combining part (ii) of Lemma 5 with Lemma 2 and the following bound, which is a particular case of Lemma 6 below: for $\beta > \log b$ and $A < 2d(\beta'/\beta)$ fixed and every non-decreasing local observable f,

$$\limsup_{h \searrow 0} \langle f \rangle^{A/h}_{\Lambda_h,-,h} \leqq \langle f \rangle_- \,.$$

The extra uniformity with which we state Lemma 6 comes for free in the proof and will only be needed for other purposes in the next section.

Lemma 6. *Suppose that $\beta > \log b$ and $A < 2d(\beta'/\beta)$ are fixed, then for every non-decreasing local observable f*

$$\limsup_{c\to\infty}\left(\sup_{\Gamma\in\mathscr{F}}\ \sup_{0\leqq h\leqq A/c}\ \langle f\rangle^{c}_{\Gamma,-,h}\right)\leqq\langle f\rangle_{-}\,.$$

Proof. Let S be the support of f and $S_n = \{x\in\mathbb{Z}^d\colon \|x-y\|_1 \leqq n \text{ for some } y\in S\}$. Let F be the set of configurations in which there is an outer contour which surrounds a site in S_1 and a site which is not in S_n. From Lemma 3 and the assumption $A < 2d(\beta'/\beta)$, it follows via a standard summation over possible contours and the counting inequality (4) that for arbitrary $c > 0$, $\Gamma\in\mathscr{F}$ and $h\in[0, A/c]$,

$$\tilde{\mu}^{c}_{\Gamma,-,h}(F)\leqq C(T,A)n^d\exp(-\alpha(T,A)n)\,, \tag{33}$$

where $\alpha(T,A)=\beta'-\beta A/(2d)>0$ and $C(T,A)$ does not depend on c, Γ and h.

Given now $\varepsilon>0$, take n so large that $C(T,A)n^d\exp(-\alpha(T,A)n)<\varepsilon/\|f\|_\infty$. Observe that given ε, n depends on T, A and f, but that those are being held fixed as we let $c\to\infty$.

Defining $E=\Omega_{\Gamma,-}\setminus F$, we obtain from (33),

$$\langle f\rangle^{c}_{\Gamma,-,h}\leqq 2\varepsilon+\langle f\,|\,E\rangle^{c}_{\Gamma,-,h}=2\varepsilon+\langle f\,|\,E\cap\mathscr{R}_c\rangle_{\Gamma,-,h}\,. \tag{34}$$

We introduce now a partition $\{G_i\colon i\in I\}$ of $E\cap\mathscr{R}_c$. This partition is generated by the equivalence relation on $E\cap\mathscr{R}_c$ in which two configurations in this set are considered to be equivalent if

i) they are identical outside of S_n,
ii) all the outer contours which surround at least one site outside of S_n are identical in these two configurations,
iii) and these configurations are also identical inside these outer contours specified in (ii) above.

To each index $i\in I$ there corresponds a (maximal) subset W_i of $\mathbb{Z}^d$ (which contains $\mathbb{Z}^d\setminus S_n$ and does not intersect S) on which all the configurations which belong to G_i are completely specified and are identical to a configuration η_i, say. It is clear that $\eta_i(x)=-1$ for every site $x\in\partial_{\text{int}}W_i$. Moreover the following crucial remark is true: for fixed n, if c is so large that $c^d>|S_n|$, then for each $i\in I$ the occurrence or not of the event G_i can be verified by looking exclusively at the spins at the sites in W_i (but observe that for small c such a statement is in general false, because to know that G_i happens we have to be sure that there is no contour γ present which surrounds more than c^d sites in $\mathbb{Z}^d\setminus W_i$). From (34) and these observations it follows, via the Markov and FKG-Holley properties of the Gibbs measures that, for large c (depending only on T, A, f and ε),

$$\begin{aligned}\langle f\rangle^{c}_{\Gamma,-,h}&\leqq 2\varepsilon+\sum_{i\in I}\langle f\,|\,G_i\rangle_{\Gamma,-,h}\,\mu_{\Gamma,-,h}(G_i\,|\,E\cap\mathscr{R}_c)\\&\leqq 2\varepsilon+\sum_{i\in I}\langle f\rangle_{\Gamma\cap S_n,-,h}\,\mu_{\Gamma,-,h}(G_i\,|\,E\cap\mathscr{R}_c)\\&=2\varepsilon+\langle f\rangle_{\Gamma\cap S_n,-,h}\,.\end{aligned}$$

Making again use of the FKG-Holley inequalities,

$$\limsup_{c\to\infty}\left(\sup_{\Gamma\in\mathscr{F}}\ \sup_{0\leqq h\leqq A/c}\langle f\rangle^{c}_{\Gamma,-,h}\right)\leqq 2\varepsilon+\limsup_{c\to\infty}\langle f\rangle_{S_n,-,A/c}$$
$$\leqq 2\varepsilon+\langle f\rangle_{S_n,-,0}$$
$$\leqq 2\varepsilon+\langle f\rangle_{-}\,.$$

Since ε is arbitrary, the lemma is proven. q.e.d.

As was said in the paragraph preceding this lemma, this finishes the proof of (25) and hence also of part (i) of Theorem 1. We turn now to other results which follow from the same techniques.

Proof of Part (i) *of Theorem 2.* Suppose that $B < B_1(T) := 2d\beta'/\beta$ and choose an arbitrary $A \in (B, B_1(T))$. Under these conditions, $\Omega_{\Lambda(B/h),-} \subset \mathscr{R}_c$ and hence $\tilde{\mu}^{A/h}_{\Lambda(B/h),-,h} = \mu_{\Lambda(B/h),-,h}$. From Lemma 6 we obtain now, for any non-increasing local observable f,

$$\limsup_{h\searrow 0}\langle f\rangle_{\Lambda(B/h),-,h}=\limsup_{h\searrow 0}\langle f\rangle^{A/h}_{\Lambda(B/h),-,h}\leqq\langle f\rangle_{-}\,.$$

The complementary inequality,

$$\liminf_{h\searrow 0}\langle f\rangle_{\Lambda(B/h),-,h}\geqq\langle f\rangle_{-}\,,$$

is an immediate consequence of the FKG-Holley inequalities. q.e.d.

Sketch of the Proof of Part (i) *of Theorem 4.* That the claimed result holds with $\lambda_1(T, D) = 2^d(d-1)^{d-1}(\beta'/\beta)^d\beta - dD$ can be proven by the same arguments used above to prove part (i) of Theorem 1, but with Λ_h replaced by $\Lambda(B/h)$ and without using Lemma 2.

If $2^d(d-1)^{d-1}(\beta'/\beta)^d\beta - dD < (2^d(d-1)^{d-1}/(d+1))(\beta'/\beta)^d\beta$, then we want to prove that the result also holds with $\lambda_1(T, D) = (2^d(d-1)^{d-1}/(d+1))(\beta'/\beta)^d\beta$. But in this case $D > (2^d(d-1)^{d-1}/(d+1))(\beta'/\beta)^d\beta$, and one can complete the proof by combining the statement in part (i) of Theorem 1 with Lemma 1. q.e.d.

We comment next on two papers which are closely related to the material in the present section. The first one is [CCO]. In that paper the authors, motivated by an approach to metastability proposed by Lebowitz and Penrose, constructed states which to some extent have the properties expected from "metastable states." For the finite system in the box $\Lambda(l)$ with $-$ boundary conditions, such states were constructed as "restricted ensembles," defined there precisely as what we denote here by $\tilde{\mu}^{c}_{\Lambda(l),-,h}$, with c not uniquely specified, but chosen properly from a certain range of values. In [CCO] the authors show that if the system starts from this state, and evolves with the (unrestricted) Glauber dynamics, then its rate of escape from $\mathscr{R}_c$, divided by the volume $|\Lambda(l)|$, can be made very small, uniformly in l, provided h is small enough and c is chosen properly, as a function of h (close to $2/h$ in 2 dimensions). To prove this result, the authors of [CCO], bounded this rate of escape per volume at time 0 by a multiple of $\tilde{\mu}^{c}_{\Lambda(l),-,h}(\partial_{\Lambda(l)}\mathscr{R}_c)$, and showed that the rate of escape is monotone non-increasing. This motivated them to derive the estimates that we reproduced here in Lemmas 3 and 4. An important difference between the use that [CCO] made of those estimates and our use of them is that in our case the "restricted ensemble" and the corresponding restricted Glauber dynamics, which we introduced, are tools rather than objects of interest in themselves. [CCO] also considered the thermodynamic limit, $l \nearrow \infty$, of $\tilde{\mu}^c_{\Lambda(l),-,h}$, and proved that such an object exists and has good mixing properties. This part of their work relies on machinery developed by Minlos and Sinai, and is technically the most difficult part of their paper. But because they considered h and c as fixed in this limiting procedure, their results here are not of the type that we needed in the present paper, and which are the content of our Lemma 6.

It is worth pointing out that an important drawback in the results of [CCO] is the fact that the rate of escape from the "metastable state," as studied there, grows proportionally to the volume of the system, for fixed values of h and c, and so diverges as the volume diverges. This is clearly so, because one considers the system to have escaped as soon as a large droplet is formed anywhere. In this connection, it is common to hear that metastability in the stochastic Ising models is a purely finite-volume phenomenon. Actually, as we see in the present paper, this is not the case if a proper characterization of metastability is adhered to. On the other hand one has to be careful with the use of the term "metastability" and differences are indeed found in the sort of "metastable behavior" found in different regimes parametrized by volume, temperature and external field. We will say more about this in Sect. 4.

The other paper closely related to part (i) of Theorem 1 is [Van]. In that paper the author is motivated by a characterization of metastability proposed by G. Sewell, who termed it "normal" (as opposed to "ideal") metastability. In [Van] the stochastic Ising models on the infinite lattice are considered (in 2 dimensions and with the specific choice of rates given in Example 2, i.e., the Heat Bath dynamics). The author then considers the evolution starting from the probability measure obtained as a thermodynamic limit of the restricted ensembles of [CCO], with a fixed h and c. Essentially, he shows that for c chosen properly as a function of h (close to $2/h$), there is a plateau in the relaxation curves of the local observables s_x, $x \in \mathbb{Z}^d$, given by $s_x(\sigma) = \sigma(x)$, when the evolution is started from those states. The length of the plateau is estimated from below by $C(T)/h$, for some $C(T) > 0$. Unfortunately, possibly due to a misprint, a much stronger conclusion seems to be claimed in [Van]; the statement in the next-to-last display in p. 2670 of that paper indicates that the plateau has length bounded below by $\exp((4 - \varepsilon)\beta/h)$, for all small enough T. Such a result would even be in contradiction with the conjectured upper bound on the relaxation time, obtained from the heuristics! Actually, looking for the origin of this displayed inequality, one finds that it does not follow from the previous inequality (because the function $i(t)$, which was defined immediately after display (3.13) in terms of the function $g(t)$, which was defined in (3.6), grows exponentially with t).

Apart from the different initial states considered, the lower bound for the length of the plateau obtained in this section may be considered as an important strengthening of the result in [Van]. It is worth pointing out also that while the analysis in [Van] uses estimates from [CCO] and can only be implemented at low enough temperatures, the simple argument that was presented in subsection 1-vi already gives a bound which in two dimensions is of the order $C(T)/h^{1/3}$, and applies to all the stochastic Ising models considered in this paper, all local observables, and all temperatures below T_c.

3. Upper Bound on the Relaxation Time

The first step to proof part (ii) of Theorem 1 was already taken, when we proved part (ii) of Proposition 1. In this section we will prove the complementary inequality: under the conditions stated in part (ii) of Theorem 1, and supposing that the local observable f is non-decreasing,

$$\liminf_{h \searrow 0} \mathbb{E}(f(\sigma^{-}_{h;t})) \geqq \int f d\mu_{+} . \tag{35}$$

Similarly to what was done to prove (25) in Sect. 2, we will prove (35) by reducing the problem to a problem referring to a finite system with $-$ boundary conditions, whose size grows as $h \searrow 0$. But this time the linear size of this system will scale as $1/h$, rather than as an exponential of a power of such a quantity. Since we want to consider the evolution of the system up to a time which actually scales as an exponential of $1/h^{d-1}$, we will not be using the "bounded speed of propagation of effects" in order to relate the finite to the infinite system, but instead the basic coupling. This comparison is the content of the following lemma, which follows from (17).

Lemma 7. *For every $\Gamma \in \mathscr{F}$ every $h \in (-h(T), h(T))$ and every non-decreasing local observable f,*

$$\mathbb{E}(f(\sigma^{-}_{\Gamma,-,h;t})) \leqq \mathbb{E}(f(\sigma^{-}_{h;t})) ,$$

for all $t \geqq 0$.

We will take $\Gamma = \Lambda(B/h)$ with $B > 2d$ fixed. When T is small enough we should have, using the notation in Theorem 2, $B > B_2(T)$, so that in equilibrium this finite system should be locally close to the plus-phase. In order to use part (ii) of Theorem 2, we will in the first subsection below prove Theorem 3, which, as we know, from Subsect. 1-iv, implies part (ii) of Theorem 2. Morally these results state that, at low temperatures, for the finite system in the box $\Lambda(B/h)$, in equilibrium the $-$ boundary conditions have little influence on the spins in the center of the box. Still another way to put it is to say that in equilibrium the center of the system is likely to be covered by a droplet of the plus-phase.

In the second subsection below, we will prove that the corresponding stochastic Ising model in the box $\Lambda(B/h)$ with $-$ boundary conditions and started with all spins down, relaxes to its equilibrium in a time of the order of an exponential of the surface of the system. The inequality (35) will follow from this result combined with Lemma 7 and part (ii) of Theorem 2.

These two steps, in each one of the subsections below, are well separated and each one interesting for its own sake; the former one is a purely static result while the latter one is basically dynamic. Before starting with the proofs, we should observe that, nevertheless, there is one important drawback to our approach for proving (35). By considering the basic-coupling type comparison given by Lemma 7, with $\Gamma = \Lambda(B/h)$, we are freezing the spins outside of the box $\Lambda(B/h)$ from the beginning, and giving up any hope of exploiting the possibility that droplets that are created far from the origin could grow and cover the neighborhood of the origin in a time much shorter than the time needed for the finite system in $\Lambda(B/h)$ to relax (this should essentially be the time for a big droplet to pop up spontaneously inside this finite system with $-$ boundary conditions). In other words, we are giving up any hope of obtaining the factor $d+1$ which should divide $\beta E_{\max}$ in the exponential rate with which $\lambda_2(T)$ grows with $1/h^{d-1}$. We are adopting the present approach because we do not know yet how to control, in a mathematically precise way, the growth of supercritical droplets (in the regime treated in the present paper, with T fixed and $h \searrow 0$); this is one of the main directions for further research. (But it is important to observe that for different regimes, in which $T \searrow 0$ with h fixed or with $h \searrow 0$ simultaneously, in a proper way, we can adapt and extend the techniques reviewed in [Sch1] to obtain better upper bounds for the relaxation time, which include the factor $d + 1$. We will say more about this point in Sect. 4).

3-i. Proof of Theorem 3. The first lemma below implies that if $B > 2d$, then, at low temperatures, big droplets are likely to be present in typical configurations selected according to the probability measure $\mu_{\Lambda(B/h), -, h}$, when h is positive but small. Recall that $m^*(T)$ is the spontaneous magnetization at temperature T and that $m^*(T) \to 1$ as $T \searrow 0$.

Lemma 8. *Suppose that $\beta > \log b$ and $A < 2d(\beta'/\beta)$ are fixed, then*

$$\limsup_{h \searrow 0} h^{d-1} \log(\mu_{\Lambda(B/h), -, h}(\mathscr{R}_{A/h})) \leqq -\beta(m^*(T)B^d - 2dB^{d-1}) .$$

Proof. Using the notation (19), we have

$$\mu_{\Lambda(B/h), -, h}(\mathscr{R}_{A/h}) = \frac{Z_{\Lambda(B/h), -, h}(\mathscr{R}_{A/h})}{Z_{\Lambda(B/h), -, 0}(\mathscr{R}_{A/h})} \frac{Z_{\Lambda(B/h), -, 0}(\mathscr{R}_{A/h})}{Z_{\Lambda(B/h), -, 0}} \frac{Z_{\Lambda(B/h), -, 0}}{Z_{\Lambda(B/h), -, h}} . \tag{36}$$

The third fraction above can be estimated by first writing it as the product of three further fractions:

$$\frac{Z_{\Lambda(B/h), -, 0}}{Z_{\Lambda(B/h), -, h}} = \frac{Z_{\Lambda(B/h), -, 0}}{Z_{\Lambda(B/h), +, 0}} \frac{Z_{\Lambda(B/h), +, 0}}{Z_{\Lambda(B/h), +, h}} \frac{Z_{\Lambda(B/h), +, h}}{Z_{\Lambda(B/h), -, h}} . \tag{37}$$

Now, we estimate each one of these three fractions, in the right-hand side of (37), as follows:

$$\log\left(\frac{Z_{\Lambda(B/h), -, 0}}{Z_{\Lambda(B/h), +, 0}}\right) = 0 , \tag{38}$$

$$\begin{aligned} \log\left(\frac{Z_{\Lambda(B/h), +, 0}}{Z_{\Lambda(B/h), +, h}}\right) &= -\int_0^h \frac{\partial}{\partial h'} \log Z_{\Lambda(B/h), +, h'}\, dh' \\ &= -(\beta/2)\int_0^h \left\langle \sum_{x \in \Lambda(B/h)} \sigma(x) \right\rangle_{\Lambda(B/h), +, h'} dh' \\ &\leqq -(\beta/2)\int_0^h m^*(T)\,|\Lambda(B/h)|\, dh' \\ &= -(\beta/2) m^*(T) B^d/h^{d-1} , \end{aligned} \tag{39}$$

and

$$\log\left(\frac{Z_{\Lambda(B/h), +, h}}{Z_{\Lambda(B/h), -, h}}\right) \leqq \beta\,|\partial_{\text{ext}}\Lambda(B/h)| \leqq \beta 2dB^{d-1}/h^{d-1} . \tag{40}$$

The equalities and inequalities (37)–(40) imply that the third fraction in the right-hand side of (36) can be estimated by

$$\log\left(\frac{Z_{\Lambda(B/h),-,0}}{Z_{\Lambda(B/h),-,h}}\right) \leqq -\beta((m^*(T)B^d/2) - 2dB^{d-1})/h^{d-1} . \tag{41}$$

The second fraction in (36) is trivially estimated above by 1, while the first one requires some extra argumentation. We start as in (39) above, by writing

$$\log\left(\frac{Z_{\Lambda(B/h),-,h}(\mathscr{R}_{A/h})}{Z_{\Lambda(B/h),-,0}(\mathscr{R}_{A/h})}\right) = (\beta/2)\int_0^h \left\langle \sum_{x\in\Lambda(B/h)} \sigma(x)\right\rangle^{A/h}_{\Lambda(B/h),-,h'} dh' . \tag{42}$$

Now, thanks to the assumptions on β and A, we can use Lemma 6 from the last section to obtain the following bound:

$$\limsup_{h\searrow 0}\left(\sup_{x\in\Lambda(B/h)}\ \sup_{0\leqq h'\leqq h} \langle\sigma(x)\rangle^{A/h}_{\Lambda(B/h),-,h'}\right) \leqq -m^*(T) . \tag{43}$$

(The uniformity in h' and in Γ in Lemma 6, which was not needed in the last section, was introduced in the statement of that lemma precisely for the purpose of being used here. Actually, we are using translation invariance to transform the uniformity in Γ into uniformity over the translates of the observable $\sigma(0)$.)

From (42) and (43) we obtain

$$\limsup_{h\searrow 0} h^{d-1}\log\left(\frac{Z_{\Lambda(B/h),-,h}(\mathscr{R}_{A/h})}{Z_{\Lambda(B/h),-,0}(\mathscr{R}_{A/h})}\right) \leqq -(\beta/2)m^*(T)B^d . \tag{44}$$

Finally, the lemma follows from (37), (41), the statement made immediately after (41), and (44). q.e.d.

Lemma 8 can be exploited by taking strictly monotone sequences as follows. First take arbitrarily $B_n \searrow 2d$. Now take $\beta_n \nearrow \infty$, so that $B_n m^*(T_n) > 2d$, where $T_n = 1/\beta_n$. Finally take $A_n \nearrow 2d$ such that $A_n < 2d(\beta_n'/\beta_n)$, where $\beta_n' = \beta_n - \log b$. Observe that as $n\to\infty$, we have $(A_n/h)^d/|\Lambda(B_n/h)| \to 1$, so that from Lemma 8 we conclude that for large n, in typical configurations chosen according to the law $\mu_{\Lambda(B_n/h),-,h}$ at inverse temperature β_n, when h is small there is typically a (necessarily unique) outer contour which surrounds a very large fraction of $\Lambda(B_n/h)$. (This fraction can be made as close to 1 as we want by taking n large enough; observe that when we refer to small h, we are considering a limit in which $h \searrow 0$, after having chosen and fixed n.) In order to complete now the proof of Theorem 3, we have to argue that such a contour is also likely to surround $\Lambda(d/h)$. This seens a very natural idea, but its implementation turned out to be very delicate in arbitrary dimension. We present below a dimension independent solution to this problem, based in part on some ideas in [Mar]. In the appendix we will present another solution, which is simpler, but can be used only in the case $d = 2$.

Given a configuration $\sigma\in\Omega$, we define $\mathscr{C}(\sigma)$ as the set of sites in $\mathbb{Z}^d$ which belong to infinite clusters of spins -1 in σ. It is clear that the sites in $\mathscr{C}(\sigma)$ cannot be surrounded by any contour. From this observation, Lemma 8 and the remarks made above, we obtain the following result, which is stated as a new lemma for future reference, in the form in which we will use it.

Lemma 9. *Assume given two strictly monotone sequences $B_n \searrow 2d$ and $\beta_n \nearrow \infty$, such that, for each n the inequality $B_n m^*(T_n) > 2d$ is satisfied. Given $\varepsilon > 0$, for all large n,*

$$\lim_{h \searrow 0} \mu_{\Lambda(B_n/h), -, h}(\{\sigma: |\mathscr{C}(\sigma) \cap \Lambda(2d/h)| > \varepsilon/h^d\}) = 0 ,$$

provided $\beta \geqq \beta_n$.

Recall that $\mathscr{B}$ is the event that the box $\Lambda(d/h)$ intersects an infinite cluster of spins -1. We will now cover this event with (possibly overlapping) events, whose probabilities will be estimated using Lemma 9 above and Lemma 11 below. Recall the following notation for boxes centered at the origin $V_i := \{x \in \mathbb{Z}^d: \|x\|_\infty \leqq i\}$. Below we will be concerned with values of i in the set $I := \{i: \Lambda((3/2)d/h) \subset V_i \subset \Lambda(2d/h)\}$. Recall also that a σ-chain is a chain of sites which in the configuration σ have all spins negative. Given a configuration $\sigma \in \Omega$, we define for each $i \in I$, $M_i(\sigma)$ as the set of sites $x \in V_i \cap \mathscr{C}(\sigma)$ which are connected to $\Lambda(d/h)$ by a σ-chain entirely contained in V_{i-1} except, possibly, for its end-point at x. Next we define $L_i(\sigma) := M_i(\sigma) \setminus V_{i-1}$.

Now we introduce some of the events whose union covers the event $\mathscr{B}$. Assume $\alpha > 0$ fixed, and define

$$\mathscr{B}^\alpha_{i,l} := \{\sigma \in \Omega: l = |L_{i+1}(\sigma)| \leqq \alpha |M_i(\sigma)|^{(d-1)/d}\} .$$

Lemma 10. *For every $\alpha > 0$, the following holds for all small enough positive h.*

$$\mathscr{B} \subset \left(\bigcup_{\substack{i \in I \\ l \geqq 1}} \mathscr{B}^\alpha_{i,l} \right) \cup \left\{ \sigma: |\mathscr{C}(\sigma) \cap \Lambda(2d/h)| \geqq \left(\frac{\alpha}{2^d - 1} \right)^d \left(\frac{d}{8h} \right)^d \right\} .$$

Proof. We will assume that $\sigma \in \mathscr{B}$ and for all $i \in I$ and $l \geqq 1$, $\sigma \notin \mathscr{B}^\alpha_{i,l}$, and we have to show that $|\mathscr{C}(\sigma) \cap \Lambda(2d/h)| \geqq (\alpha/(2^d - 1))^d ((1/8)d/h)^d$, for small enough h.

Set $m_i := |M_i(\sigma)|$, $l_i := |L_i(\sigma)|$, $r := \min I$, $s := (\max I) - r$ and $\alpha' := \alpha/(2^d - 1)$. We will prove by induction on $u \in \{1, \ldots, s\}$ that

$$m_{r+u} \geqq (\alpha')^d u^d . \tag{45}$$

For $u = 1$ we observe that, since $\sigma \in \mathscr{B}$ and $\Lambda((3/2)d/h) \subset V_{r+1}$, there must be a σ-chain contained in $M_{r+1}(\sigma)$ with length at least $(1/8)d/h$, when h is small. So we have

$$m_{r+1} \geqq (1/8)d/h \geqq (\alpha')^d u^d ,$$

for small h. Assuming now that (45) is true for a certain $u \geqq 1$, we prove it for $u + 1$ in the following way. Clearly, for each i, $M_i(\sigma) \cup L_{i+1}(\sigma) \subset M_{i+1}(\sigma)$ and $M_i(\sigma) \cap L_{i+1}(\sigma) = \emptyset$, hence

$$\begin{aligned} m_{r+u+1} &\geqq m_{r+u} + l_{r+u+1} \geqq m_{r+u} + \alpha(m_{r+u})^{(d-1)/d} \\ &\geqq (\alpha')^d u^d + ((2^d - 1)\alpha')(\alpha')^{d-1} u^{d-1} \\ &= (\alpha')^d (u^d + (2^d - 1)u^{d-1}) \geqq (\alpha'(u + 1))^d , \end{aligned}$$

where in the last step we used the fact that $u \geqq 1$. This completes the proof of (45).

The lemma now follows from the observation that $|\mathscr{C}(\sigma) \cap \Lambda(2d/h)| \geqq m_{r+s} \geqq (\alpha')^d s^d$, and that for small h, $s \geqq (1/8)d/h$. q.e.d.

Lemma 11. *Suppose $B \geqq 2d$, $\beta > 2(\log b + \log 2)$ and $\alpha < (\beta/2 - \log b - \log 2)/(2\beta d)$ are fixed, then*

i) *for each* $i \in I$ *and each* $l \geqq 1$,

$$\mu_{\Lambda(B/h),-,h}(\mathscr{B}^{\alpha}_{i,l}) \leqq C_1(\beta,\alpha,d)g(\beta,h)\exp(-C_2(\beta,\alpha,d)/h),$$

where $C_2(\beta,\alpha,d) > 0$, *and for each* β, $g(\beta,h) \to 1$ *as* $h \searrow 0$.

ii)

$$\lim_{h \searrow 0} \sum_{i \in I} \sum_{l \geqq 1} \mu_{\Lambda(B/h),-,h}(\mathscr{B}^{\alpha}_{i,l}) = 0.$$

Proof. Part (ii) clearly follows from part (i). To prove part (i), we will show that configurations in $\mathscr{B}^{\alpha}_{i,l}$ can be transformed "at little cost" into configurations in which $\Lambda(d/h)$ is surrounded by a positive contour and moreover inside of this contour there is at least one negative contour which is quite large because it has to surround sites inside $\Lambda(d/h)$ and outside $\Lambda(3d/(2h))$. Part (i) of the lemma will then follow from an estimate of the probability of such configurations.

The transformation mentioned above is F_i: $\mathscr{B}^{\alpha}_{i,l} \to \Omega$, defined by

$$(F_i(\sigma))(x) = \begin{cases} +1 & \text{if } x \in L_{i+1}(\sigma), \\ \sigma(x) & \text{otherwise}. \end{cases}$$

Define for each $\psi \in F_i(\mathscr{B}^{\alpha}_{i,l})$,

$$N_{i,l}(\psi) = |\{\sigma \in \mathscr{B}^{\alpha}_{i,l} \colon F_i(\sigma) = \psi\}|.$$

The "little cost" alluded to above is contained in the straightforward estimate

$$\begin{aligned} \mu_{\Lambda(B/h),-,h}(\mathscr{B}^{\alpha}_{i,l}) &= \sum_{\sigma \in \mathscr{B}^{\alpha}_{i,l}} \mu_{\Lambda(B/h),-,h}(\sigma) \\ &\leqq \exp(\beta 2dl) \sum_{\sigma \in \mathscr{B}^{\alpha}_{i,l}} \mu_{\Lambda(B/h),-,h}(F_i(\sigma)) \\ &= \exp(\beta 2dl) \sum_{\psi \in F_i(\mathscr{B}^{\alpha}_{i,l})} N_{i,l}(\psi)\mu_{\Lambda(B/h),-,h}(\psi). \end{aligned} \tag{46}$$

To estimate the right-hand side of (46) we will use the following facts, where we always suppose that $i \in I$, and $l \geqq 1$.

i) If $\psi \in F_i(\mathscr{B}^{\alpha}_{i,l})$, then there is no chain of spins -1 in ψ connecting any site inside $\Lambda(d/h)$ to a site outside $\Lambda(B/h)$. (Because if $\psi = F_i(\sigma)$, then such a chain would also be a σ-chain, and hence would have to cross a site in $L_{i+1}(\sigma)$, which is in contradiction with the definition of the transformation F_i). Therefore, there is a contour in ψ which surrounds $\Lambda(d/h)$.

We will use the notation $\mathscr{N}_{i,l}(\gamma_0)$ to denote the set of configurations $\psi \in F_i(\mathscr{B}^{\alpha}_{i,l})$ such that the outer contour which is present in ψ and surrounds $\Lambda(d/h)$ is γ_0.

ii) Suppose that $\psi \in \mathscr{N}_{i,l}(\gamma_0)$ and let σ be one of the configurations in $\mathscr{B}^{\alpha}_{i,l}$ such that $\psi = F_i(\sigma)$. In the configuration ψ, the spins in the interior boundary of γ_0 are positive, while all sites in $M_i(\sigma)$ have negative spins, and therefore must be surrounded by negative contours, which are surrounded by γ_0. Moreover, because $i \in I$ and $l \geqq 1$, the set of sites $M_i(\sigma)$ contains at least one chain which is contained in V_i and which has one end-point inside $\Lambda(d/h)$ and the other outside $\Lambda(3/2)d/h)$. This chain is not affected by the transformation F_i, and so it is also present in ψ. Hence, in ψ, there must be at least one negative contour which surrounds a site in $\Lambda(d/h)$ and a site outside $\Lambda((3/2)d/h)$.

We will use the notation $\mathscr{N}_{i,l}(\gamma_0; \gamma_1, \ldots, \gamma_k)$ to denote the set of configurations $\psi \in \mathscr{N}_{i,l}(\gamma_0)$ such that their negative contours which surround sites inside $\Lambda(d/h)$ and outside $\Lambda(3d/2h))$ are precisely $\gamma_1, \ldots, \gamma_k$.

iii) Suppose that $\psi \in \mathcal{N}_{i,l}(\gamma_0; \gamma_1, \ldots, \gamma_k)$, $k \geqq 1$. We will show below that if $\psi = F_i(\sigma)$, then all the sites in $L_{i+1}(\sigma)$ must be adjacent to some γ_j, $j = 1, \ldots, k$. The relevance of this observation is that it clearly implies the following bound:

$$N_{i,l}(\psi) \leqq 2^{\sum_{j=1,\ldots,k} |\gamma_j|} . \tag{47}$$

To prove the claim above observe first that if $x \in L_{i+1}(\sigma)$ then x is neighbor to one of the end-points of a σ-chain, $\mathscr{A}$, contained in V_i, and which has its other end-point inside $\Lambda(d/h)$. The chain $\mathscr{A}$ is also a ψ-chain, and we denote by $\bar{\mathscr{A}}$ the cluster of spins -1 in ψ which contains $\mathscr{A}$. Because $\psi(x) = +1$, x is separated from $\bar{\mathscr{A}}$ by a contour, and to finish the proof of the claim we only have to argue that x must be in the exterior (not the interior) boundary of this contour, since then this contour has to be a negative contour which surrounds $\mathscr{A}$, and so must be one of the γ_j, $j = 1, \ldots, k$. That x cannot be in the interior boundary of the contour which separates it from $\bar{\mathscr{A}}$ is an immediate consequence of the fact that (while $x \notin V_i$) $\bar{\mathscr{A}} \subset V_i$, which can be proved as follows. Define $\mathscr{A}' := \bar{\mathscr{A}} \cap V_i$, and observe that from the definition of the transformation F_i, in the configuration ψ, the spins at all sites outside of V_i, which are neighbors of sites in $\mathscr{A}'$ are $+1$. This means that $\bar{\mathscr{A}} = \mathscr{A}' \subset V_i$, and completes the argument.
iv) Suppose again that $\psi \in \mathcal{N}_{i,l}(\gamma_0; \gamma_1, \ldots, \gamma_k)$, $k \geqq 1$. If $\psi = F_i(\sigma)$, then $M_i(\sigma) \subset \bigcup_{j=1,\ldots,k} \Theta(\gamma_j)$. Therefore, from the definition of $\mathscr{B}^{\alpha}_{i,l}$ and the isoperimetric inequality (3), we also obtain

$$l \leqq \alpha |M_i(\sigma)|^{(d-1)/d} \leqq \alpha \sum_{j=1,\ldots,k} |\gamma_j| . \tag{48}$$

Using $\mathscr{S}_k(B, h)$ to denote the family of non-empty sets of k compatible contours inside $\Lambda(B/h)$ which have the property of surrounding sites inside $\Lambda(d/h)$ and outside $\Lambda(3d/(2h))$, we obtain now from (46), (47) and (48),

$$\begin{aligned}
&\mu_{\Lambda(B/h),-,h}(\mathscr{B}^{\alpha}_{i,l}) \\
&\leqq \exp(\beta 2dl) \sum_{\gamma_0} \sum_{k \geqq 1} \sum_{\{\gamma_1,\ldots,\gamma_k\} \in \mathscr{S}_k(B,h)} \sum_{\psi \in \mathcal{N}_{i,l}(\gamma_0;\gamma_1,\ldots,\gamma_k)} N_{i,l}(\psi) \mu_{\Lambda(B/h),-,h}(\psi) \\
&\leqq \sum_{\gamma_0} \sum_{k \geqq 1} \sum_{\{\gamma_1,\ldots,\gamma_k\} \in \mathscr{S}_k(B,h)} \sum_{\psi \in \mathcal{N}_{i,l}(\gamma_0;\gamma_1,\ldots,\gamma_k)} 2^{\sum_{j=1,\ldots,k} |\gamma_j|} \\
&\quad \times e^{2\beta d\alpha \sum_{j=1,\ldots,k} |\gamma_j|} \mu_{\Lambda(B/h),-,h}(\psi) .
\end{aligned} \tag{49}$$

Now we estimate $\mu_{\Lambda(B/h),-,h}(\psi)$ by conditioning on what γ_0 is and combining the Markov property of the Gibbs measures with Lemma 12 below (with $\Gamma = \Theta(\gamma_0) \setminus \partial_{\text{int}} \gamma_0$). In this fashion we obtain from (49),

$$\begin{aligned}
&\mu_{\Lambda(B/h),-,h}(\mathscr{B}^{\alpha}_{i,l}) \\
&\leqq \sum_{\gamma_0} \mu_{\Lambda(B/h),-,h}(\mathcal{N}_{i,l}(\gamma_0)) \\
&\quad \times \sum_{k \geqq 1} \sum_{\{\gamma_1,\ldots,\gamma_k\} \in \mathscr{S}_k(B,h)} \exp\left((-\beta + \log 2 + 2\beta d\alpha) \sum_{j=1,\ldots,k} |\gamma_j|\right) \\
&\leqq \sum_{k \geqq 1} \sum_{\{\gamma_1,\ldots,\gamma_k\} \in \mathscr{S}_k(B,h)} \exp\left((-\beta + \log 2 + 2\beta d\alpha) \sum_{j=1,\ldots,k} |\gamma_j|\right) .
\end{aligned} \tag{50}$$

We will now partition the sum over $\{\gamma_1, \ldots, \gamma_k\}$ according to the value of $v := \sum_{j=1,\ldots,k} |\gamma_j|$, and for this purpose introduce the notation $\mathscr{S}_{k,v}(B,h) := \{\{\gamma_1, \ldots, \gamma_k\} \in \mathscr{S}_k(B,h) : \sum_{j=1,\ldots,k} |\gamma_j| = v\}$. The two following inequalities will be useful in the sequel. The first one is an immediate consequence of the definition of $\mathscr{S}_{k,v}(B,h)$, while the second one can be obtained by considering first the choice of one face from each contour γ_j, $j = 1, \ldots, k$ – which can be made in no more than $(2d|\Lambda(B/h)|)^k/k!$ ways – and then using the counting inequality (4).
For all $\{\gamma_1, \ldots, \gamma_k\} \in \mathscr{S}_{k,v}(B,h)$ we have

$$v = \sum_{j=1,\ldots,k} |\gamma_j| \geqq k(d/8)/h \,. \tag{51}$$

$$|\mathscr{S}_{k,v}(B,h)| \leqq \frac{(2d|\Lambda(B/h)|)^k}{k!} b^{v+k} \leqq \frac{(2d(B/h)^d b)^k}{k!} b^v \,. \tag{52}$$

From (50), (51) and (52),

$$\begin{aligned}
&\mu_{\Lambda(B/h),-,h}(\mathscr{B}^{\alpha}_{i,l}) \\
&\quad \leqq \sum_{k \geqq 1} \sum_{v \geqq k(d/8)/h} \sum_{\{\gamma_1,\ldots,\gamma_k\} \in \mathscr{S}_{k,v}(B,h)} \exp((-\beta/2 + \log 2 + 2\beta d\alpha)v) \exp((-\beta/2)v) \\
&\quad \leqq \sum_{k \geqq 1} \frac{(2d(B/h)^d b)^k}{k!} (e^{-(\beta/2)(d/8)/h})^k \\
&\qquad \times \sum_{v \geqq k(d/8)/h} \exp((-\beta/2 + \log 2 + \log b + 2\beta d\alpha)v) \\
&\quad \leqq \exp(2d(B/h)^d b e^{-(\beta/2)(d/8)/h}) C_1(\beta, \alpha, d) \\
&\qquad \times \exp((-\beta/2 + \log 2 + \log b + 2\beta d\alpha)(d/8)/h) \,,
\end{aligned}$$

provided $\alpha < (\beta/2 - \log b - \log 2)/(2\beta d)$. This finishes the proof of part (i) of the lemma. q.e.d.

We will use below the notation $\Omega^-(\gamma_1, \ldots, \gamma_k)$ for the event that the contours $\gamma_1, \ldots, \gamma_k$ are all present as negative contours.

Lemma 12. *Suppose that $h \geqq 0$, then for every $\Gamma \in \mathscr{F}$, every boundary condition ξ, and every collection of distinct contours $\gamma_1, \ldots, \gamma_k$ inside Γ,*

$$\mu_{\Gamma,\xi,h}(\Omega^-(\gamma_1, \ldots, \gamma_k)) \leqq \exp\left(-\beta \sum_{j=1}^{k} |\gamma_j|\right).$$

Proof. First we consider the case in which there is a single contour, i.e., the case $k = 1$. Define the following sets:

$$\Gamma' := \Theta(\gamma_1) \,,$$
$$\Gamma'' := \Gamma' \backslash \partial_{\text{int}}(\gamma_1) \,.$$

From the fact that every configuration in $\Omega^-(\gamma_1)$ has the property that its restriction to $\partial_{\text{ext}}\gamma_j$ is identically $+1$ and the Markov property of the Gibbs measures we obtain, after conditioning on what the configuration restricted to the complement of Γ' is

$$\begin{aligned}
\mu_{\Gamma,\xi,h}(\Omega^-(\gamma_1)) &\leqq \mu_{\Gamma',+,h}(\Omega^-(\gamma_1)) \\
&= \frac{Z_{\Gamma',+,h}(\Omega^-(\gamma_1))}{Z_{\Gamma',+,h}} \leqq \frac{Z_{\Gamma',+,h}(\Omega^-(\gamma_1))}{Z_{\Gamma',+,h}(\Omega^{-\prime}(\gamma_1))} \,,
\end{aligned} \tag{53}$$

where we used the notation (19), and similarly to what was done in the proof of Lemma 3, $\Omega^{-\prime}(\gamma_1)$ is the set of configurations that can be obtained from configurations in $\Omega^-(\gamma_1)$ by flipping all the spins which are surrounded by γ_1. The observation that every configuration in $\Omega^-(\gamma_1)$ has the property that its restriction to $\partial_{\text{int}}\gamma_1$ is identically -1 and the fact that $h \geqq 0$ lead now to

$$\mu_{\Gamma,+,h}(\Omega^-(\gamma_1)) \leqq \exp(-\beta|\gamma_1|)\frac{Z_{\Gamma'',-,h}}{Z_{\Gamma'',+,h}}.$$

The proof, in the case $k = 1$, will be finished once we show that

$$g(h) := \frac{Z_{\Gamma'',-,h}}{Z_{\Gamma'',+,h}} \leqq 1, \tag{54}$$

whenever $h \geqq 0$. This follows from the facts that $g(0) = 1$ and, using the FKG-Holley inequalities,

$$\frac{d}{dh}\log(g(h)) = (\beta/2)\left(\left\langle \sum_{x\in\Gamma''} \sigma(x)\right\rangle_{\Gamma'',-,h} - \left\langle \sum_{x\in\Gamma''} \sigma(x)\right\rangle_{\Gamma'',+,h}\right) \leqq 0.$$

The proof is now complete when $k = 1$. The general case follows from that one by induction on k. The inductive step from $k-1$ to k is obtained by conditioning on the presence of $k-1$ of the negative contours, and realizing that, thanks to the DLR equations, (9), this conditioning transforms the Gibbs measure in Γ into a new Gibbs measure on a smaller set, with some boundary condition. After this, one just uses the result in the case of a single contour to finish the argument. (This argument can only be carried out if the contours are compatible, but if this is not the case, the lemma is trivially true.) q.e.d.

Theorem 3 follows easily from Lemmas 9, 10 and 11 and the FKG-Holley inequalities. We choose strictly monotone sequences as follows. First take arbitrarily $B_n \searrow 2d$. Now take $\beta_n \nearrow \infty$, so that $B_n m^*(T_n) > 2d$, where $T_n = 1/\beta_n$. Observe that $(\beta_n/2 - \log b - \log 2)/(2\beta_n d) \nearrow 1/(4d)$, so that we can pick $\alpha > 0$ such that $\alpha < (\beta_n/2 - \log b - \log 2)/(2\beta_n d)$ for all n. The event $\mathscr{B}$ has a non-increasing indicator function, and therefore, when $B \geqq B_n$, we have

$$\mu_{\Lambda(B/h),-,h}(\mathscr{B}) \leqq \mu_{\Lambda(B_n/h),-,h}(\mathscr{B}). \tag{55}$$

Now we can conclude from the three lemmas quoted above that for all large n,

$$\lim_{h\searrow 0} \mu_{\Lambda(B/h),-,h}(\mathscr{B}) = 0,$$

provided $\beta \geqq \beta_n$ and $B \geqq B_n$. This implies Theorem 3.

3-ii. Relaxation Times for Finite Systems. In this subsection we will derive an upper bound on the time needed for the process $(\sigma^-_{\Lambda(B/h),-,h;t})$ to reach equilibrium. A very cheap, but non-optimal estimate can be obtained by using the basic coupling between $(\sigma^-_{\Lambda(B/h),-,h;t})$ and $(\sigma^{\mu_{\Lambda(B/h),-,h}}_{\Lambda(B/h),-,h;t})$. One can readily see that these processes will hit each other in a time not larger than an exponential of the volume of the box $\Lambda(B/h)$, so that the relaxation time cannot be larger than an exponential of $1/h^d$.

The correct order of magnitude of the relaxation time of $(\sigma^-_{\Lambda(B/h),-,h;t})$, when $B > 2d$ and T is small, is actually an exponential of $1/h^{d-1}$; a lower bound of this form is contained in part (i) of Corollary 1, which was proven in Sect. 2, and an upper bound of the correct order will be obtained below. Our approach will be via an estimate of the gap in the spectrum of the generator of this process. To estimate this quantity, we will make use of an approach which received considerable attention recently, after Jerrum and Sinclair used it in [JS] and [SJ]. This approach is based on the introduction (in an arbitrary fashion) of oriented paths connecting each pair of states, and on an estimate of "how much any individual transition is essential for the system to relax to equilibrium." In [DiS], Diaconis and Stroock exploited this approach and compared it to others for obtaining estimates on relaxation times of a variety of reversible Markov processes. They also provided a greatly simplified derivation of an inequality which bounds the gap from below in terms of quantities computed from the assignment of paths on the state space. The authors in that paper provided some historic background on the use of similar techniques, which go back to Poincaré! In [Sin] the method and its applications are also discussed in detail, and the same simplified derivation presented in [DiS] (but with a different choice for a certain estimate) is provided. For our purposes the best bound turned out to be the one in [Sin], which will be rederived below, in our setting.

The results in this subsection will be stated and proved (at no extra cost) with a little bit more generality than we need in this paper. We consider a set $\Gamma \in \mathscr{F}$, which later will be considered to be a cube, and an arbitrary boundary condition $\xi \in \Omega$. To avoid a notation which would be too cumbersome, we will use the following abbreviations in this subsection: $X = \Omega_{\Gamma,\xi}$ will be the state space of the process $(\sigma_t)_{t \geqq 0} = (\sigma_{\Gamma,\xi,h;t})_{t \geqq 0}$, whose generator $L = L_{\Gamma,\xi,h}$ acts on each observable f as $(Lf)(\sigma) = \sum_{x \in \Gamma} c(x, \sigma)(f(\sigma^x) - f(\sigma))$. Here $c(x, \sigma) = c_{\Gamma,\xi,h}(x, \sigma)$ are the flip rates of the process, which are supposed to satisfy the detailed balance conditions below with respect to the Gibbs measure $\mu = \mu_{\Gamma,\xi,h}$ (see 14),

$$\mu(\sigma)c(x, \sigma) = \mu(\sigma^x)c(x, \sigma^x) ,$$

for arbitrary x and σ. Moreover we suppose also (see H(4)) that for all Γ, ξ, small $|h|$, x and σ,

$$c(x, \sigma) \geqq c_{\min}(T) = c_{\min} > 0 . \tag{56}$$

We introduce now the standard Hilbert space $L^2(X, \mu)$ with inner product $(\cdot,\cdot)_\mu$ given by

$$(f, g)_\mu := \sum_{\eta \in X} f(\eta)g(\eta)\mu(\eta) ,$$

and corresponding norm $\|f\|_\mu := (f, f)_\mu$. The Dirichlet form associated to the generator L is

$$\mathscr{D}(f, g) := (f, Lg)_\mu = (Lf, g)_\mu ,$$

where the last equality – the self-adjointness of L – follows from the detailed balance condition. Using this self-adjointness one can also write

$$\begin{aligned}\mathscr{D}(f, g) &= (1/2)((f, Lg)_\mu + (Lf, g)_\mu) \\ &= -(1/2) \sum_{x \in \Gamma} \sum_{\eta \in X} c(x, \eta)(f(\eta^x) - f(\eta))(g(\eta^x) - g(\eta))\mu(\eta) .\end{aligned} \tag{57}$$

From (57), we see that $\mathscr{D}(f,f) \leqq 0$, for each observable f, therefore the spectrum of $-L$ is contained in $[0, \infty)$. Due to (56), the process (σ_t) is irreducible and hence 0 is a simple eigenvalue, corresponding to the constant eigenfunctions. The gap in the spectrum of L is now defined as the smallest positive eigenvalue of $-L$, and will be denoted by gap(L). A standard spectral decomposition of the semigroup $S(t) := \exp(tL)$ leads to the well known L^2-exponential convergence estimate below, valid for every observable f and time t,

$$\left[\sum_{\eta \in X} ((S(t)f)(\eta) - \int f d\mu)^2 \mu(\eta)\right]^{1/2} = \| S(t)f - \int f d\mu \|_\mu$$
$$\leqq \| f - \int f d\mu \|_\mu e^{-\text{gap}(L)t} . \quad (58)$$

From this estimate, one can obtain, for an arbitrary initial configuration η, the following estimate:

$$|\mathbb{E}(f(\sigma_t^\eta)) - \int f d\mu| = |(S(t)f)(\eta) - \int f d\mu|$$
$$= \left[((S(t)f)(\eta) - \int f d\mu)^2 \frac{\mu(\eta)}{\mu(\eta)}\right]^{1/2}$$
$$\leqq (\mu(\eta))^{-1/2}\left[\sum_{\zeta \in X} ((S(t)f(\zeta) - \int f d\mu)^2 \mu(\zeta)\right]^{1/2}$$
$$\leqq (\mu(\eta))^{-1/2} \| f - \int f d\mu \|_\mu e^{-\text{gap}(L)t} . \quad (59)$$

The factor $(\mu(\eta))^{-1/2}$ in the upper bound provided by (59) may be quite large, but will not be much of a problem in our case, in spite of the fact that we want to take $\Gamma = \Lambda(B/h)$, which blows up as $h \searrow 0$. Actually the very rough bound

$$(\mu(\eta))^{-1/2} \leqq \left(\frac{\exp(-\beta H_{\Gamma,\xi,h}(\eta))}{2^{|\Gamma|}\max_{\zeta \in X}\exp(-\beta H_{\Gamma,\xi,h}(\zeta))}\right)^{-1/2}$$
$$\leqq e^{C(\beta)|\Gamma|} , \quad (60)$$

uniformly in small $|h|$, ξ and η, will be good enough for us. This is so because we are concerned with times of the order of exponentials of $1/h^{d-1}$ and the bound (59) provides an exponential decay with time, which will beat the exponential growth of $(\mu(\eta))^{-1/2}$ with $1/h^d$, thanks to Theorem 5 below.

The main estimate in this subsection is contained in the next theorem, where we make use of the following standard notation: $O(l^{d-2})$ is a quantity that satisfyes $\limsup_{l \to \infty} O(l^{d-2})/l^{d-2} < \infty$. (35) follows immediately from Lemma 7, part (ii) of Theorem 2 and the combination of Theorem 5 below (with $\Gamma = \Lambda(B/h)$ and $\xi = -\underline{1}$) with (59) and (60) above (both with these same choices of Γ and ξ and with $\eta = -\underline{1}$).

Theorem 5. *Suppose that $d \geqq 2$ and that Γ is a cube of side-length l, then for every boundary condition ξ and every value of h*

$$\text{gap}(L) \geqq c_{\min} l^{-d} \exp(-\beta(2l^{d-1} + O(l^{d-2}))) .$$

Proof. First we recall the well known variational characterization of gap(L) (for proofs see [DiS] and reference quoted there, or Theorem 2.3 in [Lig 2]),

$$\mathrm{gap}(L) = \inf\left\{\frac{-\mathscr{D}(f,f)}{(f,f)_\mu} : f \text{ is non-constant}, \sum_{\eta\in X} f(\eta)\mu(\eta) = 0\right\}.$$

Next we introduce the flux in equilibrium from η to ζ

$$Q(\eta,\zeta) := \begin{cases} \mu(\eta)c(x,\eta) & \text{if } \zeta = \eta^x, \text{ for some } x\in\Gamma, \\ 0 & \text{otherwise}. \end{cases}$$

Using (57), we can now write

$$\mathrm{gap}(L) = \inf\left\{\frac{\sum_{\eta,\zeta\in X}(f(\eta)-f(\zeta))^2 Q(\eta,\zeta)}{\sum_{\eta,\zeta\in X}(f(\eta)-f(\zeta))^2\mu(\eta)\mu(\zeta)} : f \text{ is non-constant}\right\}. \tag{61}$$

In order to estimate the right-hand side of (61), we think of X as the set of vertices of an oriented graph in which the oriented edges are the pairs (η,ζ) such that $\zeta = \eta^x$ for some $x\in\Gamma$. We will denote by $\mathscr{E}$ the set of all these oriented edges. Now, for each pair of distinct configurations $\eta,\zeta\in X$, we introduce an oriented path of configurations $\eta = \psi_1, \psi_2, \ldots, \psi_n = \zeta$, so that $(\psi_i, \psi_{i+1})\in\mathscr{E}$ for $i = 1,\ldots, n-1$. This is done in the following fashion: first we order the sites in the cube Γ in the lexicographic order, i.e., $x < y$ in case for some j, $x_i = y_i$ for $i = 1,\ldots, j-1$ and $x_j < y_j$. The configuration ψ_i is now obtained, recursively, by flipping the spin at the smallest site where ψ_{i-1} differs from ζ. This procedure is iterated until ζ is reached. The resulting path $\psi_1,\ldots,\psi_n$ from η to ζ may be identified with the set of oriented edges $\{(\psi_1,\psi_2),(\psi_2,\psi_3),\ldots,(\psi_{n-1},\psi_n)\}$, which we will denote by $\pi_{\eta,\zeta}$. We introduce now the following "measures of bottleneckness"

$$\bar{\rho} := \max_{e\in\mathscr{E}} \frac{\sum_{\pi_{\eta,\zeta}\ni e}\mu(\eta)\mu(\zeta)|\pi_{\eta,\zeta}|}{Q(e)}, \tag{62}$$

$$\rho := \max_{e\in\mathscr{E}} \frac{\sum_{\pi_{\eta,\zeta}\ni e}\mu(\eta)\mu(\zeta)}{Q(e)}, \tag{63}$$

where $Q(e)$ has the obvious meaning: $Q(e) = Q(\eta,\zeta)$, where $(\eta,\zeta) = e$, and $|\pi_{\eta,\zeta}|$ is the number of oriented edges in the path $\pi_{\eta,\zeta}$. From (61) one can obtain the following inequality:

$$\mathrm{gap}(L) \geqq 1/\bar{\rho}. \tag{64}$$

This inequality is Proposition 1′ in [DiS] and Theorem 5 in [Sin]. Since its derivation is short and sweet, we reproduce it here, for the reader's benefit. Given an edge e, define e^- and e^+ as the configurations such that $e = (e^-, e^+)$. Using the Cauchy–Schwarz inequality, we can rewrite the denominator in (61) as

$$\begin{aligned}\sum_{\eta,\zeta\in X}(f(\eta)-f(\zeta))^2\mu(\eta)\mu(\zeta) &= \sum_{\eta,\zeta\in X}\left[\sum_{e\in\pi_{\eta,\zeta}}(f(e^+)-f(e^-))\right]^2\mu(\eta)\mu(\zeta)\\ &\leqq \sum_{\eta,\zeta\in X}\left[|\pi_{\eta,\zeta}|\sum_{e\in\pi_{\eta,\zeta}}(f(e^+)-f(e^-))^2\right]\mu(\eta)\mu(\zeta)\end{aligned}$$

$$= \sum_{e \in \mathscr{E}} (f(e^+) - f(e^-))^2 \sum_{\pi_{\eta,\zeta} \ni e} |\pi_{\eta,\zeta}| \mu(\eta)\mu(\zeta)$$

$$\leqq \sum_{e \in \mathscr{E}} (f(e^+) - f(e^-))^2 Q(e)\bar{\rho}$$

$$= \bar{\rho} \sum_{\eta,\zeta \in X} (f(\eta) - f(\zeta))^2 Q(\eta, \zeta) . \tag{65}$$

Comparing the right-hand side of (65) with the numerator in (61), we obtain (64). Observe now that the maximum length that a path $\pi_{\eta,\zeta}$ can have is $|\Gamma| = l^d$. This observation allows us to derive from (64),

$$\mathrm{gap}(L) \geqq l^{-d}/\rho . \tag{66}$$

In order to estimate ρ now we will make use of another basic technique introduced in [JS] and [SJ], the so-called "injective mapping" technique. Suppose given an edge $e \in \mathscr{E}$, and suppose also that z is the site such that e^+ is obtained from e^- by flipping the spin at z. Keeping e fixed, we will introduce now a mapping F_e: $\{\pi_{\eta,\zeta}: e \in \pi_{\eta,\zeta}\} \to X$ defined by

$$(F_e(\pi_{\eta,\zeta}))(x) = \begin{cases} \eta(x) & \text{if } x < z , \\ \zeta(x) & \text{otherwise} . \end{cases} \tag{67}$$

The configuration $F_e(\pi_{\eta,\zeta})$ may be thought of as a "negative image" of the configuration e^-, since

$$e^-(x) = \begin{cases} \zeta(x) & \text{if } x < z , \\ \eta(x) & \text{otherwise} . \end{cases} \tag{68}$$

The mapping F_e has the crucial property of being injective; this follows from the fact that e^- and z are fixed, and hence, using (67) and (68), we can reconstruct η and ζ from the knowledge of $F_e(\pi_{\eta,\zeta})$. The other relevant property of the mapping F_e is expressed by the following inequality:

$$\mu(\eta)\mu(\zeta) \leqq \mu(e^-)\mu(F_e(\pi_{\eta,\zeta}))\exp(\beta(2l^{d-1} + O(l^{d-2}))) . \tag{69}$$

This inequality can be derived by using the definition

$$\mu(\sigma) = \mu_{\Gamma,\xi,h}(\sigma) = \frac{\exp(-\beta H_{\Gamma,\xi,h}(\sigma))}{\sum_{\sigma' \in \Omega_{\Gamma,\xi}} \exp(-\beta H_{\Gamma,\xi,h}(\sigma'))} ,$$

for each one of the four configurations which appear there, and then cancelling common terms in the right- and left-hand sides, using (67) and (68). In this fashion one is left with

$$\begin{aligned} \mu(\eta)\mu(\zeta) &\leqq \mu(e^-)\mu(F_e(\pi_{\eta,\zeta})) \\ &\quad \times \exp\left((\beta/2) \sum_{x < z \leqq y} (\eta(x)\eta(y) + \zeta(x)\zeta(y) - \eta(x)\zeta(y) - \zeta(x)\eta(y))\right) \\ &\leqq \mu(e^-)\mu(F_e(\pi_{\eta,\zeta}))\exp\left((\beta/2)4 \sum_{x < z \leqq y} \mathbb{I}_{\{\|x-y\|_1 = 1\}}\right) . \end{aligned}$$

The inequality (69) follows now from the simple estimate

$$|\{\{x, y\} \in \Gamma \times \Gamma: x < z \leqq y \text{ and } \|x - y\|_1 = 1\}| \leqq l^{d-1} + O(l^{d-2}) . \tag{70}$$

The inequality (70) results from the fact that we are using the lexicographic order on Γ and it can be proved by induction on the dimension, as follows. If $d = 1$, (70) is clearly satisfied. Next we will use the notation w' for the vector $(w_2, \ldots, w_n) \in \mathbb{Z}^{d-1}$, when $w = (w_1, \ldots, w_n) \in \mathbb{Z}^d$, and $\Gamma' := \{x' : x \in \Gamma\}$. There are two cases to consider. If $x < z \leqq y$, $\|x - y\|_1 = 1$ and $x_1 \neq y_1$, then $x_1 = y_1 - 1$ and $x' = y'$. It is clear that for each value of $x' = y'$ there cannot be more than one such pair $\{x, y\}$ and hence in this case we obtain at most $l - 1$ pairs. Now, if $x < z \leqq y$, $\|x - y\|_1 = 1$ and $x_1 = y_1$, then also $z_1 = x_1 = y_1$ and the number of pairs $\{x, y\}$ with the desired properties is equal to the number of pairs $\{x', y'\} \in \Gamma' \times \Gamma'$ such that $x' < z' \leqq y'$, and the inductive proof of (70) is complete.

From the definition (63), the inequality (69), the definition of $Q(e)$ and the inequality (56), we obtain

$$\begin{aligned}
\rho &= \max_{e \in \mathscr{E}} \frac{\sum_{\pi_{\eta,\zeta} \ni e} \mu(\eta)\mu(\zeta)}{Q(e)} \\
&\leqq \max_{e \in \mathscr{E}} \frac{\sum_{\pi_{\eta,\zeta} \ni e} \mu(e^-)\mu(F_e(\pi_{\eta,\zeta})) \exp(\beta(2l^{d-1} + O(l^{d-2})))}{\mu(e^-) c_{\min}} \\
&\leqq \frac{\exp(\beta(2l^{d-1} + O(l^{d-2})))}{c_{\min}} \max_{e \in \mathscr{E}} \sum_{\pi_{\eta,\zeta} \ni e} \mu(F_e(\pi_{\eta,\zeta})) \\
&\leqq \frac{\exp(\beta(2l^{d-1} + O(l^{d-2})))}{c_{\min}},
\end{aligned} \tag{71}$$

where in the last step we used the fact that for each fixed e, the mapping F_e is injective. The proof of Theorem 5 is completed by combining the estimates (66) with (71). q.e.d.

(35) can now be proved in the fashion pointed out before we stated Theorem 5. This finishes the proof of part (ii) of Theorem 1.

Other lower bounds for gap(L) are available in terms of ρ and other similar objects. For instance, [JS] and [SJ] derive from a Cheeger-type inequality a bound of the form gap$(L) > C/\rho^2$. In our case this inequality would give a weaker result, in that the value of $\lambda_2(T)$ would be doubled. We were not able to improve the bound that we obtained in Theorem 5, but we observe that from the heuristics, one predicts for the logarithm of the relaxation time of the system in the box $\Lambda(B/h)$, with $B > 2d$ at low temperatures and small external field a value which is smaller than the one obtained here by a factor $((d-1)/d)^{d-1}$.

4. Further Comments, Related Results and Directions for Future Research

There are obviously two challenging and important problems left open here: Extend the results in this paper up to the critical temperature, at least in 2 dimensions and sharpen the results by showing that actually $\lambda_1(T) = \lambda_2(T)$, preferentially expressing their common value in terms of other quantities, such as equilibrium Wulff shapes (if it is the case that they are indeed related). Below we present a few other problems which seem to us to be particularly interesting (and on which we are working). We also discuss further the relation of the present paper with some other works on metastability and describe briefly some related results, to be published elsewhere.

Different asymptotic regimes. The results that we proved in this paper are always in the form of asymptotics for positive h, when this external field vanishes. This idea that metastability phenomena should be mathematically described by considering families of processes, indexed by a parameter, and scaling the parameter to zero is not at all new. For fixed values of the parameters h and T, the stochastic Ising models do not display any clear cut, sharp, metastable behavior, but in certain limits, as the one considered here, the behavior of the system becomes closer and closer to what one identifies experimentally as metastable behavior. To some extent this is akin to many other situations in mathematical physics, in which one proves results in the form of limits, with the motivation of understanding the behavior of the system when the scaled parameter is actually fixed (but small or large enough, depending on the case). The thermodynamic limit is certainly an example which comes to mind.

We will refer below to the type of limit considered in this paper (T small fixed, $h \searrow 0$) as limit type (i). As we mentioned in the introduction, sharp results on droplet behavior and metastable behavior were obtained recently in the different regime, that we are calling "the limit of very low temperature" ($h > 0$ small fixed, $T \searrow 0$), to which we will refer as limit of type (ii). In 2 dimensions, these results include an understanding of the mechanism by which these droplets grow, and for this reason provide results on the metastable behavior of finite systems which are quite sharp. Because the size of critical droplets scales with h, but not with T, it makes sense in the case of limit (ii) to consider the system in a box $\Lambda(N)$ with fixed N (large compared with $2/h$), and in this case the metastable behavior and its decay were analyzed in great detail. Results for the infinite system, of the type of those obtained for limit (i) in this paper can also be obtained in the case of limit (ii). In this case one wants to look at the system at time $t = \exp(C\beta)$, for different values of C. One can indeed show that for small values of C (depending on h) one sees locally all spins down and for large values of C, all spins up. Efforts to identify a single critical value $C(h)$, separating the two regimes have failed so far, because it is hard to control the way the speed of growth of the supercritical droplets behaves asymptotically, when the droplet becomes very large. (The analogue of (20) seems to fail here and one has to find the value of $\lim_{\beta\to\infty}(1/\beta)\log v$.)

Something very interesting happens when one considers a third type of limit. Once one accepts as natural to scale $h \searrow 0$ or $T \searrow 0$, it becomes also reasonable to ask what happens if we let both vanish together. It turned out that in this regime, we obtained some results which are sharper than in the two other cases, in part because we could use techniques from both cases. The way in which T and h vanish is relevant here, and the analysis is easier if we let T $\searrow 0$ much faster than $h \searrow 0$, on the other hand the case in which we keep a constant ratio between h and T is particularly relevant, because, via a simple transformation, this is equivalent to keeping the temperature and external fields constant, while scaling only the coupling between spins to ∞. We call this type of limit, in which $h \searrow 0$ and $T \searrow 0$, while h/T stays constant, limit of type (iii). In this limit one wants to look at the system at a time of the form $t = \exp(\kappa\beta/h^{d-1})$. In 2 dimensions, for certain dynamics, including Heat Bath and Metropolis, we can prove that if $\kappa < 4/3$ then, in the limit, one sees locally all spins down, while if $\kappa > 4/3$ one sees locally all spins up. The value 4/3 is exactly the value predicted from the heuristics:

$2^d(d-1)^{d-1}/(d+1)$. The result in case $\kappa < 4/3$ follows from the techniques in Sect. 2 in this paper: just observe that all the estimates there become better and better as T becomes smaller and smaller. The other result relies on a careful study of the behavior of individual droplets.

From the results obtained previously on limit (ii), the ones presented in this paper on limit (i) and the one mentioned above on limit (iii), it is becoming clear that one should try to explore the relaxation patterns of stochastic Ising models parametrized by 3 quantities: h, T and the sidelength N of the box in which the system is contained (including the case $N = \infty$). From the heuristics it is usually possible to predict the correct behavior in different regimes, but there is still a substantial distance between most of these heuristic results and their rigorous counterparts.

Pathwise approach. One of the main motivations for the investigation of the limit of type (ii) in [NS1] and [NS2] was the possibility of proving that in this limit finite systems display metastable behavior in a pathwise sense (as introduced in (CGOV)). Roughly speaking, if one looks not at an average over many realizations of the evolution, as we are doing here when we take $\mathbb{E}(\cdot)$, but at a single evolution, then one should see a very sharp transition between the metastable and the stable situations. The time taken to make the transition is very short compared with the time spent in the metastable situation, so that essentially the jump is "instantaneous." The moment of the jump is nevertheless random and for this reason, when one considers an average over many realizations, one sees a much smoother evolution. For finite systems in two dimensions, this was indeed proved to be the case in the limit of type (ii), including the feature that the time of the jump is, in the proper scale, close to an exponential random variable, because it corresponds to an essentially local fluctuation. It is very interesting to see that the distinction between pathwise behavior and average behavior was also realized, apparently independently, by investigators performing simulations. In the paper [TM] the authors emphasize this distinction and observe experimentally the type of behavior predicted from limit (ii) (I am thankful to R. Kotecký for telling me about this paper; see also other references quoted in [TM]).

A detailed pathwise study of the behavior of stochastic Ising models in the limit (i) remains a challenging open problem. From the heuristics one can actually predict two different behaviors for typical paths, depending on how the size of the system grows as $h \searrow 0$. If N grows slowly, as in Corollary 1, then, when a supercritical droplet is first formed, it should cover the system in a relatively short time. In this case we should see the same type of pathwise metastable behavior described above, even if we are observing the whole system. But if N is growing as $\exp(D/h^{d-1})$, with a large D (depending on T), then while a supercritical droplet is growing, others are being formed somewhere else, and start growing and eventually coalesce. Global quantities (as, e.g., the space average of the spins) should behave quite smoothly, even for a single realization of the process. In contrast local observables should still display a jump in their pathwise behavior, reflecting the moment when they are first covered by a supercritical droplet. The only difference in this case, with regard to the smaller systems should be that the rescaled time of the jump should have a distribution that while not degenerate into a constant, should neither be an exponential. This follows from the consideration of the regions in space-time (the cones considered before) where droplets have to be formed, to cover a certain site at a certain time. (But the analysis is actually complicated by the interaction between droplets, when they touch each other.)

Gap in the spectrum, boundary conditions and droplet growth. From the discussion above it should be clear that one of the first questions to address is that of the growth of the supercritical droplets in the limit of type (i). If we admit that this growth is of the type discussed in the heuristics, similar to the movement of essentially flat interfaces, then we should try to understand this mechanism in situations which are not so complex. A simpler situation in which such a movement of a flat interface should occur is that which is produced if we have a box $\Lambda(B/h)$, with B large enough (so that in equilibrium the plus-phase dominates in the bulk), but with the boundary condition in which the spins are -1 on all the faces of $\Lambda(B/h)$, but one of them, where they are $+1$. In this situation, starting from all spins down, we should see the interface move away from this $+$ face, in a fashion similar to what we expect to happen at the border of a large droplet. From the heuristic prediction of a speed $v \sim h$ for the interface, we should expect a time of order h^{-2} for the system to relax to equilibrium. The gap in the spectrum of the generator in this case should therefore behave as h^2, as $h \searrow 0$. Proving this result (even a lower bound for the gap in the form of some large power of h) would probably be a great step towards controlling the growth of the droplets.

Curiously enough, the technique used in Subsect. 3-ii, for bounding from below the gap in the spectrum for the system in the same box, but with purely $-$ boundary condition, gives in the present case the same type of lower bound, of order $\exp(-C/h^{d-1})$. This is the correct type of result when the boundary condition is purely $-$ (as can be seen by taking in (61) f as the indicator function of $\mathscr{R}_{A/h}$, with A given by (31), so that (32) holds, and implies that the gap indeed goes to 0 exponentially with $1/h^{d-1}$), but is probably a very poor one in our present case. Even in the case in which the boundary condition is purely $+$ for the same box, it would be interesting to obtain good lower bounds for the gap.

5. Appendix: Alternative Proof of Theorem 3 in Two Dimensions

In this appendix we will present an alternative proof of Theorem 3 in two dimensions. Lemma 8 will be needed in the proof, and we suppose that the reader remembers the discussion which followed the proof of that lemma. Suppose that, as in that discussion, we have chosen three strictly monotone sequences, $B_n \searrow 2d = 4$, $\beta_n \nearrow \infty$ and $A_n \nearrow 2d = 4$, with the properties

$$B_n m^*(T_n) > 2d = 4 , \tag{72}$$

and

$$A_n < 2d(\beta_n'/\beta_n) = 4(\beta_n'/\beta_n) , \tag{73}$$

where $T_n = 1/\beta_n$ and $\beta_n' = \beta_n - \log b$. We know that for large n, it is very likely that when h is small, in typical configurations chosen according to the law $\mu_{\Lambda(B_n/h),-,h}$, at inverse temperature $\beta \geqq \beta_n$, there is a (unique) contour γ which surrounds a large fraction of $\Lambda(B_n/h)$. The next lemma below implies that at the same time, such a contour is unlikely to have a surface which is substantially larger than the surface of $\Lambda(B_n/h)$, in the sense of differing from this surface by a fraction of it. This lemma is dimension independent, but we will only be able to exploit it for our purposes if $d = 2$. The reason is that the presence of a contour as the γ above and the event $\mathscr{B}$ can occur together with the surface of γ being larger than the surface of $\Lambda(B_n/h)$ by a term of the order of $1/h$ only (think of a finger of spins $-$ penetrating a cube full of spins $+$), which only in $d=2$ is a fraction of the surface of $\Lambda(B_n/h)$. On the other hand, in $d=2$ we can show that such a difference in surfaces is indeed a necessary consequence of the simultaneous occurrence of those two events (this is the content of Lemma 14 below), and then complete the argument.

We will denote by $\mathscr{S}_l$ the set of configurations in which there is at least one contour with surface not smaller than l.

Lemma 13. *In every dimension, if $\beta > \log b$, then for arbitrary B and D,*

$$\limsup_{h \searrow 0} h^{d-1} \log \mu_{\Lambda(B/h),-,h}(\mathscr{S}_{D/h^{d-1}}) \leqq \beta((1/2)(1-m^*(T))B^d + 2dB^{d-1} - (\beta'/\beta)D) .$$

Proof. The proof is somewhat similar to that of Lemma 8. First write, using the notation (19),

$$\mu_{\Lambda(B/h),-,h}(\mathscr{S}_{D/h^{d-1}}) = \frac{Z_{\Lambda(B/h),-,h}(\mathscr{S}_{D/h^{d-1}})}{Z_{\Lambda(B/h),-,0}(\mathscr{S}_{D/h^{d-1}})} \frac{Z_{\Lambda(B/h),-,0}(\mathscr{S}_{D/h^{d-1}})}{Z_{\Lambda(B/h),-,0}} \frac{Z_{\Lambda(B/h),-,0}}{Z_{\Lambda(B/h),-,h}} . \tag{74}$$

The first fraction in the right-hand side of (74) can be estimated by the following inequality, which does not depend on what $\mathscr{S}_{D/hd-1}$ is,

$$\log\left(\frac{Z_{\Lambda(B/h),-,h}(\mathscr{S}_{D/h^{d-1}})}{Z_{\Lambda(B/h),-,0}(\mathscr{S}_{D/h^{d-1}})}\right) \leqq (\beta h/2)|\Lambda(B/h)| = \beta(1/2)B^d/h^{d-1} . \tag{75}$$

The second fraction in the right-hand side of (74) can be estimated, with the usual Peierls type of argument (i.e., Lemma 3 with $h=0$ and $c=\infty$, combined with the counting inequality (4)), by

$$\limsup_{h \searrow 0} h^{d-1} \log\left(\frac{Z_{\Lambda(B/h),-,0}(\mathscr{S}_{D/h^{d-1}})}{Z_{\Lambda(B/h),-,0}}\right) \leqq -\beta' D . \tag{76}$$

The lemma follows now from (74), (75), (76) and (41). q.e.d.

In the next lemma the fact that we are taking the D above as $17/h$ is motivated by the observation that this quantity is between the surface of $\Lambda(4/h)$ (which is $16/h$) and the surface, $18/h$, of the contour in the configuration in $\Omega_{\Lambda(4/h),-}$ which is identically $+1$ inside of $\Lambda(4/h)$, except for a straight line of spins -1 which connects $\Lambda(2/h)$ to the exterior of $\Lambda(4/h)$. The reader should keep this picture in mind when we come to point (iii) in the proof below.

Lemma 14. *Suppose that $d=2$, and that we are given two monotone sequences $B_n \searrow 2d = 4$ and $A_n \nearrow 2d = 4$. Then for large enough n, the following holds for all $h > 0$,*

$$\mathscr{B} \cap \Omega_{\Lambda(B_n/h),-} \subset \mathscr{R}_{A_n/h} \cup \mathscr{S}_{17/h} .$$

Proof. For each n and small positive δ (to be actually chosen later as $1/10$), we introduce the four distinct squares, W_n^1, W_n^2, W_n^3 and W_n^4 which are defined by saying that each one of them has side-length δ/h, is contained in $\Lambda(B_n/h)$ and shares a vertex with this square. For future reference we will suppose that the indices 1 to 4 are assigned to these squares in such a way that W_n^1 and W_n^3 contain opposite vertices of $\Lambda(B_n/h)$ (i.e., vertices which do not belong to a common side of the square).

The lemma is a consequence of the following three facts:

i) For small fixed δ, if n is so large that

$$\left(\frac{\delta}{h}\right)^2 > \left(\frac{B_n}{h}\right)^2 - \left(\frac{A_n}{h}\right)^2 , \tag{77}$$

then in each configuration in $\Omega_{\Lambda(B_n/h),-} \setminus \mathscr{R}_{A_n/h}$ there is a contour present which surrounds at least one site inside of each one of the squares $W_n^1, \ldots, W_n^4$.

ii) In any configuration in $\mathscr{B}$ there is no contour present which surrounds $\Lambda(2/h)$.

iii) A contour γ which surrounds at least one site inside of each one of the squares $W_n^1, \ldots, W_n^4$ and which does not surround $\Lambda(2/h)$ must have surface at least

$$4\left(\frac{B_n}{h} - 2\frac{\delta}{h}\right) + 2\left(\frac{B_n - 2 - 2\delta}{2h}\right) \geqq \frac{18 - 10\delta}{h} = \frac{17}{h} ,$$

where in the last step we made the choice $\delta = 1/10$. The reader can probably convince herself that this inequality is correct by drawing a few pictures, but may prefer the following formal argument. Think of the contour γ as a circuit in the graph whose vertices are the faces separating unit squares centered at the sites of $\mathbb{Z}^2$, and in which two such faces are adjacent if they have a non-empty intersection. With some abuse of language, we can say that the contour γ has to contain at least one face inside each one of the squares $W_n^1, \ldots, W_n^4$. Call these faces (choosing arbitrarily one if there are many), respectively, $F^1, \ldots, F^4$. There are two cases to be considered. First suppose that walking along the circuit γ, we can cross $F^1, \ldots, F^4$ in this order. In this case break γ into four disjoint parts each one of which connects F^i to F^{i+1}, for some $i = 1, 2, 3$ or F^4 to F^1. Clearly, each one of these parts of γ contains at least $B_n/h - 2\delta/h$ faces. Moreover, because $\Lambda(2/h)$ is not surrounded by γ, at least one of these parts of γ must contain at least $B_n/h - 2\delta/h + 2(B_n - 2 - 2\delta)/(2h)$ faces. This completes the proof in this case. If we cannot cross $F^1, \ldots, F^4$ in this order, then it must be the case that we can walk along γ from F^1 to F^3 without crossing F^2 and F^4 and also from F^2 to F^4 without crossing F^1 and F^3 (the little argument, by contradiction, needed to prove this fact is left to the reader). Moreover, we must also be able to walk along γ from F^1 to F^2, or from F^1 to F^3, without using the part of γ just mentioned. Estimating the minimum number of faces that we have to cross in each one of these three, necessarily disjoint, parts of γ, we obtain a lower bound for the size of γ which is larger than the one that we are claiming. q.e.d.

From Lemmas 8, 13 and 14 we obtain

$$\lim_{h \searrow 0} \mu_{\Lambda(B/h),-,h}(\mathscr{B}) = 0 , \tag{78}$$

if $\beta \geqq \beta_n$, provided that n is so large that all these lemmas can be applied and also

$$(1/2)(1 - m^*(T_n))B_n^2 + 4B_n - 17\beta_n'/\beta_n < 0 .$$

Theorem 3 follows now, in the case $d = 2$, from (78) and (55), which is a consequence of the FKG-Holley inequalities, when $B_n \leqq B$.

Acknowledgements I am especially indebted to two colleagues. One is E.J. Neves for his collaboration in the project related to the limit of very low temperatures, including unpublished results that we derived for infinite systems, which guided by approach in the project presented here The other is F Martinelli for sending me, prior to publication, notes from which I learned of the techniques used by Jerrum and Sinclair for estimating the gap in the spectrum of certain Markov processes and also for a joint effort to understand the work of Martirosyan

I am also very thankful to several other colleagues, particularly to M Aizenman, R Kotecký, J Lebowitz, T Mountford, E Olivieri and D Stauffer, for discussions on the subject of this paper and its possible extensions

References

[AL] Aizenman, M , Lebowitz, J : Metastability effects in bootstrap percolation J. Phys A: Math. General **21**, 3801–3813 (1988)

[BM] Binder, K , Müller-Krumbhaar, H : Investigation of metastable states and nucleation in the kinetic Ising model Phys Rev B **9**, 2328–2353 (1974)

[CCO] Capocaccia, D , Cassandro, M , Olivieri, E : A study of metastability in the Ising model Commun Math Phys. **39**, 185–205 (1974)

[CGOV] Cassandro, M , Galves, A , Olivieri, E., Vares, M E : Metastable behavior of stochastic dynamics: A pathwise approach J Stat. Phys. **35**, 603–634 (1984)

[DS] Devillard, P., Spohn, H : Kinetic shape of Ising clusters. Europhys Lett **17**, 113–118 (1992)

[DiS] Dianconis, P , Stroock, P : Geometric bounds for eigenvalues of Markov chains. Ann Appl Probab **1**, 36–61 (1991)

[GD] Gunton, J D , Droz, M : Introduction to the Theory of Metastable and Unstable States Lecture Notes in Physics **183**, Berlin, Heidelberg, New York: Springer, 1983

[GSS] Gunton, J D , San Miguel, M , Sahni, P.S : The dynamics of first order phase transitions. In: Domb, C , Lebowitz, J L (eds) Phase Transitions and Critical Phenomena New York: Academic Press, 1983, pp 269–482

[JS] Jerrum, M , Sinclair, A : Approximating the permanent SIAM J Comput **18**, 1149–1178 (1989)

[Koc] Koch, S.W : Dynamics of first order phase transitions in equilibrium and nonequilibrium systems Lecture Notes in Physics **207**, Berlin, Heidelberg, New York: Springer, 1984

[KO1] Kotecký, R., Olivieri, E.: Stochastic models for nucleation and crystal growth Proceedings of the International Workshop Probabilistic Methods in Mathematical Physics (Siena, 1991) Guerra, F , Loffredo, M I , Marchioro, C (eds) Singapore: World Scientific, 1992, pp 264–275

[KO2] Kotecký, R., Olivieri, E : Droplet dynamics for asymmetric Ising model J Stat Phys. **70**, 1121–1148 (1993)

[KO3] Kotecký, R , Olivieri, E : Shapes of growing droplets – A model of escape from a metastable phase J Stat. Phys (to appear)

[Lig1] Liggett, T M : Interacting Particle Systems Berlin, Heidelberg, New York: Springer, 1985

[Lig2] Liggett, T M : Exponential L_2 convergence of attractive reversible nearest particle systems Ann Probab **17**, 403–432 (1988)

[Mar] Martirosyan, D G : Theorems on strips in the classical Ising ferromagnetic model Sov J. Contemp Math Anal **22**, 59–83 (1987)

[NS1] Neves, E.J., Schonmann, R H : Critical droplets and metastability for a Glauber dynamics at very low temperatures Commun Math Phys **137**, 209–230 (1991)

[NS2] Neves, E J , Schonmann, R H : Behavior of droplets for a class of Glauber dynamics at very low temperature Probab Theory and Rel Fields **91**, 331–354 (1992)

[PL] Penrose, O , Lebowitz, J L : Towards a rigorous molecular theory of metastability In: Fluctuation Phenomena (second edition) Montroll, E W , Lebowitz, J L (eds) Amsterdam: North-Holland Physics Publishing, 1987

[Rue] Ruelle, D : Statistical Mechanics. Rigorous Results Reading, MA: Benjamin, 1969

[RL] Runnels, L K , Lebowitz, J L : Analyticity of a hard-core multicomponent lattice gas J Stat Phys **14**, 525–533 (1976)

[Sch1] Schonmann, R H : An approach to characterize metastability and critical droplets in stochastic Ising models. Ann Inst Henri Poincaré (Phys. Theor) **55**, 591–600 (1992)

[Sch2] Schonmann, R.H : The pattern of escape from metastability of a stochastic Ising model. Commun Math Phys **147**, 231–240 (1992)
[Sch3] Schonmann, R H : Relaxation times for stochastic Ising models in the limit of vanishing external field at fixed low temperatures In: Cellular Automata and Cooperative Systems NATO ASI Series Series C: Mathematical and Physical Sciences - Vol 396. N Boccara, E. Goles, S Martinez and P Picco, eds. (Kluwer Academic Publisher, 1993), pp 543–546
[Sco] Scoppola, E : Renormalization group for Markov chains and application to metastability J Stat Phys, to appear
[Sin] Sinclair, A : Improved bounds for mixing rates of Markov chains and multicommodity flow Combinatorics, Probability and Computing **1**, 351–370 (1992)
[SJ] Sinclair, A, Jerrum, M.: Approximate counting, uniform generation and rapidly mixing Markov chains Information and Computation **82**, 93–133 (1989)
[Sta] Stauffer, D : Ising droplets, nucleation, and stretched exponential relaxation Int J Mod Phys C **3**, 1052–1070 (1992)
[Tay] Taylor, J E : Crystalline variational problems Bull Am. Math. Soc **84**, 568–588 (1978)
[TM] Tomita, H, Miyashita, S : Statistical Properties of the relaxation processes of metastable states in the kinetic Ising model. Phys Rev B, Condensed Matter **46**, 8886–8893 (1992)
[Van] Vanheuverzwijn, P : Metastable states in the infinite Ising model J Math Phys **20**, 2665–2670 (1979)

Communicated by J L Lebowitz