EURANDOM PREPRINT SERIES
2012-012

May 25, 2012

Relaxation Height in Energy Landscapes: an Application to Multiple Metastable States

Emilio N.M. Cirillo, Francesca R. Nardi
ISSN 1389-2355

Relaxation Height in Energy Landscapes: an Application to Multiple Metastable States

Emilio N.M. Cirillo
Dipartimento di Scienze di Base e Applicate per l'Ingegneria, Sapienza Università di Roma, via A. Scarpa 16, I–00161, Roma, Italy.
E_mail: emilio.cirillo@uniroma1.it

Francesca R. Nardi
Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.
Eurandom, P.O. Box 513, 5600 MB, Eindhoven, The Netherlands.
E_mail: F.R.Nardi@tue.nl

Abstract. The study of systems with multiple (not necessarily degenerate) metastable states presents subtle difficulties from the mathematical point of view related to the variational problem that has to be solved in these cases. We introduce the notion of relaxation height in a general energy landscape and we prove sufficient conditions which are valid even in presence of multiple metastable states. We show how these results can be used to approach the problem of multiple metastable states via the use of the modern theories of metastability. We finally apply these general results to the Blume–Capel model for a particular choice of the parameters ensuring the existence of two multiple, and not degenerate in energy, metastable states.

Keywords: energy landscape, relaxation, metastability, multiple metastable states, Blume–Capel model

Appunti: May 25, 2012

1. Introduction

In many applicative problems one has to consider a stochastic system evolving in a finite not empty state space driven by an energy function. The details of the dynamics depend on the system that one has to model, but in most cases the motion is driven by the energy

This work has been partially done at Eurandom during the Stochastic Activity Month, February 2012. ENMC wants to express his thanks to the organizers, R. Fernandez, R. van der Hofstad, and M. Heydenreich, for the invitation and to Eurandom for the kind hospitality. FRN thanks F. den Hollander and A. Troiani for many interesting discussions related to the topic of the paper.

landscape [26]. The typical trajectories are those decreasing the energy; on the other hand with small probability the system can perform jumps against the energy drift.

The smallness of the probability associated to moves against the energy drift is controlled by some cost function. This cost is often given by the difference of energy between the two states involved in the move. This is the situation that one has to face when studying the evolution of Glauber dynamics associated with Statistical Mechanics lattice models [18]. Depending on the model, different cost functions can be introduced. An example, which deserved a lot of attention in the recent literature, is that of reversible [21] Probabilistic Cellular Automata [3, 12–15, 24].

The stochastic dynamics is often controlled by a parameter, say the *temperature*, whose value tunes the amount of randomness in the motion by increasing or decreasing the probability of the moves against the drift. When the temperature is very low the system freezes and tends to stick to the energy drift dynamics ending up to be trapped in the minima of the energy. When the temperature is very high, on the other hand, jumps against the drift are highly probable and the system moves almost freely in the state space.

It is then clear that a huge amount of information about the low temperature dynamics is obtained when one knows (i) the structure of the absolute minima of the energy (ground states) and (ii) the maximal barrier that has to be overcome to reach the set of the ground states starting from any possible state, namely, the *relaxation height*. The severity of problem (i) strongly depends on the particular expression of the energy function, but in most applications it is not a particularly difficult task. Problem (ii), on the other hand, is always a very laborious (and sometimes hard) one since the solution of several variational problems is involved.

The relaxation height is of basic importance in the study of this kind of dynamics. For instance, when dealing with the description of the metastable behavior of Statistical Mechanics systems, the relaxation height is the quantity controlling the typical time the system waits before nucleating the stable state starting from the metastable one [8, 22]. The problem of computing the relaxation height in specific model is often very hard, in particular when there are several states connected to the ground states via cost barriers equal to the relaxation height. This situation is found, for instance, in models with more than one metastable state [12].

In this paper we shall discuss some generic properties of the relaxation height and, in particular, we shall setup a quite general strategy for its computation. Subsequently we shall apply the theory to the very interesting case of the Blume–Capel model [4, 5, 11, 17] for a choice of the parameter ensuring the existence of two not degenerate in energy metastable states. Due to the presence of multiple metastable states this case is particularly interesting

and, a priori, very difficult to be treated.

The Blume–Capel model has been originally introduced in [4,11] to study some particular magnetic systems. The model was then generalized to the so called Blume–Emery–Griffiths model [5] in order to describe the λ transition in He^3–He^4 mixtures. Those models have been widely studied in the literature and many applications have been considered. Their particular interest is due to fact that in these three state spin systems on lattice different energies are associated with different interfaces. This is not the case in other multi–state spin systems, such as the well known Potts model, where a ferromagnetic coupling favors the presence of neighboring homologous spins and the same energy cost is assigned to any pair of differing neighboring spins.

This fact is very peculiar of the Blume–Capel model and gives rise to very interesting phenomena when the metastable behavior of the system is considered. Indeed it is seen that even if one of the three states is not favored from the energetic point of view, it can (and indeed does) behave as a bearing between the other two phases if the interfaces between this particular state and the other two ones are energetically less expensive (see, for instance, the proof of item 2 in Lemma 4.9 where this effect appears in all its importance).

In [17] the authors studied metastability in the Blume–Capel model in the so called Wentzell–Friedlin regime, that is they considered finite volume, finite positive magnetic field, and let the temperature tend to zero. In that paper a wide region of the space of parameters of the model was analyzed. More precisely, denoted by λ the chemical potential (see equation (4.1) below) and by h the magnetic field, the region $h > \lambda > 0$ was studied in detail. It was proven that in that region the Hamiltonian of the system has a single ground state $\mathbf{u}$, namely, the configuration with all the spins equal to plus one, and that the system admits a single metastable state, that is the configuration $\mathbf{d}$ with all the spins equal to minus one. It was also proven that the configuration $\mathbf{0}$, with all the spins equal to zero, plays an important role. For $2\lambda > h > \lambda$ the transition from the metastable state $\mathbf{d}$ to the stable one $\mathbf{u}$ is direct, although the plus droplet growing inside the minus phase is "protected" by a thin frame of zeros. For $h > 2\lambda$, on the other hand, during the transition the system visits the intermediate (not metastable) state $\mathbf{0}$. This kind of phenomena is not expected to be observed in multi–state spin systems such as the Potts model where all the interfaces pay the same energy cost.

In [23] the study of the metastable behavior of the Blume–Capel model was addressed in infinite volume in the regime characterized by finite λ and h and temperature tending to zero. In that paper the authors studied the regions $0 < -\lambda < h$ and $0 < \lambda < h$ of the space of parameters, see [23, Theorem 1], and proved that, starting from $\mathbf{d}$, in the first one the state $\mathbf{0}$ is observed before the transition to $\mathbf{u}$, while in the second one it is not.

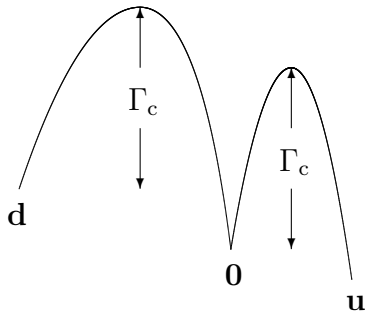

Figure 1.1: Schematic description of the energy landscape for the Blume–Capel model (4.1) at $\lambda = 0$ and $h > 0$. The quantity Γ_c represents the energy barrier that must be overcome to exit the metastable states. In the picture it is remarked that the two metastable states **d** and **0** are not degenerate in energy.

This result, looked at from the finite volume point of view, suggests that at negative λ the state **0** is metastable while at positive λ it is not. The first of these two remarks was discussed, on heuristics grounds, in [16], while the second was indeed proven in [17]. We can then conjecture that in the peculiar case $\lambda = 0$ and $h > 0$ both **d** and **0** are metastable. In other words we expect that for this choice of the parameters the Blume–Capel model shows the very interesting phenomenon of multiple not degenerate in energy metastable states (see figure 1.1).

Situations of this kind have been already studied in the case of Probabilistic Cellular Automata [12], but there an old approach [25] to metastability had been used. In that paper the model dependent study had to be very detailed and, hence, very difficult. In this paper we show that by using the more recent approach to metastability, that will be shortly reviewed at the beginning of Section 4 and the general result on relaxation height in Theorem 2.4 below, the problem can be solved relying on few model dependent properties.

The paper is organized as follows. The general results are discussed in Section 2. Their application to the problem of metastability in Statistical Mechanics systems is discussed in Section 3. In particular, their application to the study of the metastable behavior of the Blume–Capel model is discussed in Section 4. Finally we prove the theorems and the lemmas in Sections 5 and 6.

2. Maximal stability level

In this section we shall discuss some general results related to energy landscapes. Under very general assumptions on the structure of the energy landscape we shall introduce the notion of *maximal stability level* and prove some results yielding a handy recipe for its computation.

2.1. Energy landscape

An *energy landscape* is a quaternion (X, Q, H, Δ) where the finite not empty set X, $Q \subset X \times X$, $H : X \to \mathbb{R}$, and $\Delta : Q \to \mathbb{R}_+$ are respectively called *state space*, *connectivity relation*, *energy*, and *cost function*, and the relation Q on X is such that for any $x, y \in X$ there exist an integer $n \geq 2$ and $x_1, \dots, x_n \in X$ such that $x_1 = x$, $x_n = y$, and $(x_i, x_{i+1}) \in Q$ for any $i = 1, \dots, n-1$.

An energy landscape (X, Q, H, Δ) is called *reversible* if and only if the connectivity relation Q is symmetric and

$$H(x) + \Delta(x, y) = \Delta(y, x) + H(y) \tag{2.1}$$

for all $(x, y) \in Q$.

2.2. Definitions

Consider a reversible energy landscape (X, Q, H, Δ). Given $Y \subset X$ such that $H(y) = H(y')$ for any $y, y' \in Y$, we shall denote by $H(Y)$ the energy of the states in Y. For any $Y \subset X$ we shall denote by $F(Y)$ the set of the minima of the energy inside Y, that is to say $y \in F(Y)$ if and only if $H(y') \geq H(y)$ for any $y' \in Y$. We let $X_{\mathrm{s}} := F(X)$ be the set of *ground states* of H, namely, the set of the absolute minima of the energy.

For any positive integer n, $\omega \in X^n$ such that $(\omega_i, \omega_{i+1}) \in Q$ for all $i = 1, \dots, n-1$ is called a *path* joining ω_0 to ω_n; we also say that n is the length of the path. For any path ω of length n, we let

$$\Phi_\omega := \max_{i=1,\dots,n-1} H(\omega_i) + \Delta(\omega_i, \omega_{i+1}) \tag{2.2}$$

be the *height* of the path. For any $y, z \in X$ we denote by $\Omega(y, z)$ the set of the paths joining y to z and define the *communication height* between y and z as

$$\Phi(y, z) := \min_{\omega \in \Omega(y,z)} \Phi_\omega \tag{2.3}$$

From (2.1), (2.2), and (2.3) it follows immediately that

$$\Phi(y, z) = \Phi(z, y) \tag{2.4}$$

for all $y, z \in X$. For any $Y, Z \subset X$ we let

$$\Phi(Y, Z) := \min_{\omega \in \Omega(Y,Z)} \Phi_\omega = \min_{y \in Y, z \in Z} \Phi(y, z) \tag{2.5}$$

where we have used the notation $\Omega(Y, Z)$ for the set of paths joining a state in Y to a state in Z.

Since the energy landscape is reversible, the energy of the state ω_n is implicitly taken into account in (2.2), indeed (2.1) implies $H(\omega_n) \leq \Delta(\omega_{n-1}, \omega_n) + H(\omega_{n-1})$.

We say that X is *fully attracted* by $X_{\rm s}$ if and only if $\Phi(x, X_{\rm s}) - H(x) = 0$ for any $x \in X \setminus X_{\rm s}$ or (trivial case) $X = X_{\rm s}$.

For any $x \in X$ we denote by I_x the set of states $y \in X$ such that $H(y) < H(x)$. Note that $I_x = \emptyset$ if $x \in X_{\rm s}$. We then define the *stability level* of any $x \in X \setminus X_{\rm s}$

$$V_x := \Phi(x, I_x) - H(x) \geq 0 \tag{2.6}$$

Note that the stability level V_x of x is the minimal cost that, starting from x, has to be payed in order to reach states at energy lower than $H(x)$. Following [22] we now introduce the notion of maximal stability level.

Definition 2.1 *Consider a reversible energy landscape (X, Q, H, Δ). Assume $X \setminus X_{\rm s} \neq \emptyset$, we let the* maximal stability level *be*

$$\Gamma_{\rm m} := \sup_{x \in X \setminus X_{\rm s}} V_x \tag{2.7}$$

We also set

$$X_{\rm m} := \{x \in X \setminus X_{\rm s} : V_x = \Gamma\} \tag{2.8}$$

Note that, since the state space is finite, the maximal stability level $\Gamma_{\rm m}$ is a finite number. Note, also, that if X is fully attracted by $X_{\rm s}$ and $X \setminus X_{\rm s} \neq \emptyset$, then $\Gamma_{\rm m} = 0$ and $X_{\rm m} = X \setminus X_{\rm s}$.

2.3. Results

The notion of maximal stability level has been introduced by looking at the paths starting from any state of X and reaching lower energy states. This point of view is often very useful when dealing with metastability problems [22], indeed the maximal stability level controls the asymptotic of the time the system spends in the metastable state before nucleating the stable one. It is worth noting, on the other hand, that computing the maximal stability level of a concrete model is a hard task. Indeed, one has to solve the variational problem $\Phi(x, I_x)$ for any $x \in X$.

It is then clear the interest of results providing sufficient conditions, whose verification in the context of specific model is of reasonable difficulty, ensuring that a real number is the maximal stability level. We note that this question has already been debated in the pertaining literature, see for instance [22, Section 4.2], [14, Theorem 2.3], [19, Lemma 1.2 and hypothesis H2], and [20, Lemmas 1.16, 1.16 and 1.17]. We remark that in all the quoted references the authors always stated results in the case $|X_{\rm m}| = 1$.

These types of results can be obtained by looking at the problem by a different point of view, that is by looking at all the paths connecting any state not belonging to $X_{\rm s}$ to the set of ground states itself.

Theorem 2.2 *Consider a reversible energy landscape (X, Q, H, Δ). Assume $X \setminus X_{\rm s} \neq \emptyset$. If the not empty set $A \subset X \setminus X_{\rm s}$ and the positive real number $a \in \mathbb{R}_+$ are such that*

1. $\Phi(x, X_{\rm s}) - H(x) = a$ *for all* $x \in A$;
2. *either* $X \setminus (A \cup X_{\rm s}) = \emptyset$ *or* $\Phi(x, X_{\rm s}) - H(x) < a$ *for all* $x \in X \setminus (A \cup X_{\rm s})$;

then

$$\Gamma_{\rm m} = a \quad \textit{and} \quad X_{\rm m} = A$$

The above theorem is very useful in the applications, indeed it gives a general strategy to approach the problem of computing the maximal stability level in special models. The idea is that one has to figure out what is the set of configurations such that starting from them the cost to reach the ground states is precisely the maximal stability level. Once it has been proven that starting from all the configurations in this set the cost to be payed is the same, then one is just left with the proof that starting from any other configuration the cost is strictly smaller. And this is not a terrific computation since only an upper bound to the height along the paths is needed.

It is possible to prove a necessary condition in the spirit of the statement in Theorem 2.2. In other words on the basis of the Definition 2.1 we can say what is the barrier that must be overcome to visit the set of ground states $X_{\rm s}$ starting from states in $X_{\rm m}$.

Theorem 2.3 *Consider a reversible energy landscape (X, Q, H, Δ). Assume $X \setminus X_{\rm s} \neq \emptyset$. Then*

1. $\Phi(x, X_{\rm s}) - H(x) = \Gamma_{\rm m}$ *for all* $x \in X_{\rm m}$;
2. $\Phi(x, X_{\rm s}) - H(x) < \Gamma_{\rm m}$ *for all* $x \in X \setminus (X_{\rm m} \cup X_{\rm s})$.

This theorem ensures that $\Gamma_{\rm m}$ is the maximal height that must be overcome, starting from any state in the system, to reach the set $X_{\rm s}$ of ground states of the system. For this reason the quantity $\Gamma_{\rm m}$ will be also called the *relaxation height* of the system.

In many situations it is of great help combining the relaxation height and the stability level point of view to get sufficient conditions whose verification in the context of specific model is of reasonable difficulty. We then state the following theorem.

Theorem 2.4 *Consider a reversible energy landscape (X, Q, H, Δ). Assume $X \setminus X_{\rm s} \neq \emptyset$. If the not empty set $A \subset X \setminus X_{\rm s}$ and the positive real number $a \in \mathbb{R}_+$ are such that*

1. $\Phi(x, X_{\rm s}) - H(x) = a$ *for all* $x \in A$;

2. *either* $X \setminus (A \cup X_{\rm s}) = \emptyset$ *or* $V_x < a$ *for all* $x \in X \setminus (A \cup X_{\rm s})$*;*

then

$$\Gamma_{\rm m} = a \quad \textit{and} \quad X_{\rm m} = A$$

We remark that Theorem 2.2 is a corollary of Theorem 2.4. Indeed, if the set A and the real number a satisfy the hypotheses of Theorem 2.2 then they also satisfy the hypotheses of Theorem 2.4. Nevertheless, since in the proof of Theorem 2.4 the hypothesis that the state space X is finite play a crucial role while in the proof of Theorem 2.2 it does not, in Section 5 the two statements will be proven independently. The idea in the proof of Theorem 2.2 could result useful in the study of metastability of Statistical Mechanics systems in infinite volume.

3. Metastable states of Statistical Mechanics lattice models

The theory developed above can be fruitfully applied to study metastability in Statistical Mechanics lattice models when multiple metastable states are present. A natural setup in which this problem can be approached is that of Markov chains or Markov processes. In this context powerful theories [8, 22] have been developed with the aim of finding answers valid with maximal generality and to reduce to a minimum the number of model dependent inputs necessary to describe the metastable behavior of the system. In this section we shall briefly review the *pathwise* and the *potential theoretic* point of views and explain, in this context, the interest of the results stated in Section 2 in order to deal with the case of multiple metastable states. In particular in Section 3.2 we shall prove a recipe to construct metastable sets in the potential theoretic approach sense.

Consider a finite state space X and a function $q : X \times X \to [0, 1]$ called *connectivity matrix* such that for all $x, y \in X$

$$\sum_{w \in X \setminus \{x\}} q(x, w) = 1 \quad \text{and} \quad q(x, y) = q(y, x) \tag{3.1}$$

Moreover assume that for any $x, y \in X$ there exist an integer $n \geq 2$ and $x_1, \ldots, x_n \in X$ such that $x_1 = x$, $x_n = y$, and $q(x_i, x_{i+1}) > 0$ for any $i = 1, \ldots, n-1$.

Define $Q \subset X \times X$ by letting $(x, y) \in Q$ if and only if $q(x, y) > 0$. Consider the *Hamiltonian* $H : X \to \mathbb{R}$ and the *cost function* $\Delta : Q \to \mathbb{R}_+$ such that for all $(x, y) \in Q$

$$H(x) + \Delta(x, y) = \Delta(y, x) + H(y) \tag{3.2}$$

Moreover assume that the following is satisfied: (i) a state $x \in X$ is such that $q(x, x) = 0$ if and only if $\Delta(x, y) = 0$ for all $y \in X$ such that $y \neq x$ and $(x, y) \in Q$; (ii) for all $x \in X$ such that $q(x, x) > 0$ the cost function is such that $\Delta(x, x) = 0$.

Given $1/\beta > 0$, called *temperature*, consider the aperiodic ergodic Markov chain on the finite state space X and transition matrix $p_\beta : X \times X \to [0,1]$ given by

$$p_\beta(x,y) := q(x,y)e^{-\beta\Delta(x,y)} \qquad \forall x, y \in X \ \text{ and } x \neq y \tag{3.3}$$

and

$$p_\beta(x,x) := 1 - \sum_{y \in X\colon x \neq y} p_\beta(x,y) \qquad \forall x \in X \tag{3.4}$$

From the second equality in (3.1) and from (3.2) it follows that the chain is reversible with respect to the Gibbs measure

$$\mu_\beta(\sigma) := \frac{1}{Z_\beta}\, e^{-\beta H(x)} \tag{3.5}$$

where Z_β is the partition function, that is to say $\mu_\beta(x)p_\beta(x,y) = p_\beta(y,x)\mu_\beta(y)$ for any $x, y \in X$.

Noted that the quaternion (X, Q, H, Δ) is a reversible energy landscape, see Section 2.1, we use the theoretical setup introduced in Section 2 and, following [22], we call the set $X_{\rm m}$ introduced in Definition 2.1 *set of metastable states*. In the following two subsections we shall recall why such a set deserves its name.

3.1. Pathwise approach

The reason way $X_{\rm m}$ deserves its name has been widely explained in [22] in the framework of the so called *pathwise approach* to metastability. In that paper the important properties of the states in $X_{\rm m}$ have been fully studied in the case of the Metropolis dynamics, namely, when the cost function is chosen as $\Delta(x,y) := [H(y) - H(x)]_+$, where, for any real a, $[a]_+$ is equal to a if $a \geq 0$ and to 0 otherwise. Note that this choice for the function Δ is coherent with the theoretical setup introduced above since (3.2) is satisfied.

A partial generalization of the results in [22] to the present case, that is to say when the function Δ is as general as explained in Section 2, has been done in [14]. The authors had to consider this more general situation in [14] since such a setup arises naturally when Probabilistic Cellular Automata are studied.

Here, in order to justify the name of set of metastable states given to $X_{\rm m}$, we just recall the following result proven in [14]: for any $\varepsilon > 0$ and $x \in X_{\rm m}$

$$\lim_{\beta\to\infty} \mathbb{P}_x(e^{\beta(\Gamma_{\rm m}-\varepsilon)} < \tau_{X_{\rm s}} < e^{\beta(\Gamma_{\rm m}+\varepsilon)}) = 1 \tag{3.6}$$

where $\mathbb{P}_x$ is the probability measure on the space of trajectories of the Markov chain started at x and $\tau_{X_{\rm s}}$ is the first hitting time to $X_{\rm s}$ for the chain started at x. In the case of the Metropolis dynamics, this result has been proven in [22, Theorem 4.1]; in that paper a

much more detailed description of the metastable behavior of Metropolis dynamics, including results in distribution and in law [22, Theorems 4.9 and 4.15], has been given.

One of the remarkable features of the pathwise approach is that a *constructive* definition of metastable states is given. When studying a particular model, in order to find the metastable states, one just has to find the set $X_{\rm m}$, that is one has to compute the maximum stability level $\Gamma_{\rm m}$ and identify all the states whose stability level is equal to $\Gamma_{\rm m}$.

Unfortunately this is in practice a laborious task. Indeed the solution of several variational problems is required. It is then very useful the strategy suggested by the Theorems 2.2 and 2.4 for the computation of $\Gamma_{\rm m}$. This approach has been already used in some situations but always in cases in which there existed a single metastable state, see for instance [22, Section 4.2] and [14, Theorem 2.3].

3.2. Potential theoretic approach

Different interesting approaches to metastability have been developed in the recent literature, see for instance [2, 8–10, 20]. The notion of metastable states is given in different ways; but these different notions are indeed strictly related. For instance in [8] it is introduced the "metastable set" which, by using the language introduced above, can be reinterpreted in some cases as $X_{\rm m} \cup X_{\rm s}$. We discuss this issue in this subsection in detail.

Following [7, 8] we apply the theory developed in the Section 2 to the model defined at the beginning of Section 3. The potential theoretic approach to metastability, introduced in [7, 8], has been lately developed for Kawasaki dynamics in [9, 20] and for probabilistic cellular automata in [24].

In this approach the main definition of interest is that of metastable set [8]. For a generic Markov chain the *Dirichlet form* is defined as the functional

$$\mathfrak{E}_\beta[h] = \frac{1}{2}\sum_{x,y\in X} \mu_\beta(x) p_\beta(x,y)[h(x)-h(y)]^2 = \frac{1}{2}\sum_{x,y\in X} \frac{1}{Z_\beta} e^{-\beta[H(x)+\Delta(x,y)]}[h(x)-h(y)]^2 \quad (3.7)$$

where $h : X \to \mathbb{R}$ is a generic function and we have used (3.3) and the definition (3.5) of Gibbs measure.

Given two non–empty disjoint sets $A, B \subset X$ the *capacity* of the pair A and B is defined as

$$\mathrm{CAP}_\beta(A,B) := \min_{\substack{h:X\to[0,1]\\ h|_A=1, h|_B=0}} \mathfrak{E}_\beta(h) \quad (3.8)$$

Note that the capacity is a *symmetric* function of the sets A and B. It can be proven that the right hand side of (3.8) has a unique minimizer called *equilibrium potential* of the pair A and B and given by

$$h^*_{A,B}(x) = \mathbb{P}_x(\tau_A < \tau_B)$$

for any $x \in X$, where τ_A and τ_B are, respectively, the first hitting time to A and B for the chain started at x.

Definition 3.5 *A set $M \subset X$ is said to be* p.t.a.–metastable *if*

$$\lim_{\beta\to\infty} \frac{\max_{x\notin M} \mu_\beta(x)[\mathrm{CAP}_\beta(x,M)]^{-1}}{\min_{x\in M} \mu_\beta(x)[\mathrm{CAP}_\beta(x,M\setminus\{x\})]^{-1}} = 0 \tag{3.9}$$

The prefix p.t.a. stands for potential theoretic approach. We used this expression in order to avoid confusion with the set of metastable states X_{m} introduced in Definition 2.1.

In [8] and in the related papers the properties of the p.t.a.–metastable sets have been widely described. Here we just mention the following one to explain why these states are called metastable; we refer the interested reader to [7, 8]. Let M be a p.t.a.–metastable set. Let $x \in M$ and $J \subset M \setminus \{x\}$ be such that for all $y \in M \setminus (J \cup \{x\})$ either

$$\frac{\mu_\beta(y)}{\mu_\beta(x)} \ll 1 \qquad \text{or} \qquad \frac{\mathrm{CAP}_\beta(y,x)}{\mathrm{CAP}_\beta(y,J)} \ll 1$$

where by $\ll$ we mean that the ratios on the left–hand sides are uniformly bounded from above by a function of β tending to zero in the limit $\beta \to \infty$. Then

$$\mathbb{E}_x[\tau_J] = \frac{\mu_\beta(A(x))}{\mathrm{CAP}(x,J)}(1+o(1))$$

and

$$\mathbb{P}_x(\tau_J > t\mathbb{E}_x\tau_J) = [1+o(1)]e^{-t[1+o(1)]} \;\; t \geq 0$$

for β large enough, where $\mathbb{E}_x$ is the average for the chain started at x and, for any $x \in M$, the *valley* $A(x) \subset X$ is defined as

$$A(x) := \{y \in X : \mathbb{P}_y(\tau_x = \tau_M) = \sup_{z\in M} \mathbb{P}_y(\tau_z = \tau_M)\}$$

When the potential theoretic approach is applied to study metastability in specific models the main difficulties are the identification of the p.t.a.–metastable sets (e.g., the set made by the two configurations with all the spins respectively equal to minus one and to plus one in the standard Ising model) and that of giving sharp estimates to the capacities among the elements of this set. In this subsection we will show how the first of the two points above can be approached by using the notion of metastable states introduced in Definition 2.1.

It is important to note that the Definition 2.1 of set of metastable states and the Definition 3.5 of p.t.a.–metastable sets are different in spirit. The set of metastable states X_{m} is defined univocally, that is to say that all the states that are thought to deserve the name

of metastable states are collected in the same set X_{m}. In other words Definition 2.1 is constructive and the set of metastable states X_{m} is unique, even though its cardinality can be larger than one. On the other hand the p.t.a.–metastable sets are defined as those sets of states satisfying condition (3.9), so that several p.t.a.–metastable sets can exist. This subtle, but important, difference is often hidden when systems with a unique metastable state ($|X_{\mathrm{m}}| = 1$) are considered, but it emerges in all its importance when multiple metastable states ($|X_{\mathrm{m}}| \geq 2$) are present.

This aspect of the potential theoretic approach is, indeed, one of its distinguishing points. It reveals the high degree of tunability of the theory. The states in X_{m} are metastable in the sense that in order to decrease the energy, starting from them, the highest energy barrier (the relaxation height) has to be overcome; this is for sure a definition of metastable states absolutely close to the empirical physical meaning of the word metastable. In Theorem 3.6 we will show how to construct p.t.a.–metastable sets by means of states in X_{m} and X_{s}; in some sense we will construct p.t.a–metastable sets by wisely collecting "maximally stable" metastable states.

One could also enlarge its point of view by considering all the states such that starting from them the energy barrier that has to be overcome in order to lower the energy is larger than or equal to V, for some $V \in \mathbb{R}$ such that $0 < V < \Gamma_{\mathrm{m}}$. This set would obviously contain X_{m}, but, for V small enough, also some states not in X_{m} will belong to such a set. These states are not "maximally stable" metastable states, but, to some extent, they can be considered metastable since the energy barrier V has to be payed in order to decrease the energy. The potential theoretical approach is well suited to study also these states, indeed one just have to look for the p.t.a.–metastable states inside the new set just defined above. In the framework of the pathwise approach something in this sense has also been proven in [22], see, for instance, the definition [22, equation (2.13)] of metastable set at a prescribed level and the related result [22, Theorem 3.1].

In this subsection we try to clarify the connection between the "maximally stable" metastable sets in the two specified senses and, in particular, we will show how to construct the p.t.a.–metastable sets once the set of metastable states X_{m} is known. This seems to be a very smart recipe since X_{m} can be identified on the basis of the theory developed in Section 2, for instance by means of the sufficient conditions given in Theorem 2.2 or those in Theorem 2.4.

Define on X the following relation: $x, y \in X$ are $x \sim y$ if and only if $\Phi(x, y) - H(x) < \Gamma_{\mathrm{m}}$ and $\Phi(y, x) - H(y) < \Gamma_{\mathrm{m}}$. It is easy to prove that $\sim$ is an equivalence relation. Assume, now, that $X \setminus X_{\mathrm{s}} \neq \emptyset$, so that X_{m} is not empty. We denote by $X_{\mathrm{m}}^{(1)}, \ldots, X_{\mathrm{m}}^{(k_{\mathrm{m}})}$ the equivalence classes in which X_{m} is partitioned with respect to the relation $\sim$ (see figure 3.2). Analogously,

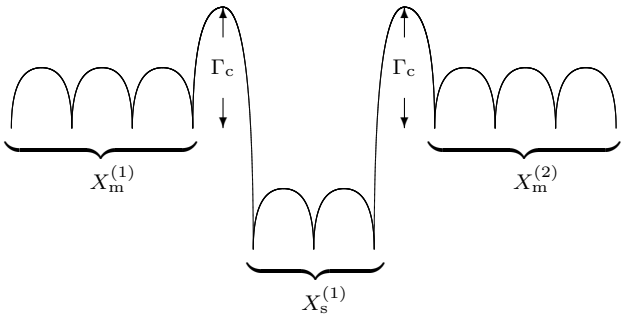

Figure 3.2: Schematic representation of the energy landscape aimed to illustrate the definition of the equivalence relation $\sim$ and the partition in equivalence classes of the sets $X_{\rm m}$ of metastable states and $X_{\rm s}$ of ground states.

we denote by $X_{\rm s}^{(1)}, \ldots, X_{\rm s}^{(k_{\rm s})}$ the equivalence classes in which the relation $\sim$ partitions $X_{\rm s}$ (see figure 3.2).

Theorem 3.6 *Assume that $X \setminus X_{\rm s} \neq \emptyset$ and $X \setminus (X_{\rm s} \cup X_{\rm m}) \neq \emptyset$. Chose arbitrarily $x_{{\rm s},i} \in X_{\rm s}^{(i)}$ for any $i = 1, \ldots, k_{\rm s}$ and $x_{{\rm m},i} \in X_{\rm m}^{(i)}$ for any $i = 1, \ldots, k_{\rm m}$. The set*

$$\{x_{{\rm s},1}, \ldots, x_{{\rm s},k_{\rm s}}, x_{{\rm m},1}, \ldots, x_{{\rm m},k_{\rm m}}\}$$

is p.t.a.–metastable.

In words we can say that any set constructed by considering one and only one element of each equivalence class $X_{\rm m}^{(i)}$ and $X_{\rm s}^{(j)}$ is a p.t.a.–metastable set.

4. Metastable behavior of the Blume–Capel model

In this section we apply the theory described above to the specific case of the Blume–Capel model with zero chemical potential.

As it has already been pointed out in the pertinent literature, in the study of the metastable behavior of a Statistical Mechanics model less details on the structure of the energy landscape of the model yield less instructive results on the metastable behavior of the system. This remark is quite obvious, but nevertheless interesting, since the study of the metastable behavior of a system is typically a very difficult task. So it is useful to understand fully to which extent the description of metastability can be pushed forward once some model dependent results are known.

Our results will be then stated according to the following main idea: we shall state model dependent lemmas containing informations on the energy landscape of the Blume–Capel model followed closely by theorems stating properties of the metastable states of the system that can be deduced by using the related lemma.

4.1. The Blume–Capel model

Consider a finite squared torus (periodic boundary condition) $\Lambda \subset \mathbb{Z}^2$ endowed with the *Euclidean* distance $\mathrm{d} : \Lambda \times \Lambda \to \mathbb{R}_+$ on the torus. As usual we shall misuse the notation and let

$$\mathrm{d}(I, J) := \min_{i \in I, j \in J} \mathrm{d}(i, j)$$

for any $I, J \subset \Lambda$. We say that two sites $i, j \in \Lambda$ are *nearest neighbor* if and only if $\mathrm{d}(i, j) = 1$.

Let $\{-1, 0, +1\}$ be the single spin state space and $\mathcal{X} := \{-1, 0, +1\}^\Lambda$ be the configuration space. The Hamiltonian of the model is

$$G(\sigma) = \sum_{<i,j>} (\sigma(i) - \sigma(j))^2 - \lambda \sum_{i \in \Lambda} \sigma(i)^2 - h \sum_{i \in \Lambda} \sigma(i) \tag{4.1}$$

for any $\sigma \in \mathcal{X}$, where the first sum is extended to the pairs of nearest neighbors, $\lambda \in \mathbb{R}$ is called *chemical potential*, and $h \in \mathbb{R}$ is called *magnetic field*. The metastable states for $h > \lambda > 0$ have been widely studied in [16, 17].

In this paper we shall attack the case $\lambda = 0$ which is particularly relevant because two metastable states will be proven to exist; in other words from now on we shall consider the zero chemical potential Blume–Capel model defined by the Hamiltonian

$$H(\sigma) = \sum_{<i,j>} (\sigma(i) - \sigma(j))^2 - h \sum_{i \in \Lambda} \sigma(i) \tag{4.2}$$

Given $\sigma \in \mathcal{X}$, $H(\sigma)$ will be also called *energy* of the configuration σ.

The stochastic version of the model is the Markov chain σ_t, with $t = 0, 1, \dots$ the discrete time variable, with transition probabilities (3.3) with $\Delta(\sigma, \eta) := [H(\eta) - H(\sigma)]_+$ and the connectivity matrix defined as

$$q(\sigma, \eta) := \begin{cases} 0 & \text{if } \sigma, \eta \text{ differ at more than one site} \\ \dfrac{1}{2|\Lambda|} & \text{if } \sigma, \eta \text{ differ at one single site} \end{cases}$$

for any $\sigma, \eta \in \mathcal{X}$ and $\sigma \neq \eta$, and $q(\sigma, \sigma) = 1$ for any $\sigma \in \mathcal{X}$. Note that, since the dynamics is Metropolis, all the results in [22] hold for this model.

It is worth nothing that from the definition of the transition matrix q given above, it follows that for any $\sigma, \eta \in \mathcal{X}$ there exist a path connecting σ to η, namely, $|\Omega(\sigma, \eta)| \geq 1$.

4.2. Metastable states of the Blume–Capel model

We shall discuss the metastable behavior of the zero chemical potential Blume–Capel model for the following choice of the parameters. For any positive real a we let $\lfloor a \rfloor$ be the largest integer smaller than or equal to a.

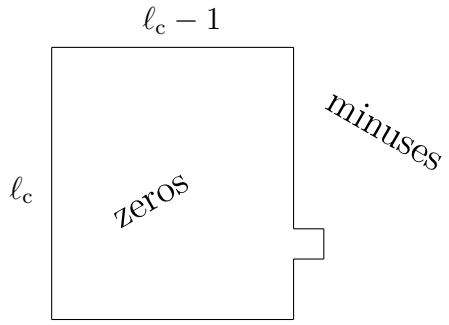

Figure 4.3: Schematic representation of the configuration $\mathcal{P}_c$. The configuration $\mathcal{Q}_c$ has the same geometry with the zeros replaced by the pluses and the minuses replaced by the zeros. Note that the protuberance can be either on the left or on the right vertical edge (the longest ones in the picture) and there it can be shifted freely.

Condition 4.7 *The magnetic field h and the torus Λ are such that $0 < h < 1$, $\lfloor 2/h \rfloor$ is not integer, and $|\Lambda| \geq 49/h^4$ finite.*

For $h > 0$ it is immediate to remark that the set of ground states of the energy is $\mathcal{X}_s = \{\mathbf{u}\}$, where $\mathbf{u} \in \mathcal{X}$ is the configuration such that $\mathbf{u}(i) = +1$ for all $i \in \Lambda$. Indeed the exchange interaction, i.e., the positive defined first term of the Hamiltonian in (4.2), gives its minimal contribution which is equal to zero. And so does the magnetic field part.

Other two very relevant configurations are $\mathbf{d}$ and $\mathbf{0}$, that is the configuration in which all the spin are minus one and the one in which all the spins are zero. Indeed in these configurations the exchange part of the energy is minimal, although the magnetic part is not. We also note that

$$H(\mathbf{u}) = -|\Lambda|h < H(\mathbf{0}) = 0 < H(\mathbf{d}) = +|\Lambda|h$$

On physical grounds, see (for instance) the zero temperature phase diagram in [17, Fig. 1], it is quite reasonable to guess that $\mathbf{d}$ and $\mathbf{0}$ are possible metastable states of the system. This is indeed true as it will be stated in the following Theorem 4.10. To get this result we shall use the strategy outlined in Section 2: in particular we will use the Theorem 2.4 with guess set $A = \{\mathbf{d}, \mathbf{0}\}$; this will yield the sufficient model dependent statements.

We now define the *critical length* of the model as

$$\ell_c := \lfloor \frac{2}{h} \rfloor + 1 \tag{4.3}$$

and remark that, by Condition 4.7, it follows that

$$\frac{2}{h} < \ell_c < \frac{2}{h} + 1 \quad \text{and} \quad \ell_c \geq 3 \tag{4.4}$$

We denote by $\mathcal{P}_c$ the set of configurations in which all the spins are minus excepted those, which are zeros, in a rectangle of sides long ℓ_c and $\ell_c - 1$ and in a site adjacent to one of the longest sides of the rectangle (see figure 4.3). We denote by $\mathcal{Q}_c$ the set of configurations in which all the spins are zeros excepted those, which are pluses, in a rectangle of sides long ℓ_c and $\ell_c - 1$ and in a site adjacent to one of the longest sides of the rectangle (see the caption of figure 4.3). From (4.2) it follows that

$$H(\mathcal{P}_c) - H(\mathbf{d}) = H(\mathcal{Q}_c) - H(\mathbf{0}) = 4\ell_c - h[\ell_c(\ell_c - 1) + 1] \tag{4.5}$$

We then set

$$\Gamma_c := H(\mathcal{P}_c) - H(\mathbf{d}) = H(\mathcal{Q}_c) - H(\mathbf{0}) \tag{4.6}$$

A simple direct computation shows that for h small one has $\Gamma_c \sim 4/h$.

We are now ready to state the model dependent lemmas on which it will be based the description of the metastable behavior of the zero chemical potential Blume–Capel model (4.2) (see Section 6 for the proof).

Lemma 4.8 *Consider the zero chemical potential Blume–Capel model (4.2) and assume that Condition 4.7 is satisfied. We have that*

1. *for any configuration $\eta \in \mathcal{P}_c$ there exists a path $\omega \in \Omega(\mathbf{d}, \mathbf{u})$ such that $\Phi_\omega - H(\mathbf{d}) = \Gamma_c$, $\omega_i = \eta$ for some $i \in \{1, \dots, n\}$ (where n is the length of the path), and $H(\omega_j) < H(\mathbf{d}) + \Gamma_c$ for all $j \in \{1, \dots, n\}$ and $j \neq i$;*

2. *for any configuration $\eta \in \mathcal{Q}_c$ there exists a path $\omega \in \Omega(\mathbf{0}, \mathbf{u})$ such that $\Phi_\omega - H(\mathbf{0}) = \Gamma_c$, $\omega_i = \eta$ for some $i \in \{1, \dots, n\}$ (where n is the legth of the path), and $H(\omega_j) < H(\mathbf{0}) + \Gamma_c$ for all $j \in \{1, \dots, n\}$ and $j \neq i$.*

Lemma 4.9 *Consider the zero chemical potential Blume–Capel model (4.2) and assume that Condition 4.7 is satisfied. We have that*

1. *$\Phi(\mathbf{d}, \mathbf{u}) - H(\mathbf{d}) = \Phi(\mathbf{0}, \mathbf{u}) - H(\mathbf{0}) = \Gamma_c$;*

2. *$V_\sigma < \Gamma_c$ for any $\sigma \in \mathcal{X} \setminus \{\mathbf{d}, \mathbf{0}, \mathbf{u}\}$.*

Note that the above lemma ensures the validity of the hypotheses of Theorem 2.4 with $A = \{\mathbf{d}, \mathbf{0}\}$. Then, by using the general results discussed in the above section, one can identify the set of metastable states and the relaxation height (see Section 6 for the proofs).

Theorem 4.10 *Consider the zero chemical potential Blume–Capel model (4.2) and assume that Condition 4.7 is satisfied. We have that*

$$\mathcal{X}_{\rm m} = \{\mathbf{d}, \mathbf{0}\} \quad \textit{and} \quad \Gamma_{\rm m} = \Gamma_{\rm c}$$

Moreover we can prove a theorem giving the asymptotic of the exit time for the system started at the metastable states.

Theorem 4.11 *Consider the zero chemical potential Blume–Capel model (4.2) and assume that Condition 4.7 is satisfied. The following holds:*

1. *considered the chain started at* $\mathbf{d}$*, for any* $\varepsilon > 0$

$$\lim_{\beta\to\infty} \mathbb{P}_{\mathbf{d}}(e^{\beta(\Gamma_{\rm c}-\varepsilon)} < \tau_{\mathbf{u}} < e^{\beta(\Gamma_{\rm c}+\varepsilon)}) = 1$$

and

$$\lim_{\beta\to\infty} \frac{1}{\beta}\mathbb{E}_{\mathbf{d}}[\tau_{\mathbf{u}}] = \Gamma_{\rm c}$$

where $\tau_{\mathbf{u}}$ *denotes the first hitting to* $\mathbf{u}$ *of the chain started at* $\mathbf{d}$ *and* $\mathbb{P}_{\mathbf{d}}$ *and* $\mathbb{E}_{\mathbf{d}}$ *denote, respectively, the probability and the average on the space of trajectories started at* $\mathbf{d}$*;*

2. *Considered the chain started at* $\mathbf{0}$*, for any* $\varepsilon > 0$

$$\lim_{\beta\to\infty} \mathbb{P}_{\mathbf{0}}(e^{\beta(\Gamma_{\rm c}-\varepsilon)} < \tau_{\mathbf{u}} < e^{\beta(\Gamma_{\rm c}+\varepsilon)}) = 1$$

and

$$\lim_{\beta\to\infty} \frac{1}{\beta}\mathbb{E}_{\mathbf{0}}[\tau_{\mathbf{u}}] = \Gamma_{\rm c}$$

where, here, $\tau_{\mathbf{u}}$ *denotes the first hitting to* $\mathbf{u}$ *of the chain started at* $\mathbf{0}$ *and* $\mathbb{P}_{\mathbf{0}}$ *and* $\mathbb{E}_{\mathbf{0}}$ *denote, respectively, the probability and the average on the space of trajectories started at* $\mathbf{0}$*.*

4.3. Escape mechanism

One important result in metastability, besides proving the asymptotic on the exit time, is that of identifying the escape mechanism. The typical result is that, in order to perform the transition to the stable state, the system has to nucleate a critical droplet of the stable phase inside the metastable one. To prove this result supplementary model dependent properties are needed.

Lemma 4.12 *Consider the zero chemical potential Blume–Capel model (4.2) and assume that Condition 4.7 is satisfied. We have that*

1. *any path* $\omega \in \Omega(\mathbf{d}, \mathbf{u})$ *such that* $\Phi_\omega - H(\mathbf{d}) = \Gamma_{\mathrm{c}}$ *visits* $\mathcal{P}_{\mathrm{c}}$*, that is there exists an integer* i *such that* $\omega_i \in \mathcal{P}_{\mathrm{c}}$*;*

2. *any path* $\omega \in \Omega(\mathbf{0}, \mathbf{u})$ *such that* $\Phi_\omega - H(\mathbf{0}) = \Gamma_{\mathrm{c}}$ *visits* $\mathcal{Q}_{\mathrm{c}}$*, that is there exists an integer* i *such that* $\omega_i \in \mathcal{Q}_{\mathrm{c}}$*.*

Theorem 4.13 *Consider the zero chemical potential Blume–Capel model (4.2) and assume that Condition 4.7 is satisfied. The following holds:*

1. *considered the chain started at* $\mathbf{d}$*, we have that*

$$\lim_{\beta\to\infty} \mathbb{P}_{\mathbf{d}}(\tau_{\mathcal{P}_{\mathrm{c}}} < \tau_{\mathbf{u}}) = 1$$

where $\tau_{\mathcal{P}_{\mathrm{c}}}$ *and* $\tau_{\mathbf{u}}$ *denote respectively the first hitting time to* $\mathcal{P}_{\mathrm{c}}$ *and* $\mathbf{u}$ *for the chain started at* $\mathbf{d}$*;*

2. *considered the chain started at* $\mathbf{0}$*, we have that*

$$\lim_{\beta\to\infty} \mathbb{P}_{\mathbf{0}}(\tau_{\mathcal{Q}_{\mathrm{c}}} < \tau_{\mathbf{u}}) = 1$$

where, here, $\tau_{\mathcal{Q}_{\mathrm{c}}}$ *and* $\tau_{\mathbf{u}}$ *denote respectively the first hitting time to* $\mathcal{Q}_{\mathrm{c}}$ *and* $\mathbf{u}$ *for the chain started at* $\mathbf{0}$*.*

4.4. Remark on the proof of the nucleation property

When studying the metastable behavior of a system, the detail of the results that one gets depends on the amount of model dependent properties that one is able to prove. In order to prove the time asymptotic in Theorem 4.11 only Lemma 4.9 is needed. As already remarked, to get the nucleation property in Theorem 4.13 it is needed the Lemma 4.12 in which it is identified the state that must necessarily be visited by a path joining the metastable state to the stable one with maximal height equal to the maximal stability level plus the energy of the starting metastable state. In particular it is proven that such a state is in $\mathcal{P}_{\mathrm{c}}$ for the paths starting at $\mathbf{d}$ and in $\mathcal{Q}_{\mathrm{c}}$ for those starting at $\mathbf{0}$.

Following the nomenclature in [22], to which we refer the interested reader for a more detailed discussion, we say that, given $\sigma, \eta \in \mathcal{X}$, a set $\mathcal{Y} \subset \mathcal{X}$ is a *gate* for the pair of configurations σ and η if and only if all the configurations in $\mathcal{Y}$ have energy equal to the communication height between σ and η and any path joining σ to η with maximal height equal to the communication height between σ and η has to pass necessarily through $\mathcal{Y}$ itself. More precisely, a set $\mathcal{Y} \subset \mathcal{X}$ is a *gate* for the pair of configurations σ and η if and only if $H(\zeta) = \Phi(\sigma, \eta)$ for any $\zeta \in \mathcal{Y}$ and for any path $\omega \in \Omega(\sigma, \eta)$ such that $\Phi_\omega = \Phi(\sigma, \eta)$ there

must be $\zeta \in \mathcal{Y}$ and i integer such that $\omega_i = \zeta$. Thus, Lemma 4.12 implies that the set $\mathcal{P}_c \subset \mathcal{X}$ is a gate for the pair of configurations $\mathbf{d}$ and $\mathbf{u}$, whereas the set $\mathcal{Q}_c \subset \mathcal{X}$ is a gate for the pair of configurations $\mathbf{0}$ and $\mathbf{u}$.

Less details on the structure of the energy landscape of the model would have yielded less instructive results. Indeed, in the proof of Lemma 4.9, in order to get the statement we shall define the set $\mathcal{X}_\mathbf{d}$ of all the configurations having precisely $\ell_c(\ell_c-1)+1$ minus spins. As a byproduct of the proof of Lemma 4.9 we shall also get (for free) that any path connecting the metastable state $\mathbf{d}$ to the stable state $\mathbf{u}$ with maximal height equal to the maximal stability level plus the energy of the starting metastable state must necessarily visit the set $F(\mathcal{X}_\mathbf{d})$ of configurations in $\mathcal{X}_\mathbf{d}$ with minimal energy. So that, without any supplementary effort with respect to the proof of Lemma 4.9, we have that $F(\mathcal{X}_\mathbf{d})$ is also a gate for the pair of configurations $\mathbf{d}$ and $\mathbf{u}$. Then, by the theory developed in [22], we get

$$\lim_{\beta\to\infty} \mathbb{P}_\mathbf{d}(\tau_{F(\mathcal{X}_\mathbf{d})} < \tau_\mathbf{u}) = 1$$

where the chain is started at $\mathbf{d}$ and $\tau_{F(\mathcal{X}_\mathbf{d})}$ and $\tau_\mathbf{u}$ denote respectively the first hitting time to $F(\mathcal{X}_\mathbf{d})$ and $\mathbf{u}$.

This result can be looked at as a nucleation statement, but less precise if compared to that in the first part of Theorem 4.13. This is somehow natural since it has been obtained with less information on the structure of the energy landscape of the system. In the language of [22] both $\mathcal{P}_c$ and $F(\mathcal{X}_\mathbf{d})$ are gates for the pair of configurations $\mathbf{d}$ and $\mathbf{u}$, but $\mathcal{P}_c$ is a so called *minimal* gate.

A similar discussion can be repeated for the pair of configuration $\mathbf{0}$ and $\mathbf{u}$ with $\mathcal{X}_\mathbf{d}$ replaced by $\mathcal{X}_\mathbf{0}$, namely, the set of configurations in which all the spins are zeros excepted for $\ell_c(\ell_c-1)+1$ which are pluses.

4.5. Remarks on gates and minimal gates

In [22, page 604] a gate $\mathcal{Y} \subset \mathcal{X}$ for the pair of configurations σ and η is said to be *minimal* if and only if for any proper subset $\mathcal{Y}'$ of $\mathcal{Y}$ there exists a path joining σ to η with maximal height equal to the communication height between σ and η which does not pass through $\mathcal{Y}'$. In other words a gate is minimal if and only if all its proper subsets are not gates. In view of this we have immediately that the gate $\mathcal{X}_\mathbf{d}$ for the pair $\mathbf{d}$ and $\mathbf{u}$ and the gate $\mathcal{X}_\mathbf{0}$ for the pair $\mathbf{0}$ and $\mathbf{u}$ are not minimal. On the other hand from Lemma 4.8 we have that $\mathcal{P}_c$ and $\mathcal{Q}_c$ are minimal gates for the specified pair of configurations.

In general minimal gates are not unique. In [22, Theorem 5.1] it is proven that the union $\mathcal{G}(\sigma,\eta)$ of all the minimal gates for the pair of configurations σ and η is the set of what the authors call *essential saddles* (the interested reader is referred to [22, page 604] for the

definition). In our model we can prove that $\mathcal{P}_c$ is the unique minimal gate for the pair $\mathbf{d}$ and $\mathbf{u}$ and $\mathcal{Q}_c$ is the unique minimal gate for the pair $\mathbf{0}$ and $\mathbf{u}$, so that we can state the following theorem.

Theorem 4.14 *Consider the zero chemical potential Blume–Capel model (4.2) and assume that Condition 4.7 is satisfied. Then*

$$\mathcal{G}(\mathbf{d}, \mathbf{u}) = \mathcal{P}_c \quad \textit{and} \quad \mathcal{G}(\mathbf{0}, \mathbf{u}) = \mathcal{Q}_c$$

We finally remark that by [22, Theorem 5.1] it follows that in our model configurations in $\mathcal{P}_c$ (resp. $\mathcal{Q}_c$) are the sole essential saddles for the pair of configurations $\mathbf{d}$ and $\mathbf{u}$ (resp. $\mathbf{0}$ and $\mathbf{u}$).

5. Proof of the general results

In this section we prove the theorems and lemmas concerning the general results stated in Section 2 and Section 3.

Proof of Theorem 2.2. First of all we note that, if $X \setminus (A \cup X_s) \neq \emptyset$ one has that

$$V_x < a \qquad \forall x \in X \setminus (A \cup X_s) \tag{5.1}$$

Indeed, by item 2 in the hypotheses it follows that there exists a path $\omega \in \Omega(x, X_s)$ such that $\Phi_\omega - H(x) < a$. Since $H(x) > H(X_s)$, we have that the same path belongs also to $\Omega(x, I_x)$ so that $V_x < \Phi_\omega - H(x)$, which proves (5.1).

We want to prove, now, that

$$V_x = a \qquad \forall x \in A \tag{5.2}$$

The same argument developed above and item 1 in the hypothesis prove that

$$V_x \leq a \qquad \forall x \in A \tag{5.3}$$

We are now left with the proof of the lower bound

$$V_x \geq a \qquad \forall x \in A \tag{5.4}$$

Pick $x \in A$ and, by absurdity, assume that $V_x < a$. Then, there exists $x_1 \in X$ and $x_1 \neq x$ such that $H(x_1) < H(x)$ and a path $\omega_1 \in \Omega(x, x_1)$ such that

$$\Phi_{\omega_1} - H(x) < a \tag{5.5}$$

If it were $x_1 \in X_s$ we would immediately get a contradiction, since we would have $\omega_1 \in \Omega(x, X_s)$ and hence $\Phi(x, X_s) - H(x) < a$ in contrast with the item 1 of the hypothesis.

Assume, finally, that $x_1 \in X \setminus X_{\rm s}$. By items 1 or 2 in the hypothesis we have that there exists a path $\omega_2 \in \Omega(x_1, X_{\rm s})$ such that

$$\Phi_{\omega_2} - H(x_1) \leq a \tag{5.6}$$

Now, note that by gluing the path ω_2 to the last configuration of ω_1 we get the path $\omega_1\omega_2 \in \Omega(x, X_{\rm s})$. Moreover,

$$\Phi_{\omega_1\omega_2} = \max\{\Phi_{\omega_1}, \Phi_{\omega_2}\}$$

and hence

$$\Phi_{\omega_1\omega_2} - H(x) = \max\{\Phi_{\omega_1} - H(x), \Phi_{\omega_2} - H(x)\} \tag{5.7}$$

Note, now, that

$$\Phi_{\omega_2} - H(x) = \Phi_{\omega_2} - H(x_1) + H(x_1) - H(x)$$

Since $H(x_1) < H(x)$, from (5.6) it follows that

$$\Phi_{\omega_2} - H(x) < a \tag{5.8}$$

In conclusion, by (5.5), (5.7), and (5.8) we have that

$$\Phi_{\omega_1\omega_2} - H(x) < a$$

which, recalling that $\omega_1\omega_2 \in \Omega(x, X_{\rm s})$ contradicts item 1 in the hypothesis. By contradiction we have that $V_x \geq a$ and hence prove (5.4).

By (5.3) and (5.4) it eventually follows (5.2). Recalling (2.7) in Definition 2.1, from (5.2) and (5.1) we get the desired equality $\Gamma_{\rm m} = a$.

We finally remark that, since $\Gamma_{\rm m} = a$, from Definition 2.1 it follows immediately that $X_{\rm m} = A$. Indeed, since $\Gamma_{\rm m} = a$ and $X_{\rm m}$ is the collection of states in $X \setminus X_{\rm s}$ such that $V_x = \Gamma_{\rm m}$, from items 1 and 2 in the hypothesis it follows that $X_{\rm m} = A$. □

Proof of Theorem 2.3. By Definition 2.1 it follows immediately that

$$\Phi(x, I_x) - H(x) = \Gamma_{\rm m} \qquad \forall x \in X_{\rm m} \tag{5.9}$$

and

$$\Phi(x, I_x) - H(x) < \Gamma_{\rm m} \qquad \forall x \in X \setminus (X_{\rm m} \cup X_{\rm s}) \tag{5.10}$$

Proof of item 1: first of all we note that from (5.9) it follows immediately that

$$\Phi(x, X_{\rm s}) - H(x) \geq \Gamma_{\rm m} \qquad \forall x \in X_{\rm m} \tag{5.11}$$

Indeed, recalling (2.5), we have that $X_{\rm s} \subset I_x$ implies $\Phi(x, I_x) - H(x) \leq \Phi(x, X_{\rm s}) - H(x)$.

To prove the upper bound, we pick $x \in X_{\rm m}$ and note that, by (5.9), (5.10), and the finiteness of the state space X, we have that there exist a sequence of $n-1$ states $x_1, \dots, x_{n-1} \in X \setminus X_{\rm s}$, a ground state $x_n \in X_{\rm s}$, and n paths $\omega_i \in \Omega(x_{i-1}, x_i)$ with $i = 1, \dots, n$ such that

$$H(x_0) > H(x_1) > H(x_2) > \cdots H(x_{n-1}) > H(x_n)$$

and

$$\Phi_{\omega_i} - H(x_{i-1}) \leq \Gamma_{\rm m}$$

for all $i = 1, \dots, n$, where we have let $x_0 = x$. Note that the bound is not strict because some of the states $x_0, x_1, \dots, x_{n-1}$ could belong to $X_{\rm m}$ (at least $x_0 = x$ does belong to $X_{\rm m}$). We then remark that the path $\omega_1\omega_2\cdots\omega_n$ obtained by gluing the above paths belongs to $\Omega(x, X_{\rm s})$ and

$$\Phi_{\omega_1\omega_2\cdots\omega_n} - H(x) = \max_{i=1,\dots,n} [\Phi_{\omega_i} - H(x)] \tag{5.12}$$

For any $i = 2, \dots, n$ we have that

$$\Phi_{\omega_1} - H(x) \leq \Gamma_{\rm m} \quad \text{and} \quad \Phi_{\omega_i} - H(x) = \Phi_{\omega_i} - H(x_{i-1}) + H(x_{i-1}) - H(x) < \Gamma_{\rm m} \tag{5.13}$$

since $H(x_{i-1}) < X(x)$ for all $i = 1, \dots, n$. In conclusion we have that

$$\Phi_{\omega_1\omega_2\cdots\omega_n} - H(x) \leq \Gamma_{\rm m}$$

which, recalling that $\omega_1\omega_2\cdots\omega_n \in \Omega(x, X_{\rm s})$ proves the upper bound

$$\Phi(x, X_{\rm s}) - H(x) \leq \Gamma_{\rm m} \qquad \forall x \in X_{\rm m} \tag{5.14}$$

Equations (5.11) and (5.14) finally yields item 1 of the theorem.

Proof of item 2: we follow the same strategy as that used in the proof of the upper bound (5.14) above. We just notice that, since $x \in X \setminus (X_{\rm m} \cup X_{\rm s})$, in equation (5.13) the first upper bound is strict, as it follows from (5.10). □

Proof of Theorem 2.4. If $X \setminus (A \cup X_{\rm s}) \neq \emptyset$, by following the same strategy described in the proof of Theorem 2.2 one easily gets

$$V_x \leq a \qquad \forall x \in A \tag{5.15}$$

Since one has to prove that $V_x = a$ for any $x \in A$, a lower bound is needed. Then, given a generic $x \in A$, we assume by absurdity that $V_x < a$. Thus, there exists $x_1 \in X$ such that $H(x_1) < H(x)$ and a path $\omega_1 \in \Omega(x, x_1)$ such that

$$\Phi_{\omega_1} - H(x) < a \tag{5.16}$$

If it were $x_1 \in X_{\rm s}$ we would immediately get a contradiction, since we would have $\omega_1 \in \Omega(x, X_{\rm s})$ and hence $\Phi(x, X_{\rm s}) - H(x) < a$ in contrast with item 1 in the hypotheses of this theorem.

If, on the other hand, $x_1 \in X \setminus X_{\rm s}$, by exploiting the fact that $X_{\rm s}$ is a finite set, we can proceed as above and find a sequence of $n-2$ states $x_2, \dots, x_{n-1} \in X \setminus X_{\rm s}$, a ground state $x_n \in X_{\rm s}$, and $n-1$ paths $\omega_i \in \Omega(x_{i-1}, x_i)$ with $i = 2, \dots, n$ such that

$$H(x_1) > H(x_2) > \cdots H(x_{n-1}) > H(x_n)$$

and

$$\Phi_{\omega_i} - H(x_{i-1}) \leq a$$

for all $i = 2, \dots, n$. Note that the bound is not strict because some of the configurations $x_1, \dots, x_{n-1}$ could belong to A. We then remark that the path $\omega_1\omega_2 \cdots \omega_n$ obtained by gluing the above paths belongs to $\Omega(x, X_{\rm s})$ and

$$\Phi_{\omega_1\omega_2\cdots\omega_n} - H(x) = \max_{i=1,\dots,n} [\Phi_{\omega_i} - H(x)] \tag{5.17}$$

Let $x_0 = x$, for any $i = 1, \dots, n$ we have that

$$\Phi_{\omega_i} - H(x) = \Phi_{\omega_i} - H(x_{i-1}) + H(x_{i-1}) - H(x) < a$$

since $H(x_{i-1}) < X(x)$ for all $i = 1, \dots, n$. In conclusion we have that

$$\Phi_{\omega_1\omega_2\cdots\omega_n} - H(x) < a$$

which, recalling that $\omega_1\omega_2 \cdots \omega_n \in \Omega(x, X_{\rm s})$ contradicts item 1 in the hypotheses of this theorem.

We have finally proven that $V_x = a$ for any $x \in A$. This, together with item 2 in the hypotheses of the theorem, implies that $\Gamma_{\rm m} = a$.

We finally remark that, since $\Gamma_{\rm m} = a$, from Definition 2.1 it follows immediately that $X_{\rm m} = A$. Indeed, since $\Gamma_{\rm m} = a$ and $X_{\rm m}$ is the collection of states in $X \setminus X_{\rm s}$ such that $V_x = \Gamma_{\rm m}$, from items 1 and 2 in the hypothesis it follows that $X_{\rm m} = A$. □

We consider now the model in Subsection 3.2 and prove the Theorem 3.6. Before proving the theorem we recall the following elementary estimate given in [9] in the context of Kawasaki dynamics.

Lemma 5.15 *For every non–empty disjoint sets $A, B \subset X$, there exist constants $0 < C_1 \leq C_2 < \infty$ (depending on A and B) such that for all $\beta > 0$*

$$C_1 \leq e^{\beta\Phi(A,B)}\, Z_\beta\, \mathrm{CAP}_\beta(A, B) \leq C_2 \tag{5.18}$$

Proof of Lemma 5.15: the lemma can be achieved via the same argument developed in the proof of [9, Lemma 3.1.1]. □

Proof of Theorem 3.6. Let $M := \{x_{\mathrm{s},1}, \dots, x_{\mathrm{s},k_\mathrm{s}}, x_{\mathrm{m},1}, \dots, x_{\mathrm{m},k_\mathrm{m}}\}$. Strategy of the proof: we bound the numerator in the left–hand side of (3.9) from above and the denominator from below. These bounds will be sufficient to prove that the ratio tends to zero for $\beta \to \infty$.

We first bound the numerator in the left–hand side of (3.9) from above. By using the lower bound in (5.18) we have that

$$\mu_\beta(x)[\mathrm{CAP}_\beta(x, M)]^{-1} \le \mu_\beta(x)\frac{1}{C_1}e^{\beta\,\Phi(x,M)}Z_\beta = \frac{1}{C_1}\,e^{\beta[\Phi(x,M)-H(x)]} \tag{5.19}$$

for any $x \notin M$.

Let $x \in X\setminus(X_\mathrm{m}\cup X_\mathrm{s})$. Since $M\cap X_\mathrm{s}^{(i)} \neq \emptyset$ for any $i = 1, \dots, k_\mathrm{s}$, item 2 of Theorem 2.3 and the definition of the equivalence relation $\sim$ given in Section 3.2 imply that $\Phi(x, M) - H(x) < \Gamma_\mathrm{m}$.

Let, now, $x \in (X_\mathrm{s}\cup X_\mathrm{m})\setminus M$ (note that this set can be possibly empty). Since $M\cap X_\mathrm{s}^{(i)} \neq \emptyset$ for any $i = 1, \dots, k_\mathrm{s}$, and $M \cap X_\mathrm{m}^{(i)} \neq \emptyset$ for any $i = 1, \dots, k_\mathrm{m}$, from the definition of the relation $\sim$ given in Section 3.2, we have that $\Phi(x, M) - H(x) < \Gamma_\mathrm{m}$.

These remarks and (5.19) yield that there exist two real numbers $C_3 < \infty$ and $\delta > 0$ such that

$$\mu_\beta(x)[\mathrm{CAP}_\beta(x, M)]^{-1} \le C_3 e^{\beta(\Gamma_m-\delta)} \tag{5.20}$$

for any $x \notin M$.

We, now, bound the denominator in the left–hand side of (3.9) from below. By using the upper bound in (5.18) we have that for any $x \in M$

$$\mu_\beta(x)[\mathrm{CAP}_\beta(x, M \setminus \{x\})]^{-1} \ge \mu_\beta(\eta)\frac{1}{C_2}e^{\beta\,\Phi(x,M\setminus\{x\})}Z_\beta = \frac{1}{C_2}\,e^{\beta[\Phi(x,M\setminus\{x\})-H(x)]} \tag{5.21}$$

First assume $x \in M \cap X_\mathrm{m}$. From the definition of the equivalence relation $\sim$ and the fact that $|M \cap X_\mathrm{m}^{(i)}| = 1$ for any $i = 1, \dots, k_\mathrm{m}$, we have that, provided $|M \cap X_\mathrm{m}| \ge 2$,

$$\Phi(x, (M \setminus \{x\}) \cap X_\mathrm{m}) - H(x) \ge \Gamma_\mathrm{m}$$

On the other hand, since $M \cap X_\mathrm{s}^{(i)} \neq \emptyset$ for any $i = 1, \dots, k_\mathrm{m}$, from item 1 in Theorem 2.3 and from the definition of the equivalence relation $\sim$ we have that

$$\Phi(x, (M \setminus \{x\}) \cap X_\mathrm{s}) - H(x) \ge \Gamma_\mathrm{m}$$

Recalling (2.5), from the two inequalities above, we have that

$$\Phi(x, M \setminus \{x\}) - H(x) = \min_{y\in M\setminus\{x\}} \Phi(x, y) - H(x) \ge \Gamma_\mathrm{m}$$

for any $x \in M \cap X_{\rm m}$.

Suppose, now, $x \in M \cap X_{\rm s}$. By proceeding as above, recalling (2.4), and using that $H(X_{\rm s}) < H(y)$ for any $y \in X_{\rm m}$, we prove that

$$\Phi(x, M \setminus \{x\}) - H(x) \geq \Gamma_{\rm m}$$

for any $x \in M \cap X_{\rm s}$.

We have eventually proven that $\Phi(x, M \setminus \{x\}) - H(x) \geq \Gamma_{\rm m}$ for any $x \in M$. This remark and (5.21) yields the lower bound

$$\mu_\beta(x)[{\rm CAP}_\beta(x, M \setminus \{x\})]^{-1} \geq \frac{1}{C_2}\, e^{\beta\Gamma_{\rm m}} \tag{5.22}$$

for any $x \in M$. Hence, from (5.20) and (5.22) we finally get

$$\frac{\max_{x\notin M} \mu_\beta(x)[{\rm CAP}_\beta(x, M)]^{-1}}{\min_{x\in M} \mu_\beta(x)[{\rm CAP}_\beta(x, M \setminus \{x\})]^{-1}} \leq \frac{C_3}{C_2}\, e^{-\beta\delta}$$

that completes the proof of the theorem. □

6. Proof of results concerning the Blume–Capel model

In this subsection we prove the Lemmas 4.8 and 4.9 and the Theorems 4.10 and 4.11. Before starting the proof of Lemmas 4.8 and 4.9 we state few technical results on *polyominoes* based on the paper [1] to which we refer the interested reader for more details.

6.1. Some results on polyominoes

Polyominoes have been widely studied both by physicists, to model crystal growth, and by combinatorialists. In combinatorics the main problem has been that of counting the number of polyominoes given their area or their perimeter [6].

In the framework of metastability polyominoes play an important role and the typical relevant problem is that of finding the polyominoes with minimal perimeter and given area. This problem has been widely studied by Alonso and Cerf [1] both in dimension two and three. In the following lemmas and in the corollary, we shall summarize some of the properties proven in that paper that we shall use in the sequel.

In order to state the lemma we need some preliminary definitions. Consider the lattice $\mathbb{Z}^2$ embedded in $\mathbb{R}^2$. Two sites of $\mathbb{Z}^2$ are said to be *nearest neighbor* if and only if their mutual Euclidean distance is equal to one. A *unit square* is a square of area one, whose center belongs to $\mathbb{Z}^2$ and whose vertices belong to its dual $\mathbb{Z}^2 + (1/2, 1/2)$. A *polyomino* is

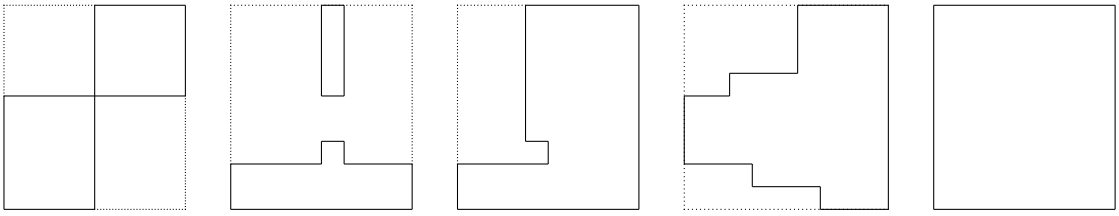

Figure 6.4: From the left to the right: two not connected convex polyominoes, a connected not convex polyomino, a connected convex polyomino which is not a convex subset of $\mathbb{R}^2$, a connected convex polyomino which is a convex subset of $\mathbb{R}^2$. The dotted lines denote the smallest surrounding rectangle.

a finite union of unit squares. Note that a polyomino is thus defined up to a translation as a subset of the plane $\mathbb{R}^2$ in which $\mathbb{Z}^2$ is embedded.

The *area of a polyomino* is the number of its squares. The *boundary of a polyomino* is the collection of unit edges of the dual lattice which belong only to one of the unit squares of the polyomino itself. The *perimeter of a polyomino* is the cardinality of its boundary, namely, the number of unit edges of the dual lattice which belong only to one of the unit squares of the polyomino itself. In other words the perimeter counts the number of interfaces on $\mathbb{Z}^2$ between the sites inside the polyomino and those outside.

A polyomino is *connected* if and only if its interior is a connected subset of $\mathbb{R}^2$. A polyomino is *convex* if and only if its intersection with any line parallel to the coordinate axes of $\mathbb{Z}^2$ is convex.

A polyomino is *monotone* if and only if its perimeter is equal to that of its smallest surrounding rectangle. A polyomino is *minimal* if and only if any other polyomino with the same area has perimeter greater or equal to that of the polyomino itself.

Lemma 6.16 *The following holds:*

1. *the perimeter of a connected polyomino is greater than or equal to that of its smallest surrounding rectangle;*

2. *a connected polyomino is convex if and only if it is monotone;*

3. *a minimal polyomino is connected and convex.*

The centers of the unit squares forming a connected polyomino are a nearest neighbor connected subset of $\mathbb{Z}^2$. As in [1] here a polyomino is not necessarily connected, different definitions, see for instance [6], can be found in the literature.

Proof of Lemma 6.16. Call *vertical* and *horizontal* the two directions defined by the coordinate axes of $\mathbb{Z}^2$. Define, also, the notion of *left*, *right*, *top*, and *bottom*. Let a *column* (resp. *row*) be the subset of $\mathbb{R}^2$ obtained by considering the union of all the unit squares whose center belong to the same vertical (resp. horizontal) line.

Item 1: since, by hypothesis, the polyomino is a connected subset of $\mathbb{R}^2$, each row and each column intersecting the smallest rectangle surrounding the polyomino intersects the boundary of the polyomino in at least two unit segments of the dual lattice. This remark yields the statement.

Item 2: by definition of convex polyominoes it follows that a connected polyomino is convex if and only if each vertical and each horizontal line passing through the sites of $\mathbb{Z}^2$ intersects the polyomino in a convex not empty subset of the line itself. Thus, a polyomino is convex if and only if each vertical and each horizontal line passing through the sites of $\mathbb{Z}^2$ intersects the boundary of the polyomino in exactly two unit segments of the dual lattice. This remark yields the statement.

Item 3: the fact that a minimal polyomino is connected is quite obvious. Indeed, it is sufficient to identify its maximal connected components and dispose them in a chain according to the following rule: excepted for the last component in the chain, the right part of the boundary of each component intersects the left part of the boundary of the component following it in the chain at least on a unit segment of the dual lattice. This construction yields a connected polyomino whose perimeter is smaller than the perimeter of the original polyomino by a quantity which at least equal to twice the number of maximal connected components minus one.

We now prove that a minimal polyomino is convex. Consider a minimal polyomino c and, by absurdity, assume it is not convex. We shall construct another polyomino with the same area and smaller perimeter. This will yield the statement by contradiction.

Since the polyomino is connected, by items 1 and 2 of this lemma we have that the perimeter P of the polyomino c is strictly greater than the perimeter R of its smallest surrounding rectangle r, that is to say $P > R$.

As a main tool in the sequel of the proof we shall use horizontal and vertical projections as those introduced in [1, Section 2, page 6], see also [19, Figure 8]. We first construct a second polyomino c_1 by projecting c vertically to its bottom, that is we consider the polyomino c_1 such that (i) the number of unit cell in each column is equal to that of the polyomino c, (ii) the intersection between c_1 and any column is a convex polyomino, and (iii) the boundary of the polyomino obtained by intersecting c_1 and any column intersects the bottom horizontal edge of r. We then construct a third polyomino c_2 by projecting c_1 horizontally to its left, that is we consider the polyomino c_2 such that (i) the number of unit cell in each row is

equal to that of the polyomino c_1, (ii) the intersection between c_2 and any row is a convex polyomino, and (iii) the boundary of the polyomino obtained by intersecting c_2 and any row intersects the left vertical edge of r.

We note, now, that c_2 is connected and convex and that the smallest rectangle r_2 surrounding c_2 is a subset of r. Since $P > R$, the perimeter P_2 of c_2 is equal to the perimeter of r_2, and $r_2 \subset r$, we have we have that $P > R \geq P_2$. Which is an absurd since the area of c is equal to that of c_2. □

Lemma 6.17 *For any n positive integer there exists two positive integers s and k, with $0 \leq k < s$, such that either (i) $n = s(s-1)+k$ or (ii) $n = s^2 + k$. The set of polyominoes of area n and minimal perimeter contains a rectangle of side lengths s and $s-1$ with a bar long k attached to one of its longest sides in the case (i) and a square of side length s with a bar long k attached to one of its sides in the case (ii).*

Proof of Lemma 6.17. The lemma is nothing but a simple restatement of Theorem 2.2 in [1]. □

Corollary 6.18 *For any n positive integer, the square of the perimeter of the polyominoes of area n is bounded from below by $16n$.*

Proof of Corollary 6.18. In the case (i) of Lemma 6.17 we have that the perimeter of the polyomino is $P = 4s$. Since the polyomino is contained in a square of side length s, we have that $n \leq s^2$. Hence $P^2 = 16s^2 \geq 16n$.

In the case (ii) of Lemma 6.17 we have that the perimeter of the polyomino is $P = 4s+2$. Since the polyomino is contained in a rectangle of side lengths s and $s+1$, we have that $n \leq s^2 + s$. Since

$$P^2 = (4s+2)^2 = 16(s^2+s) + 4 \geq 16n + 4$$

the proof of the corollary is completed. □

6.2. Metastable states of the Blume–Capel model

In this subsection we prove the Lemmas 4.8 and 4.9 and the Theorems 4.10 and 4.11 which allow to identify the metastable states of the Blume–Capel model and to deduce the exponential estimate of the exit time.

Proof of Lemma 4.8. We construct the following path. The path is defined by the following simple rule

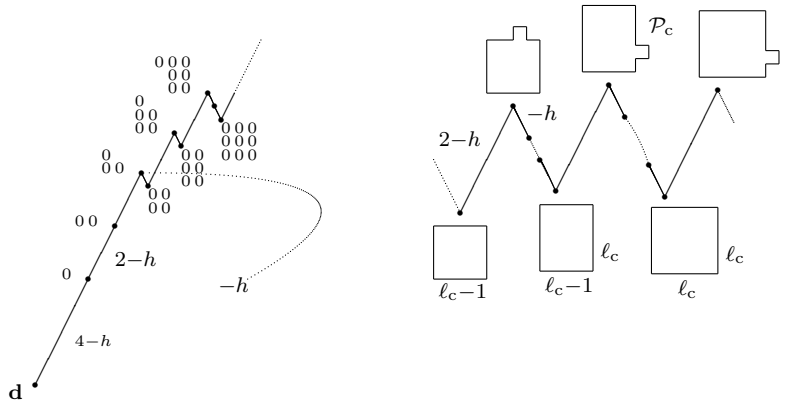

Figure 6.5: Schematic representation of path introduced at the beginning of the proof of Lemma 4.8. Only the first part of the path, the one going from $\mathbf{d}$ to few steps after $\mathcal{P}_c$, is shown.

- start from $\mathbf{d}$;
- transform one minus spin in zero;
- iteratively let the zero cluster grow until $\mathbf{0}$ is reached by adding a zero protuberance to the longest side of a rectangle and then filling the slice with zeros;
- transform one zero spin in plus;
- iteratively let the plus cluster grow until $\mathbf{u}$ is reached by adding a plus protuberance to the longest side of a rectangle and then filling the slice with pluses.

Note that the rectangles that are drawn during the above iterative procedure are squares or quasi–squares (rectangles whose side legths differ by one).

The path, see figure 6.5, starts at $\mathbf{d}$ and, passing through $\mathcal{P}_c$ reaches $\mathbf{0}$, then passes through $\mathcal{Q}_c$ and, finally, reaches $\mathbf{u}$. This path can be divided into two parts $\omega_1 \in \Omega(\mathbf{d}, \mathbf{0})$ connecting $\mathbf{d}$ to $\mathbf{0}$ and $\omega_2 \in \Omega(\mathbf{0}, \mathbf{u})$ connecting $\mathbf{0}$ to $\mathbf{u}$. By direct inspection, it is very easy to show that

$$\Phi_{\omega_1} - H(\mathbf{d}) = \Phi_{\omega_2} - H(\mathbf{0}) = \Gamma_c$$

and that ω_1 and ω_2 attain their maximal height respectively only in the states in $\mathcal{P}_c$ and $\mathcal{Q}_c$. This yields the lemma. □

Now, we introduce the set $\mathcal{X}_{\mathbf{d}} \subset \mathcal{X}$ as the set of states in $\mathcal{X}$ such that the number of minus spins is equal to

$$|\Lambda| - [\ell_c(\ell_c - 1) + 1]$$

In the following lemma we characterize the set $F(\mathcal{X}_{\mathbf{d}})$ of the minima of the energy of $\mathcal{X}_{\mathbf{d}}$.

Lemma 6.19 *Consider the zero chemical potential Blume–Capel model (4.2) and assume that Condition 4.7 is satisfied. Let $\sigma \in \mathcal{X}_{\mathbf{d}}$ and $N_\sigma \subset \Lambda$ be the set of sites $i \in \Lambda$ such that $\sigma(i) \neq -1$. We have that*

1. *the set N_σ is not a nearest neighbor connected subset of Λ winding around the torus Λ;*
2. *if $\sigma \in F(\mathcal{X}_{\mathbf{d}})$ then $\sigma(i) = 0$ for any $i \in N_\sigma$;*
3. *$\mathcal{X}_{\mathbf{d}} \supset \mathcal{P}_{\mathrm{c}}$;*
4. *$F(\mathcal{X}_{\mathbf{d}}) \supset \mathcal{P}_{\mathrm{c}}$;*
5. *$H(F(\mathcal{X}_{\mathbf{d}})) = H(\mathbf{d}) + \Gamma_{\mathrm{c}}$.*

Proof of Lemma 6.19. Item 1: recalling (4.4) we have that

$$|N_\sigma| = \ell_{\mathrm{c}}(\ell_{\mathrm{c}} - 1) + 1 < \Big(\frac{2}{h} + 1\Big)\frac{2}{h} + 1 = \frac{4}{h^2} + \frac{2}{h} + 1 \leq \frac{4}{h^2} + \frac{2}{h^2} + \frac{1}{h^2}$$

where in the last inequality we have used that, by Condition 4.7, we have $h < 1$. The item follows since, by Condition 4.7, we have that $|\Lambda| \geq 49/h^4$.

Item 2: let $\sigma \in \mathcal{X}_{\mathbf{d}}$ and let r, ℓ, m not negative integers such that r is the number of pluses in σ, ℓ is the number of plus–minus interfaces, and m is the number of plus–zero interfaces. Let σ' be the configuration obtained by flipping to zero the pluses in σ. By (4.2), it follows that

$$H(\sigma') - H(\sigma) = rh - 3\ell - m \leq rh - (\ell + m)$$

Let $C(\sigma)$ be the polyomino obtained by collecting all the unit squares associated with a plus spin of σ. Note that its perimeter is equal to $\ell + m$ and its area to r. Since the polyomino $C(\sigma)$ does not wind around the torus, we can apply Corollary 6.18 and obtain that $(\ell + m)^2 \geq 16r$. This bound on the number of interfaces between the pluses and the other two types of spins implies that

$$H(\sigma') - H(\sigma) \leq rh - 4\sqrt{r}$$

By studying the parabola of algebraic equation $y = hx^2 - 4x$ it is immediate to show that $H(\sigma') - H(\sigma) \leq 0$ provided $\sqrt{r} < 4/h$. This condition is easily proven to be true by using that $r \leq \ell_{\mathrm{c}}(\ell_{\mathrm{c}} - 1) + 1$. Indeed, by performing the same estimate as in the proof of item 1 above, we have that

$$r \leq \ell_{\mathrm{c}}(\ell_{\mathrm{c}} - 1) + 1 < \frac{7}{h^2}$$

which yields the desired bound $\sqrt{r} < 4/h$.

We can finally conclude that $H(\sigma') \leq H(\sigma)$, which completes the proof of the item. Note that the bound is not strict because, in principle, we could have no plus spin in σ, that is to say $r = 0$.

Item 3: since any configuration in $\mathcal{P}_c$ has $|\Lambda| - [\ell_c(\ell_c - 1) + 1]$ minus spins, it follows that $\mathcal{P}_c \subset \mathcal{X}_{\mathbf{d}}$.

Item 4: for any $\sigma \in F(\mathcal{X}_{\mathbf{d}})$ item 2 above states that the spins not equal to minus are equal to zero. Thus, $F(\mathcal{X}_{\mathbf{d}})$ is a subset $\mathcal{X}_{\mathbf{d},\mathbf{0}} \subset \mathcal{X}_{\mathbf{d}}$ made of those configurations such that all the spins that are not minus are zero. Consider $\sigma \in \mathcal{X}_{\mathbf{d},\mathbf{0}}$ and denote by $C(\sigma)$ the polyomino obtained by collecting all the unit squares associated with a spin zero. Remarked that the area of the polyomino $C(\sigma)$ is equal to $[\ell_c(\ell_c - 1) + 1]$, by (4.2) we have that

$$H(\sigma) - H(\mathbf{d}) = \text{Perimeter}(C(\sigma)) - h[\ell_c(\ell_c - 1) + 1]$$

This quantity is minimal if and only if the perimeter of the polyomino $C(\sigma)$ is minimal. We can then conclude that the configurations in $F(\mathcal{X}_{\mathbf{d}})$ are made of $[\ell_c(\ell_c - 1) + 1]$ zeros in the sea of minuses and that the polyomino obtained by collecting all the unit squares associated with a zero spin has area $[\ell_c(\ell_c - 1) + 1]$ and minimal perimeter.

Since item 1 above ensures that for any $\sigma \in \mathcal{X}_{\mathbf{d},\mathbf{0}}$ the polyomino $C(\sigma)$ does not wind around the torus, we can then apply Lemma 6.17 (case (i)) and conclude that $\mathcal{P}_c \subset F(\mathcal{X}_{\mathbf{d}})$.

Item 5: immediate consequence of item 4 above and the definition (4.6) of Γ_c. □

Proof of item 1 of Lemma 4.9. We have to compute $\Phi(\mathbf{d}, \mathbf{u}) - H(\mathbf{d})$ and $\Phi(\mathbf{0}, \mathbf{u}) - H(\mathbf{0})$. In order to show that they are both equal to Γ_c we show first that Γ_c is an upper bound for both. This can be done easily, indeed by Lemma 4.8 it follows

$$\Phi(\mathbf{d}, \mathbf{u}) - H(\mathbf{d}) \leq \Gamma_c \quad \text{and} \quad \Phi(\mathbf{0}, \mathbf{u}) - H(\mathbf{0}) \leq \Gamma_c \tag{6.1}$$

In order to get the lower bound the two problems have to be treated separately. We focus, first, on the problem $\Phi(\mathbf{d}, \mathbf{u})$ and proceed as follows.

Recall the definition of $\mathcal{X}_{\mathbf{d}}$ given at the beginning of this subsection. Since the connectivity matrix $q(\sigma, \eta)$ is different from zero only if the two configurations σ and η differ for the value of one single spin, we have that for any not negative integer n smaller than $|\Lambda|$, any path in $\Omega(\mathbf{d}, \mathbf{u})$ must necessarily visit the set of configurations with n minus spins. It then follows that any path in $\Omega(\mathbf{d}, \mathbf{u})$ must necessarily pass through $\mathcal{X}_{\mathbf{d}}$.

Item 5 in Lemma 6.19 states that $H(F(\mathcal{X}_{\mathbf{d}})) = H(\mathbf{d}) + \Gamma_c$. The above remarks imply that any path $\omega \in \Omega(\mathbf{d}, \mathbf{u})$ is such that $\Phi_\omega \geq H(F(\mathcal{X}_{\mathbf{d}})) = H(\mathbf{d}) + \Gamma_c$, which yields the desired lower bound $\Phi(\mathbf{d}, \mathbf{u}) - H(\mathbf{d}) \geq \Gamma_c$.

The lower bound $\Phi(\mathbf{0},\mathbf{u}) - H(\mathbf{0}) \geq H(\mathcal{Q}_{\rm c}) - H(\mathbf{0}) = \Gamma_{\rm c}$ can be achieved, "mutatis mutandis," with the same argument. We do not enter into the details, we just remark that the manifold $\mathcal{X}_{\mathbf{d}}$ must be replaced by $\mathcal{X}_{\mathbf{0}}$ defined as the set of configurations in which all the spins are zeros excepted for $\ell_{\rm c}(\ell_{\rm c}-1)+1$ which are pluses. □

Before starting the proof of item 2 of Lemma 4.9 we need to state a technical lemma on the zero chemical potential Blume–Capel Hamiltonian ensuring that the energy (4.2) of a configuration is decreased if a minus spin having at most two minuses between its nearest neighbors or a minus spin having three neighboring minus and a neighboring plus is flipped to zero. Moreover, the lemma states that the energy of a configuration is decreased if a zero spin with at least two pluses and no minus among its nearest neighbors is flipped to plus.

For any $\sigma \in \mathcal{X}$, $i \in \Lambda$, and $a \in \{-1,0,+1\}$ we let $\sigma^{i,a}$ be the configuration such that $\sigma^{i,a}(j) = \sigma(j)$ for all $j \in \Lambda$ and $j \neq i$ and $\sigma^{i,a}(i) = a$. Note that $\sigma^{i,a}$ differs from σ by at most the value of the spin associated with the site i.

Lemma 6.20 *Consider the zero chemical potential Blume–Capel model (4.2) and assume that Condition 4.7 is satisfied. Let $\sigma \in \mathcal{X}$*

1. *if there exists a site i such that $\sigma(i) = -1$, three of the nearest neighbors of i have associated spin equal to -1, and the fourth has associated spin equal to $+1$, then*

$$H(\sigma) - H(\sigma^{i,0}) = h$$

2. *if there exists a site i such that $\sigma(i) = -1$ and there are at most two nearest neighbor of i such that the associated spin in σ is -1, then*

$$H(\sigma) - H(\sigma^{i,0}) \geq h$$

3. *if there exists a site i such that $\sigma(i) = 0$, there are at most two nearest neighbors of i such that the associated spins are 0, and the remaining nearest neighbors have associated spin equal to $+1$, then*

$$H(\sigma) - H(\sigma^{i,+1}) \geq h$$

Proof of Lemma 6.20. Items 1 and 2: let $n_+, n_-, n_0 \geq 0$ be the number of plus, minus, and zero spins, respectively, among the four nearest neighbors of the site i. Note that $n_+ + n_- + n_0 = 4$. From (4.2), use also that here $\sigma(i) = -1$, it follows immediately that

$$H(\sigma) - H(\sigma^{i,0}) = 4n_+ + n_0 + h - [n_- + n_+] = h + 3n_+ - n_- + n_0$$

By using that $n_+ + n_0 = 4 - n_-$ we get

$$H(\sigma) - H(\sigma^{i,0}) = h + 4 + 2(n_+ - n_-)$$

Item 1 follows by the formula above with $n_- = 3$ and $n_+ = 1$; item 2 follows by the formula above by noticing that $n_- \leq 2$ and $n_+ \geq 0$.

Item 3: let $n_+, n_0 \geq 0$ be the number of plus and zero spins, respectively, among the four nearest neighbors of the site i. Note that, by hypothesis, $n_+ + n_0 = 4$. From (4.2), use also that here $\sigma(i) = 0$, it follows immediately that

$$H(\sigma) - H(\sigma^{i,+1}) = n_+ - [n_0 - h] = 4 - 2n_0 + h$$

where we have used $n_+ = 4 - n_0$. The item finally follows by noticing that, by hypothesis, $n_0 \leq 2$. □

Proof of item 2 of Lemma 4.9. Let $\sigma \in \mathcal{X} \setminus \{\mathbf{d}, \mathbf{0}, \mathbf{u}\}$, we have to prove that $V_\sigma < \Gamma_c$; that is we have to find a configuration at energy smaller than $H(\sigma)$ and a path connecting σ to such a configuration such that its height is smaller than $H(\sigma) + \Gamma_c$.

We first assume that σ is such that at least one spin is equal to minus one and we distinguish several different sub–cases.

Case 1. There exists a plus–minus interface in σ, that is there exist $i, j \in \Lambda$ nearest neighbors such that $\sigma(i) = -1$ and $\sigma(j) = +1$. From items 1 and 2 in Lemma 6.20 we have that $H(\sigma) - H(\sigma^{i,0}) \geq h$. It then follows $H(\sigma) > H(\sigma^{i,0})$ and, hence, $V_\sigma = 0$.

Case 2. No plus–minus interface exist in σ and there exist a site $i \in \Lambda$ such that $\sigma(i) = -1$ and at least two of its nearest neighbors are occupied by spin zero. From item 2 in Lemma 6.20 we have that $H(\sigma) - H(\sigma^{i,0}) \geq h$. It then follows $H(\sigma) > H(\sigma^{i,0})$ and, hence, $V_\sigma = 0$.

Case 3. No plus–minus interface exist in σ and for any $i \in \Lambda$ such that $\sigma(i) = -1$ there is at most one nearest neighbor occupied by spin zero. Let γ be the collection of the unit segments of the dual of Λ separating two sites with which are associated a minus and a zero spin; call those sites *adjacent* to γ. This collection γ is made of maximal connected components which are either rectangles or straight annuli winded around the torus. Note that by the characterization of σ given in this case 3 we have that each minus associated with a site adjacent to γ has one single zero (the one associated to the neighbor adjacent to γ) among its nearest neighbors.

Case 3.1. Suppose that one of the connected component of γ is a straight annulus α winding around the torus. The annulus separates a stripe of zeros by a stripe of minuses (both winding around the torus). Since each minus associated with a site adjacent to α has one

single zero among its nearest neighbors, it follows that the minus stripe occupies at least two lattice rows or columns. Then, let σ' be the configuration obtained by flipping to zero all the minuses of σ associated with sites adjacent to α and note that, by (4.2),

$$H(\sigma') - H(\sigma) = -hL \Rightarrow H(\sigma') < H(\sigma)$$

where we recall that L is the side length of Λ. Moreover, consider the $L+1$ long path connecting σ to σ' constructed by flipping to zero first one of the minus spin associated with a site adjacent to α and then the remaining $L-1$ minuses so that at each step one minus with two zeros among its nearest neighbors is flipped. The height of this path is $H(\sigma)+2-h$. Since $\Gamma_{\rm c} > 2-h$ we have that $V_\sigma < \Gamma_{\rm c}$.

Case 3.2. Suppose that none of the connected component of γ is a straight annulus winding around the torus and assume that there exists a rectangular component such that one of its four sides, say ρ, has length $|\rho|$ larger or equal to $\ell_{\rm c}$. Let σ' be the configuration obtained by flipping to zero all the minuses of σ associated with sites adjacent to ρ. Since each minus associated with a site adjacent to ρ has one single zero among its nearest neighbors, by (4.2), we have that

$$H(\sigma') - H(\sigma) = 2 - h|\rho| \leq 2 - h\ell_{\rm c} < 2 - h\frac{2}{h} \Rightarrow H(\sigma') < H(\sigma)$$

where we have used that $\ell_{\rm c} > 2/h$ (recall (4.3)). Moreover, consider the $|\rho|+1$ long path connecting σ to σ' constructed by flipping to zero first one of the minus spin associated with a site adjacent to ρ and then the remaining $|\rho|-1$ minuses so that at each step one minus with two zeros among its nearest neighbors is flipped. The height of this path is $H(\sigma)+2-h$. Since $\Gamma_{\rm c} > 2-h$ we have that $V_\sigma < \Gamma_{\rm c}$.

Case 3.3. Suppose that none of the connected component of γ is a straight annulus winding around the torus and assume that there exists a rectangular component such that one of its four sides, say ρ, has length $|\rho| = 1$. Let σ' be the configuration obtained by flipping to minus the zero of σ associated with sites adjacent to ρ. By considering the two possible situations, that is to say the 1×1 square and the $1 \times \ell$ rectangle, by (4.2) we have that

$$H(\sigma') - H(\sigma) \leq \max\{-4+h, -2+h\} = -2+h < 0 \Rightarrow H(\sigma') < H(\sigma)$$

where we have used that $h < 1$. Since σ is transformed in σ' in one step, we have that $V_\sigma = 0 < \Gamma_{\rm c}$.

Case 3.4. Suppose that none of the connected component of γ is a straight annulus winding around the torus and assume that there exists a rectangular component such that one of

its four sides, say ρ, has length $|\rho| = 2$. Let σ' be the configuration obtained by flipping to minus the two zeros of σ associated with sites adjacent to ρ. By (4.2) we have that

$$H(\sigma') - H(\sigma) = -2 + 2h < 0 \Rightarrow H(\sigma') < H(\sigma)$$

where we have used that $h < 1$. Moreover, consider the path long three connecting σ to σ' constructed by flipping to minus the two zero spins associated with the sites adjacent to ρ. The height of this path is $H(\sigma) + h$. Since $\Gamma_c > h$ we have that $V_\sigma < \Gamma_c$.

Case 3.5. Suppose that none of the connected component of γ is a straight annulus winding around the torus and assume that all the rectangular components have side lengths larger or equal to three and smaller or equal to $\ell_c - 1$. Note that this case can be empty if $\ell_c = 3$. Consider one of this rectangular component and let ρ be one of its four sides; note that $3 \leq |\rho| \leq \ell_c - 1$. For the sake of clearness assume that ρ is vertical.

Case 3.5.1. Suppose that the column of sites at distance one from those occupied by zeros and adjacent to ρ (in other less precise words the "second" column inside the rectangle starting from ρ) is occupied only by spin zero. Let σ' be the configuration obtained by flipping to minus the $|\rho|$ zeros of σ associated with sites adjacent to ρ. By (4.2) we have that

$$H(\sigma') - H(\sigma) = -2 + h|\rho| \leq -2 + h(\ell_c - 1) < -2 + h\frac{2}{h} < 0 \Rightarrow H(\sigma') < H(\sigma)$$

where we have used that $|\rho| \leq \ell_c - 1$. Moreover, consider the $|\rho| + 1$ long path connecting σ to σ' constructed by flipping to minus, one after the others, the zero spins associated with the sites adjacent to ρ and by flipping, in the first $|\rho| - 1$ steps, one of the two zeros with two minuses among its four nearest neighbors; in the final step the remaining spin zero is finally flipped to minus. The height of this path is $H(\sigma) + h(|\rho| - 1)$. Since $\Gamma_c > h(\ell_c - 2)$ we have that $V_\sigma < \Gamma_c$.

Case 3.5.2. Suppose that the column of sites at distance one from those occupied by zeros and adjacent to ρ (in other less precise words the "second" column inside the rectangle starting from ρ) is occupied at least by one spin plus. Let σ' be the configuration obtained by flipping to zero the pluses of σ associated with sites at distance one from those adjacent to ρ and occupied by zeros.

Case 3.5.2.1. Suppose $|\rho| - 2 = 1$. The configuration σ is transformed into σ' by flipping to zero a plus spin having at least three zeros among its nearest neighbors and no minus. It then follows that

$$H(\sigma') - H(\sigma) \leq \max\{-4 + h, -2 + h\} = -2 + h < 0 \Rightarrow H(\sigma') < H(\sigma)$$

where we have used that $h < 1$. Since σ is transformed into σ' in one step we have that $V_\sigma < 0 < \Gamma_c$.

Case 3.5.2.2. Suppose $|\rho| - 2 \geq 2$. The configuration σ is transformed into σ' by flipping to zero at most $|\rho| - 3$ plus spin having two zeros and two pluses as nearest neighbors and at least one plus spin having no minus and at most one plus among its nearest neighbors. It then follows that

$$H(\sigma') - H(\sigma) \leq \max\{-4 + h, -2 + h\} + (|\rho| - 3)h \leq -2 + (|\rho| - 2)h$$

Hence, recalling that $|\rho| \leq \ell_{\rm c} - 1$, we have

$$H(\sigma') - H(\sigma) \leq -2 + (\ell_{\rm c} - 3)h < -2 + \frac{2}{h}\, h - 2h < -2h \Rightarrow H(\sigma') < H(\sigma)$$

Moreover, consider the path connecting σ to σ' constructed by flipping to zero, one after the others, the plus spins with two neighboring pluses and then those with at most one neighboring plus. The height of this path is smaller that $H(\sigma)+h(|\rho|-3)$. Since $\Gamma_{\rm c} > h(\ell_{\rm c}-4)$ we have that $V_\sigma < \Gamma_{\rm c}$.

In order to complete the proof, we finally have to consider those configurations $\sigma \in \mathcal{X} \setminus \{\mathbf{d}, \mathbf{0}, \mathbf{u}\}$ such that none of the spins is minus. The proof can be achieved with arguments very similar to those developed above; we just sketch the idea. First of all we consider the collection of the unit segments of the dual of Λ separating two sites with which are associated a zero and a plus spin. From item 3 in Lemma 6.20 it follows that this collection is made of maximal connected components which are either rectangles or straight annuli winded around the torus. We then have to consider all the cases analogous to those taken into account above. □

Proof of Theorem 4.10. The theorem follows from Lemma 4.9 and Theorem 2.4. □

Proof of Theorem 4.11. The theorem follows from Theorem 4.10 above and [22, Theorems 4.1 and 4.9]. □

6.3. Escape mechanism

In this subsection we prove Lemma 4.12 and Theorem 4.13 allowing the identification of the escape mechanism as the nucleation of the critical droplet.

Proof of Lemma 4.12. Item 1: the proof of this item is divided into two parts. In the first part we characterize the set $F(\mathcal{X}_{\mathbf{d}})$ of the minima of the energy of the set $\mathcal{X}_{\mathbf{d}} \subset \mathcal{X}$ defined at the beginning of Subsection 6.2. In the second part we show that any path connecting $\mathbf{d}$ to $\mathbf{u}$ must necessarily pass through $\mathcal{P}_{\rm c}$.

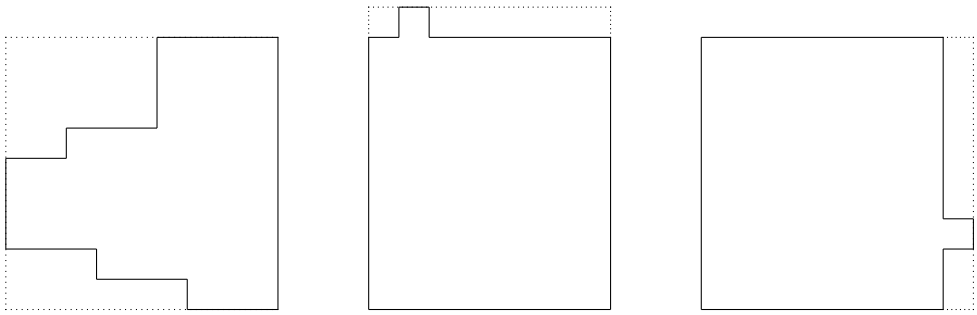

Figure 6.6: Typical configuration of $F_1(\mathcal{X}_\mathbf{d})$ (left) and $F_2(\mathcal{X}_\mathbf{d})$ (center and right). Zeros are associated with the sites inside the solid lines and minuses outside.

First part of the proof. From items 1 and 2 of Lemma 6.19 it follows that for any $\sigma \in F(\mathcal{X}_\mathbf{d})$ the polyomino $C(\sigma)$ associated with the zero spins does not wind around the torus. We can then apply item 3 in Lemma 6.16 and deduce that $C(\sigma)$ is convex. Thus, we partition the set $F(\mathcal{X}_\mathbf{d})$ into two disjoint subsets $F_1(\mathcal{X}_\mathbf{d})$ and $F_2(\mathcal{X}_\mathbf{d})$ defined as follows (see, also, figure 6.6): the set $F_1(\mathcal{X}_\mathbf{d})$ is the set of configurations $\sigma \in F(\mathcal{X}_\mathbf{d})$ such that the boundary of the polyomino $C(\sigma)$ intersects each side of the boundary of its smallest surrounding rectangle on a set of the dual lattice $\mathbb{Z}^2 + (1/2, 1/2)$ made by at least two pairwise consecutive unit segments. The set $F_2(\mathcal{X}_\mathbf{d})$ is the set of configurations $\sigma \in F(\mathcal{X}_\mathbf{d})$ such that the boundary of the polyomino $C(\sigma)$ intersects at least one of the four sides of the boundary of its smallest surrounding rectangle on a set made of a single unit segments. Note that, since $C(\sigma)$ is convex, there cannot exist multiple unit protuberances intersecting the same side of the smallest surrounding rectangle.

In figure 6.6 the two configurations in $F_2(\mathcal{X}_\mathbf{d})$ have been represented by a rectangle of zeros with a unit zero protuberance placed either on the longest (picture on the right) or on the shortest side (center picture). In principle other situations should be taken into account, indeed the boundary of the polyomino associated with the zero component of the considered configuration could intersect the other three sides of the boundary of its smallest surrounding rectangle in proper subsets of the side itself. In other words the rectangle on which the unit protuberance is placed could be not fully occupied by zeros.

We prove, now, that the two situations depicted in the picture indeed cover all the possible cases. Consider $\sigma \in F_2(\mathcal{X}_\mathbf{d})$. Let $\ell_\mathrm{c} + k$ and $\ell_\mathrm{c} + s$ be the side lengths of the smallest rectangle surrounding the polyomino $C(\sigma)$ associated with σ with $k, s \in \mathbb{Z}$. Assume, also, that a single protuberance U is placed on the side of length $\ell_\mathrm{c} + s$. Let $R(\sigma)$ be the smallest rectangle surrounding the polyomino obtained by removing the unit cell U from $C(\sigma)$; note

that $R(\sigma)$ is a rectangle with side lengths $\ell_c + s$ and $\ell_c + k - 1$. Since the polyomino $C(\sigma)$ is convex, we have that its perimeter is equal to that of the smallest surrounding rectangle, namely,

$$\text{Perimeter}(C(\sigma)) = 2(\ell_c + k) + 2(\ell_c + s) = 4\ell_c + 2(k + s)$$

Moreover, from item 4 in Lemma 6.19 it follows that $\text{Perimeter}(C(\sigma)) = 4\ell_c$ and, hence, $k + s = 0$.

Note, now, that the area of the rectangle $R(\sigma)$ is

$$(\ell_c + s)(\ell_c + k - 1) = \ell_c(\ell_c - 1) + (k + s)\ell_c + ks - s = \ell_c(\ell_c - 1) - s^2 - s$$

where we have used $k + s = 0$. Since the area of the polyomino $C(\sigma)$ is $\ell_c(\ell_c - 1) + 1$, the number of zeros inside $R(\sigma)$ is $\ell_c(\ell_c - 1)$. Hence the area of $R(\sigma)$ must be larger or equal to $\ell_c(\ell_c - 1)$; it then follows that

$$\ell_c(\ell_c - 1) - s^2 - s \geq \ell_c(\ell_c - 1) \Rightarrow -s^2 - s \geq 0$$

Since s is an integer, the above inequality is satisfied only for $s = 0$ and $s = -1$; note that in both cases the inequality is indeed an equality. We can then give a full characterization of the set $F_2(\mathcal{X}_\mathbf{d})$. Indeed we can write

$$F_2(\mathcal{X}_\mathbf{d}) = F_{2,a}(\mathcal{X}_\mathbf{d}) \cup F_{2,b}(\mathcal{X}_\mathbf{d})$$

with $F_{2,a}(\mathcal{X}_\mathbf{d}) = \mathcal{P}_c$ and $F_{2,b}(\mathcal{X}_\mathbf{d})$ the set of configurations of $F(\mathcal{X}_\mathbf{d})$ such that the zeros occupy a rectangle of side lengths $\ell_c - 1$ and ℓ_c and a unit protuberance placed on one of the two shortest sides.

Second part of the proof. We prove, now, that any path connecting $\mathbf{d}$ to $\mathbf{u}$ necessarily pass through $\mathcal{P}_c$. The proof is organized in several steps.

Step 0: for any path $\omega \in \Omega(\mathbf{d}, \mathbf{u})$ there exists a positive integer i such that $\omega_i \in \mathcal{X}_\mathbf{d}$. See the proof (lower bound) of item 1 of Lemma 4.9.

Step 1: for any path $\omega \in \Omega(\mathbf{d}, \mathbf{u})$ such that $\Phi_\omega - H(\mathbf{d}) = \Gamma_c$ there exists a positive integer i such that $\omega_i \in F(\mathcal{X}_\mathbf{d})$. This property follows from step 0 and item 4 in Lemma 6.19.

Step 2: for any path $\omega \in \Omega(\mathbf{d}, \mathbf{u})$ such that $\Phi_\omega - H(\mathbf{d}) = \Gamma_c$ we denote by $f(\omega)$ the set of positive integers such that for any $i \in f(\omega)$ one has that $\omega_i \in F(\mathcal{X}_\mathbf{d})$ and ω_{i-1} is made of $|\Lambda| - \ell_c(\ell_c - 1)$ minus spins and $\ell_c(\ell_c - 1)$ zeros. From step 1 above it follows that $f(\omega)$ is not empty.

Step 3: for any path $\omega \in \Omega(\mathbf{d}, \mathbf{u})$ such that $\Phi_\omega - H(\mathbf{d}) = \Gamma_c$ one has that $\omega_i \in F_2(\mathcal{X}_\mathbf{d})$ for any $i \in f(\omega)$. Indeed, assume by absurdity that that there exists $i \in f(\omega)$ such that $\omega_i \in F_1(\mathcal{X}_\mathbf{d})$. By definition of $f(\omega)$ the configuration ω_i is transformed into ω_{i-1} by flipping

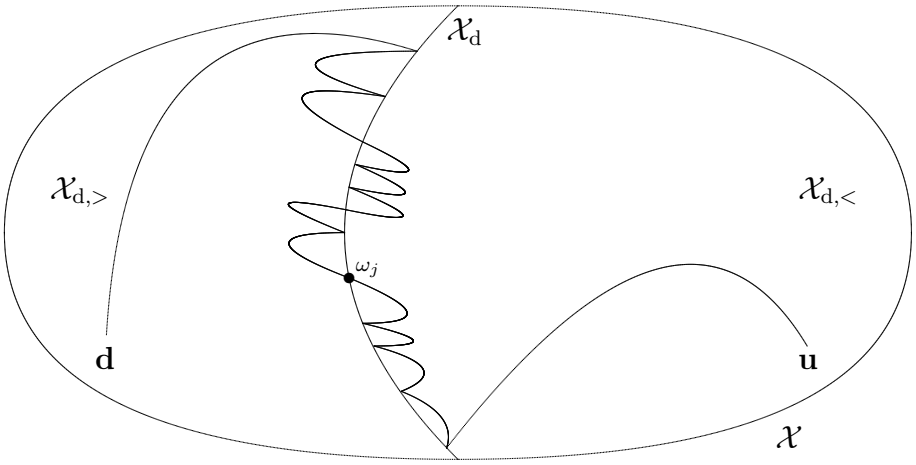

Figure 6.7: Paths constructed in the second part of the proof of Lemma 4.12.

to minus one of the zeros in ω_i. From the definition of $F_1(\mathcal{X}_{\mathbf{d}})$ it follows that all the zero spins in ω_i have at least two zeros among their nearest neighbors. Then by (4.2) it follows easily that $H(\omega_{i-1}) > H(\omega_i)$. Hence,

$$\Phi_\omega \geq H(\omega_{i-1}) > H(\omega_i) = H(F(\mathcal{X}_{\mathbf{d}})) = H(\mathbf{d}) + \Gamma_{\rm c}$$

which is an absurd. We can then conclude that $\omega_i \in F_2(\mathcal{X}_{\mathbf{d}})$ for any $i \in f(\omega)$.

Step 4: for any path $\omega \in \Omega(\mathbf{d}, \mathbf{u})$ such that $\Phi_\omega - H(\mathbf{d}) = \Gamma_{\rm c}$, if i is a positive integer such that $\omega_i \in F(\mathcal{X}_{\mathbf{d}})$ then $\omega_{i-1}, \omega_{i+1} \notin \mathcal{X}_{\mathbf{d}}$. Indeed, assume by absurdity that $\omega_{i+1} \in \mathcal{X}_{\mathbf{d}}$. Then ω_{i+1} is necessarily obtained by ω_i by flipping to plus a zero spin, otherwise the number of minus spins would be changed (recall that all the configurations in $\mathcal{X}_{\mathbf{d}}$ have the same number of minuses). Since any zero spin of ω_i has no plus among its nearest neighbors, from (4.2) it follows that $H(\omega_{i+1}) > H(\omega_i)$. Hence,

$$\Phi_\omega \geq H(\omega_{i+1}) > H(\omega_i) = H(F(\mathcal{X}_{\mathbf{d}})) = H(\mathbf{d}) + \Gamma_{\rm c}$$

which is an absurd. We can then conclude that $\omega_{i+i} \notin \mathcal{X}_{\mathbf{d}}$. Similarly we also prove that $\omega_{i-i} \notin \mathcal{X}_{\mathbf{d}}$.

Step 5: consider a path $\omega \in \Omega(\mathbf{d}, \mathbf{u})$ such that $\Phi_\omega - H(\mathbf{d}) = \Gamma_{\rm c}$ and assume, by absurdity, that it does not pass through $\mathcal{P}_{\rm c}$, that is to say $\omega_i \notin \mathcal{P}_{\rm c}$ for any integer i.

From step 4 (see figure 6.7) it follows that there exists n not consecutive integers $i_1 < i_2 < \cdots < i_{n-1} < i_n$ such that $\omega_{i_k} \in F(\mathcal{X}_{\mathbf{d}})$, for $k = 1, \ldots, n$. If $n \geq 2$, each sub–path $(\omega_{i_k+1}, \ldots, \omega_{i_{k+1}-1})$, for $k = 1, \ldots, n-1$, belongs either to the subset $\mathcal{X}_{\mathbf{d},<}$ of $\mathcal{X}$ made of all

the configurations with number of minus spins smaller or equal to $|\Lambda| - [\ell_c(\ell_c - 1) + 1] - 1$ or to the subset $\mathcal{X}_{\mathbf{d},>}$ of $\mathcal{X}$ made of all the configurations with number of minus spins larger or equal to $|\Lambda| - [\ell_c(\ell_c - 1) + 1] + 1$.

Let j be the smallest integer in $\{1, \dots, n\}$ such that the path $(\omega_{i_j}, \omega_{i_j+1}, \dots, \mathbf{u})$ does not visit $\mathcal{X}_{\mathbf{d},>}$. Since $\omega_{i_j-1} \in \mathcal{X}_{\mathbf{d},>}$, from step 3 and from the absurd hypothesis it follows that $\omega_{i_j} \in F_{2,b}(\mathcal{X}_{\mathbf{d}})$.

Since $H(\omega_{i_j}) = H(\mathbf{d}) + \Gamma_c$, starting from ω_{i_j} the path must necessarily decrease the energy. Since it is not possible to flip to minus the unit zero protuberance (otherwise the path would enter $\mathcal{X}_{\mathbf{d},>}$), the only move that decreases the energy is flipping to zero a minus with two zeros among its nearest neighbors. So that $H(\omega_{i_j+1}) = H(\mathbf{d}) + \Gamma_c - h$.

Starting from ω_{i_j+1} only moves which increase the energy of at most h are allowed. So that the sole possible moves are: flipping to zero a minus with two zeros among its nearest neighbors or flipping to minus a zero with two minuses among its nearest neighbors.

With these types of moves only configurations σ such that the polyomino $C(\sigma)$ associated with the zero spin component is convex and its smallest surrounding rectangle has side lengths $\ell_c + 1$ and $\ell_c - 1$. In this set the smallest energy is that of the configuration in which $C(\sigma)$ is precisely the rectangle with side lengths $\ell_c + 1$ and $\ell_c - 1$.

Besides the already mentioned moves, the one to which competes the smallest energy increase $2 - h$ is flipping to zero a minus with one zero and three minuses among its nearest neighbors. Thus, we have that

$$\Phi_\omega - H(\mathbf{d}) \geq 2(\ell_c + 1) + 2(\ell_c - 1) - h(\ell_c + 1)(\ell_c - 1) + 2 - h$$

Hence

$$\Phi_\omega - H(\mathbf{d}) \geq 4\ell_c - h[\ell_c(\ell_c - 1) + 1] - h(\ell_c - 1) + 2 > \Gamma_c$$

where, in the last inequality, we have used (4.5), (4.6), and the bound

$$2 - h(\ell_c - 1) > 2 - h\frac{2}{h} > 0$$

where we recalled the upper bound (4.4).

We finally got an absurd. We then have that the path ω has to visit $\mathcal{P}_c$. This completes the proof of item 1 of the lemma.

Item 2: the proof can be achieved, "mutatis mutandis," with the same arguments used in the proof of item 1. We do not enter into the details, we just remark that the manifold $\mathcal{X}_{\mathbf{d}}$ must be replaced by $\mathcal{X}_{\mathbf{0}}$ defined as the set of configurations in which all the spins are zeros excepted for $\ell_c(\ell_c - 1) + 1$ which are pluses. □

Proof of Theorem 4.13. In the proof of this theorem we use results in [22], in particular the definition of gate on page 603 and the Theorem 5.4. The Lemma 4.12 in Section 4 above

implies that $\mathcal{P}_c$ is a gate for the pair of configurations $\mathbf{d}$ and $\mathbf{u}$ and $\mathcal{Q}_c$ is a gate for the pair of configurations $\mathbf{0}$ and $\mathbf{u}$. Then the theorem follows from Theorem 5.4 in [22]. □

Proof of Theorem 4.14. As stated in item 1 of Lemma 4.8 (see also figure 6.5), there exists a path joining $\mathbf{d}$ to $\mathbf{u}$ passing through any state in $\mathcal{P}_c$ and attaining its maximal height only in this point. This implies that any subset of $\mathcal{X}$ not containing $\mathcal{P}_c$ is not a gate, in particular it is not minimal.

The second statement of the theorem can be prove similarly. □

References

[1] L. Alonso, R. Cerf, "The three dimensional polyominoes of minimal area." *The Electronic Journal of Combinatorics* **3**, #R27 (1996).

[2] J. Beltrán, C. Landim, "Tunneling and Metastability of Continuous Time Markov Chains." *Journ. Stat. Phys.* **140**, 1065–1114 (2010).

[3] S. Bigelis, E.N.M. Cirillo, J.L. Lebowitz, E.R. Speer, "Critical droplets in metastable probabilistic cellular automata," *Phys. Rev. E* **59**, 3935 (1999).

[4] M. Blume, "Theory of the First–Order Magnetic Phase Change in UO_2." *Phys. Rev.* **141**, 517 (1966).

[5] M. Blume, V.J. Emery, R.B. Griffiths, "Ising Model for the λ Transition and Phase Separation in He^3–He^4 Mixtures." *Phys. Rev. A* **4**, 1071–1077 (1971).

[6] M. Bousquet–Mélou, J.M. Fédou, "The generating function of convex polyominoes: the resolution of a q–differential system." *Discrete Mathematics* **137**, 53–75 (1995).

[7] A. Bovier, "Metastability: a potential theoretic approach." Proceedings of ICM 2006, EMS Publishing House, 2006, pp. 499–518.

[8] A. Bovier, M. Eckhoff, V. Gayrard, M. Klein, "Metastability and low lying spectra in reversible Markov chains." *Comm. Math. Phys.* **228**, 219–255 (2002).

[9] A. Bovier, F. den Hollander, and F.R. Nardi, "Sharp asymptotics for Kawasaki dynamics on a finite box with open boundary." *Probab. Theory Relat. Fields* **135**, 265–310 (2006).

[10] A. Bovier, F. Manzo, "Metastability in Glauber dynamics in the low–temperature limit: beyond exponential asymptotics." *Journ. Stat. Phys.* **107**, 757–779 (2002).

[11] H.W. Capel, "On possibility of first–order phase transitions in ising systems of triplet ions with zero–field splitting." *Physica* **32**, 966 (1966).

[12] E.N.M. Cirillo, F.R. Nardi, "Metastability for the Ising model with a parallel dynamics." *Journ. Stat. Phys.* **110**, 183–217 (2003).

[13] E.N.M. Cirillo, F.R. Nardi, A.D. Polosa, "Magnetic order in the Ising model with parallel dynamics." *Phys. Rev. E* **64**, 57103 (2001).

[14] E.N.M. Cirillo, F.R. Nardi, C. Spitoni, "Metastability for reversible probabilistic cellular automata with self–interaction." *Journ. Stat. Phys.* **132**, 431–471 (2008).

[15] E.N.M. Cirillo, F.R. Nardi, C. Spitoni, "Competitive nucleation in reversible Probabilistic Cellular Automata." *Phys. Rev. E* **78**, 040601 (2008).

[16] E.N.M. Cirillo, F.R. Nardi, C. Spitoni, "Competitive nucleation in metastable systems." Applied and Industrial Mathematics in Italy III, Series on Advances in Mathematics for Applied Sciences, Vol. 82, Pages 208-219, 2010.

[17] E.N.M. Cirillo, E. Olivieri, "Metastability and nucleation for the Blume-Capel model. Different mechanisms of transition," *Journ. Stat. Phys.* **83**, 473–554 (1996).

[18] R.J. Glauber, "Time-Dependent Statistics of the Ising Model." *J. Math. Phys.* **4**, 294 (1963).

[19] F. den Hollander, F.R. Nardi, A. Troiani, "Kawasaki Dynamics with Two Types of Particles: Stable/Metastable Configurations and Communication Heights." *Journ. Stat. Phys.* **145**, 1423–1457 (2011).

[20] F. den Hollander, F.R. Nardi, A. Troiani, " Metastability for Kawasaki dynamics at low temperature with two types of particles." *Electronic Journal of Probability* **17**, 1–26 (2012).

[21] G. Grinstein, C. Jayaprakash, Y. He, "Statistical Mechanics of Probabilistic Cellular Automata." *Phys. Rev. Lett.* **55**, 2527–2530 (1985).

[22] F. Manzo, F.R. Nardi, E. Olivieri, E. Scoppola, "On the essential features of metastability: tunnelling time and critical configurations." *Journ. Stat. Phys.* **115**, 591–642 (2004).

[23] F. Manzo, E. Olivieri, "Dynamical Blume–Capel model: Competing metastable states at infinite volume." *Journ. Stat. Phys.* **104**, 1029–1090 (2001).

[24] F.R. Nardi, C. Spitoni, "Sharp Asymptotics for Stochastic Dynamics with Parallel Updating Rule." *Journ. Stat. Phys.* **146**, 701–718 (2012).

[25] E. Olivieri, E. Scoppola, "Markov chains with exponentially small transition probabilities: First exit problem from a general domain. I. The reversible case," *Journ. Stat. Phys.* **79**, 613–647 (1995).

[26] D.J. Wales, Energy landscapes. Cambridge University Press, Cambridge, UK, 2003.