

# Energy landscape and metastability of stochastic Ising and Potts models on three-dimensional lattices without external fields\*

Seonwoo Kim<sup>†</sup>      Insuk Seo<sup>‡</sup>

## Abstract

In this study, we investigate the energy landscape of the Ising and Potts models on fixed and finite but large three-dimensional (3D) lattices where no external field exists and quantitatively characterize the metastable behavior of the associated Glauber dynamics in the very low temperature regime. Such analyses for the models with non-zero external magnetic fields have been extensively performed over the past two decades; however, models without external fields remained uninvestigated. Recently, the corresponding investigation has been conducted for the two-dimensional (2D) model without an external field, and in this study, we further extend these successes to the 3D model, which has a far more complicated energy landscape than the 2D one. In particular, we provide a detailed description of the highly complex plateau structure of saddle configurations between ground states and then analyze the typical behavior of the Glauber dynamics thereon. Thus, we achieve a quantitatively precise analysis of metastability, including the Eyring–Kramers law, the Markov chain model reduction, and a full characterization of metastable transition paths.

**Keywords:** energy landscape; metastability; Ising model; Potts model; three dimension; Eyring–Kramers law; Markov chain model reduction; transition paths.

**MSC2020 subject classifications:** Primary 60J28, Secondary 60K35; 82C22.

Submitted to EJP on October 21, 2023, final version accepted on March 1, 2024.

---

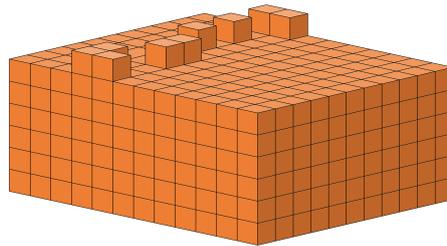
\*S.K. was supported by NRF-2019-Fostering Core Leaders of the Future Basic Science Program/Global Ph.D. Fellowship Program, the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2022R1F1A106366811, No. 2022R1A5A6000840) and KIAS Individual Grant (HP095101) at Korea Institute for Advanced Study. I.S. was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2022R1F1A106366811, 2022R1A5A600084012, 2023R1A2C100517311) and Samsung Science and Technology Foundation (Project Number SSTF-BA1901-03).

<sup>†</sup>Korea Institute for Advanced Study (Korea). E-mail: [seonwookim@kias.re.kr](mailto:seonwookim@kias.re.kr)

<sup>‡</sup>Seoul National University (Korea). E-mail: [insuk.seo@snu.ac.kr](mailto:insuk.seo@snu.ac.kr)

**Contents**

<b>1 Introduction</b>	<b>3</b>
<b>2 Main results</b>	<b>5</b>
2.1 Models . . . . .	5
2.2 Main results: Large deviation-type results . . . . .	8
2.3 Main results: Eyring–Kramers law and Markov chain model reduction . . . . .	11
<b>3 Outline of the proof</b>	<b>14</b>
3.1 Capacity estimate and proof of Theorems 2.10 and 2.12 . . . . .	14
3.2 $H^1$ -approximation of equilibrium potential and proof of Theorem 3.1 . . . . .	18
<b>4 Neighborhood of configurations</b>	<b>19</b>
<b>5 Review of two-dimensional model</b>	<b>19</b>
5.1 2D stochastic Ising and Potts models with periodic boundary conditions . . . . .	20
5.2 Energy barrier and canonical transition paths . . . . .	20
5.3 Saddle structure . . . . .	24
5.4 Test function . . . . .	25
5.5 Auxiliary results . . . . .	26
<b>6 Canonical configurations and paths</b>	<b>27</b>
6.1 Canonical configurations . . . . .	28
6.2 Energy of canonical configurations . . . . .	30
6.3 Canonical paths . . . . .	31
6.4 Characterization of the deepest valleys . . . . .	32
6.5 Auxiliary result on saddle configurations . . . . .	33
<b>7 Gateway configurations</b>	<b>35</b>
7.1 Gateway configurations . . . . .	35
7.2 Properties of gateway configurations . . . . .	37
<b>8 Energy barrier between ground states</b>	<b>39</b>
8.1 Preliminary analysis on energy landscape . . . . .	39
8.2 Proof of Proposition 8.1 . . . . .	44
8.3 Proof of Theorem 2.7 . . . . .	50
<b>9 Typical configurations and optimal paths</b>	<b>50</b>
9.1 Typical configurations . . . . .	51
9.2 Properties of typical configurations . . . . .	52
9.3 Structure of edge typical configurations . . . . .	53
9.4 Analysis of 3D transition paths . . . . .	58
<b>10 Construction of test function</b>	<b>59</b>
10.1 Construction of test function . . . . .	59
10.2 Dirichlet form of test function . . . . .	60
10.3 $H^1$ -approximation . . . . .	61
<b>11 Remarks on open boundary condition</b>	<b>66</b>
<b>References</b>	<b>68</b>



*Example of a three-dimensional saddle configuration*

## 1 Introduction

Metastability is a ubiquitous phenomenon that arises when a stochastic system has several locally stable sets; it is observed in a wide class of models, e.g., in the small random perturbations of dynamical systems (e.g., [16, 17, 23, 33, 35, 38, 39, 36, 48]), interacting particle systems consisting of sticky particles (e.g., [3, 12, 24, 26, 31, 32, 47, 49]), and spin systems in the low temperature regime (e.g., [1, 6, 10, 11, 13, 14, 15, 18, 19, 21, 22, 25, 29, 30, 34, 37, 44, 45, 42]). Numerous important works are not listed here; we direct the references of the monographs [13, 46], which provide a comprehensive introduction to this broad topic.

### Metastable behaviors of stochastic Ising and Potts models

In this study, we consider the metastability of the stochastic Ising and Potts models evolving according to Metropolis–Hastings-type Glauber dynamics on a large, but fixed three-dimensional (3D) lattice. For such models, the Gibbs invariant measure is exponentially concentrated on monochromatic configurations (i.e., the configurations consisting of a single spin, which are the ground states of the Ising and Potts Hamiltonians) in the very low temperature regime. Hence, in such regimes, the dynamics exhibits metastable behavior between the monochromatic configurations: It starts from a monochromatic configuration, remains in a certain neighborhood of the starting configuration for an exponentially long time, and finally overcomes the energy barrier between monochromatic configurations to reach another monochromatic one.

Several mathematical questions persist regarding the metastable behavior explained above. For instance, in the transition from one monochromatic configuration to another, the mean transition time, the asymptotic law of the rescaled transition time, and the typical transition paths are all points of interest. We are also interested in the characterization of the energy barrier and the saddle configurations that realize this energy barrier via optimal paths between monochromatic configurations. The final issue is particularly important and challenging for the model considered in the present article and has remained open for a long time. It is also important to estimate the mixing time or spectral gap of the associated dynamics; this allows us to measure the effects of metastable behavior on the global mixing properties of the associated Markovian dynamics. *In this article, we answer all these questions for the stochastic Ising and Potts models on finite three-dimensional lattices in the absence of external fields.*

### Model with non-zero external field

The first rigorous mathematical treatment of the metastable behavior of the Ising model was performed in [44, 45], where the authors considered the Ising model on a two-dimensional (2D) lattice in the presence of a *non-zero* external field. These studies

verified that the transition from a metastable monochromatic configuration to a stable one is essentially equivalent to the formation of a certain type of critical droplet. From this observation, precise information regarding the transition path was obtained, as well as large deviation-type estimates for the transition time and mixing behavior associated with the Metropolis–Hastings dynamics. This result was extended to the 3D Ising model presented in [1, 6]. Similar results for four- or higher-dimensional models remain to be found, because the variational problems related to the analysis of the energy landscape and critical droplet are highly complicated.

In [18], the aforementioned analyses were further refined via the *potential-theoretic approach* developed in [16]. In [18], the authors obtained the Eyring–Kramers law for the transition time between monochromatic configurations, as well as the spectral gap of the associated dynamics. This new technology does not provide information on the transition path; however, it provides precise asymptotics for the mean metastable transition time and spectral gap. The same model on growing lattice boxes, rather than fixed ones, was investigated in [15], and the Kawasaki-type (instead of Glauber-type) dynamics for the same model were studied in [14].

### Model without external field

When studying the metastability of the stochastic Ising model with a non-zero external field (as described above), the crucial object is the critical droplet, which provides a sharp saddle structure for the energy landscape. However, in the zero external field case, the critical droplet does not exist. Instead, the saddle structure is flat, structurally complex, and composed of a large set of saddle configurations. This is the crucial challenge in the zero external field case, which has left the problem unsolved for a long time. *In the present study, we solve this problem by comprehensively analyzing the energy landscape.*

Recently, [42] analyzed for the first time the 2D Ising and Potts models in the absence of external fields. More precisely, they characterized (1) the energy barrier between ground states and (2) the deepest metastable valleys in the landscape. Using the energy landscape results and a general tool referred to as the *pathwise approach* to metastability (developed in [19, 20, 41, 43]), they obtained large deviation-type results for the metastable behaviors of the 2D models in the absence of external fields.

In [27], which is a companion article of the present one, we improved on the refinement of results in the previous studies for the 2D model using the potential-theoretic approach, thereby making the following contributions:

- the Eyring–Kramers law for metastable transitions between monochromatic configurations,
- the Markov chain model reduction of metastable behavior (cf. [28] for a comprehensive review on this method), and
- the full characterization of typical transition paths.

To this end, we derive a highly detailed analysis of the energy landscape and characterize **all** saddle configurations. In particular, we comprehensively and precisely describe the large and complicated saddle structure of the model. Our analysis is sufficiently accurate to allow the transition paths between ground states to be characterized explicitly.

### Main achievement

In the current article, ***we extend all these analyses to the 3D Ising and Potts models by combining the pathwise approach and the potential-theoretic approach.***

Indeed, the energy landscape of the 3D model is significantly more complicated than that of the 2D model. For both the 2D and 3D models, there are numerous saddle configurations between ground states, and they form a plateau structure. For the 2D model, at least the bulk part of this plateau structure is relatively simple, because each saddle configuration can only move forward or backward to reach another saddle configuration. In contrast, for the 3D model, we cannot expect such a simplification, because there exist certain configurations for which the legitimate movements between saddle configurations can occur in a substantially more complex manner. We refer to the figure at the front page for an example of a highly complicated saddle configuration in the 3D case (which should be characterized in some way to answer all the questions above). Readers who are familiar with the results on the non-zero external field model can notice from this figure that the saddle configurations for the zero external model may not have a clear structure as in the non-zero external field case.

### Approximation method to metastability

In our companion paper [27], we introduced a new approximation method to prove the Eyring–Kramers law and Markov chain model reduction. This method relies on the approximation of the equilibrium potential function (refer to Section 3.1 for the precise definition) in a Sobolev space defined via the Dirichlet norm associated with the Markov chain. It is robust and particularly suitable if the energy landscape is too complex to apply the potential-theoretic approach [16] via variational principles (the Dirichlet and Thomson principles), because it effectively avoids these variational principles via an approximation in the Sobolev space. We apply this method to the 3D model to achieve our main result.

The main mathematical difficulty of applying this method lies in the fact that we must construct a test function that accurately approximates the equilibrium potential function so that we can obtain the precise Sobolev norm. For this procedure, we need a comprehensive understanding of the whole energy landscape regarding the metastable transitions. Thus, compared with the 2D model, the corresponding construction for the 3D model is far more complicated. Overcoming this difficulty is the main contribution of the present study.

## 2 Main results

### 2.1 Models

In this subsection, we introduce the stochastic Ising and Potts models on a fixed 3D lattice and review their basic features.

#### Ising and Potts models

We fix three positive integers  $K \leq L \leq M$ . Then, we denote by

$$\Lambda = \llbracket 1, K \rrbracket \times \llbracket 1, L \rrbracket \times \llbracket 1, M \rrbracket$$

the 3D lattice box. We use the notation  $\llbracket a, b \rrbracket = [a, b] \cap \mathbb{Z}$  throughout this article. We impose either open or periodic boundary conditions upon the lattice box  $\Lambda$ . For the latter boundary condition, we can write

$$\Lambda = \mathbb{T}_K \times \mathbb{T}_L \times \mathbb{T}_M, \quad (2.1)$$

where  $\mathbb{T}_k = \mathbb{Z}/(k\mathbb{Z})$  represents the discrete one-dimensional torus.

For an integer  $q \geq 2$ , we use  $S = \{1, \dots, q\}$  to represent the set of spins and  $\mathcal{X} = S^\Lambda$  to represent the space of spin configurations in the 3D box  $\Lambda$ . We express a configuration  $\sigma \in \mathcal{X}$  as  $\sigma = (\sigma(x))_{x \in \Lambda}$ , where  $\sigma(x) \in S$  represents the spin of  $\sigma$  at site  $x \in \Lambda$ .

For  $x, y \in \Lambda$ , we write  $x \sim y$  if they are neighboring sites; that is,  $\|x - y\| = 1$  where  $\|\cdot\|$  denotes the Euclidean distance in  $\Lambda$ . With this notation, we define the Hamiltonian  $H : \mathcal{X} \rightarrow \mathbb{R}$  as

$$H(\sigma) = \sum_{\{x, y\} \subseteq \Lambda: x \sim y} \mathbf{1}\{\sigma(x) \neq \sigma(y)\} - h \sum_{x \in \Lambda} \sigma(x) \quad ; \quad \sigma \in \mathcal{X}, \tag{2.2}$$

where  $h \in \mathbb{R}$  denotes the magnitude of the external magnetic field. Thus, the first summation on the right-hand side represents the spin-spin interactions, and the second one corresponds to the effect of the external magnetic field. We use  $\mu_\beta(\cdot)$  to denote the Gibbs measure on  $\mathcal{X}$  associated with the Hamiltonian  $H$  at inverse temperature  $\beta > 0$ ; that is,

$$\mu_\beta(\sigma) = \frac{1}{Z_\beta} e^{-\beta H(\sigma)} \quad ; \quad \sigma \in \mathcal{X}, \tag{2.3}$$

where  $Z_\beta = \sum_{\zeta \in \mathcal{X}} e^{-\beta H(\zeta)}$  is the partition function. The random spin configuration on box  $\Lambda$  corresponds to the probability measure  $\mu_\beta(\cdot)$  on  $\mathcal{X}$ ; it is referred to as the *Ising model* if  $q = 2$  and the *Potts model* if  $q \geq 3$ . Henceforth, we treat  $q$  as a fixed parameter. Our primary concern is the metastability analyses of these models as  $\beta \rightarrow \infty$  under Metropolis–Hastings dynamics, which will be defined precisely below.

**Remark 2.1** (Results for non-zero external field). Comprehensive analyses of the energy landscape and the metastability of the Ising model with a non-zero external field i.e.,  $h \neq 0$ , were performed in [18, 44, 45] for the 2D case, and in [1, 6] for the 3D one. For these models, the characterization of the *critical droplet* comprehensively explains the metastable behavior. We remark that analysis for cases of more than three dimensions has yet to be undertaken, because the energy landscape is too complex to allow critical droplets to be characterized. Recently, the 2D Potts model with an external field toward one specific spin has been studied [7, 8, 9].

In this study, we consider the zero external field case (i.e.,  $h = 0$ ); thus, **we henceforth assume that  $h = 0$** . This case differs from those involving non-zero external fields, in the sense that the energy landscape is not characterized by critical droplets. Instead, we must tackle a large and complex landscape of saddle configurations via complicated combinatorial and probabilistic arguments.

### Ground states

For each  $a \in S$ , denote by  $\mathbf{s}_a \in \mathcal{X}$  the monochromatic configuration in which all spins are  $a$ , i.e.,  $\mathbf{s}_a(x) = a$  for all  $x \in \Lambda$ . We write

$$\mathcal{S} = \{\mathbf{s}_1, \dots, \mathbf{s}_q\}. \tag{2.4}$$

It is precisely upon  $\mathcal{S}$  that the Hamiltonian  $H(\cdot)$  attains its minimum 0; hence,  $\mathcal{S}$  represents the set of ground states of the model. Accordingly, we obtain the following characterization of the partition function  $Z_\beta$  that appears in (2.3), as well as the Gibbs measure  $\mu_\beta$  as  $\beta \rightarrow \infty$ .

**Theorem 2.2.** We have<sup>1</sup>

$$Z_\beta = q + O_\beta(e^{-3\beta}) . \tag{2.5}$$

Thus, we obtain

$$\lim_{\beta \rightarrow \infty} \mu_\beta(\mathbf{s}) = \frac{1}{q} \text{ for all } \mathbf{s} \in \mathcal{S} \quad \text{and} \quad \lim_{\beta \rightarrow \infty} \mu_\beta(\mathcal{S}) = 1 .$$

*Proof.* The estimate (2.5) of the partition function comes directly from the expression of the partition function given right after (2.3) and the fact that  $H(\sigma) \geq 3$  for  $\sigma \notin \mathcal{S}$ . The second assertion of the theorem is directly derived from the first one and the expression (2.3) of  $\mu_\beta$ .  $\square$

**Metropolis–Hastings dynamics and metastability**

We give a continuous version of the Metropolis–Hastings dynamics, which is the standard heat-bath Glauber dynamics used for studying the metastability of the Ising model [44]. For  $x \in \Lambda$  and  $a \in S$ , we use  $\sigma^{x,a} \in \mathcal{X}$  to denote the configuration obtained from  $\sigma$  by updating the spin at site  $x$  to  $a$ . Then, the continuous version of the Metropolis–Hastings dynamics is defined as a continuous-time Markov chain  $\{\sigma_\beta(t)\}_{t \geq 0}$  on  $\mathcal{X}$ , whose transition rates are given by

$$r_\beta(\sigma, \zeta) = \begin{cases} e^{-\beta[H(\zeta)-H(\sigma)]_+} & \text{if } \zeta = \sigma^{x,a} \neq \sigma \text{ for some } x \in \Lambda \text{ and } a \in S , \\ 0 & \text{otherwise ,} \end{cases}$$

where  $[\alpha]_+ = \max\{\alpha, 0\}$ . We notice from this definition of the rate  $r_\beta(\cdot, \cdot)$  that the Metropolis–Hastings dynamics tends to lower the energy, particularly when  $\beta$  is large, because the jump rate from one configuration to another one with higher energy is exponentially small, whereas the jump rate to another one with lower or equal energy is 1. We let  $\mathbb{P}_\sigma^\beta$  and  $\mathbb{E}_\sigma^\beta$  represent the law and expectation, respectively, of the process  $\sigma_\beta(\cdot)$  starting from  $\sigma$ .

For  $\sigma, \zeta \in \mathcal{X}$ , we write  $\sigma \sim \zeta$  if  $r_\beta(\sigma, \zeta) > 0$ . Note that  $\sigma \sim \zeta$  if and only if  $\zeta \sim \sigma$ , and that the relation  $\sigma \sim \zeta$  does not depend on  $\beta$ . A crucial observation regarding the rate  $r_\beta(\cdot, \cdot)$  defined above is that

$$\mu_\beta(\sigma) r_\beta(\sigma, \zeta) = \mu_\beta(\zeta) r_\beta(\zeta, \sigma) = \begin{cases} \min\{\mu_\beta(\sigma), \mu_\beta(\zeta)\} & \text{if } \sigma \sim \zeta , \\ 0 & \text{otherwise .} \end{cases} \tag{2.6}$$

From this detailed balance condition, we observe that the invariant measure for the Metropolis–Hastings dynamics  $\sigma_\beta(\cdot)$  is  $\mu_\beta(\cdot)$  and that  $\{\sigma_\beta(t)\}_{t \geq 0}$  is reversible with respect to  $\mu_\beta(\cdot)$ . We also note that the Markov chain  $\sigma_\beta(\cdot)$  is irreducible.

In view of Theorem 2.2, we anticipate that the process  $\sigma_\beta(\cdot)$  will exhibit *metastable behavior between ground states, provided that  $\beta$  is sufficiently large*. More precisely, the process  $\sigma_\beta(\cdot)$  starting from configuration  $\mathbf{s} \in \mathcal{S}$  remains in a certain neighborhood of  $\mathbf{s}$  for a sufficiently long time, and then undergoes a rare but rapid transition to another ground state. Our main concern is to precisely analyze such metastability of the stochastic Ising and Potts models under the Metropolis–Hastings dynamics (defined above) in the very low temperature regime; that is, when  $\beta \rightarrow \infty$ . We explain these results in the following subsection.

<sup>1</sup>For two collections  $(a_\beta)_{\beta > 0} = (a_\beta(K, L, M))_{\beta > 0}$  and  $(b_\beta)_{\beta > 0} = (b_\beta(K, L, M))_{\beta > 0}$  of real numbers, we write  $a_\beta = O_\beta(b_\beta)$  if there exists some  $C = C(K, L, M) > 0$  such that

$$|a_\beta| \leq C b_\beta \text{ for all } \beta > 0 \text{ and } K, L, M .$$

**Remark 2.3.** We employ the continuous-time dynamics (as applied in numerous previous studies) because it offers a simpler presentation than the corresponding discrete dynamics (as demonstrated in [6, 18, 42]), for which the jump probability is given by

$$p_\beta(\sigma, \zeta) = \begin{cases} \frac{1}{q|\Lambda|} e^{-\beta[H(\zeta)-H(\sigma)]_+} & \text{if } \zeta = \sigma^{x,a} \neq \sigma \text{ for some } x \in \Lambda, a \in S, \\ 1 - \sum_{x \in \Lambda, a \in S: \sigma^{x,a} \neq \sigma} p_\beta(\sigma, \sigma^{x,a}) & \text{if } \zeta = \sigma, \\ 0 & \text{otherwise.} \end{cases} \tag{2.7}$$

However, our computations can be applied to this model as well. See also Remark 2.15.

### 2.2 Main results: Large deviation-type results

Hereafter, we explain our results regarding the metastability of the stochastic Ising and Potts models. In the current subsection, we explain the large deviation-type results obtained for the metastable behavior.

#### Energy barrier between ground states

First, we introduce the energy barrier associated with the Ising and Potts models considered in this study. This is important for the analysis of metastable behaviors, in that the Metropolis–Hastings dynamics must overcome this energy barrier to make a transition from one ground state to another.

A sequence of configurations  $(\omega_t)_{t=0}^T = (\omega_0, \omega_1, \dots, \omega_T) \subseteq \mathcal{X}$  for some integer  $T \geq 0$  is called a *path* if  $\omega_t \sim \omega_{t+1}$  (i.e.,  $r_\beta(\omega_t, \omega_{t+1}) > 0$ ) for all  $t \in \llbracket 0, T-1 \rrbracket$ . We say that this path connects two configurations  $\sigma$  and  $\zeta$  if  $\omega_0 = \sigma$  and  $\omega_T = \zeta$  or vice versa. The *communication height* between two configurations  $\sigma, \zeta \in \mathcal{X}^2$  is defined as

$$\Phi(\sigma, \zeta) = \min_{(\omega_t)_{t=0}^T} \max_{t \in \llbracket 0, T \rrbracket} H(\omega_t),$$

where the minimum is taken over all paths connecting  $\sigma$  and  $\zeta$ . Moreover, for two disjoint subsets  $\mathcal{P}$  and  $\mathcal{Q}$  of  $\mathcal{X}$ , we define

$$\Phi(\mathcal{P}, \mathcal{Q}) = \min_{\sigma \in \mathcal{P}, \zeta \in \mathcal{Q}} \Phi(\sigma, \zeta).$$

Then, we define

$$\Gamma = \Gamma(K, L, M) = \Phi(\mathbf{s}, \mathbf{s}') \quad ; \quad \mathbf{s}, \mathbf{s}' \in \mathcal{S}.$$

Note that  $\Phi(\mathbf{s}, \mathbf{s}')$  does not depend on the selections of  $\mathbf{s}, \mathbf{s}' \in \mathcal{S}$ , owing to the model symmetry. Additionally, note that  $\Gamma$  represents the *energy barrier* between ground states, because the dynamics must overcome this energy level to make a transition from one ground state to another.

To characterize the energy barrier, we must check the maximum energy of all paths connecting the ground states. Thus, the energy barrier is a global feature of the energy landscape, and characterizing it is a non-trivial task. For the current model, we can identify the exact value of the energy barrier. Recall that we assumed  $K \leq L \leq M$ .

**Theorem 2.4.** *For all sufficiently large  $K$ , it holds that*

$$\Gamma = \begin{cases} 2KL + 2K + 2 & \text{under periodic boundary conditions,} \\ KL + K + 1 & \text{under open boundary conditions.} \end{cases} \tag{2.8}$$

<sup>2</sup>By writing  $a, b \in A$ , we implicitly state that  $a$  and  $b$  are different.

**Remark 2.5.** Our arguments state that this theorem holds for  $K \geq 2829$ , where the threshold 2829 may be sub-optimal (cf. Remark 8.4). However, the optimality of this threshold is a minor issue, because our main concern is the spin system on large boxes. Henceforth, we assume that  $K$  satisfies this condition, i.e.,  $K \geq 2829$ .

Theorem 2.4 is proved in Section 8.

**Remark 2.6.** Several remarks regarding the previous theorem are in order.

1. Note that Theorem 2.4 does not depend on the value of  $q$ , because in the transition from  $s_a$  to  $s_b$  for  $a, b \in S$ , no spins besides  $a$  and  $b$  play a significant role.
2. Suppose temporarily that  $\Gamma_d$  is the energy barrier, defined in the same way as above, subjected to Ising/Potts models defined on a  $d$ -dimensional lattice box of size  $K_1 \times \dots \times K_d$  with  $K_1 \leq \dots \leq K_d$ . Then, we expect that  $\Gamma_d = 2 + 2 \sum_{n=1}^{d-1} \prod_{i=1}^n K_i$  under periodic boundary conditions and  $\Gamma_d = 1 + \sum_{n=1}^{d-1} \prod_{i=1}^n K_i$  under open boundary conditions for all  $d \geq 2$ . Notice that the case of  $d = 2$  is handled in [42, Theorem 1.1] and the case of  $d = 3$  is handled in Theorem 2.4. We leave the verification of this conjecture for the case of  $d \geq 4$  as a future research problem.

### Comparison with non-zero external field case

We conclude this energy barrier discussion by comparing our results for the zero external field case with those for the non-zero external field case obtained in [44] and [6] for the Ising model (i.e.,  $q = 2$ ) in two or three dimensions, respectively. More precisely, they showed that the energy barrier is given by (under some technical assumptions regarding  $h$ )

$$\begin{aligned} \Gamma_2(h) &= 4\ell_h - h[\ell_h(\ell_h - 1) + 1], \\ \Gamma_3(h) &= 2m_h(2m_h - \delta_h - 1) + 2(m_h - \delta_h)(m_h - 1) + 4\ell_h \\ &\quad - h[m_h(m_h - \delta_h)(m_h - 1) + \ell_h(\ell_h - 1) + 1], \end{aligned}$$

where  $\Gamma_d$  represents the  $d$ -dimensional energy barrier,  $\ell_h = \lceil 2/|h| \rceil$ ,  $m_h = \lceil 4/|h| \rceil$ , and  $\delta_h \in \{0, 1\}$  is a constant depending only on  $h$  (provided that the lattice is sufficiently large). We refer to [13, Chapter 17] for details. These energy barriers are characterized by the energy of the critical droplet, and *their values do not depend on the size of the box but are determined solely by the magnitude  $h$  of the external field*. This is primarily because the size of the critical droplet is determined solely by  $|h|$ , and the size of the box plays no role provided that the box is sufficiently large to contain a single droplet. In contrast, the zero external field case does not feature such a critical droplet; hence, the magnitude of the energy barrier depends crucially on the box size. This is the key difference between the zero external field and non-zero external field cases.

### Large deviation-type results based on pathwise approach

Here, we explain the large deviation-type analysis of the metastable behavior of the Metropolis–Hastings dynamics. These results can be obtained via the pathwise approach developed in [19], provided that we can analyze the model energy landscape to a certain degree of precision. We refer to the monograph [46] for an extensive summary of the pathwise approach. This approach allows us to analyze the metastability from three different perspectives: transition time, spectral gap, and mixing time. All these quantities are crucial for quantifying the metastable behavior. First, we explicitly define them as follows:

- For  $\mathcal{A} \subseteq \mathcal{X}$ , we denote by  $\tau_{\mathcal{A}} = \inf \{t \geq 0 : \sigma_{\beta}(t) \in \mathcal{A}\}$  the hitting time of the set  $\mathcal{A}$ . If  $\mathcal{A} = \{\sigma\}$  is a singleton, we write  $\tau_{\{\sigma\}} = \tau_{\sigma}$ .

- For  $s \in \mathcal{S}$ , we write  $\check{s} = \mathcal{S} \setminus \{s\}$ . Then, our primary concern is the hitting time  $\tau_{\check{s}}$  or  $\tau_{s'}$  for  $s' \in \check{s}$  when the dynamics starts from  $s \in \mathcal{S}$ . We refer to this as the (*metastable*) *transition time*, because it expresses the time required for a transition to proceed from the ground state to another one.
- The *mixing time* corresponding to the level  $\epsilon \in (0, 1)$  is defined as

$$t_{\beta}^{\text{mix}}(\epsilon) = \min \left\{ t \geq 0 : \max_{\sigma \in \mathcal{X}} \|\mathbb{P}_{\sigma}^{\beta}[\sigma_{\beta}(t) \in \cdot] - \mu_{\beta}(\cdot)\|_{\text{TV}} \leq \epsilon \right\},$$

where  $\|\cdot\|_{\text{TV}}$  represents the total variation distance between measures (cf. [40, Chapter 4]).

- We denote by  $\lambda_{\beta}$  the *spectral gap* of the Metropolis–Hastings dynamics defined in Section 2.1.

The 2D version of the following theorem was established in [42] using the refined pathwise approach developed in [20, 41, 43]. We extend their results to the 3D model.

**Theorem 2.7.** *The following statements hold.*

1. (**Transition time**) *For all  $s, s' \in \mathcal{S}$  and  $\epsilon > 0$ , we have*

$$\lim_{\beta \rightarrow \infty} \mathbb{P}_{\mathbf{s}}^{\beta} [e^{\beta(\Gamma-\epsilon)} < \tau_{\check{s}} \leq \tau_{s'} < e^{\beta(\Gamma+\epsilon)}] = 1, \tag{2.9}$$

$$\lim_{\beta \rightarrow \infty} \frac{1}{\beta} \log \mathbb{E}_{\mathbf{s}}^{\beta} [\tau_{\check{s}}] = \lim_{\beta \rightarrow \infty} \frac{1}{\beta} \log \mathbb{E}_{\mathbf{s}}^{\beta} [\tau_{s'}] = \Gamma. \tag{2.10}$$

Moreover, under  $\mathbb{P}_{\mathbf{s}}^{\beta}$ , as  $\beta \rightarrow \infty$ ,

$$\frac{\tau_{\check{s}}}{\mathbb{E}_{\mathbf{s}}^{\beta} [\tau_{\check{s}}]}, \frac{\tau_{s'}}{\mathbb{E}_{\mathbf{s}}^{\beta} [\tau_{s'}]} \rightarrow \text{Exp}(1), \tag{2.11}$$

where  $\text{Exp}(1)$  is the exponential random variable with a mean value of 1.

2. (**Mixing time**) *For all  $\epsilon \in (0, 1/2)$ , the mixing time satisfies*

$$\lim_{\beta \rightarrow \infty} \frac{1}{\beta} \log t_{\beta}^{\text{mix}}(\epsilon) = \Gamma.$$

3. (**Spectral gap**) *There exist two constants  $0 < c_1 \leq c_2$  such that*

$$c_1 e^{-\beta\Gamma} \leq \lambda_{\beta} \leq c_2 e^{-\beta\Gamma}.$$

**Remark 2.8.** The above theorem holds under both open and periodic boundary conditions.

Theorem 2.7 states that the metastable transition time, mixing time, and inverse spectral gap become exponentially large as  $\beta \rightarrow \infty$ , and their exponential growth rates are determined by the energy barrier  $\Gamma$ .

The robust methodology developed in [20, 41, 43] implies that characterizing the energy barrier between ground states and identifying all the deepest valleys suffice (up to several technical issues) to confirm the results presented in Theorem 2.7. In [42], the authors performed corresponding analyses of the energy landscape; then, they used this robust methodology to prove Theorem 2.7 for two dimensions. We perform the corresponding analysis of the energy landscape for the 3D model as well in Sections 6, 7, and 8. The proof of Theorem 2.7 is given in Section 8.3. Analysis of the energy landscape is far more difficult than that of the 2D one considered in [27] for several reasons. Details are presented at the beginning of Section 6.

**Characterization of transition path**

Our analysis of the energy landscape is sufficiently precise to characterize all the possible transition paths between ground states in a high level of detail. The *transition paths* are rigorously defined in Definition 9.13; we do not present explicit definitions here, because we would have to define a large amount of notation. The following theorem asserts that, with dominating probability, the Metropolis–Hastings dynamics evolves along one of the transition paths when a transition occurs from one ground state to another.

**Theorem 2.9.** *For all  $s \in \mathcal{S}$ , we have<sup>3</sup>*

$$\mathbb{P}_s^\beta \left[ \exists 0 < t_1 < \dots < t_N < \tau_s \text{ such that } (\sigma_\beta(t_n))_{n=1}^N \text{ is a transition path between } s \text{ and } \check{s} \right] = 1 - o_\beta(1).$$

The characterization of the transition paths and the proof of this theorem are given in Section 9.4.

**2.3 Main results: Eyring–Kramers law and Markov chain model reduction**

The following results constitute more quantitative analyses of the metastable behavior obtained using potential-theoretic methods. In particular, we obtain the Eyring–Kramers law (which is a considerable refinement of (2.10)) and the Markov chain model reduction of metastable behavior in the sense of [2, 4].

For these results, we require an accurate understanding of the energy landscape and the behavior of the Metropolis–Hastings dynamics on a large set of saddle configurations between ground states. We conduct these analyses in Sections 9 and 10.

We further remark that the quantitative results given below depend on the selection of boundary condition, in contrast to Theorems 2.7 and 2.9 (cf. Remark 2.8). For brevity, we assume periodic boundary conditions throughout this subsection. We can treat the open boundary case in a similar manner; the results and a sketch of the proof are presented in Section 11.

**Eyring–Kramers law**

The following result constitutes a refinement of (2.10) (and hence of (2.11)) that allows us to pin down the sub-exponential prefactor associated with the large deviation-type exponential estimates of the mean transition time between ground states.

**Theorem 2.10.** *There exists a constant  $\kappa = \kappa(K, L, M) > 0$  such that for all  $s, s' \in \mathcal{S}$ ,*

$$\mathbb{E}_s^\beta [\tau_{\check{s}}] = (1 + o_\beta(1)) \frac{\kappa}{q - 1} e^{\Gamma\beta} \quad \text{and} \quad \mathbb{E}_{s'}^\beta [\tau_s] = (1 + o_\beta(1)) \kappa e^{\Gamma\beta}. \tag{2.12}$$

Moreover, the constant  $\kappa$  satisfies

$$\lim_{K \rightarrow \infty} \kappa(K, L, M) = \begin{cases} 1/8 & \text{if } K < L < M, \\ 1/16 & \text{if } K = L < M \text{ or } K < L = M, \\ 1/48 & \text{if } K = L = M. \end{cases} \tag{2.13}$$

In particular, the quantity  $\mathbb{E}_s^\beta [\tau_{\check{s}}]$  represents the mean time required to jump from  $s$  to another ground state; hence, the first formula of (2.12) corresponds to the so-called *Eyring–Kramers law* for the Metropolis–Hastings dynamics.

<sup>3</sup>A collection  $(a_\beta)_{\beta>0} = (a_\beta(K, L, M))_{\beta>0}$  of real numbers is written as  $a_\beta = o_\beta(1)$  if

$$\lim_{\beta \rightarrow \infty} a_\beta = 0 \text{ for all } K, L, M.$$

**Remark 2.11.** Here, we make several comments regarding Theorem 2.10.

1. Although we do not present the exact formula for the constant  $\kappa$  in the theorem, they can be explicitly expressed in terms of potential-theoretic notions relevant to a random walk defined in a complicated space (cf. (3.10) and (3.11) for the formulas). This random walk is vague (cf. Proposition 9.9) compared with the corresponding random walk identified in [27, Proposition 6.22] for the 2D model, which reflects the complexity of the energy landscape of the 3D model compared with that of the 2D one.
2. The constant  $\kappa$  is model-dependent. For different Glauber dynamics (even with identical boundary conditions), this constant may differ.
3. If  $K < L < M$ , the transition between ground states must occur in a specific direction; meanwhile, if  $K = L < M$  or  $K < L = M$ , there are two possible directions for the transition. If  $K = L = M$ , there are six possible directions. This explains the dependence of the asymptotics of  $\kappa$  on the relationships among  $K$ ,  $L$ , and  $M$ .

The proof of Theorem 2.10 is conducted via the potential-theoretic approach, which originates from [16]. Using this approach, we can estimate the mean transition time  $\mathbb{E}_s^\beta [\tau_{\bar{s}}]$  by obtaining a precise estimate of the *capacity* between ground states (cf. [2, Proposition 6.10]). This estimate is typically obtained from variational principles for capacities, such as the Dirichlet and Thomson principles. In contrast, we use the  $H^1$ -approximation technique developed in our companion article [27], which considerably simplifies the proof but still points out the gist of the logical structure needed to estimate the capacity.

To this end, we require precise analyses of the energy landscape and the behavior of the underlying metastable processes on a certain neighborhood of saddle configurations between metastable sets. In most other models for which the Eyring–Kramers law can be obtained via such robust strategies, the energy landscape is relatively simple; hence, the landscape only marginally presents serious mathematical issues. However, in the current model, the saddle consists of a very large collection of saddle configurations, which form a complex structure. Analyzing this structure is a highly complicated task; moreover, it is difficult to assess the behavior of the dynamics in the neighborhood of this large set with adequate precision. The achievement of these tasks is one of the main contributions of this study. We emphasize here that the  $H^1$ -approximation technique, which is used in the proof of the main results in a critical manner, is particularly handy for models with complicated landscapes, such as the one considered in this study.

### Markov chain model reduction of metastable behavior

Because the transitions between ground states occur successively, analyzing all these transitions together is also an important problem in the study of metastability. The general method used is Markov chain model reduction [2, 4, 5]. In this methodology, one proves that the metastable process (accelerated by a certain scale) converges, in a suitable sense, to a Markov chain on the set of metastable sets. For our model, the target Markov chain must be a Markov chain on the collection of ground states, because each ground state corresponds to a metastable set.

To explain this result in the context of our model, we introduce trace process on ground states. In view of Theorem 2.10, we must accelerate the process by a factor  $e^{\Gamma\beta}$  to observe transitions between ground states in the ordinary time scale; hence, let us denote by  $\hat{\sigma}_\beta(t) = \sigma_\beta(e^{\Gamma\beta}t)$ ,  $t \geq 0$  the accelerated process. Then, we define a random

time  $T(t)$ ,  $t \geq 0$  as

$$T(t) = \int_0^t \mathbf{1}\{\widehat{\sigma}_\beta(u) \in \mathcal{S}\} du \quad ; \quad t \geq 0 ,$$

which measures the amount of time (up to  $t$ ) the accelerated process spends on the ground states. Let  $S(\cdot)$  be the generalized inverse of  $T(\cdot)$ ; that is,

$$S(t) = \sup \{u \geq 0 : T(u) \leq t\} \quad ; \quad t \geq 0 .$$

Then, the (accelerated) trace process  $\{X_\beta(t)\}_{t \geq 0}$  on the set  $\mathcal{S}$  of ground states is defined by

$$X_\beta(t) = \widehat{\sigma}_\beta(S(t)) \text{ for } t \geq 0 . \tag{2.14}$$

We observe that the trace process  $X_\beta(\cdot)$  is obtained from the accelerated process  $\widehat{\sigma}_\beta(\cdot)$  by turning off the clock whenever it is not on a ground state; thus, the process  $X_\beta(\cdot)$  extracts information regarding the hopping dynamics on ground states. It is well known that the trace process  $X_\beta(\cdot)$  is a continuous-time, irreducible Markov chain on  $\mathcal{S}$ ; see [2, Proposition 6.1] for a rigorous proof.

Here, in view of the second estimate of (2.12), we define the limiting Markov chain  $\{X(t)\}_{t \geq 0}$  on  $\mathcal{S}$ , which expresses the asymptotic behavior of the accelerated process  $\widehat{\sigma}_\beta(\cdot)$  between the ground states as a continuous-time Markov chain with jump rate

$$r_X(\mathbf{s}, \mathbf{s}') = \kappa^{-1} \text{ for all } \mathbf{s}, \mathbf{s}' \in \mathcal{S} . \tag{2.15}$$

**Theorem 2.12.** *The following statements hold.*

1. *The law of the Markov chain  $X_\beta(\cdot)$  converges to that of the limiting Markov chain  $X(\cdot)$  as  $\beta \rightarrow \infty$ , in the usual Skorokhod topology.*
2. *It holds that*

$$\lim_{\beta \rightarrow \infty} \max_{\mathbf{s} \in \mathcal{S}} \mathbb{E}_{\mathbf{s}}^\beta \left[ \int_0^t \mathbf{1}\{\widehat{\sigma}_\beta(u) \notin \mathcal{S}\} du \right] = 0 .$$

The second part of this theorem implies that the accelerated process spends a negligible amount of time in the set  $\mathcal{X} \setminus \mathcal{S}$ . Therefore, the trace process  $X_\beta(\cdot)$  of  $\widehat{\sigma}_\beta(\cdot)$  on the set  $\mathcal{S}$ , which is essentially obtained by neglecting the excursion of  $\widehat{\sigma}_\beta(\cdot)$  on the set  $\mathcal{X} \setminus \mathcal{S}$ , is indeed a reasonable object for approximating the process  $\widehat{\sigma}_\beta(\cdot)$ . Combining this observation with the first part of the theorem implies that the limiting Markov chain  $X(\cdot)$  describes the successive metastable transitions of the Metropolis–Hastings dynamics.

**Remark 2.13.** The proofs of Theorems 2.10 and 2.12 are based on the potential-theoretic argument, and we present the arguments in Section 3. We conjecture that these results also hold for the cases of more than three dimensions.

**Remark 2.14.** Temporarily, we denote by  $E_{\mathbf{s}}$  the law of the limiting Markov chain  $X(\cdot)$  starting at  $\mathbf{s} \in \mathcal{S}$ . Theorem 2.12 is consistent with Theorem 2.10, in that for any  $\mathbf{s}' \in \mathcal{S}$ , we have  $E_{\mathbf{s}}[\tau_{\mathbf{s}'}] = \kappa$ .

**Remark 2.15** (Discrete Metropolis–Hastings dynamics). The only difference in the discrete dynamics defined by (2.7) is that it is  $q|\Lambda|$  times slower than the continuous dynamics (in the average sense). Therefore, Theorems 2.4, 2.7, and 2.9 are valid for this dynamics without any modification. Theorems 2.10 and 2.12 hold provided that we replace the constant  $\kappa$  with  $\kappa' = q|\Lambda|\kappa$ . The rigorous verification of the result proceeds in a similar way; thus, we do not repeat it here.

## Outlook of proofs of main results

To prove Theorems 2.4 and 2.7, which fall into the category of pathwise-type metastability results, we investigate the energy landscape of the Ising/Potts models on the 3D lattice  $\Lambda$ , as described in Sections 6, 7, and 8. Along the investigation, we present proofs of Theorems 2.4 and 2.7 in Section 8. Then, we proceed to the proofs of Theorems 2.10 and 2.12, which require more accurate analyses of the energy landscape than the previous theorems. These detailed analyses are presented in Section 9, and as a byproduct we present the proof of Theorem 2.9 in Section 9.4. Then, we present the proofs of Theorems 2.10 and 2.12 in Section 10.

## Non-reversible models

The stochastic system considered in this study is the continuous-time Metropolis–Hastings spin-updating dynamics, which is reversible with respect to the Gibbs measure  $\mu_\beta(\cdot)$ . In fact, as in our companion paper [27], we can consider various dynamics with invariant measure  $\mu_\beta(\cdot)$  but are non-reversible with respect to this measure. Since the approximation method and the pathwise approach used in the proof of the main results presented above are robust and can be used in the non-reversible setting as well, we can analyze the 3D version of the non-reversible models introduced in [27] for the 2D model and obtain similar results. However, for simplicity (as analysis of the energy landscape of the 3D model is very complicated itself), we decided not to include the non-reversible content in the current article. Readers who are interested in non-reversible generalizations can refer to [27, Sections 2.2 and 5] for details.

## 3 Outline of the proof

In this section, we provide a brief summary of proof of the main results. We emphasize again that in the remainder of this article (except in Section 11), we assume periodic boundary conditions; that is,  $\Lambda = \mathbb{T}_K \times \mathbb{T}_L \times \mathbb{T}_M$ . In addition, we always assume that  $K$  satisfies the condition given in Remark 2.5.

We reduce the proofs of Theorems 2.10 and 2.12 (which are the final destinations of the current article) to an estimate of the capacity between ground states (cf. Theorem 3.1), and then we reduce the proof of this capacity estimate to the construction of a certain test function (cf. Proposition 3.2) which is a proper approximation of the equilibrium potential function defined in (3.4). The construction and verification of Proposition 3.2 are done in Section 10. This procedure takes into advantage all the information on the energy landscape, analyzed in Sections 6-9.

General strategy to prove such results, which works also in non-reversible cases, was developed in our companion article [27, Section 4]. Thus, we state here only the essential ingredients in a self-contained manner and refer the interested readers to [27, Section 4] for more detail.

### 3.1 Capacity estimate and proof of Theorems 2.10 and 2.12

The *Dirichlet form*  $D_\beta(\cdot)$  associated with the (reversible) Metropolis–Hastings dynamics  $\sigma_\beta(\cdot)$  is given by, for  $f : \mathcal{X} \rightarrow \mathbb{R}$ ,

$$D_\beta(f) = \frac{1}{2} \sum_{\sigma, \zeta \in \mathcal{X}} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) \{f(\zeta) - f(\sigma)\}^2. \quad (3.1)$$

An alternative expression for the Dirichlet form is given as

$$D_\beta(f) = \langle f, -\mathcal{L}_\beta f \rangle_{\mu_\beta}, \quad (3.2)$$

where  $\langle \cdot, \cdot \rangle_{\mu_\beta}$  is the inner product on  $L^2(\mu_\beta)$  and  $\mathcal{L}_\beta$  is the generator of the original process, that is,

$$(\mathcal{L}_\beta f)(\sigma) = \sum_{\zeta \in \mathcal{X}} r_\beta(\sigma, \zeta) [f(\zeta) - f(\sigma)] . \tag{3.3}$$

For two disjoint and non-empty subsets  $\mathcal{P}$  and  $\mathcal{Q}$  of  $\mathcal{X}$ , the *equilibrium potential* between  $\mathcal{P}$  and  $\mathcal{Q}$  is the function  $h_{\mathcal{P}, \mathcal{Q}}^\beta : \mathcal{X} \rightarrow \mathbb{R}$  defined by

$$h_{\mathcal{P}, \mathcal{Q}}^\beta(\sigma) = \mathbb{P}_\sigma^\beta [\tau_{\mathcal{P}} < \tau_{\mathcal{Q}}] . \tag{3.4}$$

By definition, it readily follows that  $h_{\mathcal{P}, \mathcal{Q}}^\beta \equiv 1$  on  $\mathcal{P}$  and  $h_{\mathcal{P}, \mathcal{Q}}^\beta \equiv 0$  on  $\mathcal{Q}$ . Then, we define the *capacity* between  $\mathcal{P}$  and  $\mathcal{Q}$  as

$$\text{Cap}_\beta(\mathcal{P}, \mathcal{Q}) = D_\beta(h_{\mathcal{P}, \mathcal{Q}}^\beta) . \tag{3.5}$$

It is well known that the equilibrium potential is the unique solution to the following equation:

$$\begin{cases} f \equiv 1 & \text{on } \mathcal{P} , \\ f \equiv 0 & \text{on } \mathcal{Q} , \\ \mathcal{L}_\beta f \equiv 0 & \text{on } \mathcal{X} \setminus (\mathcal{P} \cup \mathcal{Q}) . \end{cases} \tag{3.6}$$

Next, we define the constant  $\kappa = \kappa(K, L, M)$  that appears in Theorems 2.10 and 2.12.

- Let  $m_K = \lfloor K^{2/3} \rfloor$ , where  $\lfloor \alpha \rfloor$  is the biggest integer not bigger than  $\alpha$ , and let  $\kappa^{2D} = \kappa^{2D}(K, L)$  be the constant that appeared in [27, (4.13)], which is defined later in (5.11) explicitly and satisfies

$$\lim_{K \rightarrow \infty} \kappa^{2D}(K, L) = \begin{cases} 1/4 & \text{if } K < L , \\ 1/8 & \text{if } K = L . \end{cases} \tag{3.7}$$

Then, for  $n \in \llbracket 1, q-1 \rrbracket$ , the bulk constant  $\mathfrak{b}(n)$  is defined explicitly as

$$\mathfrak{b}(n) = \begin{cases} \frac{1}{n(q-n)} \cdot \frac{M-2m_K}{2M} \cdot \kappa^{2D}(K, L) & \text{if } K < L < M , \\ \frac{1}{n(q-n)} \cdot \frac{M-2m_K}{2M} \cdot \kappa^{2D}(K, L) & \text{if } K = L < M , \\ \frac{1}{n(q-n)} \cdot \frac{M-2m_K}{4M} \cdot \kappa^{2D}(K, L) & \text{if } K < L = M , \\ \frac{1}{n(q-n)} \cdot \frac{M-2m_K}{6M} \cdot \kappa^{2D}(K, L) & \text{if } K = L = M . \end{cases} \tag{3.8}$$

- The edge constant  $\mathfrak{c}(n)$ ,  $n \in \llbracket 1, q-1 \rrbracket$ , is defined in (9.12). Furthermore, it is verified in Proposition 9.9 that

$$0 < \mathfrak{c}(n) \leq \frac{1}{K^{1/3}} \text{ for all } n \in \llbracket 1, q-1 \rrbracket . \tag{3.9}$$

- Then, for  $n \in \llbracket 1, q-1 \rrbracket$ , we define the constant

$$\mathfrak{c}(n) = \mathfrak{b}(n) + \mathfrak{c}(n) + \mathfrak{c}(q-n) . \tag{3.10}$$

We remark that by definition,  $\mathfrak{b}(n) = \mathfrak{b}(q-n)$  for  $n \in \llbracket 1, q-1 \rrbracket$ ; therefore, we have  $\mathfrak{c}(n) = \mathfrak{c}(q-n)$ . Finally, we define the constant  $\kappa$  that appears in Theorem 2.10 as

$$\kappa = (q-1)\mathfrak{c}(1) . \tag{3.11}$$

For  $A \subseteq S$ , we define (cf. (2.4))

$$\mathcal{S}(A) = \{s_a : a \in A\}.$$

A pair  $(A, B)$  of two subsets  $A$  and  $B$  of  $S$  is referred to as a *proper partition* of  $S$  if  $A$  and  $B$  are non-empty subsets of  $S$  satisfying  $A \cup B = S$  and  $A \cap B = \emptyset$ . Our aim is to estimate the capacity between  $\mathcal{S}(A)$  and  $\mathcal{S}(B)$  for proper partitions  $(A, B)$  of  $S$ . The following theorem expresses the key capacity estimate:

**Theorem 3.1.** *It holds for any proper partition  $(A, B)$  of  $S$  that*

$$\text{Cap}_\beta(\mathcal{S}(A), \mathcal{S}(B)) = \frac{1 + o_\beta(1)}{q\mathfrak{c}(|A|)} e^{-\Gamma\beta}, \tag{3.12}$$

where  $\mathfrak{c}(|A|)$  is the constant defined in (3.10).

We explain the strategy used to prove this theorem in Section 3.2. Here, we conclude the proofs of Theorems 2.10 and 2.12 by assuming Theorem 3.1.

*Proof of Theorem 2.10.* By [2, Proposition 6.10], we have the following formula for the mean transition time:

$$\mathbb{E}_s^\beta [\tau_{\check{s}}] = \frac{1}{\text{Cap}_\beta(s, \check{s})} \sum_{\sigma \in \mathcal{X}} \mu_\beta(\sigma) h_{s, \check{s}}^\beta(\sigma).$$

Using Theorem 2.2 and the fact that  $h_{s, \check{s}}^\beta(s) = 1$  and  $h_{s, \check{s}}^\beta \equiv 0$  on  $\check{s}$ , we can rewrite the last summation as

$$\frac{1}{q} + o_\beta(1) + \sum_{\sigma \in \mathcal{X} \setminus \mathcal{S}} \mu_\beta(\sigma) h_{s, \check{s}}^\beta(\sigma) = \frac{1}{q} + o_\beta(1),$$

where the identity follows from the trivial bound  $|h_{s, \check{s}}^\beta| \leq 1$  (cf. (3.4)). Summing up the computations above and applying Theorem 3.1, we obtain

$$\mathbb{E}_s^\beta [\tau_{\check{s}}] = \frac{1}{\text{Cap}_\beta(s, \check{s})} \left[ \frac{1}{q} + o_\beta(1) \right] = (1 + o_\beta(1)) \frac{\kappa}{q-1} e^{\Gamma\beta}. \tag{3.13}$$

We next address the second estimate of (2.12). Assume that the process  $\sigma_\beta(\cdot)$  starts at  $s$  and that  $s \neq s'$ . We define a sequence of stopping times  $(J_n)_{n=0}^\infty$  by  $J_0 = 0$  and

$$J_{n+1} = \inf \{t \geq J_n : \sigma_\beta(t) \in \mathcal{S} \setminus \sigma_\beta(J_n)\} \quad ; \quad n \geq 0.$$

In other words,  $(J_n)_{n=0}^\infty$  is the sequence of random times at which the process  $\sigma_\beta(\cdot)$  visits a new ground state. By (3.13) and the strong Markov property, we have for all  $n \geq 0$  that

$$\mathbb{E}_s^\beta [J_{n+1} - J_n] = (1 + o_\beta(1)) \frac{\kappa}{q-1} e^{\Gamma\beta}. \tag{3.14}$$

Then, we define

$$n(s') = \inf \{n \geq 0 : \sigma_\beta(J_n) = s'\}$$

such that  $\tau_{s'} = J_{n(s')}$ ; thus, we can write

$$\tau_{s'} = \sum_{i=0}^{n(s')-1} (J_{i+1} - J_i). \tag{3.15}$$

Note that because we have assumed  $\mathbf{s} \neq \mathbf{s}'$ , it holds that  $n(\mathbf{s}') \geq 1$ . By symmetry, we observe that  $n(\mathbf{s}')$  is a geometric random variable with success probability  $\frac{1}{q-1}$  that is independent of the sequence  $(J_n)_{n=0}^\infty$ . Thus, we get from (3.14) and (3.15) that

$$\mathbb{E}_{\mathbf{s}}^\beta [\tau_{\mathbf{s}'}] = (1 + o_\beta(1)) \frac{\kappa}{q-1} e^{\Gamma\beta} \times (q-1) = (1 + o_\beta(1)) \kappa e^{\Gamma\beta} .$$

Finally, from (3.7), (3.8), (3.9), and (3.10), we can easily see that  $\kappa$  satisfies the asymptotics (2.13). This completes the proof.  $\square$

Next, we consider Theorem 2.12. The general methodology used to prove this type of Markov chain model reduction, based on potential-theoretic computations, was developed in [2, 4]. Our proof also uses the potential-theoretic approach; however, the computation is slightly simpler because the metastable sets are singletons. Before stating the proof, we remark that two alternative approaches are available for the Markov chain model reduction in the context of metastability: an approach based on the Poisson equation [33, 35, 47, 48], and one based on the resolvent equation [32, 39].

*Proof of Theorem 2.12.* We first consider part (1). We denote by  $r_\beta^{\text{tr}} : \mathcal{S} \times \mathcal{S} \rightarrow [0, \infty)$  the transition rate of the trace process  $X_\beta(\cdot)$ . In view of the rate (2.15) of the limiting Markov chain, it suffices to prove that  $r_\beta^{\text{tr}}(\mathbf{s}, \mathbf{s}') = (1 + o_\beta(1)) \frac{1}{\kappa}$  for all  $\mathbf{s}, \mathbf{s}' \in \mathcal{S}$  [2, Theorem 2.4]. Since  $r_\beta(\mathbf{s}, \mathbf{s}')$  does not depend on the selections of  $\mathbf{s}, \mathbf{s}' \in \mathcal{S}$  by the symmetry of the model, it remains to prove that

$$r_\beta^{\text{tr}}(\mathbf{s}, \check{\mathbf{s}}) = (1 + o_\beta(1)) \frac{q-1}{\kappa} \quad \text{for all } \mathbf{s} \in \mathcal{S} . \tag{3.16}$$

We denote by  $\mathbb{E}_{\mathbf{s}}^\beta$  the law of the trace process  $X_\beta(\cdot)$  starting at  $\mathbf{s}$ . Then,

$$\frac{1}{r_\beta^{\text{tr}}(\mathbf{s}, \check{\mathbf{s}})} = \mathbb{E}_{\mathbf{s}}^\beta [\tau_{\check{\mathbf{s}}}] = e^{-\Gamma\beta} \mathbb{E}_{\mathbf{s}}^\beta \left[ \int_0^{\tau_{\check{\mathbf{s}}}} \mathbf{1}\{\sigma_\beta(t) \in \mathcal{S}\} dt \right] , \tag{3.17}$$

where the factor  $e^{-\Gamma\beta}$  is included because we accelerated the process by the factor  $e^{\Gamma\beta}$  when defining the trace process; the integrand  $\mathbf{1}\{\sigma_\beta(t) \in \mathcal{S}\}$  arises because the trace process is obtained from the accelerated process by turning off the clock when the process resides outside  $\mathcal{S}$ . Then, by [2, Proposition 6.10], we can write

$$\mathbb{E}_{\mathbf{s}}^\beta \left[ \int_0^{\tau_{\check{\mathbf{s}}}} \mathbf{1}\{\sigma_\beta(t) \in \mathcal{S}\} dt \right] = \frac{1}{\text{Cap}_\beta(\mathbf{s}, \check{\mathbf{s}})} \sum_{\sigma \in \mathcal{X}} \mu_\beta(\sigma) \mathbf{1}\{\sigma \in \mathcal{S}\} h_{\mathbf{s}, \check{\mathbf{s}}}^\beta(\sigma) = \frac{\mu_\beta(\mathbf{s})}{\text{Cap}_\beta(\mathbf{s}, \check{\mathbf{s}})} ,$$

where the second identity follows from the fact that  $h_{\mathbf{s}, \check{\mathbf{s}}}^\beta(\mathbf{s}) = 1$  and  $h_{\mathbf{s}, \check{\mathbf{s}}}^\beta \equiv 0$  on  $\check{\mathcal{S}}$ . Therefore, by Theorems 2.2 and 3.1, we obtain

$$\mathbb{E}_{\mathbf{s}}^\beta \left[ \int_0^{\tau_{\check{\mathbf{s}}}} \mathbf{1}\{\sigma_\beta(t) \in \mathcal{S}\} dt \right] = (1 + o_\beta(1)) \frac{\kappa}{q-1} e^{\Gamma\beta} .$$

Inserting this into (3.17) yields (3.16).

Here, we address part (2). Denote by  $\mathbb{P}_{\mu_\beta}^\beta$  the law of the Metropolis–Hastings dynamics  $\sigma_\beta(\cdot)$  for which the initial distribution is  $\mu_\beta$ . Then, for any  $u > 0$ , we obtain

$$\mathbb{P}_{\mathbf{s}}^\beta [\sigma_\beta(u) \notin \mathcal{S}] \leq \frac{1}{\mu_\beta(\mathbf{s})} \mathbb{P}_{\mu_\beta}^\beta [\sigma_\beta(u) \notin \mathcal{S}] = \frac{\mu_\beta(\mathcal{X} \setminus \mathcal{S})}{\mu_\beta(\mathbf{s})} , \tag{3.18}$$

where the final identity holds because  $\mu_\beta$  is the invariant distribution. Therefore by the Fubini theorem,

$$\mathbb{E}_{\mathbf{s}}^\beta \left[ \int_0^t \mathbf{1}\{\widehat{\sigma}_\beta(u) \notin \mathcal{S}\} du \right] = \int_0^t \mathbb{P}_{\mathbf{s}}^\beta [\sigma_\beta(e^{\Gamma\beta} u) \notin \mathcal{S}] du \leq t \cdot \frac{\mu_\beta(\mathcal{X} \setminus \mathcal{S})}{\mu_\beta(\mathbf{s})} ,$$

which vanishes as  $\beta \rightarrow \infty$  by Theorem 2.2.  $\square$

**3.2  $H^1$ -approximation of equilibrium potential and proof of Theorem 3.1**

We fix a proper partition  $(A, B)$  of  $S$ , and explain the general strategy to prove Theorem 3.1, that is, to estimate the capacity  $\text{Cap}_\beta(\mathcal{S}(A), \mathcal{S}(B))$ .

The methodology explained here is based on [27, Section 4.5], in which it is demonstrated that finding a suitable  $H^1$ -approximation of the equilibrium potential  $h_{\mathcal{S}(A), \mathcal{S}(B)}^\beta$  between  $\mathcal{S}(A)$  and  $\mathcal{S}(B)$  is sufficient to establish the capacity estimate. The following proposition states this result.

**Proposition 3.2** ( $H^1$ -approximation of the equilibrium potential). *For any proper partition  $(A, B)$  of  $S$ , there exists a function  $\tilde{h} = \tilde{h}_{\mathcal{S}(A), \mathcal{S}(B)}^\beta : \mathcal{X} \rightarrow \mathbb{R}$  such that the following properties hold.*

1. The function  $\tilde{h}$  approximates  $h_{\mathcal{S}(A), \mathcal{S}(B)}^\beta$  in the sense that

$$D_\beta(h_{\mathcal{S}(A), \mathcal{S}(B)}^\beta - \tilde{h}) = o_\beta(e^{-\beta\Gamma}) . \tag{3.19}$$

2. It holds that

$$D_\beta(\tilde{h}) = \frac{1 + o_\beta(1)}{qc(|A|)} e^{-\Gamma\beta} . \tag{3.20}$$

**Remark 3.3.** The following statements are remarks on the previous proposition.

1. Since the (square root of the) Dirichlet form can be regarded as an  $H^1$ -seminorm, by (3.19), the test function  $\tilde{h}$  approximates  $h_{\mathcal{S}(A), \mathcal{S}(B)}^\beta$  in the  $H^1$ -sense.
2. Property (3.20) is the one that should be satisfied by the equilibrium potential, provided that Theorem 3.1 holds in view of (3.5).
3. Proposition 3.2 has a simpler form compared to the original one [27, Proposition 4.4], because here we only need to consider the case when  $(A, B)$  is a proper partition of  $S$ .
4. Finding the test function  $\tilde{h}$  requires precise information on the energy landscape and a deep insight into typical patterns of the Metropolis–Hastings dynamics in a suitable neighborhood of saddle configurations. We derive this in Sections 6-9. Then, the construction of the test function  $\tilde{h}$  and the proof of Proposition 3.2 are given in Section 10.
5. We remark that the strategies explained in Section 3, which are the proofs of the Eyring–Kramers law and the Markov chain model reduction explained in Section 3.1 and the  $H^1$ -approximation method explained in Section 3.2, are model-independent and robust against specific models. Whereas, the remaining parts of the article are dedicated to construct the test function  $\tilde{h}$  satisfying the assumptions in Proposition 3.2 and highly depend on the complicated structure of the energy landscape of the 3D Ising/Potts models.

Finally, provided that Proposition 3.2 holds, we prove Theorem 3.1.

*Proof of Theorem 3.1.* By the triangle inequality for the seminorm  $D_\beta(\cdot)^{1/2}$ , it holds that  $D_\beta(\tilde{h})^{1/2} - D_\beta(h_{\mathcal{S}(A), \mathcal{S}(B)}^\beta - \tilde{h})^{1/2} \leq D_\beta(h_{\mathcal{S}(A), \mathcal{S}(B)}^\beta)^{1/2} \leq D_\beta(\tilde{h})^{1/2} + D_\beta(h_{\mathcal{S}(A), \mathcal{S}(B)}^\beta - \tilde{h})^{1/2}$ .

Hence, by (3.19) and (3.20), we obtain

$$D_\beta(h_{\mathcal{S}(A), \mathcal{S}(B)}^\beta) = \frac{1 + o_\beta(1)}{qc(|A|)} e^{-\Gamma\beta} .$$

By (3.5), the last display completes the proof. □

Hence, to prove the main results given in Theorems 2.10 and 2.12, it remains to prove Proposition 3.2. The proof is given in Section 10.

#### 4 Neighborhood of configurations

In this section, we introduce several notions of neighborhoods of configurations, which are analogues of the same concepts defined in [27, Section 6.1]. These notions will be crucially used in the characterization of energy landscape and in the construction of test objects.

For  $c \in \mathbb{R}$ , a path  $(\omega_t)_{t=0}^T$  in  $\mathcal{X}$  is called a  $c$ -path if it is a path in the sense of Section 2.2, and moreover satisfies  $H(\omega_t) \leq c$  for all  $t \in \llbracket 0, T \rrbracket$ . Moreover, we say that this path is in  $\mathcal{P} \subseteq \mathcal{X}$  if  $\omega_t \in \mathcal{P}$  for all  $t \in \llbracket 0, T \rrbracket$ .

**Definition 4.1** (Neighborhood of configurations). 1. For  $\sigma \in \mathcal{X}$ , the neighborhood  $\mathcal{N}(\sigma)$  and the extended neighborhood  $\widehat{\mathcal{N}}(\sigma)$  are defined as

$$\begin{aligned} \mathcal{N}(\sigma) &= \{\zeta \in \mathcal{X} : \text{There exists a } (\Gamma - 1)\text{-path } (\omega_t)_{t=0}^T \text{ connecting } \sigma \text{ and } \zeta\} \text{ and} \\ \widehat{\mathcal{N}}(\sigma) &= \{\zeta \in \mathcal{X} : \text{There exists a } \Gamma\text{-path } (\omega_t)_{t=0}^T \text{ connecting } \sigma \text{ and } \zeta\}. \end{aligned}$$

We set  $\mathcal{N}(\sigma) = \emptyset$  (resp.  $\widehat{\mathcal{N}}(\sigma) = \emptyset$ ) if  $H(\sigma) > \Gamma - 1$  (resp.  $H(\sigma) > \Gamma$ ). Then for  $\mathcal{P} \subseteq \mathcal{X}$ , we define

$$\mathcal{N}(\mathcal{P}) = \bigcup_{\sigma \in \mathcal{P}} \mathcal{N}(\sigma) \quad \text{and} \quad \widehat{\mathcal{N}}(\mathcal{P}) = \bigcup_{\sigma \in \mathcal{P}} \widehat{\mathcal{N}}(\sigma).$$

2. Let  $\mathcal{Q} \subseteq \mathcal{X}$ . For  $\sigma \in \mathcal{X}$  such that  $\sigma \notin \mathcal{Q}$ , we define

$$\widehat{\mathcal{N}}(\sigma; \mathcal{Q}) = \{\zeta \in \mathcal{X} : \text{There exists a } \Gamma\text{-path in } \mathcal{X} \setminus \mathcal{Q} \text{ connecting } \sigma \text{ and } \zeta\}.$$

As before, we set  $\widehat{\mathcal{N}}(\sigma; \mathcal{Q}) = \emptyset$  if  $H(\sigma) > \Gamma$ . Then for  $\mathcal{P} \subseteq \mathcal{X}$  disjoint with  $\mathcal{Q}$ , define

$$\widehat{\mathcal{N}}(\mathcal{P}; \mathcal{Q}) = \bigcup_{\sigma \in \mathcal{P}} \widehat{\mathcal{N}}(\sigma; \mathcal{Q}).$$

With this notation, by the definition of  $\Gamma$ , it holds that  $\mathcal{N}(s) \cap \mathcal{N}(s') = \emptyset$  and  $\widehat{\mathcal{N}}(s) = \widehat{\mathcal{N}}(s')$  for any  $s, s' \in \mathcal{S}$ . Moreover, in the spirit of the large deviation principle, the only configurations relevant to the study of metastability are the ones in  $\widehat{\mathcal{N}}(\mathcal{S})$ . Hence, it is crucial to understand the structure of the set  $\widehat{\mathcal{N}}(\mathcal{S})$ . That is the content of Proposition 9.6.

We conclude this section with an elementary lemma which will be used in several instances of our discussion. The proof is well explained in [27, Lemma A.1], and thus we omit the detail.

**Lemma 4.2.** Suppose that  $\mathcal{P}$  and  $\mathcal{Q}$  are disjoint subsets of  $\mathcal{X}$ . Then, it holds that

$$\widehat{\mathcal{N}}(\mathcal{P} \cup \mathcal{Q}) = \widehat{\mathcal{N}}(\mathcal{Q}; \mathcal{P}) \cup \widehat{\mathcal{N}}(\mathcal{P}; \mathcal{Q}).$$

#### 5 Review of two-dimensional model

In this section, we recall some crucial 2D results on the energy landscape from [27, Sections 6, 7 and Appendices B, C], which are needed in our investigation of the 3D model. Since all the results that appear in the current section are proved in [27], we refer to the proofs therein.

**Notation 5.1.** Greek letters  $\eta$  and  $\xi$  are used to denote the spin configurations of the 2D model, while letters  $\sigma$  and  $\zeta$  are used to denote the 3D configurations. We use the superscript 2D to stress the notation for the 2D model; for example, we shall denote by  $H^{2D}(\cdot)$  the Hamiltonian of the 2D model to distinguish with  $H(\cdot)$  which denotes the Hamiltonian of the 3D model.

**5.1 2D stochastic Ising and Potts models with periodic boundary conditions**

We denote by  $\Lambda^{2D} = \mathbb{T}_K \times \mathbb{T}_L$  the 2D lattice with periodic boundary conditions. Recall that  $S = \{1, 2, \dots, q\}$  denotes the set of spins, and denote by  $\mathcal{X}^{2D} = S^{\Lambda^{2D}}$  the space of spin configurations on the 2D lattice. Then, the 2D Ising/Potts Hamiltonian function  $H^{2D} : \mathcal{X}^{2D} \rightarrow \mathbb{R}$  (without external field) is defined by

$$H^{2D}(\eta) = \sum_{\{x,y\} \subseteq \Lambda^{2D}: x \sim y} \mathbf{1}\{\eta(x) \neq \eta(y)\} \quad ; \quad \eta \in \mathcal{X}^{2D} . \tag{5.1}$$

We denote by  $\mathbf{s}_a^{2D}$ ,  $a \in S$  the 2D monochromatic configurations of spin  $a$ , that is,  $\mathbf{s}_a^{2D}(x) = a$  for all  $x \in \Lambda^{2D}$ . Then, it is straightforward that the ground states of this Hamiltonian is also the monochromatic configurations, i.e., the collection  $\mathcal{S}^{2D}$  of the ground states is given as

$$\mathcal{S}^{2D} = \{\mathbf{s}_1^{2D}, \mathbf{s}_2^{2D}, \dots, \mathbf{s}_q^{2D}\} .$$

Then, we write  $\mu_\beta^{2D}(\cdot)$  the associated 2D Gibbs measure, i.e.,

$$\mu_\beta^{2D}(\eta) = \frac{1}{Z_\beta^{2D}} e^{-\beta H^{2D}(\eta)} \quad ; \quad \eta \in \mathcal{X}^{2D} .$$

Here,  $Z_\beta^{2D}$  is the 2D partition function with the property that (cf. [27, Theorem 2.1])

$$\lim_{\beta \rightarrow \infty} Z_\beta^{2D} = q . \tag{5.2}$$

In the 2D model, we also consider the continuous-time Metropolis–Hastings dynamics whose transition rate is defined as

$$r_\beta^{2D}(\eta, \xi) = \begin{cases} e^{-\beta[H^{2D}(\xi) - H^{2D}(\eta)]_+} & \text{if } \xi = \eta^{x,a} \neq \eta \text{ for some } x \in \Lambda^{2D} \text{ and } a \in S , \\ 0 & \text{otherwise .} \end{cases}$$

This 2D stochastic Ising/Potts model is thoroughly analyzed in our companion article [27]. The remainder of this section presents a review of our analysis.

**5.2 Energy barrier and canonical transition paths**

It is verified in [42, Theorem 2.1] that the energy barrier between the ground states of the 2D model is given by

$$\Gamma^{2D} = 2K + 2 .$$

Then, by replacing  $\Gamma$  that appears in Definition 4.1 with  $\Gamma^{2D}$ , we get two types of neighborhoods  $\mathcal{N}^{2D}$  and  $\widehat{\mathcal{N}}^{2D}$  for the 2D model. In this subsection, we explain a class of natural optimal transition paths that achieve this energy level. These paths are denoted as canonical paths. To define these paths, we first define the so-called canonical configurations. We note that the constructions given here is a brief survey of [27, Section 6.2].

**Canonical configurations**

The following notation is used throughout the article (also for the 3D model).

**Notation 5.2.** *Suppose that  $N \geq 2$  is a positive integer.*

- Define  $\mathfrak{S}_N$  as the collection of connected subsets of  $\mathbb{T}_N$ . For example, if  $N = 6$ , few examples of the elements of  $\mathfrak{S}_6$  are  $\emptyset$ ,  $\{2, 3\}$ ,  $\{5, 6, 1\}$ ,  $\{4, 5, 6, 1, 2\}$ ,  $\mathbb{T}_6$ , etc.
- For  $P, P' \in \mathfrak{S}_N$ , we write  $P \prec P'$  if  $P \subseteq P'$  and  $|P'| = |P| + 1$ .

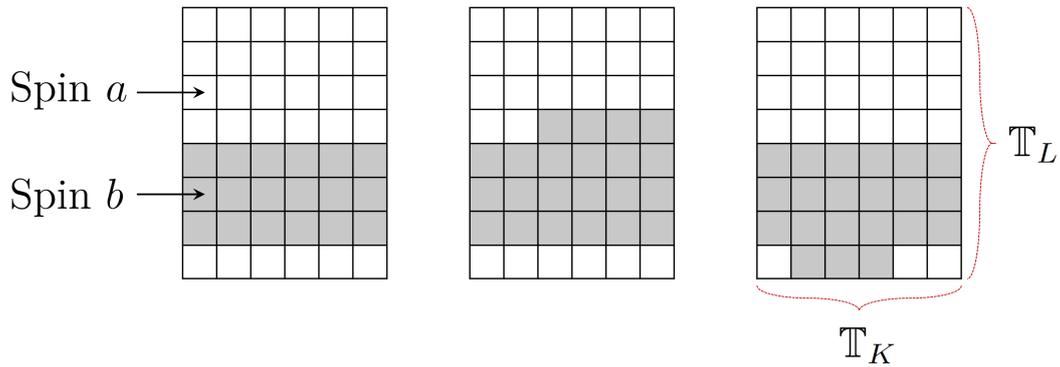


Figure 1: **2D pre-canonical configurations.** These configurations illustrate  $\xi_{2,3}^{a,b}$ ,  $\xi_{2,3;3,4}^{a,b,+}$ , and  $\xi_{2,3;2,3}^{a,b,-}$ , respectively.

- A sequence  $(P_m)_{m=0}^N$  of sets in  $\mathfrak{S}_N$  is called an increasing sequence if it satisfies

$$\emptyset = P_0 \prec P_1 \prec \dots \prec P_N = \mathbb{T}_N$$

so that  $|P_m| = m$  for all  $m \in \llbracket 0, N \rrbracket$ .

We first introduce the pre-canonical configurations which are illustrated in Figure 1.

**Definition 5.3** (2D pre-canonical configurations). Fix two spins  $a, b \in S$ .

- For  $\ell \in \mathbb{T}_L$  and  $v \in \llbracket 0, L \rrbracket$ , we denote by  $\xi_{\ell,v}^{a,b} \in \mathcal{X}^{2D}$  the configuration whose spins are  $b$  on

$$\mathbb{T}_K \times \{\ell + n \in \mathbb{T}_L : n \in \llbracket 0, v - 1 \rrbracket \subseteq \mathbb{Z}\} .$$

and  $a$  on the remainder.

- For  $\ell \in \mathbb{T}_L$ ,  $v \in \llbracket 0, L - 1 \rrbracket$ ,  $k \in \mathbb{T}_K$ , and  $h \in \llbracket 0, K \rrbracket$ , we denote by  $\xi_{\ell,v;k,h}^{a,b,+} \in \mathcal{X}^{2D}$  the configuration whose spins are  $b$  on

$$\{x \in \Lambda^{2D} : \xi_{\ell,v}^{a,b}(x) = b\} \cup [\{k + n \in \mathbb{T}_K : n \in \llbracket 0, h - 1 \rrbracket \subseteq \mathbb{Z}\} \times \{\ell + v\}]$$

and  $a$  on the remainder. Similarly,  $\xi_{\ell,v;k,h}^{a,b,-} \in \mathcal{X}^{2D}$  is the configuration whose spins are  $b$  on

$$\{x \in \Lambda^{2D} : \xi_{\ell,v}^{a,b}(x) = b\} \cup [\{k + n \in \mathbb{T}_K : n \in \llbracket 0, h - 1 \rrbracket \subseteq \mathbb{Z}\} \times \{\ell - 1\}]$$

and  $a$  on the remainder. The configurations defined here are 2D pre-canonical configurations.

Based on this definition, the 2D canonical and regular configurations are defined.

**Definition 5.4** (2D canonical and regular configurations). Fix  $a, b \in S$ . The definitions are slightly different for the case of  $K < L$  and the case of  $K = L$ .

- **(Case  $K < L$ )** Collection  $\mathcal{C}^{a,b,2D}$  of 2D canonical configurations between  $\mathfrak{s}_a^{2D}$  and  $\mathfrak{s}_b^{2D}$  is defined by

$$\mathcal{C}^{a,b,2D} = \bigcup_{\ell \in \mathbb{T}_L} \bigcup_{v \in \llbracket 0, L \rrbracket} \{\xi_{\ell,v}^{a,b}\} \cup \bigcup_{\ell \in \mathbb{T}_L} \bigcup_{v \in \llbracket 0, L-1 \rrbracket} \bigcup_{k \in \mathbb{T}_K} \bigcup_{h \in \llbracket 1, K-1 \rrbracket} \{\xi_{\ell,v;k,h}^{a,b,+}, \xi_{\ell,v;k,h}^{a,b,-}\} .$$

Then, the collection of canonical configurations is given as

$$\mathcal{C}^{2D} = \bigcup_{a,b \in S} \mathcal{C}^{a,b,2D} . \tag{5.3}$$

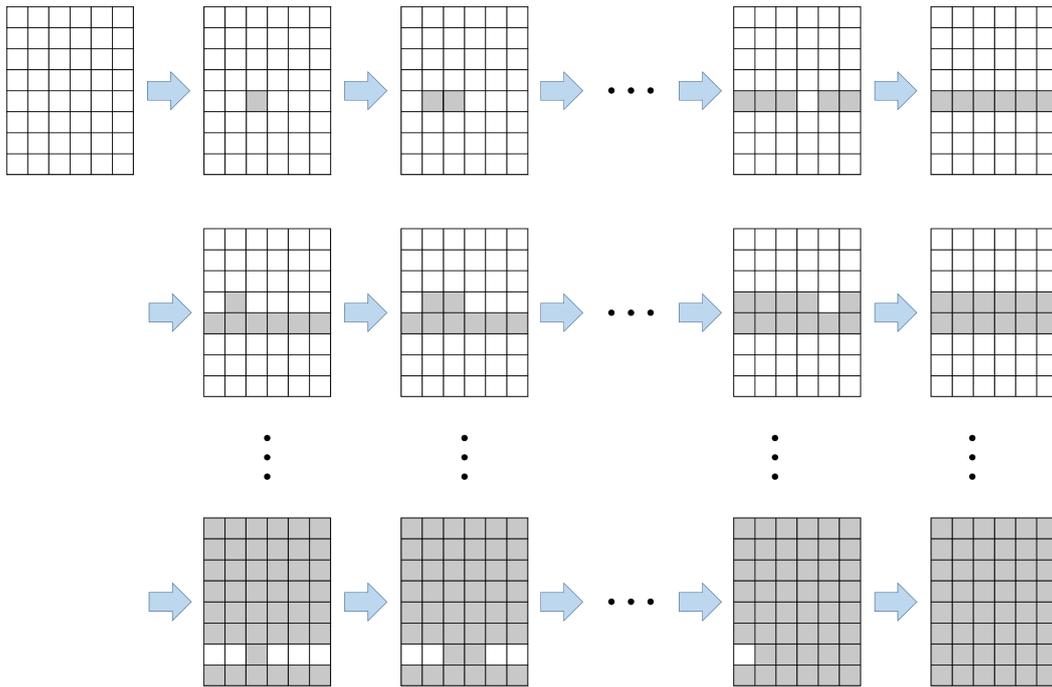


Figure 2: **Example of a 2D canonical path from  $s_a^{2D}$  to  $s_b^{2D}$ .**

Similarly,

$$\mathcal{R}_v^{a,b,2D} = \bigcup_{\ell \in \mathbb{T}_L} \{\xi_{\ell,v}^{a,b}\} \quad \text{and} \quad \mathcal{Q}_v^{a,b,2D} = \bigcup_{\ell \in \mathbb{T}_L} \bigcup_{k \in \mathbb{T}_K} \bigcup_{h \in [1, K-1]} \{\xi_{\ell,v;k,h}^{a,b,\pm}\},$$

and then define  $\mathcal{R}_v^{2D} = \bigcup_{a,b \in S} \mathcal{R}_v^{a,b,2D}$  and  $\mathcal{Q}_v^{2D} = \bigcup_{a,b \in S} \mathcal{Q}_v^{a,b,2D}$ . A configuration in  $\mathcal{R}_v^{2D}$  is called a 2D regular configuration.

- **(Case  $K = L$ )** Define an operator  $\Theta : \mathcal{X}^{2D} \rightarrow \mathcal{X}^{2D}$  as a transpose operator, i.e.,

$$(\Theta(\sigma))(k, \ell) = \sigma(\ell, k) \quad ; \quad k \in \mathbb{T}_K \quad \text{and} \quad \ell \in \mathbb{T}_L. \quad (5.4)$$

Denote temporarily by  $\tilde{\mathcal{C}}^{a,b,2D}$  the collection  $\mathcal{C}^{a,b,2D}$  defined in the case of  $K < L$  above. Then for  $a, b \in S$ , we define the collections of 2D canonical configurations between  $s_a^{2D}$  and  $s_b^{2D}$  as

$$\mathcal{C}^{a,b,2D} = \tilde{\mathcal{C}}^{a,b,2D} \cup \Theta(\tilde{\mathcal{C}}^{a,b,2D}) \quad \text{and} \quad \mathcal{C}^{2D} = \bigcup_{a,b \in S} \mathcal{C}^{a,b,2D}.$$

Similarly, we may define the collections  $\mathcal{R}_v^{a,b,2D}$ ,  $\mathcal{R}_v^{2D}$ ,  $\mathcal{Q}_v^{a,b,2D}$ , and  $\mathcal{Q}_v^{2D}$ .

### Canonical paths

Now, we explain natural optimal paths between monochromatic configurations (illustrated in Figure 2) that consist of canonical configurations.

**Definition 5.5** (2D canonical paths). *The definition below relies on Notation 5.2.*

1. For  $P, P' \in \mathfrak{S}_L$  with  $P \prec P'$ , a sequence  $(A_k)_{k=0}^K$  of subsets of  $\Lambda^{2D}$  is a standard sequence connecting  $\mathbb{T}_K \times P$  and  $\mathbb{T}_K \times P'$  if there exists an increasing sequence  $(Q_k)_{k=0}^K$  in  $\mathfrak{S}_K$  such that

$$A_k = (\mathbb{T}_K \times P) \cup [Q_k \times (P' \setminus P)] \quad ; \quad k \in [0, K].$$

2. A sequence  $(A_n)_{n=0}^{KL}$  of subsets of  $\Lambda^{2D}$  is a standard sequence connecting  $\emptyset$  and  $\Lambda^{2D}$  if there exists an increasing sequence  $(P_\ell)_{\ell=0}^L$  in  $\mathfrak{S}_L$  such that  $A_{K\ell} = \mathbb{T}_K \times P_\ell$  for all  $\ell \in \llbracket 0, L \rrbracket$ , and furthermore for each  $\ell \in \llbracket 0, L-1 \rrbracket$  the subsequence  $(A_k)_{k=K\ell}^{K(\ell+1)}$  is a standard sequence connecting  $\mathbb{T}_K \times P_\ell$  and  $\mathbb{T}_K \times P_{\ell+1}$ .
3. For  $a, b \in S$ , a sequence  $(\omega_n)_{n=0}^{KL}$  of 2D configurations is called a pre-canonical path from  $s_a^{2D}$  to  $s_b^{2D}$  if there exists a standard sequence  $(A_n)_{n=0}^{KL}$  connecting  $\emptyset$  and  $\Lambda^{2D}$  such that

$$\omega_n(x) = \begin{cases} a & \text{if } x \notin A_n, \\ b & \text{if } x \in A_n. \end{cases}$$

4. Moreover, a sequence  $(\omega_n)_{n=0}^{KL}$  of 2D configurations is called a canonical path (cf. Figure 2) connecting  $s_a^{2D}$  and  $s_b^{2D}$  if there exists a pre-canonical path  $(\tilde{\omega}_n)_{n=0}^{KL}$  such that

- (a) **(Case  $K < L$ )**  $\omega_n = \tilde{\omega}_n$  for all  $n \in \llbracket 0, KL \rrbracket$ ,
- (b) **(Case  $K = L$ )**  $\omega_n = \tilde{\omega}_n$  for all  $n \in \llbracket 0, KL \rrbracket$  or  $\omega_n = \Theta(\tilde{\omega}_n)$  for all  $n \in \llbracket 0, KL \rrbracket$ .

It holds that  $H^{2D}(\eta) \leq 2K + 2$  for all  $\eta \in \mathcal{C}^{a,b,2D}$  and

$$H^{2D}(\eta) = \begin{cases} 2K & \text{if } \eta \in \mathcal{R}_v^{a,b,2D} \text{ for } v \in \llbracket 1, L-1 \rrbracket, \\ 2K + 2 & \text{if } \eta \in \mathcal{Q}_v^{a,b,2D} \text{ for } v \in \llbracket 1, L-2 \rrbracket. \end{cases} \quad (5.5)$$

Moreover, the following lemma is immediate.

**Lemma 5.6** ([27, Lemma 6.12]). For a 2D canonical path  $(\omega_n)_{n=0}^{KL}$  connecting  $s_a^{2D}$  and  $s_b^{2D}$ , it holds that

$$\max_{n \in \llbracket 0, KL \rrbracket} H^{2D}(\omega_n) = \Gamma^{2D} = 2K + 2.$$

### Comment on depth of valleys

We conclude this subsection with an application of Definition 5.5 and Lemma 5.6 that is crucially used later to calculate the 3D valley depths.

**Lemma 5.7** ([27, Lemma B.4]). Let  $\eta \in \mathcal{X}^{2D}$  and  $a \in S$ . For any standard sequence  $(A_k)_{k=0}^{KL}$  of sets connecting  $\emptyset$  and  $\Lambda^{2D}$  and for  $n \in \llbracket 0, KL \rrbracket$ , we define  $\omega_n \in \mathcal{X}^{2D}$  as

$$\omega_n(x) = \begin{cases} a & \text{if } x \in A_n, \\ \eta(x) & \text{if } x \in \Lambda^{2D} \setminus A_n. \end{cases}$$

Then, we have that  $H^{2D}(\omega_n) \leq H^{2D}(\eta) + \Gamma^{2D}$  for all  $n \in \llbracket 0, KL \rrbracket$ .

In Lemma 5.7, we have  $\omega_{KL} = s_a^{2D} \in \mathcal{S}^{2D}$  which implies that every  $\eta \in \mathcal{X}^{2D}$  is connected to each ground state in  $\mathcal{S}^{2D}$  with maximum energy  $H^{2D}(\eta) + \Gamma^{2D}$ . This fact implies that the maximum depth of valleys in the 2D energy landscape is  $\Gamma^{2D}$ .

It can be further proved that only the valleys containing the ground states have maximum depth  $\Gamma^{2D}$ , and all the other valleys have depth strictly less than  $\Gamma^{2D}$ . Indeed, this is a necessary condition for the pathwise approach technique to metastability; however, this level of precision is not necessarily needed in our investigation of the 3D energy landscape. Thus, we do not go further into this direction and refer the interested readers to [42, Theorem 2.1-(ii)].

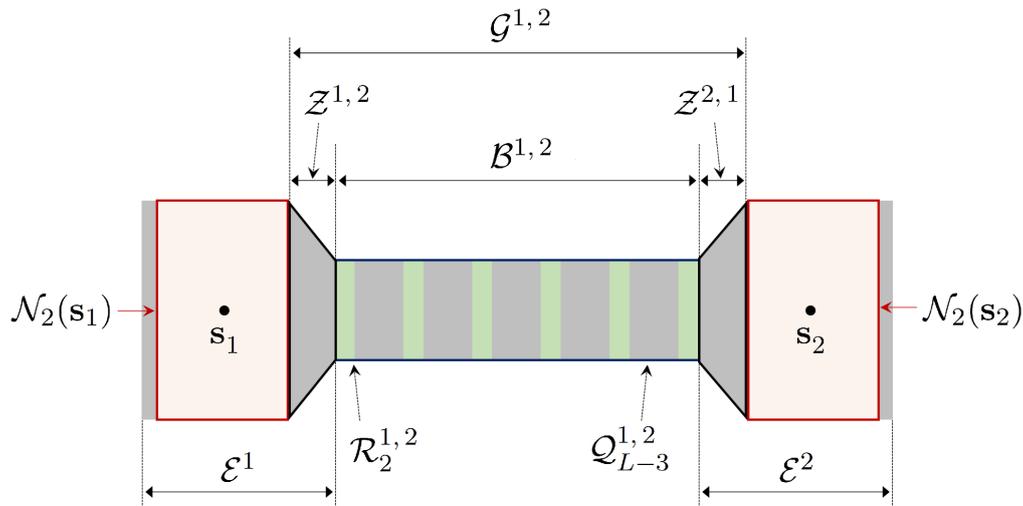


Figure 3: **Saddle structure of the 2D Ising model with  $S = \{1, 2\}$ .** For simplicity, we drop the superscripts 2D in this figure.

### 5.3 Saddle structure

Crucial configurations in the description of the saddle structure of the 2D model is the so-called typical configurations, which turn out to be the elements of the extended neighborhood  $\widehat{\mathcal{N}}^{2D}(\mathcal{S}^{2D})$  (cf. Proposition 5.9 below). We present in Figure 3 an illustration of the saddle structure explained in this subsection.

**Definition 5.8** (2D typical configurations). *There are two different types of typical configurations: the bulk and edge typical configurations.*

- For  $a, b \in S$ , the collection of bulk typical configurations (between  $s_a^{2D}$  and  $s_b^{2D}$ ) is defined by

$$\mathcal{B}^{a,b,2D} = \bigcup_{v \in \llbracket 2, L-2 \rrbracket} \mathcal{R}_v^{a,b,2D} \cup \bigcup_{v \in \llbracket 2, L-3 \rrbracket} \mathcal{Q}_v^{a,b,2D}. \quad (5.6)$$

Then, we write  $\mathcal{B}^{2D} = \bigcup_{a,b \in S} \mathcal{B}^{a,b,2D}$ .

- Next, define

$$\mathcal{B}_\Gamma^{a,b,2D} = \bigcup_{v \in \llbracket 2, L-3 \rrbracket} \mathcal{Q}_v^{a,b,2D} \quad \text{and} \quad \mathcal{B}_\Gamma^{2D} = \bigcup_{a,b \in S} \mathcal{B}_\Gamma^{a,b,2D} = \bigcup_{a,b \in S} \bigcup_{v \in \llbracket 2, L-3 \rrbracket} \mathcal{Q}_v^{a,b,2D}. \quad (5.7)$$

Then, for  $a \in S$ , the collection of edge typical configurations with respect to  $s_a^{2D}$  is defined by

$$\mathcal{E}^{a,2D} = \widehat{\mathcal{N}}^{2D}(s_a^{2D}; \mathcal{B}_\Gamma^{2D}). \quad (5.8)$$

Finally, we write  $\mathcal{E}^{2D} = \bigcup_{a \in S} \mathcal{E}^{a,2D}$ .

Then, the following crucial proposition provides the picture of the saddle structure of the 2D model. We shall provide a similar result for the 3D model in Proposition 9.6.

**Proposition 5.9** ([27, Proposition 6.16]). 1. For spins  $a, b, c \in S$ , we have

$$\mathcal{E}^{a,2D} \cap \mathcal{E}^{b,2D} = \emptyset, \quad \mathcal{E}^{a,2D} \cap \mathcal{B}^{a,b,2D} = \mathcal{R}_2^{a,b,2D} \quad \text{and} \quad \mathcal{E}^{a,2D} \cap \mathcal{B}^{b,c,2D} = \emptyset.$$

2. It holds that  $\mathcal{E}^{2D} \cup \mathcal{B}^{2D} = \widehat{\mathcal{N}}^{2D}(\mathcal{S}^{2D})$ .

**Gateway configurations**

Next, we introduce the gateway configurations.

**Definition 5.10** (2D gateway configurations). *Fix  $a, b \in S$ . Define*

$$\mathcal{Z}^{a,b,2D} := \{ \eta \in \mathcal{X}^{2D} : \text{There exists a path } (\omega_t)_{t=0}^T \text{ in } \mathcal{X}^{2D} \setminus \mathcal{B}_\Gamma^{2D} \text{ with } T \geq 1 \text{ such that } \omega_0 \in \mathcal{R}_2^{a,b,2D}, \omega_T = \eta \text{ and } H^{2D}(\omega_t) = \Gamma^{2D} \text{ for all } t \in \llbracket 1, T \rrbracket \} . \tag{5.9}$$

*Intuitively, this set is the collection of saddle configurations between  $\mathcal{R}_2^{a,b,2D}$  and  $s_a^{2D}$ . Then, we recall the 2D gateway configurations [27, Section B.5]. The gateway between  $s_a^{2D}$  and  $s_b^{2D}$  is denoted as*

$$\mathcal{G}^{a,b,2D} = \mathcal{Z}^{a,b,2D} \cup \mathcal{B}^{a,b,2D} \cup \mathcal{Z}^{b,a,2D} , \tag{5.10}$$

*which is a decomposition of  $\mathcal{G}^{a,b,2D}$ . A configuration belonging to  $\mathcal{G}^{a,b,2D}$  is called a gateway configuration between  $s_a^{2D}$  and  $s_b^{2D}$ .*

Here,  $\mathcal{G}^{a,b,2D}$  is named the collection of gateway configurations because of the following lemma, which indicates that it indeed contains the saddle configurations between  $s_a^{2D}$  and  $s_b^{2D}$ .

**Lemma 5.11** ([27, Lemma B.10]). *For  $a, b \in S$ , suppose that two 2D configurations  $\eta$  and  $\xi$  satisfy*

$$\eta \in \mathcal{G}^{a,b,2D}, \xi \notin \mathcal{G}^{a,b,2D}, \eta \sim \xi, \text{ and } H^{2D}(\xi) \leq \Gamma^{2D} .$$

*Then, we have either  $\xi \in \mathcal{N}^{2D}(s_a^{2D})$  and  $\eta \in \mathcal{Z}^{a,b,2D}$  or  $\xi \in \mathcal{N}^{2D}(s_b^{2D})$  and  $\eta \in \mathcal{Z}^{b,a,2D}$ . In particular,  $\eta \notin \mathcal{B}^{a,b,2D}$ .*

We note that the construction of regular, canonical, typical, and gateway configurations, as well as canonical paths for the 2D model, will be extended to the 3D model in the remainder of the article.

**5.4 Test function**

We also recall the 2D test function defined in [27, Section 7]. Although the construction therein was carried out for both Ising and Potts models, we only need the objects for the Ising model in this article. Hence, in this subsection, **we assume that  $q = 2$** .

Recall that we always assume  $K \leq L$ . We recall a constant

$$\kappa^{2D} = \kappa^{2D}(K, L) \tag{5.11}$$

from [27, (4.13)], which plays the role of  $\kappa$  in the current article and also satisfies

$$\lim_{K \rightarrow \infty} \kappa^{2D}(K, L) = \begin{cases} 1/4 & \text{if } K < L , \\ 1/8 & \text{if } K = L . \end{cases} \tag{5.12}$$

In [27, Definition 7.2], a test function  $\tilde{h}^{2D} : \mathcal{X}^{2D} \rightarrow \mathbb{R}$  (corresponding to  $\tilde{h}$  of the 3D model introduced in Proposition 3.2) is constructed as an  $H^1$ -approximation of the equilibrium potential between two ground states. We proclaim that this function is crucially used in the construction of the 3D test function  $\tilde{h}$ . In the proof of Proposition 3.2, some estimates of  $\tilde{h}^{2D}$  are crucially used. The next estimate is used in the proof of (3.20).

**Proposition 5.12** ([27, Proposition C.1]). *There exists a function  $\tilde{h}^{2D} : \mathcal{X}^{2D} \rightarrow \mathbb{R}$  such that*

$$\sum_{\{\eta, \xi\} \subseteq \mathcal{X}^{2D} : \{\eta, \xi\} \cap \mathcal{G}^{1,2,2D} \neq \emptyset} \mu_\beta^{2D}(\eta) r_\beta^{2D}(\eta, \xi) \{ \tilde{h}^{2D}(\xi) - \tilde{h}^{2D}(\eta) \}^2 = \frac{1 + o_\beta(1)}{2\kappa^{2D}} e^{-\Gamma^{2D}\beta} .$$

The next one is crucially used in the proof of (3.19).

**Proposition 5.13** ([27, Lemmas 7.10-7.16]). *1. For all  $\eta \in \mathcal{X}^{2D} \setminus \mathcal{N}^{2D}(S^{2D})$ , it holds that*

$$\sum_{\xi \in \mathcal{X}^{2D}} \mu_{\beta}^{2D}(\eta) r_{\beta}^{2D}(\eta, \xi) [\tilde{h}^{2D}(\eta) - \tilde{h}^{2D}(\xi)] = o_{\beta}(e^{-\Gamma^{2D}\beta}).$$

*2. We have that*

$$\begin{aligned} \sum_{\eta \in \mathcal{N}^{2D}(s_1^{2D})} \sum_{\xi \in \mathcal{X}^{2D}} \mu_{\beta}^{2D}(\eta) r_{\beta}^{2D}(\eta, \xi) [\tilde{h}^{2D}(\eta) - \tilde{h}^{2D}(\xi)] &= (1 + o_{\beta}(1)) \times \frac{1}{2\kappa^{2D}} e^{-\Gamma^{2D}\beta}, \\ \sum_{\eta \in \mathcal{N}^{2D}(s_2^{2D})} \sum_{\xi \in \mathcal{X}^{2D}} \mu_{\beta}^{2D}(\eta) r_{\beta}^{2D}(\eta, \xi) [\tilde{h}^{2D}(\eta) - \tilde{h}^{2D}(\xi)] &= -(1 + o_{\beta}(1)) \times \frac{1}{2\kappa^{2D}} e^{-\Gamma^{2D}\beta}. \end{aligned}$$

### 5.5 Auxiliary results

In this subsection, we summarize two auxiliary results of the 2D model that are crucially used in our arguments.

#### Bridges, crosses and a bound on 2D Hamiltonian

For a configuration  $\eta \in \mathcal{X}^{2D}$ , a *bridge*, which is a *horizontal* or *vertical bridge*, is a row or column, respectively, in which all spins are the same. If a bridge consists of spin  $a \in S$ , we call this bridge an  $a$ -bridge. Then, we denote by  $B_a(\eta)$  the number of  $a$ -bridges with respect to  $\eta$ . A *cross* (resp.  $a$ -*cross*) is the union of a horizontal bridge and a vertical bridge (resp.  $a$ -bridges). With this notation, we have the following lower bound.

**Lemma 5.14** ([27, Lemma B.2]). *It holds that*

$$H^{2D}(\eta) \geq 2 \left[ K + L - \sum_{a \in S} B_a(\eta) \right].$$

#### Characterization of configurations with low energy

Let  $a \in S$ . For  $\eta \in \mathcal{X}^{2D}$  and  $\sigma \in \mathcal{X}$  (a 3D configuration), we write

$$\|\eta\|_a = \sum_{x \in \Lambda^{2D}} \mathbf{1}\{\eta(x) = a\} \quad \text{and} \quad \|\sigma\|_a = \sum_{x \in \Lambda} \mathbf{1}\{\sigma(x) = a\}. \quad (5.13)$$

The following proposition characterizes all the 2D configurations with energy less than  $\Gamma^{2D}$ .

**Proposition 5.15** ([27, Proposition B.3]). *Suppose that  $\eta \in \mathcal{X}^{2D}$  satisfies  $H^{2D}(\eta) < \Gamma^{2D}$ . Then,  $\eta$  satisfies exactly one of the following properties.*

- **(L1)** *There exist  $a, b \in S$  and  $v \in \llbracket 2, L - 2 \rrbracket$  such that  $\eta \in \mathcal{R}_v^{a, b, 2D}$ . Here,  $\mathcal{N}^{2D}(\eta) = \{\eta\}$ .*
- **(L2)** *There exist  $a, b \in S$  such that  $\eta \in \mathcal{R}_1^{a, b, 2D}$ . In this case,  $\mathcal{N}^{2D}(\eta) = \mathcal{N}^{2D}(s_a^{2D})$ .*
- **(L3)** *For some  $a \in S$ ,  $\eta$  has an  $a$ -cross. Then,  $\mathcal{N}^{2D}(\eta) = \mathcal{N}^{2D}(s_a^{2D})$  and*

$$\sum_{b \neq a} \|\eta\|_b \leq \frac{H^{2D}(\eta)^2}{16} \leq \frac{(2K + 1)^2}{16}. \quad (5.14)$$

## 6 Canonical configurations and paths

Analyzing the energy landscape of the 3D model is far more complex than that of the 2D model; below, we briefly list the main differences between them that serve to complexify the problem.

1. In the 2D model, the energy of the gateway configuration is either  $\Gamma^{2D}$  or  $\Gamma^{2D} - 2$ . Thus, a  $\Gamma^{2D}$ -path on the gateway configurations does not have the freedom to move. On the other hand, in the 3D model, the energy of the gateway configuration ranges from  $\Gamma - 2K - 2$  to  $\Gamma$ . This implies that the behavior of a  $\Gamma$ -path around a gateway configuration of energy  $\Gamma - 2K - 2$  (which is a regular configuration) cannot be characterized precisely.
2. In the 2D model, a  $\Gamma^{2D}$ -path from  $s_a^{2D}$  to  $s_b^{2D}$  must visit a configuration in  $\mathcal{R}_2^{a,b,2D}$ . Then, it successively visits  $\mathcal{R}_3^{a,b,2D}, \dots, \mathcal{R}_{L-2}^{a,b,2D}$  and finally arrives at  $s_b^{2D}$ . Remarkably, this path does not need to visit a configuration in  $\mathcal{R}_1^{a,b,2D}$  and in  $\mathcal{R}_{L-1}^{a,b,2D}$ ; this fact essentially arises from the features of the 2D geometry. In the 3D model, we observe a similar phenomenon. To explain this, let us temporarily denote by  $\mathcal{R}_v^{a,b}$ ,  $v \in \llbracket 1, L-1 \rrbracket$  the collection of 3D configurations such that there are  $v$  consecutive  $K \times L$  slabs of spins  $b$  and such that the spins at the remaining sites are  $a$ . Then, there exists an integer  $n = n_{K,L,M}$  such that any  $\Gamma$ -path connecting  $s_a$  and  $s_b$  must successively visit configurations in  $\mathcal{R}_n^{a,b}, \mathcal{R}_{n+1}^{a,b}, \dots, \mathcal{R}_{M-n}^{a,b}$  but need not visit  $\mathcal{R}_i^{a,b}$  for  $i \in \llbracket 1, n-1 \rrbracket$  and  $i \in \llbracket M-n+1, M-1 \rrbracket$ . In the 2D model, the number corresponding to this  $n = n_{K,L,M}$  is 2. We guess that in the 3D model,  $n \sim K^{1/2}$ ; however, we cannot determine the exact value of  $n$ . This fact reveals the complex structure of the energy landscape in the 3D model. Instead, we prove below (cf. Propositions 6.14 and 8.1) that

$$\lfloor K^{1/2} \rfloor \leq n \leq \lfloor K^{2/3} \rfloor.$$

Fortunately, this bound suffices to complete our analysis without identifying the exact value of  $n$ .

3. In the 2D model, the  $\mathcal{N}^{2D}$ -neighborhoods are fully characterized in Proposition 5.15; meanwhile, in the 3D case, we cannot obtain such a specific and simple result. We overcome the absence of this result by using the 2D result obtained in Proposition 5.15, through suitably applying it to the analysis of the 3D model. *Indeed, this absence is a crucial difficulty in extending the analysis to the four- or higher-dimensional models.*
4. Because of the aforementioned complexity of the energy landscape, the transition may encounter a dead-end with energy  $\Gamma$ , even in the bulk part of the transition; this is not the case in the 2D model. Therefore, another technical challenge is that of carefully characterizing these dead-ends and appropriately excluding them from the computation.

As explained above, the energy landscape of the 3D model is more complex than that of the 2D one, and we are unable to present a complete description of the energy landscape for the former. Nevertheless, we analyze the landscape with the precision required to prove our main results.

In Section 6, we introduce canonical configurations and paths. Their definitions are direct generalizations of those in the 2D model. Then, we explain several applications of these canonical objects. For simplicity, we temporarily denote by

$$\Gamma^* := 2KL + K + 2.$$

We first collect several notation which will be frequently used throughout the remainder of the article.

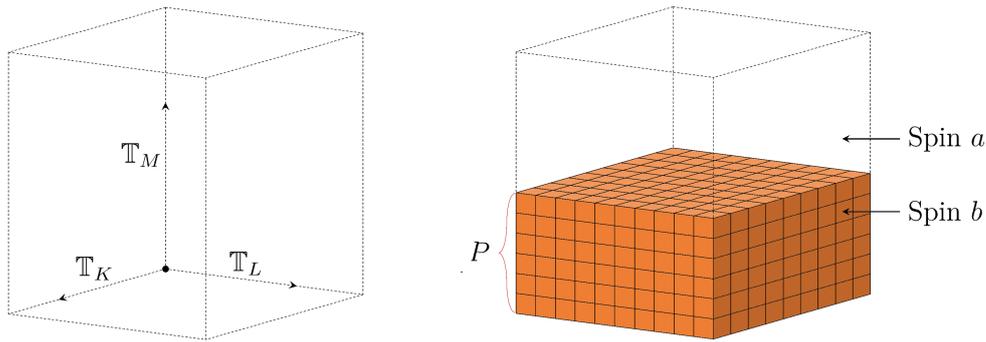


Figure 4: **Figures on Notation 6.1.** This form of figure is used throughout the remainder of the article to illustrate a 3D configuration consisting of two types of spins only. The large dotted box denotes  $\Lambda = \mathbb{T}_K \times \mathbb{T}_L \times \mathbb{T}_M$ . The orange unit boxes denote the sites with spin  $b$ , and the empty part denotes the cluster of spin  $a$ . For some cases when we only concern the shape of the cluster of spin  $b$  (e.g. in Figure 5), we omit the dotted box representing  $\Lambda$ .

**Notation 6.1.** We refer to Figure 4<sup>4</sup> for an illustration of the notation below.

- For  $m \in \mathbb{T}_M$ , the slab  $\mathbb{T}_K \times \mathbb{T}_L \times \{m\} \subseteq \Lambda$  is called an  $m$ -th floor. For each configuration  $\sigma \in \mathcal{X}$ , we denote by  $\sigma^{(m)}$  the configuration of  $\sigma$  at the  $m$ -th floor, i.e.,

$$\sigma^{(m)}(k, \ell) = \sigma(k, \ell, m) \quad ; \quad k \in \mathbb{T}_K, \ell \in \mathbb{T}_L. \quad (6.1)$$

Thus,  $\sigma^{(m)} \in \mathcal{X}^{2D}$  is a spin configuration in  $\Lambda^{2D} = \mathbb{T}_K \times \mathbb{T}_L$ .

- For  $a, b \in S$  and  $P \subseteq \mathbb{T}_M$ , we denote by  $\sigma_P^{a,b} \in \mathcal{X}$  the configuration satisfying

$$\sigma_P^{a,b}(k, \ell, m) = b \cdot \mathbf{1}\{m \in P\} + a \cdot \mathbf{1}\{m \notin P\}. \quad (6.2)$$

### 6.1 Canonical configurations

The following notation is used frequently.

**Notation 6.2.** We first introduce several maps on  $\mathcal{X}$ . If  $K = L$ , we define a bijection  $\Theta^{(12)} : \mathcal{X} \rightarrow \mathcal{X}$  as the map switching the first and second coordinates, i.e., for all  $\sigma \in \mathcal{X}$  and  $(k, \ell, m) \in \Lambda$ ,

$$(\Theta^{(12)}(\sigma))(k, \ell, m) = \sigma(\ell, k, m).$$

If  $L = M$ , we can similarly define a bijection  $\Theta^{(23)}$  on  $\mathcal{X}$  switching the second and third coordinates. Finally, for the case of  $K = L = M$ , we can even define the bijection  $\Theta^{(13)}$  on  $\mathcal{X}$  switching the first and third coordinates.

Then, for  $\mathcal{A} \subseteq \mathcal{X}$ , we define  $\Upsilon(\mathcal{A})$  as

$$\Upsilon(\mathcal{A}) = \begin{cases} \mathcal{A} & \text{if } K < L < M, \\ \mathcal{A} \cup \Theta^{(12)}(\mathcal{A}) & \text{if } K = L < M, \\ \mathcal{A} \cup \Theta^{(23)}(\mathcal{A}) & \text{if } K < L = M, \\ \mathcal{A} \cup \Theta^{(12)}(\mathcal{A}) \cup \Theta^{(23)}(\mathcal{A}) \cup \Theta^{(13)}(\mathcal{A}) & \text{if } K = L = M. \\ \cup (\Theta^{(12)} \circ \Theta^{(23)})(\mathcal{A}) \cup (\Theta^{(23)} \circ \Theta^{(12)})(\mathcal{A}) & \end{cases}$$

Note that the set  $\Upsilon(\mathcal{A})$  for the case of  $K = L = M$  denotes the set of all configurations obtained by permuting the coordinates of the configurations in  $\mathcal{A}$ .

<sup>4</sup>In fact, this figure and all the 3D figures below contradict our assumption that  $K \geq 2829$ . However, we believe that there will be absolutely no confusion with these figures which only provide simple illustrations of complicated notions.

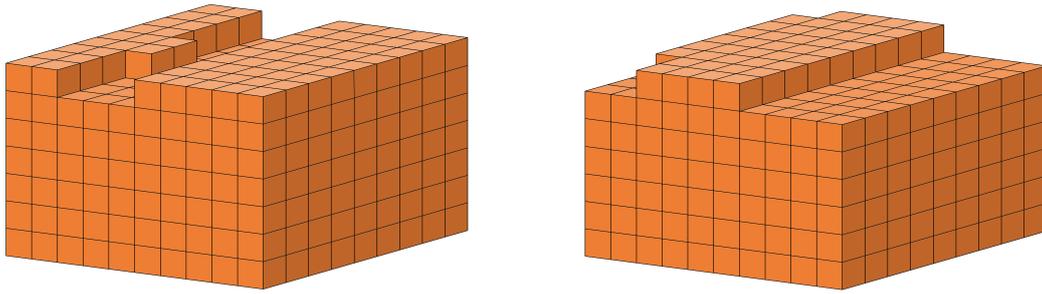


Figure 5: **Canonical configurations.** These two configurations belong to  $\mathcal{C}_6^{a,b}$  (if the orange boxes represent the sites with spin  $b$  as in Figure 4), since the 2D configurations at the 7-th floor are 2D canonical configurations  $\xi_{6,7;5,2}^{a,b,+}$  and  $\xi_{3,4;2,6}^{a,b,-}$ , respectively.

Now, we define canonical configurations of our 3D model.

**Definition 6.3** (Canonical configurations). We refer to Figure 5 for a visualization of the objects introduced below. Recall Notation 5.2.

1. We first introduce some building blocks in the definition of canonical and gateway configurations. For  $a, b \in S$  and  $P, Q \in \mathfrak{S}_M$  with  $P \prec Q$ , we define  $\tilde{\mathcal{C}}_{P,Q}^{a,b} \subseteq \mathcal{X}$  as

$$\sigma \in \tilde{\mathcal{C}}_{P,Q}^{a,b} \Leftrightarrow \begin{cases} \sigma^{(m)} = \mathbf{s}_b^{2D} & \text{if } m \in P, \\ \sigma^{(m)} = \mathbf{s}_a^{2D} & \text{if } m \in Q^c, \\ \sigma^{(m)} \in \mathcal{C}^{a,b,2D} & \text{if } m \in Q \setminus P, \end{cases}$$

where the 2D objects are defined in Section 5.2. Then, we set

$$\mathcal{C}_{P,Q}^{a,b} = \Upsilon(\tilde{\mathcal{C}}_{P,Q}^{a,b}). \tag{6.3}$$

We then define, for  $i \in \llbracket 0, M-1 \rrbracket$ ,

$$\mathcal{C}_i^{a,b} = \bigcup_{P, Q \in \mathfrak{S}_M: |P|=i \text{ and } P \prec Q} \mathcal{C}_{P,Q}^{a,b} \quad \text{and} \quad \mathcal{C}^{a,b} = \bigcup_{i=0}^{M-1} \mathcal{C}_i^{a,b}. \tag{6.4}$$

Finally, for a proper partition  $(A, B)$  of  $S$ , we write

$$\mathcal{C}_i^{A,B} = \bigcup_{a \in A} \bigcup_{b \in B} \mathcal{C}_i^{a,b} \quad \text{and} \quad \mathcal{C}^{A,B} = \bigcup_{a \in A} \bigcup_{b \in B} \mathcal{C}^{a,b}.$$

A configuration belonging to  $\mathcal{C}^{a,b}$  for some  $a, b \in S$  is called a canonical configuration between  $\mathbf{s}_a$  and  $\mathbf{s}_b$ .

In view of the definition above, the role of the map  $\Upsilon$  is clear. When  $K < L < M$  there is only one direction of transition, if  $K = L < M$  or  $K < L = M$  there are  $2 = 2!$  possible directions, while if  $K = L = M$  there are  $6 = 3!$  possible directions. The map  $\Upsilon$  reflects this observation into the definition. Next, let us define regular configurations which are the special ones among the canonical configurations.

**Definition 6.4** (Regular configurations). For  $a, b \in S$  and  $P \in \mathfrak{S}_M$ , recall the configuration  $\sigma_P^{a,b}$  from (6.2) and define

$$\tilde{\mathcal{R}}_i^{a,b} = \{\sigma_P^{a,b} : P \in \mathfrak{S}_M, |P| = i\} \quad ; \quad i \in \llbracket 0, M \rrbracket. \tag{6.5}$$

Note that  $\tilde{\mathcal{R}}_i^{a,b}$  is a collection of configurations consisting of spins  $a$  and  $b$  only, where spins  $a$  and  $b$  are located at slabs  $\mathbb{T}_K \times \mathbb{T}_L \times (\mathbb{T}_M \setminus P)$  and  $\mathbb{T}_K \times \mathbb{T}_L \times P$ , respectively, for  $P \in \mathfrak{S}_M$  with  $|P| = i$ . Then, define (cf. Notation 6.2)

$$\mathcal{R}_i^{a,b} = \Upsilon(\tilde{\mathcal{R}}_i^{a,b}) . \tag{6.6}$$

A configuration belonging to  $\mathcal{R}_i^{a,b}$  for some  $i \in \llbracket 0, M \rrbracket$  is called a regular configuration. Clearly, we have  $\mathcal{R}_0^{a,b} = \{s_a\}$  and  $\mathcal{R}_M^{a,b} = \{s_b\}$ . For a proper partition  $(A, B)$  of  $S$ , we write

$$\mathcal{R}_i^{A,B} = \bigcup_{a \in A} \bigcup_{b \in B} \mathcal{R}_i^{a,b} . \tag{6.7}$$

### 6.2 Energy of canonical configurations

One can compute the energy of canonical configurations readily by elementary computations, but we provide a more systematic approach which will be frequently used in later computations. To this end, we first introduce a notation.

**Notation 6.5.** For  $(k, \ell) \in \mathbb{T}_K \times \mathbb{T}_L$ , we denote by  $\sigma^{(k,\ell)} \in S^{\mathbb{T}_M}$  the configuration of  $\sigma$  on the  $(k, \ell)$ -th pillar  $\{k\} \times \{\ell\} \times \mathbb{T}_M$ , i.e.,

$$\sigma^{(k,\ell)}(m) = \sigma(k, \ell, m) \quad ; \quad m \in \mathbb{T}_M . \tag{6.8}$$

The energy of the one-dimensional (1D) configuration  $\sigma^{(k,\ell)}$  is denoted by

$$H^{1D}(\sigma^{(k,\ell)}) = \sum_{m \in \mathbb{T}_M} \mathbf{1}\{\sigma(k, \ell, m) \neq \sigma(k, \ell, m+1)\} . \tag{6.9}$$

In the following lemma, we decompose the 3D energy into lower-dimensional ones.

**Lemma 6.6.** For each  $\sigma \in \mathcal{X}$ , it holds that

$$H(\sigma) = \sum_{m \in \mathbb{T}_M} H^{2D}(\sigma^{(m)}) + \sum_{(k,\ell) \in \mathbb{T}_K \times \mathbb{T}_L} H^{1D}(\sigma^{(k,\ell)}) . \tag{6.10}$$

*Proof.* We can write  $H(\sigma)$  as

$$\begin{aligned} & \sum_{m \in \mathbb{T}_M} \left[ \sum_{k \in \mathbb{T}_K} \sum_{\ell \in \mathbb{T}_L} \mathbf{1}\{\sigma(k+1, \ell, m) \neq \sigma(k, \ell, m)\} + \mathbf{1}\{\sigma(k, \ell+1, m) \neq \sigma(k, \ell, m)\} \right] \\ & + \sum_{k \in \mathbb{T}_K} \sum_{\ell \in \mathbb{T}_L} \left[ \sum_{m \in \mathbb{T}_M} \mathbf{1}\{\sigma(k, \ell, m) \neq \sigma(k, \ell, m+1)\} \right] . \end{aligned}$$

The first and second lines correspond to the first and second terms at the right-hand side of (6.10), respectively.  $\square$

Based on the previous expression, we deduce the following proposition.

**Proposition 6.7** (Energy of canonical configurations). *The following properties hold.*

1. For each canonical configuration  $\sigma$ , we have  $H(\sigma) \leq \Gamma^*$ .
2. For each configuration  $\sigma \in \mathcal{C}_i^{a,b}$  for some  $a, b \in S$  and  $i \in \llbracket 1, M-2 \rrbracket$ , we have

$$H(\sigma) \in \llbracket \Gamma^* - 2K - 2, \Gamma^* \rrbracket .$$

*Proof.* Observe that for a canonical configuration  $\sigma$ , we have  $H^{1D}(\sigma^{(k,\ell)}) \leq 2$  for all  $(k, \ell) \in \mathbb{T}_K \times \mathbb{T}_L$ , and  $H^{2D}(\sigma^{(m)}) = 0$  for all  $m \in \mathbb{T}_M \setminus \{m_0\}$  for some  $m_0 \in \mathbb{T}_M$ , at which it holds that  $H^{2D}(\sigma^{(m_0)}) \leq 2K + 2$  (cf. (5.5)). Thus, by Lemma 6.6,

$$H(\sigma) \leq (2K + 2) + 2KL = \Gamma^* .$$

## Energy landscape and metastability of 3D Ising/Potts models

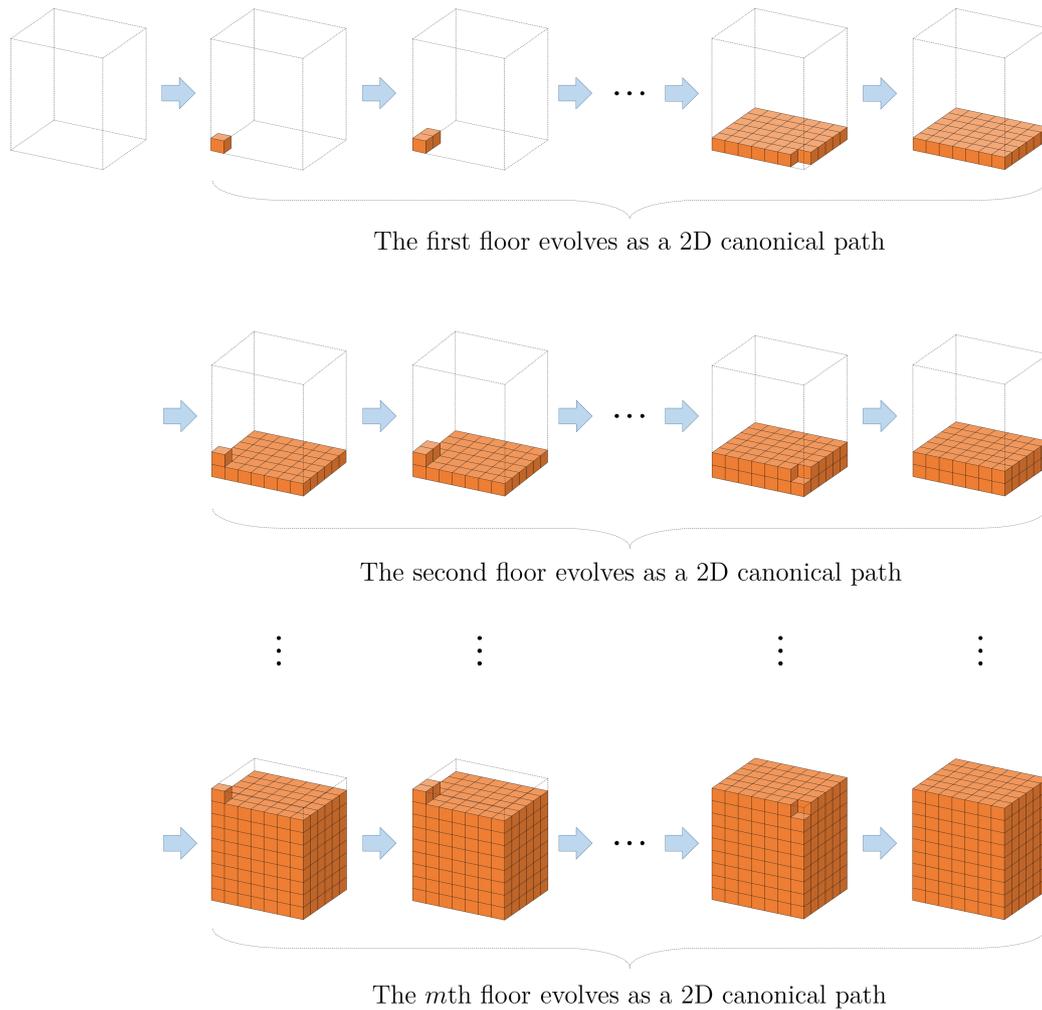


Figure 6: **Canonical path connecting  $s_a$  and  $s_b$ .**

For part (2), it suffices to additionally observe that  $H^{1D}(\sigma^{(k,\ell)}) = 2$  for all  $(k, \ell) \in \mathbb{T}_K \times \mathbb{T}_L$  if  $i \in \llbracket 1, M-2 \rrbracket$  and thus

$$H(\sigma) \geq 2KL = \Gamma^* - 2K - 2.$$

□

**Remark 6.8.** In particular, we have  $H(\sigma) = \Gamma^* - 2K - 2 = 2KL$  for any  $\sigma \in \mathcal{R}_i^{A,B}$ ,  $i \in \llbracket 1, M-1 \rrbracket$ . Hence, a  $\Gamma^*$ -path at a regular configuration can evolve in a non-canonical way, since we still have a spare of  $2K + 2$  to reach the energy barrier  $\Gamma$ . Incorporating all these behaviors in the metastability analysis is a demanding part of the 3D model. For this reason, the regular configuration plays a crucial role. We remark that for the 2D case [7, 27, 42], any optimal path at a regular configuration does not have freedom, and that helped a lot simplifying the arguments.

### 6.3 Canonical paths

In this subsection, we define 3D canonical paths between ground states. They generalize the 2D paths recalled in Definition 5.5. Refer to Figure 6 for an illustration.

**Definition 6.9** (Canonical paths). We recall Notation 5.2. Let us fix  $a, b \in S$ . A path  $(\omega_t)_{t=0}^{KLM}$  is called a pre-canonical path connecting  $s_a$  and  $s_b$  if there exists an increasing sequence  $(P_i)_{i=0}^M$  in  $\mathfrak{S}_M$  such that

- for each  $i \in \llbracket 0, M \rrbracket$ , we have that  $\omega_{KLi} = \sigma_{P_i}^{a,b}$  (cf. (6.2)), and
- for each  $i \in \llbracket 0, M - 1 \rrbracket$ , there exists a 2D canonical path  $(\gamma_t^i)_{t=0}^{KL}$  from  $s_a^{2D}$  to  $s_b^{2D}$  defined in Definition 5.5 such that

$$\omega_t^{(m)} = \begin{cases} s_b^{2D} & \text{if } m \in P_i, \\ s_a^{2D} & \text{if } m \in \mathbb{T}_M \setminus P_{i+1}, \\ \gamma_{t-KLi}^i & \text{if } m \in P_{i+1} \setminus P_i, \end{cases} \quad \text{for all } t \in \llbracket KLi, KL(i+1) \rrbracket.$$

If  $K < L < M$ , a path is called a canonical path if it is a pre-canonical path. If  $K = L < M$ , a path is called a canonical path if it is either a pre-canonical one or the image of a pre-canonical one with respect to the map  $\Theta^{(12)}$ . We can define canonical paths for the cases of  $K < L = M$  and  $K = L = M$  in a similar manner.

**Remark 6.10.** We emphasize that for a canonical path  $(\omega_t)_{t=0}^{KLM}$ , all configurations  $\omega_t$ ,  $t \in \llbracket 0, KLM \rrbracket$ , are canonical configurations, and hence any canonical path is a  $\Gamma$ -path by part (1) of Proposition 6.7.

Canonical paths provide optimal paths between two ground states, and hence we can confirm the following upper bound for the energy barrier.

**Proposition 6.11.** For  $s, s' \in S$ , we have that  $\Phi(s, s') \leq \Gamma^*$ .

*Proof.* By Remark 6.10, it suffices to take a canonical path connecting  $s$  and  $s'$ . □

We prove  $\Phi(s, s') \geq \Gamma$  in Section 8 to verify  $\Phi(s, s') = \Gamma^*$ . This reversed inequality requires a much more complicated proof.

### 6.4 Characterization of the deepest valleys

We show in this subsection that using the canonical paths, the valleys in the energy landscape, except for the ones associated to the ground states, have depths less than  $\Gamma$ . Note that Theorem 2.4, although not yet proved, indicates that the valleys associated to the ground states have depth  $\Gamma$ . This characterization of the depths of other valleys is essentially required since we have to reject the possibility of being trapped in a deeper valley in the course of transition. This fact is crucially used in the application of the pathwise approach to metastability.

**Notation 6.12.** For the convenience of notation, we call  $(\omega_t)_{t=0}^T$  a pseudo-path if either  $\omega_t \sim \omega_{t+1}$  or  $\omega_t = \omega_{t+1}$  for all  $t \in \llbracket 0, T - 1 \rrbracket$ .

**Proposition 6.13.** For  $\sigma \in \mathcal{X} \setminus S$ , we have

$$\Phi(\sigma, S) - H(\sigma) \leq \Gamma^* - 2 < \Gamma^*.$$

*Proof.* Main idea of the proof is inherited from the proof of [42, Theorem 2.1]. Let us find two spins  $a, b \in S$  so that  $\sigma$  has spins  $a$  and  $b$  at some sites, which is clearly possible since  $\sigma \notin S$ . Let us fix a canonical path  $(\omega_t)_{t=0}^{KLM}$  connecting  $s_a$  and  $s_b$ . Then, we write

$$A_t = \{x \in \Lambda : \omega_t(x) = b\} \quad ; \quad t \in \llbracket 0, KLM \rrbracket,$$

so that we have  $\emptyset = A_0 \subseteq A_1 \subseteq \dots \subseteq A_{KLM} = \Lambda$  and  $|A_t| = t$  for all  $t \in \llbracket 0, KLM \rrbracket$ . We can take the path  $(\omega_t)_{t=0}^{KLM}$  in a way that

$$A_1 = \{x_0\} \quad \text{and} \quad \sigma(x_0) = b. \tag{6.11}$$

Now, we define a pseudo-path (cf. Notation 6.12)  $(\tilde{\omega}_t)_{t=0}^{KLM}$  connecting  $\sigma$  and  $s_b$  as

$$\tilde{\omega}_t(x) = \begin{cases} \sigma(x) & \text{if } x \notin A_t, \\ b & \text{if } x \in A_t. \end{cases}$$

In other words, we update the spins in an exactly same manner with the canonical path  $(\omega_t)_{t=0}^{KLM}$ . We claim that

$$H(\tilde{\omega}_t) - H(\sigma) \leq 2KL + 2K = \Gamma^* - 2 \quad \text{for all } t \in \llbracket 0, KLM \rrbracket. \quad (6.12)$$

It is immediate that this claim concludes the proof. To prove this claim, we recall the decomposition obtained in Lemma 6.6 and write  $\tilde{\omega}_t = \zeta$ . Then, we can write  $H(\zeta) - H(\sigma)$  as

$$\sum_{m \in \mathbb{T}_M} [H^{2D}(\zeta^{(m)}) - H^{2D}(\sigma^{(m)})] + \sum_{(k, \ell) \in \mathbb{T}_K \times \mathbb{T}_L} [H^{1D}(\zeta^{(k, \ell)}) - H^{1D}(\sigma^{(k, \ell)})]. \quad (6.13)$$

Let us first consider the first summation of (6.13). We suppose that  $t \in \llbracket KLi, KL(i+1) \rrbracket$  and write  $\omega_{KLi} = \sigma_P^{a, b}$  and  $\omega_{KL(i+1)} = \sigma_Q^{a, b}$  where  $P \prec Q$ . Write  $Q \setminus P = \{m'\}$ . Then, we have that

$$H^{2D}(\zeta^{(m)}) - H^{2D}(\sigma^{(m)}) = \begin{cases} -H^{2D}(\sigma^{(m)}) \leq 0 & \text{if } m \in P, \\ 0 & \text{if } m \in Q^c, \end{cases} \quad (6.14)$$

since  $\zeta^{(m)} = s_b^{2D}$  for  $m \in P$  and  $\zeta^{(m)} = \sigma^{(m)}$  for  $m \in Q^c$ . On the other hand, by Lemma 5.7, we have that

$$H^{2D}(\zeta^{(m')}) - H^{2D}(\sigma^{(m')}) \leq 2K + 2. \quad (6.15)$$

By (6.14) and (6.15), we conclude that

$$\sum_{m \in \mathbb{T}_M} [H^{2D}(\zeta^{(m)}) - H^{2D}(\sigma^{(m)})] \leq 2K + 2. \quad (6.16)$$

Now, we turn to the second summation of (6.13). Note that  $\zeta^{(k, \ell)}$  is obtained from  $\sigma^{(k, \ell)}$  by flipping the spins in consecutive sites in  $Q$  to  $b$ . From this, we can readily deduce that

$$H^{1D}(\zeta^{(k, \ell)}) - H^{1D}(\sigma^{(k, \ell)}) \leq 2 \quad \text{for all } k \in \mathbb{T}_K \text{ and } \ell \in \mathbb{T}_L. \quad (6.17)$$

Moreover, if  $x_0 = (k_0, \ell_0, m_0)$ , we can check that

$$H^{1D}(\zeta^{(k_0, \ell_0)}) - H^{1D}(\sigma^{(k_0, \ell_0)}) \leq 0. \quad (6.18)$$

By (6.17) and (6.18), we get

$$\sum_{(k, \ell) \in \mathbb{T}_K \times \mathbb{T}_L} [H^{1D}(\zeta^{(k, \ell)}) - H^{1D}(\sigma^{(k, \ell)})] \leq 2(KL - 1). \quad (6.19)$$

Now, the claim (6.12) follows from (6.13), (6.16), and (6.19).  $\square$

### 6.5 Auxiliary result on saddle configurations

In the 2D case, in the analysis of the energy landscape, the collection  $\mathcal{R}_2^{2D}$  plays a significant role since to make an optimal transition (not exceeding the energy barrier  $2K + 2$ ), we may skip the collection  $\mathcal{R}_1^{2D}$  but must pass through  $\mathcal{R}_2^{2D}$ . Thus, the integer 2 worked as some kind of a threshold for metastable transitions. We expect a similar pattern in the 3D case, and we briefly explain this phenomenon in this subsection.

Let us define

$$\mathfrak{m}_K = \lfloor K^{2/3} \rfloor. \quad (6.20)$$

Then, we shall prove in Corollary 8.5 below that

$$\Phi(\mathbf{s}_a, \sigma_{\llbracket 1, n \rrbracket}^{a,b}) = \Gamma^* \quad \text{for all } n \in \llbracket \mathfrak{m}_K, M - \mathfrak{m}_K \rrbracket. \quad (6.21)$$

Thus, we can define (cf. Figure 8 below)

$$n_{K,L,M} = \min \{n \in \llbracket 1, M - 1 \rrbracket : \Phi(\mathbf{s}_a, \sigma_{\llbracket 1, n \rrbracket}^{a,b}) = \Gamma^*\}. \quad (6.22)$$

We strongly believe that this quantity does not depend on  $M$ , but we do not have a proof for it at the moment. Note that this number was just 2 in the 2D case. In the 3D model, we do not know this number exactly, since non-canonical movements at the early stage of transitions are hard to characterize. However, the upper bound  $n_{K,L,M} \leq \mathfrak{m}_K = \lfloor K^{2/3} \rfloor$  obtained from (6.21) is enough for our purpose, as we shall see later.

The main result of this subsection is the corresponding lower bound. This result will not be used in the proofs later, but emphasizes the complexity of the energy landscape near ground states.

**Proposition 6.14.** *We have  $n_{K,L,M} \geq \lfloor K^{1/2} \rfloor$ .*

*Proof.* It suffices to prove that

$$\Phi(\mathbf{s}_1, \sigma_{\llbracket 1, n \rrbracket}^{1,2}) \leq \Gamma^* - 2 \quad \text{for all } n \in \llbracket 1, \lfloor K^{1/2} \rfloor - 1 \rrbracket.$$

We fix such an  $n$  and write  $\sigma = \sigma_{\llbracket 1, n \rrbracket}^{1,2}$ . We now construct an explicit path from  $\sigma$  to  $\mathbf{s}_1$  without exceeding the energy  $\Gamma^* - 2$ . Note that  $\sigma_{\llbracket 1, n \rrbracket}^{1,2}$  has spins 2 at  $\mathbb{T}_K \times \mathbb{T}_L \times \llbracket 1, n \rrbracket$  and spins 1 at all the other sites. In this proof, we regard  $\mathbb{T}_K = \llbracket 1, K \rrbracket$  and  $\mathbb{T}_L = \llbracket 1, L \rrbracket$  in order to simplify the explanation of the order of spin flips in a lexicographic manner.

- First, starting from  $\sigma$ , we change spins 2 to 1 in  $\llbracket 1, K \rrbracket \times \llbracket 1, n \rrbracket \times \llbracket 1, n \rrbracket$  in ascending lexicographic order. Denote by  $\zeta \in \mathcal{X}$  the obtained spin configuration, which has spins 2 only on  $\llbracket 1, K \rrbracket \times \llbracket n + 1, L \rrbracket \times \llbracket 1, n \rrbracket$ . Then, the variation of the Hamiltonian from  $\sigma$  to  $\zeta$  can be expressed by the following  $n \times n$  matrices:

$$\begin{bmatrix} +2 & +0 & \cdots & +0 \\ +4 & +2 & \cdots & +2 \\ & & \vdots & \\ +4 & +2 & \cdots & +2 \end{bmatrix}, \quad \begin{bmatrix} +0 & -2 & \cdots & -2 \\ +2 & +0 & \cdots & +0 \\ & & \vdots & \\ +2 & +0 & \cdots & +0 \end{bmatrix} \times (K - 2), \quad \text{and} \quad \begin{bmatrix} -2 & -4 & \cdots & -4 \\ +0 & -2 & \cdots & -2 \\ & & \vdots & \\ +0 & -2 & \cdots & -2 \end{bmatrix}.$$

Here, each  $n \times n$  matrix represents  $\{i\} \times \llbracket 1, n \rrbracket \times \llbracket 1, n \rrbracket$  for  $1 \leq i \leq K$ , in which the numbers represent the variation of the energy which should be read in ascending lexicographic order. From this path, we obtain

$$\Phi(\sigma, \zeta) \leq 2KL + 2n^2 + 2n - 2, \quad (6.23)$$

where the maximum of the energy is obtained right after flipping the spin at  $(2, 1, n - 1)$ , which is denoted by bold font at the matrices above.

- Next, starting from  $\zeta$ , we change spins 2 to 1 in  $\llbracket 1, K \rrbracket \times \{i\} \times \llbracket 1, n \rrbracket$  in the ascending lexicographic order for  $i \in \llbracket n + 1, L - 1 \rrbracket$ , from  $i = n + 1$  to  $i = L - 1$ . Denote by  $\zeta' \in \mathcal{X}$  the obtained spin configuration, which has spins 2 only on  $\llbracket 1, K \rrbracket \times \{L\} \times \llbracket 1, n \rrbracket$ . In each step, the variation of the Hamiltonian is represented by the  $n \times K$  matrix

$$\begin{bmatrix} +0 & -2 & \cdots & -2 & -4 \\ +2 & +0 & \cdots & +0 & -2 \\ & & \vdots & & \\ +2 & +0 & \cdots & +0 & -2 \end{bmatrix}.$$

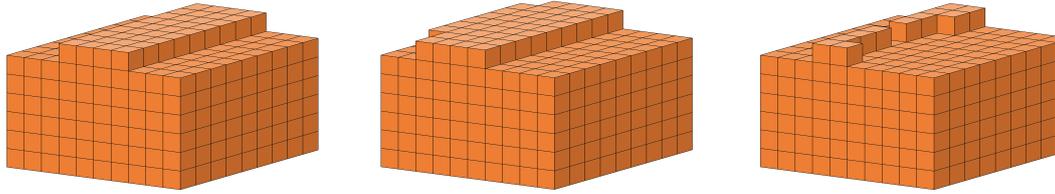


Figure 7: **Examples of gateway configurations.** Each configuration above represents a gateway configuration of type 1 (left), type 2 (middle), or type 3 (right), respectively.

Since  $H(\zeta) = 2KL$ , we can verify that

$$\Phi(\zeta, \zeta') \leq 2KL + 2, \tag{6.24}$$

where the maximum is obtained right after flipping the spin at  $(1, n + 1, 1)$  (cf. bold font  $+2$ ).

- Finally, starting from  $\zeta'$ , we change spins 2 to 1 in the ascending lexicographic order. The variation of the Hamiltonian is represented by

$$\begin{bmatrix} -2 & -4 & \cdots & -4 & -6 \\ +0 & -2 & \cdots & -2 & -4 \\ & & \vdots & & \\ +0 & -2 & \cdots & -2 & -4 \end{bmatrix}.$$

Hence, the Hamiltonian monotonically decreases from  $H(\zeta') = 2K(n + 1)$  to arrive at  $H(s_1) = 0$ . Hence, we have

$$\Phi(\zeta', s_1) \leq 2K(n + 1). \tag{6.25}$$

Therefore, by (6.23), (6.24), and (6.25), we have

$$\Phi(\sigma, s_1) \leq 2KL + 2n^2 + 2n - 2.$$

Since  $n \in \llbracket 1, \lfloor K^{1/2} \rfloor - 1 \rrbracket$ , it holds that  $2n^2 + 2n - 2 \leq 2K$ . This concludes the proof.  $\square$

## 7 Gateway configurations

In the analysis of the 3D model, a crucial notion is the concept of gateway configurations. The gateway configurations of the 3D model play a far more significant role than those of the 2D model.

We fix a proper partition  $(A, B)$  of  $S$  throughout this section.

### 7.1 Gateway configurations

We refer to Figure 7 for an illustration of gateway configurations defined below.

**Definition 7.1** (Gateway configurations). For  $a, b \in S$  and  $P, Q \in \mathfrak{S}_M$  with  $P \prec Q$ , we define  $\tilde{\mathcal{G}}_{P,Q}^{a,b} \subseteq \tilde{\mathcal{C}}_{P,Q}^{a,b}$  as

$$\sigma \in \tilde{\mathcal{G}}_{P,Q}^{a,b} \Leftrightarrow \begin{cases} \sigma^{(m)} = s_b^{2D} & \text{if } m \in P, \\ \sigma^{(m)} = s_a^{2D} & \text{if } m \in Q^c, \\ \sigma^{(m)} \in \mathcal{G}^{a,b,2D} & \text{if } m \in Q \setminus P, \end{cases}$$

where  $\mathcal{G}^{a,b,2D}$  is defined in Definition 5.10. Then, we define (cf. Notation 6.2)

$$\mathcal{G}_{P,Q}^{a,b} = \Upsilon(\tilde{\mathcal{G}}_{P,Q}^{a,b}).$$

Then, recall  $m_K$  from (6.20) and define, for  $i \in \llbracket 0, M - 1 \rrbracket$ ,

$$\mathcal{G}_i^{a,b} = \bigcup_{P,Q \in \mathfrak{S}_M: |P|=i \text{ and } P \prec Q} \mathcal{G}_{P,Q}^{a,b} \quad \text{and} \quad \mathcal{G}^{a,b} = \bigcup_{i=m_K-1}^{M-m_K} \mathcal{G}_i^{a,b}. \quad (7.1)$$

Notice that the crucial difference between (7.1) and (6.4) is the fact that the second union in (7.1) is taken only over  $i \in \llbracket m_K - 1, M - m_K \rrbracket$ . This is related to (6.21), and we give a more detailed reasoning in Section 7.2. A configuration belonging to  $\mathcal{G}^{a,b}$  for some  $a, b \in S$  is called a gateway configuration.

Finally, for a proper partition  $(A, B)$  of  $S$  (which is fixed throughout the current section), we write for  $i \in \llbracket 0, M - 1 \rrbracket$ ,

$$\mathcal{G}_i^{A,B} = \bigcup_{a \in A} \bigcup_{b \in B} \mathcal{G}_i^{a,b} \quad \text{and} \quad \mathcal{G}^{A,B} = \bigcup_{a \in A} \bigcup_{b \in B} \mathcal{G}^{a,b}. \quad (7.2)$$

**Notation 7.2.** For  $a, b \in S$  and  $P, Q \in \mathfrak{S}_M$  with  $P \prec Q$ ,  $Q \setminus P = \{m_0\}$ , and  $|P| \in \llbracket m_K - 1, M - m_K \rrbracket$ , we decompose

$$\tilde{\mathcal{G}}_{P,Q}^{a,b} = \tilde{\mathcal{G}}_{P,Q}^{a,b,[1]} \cup \tilde{\mathcal{G}}_{P,Q}^{a,b,[2]} \cup \tilde{\mathcal{G}}_{P,Q}^{a,b,[3]},$$

where (cf. (5.6), (5.7), and (5.10))

$$\begin{aligned} \tilde{\mathcal{G}}_{P,Q}^{a,b,[1]} &= \{\sigma \in \tilde{\mathcal{G}}_{P,Q}^{a,b} : \sigma^{(m_0)} \in \mathcal{B}^{a,b,2D} \setminus \mathcal{B}_\Gamma^{a,b,2D}\}, \\ \tilde{\mathcal{G}}_{P,Q}^{a,b,[2]} &= \{\sigma \in \tilde{\mathcal{G}}_{P,Q}^{a,b} : \sigma^{(m_0)} \in \mathcal{B}_\Gamma^{a,b,2D}\}, \\ \tilde{\mathcal{G}}_{P,Q}^{a,b,[3]} &= \{\sigma \in \tilde{\mathcal{G}}_{P,Q}^{a,b} : \sigma^{(m_0)} \in \mathcal{Z}^{a,b,2D} \cup \mathcal{Z}^{b,a,2D}\}. \end{aligned}$$

Then, write  $\mathcal{G}_{P,Q}^{a,b,[n]} = \Upsilon(\tilde{\mathcal{G}}_{P,Q}^{a,b,[n]})$ ,  $n \in \{1, 2, 3\}$ . A configuration  $\sigma \in \mathcal{G}^{A,B}$  is called a gateway configuration of type  $n$ ,  $n \in \{1, 2, 3\}$ , if  $\sigma \in \mathcal{G}_{P,Q}^{a,b,[n]}$  for some  $a \in A$ ,  $b \in B$  and  $P, Q \in \mathfrak{S}_M$  with  $P \prec Q$ .

The following proposition is direct from the definition of gateway configurations.

**Proposition 7.3.** For  $\sigma \in \mathcal{G}^{A,B}$ , we have  $H(\sigma) \in \{\Gamma^* - 2, \Gamma^*\}$ . Moreover, we have  $H(\sigma) = \Gamma^* - 2$  if and only if  $\sigma$  is a gateway configuration of type 1 and  $H(\sigma) = \Gamma^*$  if and only if  $\sigma$  is a gateway configuration of type 2 or 3.

*Proof.* Let  $\sigma \in \tilde{\mathcal{G}}_{P,Q}^{a,b}$  for some  $a \in A$ ,  $b \in B$  and  $P, Q \in \mathfrak{S}_M$  with  $P \prec Q$ ,  $Q \setminus P = \{m_0\}$ , and  $|P| \in \llbracket m_K - 1, M - m_K \rrbracket$ . Then, by Lemma 6.6, we can write

$$H(\sigma) = H^{2D}(\sigma^{(m_0)}) + 2KL$$

since  $H^{2D}(\sigma^{(m)}) = 0$  for all  $m \neq m_0$  and  $H^{1D}(\sigma^{(k,\ell)}) = 2$  for all  $k \in \mathbb{T}_K$  and  $\ell \in \mathbb{T}_L$ . Hence, by definition, we have

$$H(\sigma) = \begin{cases} 2KL + 2K = \Gamma^* - 2 & \text{if } \sigma \in \tilde{\mathcal{G}}_{P,Q}^{a,b,[1]}, \\ 2KL + 2K + 2 = \Gamma^* & \text{if } \sigma \in \tilde{\mathcal{G}}_{P,Q}^{a,b,[2]} \cup \tilde{\mathcal{G}}_{P,Q}^{a,b,[3]}. \end{cases}$$

Since the Hamiltonian is invariant under  $\Upsilon$ , the proof is completed.  $\square$

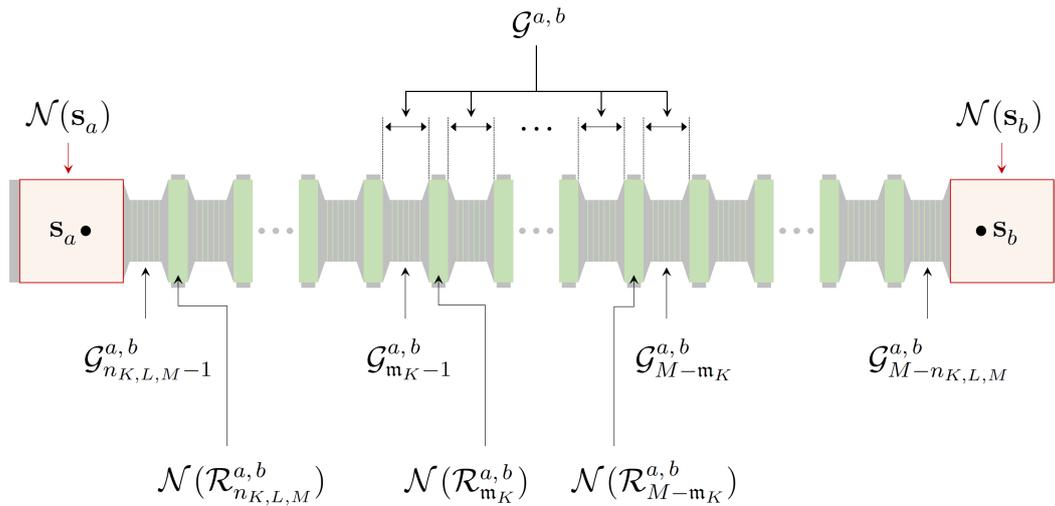


Figure 8: **Structure of gateway configurations between  $s_a$  and  $s_b$ .** The grey regions consist of configurations of energy  $\Gamma^*$ . The green boxes denote the sets of the form  $\mathcal{N}(\mathcal{R}_i^{a,b})$  for  $i \in \llbracket n_{K,L,M}, M - n_{K,L,M} \rrbracket$  (cf. (6.22)), while the green lines denote the gateway configurations of type 1 whose energy is  $\Gamma^* - 2$  (cf. Proposition 7.3). Later in Proposition 8.1, we shall show that  $m_K \geq n_{K,L,M}$ . The structure given in this figure (especially between  $\mathcal{G}_{m_K-1}^{a,b}$  and  $\mathcal{G}_{M-m_K}^{a,b}$ ) is confirmed in Lemma 7.5. We remark that the dead-ends are attached to  $\mathcal{N}(s_a)$ ,  $\mathcal{N}(s_b)$ , and  $\mathcal{N}(\mathcal{R}_i^{a,b})$ ,  $i \in \llbracket n_{K,L,M}, M - n_{K,L,M} \rrbracket$ . In particular, the configurations in  $\mathcal{G}_i^{a,b}$  with  $i < n_{K,L,M} - 1$  belong to the dead-ends attached to the set  $\mathcal{N}(s_a)$ .

### 7.2 Properties of gateway configurations

Next, we investigate several crucial properties of the gateway configurations which will be used frequently in the following discussions. The following notation will be useful in the remaining parts of the article.

**Notation 7.4.** For any integers  $u, v$  such that  $0 \leq u < v \leq M$ , we write

$$\mathcal{K}_{[u,v]}^{a,b} = \bigcup_{i=u}^v \mathcal{K}_i^{a,b} \quad \text{and} \quad \mathcal{K}_{[u,v]}^{A,B} = \bigcup_{i=u}^v \mathcal{K}_i^{A,B},$$

where  $\mathcal{K} \in \{\mathcal{C}, \mathcal{G}, \mathcal{R}\}$ . In particular, by (7.1) and (7.2), we can write

$$\mathcal{G}^{A,B} = \mathcal{G}_{[m_K-1, M-m_K]}^{A,B}. \tag{7.3}$$

In this section, we focus on the relation between gateway configurations and neighborhoods of regular configurations. We refer to Figure 8 for an illustration of the relations obtained in the current subsection.

The first one below states that we have to escape from a gateway configuration via a neighborhood of regular configurations, unless we touch a configuration with energy higher than  $\Gamma^*$ .

**Lemma 7.5.** For a proper partition  $(A, B)$  of  $S$ , the following statements hold.

1. For  $a \in A$ ,  $b \in B$ , and  $i \in \llbracket m_K - 1, M - m_K \rrbracket$ , we suppose that  $\sigma \in \mathcal{G}_i^{a,b}$  and  $\zeta \in \mathcal{X} \setminus \mathcal{G}_i^{a,b}$  satisfy  $\sigma \sim \zeta$  and  $H(\zeta) \leq \Gamma^*$ . Then, we have  $\zeta \in \mathcal{N}(\mathcal{R}_{[i, i+1]}^{a,b})$ , and moreover  $\sigma$  is a gateway configuration of type 3.

2. Suppose that  $\sigma \in \mathcal{G}^{A,B}$  and  $\zeta \in \mathcal{X} \setminus \mathcal{G}^{A,B}$  satisfy  $\sigma \sim \zeta$  and  $H(\zeta) \leq \Gamma^*$ . Then, we have  $\zeta \in \mathcal{N}(\mathcal{R}_{[\mathfrak{m}_K-1, M-\mathfrak{m}_K+1]}^{A,B})$ , and moreover  $\sigma$  is a gateway configuration of type 3.

*Proof.* We first suppose that  $\sigma \in \tilde{\mathcal{G}}_{P,Q}^{a,b}$  and  $\zeta \in \mathcal{X} \setminus \tilde{\mathcal{G}}_{P,Q}^{a,b}$  for some  $a \in A, b \in B$  and  $P, Q \in \mathfrak{S}_M$  with  $P \prec Q$  and  $|P| \in [\mathfrak{m}_K - 1, M - \mathfrak{m}_K]$ . We write  $Q \setminus P = \{m_0\}$ . Then, we claim that  $\zeta \in \mathcal{N}(\{\sigma_P^{a,b}, \sigma_Q^{a,b}\})$ , and  $\sigma$  is of type 3.

Let us first show that  $\sigma$  is a gateway configuration of type 3. If  $\sigma$  is of type 1, then we have  $H(\sigma) = \Gamma^* - 2, H^{2D}(\sigma^{(m_0)}) = 2K$ , and  $\sigma^{(m_0)} \in \mathcal{B}^{a,b,2D}$ . To update a spin in  $\sigma$  without increasing the energy by 3 or more, it can be readily observed that we have to update a spin of  $\sigma$  at the  $m_0$ -th floor to get  $\zeta$  with  $H^{2D}(\zeta^{(m_0)}) \leq 2K + 2$ . In such a situation, Lemma 5.11 asserts that  $\sigma^{(m_0)} \notin \mathcal{B}^{a,b,2D}$  and we get a contradiction. A similar argument can be applied if  $\sigma$  is of type 2, and hence we can conclude that  $\sigma$  is of type 3.

Now, since  $\sigma$  is of type 3, we have  $H(\sigma) = \Gamma^*, H^{2D}(\sigma^{(m_0)}) = 2K + 2$ , and  $\sigma^{(m_0)} \in \mathcal{Z}^{a,b,2D} \cup \mathcal{Z}^{b,a,2D}$  (cf. (5.9)). In order not to increase the energy by flipping a site of  $\sigma$ , it is clear that we have to flip a spin at the  $m_0$ -th floor (cf. Figure 7). This means that, by Lemma 5.11, we have  $\zeta^{(m_0)} \in \mathcal{N}^{2D}(\mathfrak{s}_a^{2D}) \cup \mathcal{N}^{2D}(\mathfrak{s}_b^{2D})$ . Now, we suppose first that  $\zeta^{(m_0)} \in \mathcal{N}^{2D}(\mathfrak{s}_a^{2D})$ . Then, there exists a 2D  $(2K + 1)$ -path  $(\omega_t)_{t=0}^T$  in  $\mathcal{X}^{2D} = S^{\Lambda^{2D}}$  such that  $\omega_0 = \mathfrak{s}_a^{2D}$  and  $\omega_T = \zeta^{(m_0)}$ . Define a 3D path  $(\tilde{\omega}_t)_{t=0}^T$  as

$$\tilde{\omega}_t^{(m)} = \begin{cases} \omega_t^{(m)} & \text{if } m = m_0, \\ \zeta^{(m)} = \sigma^{(m)} & \text{if } m \neq m_0. \end{cases}$$

Then,  $(\tilde{\omega}_t)_{t=0}^T$  is a  $(\Gamma^* - 1)$ -path connecting  $\sigma_P^{a,b}$  and  $\zeta$ , and thus we get  $\zeta \in \mathcal{N}(\sigma_P^{a,b})$ . Similarly, we can deduce that  $\zeta^{(m_0)} \in \mathcal{N}^{2D}(\mathfrak{s}_b^{2D})$  implies  $\zeta \in \mathcal{N}(\sigma_Q^{a,b})$ . This concludes the proof of the claim.

Now, we return to the lemma. For part (1), suppose that  $\sigma \in \mathcal{G}_{P,Q}^{a,b}$  for some  $a \in A, b \in B$  and  $P, Q \in \mathfrak{S}_M$  with  $|P| = i \in [\mathfrak{m}_K - 1, M - \mathfrak{m}_K]$  and  $P \prec Q$ . If  $\sigma \in \tilde{\mathcal{G}}_{P,Q}^{a,b}$ , then by the claim above, we get

$$\zeta \in \mathcal{N}(\{\sigma_P^{a,b}, \sigma_Q^{a,b}\}) \subseteq \mathcal{N}(\mathcal{R}_{[i, i+1]}^{a,b}),$$

and moreover  $\sigma$  is a gateway configuration of type 3. On the other hand, if  $\sigma \in \Theta(\tilde{\mathcal{G}}_{P,Q}^{a,b})$  for some permutation operator  $\Theta$  that appears in Notation 6.2, then by the same logic as above, we obtain that

$$\zeta \in \mathcal{N}(\{\Theta(\sigma_P^{a,b}), \Theta(\sigma_Q^{a,b})\}) \subseteq \mathcal{N}(\Theta(\tilde{\mathcal{R}}_{[i, i+1]}^{a,b})) \subseteq \mathcal{N}(\mathcal{R}_{[i, i+1]}^{a,b}),$$

and that  $\sigma$  is a gateway configuration of type 3. This completes the proof of part (1). Part (2) is direct from part (1).  $\square$

Next, we establish a relation between  $\mathcal{G}^{A,B}$  and  $\mathcal{N}(\mathcal{R}_{[0, M]}^{A,B})$  for proper partitions  $(A, B)$  of  $S$ . Recall Definition 4.1.

**Lemma 7.6.** *For a proper partition  $(A, B)$  of  $S$ , the two sets  $\mathcal{G}^{A,B}$  and  $\mathcal{N}(\mathcal{R}_{[0, M]}^{A,B})$  are disjoint and moreover, it holds that*

$$\widehat{\mathcal{N}}(\mathcal{G}^{A,B}; \mathcal{N}(\mathcal{R}_{[0, M]}^{A,B})) = \mathcal{G}^{A,B}. \tag{7.4}$$

*Proof.* We first claim that, for any  $a \in A, b \in B$ , and  $P, Q \in \mathfrak{S}_M$  with  $P \prec Q$  and  $|P| \in [\mathfrak{m}_K - 1, M - \mathfrak{m}_K]$ <sup>5</sup>,

$$\tilde{\mathcal{G}}_{P,Q}^{a,b} \cap \mathcal{N}(\mathcal{R}_{[0, M]}^{A,B}) = \emptyset. \tag{7.5}$$

<sup>5</sup>In fact, it holds even if  $|P| \in [0, M - 1]$ .

Suppose the contrary that we can take a configuration  $\sigma \in \tilde{\mathcal{G}}_{P,Q}^{a,b} \cap \mathcal{N}(\mathcal{R}_{[0,M]}^{A,B})$ . Then, since  $\sigma \in \tilde{\mathcal{G}}_{P,Q}^{a,b}$  and since  $H(\sigma) < \Gamma^*$  as  $\sigma \in \mathcal{N}(\mathcal{R}_{[0,M]}^{A,B})$ , the configuration  $\sigma$  must be a gateway configuration of type 1 by Proposition 7.3. Since  $\sigma \in \mathcal{N}(\mathcal{R}_{[0,M]}^{A,B})$ , there exists a  $(\Gamma^* - 1)$ -path connecting  $\sigma$  and  $\mathcal{R}_{[0,M]}^{A,B}$ . However, it is clear that (cf. Figure 7) any configuration  $\zeta$  such that  $\zeta \sim \sigma$  has energy at least  $\Gamma^*$ . This yields a contradiction. By the same argument, we can show that  $\Theta(\tilde{\mathcal{G}}_{P,Q}^{a,b})$  is also disjoint with  $\mathcal{N}(\mathcal{R}_{[0,M]}^{A,B})$  where  $\Theta$  is one of the permutation operators introduced in Notation 6.2, and hence it holds that  $\mathcal{G}_{P,Q}^{a,b}$  is disjoint with  $\mathcal{N}(\mathcal{R}_{[0,M]}^{A,B})$ . Hence, the two sets  $\mathcal{G}^{A,B}$  and  $\mathcal{N}(\mathcal{R}_{[0,M]}^{A,B})$  are disjoint.

Next, we turn to (7.4). Since  $\mathcal{G}^{A,B} \subseteq \widehat{\mathcal{N}}(\mathcal{G}^{A,B}; \mathcal{N}(\mathcal{R}_{[0,M]}^{A,B}))$  easily follows from (7.5), it suffices to show that

$$\widehat{\mathcal{N}}(\mathcal{G}^{A,B}; \mathcal{N}(\mathcal{R}_{[0,M]}^{A,B})) \subseteq \mathcal{G}^{A,B}.$$

Suppose the contrary that we can take  $\sigma \in \widehat{\mathcal{N}}(\mathcal{G}^{A,B}; \mathcal{N}(\mathcal{R}_{[0,M]}^{A,B}))$  which does not belong to  $\mathcal{G}^{A,B}$ . Let  $(\omega_t)_{t=0}^T$  be a  $\Gamma^*$ -path in  $\mathcal{X} \setminus \mathcal{N}(\mathcal{R}_{[0,M]}^{A,B})$  connecting  $\mathcal{G}^{A,B}$  and  $\sigma$ . Since we have assumed that  $\sigma \notin \mathcal{G}^{A,B}$ , we can take

$$t_0 = \min \{t : \omega_t \notin \mathcal{G}^{A,B}\}.$$

Since  $\omega_{t_0-1} \in \mathcal{G}^{A,B}$ ,  $\omega_{t_0} \notin \mathcal{G}^{A,B}$ , and  $\omega_{t_0-1} \sim \omega_{t_0}$ , by Lemma 7.5, we have  $\omega_{t_0-1} \in \mathcal{N}(\mathcal{R}_{[m_K-1, M-m_K+1]}^{A,B})$ . This contradicts the fact that  $(\omega_t)_{t=0}^T$  is a path in  $\mathcal{X} \setminus \mathcal{N}(\mathcal{R}_{[0,M]}^{A,B})$ .  $\square$

## 8 Energy barrier between ground states

The main objective of the current section is to analyze the energy barrier and optimal paths between ground states. In this section, we fix a proper partition  $(A, B)$  of  $S$ . The main result of the current section is the following result regarding the energy barrier between the ground states.

**Proposition 8.1.** *The following statements hold.*

1. For  $s, s' \in \mathcal{S}$ , we have that  $\Phi(s, s') \geq \Gamma^*$ .
2. Let  $(\omega_t)_{t=0}^T$  be a path in  $\mathcal{X} \setminus \mathcal{G}^{A,B}$  connecting  $\mathcal{S}(A)$  and  $\mathcal{S}(B)$ . Then, there exists  $t \in \llbracket 0, T \rrbracket$  such that  $H(\omega_t) \geq \Gamma^* + 1$ .

Part (1) of the previous proposition gives an opposite bound of Proposition 6.11 and hence completes the proof of the characterization of the energy barrier. *Moreover, in part (2), it is verified that any optimal path connecting  $\mathcal{S}(A)$  and  $\mathcal{S}(B)$  must visit a gateway configuration between them.* Before proceeding further, we officially conclude the proof of Theorem 2.4 by assuming Proposition 8.1.

*Proof of Theorem 2.4.* The conclusion of the theorem holds by Proposition 6.11 and part (1) of Proposition 8.1.  $\square$

We provide the proof of Proposition 8.1 in Sections 8.1 and 8.2. Then, in Section 8.3, we prove the large deviation-type results, namely Theorem 2.7, based on the analysis of energy landscape that we carried out so far.

### 8.1 Preliminary analysis on energy landscape

The purpose of this subsection is to provide a lemma (cf. Lemma 8.3 below) regarding the communication height between two far away configurations, which will be the crucial tool in the proof of Proposition 8.1.

Before proceeding to this result, we first introduce a lower bound on the Hamiltonian  $H$  which will be used frequently in the remaining computations of the current section.

For  $\sigma \in \mathcal{X}$  and  $a \in S$ , denote by  $\mathcal{D}_a(\sigma) \subseteq \mathbb{T}_K \times \mathbb{T}_L$  the collection of monochromatic pillars in  $\sigma$  of spin  $a$ :

$$\mathcal{D}_a(\sigma) = \{(k, \ell) \in \mathbb{T}_K \times \mathbb{T}_L : \sigma^{(k, \ell)}(m) = a \text{ for all } m \in \mathbb{T}_M\}.$$

Then, let  $\mathcal{D}(\sigma) = \bigcup_{a \in S} \mathcal{D}_a(\sigma)$  and write

$$d_a(\sigma) = |\mathcal{D}_a(\sigma)| \quad \text{and} \quad d(\sigma) = |\mathcal{D}(\sigma)| = \sum_{a \in S} d_a(\sigma). \tag{8.1}$$

Now, we derive a lower bound on  $H$ . Recall the 1D and 2D Hamiltonians from (6.9) and (5.1), respectively.

**Lemma 8.2.** *For each  $\sigma \in \mathcal{X}$ , it holds that*

$$H(\sigma) \geq 2KL - 2d(\sigma) + \sum_{m \in \mathbb{T}_M} H^{2D}(\sigma^{(m)}), \tag{8.2}$$

and the equality holds if and only if  $H^{1D}(\sigma^{(k, \ell)}) = 2$  for all  $(k, \ell) \in (\mathbb{T}_K \times \mathbb{T}_L) \setminus \mathcal{D}(\sigma)$ .

*Proof.* Since  $H^{1D}(\sigma^{(k, \ell)}) = 0$  if  $(k, \ell) \in \mathcal{D}(\sigma)$  and  $H^{1D}(\sigma^{(k, \ell)}) \geq 2$  otherwise, we have that

$$\sum_{(k, \ell) \in \mathbb{T}_K \times \mathbb{T}_L} H^{1D}(\sigma^{(k, \ell)}) \geq 2(KL - d(\sigma)). \tag{8.3}$$

Hence, we can deduce (8.2) from Lemma 6.6. The conclusion on the equality condition is immediate from the argument above.  $\square$

Now, we proceed to the main result of this subsection. For the simplicity of notation, we write, for  $a \in S$ ,

$$\mathcal{V}^a := \mathcal{N}^{2D}(\mathbf{s}_a^{2D}) \subseteq \mathcal{X}^{2D} \quad \text{and} \quad \Delta^{2D} := \mathcal{X}^{2D} \setminus \bigcup_{a=1}^q \mathcal{V}^a \tag{8.4}$$

so that we have the following natural decomposition of the set  $\mathcal{X}^{2D}$ :

$$\mathcal{X}^{2D} = \left( \bigcup_{a=1}^q \mathcal{V}^a \right) \cup \Delta^{2D}. \tag{8.5}$$

Note that the set  $\Delta^{2D}$  is non-empty by the definition of  $\mathcal{N}^{2D}$ . Recall  $\mathfrak{m}_K \in \mathbb{N}$  from (6.20). The following lemma, which is the main technical result in the analysis of the energy landscape, asserts that we have to overcome an energy barrier of  $\Gamma^*$  in order to change a 2D configuration at a certain floor from a neighborhood of a ground state to a neighborhood of another ground state.

**Lemma 8.3.** *Suppose that  $a, b \in S$ . Moreover, let  $U$  and  $V$  be two disjoint subsets of  $\mathbb{T}_M$  satisfying  $|U|, |V| \geq \mathfrak{m}_K$ , and let  $\sigma \in \mathcal{X}$  be a configuration satisfying*

$$\sigma^{(m)} \in \mathcal{V}^a \text{ for all } m \in U \quad \text{and} \quad \sigma^{(m)} \in \mathcal{V}^b \text{ for all } m \in V.$$

*Suppose that another configuration  $\zeta \in \mathcal{X}$  satisfies either  $\zeta^{(m)} \in \mathcal{V}^{a_1}$  for some  $m \in U$  and  $a_1 \neq a$  or  $\zeta^{(m)} \in \mathcal{V}^{b_1}$  for some  $m \in V$  and  $b_1 \neq b$ . Finally, we assume that  $\sigma$  satisfies*

$$d(\sigma) < 200. \tag{8.6}$$

*Then, both of the following statements hold.*

1. *It holds that  $\Phi(\sigma, \zeta) \geq \Gamma^*$ .*

2. For any path  $(\omega_t)_{t=0}^T$  in  $\mathcal{X} \setminus \mathcal{G}^{a,b}$  connecting  $\sigma$  and  $\zeta$ , there exists  $t \in \llbracket 0, T \rrbracket$  such that  $H(\omega_t) \geq \Gamma^* + 1$ .

*Proof.* We first consider part (1). Let  $(\omega_t)_{t=0}^T$  be a path connecting  $\sigma$  and  $\zeta$ . For convenience of notation, we define a collection  $(c_m)_{m \in U \cup V}$  such that

$$c_m = \begin{cases} a & \text{if } m \in U, \\ b & \text{if } m \in V. \end{cases} \tag{8.7}$$

Then, we define

$$T_0 = \min \{t : H^{2D}(\omega_t^{(m)}) \notin \mathcal{V}^{c_m} \text{ for some } m \in U \cup V\},$$

where the existence of  $t \in \llbracket 1, T - 1 \rrbracket$  such that  $H^{2D}(\omega_t^{(m)}) \notin \mathcal{V}^{c_m}$  for some  $m \in U \cup V$  is guaranteed by the conditions on  $\sigma$  and  $\zeta$ . Now, we find  $m_0 \in U \cup V$  such that

$$H^{2D}(\omega_{T_0}^{(m_0)}) \notin \mathcal{V}^{c_{m_0}}. \tag{8.8}$$

By the definitions of  $\mathcal{V}^a$  and  $T_0$ , we have that

$$H^{2D}(\omega_{T_0}^{(m_0)}) \geq \Gamma^{2D} = 2K + 2. \tag{8.9}$$

If  $H(\omega_{T_0}) \geq \Gamma^*$ , there is nothing to prove. Hence, let us assume from now on that

$$H(\omega_{T_0}) < \Gamma^*. \tag{8.10}$$

Then, by Lemma 8.2 with  $\sigma = \omega_{T_0}$  and by recalling the definition (8.1) of  $d(\sigma)$ , we have

$$2 \sum_{n \in S} d_n(\omega_{T_0}) + 2K + 2 > \sum_{m \in \mathbb{T}_M} H^{2D}(\omega_{T_0}^{(m)}). \tag{8.11}$$

Since we get a contradiction to (8.9) if  $\mathcal{D}_n(\omega_{T_0}) = \emptyset$  for all  $n \in S$ , there exists  $n_0 \in S$  such that  $\mathcal{D}_{n_0}(\omega_{T_0}) \neq \emptyset$ . Suppose first that  $n_0 \in S \setminus \{b\}$ . For this case, we claim that

$$H^{2D}(\omega_{T_0}^{(m)}) \geq 4 \text{ for all } m \in V. \tag{8.12}$$

Assume not, so that we have  $\omega_{T_0}^{(m)} = \mathbf{s}_{n_0}^{2D}$  for some  $m \in V$ . If  $m = m_0$ , this obviously cannot happen. On the other hand, if  $m \in V \setminus \{m_0\}$ , we have  $\omega_{T_0}^{(m)} \in \mathcal{V}^b$  by the definition of  $T_0$  and thus  $\omega_{T_0}^{(m)}$  cannot be  $\mathbf{s}_{n_0}^{2D}$  as  $b \neq n_0$ . Therefore, we verified (8.12). Similarly, if  $n_0 \in S \setminus \{a\}$ , we obtain

$$H^{2D}(\omega_{T_0}^{(m)}) \geq 4 \text{ for all } m \in U. \tag{8.13}$$

Since either (8.12) or (8.13) must happen, and since  $|U|, |V| \geq m_K$ , we get from (8.9) and (8.11) that

$$2 \sum_{n \in S} d_n(\omega_{T_0}) + 2K + 2 > (2K + 2) + 4(m_K - 1), \tag{8.14}$$

and hence

$$\sum_{n \in S} d_n(\omega_{T_0}) \geq 2m_K - 1. \tag{8.15}$$

Thus, we have either

$$\sum_{n \in S \setminus \{a\}} d_n(\omega_{T_0}) \geq m_K \quad \text{or} \quad \sum_{n \in S \setminus \{b\}} d_n(\omega_{T_0}) \geq m_K.$$

Then for  $K$  satisfying the condition in Theorem 2.4, we have  $m_K \geq 200$  and thus by the condition (8.6), we can take  $T_1 < T_0$  such that

$$T_1 = \min \left\{ t : \sum_{n \in S \setminus \{a\}} d_n(\omega_t) = h_K^2 \quad \text{or} \quad \sum_{n \in S \setminus \{b\}} d_n(\omega_t) = h_K^2 \right\} \quad (8.16)$$

where  $h_K = \lfloor \sqrt{m_K - 1} \rfloor$ . Since  $T_1 < T_0$ , by the definition of  $T_0$ , we have

$$\omega_{T_1}^{(m)} \in \mathcal{V}^a, \quad \forall m \in U \quad \text{and} \quad \omega_{T_1}^{(m)} \in \mathcal{V}^b, \quad \forall m \in V. \quad (8.17)$$

We first suppose that  $\sum_{n \in S \setminus \{a\}} d_n(\omega_{T_1}) = h_K^2$ . Since (cf. (5.13))

$$\|\omega_{T_1}^{(m)}\|_n \geq d_n(\omega_{T_1}) \quad \text{for all } m \in \mathbb{T}_M, n \in S,$$

we can assert from (8.17) and **(L2)**, **(L3)** of Proposition 5.15 that

$$H^{2D}(\omega_{T_1}^{(m)}) \geq 4 \left( \sum_{n \in S \setminus \{a\}} d_n(\omega_{T_1}) \right)^{1/2} = 4h_K \quad \text{for all } m \in U. \quad (8.18)$$

Therefore, by Lemma 8.2 with  $\sigma = \omega_{T_1}$ , the definition of  $T_1$ , and (8.18), we get

$$H(\omega_{T_1}) \geq 2KL - 4h_K^2 + 4h_K|U| \geq 2KL - 4h_K^2 + 4h_K m_K > 2KL + 2K + 2 = \Gamma^*,$$

where the last inequality holds for  $K \geq 32$ . Of course, we get the same conclusion for the case of  $\sum_{n \in S \setminus \{b\}} d_n(\omega_{T_1}) = h_K^2$  by an identical argument. Therefore, we can conclude that  $H(\omega_{T_1}) > \Gamma^*$ , and thus part (1) is verified.

Now, we turn to part (2). We now assume that, for some  $\sigma$  and  $\zeta$  satisfying the assumptions of the lemma, there exists a path  $(\omega_t)_{t=0}^T$  in  $\mathcal{X} \setminus \mathcal{G}^{a,b}$  connecting  $\sigma$  and  $\zeta$  with

$$H(\omega_t) \leq \Gamma^* \quad \text{for all } t \in \llbracket 0, T \rrbracket. \quad (8.19)$$

Without loss of generality, we can assume that the triple  $(\sigma, \zeta, (\omega_t)_{t=0}^T)$  that we selected has the smallest path length  $T$  among all such triples.

Recall  $T_0$  from the proof of the first part. If  $\mathcal{D}_n(\omega_{T_0}) \neq \emptyset$  for some  $n \in S$ , we can repeat the same argument with part (1) to deduce  $H(\omega_{T_1}) > \Gamma^*$ , where  $T_1$  is defined in (8.16). This contradicts (8.19).

Next, we consider the case when  $\mathcal{D}_n(\omega_{T_0}) = \emptyset$  for all  $n \in S$ . The contradiction for this case is more involved than that of the corresponding case of part (1). By Lemma 8.2, we have that

$$2K + 2 \geq \sum_{m \in \mathbb{T}_M} H^{2D}(\omega_{T_0}^{(m)}). \quad (8.20)$$

Recall  $m_0$  from (8.8). Since  $H^{2D}(\omega_{T_0}^{(m_0)}) = 2K + 2$  by (8.9), we not only have

$$H^{2D}(\omega_{T_0}^{(m)}) = 0 \quad \text{for all } m \in \mathbb{T}_M \setminus \{m_0\}, \quad (8.21)$$

but also the equality in (8.20) holds, i.e.,

$$\sum_{m \in \mathbb{T}_M} H^{2D}(\omega_{T_0}^{(m)}) = 2K + 2. \quad (8.22)$$

Hence, by the last part of Lemma 8.2, we must have

$$H^{1D}(\omega_{T_0}^{(k,\ell)}) = 2 \quad \text{for all } (k, \ell) \in \mathbb{T}_K \times \mathbb{T}_L. \quad (8.23)$$

From these observations, we can deduce the following facts:

- By (8.22), (8.23), and Lemma 8.2, we have  $H(\omega_{T_0}) = \Gamma^*$ .
- By (8.21) and (8.23), we have  $\omega_{T_0}^{(m)} \in \{\mathbf{s}_a^{2D}, \mathbf{s}_b^{2D}\}$  for all  $m \in \mathbb{T}_M \setminus \{m_0\}$ .

Moreover, the spins must be aligned so that (8.23) holds. Without loss of generality, we assume that  $m_0 \in U$ , since the case  $m_0 \in V$  can be handled in an identical manner. Starting from  $\omega_{T_0}$ , suppose that we flip a spin at  $m$ -th floor,  $m \neq m_0$ , without decreasing the 2D energy of the  $m_0$ -th floor. Then, since each non- $m_0$ -th floor is monochromatic and (8.23) holds, the 3D energy of  $\sigma$  increases by at least four and we obtain a contradiction to the fact that  $(\omega_t)_{t=0}^T$  is a  $\Gamma^*$ -path. Thus, we must decrease the 2D energy of the  $m_0$ -th floor before modifying the other floors. Define

$$T_2 = \min \{t > T_0 : H^{2D}(\omega_t^{(m_0)}) < 2K + 2\} .$$

Then, by Proposition 5.15, it suffices to consider the following two cases:

- **(Case 1:  $\omega_{T_2}^{(m_0)} \in \mathcal{V}^n$  for some  $n \in S$ )** Since  $\omega_{T_0}^{(m_0)} \in \mathcal{X}^{2D}$  is the first escape from the valley  $\mathcal{V}^a$ , it holds from the minimality of  $T_2$  that  $\omega_{T_2}^{(m_0)} \notin \mathcal{V}^n$  for  $n \in S \setminus \{a\}$  (the 2D path must visit a number of regular configurations first; see part (1) of Proposition 5.9). On the other hand, if  $\omega_{T_2}^{(m_0)} \in \mathcal{V}^a$ , then we obtain a contradiction from the minimality of the length of  $(\omega_t)_{t=0}^T$ , as we have a shorter path from  $\omega_{T_2}$  to  $\zeta$  where  $\omega_{T_2}$  clearly satisfies the conditions imposed to  $\sigma$ .
- **(Case 2:  $\omega_{T_2}^{(m_0)}$  is a 2D regular configuration)** Since we have assumed that  $m_0 \in U$ , we have  $\omega_{T_2}^{(m_0)} \in \mathcal{R}_2^{a,b'}$  for some  $b' \in S \setminus \{a\}$  (by the minimality of  $T_2$  and part (1) of Proposition 5.9). Now, we claim that  $b' = b$ . To this end, let us suppose that  $b' \neq b$ . Then as  $\omega_{T_2}^{(m)} \in \{\mathbf{s}_a^{2D}, \mathbf{s}_b^{2D}\}$  for  $m \neq m_0$ , we have  $H^{1D}(\omega_{T_2}^{(k,\ell)}) \geq 3$  for  $(k, \ell) \in \mathbb{T}_K \times \mathbb{T}_L$  satisfying  $\omega_{T_2}^{(m_0)}(k, \ell) = b'$ . Because there are exactly  $2K$  such  $(k, \ell)$ , by Lemma 8.2, we have

$$\begin{aligned} H(\omega_{T_2}) &= \sum_{(k,\ell) \in \mathbb{T}_K \times \mathbb{T}_L} H^{1D}(\omega_{T_2}^{(k,\ell)}) + \sum_{m \in \mathbb{T}_M} H^{2D}(\omega_{T_2}^{(m)}) \\ &\geq 3 \times 2K + 2 \times (KL - 2K) + 2K > \Gamma^* , \end{aligned}$$

where at the first inequality we used the fact that  $H^{2D}(\omega_{T_2}^{(m_0)}) = 2K$ . This contradicts the fact that  $(\omega_t)_{t=0}^T$  is a  $\Gamma^*$ -path. Therefore, we must have  $b' = b$ , which implies along with (8.23) that  $\omega_{T_2} \in \mathcal{G}^{a,b}$ . Hence, we get a contradiction as we assumed that  $(\omega_t)_{t=0}^T$  is a path in  $\mathcal{X} \setminus \mathcal{G}^{a,b}$ .

Since we get a contradiction for both cases, we completed the proof of part (2). □

**Remark 8.4.** We remark that (8.16) is exactly the place from which the lower bound 2829 of  $K$  in Theorem 2.4 originates.

The following is a direct consequence of the previous lemma which will be used later.

**Corollary 8.5.** *Suppose that  $P, Q \in \mathfrak{S}_M$  and  $|P| \in \llbracket \mathfrak{m}_K, M - \mathfrak{m}_K \rrbracket$ . Then for  $a, b \in S$ , we have  $\Phi(\sigma_P^{a,b}, \sigma_Q^{a,b}) = \Gamma^*$ . In particular, we have  $\Phi(\sigma_P^{a,b}, \mathbf{s}_a) = \Gamma^*$ .*

*Proof.* We can apply Lemma 8.3 with  $\sigma = \sigma_P^{a,b}$  and  $\zeta = \sigma_Q^{a,b}$  to get

$$\Phi(\sigma_P^{a,b}, \sigma_Q^{a,b}) \geq \Gamma^* . \tag{8.24}$$

On the other hand, by taking a canonical path connecting  $\mathbf{s}_a$  and  $\sigma_P^{a,b}$ , we get  $\Phi(\mathbf{s}_a, \sigma_P^{a,b}) \leq \Gamma^*$ . Similarly, we get  $\Phi(\mathbf{s}_a, \sigma_Q^{a,b}) \leq \Gamma^*$ . Hence, we obtain

$$\Phi(\sigma_P^{a,b}, \sigma_Q^{a,b}) \leq \max \{ \Phi(\mathbf{s}_a, \sigma_P^{a,b}), \Phi(\mathbf{s}_a, \sigma_Q^{a,b}) \} \leq \Gamma^* . \tag{8.25}$$

Combining (8.24) and (8.25) proves  $\Phi(\sigma_P^{a,b}, \sigma_Q^{a,b}) = \Gamma^*$ . By inserting  $Q = \emptyset$ , we get  $\Phi(\sigma_P^{a,b}, \mathbf{s}_a) = \Gamma^*$ . □

**8.2 Proof of Proposition 8.1**

Recall (5.13). Note that

$$\|\sigma\|_a = \sum_{m \in \mathbb{T}_M} \|\sigma^{(m)}\|_a. \tag{8.26}$$

We are now ready to prove Proposition 8.1. We first prove this proposition when  $q = 2$ . Then, the general case can be verified from this result via a projection-type argument.

*Proof of Proposition 8.1:*  $q = 2$ . Since  $q = 2$ , we only have two spins 1 and 2 and hence we let  $s = s_1$  and  $s' = s_2$ . We fix an arbitrary path  $(\omega_t)_{t=0}^T$  connecting  $s$  and  $s'$ , and take  $\sigma \in (\omega_t)_{t=0}^T$  such that

$$\|\sigma\|_1 = \lfloor KLM/2 \rfloor + 1. \tag{8.27}$$

Since there is nothing to prove if  $H(\sigma) \geq \Gamma^* + 1$ , we assume that

$$H(\sigma) \leq \Gamma^*. \tag{8.28}$$

Then, we claim that there exists  $t \in \llbracket 0, T \rrbracket$  such that  $H(\omega_t) = \Gamma^*$ . Moreover, we claim that if  $(\omega_t)_{t=0}^T$  is a path in  $\mathcal{X} \setminus \mathcal{G}^{1,2}$ , there exists  $t \in \llbracket 0, T \rrbracket$  such that  $H(\omega_t) = \Gamma^* + 1$ . It is clear that a verification of these claims immediately proves the case of  $q = 2$ .

We recall the decomposition (8.5) of  $\mathcal{X}^{2D}$  and write

$$\begin{aligned} P_n = P_n(\sigma) &= \{m \in \mathbb{T}_M : \sigma^{(m)} \in \mathcal{V}^n\} \quad ; \quad n \in \{1, 2\}, \\ R = R(\sigma) &= \{m \in \mathbb{T}_M : \sigma^{(m)} \in \Delta^{2D}\}, \end{aligned}$$

so that  $\mathbb{T}_M$  can be decomposed into  $\mathbb{T}_M = P_1 \cup P_2 \cup R$ . Write  $p_1 = |P_1|$ ,  $p_2 = |P_2|$ , and  $r = |R|$  so that the previous decomposition of  $\mathbb{T}_M$  implies

$$p_1 + p_2 + r = M. \tag{8.29}$$

We also write  $d_1 = d_1(\sigma)$ ,  $d_2 = d_2(\sigma)$ , and  $d = d(\sigma)$  so that  $d = d_1 + d_2$ . The following facts are crucially used:

- By Lemma 8.2 and (8.28), it holds that

$$d_1 + d_2 + K + 1 \geq \frac{1}{2} \sum_{m \in \mathbb{T}_M} H^{2D}(\sigma^{(m)}). \tag{8.30}$$

- We have

$$\begin{cases} H^{2D}(\sigma^{(m)}) \geq 4 \|\sigma^{(m)}\|_2^{1/2} \geq 4d_2^{1/2} & \text{if } m \in P_1, \\ H^{2D}(\sigma^{(m)}) \geq 4 \|\sigma^{(m)}\|_1^{1/2} \geq 4d_1^{1/2} & \text{if } m \in P_2, \\ H^{2D}(\sigma^{(m)}) \geq 2K & \text{if } m \in R, \end{cases} \tag{8.31}$$

where the first two bounds follow from **(L2)** and **(L3)** of Proposition 5.15, while the last one follows from **(L1)** of Proposition 5.15.

- By inserting (8.31) to (8.30), we get

$$d_1 + d_2 + K + 1 \geq 2p_1d_2^{1/2} + 2p_2d_1^{1/2} + Kr. \tag{8.32}$$

We consider four cases separately based on the conditions on  $p_1$ ,  $p_2$ , and  $r$ . Recall that we assumed  $K \geq 2829$ ; several arguments below require  $K$  to be large enough, and they indeed hold for  $K$  in this range.

**(Case 1:  $p_1, p_2 \geq 1$ )** Since both  $P_1$  and  $P_2$  are non-empty, the first two bounds in (8.31) activate and thus

$$d_1, d_2 \leq \frac{(2K + 1)^2}{16}. \tag{8.33}$$

We note that, since the function  $f(x) = x - 2ax^{1/2}$  is convex on  $[0, \infty)$  for  $a > 0$ , by (8.33) we have

$$\begin{aligned} d_1 - 2p_2d_1^{1/2} &\leq \max \left\{ 0, \frac{(2K+1)^2}{16} - \frac{2K+1}{2}p_2 \right\}, \\ d_2 - 2p_1d_2^{1/2} &\leq \max \left\{ 0, \frac{(2K+1)^2}{16} - \frac{2K+1}{2}p_1 \right\}. \end{aligned} \tag{8.34}$$

Inserting (8.34) to (8.32), we get

$$Kr \leq K + 1 + \sum_{i=1}^2 \max \left\{ 0, \frac{(2K+1)^2}{16} - \frac{2K+1}{2}p_i \right\}. \tag{8.35}$$

We now consider three sub-cases:

- $p_1, p_2 \leq (2K+1)/8$ : For this case, we can rewrite (8.35) as

$$Kr \leq K + 1 + \frac{(2K+1)^2}{8} - \frac{2K+1}{2}(p_1 + p_2) < K + 1 + \frac{(2K+1)^2}{8} - K(p_1 + p_2).$$

Inserting (8.29) yields a contradiction since  $K \leq M$ .

- $p_1 \leq (2K+1)/8 < p_2$  or  $p_2 \leq (2K+1)/8 < p_1$ : By symmetry, it suffices to consider the former case, for which we can rewrite (8.35) as

$$Kr \leq K + 1 + \frac{(2K+1)^2}{16} - \frac{2K+1}{2}p_1 < 2K + \frac{(2K+1)^2}{16} - Kp_1.$$

Thus, we get

$$p_1 + r \leq 2 + \frac{(2K+1)^2}{16K},$$

and thus by the second bound in (8.31),

$$\|\sigma\|_2 \geq p_2 \left( KL - \frac{(2K+1)^2}{16} \right) \geq \left( M - 2 - \frac{(2K+1)^2}{16K} \right) \left( KL - \frac{(2K+1)^2}{16} \right).$$

We get a contradiction to (8.27) since the right-hand side is greater than  $\lfloor KLM/2 \rfloor + 1$ .

- $p_1, p_2 > (2K+1)/8$ : By (8.35), we can notice that  $r = 0$  or  $1$ . By (8.27), the first bound in (8.31), and (8.26), we get

$$\left\lfloor \frac{KLM}{2} \right\rfloor + 1 = \|\sigma\|_1 \geq \sum_{m \in P_1} \|\sigma^{(m)}\|_1 \geq p_1 \left( KL - \frac{(2K+1)^2}{16} \right),$$

and thus,

$$p_2 \geq M - 1 - p_1 \geq M - 1 - \frac{\lfloor KLM/2 \rfloor + 1}{KL - (2K+1)^2/16} \geq \frac{2K+1}{7}. \tag{8.36}$$

Similarly, we get

$$p_1 \geq \frac{2K+1}{7}. \tag{8.37}$$

Now, by (8.33), (8.36) and (8.37), it holds that

$$p_2 \geq \frac{4}{7}d_1^{1/2} \quad \text{and} \quad p_1 \geq \frac{4}{7}d_2^{1/2}.$$

Inserting this along with (8.36) and (8.37) to the right-hand side of (8.32), we get

$$2p_1d_2^{1/2} + 2p_2d_1^{1/2} + Kr \geq \left( d_2 + \frac{1}{4}p_1d_2^{1/2} \right) + \left( d_1 + \frac{1}{4}p_2d_1^{1/2} \right) \geq d + \frac{2K+1}{28}d^{1/2},$$

where the last inequality follows from the inequality  $x^{1/2} + y^{1/2} \geq (x + y)^{1/2}$ . Applying this to (8.32), we conclude that

$$d \leq \left( \frac{28(K+1)}{2K+1} \right)^2 < 200 .$$

This proves the condition (8.6) for  $\sigma$ . Moreover, since  $p_1, p_2 > (2K+1)/8 \geq m_K$ , we can now apply part (1) of Lemma 8.3 to deduce  $\Phi(\sigma, s') \geq \Gamma^*$ , and this proves the first part of the claim. Moreover, if  $(\omega_t)_{t=0}^T$  is a path in  $\mathcal{X} \setminus \mathcal{G}^{1,2}$ , then the sub-path from  $\sigma$  to  $\omega_T = s'$  is also in  $\mathcal{X} \setminus \mathcal{G}^{1,2}$ , and thus part (2) of Lemma 8.3 verifies the second assertion of the claim as well.

**(Case 2:**  $p_1 \geq 1, p_2 = 0, r \geq 1$  or  $p_1 \geq 1, p_2 = 0, r \geq 1$ ) By symmetry, it suffices to consider the former case. Similarly as in **(Case 1)**, we can apply the first bound in (8.31) to deduce

$$d_2 \leq \frac{(2K+1)^2}{16} . \tag{8.38}$$

Again by the first bound in (8.31), we have

$$\|\sigma^{(m)}\|_1 \geq KL - \frac{(2K+1)^2}{16} \quad \text{for all } m \in P_1 ,$$

and thus we get

$$\sum_{m \in R} \|\sigma^{(m)}\|_1 = \|\sigma\|_1 - \sum_{m \in P_1} \|\sigma^{(m)}\|_1 \leq \frac{KLM}{2} + 1 - p_1 \left( KL - \frac{(2K+1)^2}{16} \right) .$$

Therefore, there exists  $m_0 \in R$  such that

$$\begin{aligned} \|\sigma^{(m_0)}\|_1 &\leq \frac{1}{r} \left[ \frac{KLM}{2} + 1 - p_1 \left( KL - \frac{(2K+1)^2}{16} \right) \right] \\ &= KL - \frac{(2K+1)^2}{16} + \frac{1}{r} \left[ -\frac{KLM}{2} + \frac{(2K+1)^2 M}{16} + 1 \right] \\ &\leq KL - \frac{(2K+1)^2}{16} - \frac{1}{r} \left[ \frac{KLM}{4} - \frac{K^2 M}{20} \right] , \end{aligned}$$

where at the second line we used  $p_1 = M - r$ . Thus, we have

$$d_1 \leq \|\sigma^{(m_0)}\|_1 \leq KL - \frac{(2K+1)^2}{16} - \frac{1}{r} \left[ \frac{KLM}{4} - \frac{K^2 M}{20} \right] . \tag{8.39}$$

Inserting this to (8.32), we get

$$2p_1 d_2^{1/2} + Kr \leq d_2 + \left[ KL - \frac{(2K+1)^2}{16} - \frac{1}{r} \left( \frac{KLM}{4} - \frac{K^2 M}{20} \right) \right] + K + 1 .$$

Reorganizing and applying a similar estimate as in (8.34), we get

$$\begin{aligned} Kr + \frac{1}{r} \left[ \frac{KLM}{4} - \frac{K^2 M}{20} \right] &\leq KL - \frac{(2K+1)^2}{16} + K + 1 + (d_2 - 2p_1 d_2^{1/2}) \\ &\leq KL - \frac{(2K+1)^2}{16} + K + 1 + \max \left\{ 0, \frac{(2K+1)^2}{16} - \frac{2K+1}{2} p_1 \right\} \\ &= KL + K + 1 - \min \left\{ \frac{(2K+1)^2}{16}, \frac{2K+1}{2} p_1 \right\} . \end{aligned} \tag{8.40}$$

Now, we analyze two sub-cases separately.

- $p_1 \leq (2K + 1)/8$ : Then, we can rewrite (8.40) as

$$Kr + \frac{1}{r} \left[ \frac{KLM}{4} - \frac{K^2M}{20} \right] \leq KL + K + 1 - \frac{2K + 1}{2} p_1 \leq KL + K.$$

Multiplying  $r/K$  in both sides, we reorganize the previous inequality as

$$\left( r - \frac{L + 1}{2} \right)^2 \leq \frac{(L + 1)^2}{4} - \frac{LM}{4} + \frac{KM}{20} \leq \frac{L^2 + 10L + 5}{20}. \tag{8.41}$$

Since  $p_1 \leq (2K + 1)/8$ , we have

$$r \geq M - \frac{2K + 1}{8} \geq \frac{3}{4}L - 1. \tag{8.42}$$

Inserting (8.42) to (8.41) yields a contradiction for  $L \geq K \geq 2829$ .

- $p_1 > (2K + 1)/8$ : For this case, (8.40) becomes

$$Kr + \frac{1}{r} \left[ \frac{KLM}{4} - \frac{K^2M}{20} \right] \leq KL + K + 1 - \frac{(2K + 1)^2}{16} \leq KL - \frac{K^2}{4} + K.$$

Multiplying both sides by  $r/K$  and reorganizing, we get

$$\left( r - \frac{L}{2} + \frac{K}{8} - \frac{1}{2} \right)^2 \leq \frac{1}{64} (4L - K + 4)^2 - \frac{LM}{4} + \frac{KM}{20}.$$

Since the right-hand side is negative for  $K \geq 9$ , we get a contradiction.

**(Case 3:**  $p_1 \geq 1, p_2 = 0, r = 0$  or  $p_1 = 0, p_2 \geq 1, r = 0$ ) As before, we only consider the former case. In this case, indeed  $P_1 = \mathbb{T}_M$ . Thus, by the first bound in (8.31) we have

$$\|\sigma\|_1 = \sum_{m \in \mathbb{T}_M} \|\sigma^{(m)}\|_1 \geq M \left( KL - \frac{(2K + 1)^2}{16} \right) > \frac{KLM}{2} + 2.$$

This contradicts (8.27).

**(Case 4:**  $p_1 = p_2 = 0$ ) For this case, we have  $\sigma^{(m)} \in \Delta^{2D}$  for all  $m \in \mathbb{T}_M$ . Hence,  $H^{2D}(\sigma^{(m)}) \geq 2K$  for all  $m \in \mathbb{T}_M$  by **(L1)** of Proposition 5.15, and thus by (8.30) we get

$$d_1 + d_2 \geq K(M - 1) - 1. \tag{8.43}$$

Since  $d_1 + d_2 = d \leq KL$ , we get  $M = L + 1$  or  $L$ .

If  $M = L + 1$ , we must have  $d_1 + d_2 = KL$  or  $KL - 1$ . If this is  $KL$ , then all floors should have the same configuration, which is impossible since  $\|\sigma\|_1 = \lfloor KLM/2 \rfloor + 1$  cannot be a multiple of  $M$ . If this is  $KL - 1$ , then the equality in (8.43) must hold and thus we have  $H^{2D}(\sigma^{(m)}) = 2K$  for all  $m \in \mathbb{T}_M$ . Hence, by **(L1)** of Proposition 5.15,  $\|\sigma^{(m)}\|_1, m \in \mathbb{T}_M$ , is a multiple of  $K$ , and thus  $\|\sigma\|_1 = \sum_{m \in \mathbb{T}_M} \|\sigma^{(m)}\|_1$  is also a multiple of  $K$ . This is impossible since  $\|\sigma\|_1 = \lfloor KLM/2 \rfloor + 1$  is not a multiple of  $K$ .

It remains to consider the case of  $M = L$ . For this case, (8.43) becomes

$$d_1 + d_2 \geq K(L - 1) - 1. \tag{8.44}$$

Define

$$\mathcal{E}(\sigma) = (\mathbb{T}_K \times \mathbb{T}_L) \setminus (\mathcal{D}_1(\sigma) \cup \mathcal{D}_2(\sigma)) \tag{8.45}$$

so that we have  $|\mathcal{E}(\sigma)| \leq K + 1$  by (8.44). We now have three sub-cases. We note that  $H^{2D}(\sigma^{(m)})$  is an even integer for each  $m \in \mathbb{T}_M$ , as  $q = 2$ .

- First, we assume that  $H^{2D}(\sigma^{(m)}) = 2K$  for all  $m \in \mathbb{T}_M$ . Then, as in the previous discussion on the case of  $M = L + 1$ , we get a contradiction since  $\|\sigma\|_1$  must be a multiple of  $K$  for this case.
- Next, we assume that  $H^{2D}(\sigma^{(m)}) \geq 2K + 2$  for all  $m \in \mathbb{T}_M$ . Then, by (8.30),

$$d = d_1 + d_2 \geq (K + 1)(L - 1) . \tag{8.46}$$

If  $d = KL$ , then as in the case of  $M = L + 1$ , we get a contradiction since  $\|\sigma\|_1 = \lfloor KLM/2 \rfloor + 1$  cannot be a multiple of  $M$ . Hence, we have  $d \leq KL - 1$ , and combining this with (8.46) implies that we must have  $K = L$  (and thus  $K = L = M$ ), and moreover

$$d = KL - 1 \quad \text{and} \quad H^{2D}(\sigma^{(m)}) = 2K + 2 \quad \text{for all } m \in \mathbb{T}_M .$$

Hence, we have  $|\mathcal{E}(\sigma)| = KL - d = 1$ . Write  $\mathcal{E}(\sigma) = \{(k_0, \ell_0)\}$ . By Lemma 5.14, we can deduce that the configuration  $\sigma^{(m)}$  has at least  $L - 1 \geq 3$  monochromatic bridges, and thus we have at least one monochromatic bridge of the form  $\mathbb{T}_K \times \{\ell\}$  or  $\{k\} \times \mathbb{T}_L$  that does not touch  $\mathcal{E}(\sigma)$ , so that it is a subset of either  $\mathcal{D}_1(\sigma)$  or  $\mathcal{D}_2(\sigma)$ . Suppose first that this bridge is  $\mathbb{T}_K \times \{\ell\}$  for some  $\ell \in \mathbb{T}_L$ . Then, the slab  $\mathbb{T}_K \times \{\ell\} \times \mathbb{T}_M$  is monochromatic. Therefore, by replacing the role of the second and third coordinates, which is possible since  $K = L = M$ , the proof is reduced to one of **(Case 1)**, **(Case 2)**, and **(Case 3)** as there is a monochromatic floor so that either  $p_1$  or  $p_2$  is positive. This completes the proof. Similarly, if the monochromatic bridge is  $\{k\} \times \mathbb{T}_L$ , then we replace the role of the first and third coordinates to complete the proof.

- Now, we lastly assume that  $H^{2D}(\sigma^{(i_0)}) = 2K$  for some  $i_0 \in \mathbb{T}_M$  and  $H^{2D}(\sigma^{(j_0)}) \geq 2K + 2$  for some  $j_0 \in \mathbb{T}_M$ . By (8.30), we get

$$d \geq KL - K , \tag{8.47}$$

and hence, we have  $|\mathcal{E}(\sigma)| \leq K$  (cf. (8.45)). Now, we consider two sub-sub-cases separately.

- $|\mathcal{E}(\sigma)| \leq K - 1$ : First, suppose that  $K < L$ . By **(L1)** of Proposition 5.15, we have  $\sigma^{(i_0)} \in \mathcal{R}_v^{1,2}$  for some  $v \in \llbracket 2, L - 2 \rrbracket$ . Since  $|\mathcal{E}(\sigma)| \leq K - 1 \leq L - 1$ , there exists  $\ell_1 \in \mathbb{T}_L$  such that

$$(\mathbb{T}_K \times \{\ell_1\}) \cap \mathcal{E}(\sigma) = \emptyset .$$

We further have  $\mathbb{T}_K \times \{\ell_1\} \subseteq \mathcal{D}_1(\sigma)$  or  $\mathbb{T}_K \times \{\ell_1\} \subseteq \mathcal{D}_2(\sigma)$  since  $\sigma^{(i_0)} \in \mathcal{R}_v^{1,2}$ . This implies that all sites in the slab  $\mathbb{T}_K \times \{\ell_1\} \times \mathbb{T}_M$  have the same spin  $n$  under  $\sigma$ . Since  $L = M$ , we can replace the role of the second and third coordinates to reduce the proof to one of **(Case 1)**, **(Case 2)** and **(Case 3)**. This completes the proof. Next, if  $K = L$ , then since there further exists  $k_1 \in \mathbb{T}_K$  such that  $(\{k_1\} \times \mathbb{T}_L) \cap \mathcal{E}(\sigma) = \emptyset$ , we can use the same argument as above to handle this case as well.

- $|\mathcal{E}(\sigma)| = K$ : The equality in (8.47) must hold, and thus we get  $H^{2D}(\sigma^{(j_0)}) = 2K + 2$  and  $H^{2D}(\sigma^{(m)}) = 2K$  for all  $m \in \mathbb{T}_M \setminus \{j_0\}$ . We first suppose that  $K < L$ . By **(L1)** of Proposition 5.15, we get  $\sigma^{(m)} = \xi_{\ell_m, v_m}^{1,2}$  for some  $\ell_m \in \mathbb{T}_L$  and  $v_m \in \llbracket 2, L - 2 \rrbracket$  for all  $m \in \mathbb{T}_M \setminus \{j_0\}$ . Then, since  $L$  is strictly bigger than  $K = |\mathcal{E}(\sigma)|$ , we can always find a row in  $\mathbb{T}_K \times \mathbb{T}_L$  which is either a subset of  $\mathcal{D}_1(\sigma)$  or  $\mathcal{D}_2(\sigma)$ . Thus, by changing the role of the second and third coordinates, which is possible since  $L = M$ , we find a monochromatic floor and the proof is reduced to one of **(Case 1)**, **(Case 2)** and **(Case 3)**. Next,

we handle the case  $K = L$ , so that for all  $m \in \mathbb{T}_M \setminus \{j_0\}$ ,  $\sigma^{(m)} = \xi_{\ell_m, v_m}^{1,2}$  or  $\Theta(\xi_{\ell_m, v_m}^{1,2})$  (cf. Definition 5.4) for some  $\ell_m \in \mathbb{T}_L$  and  $v_m \in \llbracket 2, L - 2 \rrbracket$ . First of all, assume that all of them are of the same direction. Without loss of generality, assume that  $\sigma^{(m)} = \xi_{\ell_m, v_m}^{1,2}$  for all  $m \in \mathbb{T}_M \setminus \{j_0\}$ . If  $\sigma^{(m_1)} \neq \sigma^{(m_2)}$  for some  $m_1, m_2 \in \mathbb{T}_M \setminus \{j_0\}$ , then  $\mathcal{E}(\sigma)$  must be exactly the line where they differ and hence we can write  $\mathcal{E}(\sigma) = \mathbb{T}_K \times \{\ell_0\}$  for some  $\ell_0 \in \mathbb{T}_L$ . Then, by taking any  $\ell \in \mathbb{T}_L \setminus \{\ell_0\}$ , we notice that  $\mathbb{T}_K \times \{\ell\}$  is not only monochromatic in  $\sigma^{(m)}$  with  $m \in \mathbb{T}_M \setminus \{j_0\}$ , but also a subset of either  $\mathcal{D}_1(\sigma)$  or  $\mathcal{D}_2(\sigma)$ ; hence,  $\mathbb{T}_K \times \{\ell\} \times \mathbb{T}_M$  is a monochromatic slab. By replacing the role of the second and third coordinates, which is possible since  $L = M$ , we find a monochromatic floor and the proof is reduced to one of **(Case 1)**, **(Case 2)** and **(Case 3)**. On the contrary, suppose that  $\sigma^{(m_1)} = \sigma^{(m_2)}$  for all  $m_1, m_2 \in \mathbb{T}_M \setminus \{j_0\}$ . If there exists a row or column which is disjoint with  $\mathcal{E}(\sigma)$ , then we can argue as above. If not, then we can easily deduce that for the  $j_0$ -th floor,

$$H^{2D}(\sigma^{(j_0)}) \geq 4(K - 4) > 2K + 2,$$

which contradicts the assumption that  $H^{2D}(\sigma^{(j_0)}) = 2K + 2$ . Finally, we consider the case when  $\sigma^{(m)} = \xi_{\ell, v}^{1,2}$  and  $\sigma^{(m')} = \Theta(\xi_{\ell', v'}^{1,2})$  for some  $m, m' \in \mathbb{T}_M \setminus \{j_0\}$  simultaneously. In this case, we have

$$d_1 \leq (K - v)(K - v') \quad \text{and} \quad d_2 \leq vv'.$$

Thus, we get a contradiction since

$$\begin{aligned} |\mathcal{E}(\sigma)| &\geq K^2 - (K - v)(K - v') - vv' \\ &= \frac{1}{2}K^2 - \frac{1}{2}(K - 2v)(K - 2v') \geq \frac{1}{2}K^2 - \frac{1}{2}(K - 4)^2 = 4K - 8 > K, \end{aligned}$$

where the second inequality holds since  $v, v' \in \llbracket 2, K - 2 \rrbracket$ . □

Now, we consider the general case of Proposition 8.1.

*Proof of Proposition 8.1: general case.* We fix a proper partition  $(A, B)$  of  $S$  and then fix  $a \in A$  and  $b \in B$ . Let  $(\omega_t)_{t=0}^T$  be a path connecting  $s_a$  and  $s_b$ . For each  $\sigma \in \mathcal{X}$ , we denote by  $\tilde{\sigma}$  the configuration obtained from  $\sigma$  by changing all spins in  $A$  to 1 and spins in  $B$  to 2. Thus,  $\tilde{\sigma}$  becomes an Ising configuration, i.e. a spin configuration for  $q = 2$ . Note that

$$\begin{aligned} H(\tilde{\sigma}) &= \sum_{\{x, y\} \subseteq \Lambda: x \sim y} \mathbf{1}\{\tilde{\sigma}(x) \neq \tilde{\sigma}(y)\} \\ &= \sum_{\{x, y\} \subseteq \Lambda: x \sim y} \mathbf{1}\{\sigma(x) \in A, \sigma(y) \in B \text{ or } \sigma(x) \in B, \sigma(y) \in A\} \\ &\leq \sum_{\{x, y\} \subseteq \Lambda: x \sim y} \mathbf{1}\{\sigma(x) \neq \sigma(y)\} = H(\sigma). \end{aligned} \tag{8.48}$$

Now, we consider the induced pseudo-path  $(\tilde{\omega}_t)_{t=0}^T$  of  $(\omega_t)_{t=0}^T$  (cf. Notation 6.12). Thus, by the proof above for  $q = 2$ , there exists  $t_1 \in \llbracket 0, T \rrbracket$  such that  $H(\tilde{\omega}_{t_1}) \geq \Gamma^*$ . Thus, we get from (8.48) that

$$\Gamma^* \leq H(\tilde{\omega}_{t_1}) \leq H(\omega_{t_1}),$$

and we complete the proof for part (1).

For part (2), suppose that  $(\omega_t)_{t=0}^T$  is a path such that

$$H(\omega_t) \leq \Gamma^* \quad \text{for all } t \in \llbracket 0, T \rrbracket. \tag{8.49}$$

Then, by (8.48), we have  $H(\tilde{\omega}_t) \leq \Gamma^*$  for all  $t \in \llbracket 0, T \rrbracket$ . Thus, by the proof above for  $q = 2$ , there exists  $s \in \llbracket 0, T \rrbracket$  such that  $\tilde{\omega}_s \in \mathcal{G}^{1,2}$ . We now claim that  $\omega_s \in \mathcal{G}^{A,B}$ . It is immediate that this claim finishes the proof.

To prove this claim, we write

$$\mathcal{U}_n = \{x \in \Lambda : \tilde{\omega}_s(x) = n\} \quad ; \quad n = 1, 2 .$$

Then, we have

$$\omega_s(x) \in A \text{ for } x \in \mathcal{U}_1 \quad \text{and} \quad \omega_s(x) \in B \text{ for } x \in \mathcal{U}_2 . \tag{8.50}$$

Now, we assume that

$$\omega_s(x) \neq \omega_s(y) \text{ for some } x, y \in \mathcal{U}_1 \text{ or } x, y \in \mathcal{U}_2 \text{ with } x \sim y . \tag{8.51}$$

We now express the energy  $H(\omega_s)$  as

$$H(\omega_s) = \left[ \sum_{\{x,y\} \subseteq \mathcal{U}_1 \text{ or } \{x,y\} \subseteq \mathcal{U}_2} + \sum_{x \in \mathcal{U}_1, y \in \mathcal{U}_2} \right] \mathbf{1}\{\omega_s(x) \neq \omega_s(y)\} , \tag{8.52}$$

where the summation is carried over  $x, y$  satisfying  $x \sim y$ . Note that the second summation is equal to  $H(\tilde{\omega}_s)$  by (8.50). On the other hand, we can readily deduce from Figure 7 that the first summation of (8.52) is at least 4 if  $\tilde{\omega}_s$  is a gateway configuration of type 1, and at least 2 if  $\tilde{\omega}_s$  is a gateway configuration of type 2 or 3 (cf. Notation 7.2). Thus, by Proposition 7.3, we can conclude that the right-hand side of (8.52) is at least  $\Gamma^* + 2$ ; i.e., we get  $H(\omega_s) \geq \Gamma^* + 2$ . This contradicts (8.49) and hence, we cannot have (8.51). This finally implies that there exist  $a_0 \in A$  and  $b_0 \in B$  such that

$$\omega_s(x) = \begin{cases} a_0 & \text{if } x \in \mathcal{U}_1 , \\ b_0 & \text{if } x \in \mathcal{U}_2 , \end{cases}$$

and thus we have  $\omega_s \in \mathcal{G}^{a_0, b_0} \subseteq \mathcal{G}^{A,B}$  as claimed. □

### 8.3 Proof of Theorem 2.7

Theorem 2.7 is now a consequence of our analysis on the energy landscape and the general theory developed in [42, 43].

*Proof of Theorem 2.7.* We have two results on the energy barrier; Theorem 2.4 and Proposition 6.13. The theory developed in [43] implies that these two are sufficient to conclude Theorem 2.7<sup>6</sup>. This implication has been rigorously verified in [42] for the case of  $d = 2$ , and this argument extends to the case of  $d = 3$  without a modification. Hence, we do not repeat the argument here, and refer the readers to [42, Section 3] for a detailed proof. □

## 9 Typical configurations and optimal paths

In the previous sections, we proved large deviation-type results regarding the metastable behavior by analyzing the energy barrier in terms of canonical and gateway configurations. In order to get precise quantitative results such as Theorems 2.10 and 2.12 or to get a characterization of optimal paths, we need a more refined analysis of the energy landscape based on the typical configurations which will be introduced and analyzed in the current section.

We fix a proper partition  $(A, B)$  of  $S$  throughout the section.

<sup>6</sup>We remark that the second convergence of (2.11) is not a consequence of an analysis of the energy barrier, but of the first convergence of (2.11) and the symmetry of the model. This argument is also given in [42, Section 3] for  $d = 2$ , and an identical one works for  $d = 3$ .

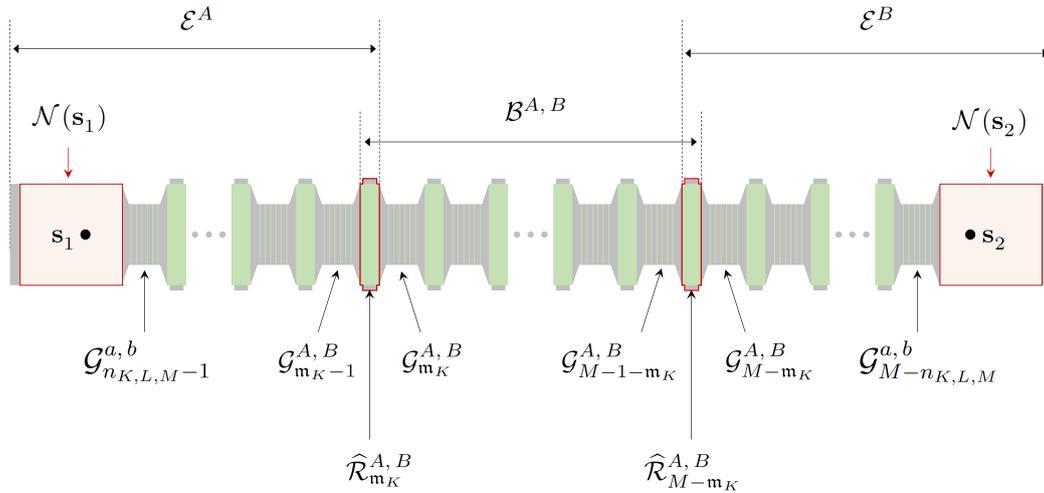


Figure 9: **3D typical configurations for the Ising model.** Suppose that  $q = 2$ ,  $A = \{1\}$ , and  $B = \{2\}$  (one can compare this figure with Figure 8). The given figure provides an illustration of the complete structure of the set  $\widehat{\mathcal{N}}(S)$ . This characterization is verified in Proposition 9.6. We take bulk ones only from  $m_K$  to  $M - m_K$  instead of  $n_{K,L,M}$  (cf. (6.22)) to  $M - n_{K,L,M}$  since we do not know the exact value of  $n_{K,L,M}$ . Because of this, the structure of the edge typical configurations is a little bit more complicated than in the 2D case.

### 9.1 Typical configurations

Let us start by defining the typical configurations. We consistently refer to Figure 9 for an illustration of our construction.

For  $a, b \in S$  and  $i \in \llbracket 0, M \rrbracket$ , we define

$$\widehat{\mathcal{R}}_i^{a,b} = \widehat{\mathcal{N}}(\mathcal{R}_i^{a,b}; \mathcal{G}^{a,b}). \tag{9.1}$$

We also define

$$\widehat{\mathcal{R}}_i^{A,B} = \widehat{\mathcal{N}}(\mathcal{R}_i^{A,B}; \mathcal{G}^{A,B}) \quad ; \quad i \in \llbracket 0, M \rrbracket .$$

**Remark 9.1.** For  $i \in \llbracket m_K, M - m_K \rrbracket$ , we have that

$$\widehat{\mathcal{R}}_i^{A,B} = \bigcup_{a \in A, b \in B} \widehat{\mathcal{R}}_i^{a,b} .$$

To check this, it suffices to check  $\widehat{\mathcal{N}}(\mathcal{R}_i^{a,b}; \mathcal{G}^{a,b}) = \widehat{\mathcal{N}}(\mathcal{R}_i^{a,b}; \mathcal{G}^{A,B})$  provided that  $a \in A$  and  $b \in B$ . This follows from Lemma 8.3 since  $\mathcal{R}_i^{a,b}$  cannot be connected to a configuration in  $\mathcal{G}^{A,B} \setminus \mathcal{G}^{a,b}$  via a  $\Gamma$ -path in  $\mathcal{X} \setminus \mathcal{G}^{a,b}$  by part (2) of Lemma 8.3.

**Remark 9.2.** By Lemma 8.3, two sets  $\widehat{\mathcal{R}}_i^{A,B}$  and  $\widehat{\mathcal{R}}_j^{A,B}$  for different  $i, j$  are disjoint if either  $i \in \llbracket m_K, M - m_K \rrbracket$  or  $j \in \llbracket m_K, M - m_K \rrbracket$ . Moreover by Proposition 8.1, they are disjoint if  $i \in \llbracket 0, m_K - 1 \rrbracket$  and  $j \in \llbracket M - m_K + 1, M \rrbracket$  or vice versa. On the other hand, they might be the same set if  $i, j \in \llbracket 0, m_K - 1 \rrbracket$  or  $i, j \in \llbracket M - m_K + 1, M \rrbracket$ . In particular, we have

$$\widehat{\mathcal{R}}_0^{A,B} = \widehat{\mathcal{R}}_1^{A,B} = \dots = \widehat{\mathcal{R}}_{n_{K,L,M}-1}^{A,B} ,$$

where  $n_{K,L,M}$  is defined in (6.22). The same result holds for  $\widehat{\mathcal{R}}_i^{a,b}$  instead of  $\widehat{\mathcal{R}}_i^{A,B}$ .

Now, we define the typical configurations. We recall Notation 7.4.

**Definition 9.3** (Typical configurations). For a proper partition  $(A, B)$  of  $S$ , we define the typical configurations as follows.

- **Bulk typical configurations:** We define, for  $a, b \in S$ ,

$$\mathcal{B}^{a,b} = \mathcal{G}_{[m_K, M-m_K-1]}^{a,b} \cup \widehat{\mathcal{R}}_{[m_K, M-m_K]}^{a,b},$$

and then define

$$\mathcal{B}^{A,B} = \bigcup_{a \in A} \bigcup_{b \in B} \mathcal{B}^{a,b} = \mathcal{G}_{[m_K, M-m_K-1]}^{A,B} \cup \widehat{\mathcal{R}}_{[m_K, M-m_K]}^{A,B}, \quad (9.2)$$

where the second identity holds because of Remark 9.1. A configuration belonging to  $\mathcal{B}^{A,B}$  is called a bulk typical configuration between  $S(A)$  and  $S(B)$ .

- **Edge typical configurations:** We define

$$\mathcal{E}^A = \mathcal{G}_{m_K-1}^{A,B} \cup \widehat{\mathcal{R}}_{[0, m_K]}^{A,B} \quad \text{and} \quad \mathcal{E}^B = \mathcal{G}_{M-m_K}^{A,B} \cup \widehat{\mathcal{R}}_{[M-m_K, M]}^{A,B}. \quad (9.3)$$

Finally, we define  $\mathcal{E}^{A,B} = \mathcal{E}^A \cup \mathcal{E}^B$ . A configuration belonging to  $\mathcal{E}^{A,B}$  is called an edge typical configuration between  $S(A)$  and  $S(B)$ .

Later in Proposition 9.6, we shall show that  $\mathcal{B}^{A,B} \cup \mathcal{E}^{A,B} = \widehat{\mathcal{N}}(S)$  and hence all relevant configurations in the analysis of metastable behavior between  $S(A)$  and  $S(B)$  belong to either  $\mathcal{B}^{A,B}$  or  $\mathcal{E}^{A,B}$ .

**Remark 9.4.** Since  $\mathcal{R}_0^{A,B} = S(A)$  and  $\mathcal{R}_M^{A,B} = S(B)$  (cf. (6.7)), we can readily observe that  $S(A) \subseteq \mathcal{E}^A$  and  $S(B) \subseteq \mathcal{E}^B$ .

## 9.2 Properties of typical configurations

In this subsection, we analyze some properties of the edge and bulk typical configurations. In fact, we have to take  $K$  large enough (i.e.,  $K \geq 2829$ ) in order to get the structural properties of edge and bulk typical configurations given in the current section.

The first property asserts that  $\mathcal{E}^A$  and  $\mathcal{E}^B$  are disjoint.

**Proposition 9.5.** The two sets  $\mathcal{E}^A$  and  $\mathcal{E}^B$  are disjoint.

*Proof.* By part (2) of Lemma 8.3 (cf. Remark 9.2), the set  $\widehat{\mathcal{R}}_{m_K}^{A,B}$  is disjoint with  $\mathcal{E}^B$ ; similarly, the set  $\widehat{\mathcal{R}}_{M-m_K}^{A,B}$  is disjoint with  $\mathcal{E}^A$ . It is direct from the definition that  $\mathcal{G}_{m_K-1}^{A,B}$  and  $\mathcal{G}_{M-m_K}^{A,B}$  are disjoint. By definition,  $\mathcal{G}_{m_K-1}^{A,B}, \mathcal{G}_{M-m_K}^{A,B} \subseteq \mathcal{G}^{A,B}$  are mutually disjoint with  $\widehat{\mathcal{R}}_{[0, m_K]}^{A,B}$  and  $\widehat{\mathcal{R}}_{[M-m_K, M]}^{A,B}$ . Hence, it suffices to prove that  $\widehat{\mathcal{R}}_{[0, m_K-1]}^{A,B}$  and  $\widehat{\mathcal{R}}_{[M-m_K+1, M]}^{A,B}$  are disjoint. Otherwise, we can take a configuration  $\sigma$  such that

$$\sigma \in \widehat{\mathcal{R}}_{[0, m_K-1]}^{A,B} \cap \widehat{\mathcal{R}}_{[M-m_K+1, M]}^{A,B}.$$

Since  $\sigma \in \widehat{\mathcal{R}}_{[0, m_K-1]}^{A,B}$ , there exists a  $\Gamma$ -path in  $\mathcal{X} \setminus \mathcal{G}^{A,B}$  (which is indeed a part of a canonical path) connecting  $\sigma$  and  $S(A)$ . Similarly, there exists a  $\Gamma$ -path in  $\mathcal{X} \setminus \mathcal{G}^{A,B}$  connecting  $\sigma$  and  $S(B)$ . By concatenating them, we can find a  $\Gamma$ -path  $(\omega_t)_{t=0}^T$  in  $\mathcal{X} \setminus \mathcal{G}^{A,B}$  connecting  $S(A)$  and  $S(B)$ . This contradicts part (2) of Proposition 8.1.  $\square$

Now, we analyze the crucial features of the typical configurations. Note that this is a 3D version of Proposition 5.9.

**Proposition 9.6.** For a proper partition  $(A, B)$  of  $S$ , the following properties hold for the typical configurations.

1. It holds that  $\mathcal{E}^A \cap \mathcal{B}^{A,B} = \widehat{\mathcal{R}}_{[0, m_K]}^{A,B}$  and  $\mathcal{E}^B \cap \mathcal{B}^{A,B} = \widehat{\mathcal{R}}_{[M-m_K]}^{A,B}$ .

2. We have  $\mathcal{E}^{A,B} \cup \mathcal{B}^{A,B} = \widehat{\mathcal{N}}(\mathcal{S})$ .

*Proof.* (1) It suffices to prove the first identity, as the second one follows similarly. One can observe that the set  $\mathcal{G}_{\mathfrak{m}_K-1}^{A,B} \subseteq \mathcal{G}^{A,B}$  is disjoint with  $\mathcal{B}^{A,B}$  from (9.2), and the set  $\mathcal{G}_{[\mathfrak{m}_K, M-\mathfrak{m}_K-1]}^{A,B} \subseteq \mathcal{G}^{A,B}$  is disjoint with  $\mathcal{E}^A$  in view of the expression (9.3). Therefore, by (9.2) and (9.3), we get

$$\mathcal{E}^A \cap \mathcal{B}^{A,B} = \widehat{\mathcal{R}}_{[0, \mathfrak{m}_K]}^{A,B} \cap \widehat{\mathcal{R}}_{[\mathfrak{m}_K, M-\mathfrak{m}_K]}^{A,B}. \tag{9.4}$$

By Lemma 8.3, the two sets  $\mathcal{R}_{[0, \mathfrak{m}_K-1]}^{A,B}$  and  $\mathcal{R}_{[\mathfrak{m}_K, M-\mathfrak{m}_K]}^{A,B}$  cannot be connected by a  $\Gamma$ -path in  $\mathcal{X} \setminus \mathcal{G}^{A,B}$ , and hence  $\widehat{\mathcal{R}}_{[0, \mathfrak{m}_K-1]}^{A,B}$  and  $\widehat{\mathcal{R}}_{[\mathfrak{m}_K, M-\mathfrak{m}_K]}^{A,B}$  are disjoint. Therefore, we have

$$\widehat{\mathcal{R}}_{[0, \mathfrak{m}_K]}^{A,B} \cap \widehat{\mathcal{R}}_{[\mathfrak{m}_K, M-\mathfrak{m}_K]}^{A,B} = \widehat{\mathcal{R}}_{\mathfrak{m}_K}^{A,B} \cap \widehat{\mathcal{R}}_{[\mathfrak{m}_K, M-\mathfrak{m}_K]}^{A,B} = \widehat{\mathcal{R}}_{\mathfrak{m}_K}^{A,B}. \tag{9.5}$$

The proof is completed by (9.4) and (9.5).

(2) We will first prove that

$$\widehat{\mathcal{N}}(\mathcal{S}) = \widehat{\mathcal{N}}(\mathcal{N}(\mathcal{R}_{[0, M]}^{A,B}) \cup \mathcal{G}^{A,B}). \tag{9.6}$$

Since it is immediate that  $\mathcal{S}$  is a subset of the right-hand side, we have  $\widehat{\mathcal{N}}(\mathcal{S}) \subseteq \widehat{\mathcal{N}}(\mathcal{N}(\mathcal{R}_{[0, M]}^{A,B}) \cup \mathcal{G}^{A,B})$ . On the other hand, since  $\mathcal{N}(\mathcal{R}_{[0, M]}^{A,B}) \cup \mathcal{G}^{A,B} \subseteq \widehat{\mathcal{N}}(\mathcal{S})$  clearly holds, we also have  $\widehat{\mathcal{N}}(\mathcal{N}(\mathcal{R}_{[0, M]}^{A,B}) \cup \mathcal{G}^{A,B}) \subseteq \widehat{\mathcal{N}}(\mathcal{S})$ . This proves (9.6). Since the sets  $\mathcal{N}(\mathcal{R}_{[0, M]}^{A,B})$  and  $\mathcal{G}^{A,B}$  are disjoint by Lemma 7.6, we can apply Lemmas 4.2 and 7.6 to deduce

$$\begin{aligned} \widehat{\mathcal{N}}(\mathcal{S}) &= \widehat{\mathcal{N}}(\mathcal{N}(\mathcal{R}_{[0, M]}^{A,B}); \mathcal{G}^{A,B}) \cup \widehat{\mathcal{N}}(\mathcal{G}^{A,B}; \mathcal{N}(\mathcal{R}_{[0, M]}^{A,B})) \\ &= \widehat{\mathcal{N}}(\mathcal{R}_{[0, M]}^{A,B}; \mathcal{G}^{A,B}) \cup \mathcal{G}^{A,B} = \widehat{\mathcal{R}}_{[0, M]}^{A,B} \cup \mathcal{G}^{A,B}. \end{aligned}$$

This completes the proof since by (7.3), (9.2) and (9.3), we have  $\mathcal{E}^{A,B} \cup \mathcal{B}^{A,B} = \mathcal{G}^{A,B} \cup \widehat{\mathcal{R}}_{[0, M]}^{A,B}$ . □

### 9.3 Structure of edge typical configurations

As in the 2D case [27, Sections 6.4 and 6.5], we analyze the structure of edge typical configurations.

We remark that we fixed a proper partition  $(A, B)$  of  $S$ . Decompose

$$\mathcal{E}^A = \mathcal{O}^A \cup \mathcal{I}^A \tag{9.7}$$

where

$$\mathcal{O}^A = \{\sigma \in \mathcal{E}^A : H(\sigma) = \Gamma\} \quad \text{and} \quad \mathcal{I}^A = \{\sigma \in \mathcal{E}^A : H(\sigma) < \Gamma\}.$$

We now take a subset  $\overline{\mathcal{I}}^A$  of  $\mathcal{I}^A$  so that we can decompose  $\mathcal{I}^A$  into the following disjoint union:

$$\mathcal{I}^A = \bigcup_{\sigma \in \overline{\mathcal{I}}^A} \mathcal{N}(\sigma).$$

Consequently, we get the following decomposition of  $\mathcal{E}^A$ :

$$\mathcal{E}^A = \mathcal{O}^A \cup \left( \bigcup_{\sigma \in \overline{\mathcal{I}}^A} \mathcal{N}(\sigma) \right). \tag{9.8}$$

**Notation 9.7.** For  $\sigma \in \mathcal{I}^A$ , we denote by  $\overline{\sigma} \in \overline{\mathcal{I}}^A$  the unique configuration satisfying  $\sigma \in \mathcal{N}(\overline{\sigma})$ .

By part (1) of Lemma 8.3, for  $\sigma, \sigma' \in \mathcal{R}_{m_K}^{A,B}$ , the two sets  $\mathcal{N}(\sigma)$  and  $\mathcal{N}(\sigma')$  are disjoint. By a similar reasoning, we know that for any  $\sigma \in \mathcal{R}_{m_K}^{A,B}$  and  $a \in A$ , the sets  $\mathcal{N}(\sigma)$  and  $\mathcal{N}(s_a)$  are disjoint. Thus, we can assume that

$$\mathcal{R}_{m_K}^{A,B} \cup \mathcal{S}(A) \subseteq \bar{\mathcal{I}}^A. \tag{9.9}$$

The following construction of an auxiliary Markov chain is an analogue of [27, Definition 6.20].

**Definition 9.8.** For a proper partition  $(A, B)$  of  $S$ , we define a Markov chain  $Z^A(\cdot)$  on  $\bar{\mathcal{I}}^A \cup \mathcal{O}^A$ .

- **(Graph)** We define the graph structure  $\mathcal{G}^A = (\mathcal{V}^A, \mathcal{E}^A)$  for  $\mathcal{V}^A = \mathcal{O}^A \cup \bar{\mathcal{I}}^A$ . The edge set  $\mathcal{E}^A$  is defined by declaring that  $\{\sigma, \sigma'\} \in \mathcal{E}^A$  for  $\sigma, \sigma' \in \mathcal{V}^A$  if

$$\begin{cases} \sigma, \sigma' \in \mathcal{O}^A \text{ and } \sigma \sim \sigma' \text{ or} \\ \sigma \in \mathcal{O}^A, \sigma' \in \bar{\mathcal{I}}^A, \text{ and } \sigma \sim \zeta \text{ for some } \zeta \in \mathcal{N}(\sigma'). \end{cases}$$

- **(Markov chain)** We first define a rate  $r^A : \mathcal{V}^A \times \mathcal{V}^A \rightarrow [0, \infty)$ . If  $\{\sigma, \sigma'\} \notin \mathcal{E}^A$ , we set  $r^A(\sigma, \sigma') = 0$ , and if  $\{\sigma, \sigma'\} \in \mathcal{E}^A$ , we set

$$r^A(\sigma, \sigma') = \begin{cases} 1 & \text{if } \sigma, \sigma' \in \mathcal{O}^A, \\ |\{\zeta \in \mathcal{N}(\sigma) : \zeta \sim \sigma'\}| & \text{if } \sigma \in \bar{\mathcal{I}}^A, \sigma' \in \mathcal{O}^A, \\ |\{\zeta \in \mathcal{N}(\sigma') : \zeta \sim \sigma\}| & \text{if } \sigma \in \mathcal{O}^A, \sigma' \in \bar{\mathcal{I}}^A. \end{cases} \tag{9.10}$$

We now let  $(Z^A(t))_{t \geq 0}$  be the continuous-time Markov chain on  $\mathcal{V}^A$  with rate  $r^A(\cdot, \cdot)$ . Note that the uniform distribution on  $\mathcal{V}^A$  is the invariant measure for the chain  $Z^A(\cdot)$ , and indeed this chain is reversible with respect to this measure.

- **(Potential-theoretic objects)** Denote by  $L^A, h_{\cdot}^A(\cdot)$ , and  $\text{cap}^A(\cdot, \cdot)$  the generator, equilibrium potential, and capacity with respect to the Markov chain  $Z^A(\cdot)$ , respectively.

We now give three important propositions regarding the objects constructed above. These propositions play fundamental roles in the construction of the test function on the edge typical configurations.

We remark from (9.9) that  $\mathcal{S}(A), \mathcal{R}_{m_K}^{A,B} \subseteq \bar{\mathcal{I}}^A \subseteq \mathcal{V}^A$ . Potential-theoretic objects between these two sets are crucially used in our discussion. We define

$$\epsilon(A) = \frac{1}{|\mathcal{V}^A| \text{cap}^A(\mathcal{S}(A), \mathcal{R}_{m_K}^{A,B})}. \tag{9.11}$$

For  $n \in \llbracket 1, q-1 \rrbracket$  (with a slight abuse of notation) we can write

$$\epsilon(n) = \frac{1}{|\mathcal{V}^{A_n}| \text{cap}^{A_n}(\mathcal{S}(A_n), \mathcal{R}_{m_K}^{A_n, B_n})}, \tag{9.12}$$

where  $A_n = \{1, \dots, n\}$  and  $B_n = \{n+1, \dots, q\}$ . Since  $\epsilon(A)$  depends on  $A$  only through  $|A|$ , it holds that  $\epsilon(A) = \epsilon(|A|)$ . We next derive a rough bound of  $\epsilon(n)$  via the Thomson principle. We refer to e.g., [13] for the flow structure and the Thomson principle.

**Proposition 9.9.** For all  $n \in \llbracket 1, q-1 \rrbracket$ , we have that  $\epsilon(n) \leq \frac{1}{K^{1/3}}$ .

*Proof.* We recall that an anti-symmetric function  $\phi : \mathcal{V}^A \times \mathcal{V}^A \rightarrow \mathbb{R}$  is called a flow associated with the Markov chain  $Z^A(\cdot)$ , provided that  $\phi(x, y) \neq 0$  if and only if  $\{x, y\} \in \mathcal{E}^A$ . For each flow  $\phi$ , the associated flow norm is defined by

$$\|\phi\|^2 := \sum_{\{x, y\} \subseteq \mathcal{V}^A: \{x, y\} \in \mathcal{E}^A} \frac{\phi(x, y)^2}{|\mathcal{V}^A|^{-1} r^A(x, y)}.$$

For each  $x \in \mathcal{V}^A$ , the divergence of a flow  $\phi$  at  $x$  is defined by

$$(\operatorname{div} \phi)(x) := \sum_{y \in \mathcal{V}^A} \phi(x, y).$$

Finally, for two disjoint non-empty subsets  $\mathcal{U}, \mathcal{V}$  of  $\mathcal{V}^A$ , a flow  $\phi$  is called a *unit flow* from  $\mathcal{U}$  to  $\mathcal{V}$  if

$$\sum_{x \in \mathcal{U}} (\operatorname{div} \phi)(x) = - \sum_{x \in \mathcal{V}} (\operatorname{div} \phi)(x) = 1 \quad \text{and} \quad (\operatorname{div} \phi)(z) = 0 \quad \text{for all } z \notin \mathcal{U} \cup \mathcal{V}.$$

Then, by the Thomson principle (cf. [13, Theorem 7.37]), for any *unit flow*  $\psi$  from  $\mathcal{S}(A)$  to  $\mathcal{R}_{\mathfrak{m}_K}^{A,B}$ , we have

$$\operatorname{cap}^A(\mathcal{S}(A), \mathcal{R}_{\mathfrak{m}_K}^{A,B}) \geq \frac{1}{\|\psi\|^2}. \tag{9.13}$$

We shall construct below a unit flow  $\psi$  from  $\mathcal{S}(A)$  to  $\mathcal{R}_{\mathfrak{m}_K}^{A,B}$  that satisfies

$$\|\psi\|^2 < \frac{\mathfrak{m}_K |\mathcal{V}^A|}{2M}. \tag{9.14}$$

Then, by combining (9.13) and (9.14), we have (recalling the definition (6.20) of  $\mathfrak{m}_K$ )

$$\operatorname{cap}^A(\mathcal{S}(A), \mathcal{R}_{\mathfrak{m}_K}^{A,B}) \geq \frac{1}{\|\psi\|^2} > \frac{2M}{|\mathcal{V}^A| \mathfrak{m}_K} \geq \frac{K^{1/3}}{|\mathcal{V}^A|}.$$

Recalling the definition (9.11), this completes the proof.

Now, it remains to construct a unit flow  $\psi$  from  $\mathcal{S}(A)$  to  $\mathcal{R}_{\mathfrak{m}_K}^{A,B}$  satisfying bound (9.14). To this end, let us first fix  $a \in A$  and  $b \in B$ . Define

$$i_{K,L,M} = \max \{m \geq 1 : \Phi(\mathcal{S}(A), \mathcal{R}_m^{A,B}) < \Gamma\}. \tag{9.15}$$

By Corollary 8.5, we know that  $i_{K,L,M} < \mathfrak{m}_K$ .

Let us start by fixing  $P, Q \in \mathfrak{S}_M$  such that  $P \prec Q$ ,  $Q \setminus P = \{m\}$ ,  $i_{K,L,M} \leq |P| < |Q| \leq \mathfrak{m}_K$ ,  $a \in A$ , and  $b \in B$ . Then, we first define a flow  $\psi_{P,Q}$  connecting  $\bar{\sigma}_P^{a,b} = \bar{\sigma}_P^{a,b}$  and  $\bar{\sigma}_Q^{a,b} = \bar{\sigma}_Q^{a,b}$  (cf. Notation 9.7). First, we set

$$\psi_{P,Q}(\sigma, \zeta) = -\psi_{P,Q}(\zeta, \sigma) = \frac{1}{2KLM} \tag{9.16}$$

if  $\sigma, \zeta \in \mathcal{C}_{P,Q}^{a,b}$  satisfy, for some  $\ell \in \mathbb{T}_L$ ,  $k \in \mathbb{T}_K$ ,  $v \in \llbracket 1, L-2 \rrbracket$ , and  $h \in \llbracket 1, K-2 \rrbracket$ ,

$$\begin{cases} \sigma^{(m)} = \xi_{\ell,v}^{a,b} \quad \text{and} \quad \zeta^{(m)} = \xi_{\ell,v;k,1}^{a,b,+} & \text{or,} \\ \sigma^{(m)} = \xi_{\ell,v;k,h}^{a,b,+} \quad \text{and} \quad \zeta^{(m)} = \xi_{\ell,v;k,h+1}^{a,b,+} & \text{or,} \\ \sigma^{(m)} = \xi_{\ell,v;k,K-1}^{a,b,+} \quad \text{and} \quad \zeta^{(m)} = \xi_{\ell,v+1}^{a,b}. \end{cases} \tag{9.17}$$

Now, we claim that all configurations that appear in (9.17) except for the ones corresponding to  $\xi_{\ell,1}^{a,b}$  and  $\xi_{\ell,L-1}^{a,b}$  belong to  $\mathcal{V}^A$ . To check this, observe first that if the  $m$ -th floor of  $\sigma \in \mathcal{C}_{P,Q}^{a,b}$  is of the form  $\sigma^{(m)} = \xi_{\ell,v;k,h}^{a,b,+}$ , we have  $H(\sigma) = \Gamma$  and hence  $\sigma \in \mathcal{O}^A$ . On the other hand, if the  $m$ -th floor of  $\sigma \in \mathcal{C}_{P,Q}^{a,b}$  is of the form  $\sigma^{(m)} = \xi_{\ell,v}^{a,b}$  for  $v \in \llbracket 2, L-2 \rrbracket$ , we have  $H(\sigma) = \Gamma - 2$  and moreover  $\mathcal{N}(\sigma) = \{\sigma\}$ . This implies that  $\sigma \in \bar{\mathcal{I}}^A$ . This proves the claim. On the other hand, since if  $\sigma^{(m)} = \xi_{\ell,1}^{a,b}$  then  $\sigma \in \mathcal{N}(\bar{\sigma}_P^{a,b})$ , and if  $\sigma^{(m)} = \xi_{\ell,L-1}^{a,b}$  then  $\sigma \in \mathcal{N}(\bar{\sigma}_Q^{a,b})$  (cf. the canonical paths provide  $(\Gamma - 1)$ -paths), we can replace the configurations corresponding to  $\xi_{\ell,1}^{a,b}$  and  $\xi_{\ell,L-1}^{a,b}$  that appear in (9.17) with  $\bar{\sigma}_P^{a,b}$  and

$\bar{\sigma}_Q^{a,b}$ , respectively, to get a flow connecting  $\bar{\sigma}_P^{a,b}$  and  $\bar{\sigma}_Q^{a,b}$ . We remark that we may have  $\bar{\sigma}_P^{a,b} = \bar{\sigma}_Q^{a,b}$ .

We deduce from the definition of the flow norm that

$$\|\psi_{P,Q}\|^2 = \frac{|\mathcal{V}^A|}{(2KLM)^2} \times K^2L(L-2) < \frac{|\mathcal{V}^A|}{4M^2}, \tag{9.18}$$

where  $K^2L(L-2)$  is the number of edges that appear in (9.17). Next, we define

$$\psi = \sum_{r=i_{K,L,M}}^{m_K-1} \sum_{P,Q \in \mathfrak{S}_M: |P|=r, P \prec Q} \psi_{P,Q}.$$

Notice from (9.15) that a configuration of the form  $\bar{\sigma}_{P,Q}^{a,b}$  with  $|P| = i_{K,L,M}$  is indeed an element of  $\mathcal{S}(A)$ . Then, from the definition (9.16), we can readily check that  $\psi(x) = 0$  for all  $x \notin \mathcal{S}(A) \cup \mathcal{R}_{m_K}^{A,B}$  (by using the fact that the flow on each edge has a constant magnitude  $\frac{1}{2KLM}$ ), and moreover it holds that (cf. (9.9))

$$\sum_{x \in \mathcal{S}(A)} \sum_{y \in \mathcal{V}^A} \psi(x, y) = 1. \tag{9.19}$$

Indeed, to prove the last assertion, it suffices to observe that

$$\begin{aligned} \sum_{x \in \mathcal{S}(A)} \sum_{y \in \mathcal{V}^A} \psi(x, y) &= \sum_{P,Q \in \mathfrak{S}_M: |P|=i_{K,L,M}, P \prec Q} \sum_{\zeta \in \mathcal{V}^A: \bar{\sigma}_P^{a,b} \sim \zeta} \psi_{P,Q}(\bar{\sigma}_P^{a,b}, \zeta) \\ &= \frac{1}{2KLM} \times KL \times 2M = 1, \end{aligned}$$

where  $KL$  is the number of configurations in  $\mathcal{C}_{P,Q}^{a,b}$  connected to  $\bar{\sigma}_P^{a,b}$ , and  $2M$  is the number of possible choices of  $P$  and  $Q$ . Consequently, the flow  $\psi$  is a unit flow from  $\mathcal{S}(A)$  to  $\mathcal{R}_{m_K}^{A,B}$ .

Thus, it suffices to verify (9.14). Since the support of  $\psi_{P,Q}$  (which is the collection of edges on which  $\psi_{P,Q}$  is non-zero) for different pairs  $(P, Q)$  are disjoint, we deduce from (9.18) that

$$\|\psi\|^2 = \sum_{r=i_{K,L,M}}^{m_K-1} \sum_{P,Q \in \mathfrak{S}_M: |P|=r, P \prec Q} \|\psi_{P,Q}\|^2 < m_K \times 2M \times \frac{|\mathcal{V}^A|}{4M^2} = \frac{m_K |\mathcal{V}^A|}{2M},$$

and therefore  $\psi$  satisfies (9.14). □

For simplicity, we write (cf. (9.9))

$$\mathfrak{h}^A(\cdot) = h_{\mathcal{S}(A), \mathcal{R}_{m_K}^{A,B}}^A(\cdot) \tag{9.20}$$

where  $h^A$  is the equilibrium potential defined in Definition 9.8. This function is a fundamental object in the construction of the test function in Section 10.

**Proposition 9.10.** *For  $\sigma \in \widehat{\mathcal{R}}_{m_K}^{A,B} \cap \mathcal{O}^A \subseteq \mathcal{V}^A$ , we have  $\mathfrak{h}^A(\sigma) = 0$ .*

*Proof.* We fix  $\sigma \in \widehat{\mathcal{R}}_{m_K}^{A,B} \cap \mathcal{O}^A$ . It suffices to prove that any  $\Gamma$ -path  $(\omega_t)_{t=0}^T$  from  $\sigma$  to  $\mathcal{S}(A)$  must visit  $\mathcal{N}(\mathcal{R}_{m_K}^{A,B})$ . Suppose first that the path  $(\omega_t)_{t=0}^T$  does not visit  $\mathcal{G}^{A,B}$ . Since  $\sigma \in \widehat{\mathcal{R}}_{m_K}^{A,B}$ , there exists a  $\Gamma$ -path in  $\mathcal{X} \setminus \mathcal{G}^{A,B}$  connecting  $\mathcal{R}_{m_K}^{A,B}$  and  $\sigma$ , and therefore by concatenating this path with  $(\omega_t)_{t=0}^T$ , we get a  $\Gamma$ -path in  $\mathcal{X} \setminus \mathcal{G}^{A,B}$  connecting  $\mathcal{R}_{m_K}^{A,B}$  and  $\mathcal{S}(A)$ . This contradicts part (2) of Lemma 8.3. Thus, the path  $(\omega_t)_{t=0}^T$  must visit  $\mathcal{G}^{A,B}$  and we let

$$t_0 = \min \{t : \omega_t \in \mathcal{G}^{A,B}\}.$$

By part (2) of Lemma 7.5, we have  $\omega_{t_0-1} \in \mathcal{N}(\mathcal{R}_{[m_K-1, M-m_K+1]}^{A,B})$ . If  $\omega_{t_0-1} \in \mathcal{N}(\mathcal{R}_i^{A,B})$  for some  $i \in [m_K - 1, M - m_K + 1] \setminus \{m_K\}$ , then  $(\omega_t)_{t=0}^{t_0-1}$  induces a  $\Gamma$ -path from  $\mathcal{R}_{m_K}^{A,B}$  from  $\mathcal{R}_i^{A,B}$  avoiding  $\mathcal{G}^{A,B}$ , which contradicts part (2) of Lemma 8.3. Hence, we can conclude that  $\omega_{t_0-1} \in \mathcal{N}(\mathcal{R}_{m_K}^{A,B})$ , as desired.  $\square$

**Remark 9.11.** The previous proposition implies that configurations  $\sigma$  that belong to  $\widehat{\mathcal{R}}_{m_K}^{A,B} \cap \mathcal{O}^A$  are dead-ends attached to  $\mathcal{N}(\mathcal{R}_{m_K}^{A,B})$  (cf. grey protuberances attached to green boxes in Figures 8 and 9).

The next proposition highlights the fact that the auxiliary process  $Z^A(\cdot)$  defined in Definition 9.8 approximates the behavior of the Metropolis–Hastings dynamics at the edge typical configurations.

**Proposition 9.12.** Define a projection map  $\Pi^A : \mathcal{E}^A \rightarrow \mathcal{V}^A$  by (cf. Notation 9.7)

$$\Pi^A(\sigma) = \begin{cases} \bar{\sigma} & \text{if } \sigma \in \mathcal{I}^A, \\ \sigma & \text{if } \sigma \in \mathcal{O}^A. \end{cases}$$

Then, there exists  $C = C(K, L, M) > 0$  such that

1. for  $\sigma_1, \sigma_2 \in \mathcal{O}^A$ , we have

$$\left| \frac{1}{q} e^{-\Gamma\beta} r^A(\Pi^A(\sigma_1), \Pi^A(\sigma_2)) - \mu_\beta(\sigma_1) r_\beta(\sigma_1, \sigma_2) \right| \leq C e^{-(\Gamma+1)\beta}, \quad (9.21)$$

2. for  $\sigma_1 \in \mathcal{O}^A$  and  $\sigma_2 \in \bar{\mathcal{I}}^A$ , we have

$$\left| \frac{1}{q} e^{-\Gamma\beta} r^A(\Pi^A(\sigma_1), \Pi^A(\sigma_2)) - \sum_{\zeta \in \mathcal{N}(\sigma_2)} \mu_\beta(\sigma_1) r_\beta(\sigma_1, \zeta) \right| \leq C e^{-(\Gamma+1)\beta}. \quad (9.22)$$

*Proof.* (1) Suppose that  $\sigma_1, \sigma_2 \in \mathcal{O}^A$ . Since if  $\sigma_1 \not\sim \sigma_2$  then the left-hand side of (9.21) is 0, we may assume that  $\sigma_1 \sim \sigma_2$ . In this case,  $\{\sigma_1, \sigma_2\} \in \mathcal{E}^A$ , and thus

$$\left| \frac{1}{q} e^{-\Gamma\beta} r^A(\Pi^A(\sigma_1), \Pi^A(\sigma_2)) - \mu_\beta(\sigma_1) r_\beta(\sigma_1, \sigma_2) \right| = \left| \frac{1}{q} e^{-\Gamma\beta} - \frac{1}{Z_\beta} e^{-\Gamma\beta} \right|$$

since  $\mu_\beta(\sigma_1) = \mu_\beta(\sigma_2) = \frac{1}{Z_\beta} e^{-\Gamma\beta}$  by the definition of  $\mathcal{O}^A$ . By (2.5), the right-hand side of the previous display is  $O_\beta(e^{-(\Gamma+1)\beta})$ .

(2) Let  $\sigma_1 \in \mathcal{O}^A$  and  $\sigma_2 \in \bar{\mathcal{I}}^A$ . Similarly, we may assume that  $\sigma_1 \sim \sigma_2$ . Then, we can write

$$\begin{aligned} & \left| \frac{1}{q} e^{-\Gamma\beta} r^A(\Pi^A(\sigma_1), \Pi^A(\sigma_2)) - \sum_{\zeta \in \mathcal{N}(\sigma_2)} \mu_\beta(\sigma_1) r_\beta(\sigma_1, \zeta) \right| \\ &= \left| \frac{1}{q} e^{-\Gamma\beta} |\{\zeta \in \mathcal{N}(\sigma_2) : \zeta \sim \sigma_1\}| - \sum_{\zeta \in \mathcal{N}(\sigma_2) : \zeta \sim \sigma_1} \min\{\mu_\beta(\sigma_1), \mu_\beta(\zeta)\} \right| \\ &= |\{\zeta \in \mathcal{N}(\sigma_2) : \zeta \sim \sigma_1\}| \times \left| \frac{1}{q} e^{-\Gamma\beta} - \frac{1}{Z_\beta} e^{-\Gamma\beta} \right|, \end{aligned}$$

since  $\min\{\mu_\beta(\sigma_1), \mu_\beta(\zeta)\} = \mu_\beta(\sigma_1)$  for all  $\zeta \in \mathcal{N}(\sigma_2)$ . Again by (2.5), the last line is clearly bounded from above by  $KL \times O_\beta(e^{-\Gamma\beta} e^{-\beta}) = O_\beta(e^{-(\Gamma+1)\beta})$ . This concludes the proof.  $\square$

### 9.4 Analysis of 3D transition paths

In this section, we finally define the collection of transition paths between ground states that appear in Theorem 2.9.

**Definition 9.13** (Transition paths). Write

$$\mathcal{H}^{A,B} = \mathcal{G}^{A,B} \cup \widehat{\mathcal{R}}_{[\mathfrak{m}_K, M-\mathfrak{m}_K]}^{A,B}. \quad (9.23)$$

A path  $(\omega_t)_{t=0}^T$  is called a transition path between  $\mathcal{S}(A)$  and  $\mathcal{S}(B)$  if

$$\omega_0 \in \widehat{\mathcal{N}}(\mathcal{S}(A); \mathcal{G}^{A,B}), \omega_T \in \widehat{\mathcal{N}}(\mathcal{S}(B); \mathcal{G}^{A,B}), \text{ and} \\ \omega_t \in \mathcal{H}^{A,B} \text{ for all } t \in \llbracket 1, T-1 \rrbracket.$$

In particular, we have  $\omega_0 \in \mathcal{N}(\mathcal{R}_{\mathfrak{m}_K-1}^{A,B})$ ,  $\omega_1 \in \mathcal{G}_{\mathfrak{m}_K-1}^{A,B}$ ,  $\omega_{T-1} \in \mathcal{G}_{M-\mathfrak{m}_K}^{A,B}$ , and  $\omega_T \in \mathcal{N}(\mathcal{R}_{M-\mathfrak{m}_K+1}^{A,B})$  by part (1) of Lemma 7.5.

**Remark 9.14.** The two sets  $\widehat{\mathcal{N}}(\mathcal{S}(A); \mathcal{G}^{A,B})$  and  $\widehat{\mathcal{N}}(\mathcal{S}(B); \mathcal{G}^{A,B})$  are disjoint thanks to part (2) of Proposition 8.1.

Now, we characterize all the optimal paths between ground states in terms of the transition paths.

**Theorem 9.15.** Let  $(\omega_t)_{t=0}^T$  be a  $\Gamma$ -path connecting  $\mathcal{S}(A)$  and  $\mathcal{S}(B)$ . Then,  $(\omega_t)_{t=0}^T$  has a transition path between  $\mathcal{S}(A)$  and  $\mathcal{S}(B)$  as a sub-path.

*Proof.* Let  $(\omega_t)_{t=0}^T$  be a  $\Gamma$ -path connecting  $\mathcal{S}(A)$  and  $\mathcal{S}(B)$ , and define

$$T' = \min \{t : \omega_t \in \widehat{\mathcal{N}}(\mathcal{S}(B); \mathcal{G}^{A,B})\}.$$

Then, define

$$t' = \max \{t < T' : \omega_t \in \widehat{\mathcal{N}}(\mathcal{S}(A); \mathcal{G}^{A,B})\}.$$

We claim that the sub-path  $(\omega_t)_{t=t'}^{T'}$  is a transition path between  $\mathcal{S}(A)$  and  $\mathcal{S}(B)$ . By part (1) of Lemma 7.5, we have

$$\omega_{t'} \in \mathcal{N}(\mathcal{R}_{\mathfrak{m}_K-1}^{A,B}), \omega_{t'+1} \in \mathcal{G}_{\mathfrak{m}_K-1}^{A,B}, \omega_{T'-1} \in \mathcal{G}_{M-\mathfrak{m}_K}^{A,B}, \text{ and } \omega_{T'} \in \mathcal{N}(\mathcal{R}_{M-\mathfrak{m}_K+1}^{A,B}).$$

In particular, we get  $\omega_{t'+1}, \omega_{T'-1} \in \mathcal{H}^{A,B}$ . To complete the proof of the claim, it suffices to check that, if  $\sigma \in \mathcal{H}^{A,B}$  and  $\zeta \notin \mathcal{H}^{A,B}$  satisfy  $\sigma \sim \zeta$  and  $H(\zeta) \leq \Gamma$ , then  $\zeta \in \mathcal{N}(\mathcal{R}_{\mathfrak{m}_K-1}^{A,B}) \cup \mathcal{N}(\mathcal{R}_{M-\mathfrak{m}_K+1}^{A,B})$ . To prove this, let us first assume that  $\sigma \in \widehat{\mathcal{R}}_i^{a,b}$  for some  $a \in A, b \in B$ , and  $i \in \llbracket \mathfrak{m}_K, M-\mathfrak{m}_K \rrbracket$  (cf. (9.23)). Then, since  $\zeta \notin \widehat{\mathcal{R}}_i^{a,b}$  and  $H(\zeta) \leq \Gamma$ , by the definition of  $\widehat{\mathcal{R}}_i^{a,b}$  we must have  $\zeta \in \mathcal{G}^{A,B}$ , and hence we get a contradiction to the fact that  $\zeta \notin \mathcal{H}^{A,B}$ . Next, we assume that  $\sigma \in \mathcal{G}_i^{A,B}$  for some  $i \in \llbracket \mathfrak{m}_K-1, M-\mathfrak{m}_K \rrbracket$ . Since  $\zeta \in \mathcal{X} \setminus \mathcal{H}^{A,B}$ , by Lemma 7.5, we have  $\zeta \in \mathcal{N}(\mathcal{R}_{\mathfrak{m}_K-1}^{A,B}) \cup \mathcal{N}(\mathcal{R}_{M-\mathfrak{m}_K+1}^{A,B})$ . This completes the proof.  $\square$

Therefore, we can now say that the set  $\mathcal{G}^{A,B} \cup \widehat{\mathcal{R}}_{[\mathfrak{m}_K, M-\mathfrak{m}_K]}^{A,B}$  consists of a saddle plateau between  $\mathcal{S}(A)$  and  $\mathcal{S}(B)$ , which is a huge set of saddle configurations.

Now, we can prove Theorem 2.9.

*Proof of Theorem 2.9.* Denote by  $\widehat{\tau}$  the hitting time of the set  $\{\sigma \in \mathcal{X} : H(\sigma) \geq \Gamma + 1\}$ . Then, by the large deviation principle (e.g. [43, Theorem 3.2]), we have that

$$\mathbb{P}_s^\beta [\widehat{\tau} < e^{\beta(\Gamma+1/2)}] = o_\beta(1).$$

Hence, by part (1) of Theorem 2.7, we have that  $\mathbb{P}_s^\beta [\tau_s^\beta < \widehat{\tau}] = 1 - o_\beta(1)$ . Thus, the conclusion of the theorem follows immediately from Theorem 9.15.  $\square$

### 10 Construction of test function

We fix in this section a proper partition  $(A, B)$  of  $S$ . The main purpose of the current section is to construct a test function  $\tilde{h} = \tilde{h}_{A,B}^B : \mathcal{X} \rightarrow \mathbb{R}$  that satisfies the two requirements of Proposition 3.2.

**Notation 10.1.** *Since the partition  $(A, B)$  is fixed, we simply write  $\mathfrak{b} = \mathfrak{b}(|A|)$ ,  $\epsilon_A = \epsilon(|A|)$ ,  $\epsilon_B = \epsilon(|B|)$ , and  $\mathfrak{c} = \mathfrak{c}(|A|)$  so that  $\mathfrak{c} = \mathfrak{b} + \epsilon_A + \epsilon_B$  throughout the current section.*

#### 10.1 Construction of test function

We now define a function  $\tilde{h} : \mathcal{X} \rightarrow \mathbb{R}$  which indeed fulfills all requirements in Proposition 3.2, as we shall verify later.

**Definition 10.2** (Test function). *We construct the test function  $\tilde{h}$  on  $\mathcal{E}^{A,B}$ ,  $\mathcal{B}^{A,B}$ , and  $(\mathcal{E}^{A,B} \cup \mathcal{B}^{A,B})^c$  separately. Recall Notation 9.7.*

1. **Construction of  $\tilde{h}$  on edge typical configurations  $\mathcal{E}^{A,B} = \mathcal{E}^A \cup \mathcal{E}^B$ .**

- For  $\sigma \in \mathcal{E}^A$ , we recall the decomposition (9.8) of  $\mathcal{E}^A$  and define

$$\tilde{h}(\sigma) = \begin{cases} 1 - \frac{\epsilon_A}{\mathfrak{c}}(1 - \mathfrak{h}^A(\sigma)) & \text{if } \sigma \in \mathcal{O}^A, \\ 1 - \frac{\epsilon_A}{\mathfrak{c}}(1 - \mathfrak{h}^A(\bar{\sigma})) & \text{if } \sigma \in \mathcal{I}^A. \end{cases} \tag{10.1}$$

- For  $\sigma \in \mathcal{E}^B$ , we similarly define

$$\tilde{h}(\sigma) = \begin{cases} \frac{\epsilon_B}{\mathfrak{c}}(1 - \mathfrak{h}^B(\sigma)) & \text{if } \sigma \in \mathcal{O}^B, \\ \frac{\epsilon_B}{\mathfrak{c}}(1 - \mathfrak{h}^B(\bar{\sigma})) & \text{if } \sigma \in \mathcal{I}^B. \end{cases} \tag{10.2}$$

2. **Construction of  $\tilde{h}$  on bulk typical configurations  $\mathcal{B}^{A,B}$ .** Recall the 2D test function  $\tilde{h}^{2D}$  explained in Proposition 5.12. We define the test function on each component of the decomposition (9.2) of  $\mathcal{B}^{A,B}$ .

- Construction on  $\mathcal{G}_{[\mathfrak{m}_K, M - \mathfrak{m}_K - 1]}^{A,B}$ : Let us first fix  $P, Q \in \mathfrak{S}_M$  such that  $P \prec Q$  and  $|P| \in [\mathfrak{m}_K, M - \mathfrak{m}_K - 1]$ . Write

$$\mathcal{G}_{P,Q}^{A,B} = \bigcup_{a \in A, b \in B} \mathcal{G}_{P,Q}^{a,b}.$$

The test function  $\tilde{h}$  is defined on  $\mathcal{G}_{P,Q}^{A,B}$  by

$$\tilde{h}(\sigma) = \frac{1}{\mathfrak{c}} \left[ \frac{M - \mathfrak{m}_K - |P| - (1 - \tilde{h}^{2D}(\sigma^{(m)}))}{M - 2\mathfrak{m}_K} \mathfrak{b} + \epsilon_B \right] ; \sigma \in \mathcal{G}_{P,Q}^{A,B}, \tag{10.3}$$

where  $\{m\} = Q \setminus P$  so that  $\sigma^{(m)}$  is a 2D gateway configuration between  $\mathfrak{s}_a^{2D}$  and  $\mathfrak{s}_b^{2D}$  for some  $(a, b) \in A \times B$ . Since  $\mathcal{G}_{[\mathfrak{m}_K, M - \mathfrak{m}_K - 1]}^{A,B}$  can be decomposed into

$$\mathcal{G}_{[\mathfrak{m}_K, M - \mathfrak{m}_K - 1]}^{A,B} = \bigcup_{i=\mathfrak{m}_K}^{M - \mathfrak{m}_K - 1} \bigcup_{P, Q \in \mathfrak{S}_M : P \prec Q \text{ and } |P|=i} \mathcal{G}_{P,Q}^{A,B},$$

we can combine the constructions (10.3) to define the test function on  $\mathcal{G}_{[\mathfrak{m}_K, M - \mathfrak{m}_K - 1]}^{A,B}$ .

- Construction on  $\widehat{\mathcal{R}}_i^{A,B}$  for  $i \in [\mathfrak{m}_K, M - \mathfrak{m}_K]$ : We set

$$\tilde{h}(\sigma) = \frac{1}{\mathfrak{c}} \left[ \frac{M - \mathfrak{m}_K - i}{M - 2\mathfrak{m}_K} \mathfrak{b} + \epsilon_B \right] ; \sigma \in \widehat{\mathcal{R}}_i^{A,B}, \tag{10.4}$$

so that the function  $\tilde{h}$  is constant on each  $\widehat{\mathcal{R}}_i^{A,B}$ ,  $i \in [\mathfrak{m}_K, M - \mathfrak{m}_K]$ .

3. Construction of  $\tilde{h}$  on the remainder set  $\mathcal{X} \setminus (\mathcal{E}^{A,B} \cup \mathcal{B}^{A,B})$ : We define  $\tilde{h}(\sigma) = 1$  for all  $\sigma \in \mathcal{X} \setminus (\mathcal{E}^{A,B} \cup \mathcal{B}^{A,B})$ .

**Remark 10.3.** From the definition above, we can readily observe the following properties of the test function  $\tilde{h}$ .

1. In view of part (1) of Proposition 9.6, we should check that the definitions of  $\tilde{h}$  on  $\mathcal{E}^{A,B}$  and  $\mathcal{B}^{A,B}$  agree on  $\widehat{\mathcal{R}}_{m_K}^{A,B}$  and  $\widehat{\mathcal{R}}_{M-m_K}^{A,B}$ . This can be verified from (10.1), (10.2), (10.4), and Proposition 9.10. In particular, both definitions imply that the value of  $\tilde{h}$  on the former set is constantly  $\frac{b+\epsilon_B}{c}$ , while the value of  $\tilde{h}$  on the latter set is constantly  $\frac{\epsilon_B}{c}$ .
2. It is obvious that  $\tilde{h} \equiv 1$  on  $\mathcal{S}(A)$  and  $\tilde{h} \equiv 0$  on  $\mathcal{S}(B)$ , and moreover we can readily verify from the definition that  $0 \leq \tilde{h} \leq 1$ .

The remainder of this section is devoted to proving parts (1) and (2) of Proposition 3.2. In the remainder of the current section, we assume for simplicity that  $K < L < M$ . The other cases,  $K = L < M$ ,  $K < L = M$ , or  $K = L = M$ , can be handled in the exact same manner.

### 10.2 Dirichlet form of test function

We first prove that the test function  $\tilde{h}$  satisfies property (2) of Proposition 3.2.

*Proof of part (2) of Proposition 3.2.* We divide the Dirichlet form into three parts as

$$\left[ \sum_{\{\sigma, \zeta\} \subseteq \mathcal{E}^{A,B} \cup \mathcal{B}^{A,B}} + \sum_{\substack{\sigma \in \mathcal{E}^{A,B} \cup \mathcal{B}^{A,B} \\ \zeta \in (\mathcal{E}^{A,B} \cup \mathcal{B}^{A,B})^c}} + \sum_{\{\sigma, \zeta\} \subseteq (\mathcal{E}^{A,B} \cup \mathcal{B}^{A,B})^c} \right] \mu_\beta(\sigma) r_\beta(\sigma, \zeta) \{\tilde{h}(\zeta) - \tilde{h}(\sigma)\}^2.$$

We first consider the second summation. Observe first that, by part (2) of Proposition 9.6, we have  $\mathcal{E}^{A,B} \cup \mathcal{B}^{A,B} = \widehat{\mathcal{N}}(\mathcal{S})$  and thus we get  $H(\zeta) \geq \Gamma + 1$  if  $\sigma \sim \zeta$ . Hence, by (2.6) and Theorem 2.2, we get

$$\mu_\beta(\sigma) r_\beta(\sigma, \zeta) = \min \{ \mu_\beta(\sigma), \mu_\beta(\zeta) \} = \mu_\beta(\zeta) \leq C e^{-(\Gamma+1)\beta}.$$

From the fact that  $0 \leq \tilde{h} \leq 1$  (cf. part (2) of Remark 10.3), we can conclude that the second summation is  $o_\beta(1) e^{-\Gamma\beta}$ . The third summation is trivially 0 by the definition of the test function on  $(\mathcal{E}^{A,B} \cup \mathcal{B}^{A,B})^c$ . Therefore, it remains to show that

$$\sum_{\{\sigma, \zeta\} \subseteq \mathcal{E}^{A,B} \cup \mathcal{B}^{A,B}} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) \{\tilde{h}(\zeta) - \tilde{h}(\sigma)\}^2 = \frac{1 + o_\beta(1)}{cq} e^{-\Gamma\beta}. \quad (10.5)$$

By part (1) of Proposition 9.6 and the fact that  $\tilde{h}$  is constant on each  $\widehat{\mathcal{R}}_i^{A,B}$ ,  $i \in \llbracket m_K, M - m_K \rrbracket$  (cf. (10.4)), we can decompose the left-hand side into

$$\left[ \sum_{\{\sigma, \zeta\} \subseteq \mathcal{B}^{A,B}} + \sum_{\{\sigma, \zeta\} \subseteq \mathcal{E}^A} + \sum_{\{\sigma, \zeta\} \subseteq \mathcal{E}^B} \right] \mu_\beta(\sigma) r_\beta(\sigma, \zeta) \{\tilde{h}(\zeta) - \tilde{h}(\sigma)\}^2. \quad (10.6)$$

Again by the fact that  $\tilde{h}$  is constant on each  $\widehat{\mathcal{R}}_i^{A,B}$ , we can express the first summation as

$$\sum_{i=m_K}^{M-m_K-1} \sum_{a \in A, b \in B} \sum_{\substack{P, Q \in \mathfrak{S}_M: \\ P \prec Q, |P|=i}} \sum_{\substack{\{\sigma, \zeta\} \subseteq \mathcal{B}^{a,b}: \\ \{\sigma, \zeta\} \cap \mathcal{G}_{P,Q}^{a,b} \neq \emptyset}} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) \{\tilde{h}(\zeta) - \tilde{h}(\sigma)\}^2. \quad (10.7)$$

By (10.3) and (10.4), Theorem 2.2 and (5.2), we can write the last summation as  $(1+o_\beta(1))$  times

$$\frac{2\mathfrak{b}^2 e^{-2KL\beta}}{q\mathfrak{c}^2(M-2\mathfrak{m}_K)^2} \sum_{\{\sigma, \zeta\} \subseteq \mathcal{B}^{\mathfrak{a}, \mathfrak{b}}: \{\sigma, \zeta\} \cap \mathcal{G}_{P, Q}^{\mathfrak{a}, \mathfrak{b}} \neq \emptyset} \mu_\beta^{2D}(\sigma^{(m)}) r_\beta^{2D}(\sigma^{(m)}, \zeta^{(m)}) \{\tilde{h}^{2D}(\zeta^{(m)}) - \tilde{h}^{2D}(\sigma^{(m)})\}^2, \tag{10.8}$$

where  $\{m\} = Q \setminus P$  and  $\sigma^{(m)}$  and  $\zeta^{(m)}$  are regarded as 2D Ising configurations. By Proposition 5.12, the last summation is  $\frac{1+o_\beta(1)}{2\kappa^{2D}} e^{-\Gamma^{2D}\beta}$ . Therefore, display (10.8) equals

$$\frac{\mathfrak{b}^2 e^{-2KL\beta}}{\mathfrak{c}^2(M-2\mathfrak{m}_K)^2} \times \frac{(1+o_\beta(1)) e^{-\Gamma^{2D}\beta}}{q\kappa^{2D}}. \tag{10.9}$$

Inserting this to (10.7) (and recalling (3.8)), we get

$$\sum_{\{\sigma, \zeta\} \subseteq \mathcal{B}^{A, B}} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) \{\tilde{h}(\zeta) - \tilde{h}(\sigma)\}^2 = \frac{\mathfrak{b} + o_\beta(1)}{\mathfrak{c}^2 q} e^{-\Gamma\beta}. \tag{10.10}$$

Next, we deal with the second and third summations of (10.6). By (10.1) and Proposition 9.12, the second summation equals

$$\frac{e^{-\Gamma\beta}}{q} \sum_{\{\sigma, \zeta\} \subseteq \mathcal{Y}^A} \frac{\mathfrak{e}_A^2 r^A(\sigma, \zeta) \{\mathfrak{h}^A(\zeta) - \mathfrak{h}^A(\sigma)\}^2}{\mathfrak{c}^2} + o_\beta(1) e^{-\Gamma\beta} = \frac{\mathfrak{e}_A + o_\beta(1)}{\mathfrak{c}^2 q} e^{-\Gamma\beta}. \tag{10.11}$$

Similarly, we get

$$\sum_{\{\sigma, \zeta\} \subseteq \mathcal{E}^B} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) \{\tilde{h}(\zeta) - \tilde{h}(\sigma)\}^2 = \frac{\mathfrak{e}_B + o_\beta(1)}{\mathfrak{c}^2 q} e^{-\Gamma\beta}. \tag{10.12}$$

Therefore, by (10.10), (10.11), and (10.12), we can conclude that the left-hand side of (10.5) is equal to

$$(1+o_\beta(1)) \times \frac{\mathfrak{b} + \mathfrak{e}_A + \mathfrak{e}_B}{\mathfrak{c}^2 q} e^{-\Gamma\beta} = \frac{1+o_\beta(1)}{q\mathfrak{c}} e^{-\Gamma\beta}.$$

This concludes the proof. □

### 10.3 $H^1$ -approximation

Now it remains to prove that the test function  $\tilde{h}$  satisfies part (1) of Proposition 3.2. We shall carry this out in the current section to conclude the proof of Proposition 3.2.

We abbreviate  $h = h_{\mathcal{S}(A), \mathcal{S}(B)}^\beta$  in the remainder of the section. Then, the next lemma asserts that the equilibrium potential is nearly constant on each  $\mathcal{N}$ -neighborhood. Since this lemma can be proved in the exact same manner as [27, Lemma 7.8], we omit the proof.

**Lemma 10.4.** *For any  $\sigma \in \mathcal{X}$  such that  $H(\sigma) < \Gamma$ , it holds that*

$$\max_{\zeta \in \mathcal{N}(\sigma)} |h(\zeta) - h(\sigma)| = o_\beta(1).$$

Now we proceed to the proof of (3.19). By (3.2), we can write

$$\begin{aligned} D_\beta(h - \tilde{h}) &= \langle h - \tilde{h}, -\mathcal{L}_\beta h + \mathcal{L}_\beta \tilde{h} \rangle_{\mu_\beta} \\ &= D_\beta(h) + D_\beta(\tilde{h}) - \langle h, -\mathcal{L}_\beta \tilde{h} \rangle_{\mu_\beta} - \langle \tilde{h}, -\mathcal{L}_\beta h \rangle_{\mu_\beta}. \end{aligned}$$

Since  $\tilde{h} \equiv h \equiv 1$  on  $\mathcal{S}(A)$ ,  $\tilde{h} \equiv h \equiv 0$  on  $\mathcal{S}(B)$  (cf. Remark 10.3-(2)), and  $\mathcal{L}_\beta h \equiv 0$  on  $\mathcal{X} \setminus \mathcal{S}$  (cf. (3.6)), we have

$$\langle \tilde{h}, -\mathcal{L}_\beta h \rangle_{\mu_\beta} = \sum_{\mathbf{s} \in \mathcal{S}(A)} \tilde{h}(\mathbf{s}) (-\mathcal{L}_\beta h)(\mathbf{s}) \mu_\beta(\mathbf{s}) = \sum_{\mathbf{s} \in \mathcal{S}(A)} h(\mathbf{s}) (-\mathcal{L}_\beta h)(\mathbf{s}) \mu_\beta(\mathbf{s}) = D_\beta(h).$$

By the last two displayed equations, we obtain that

$$D_\beta(h - \tilde{h}) = D_\beta(\tilde{h}) - \sum_{\sigma \in \mathcal{X}} h(\sigma) (-\mathcal{L}_\beta \tilde{h})(\sigma) \mu_\beta(\sigma). \tag{10.13}$$

Therefore, by part (2) of Proposition 3.2 proved in the previous subsection and the definition of  $\mathcal{L}_\beta$  (cf. (3.3)), we are left to prove that

$$\sum_{\sigma \in \mathcal{X}} h(\sigma) \sum_{\zeta \in \mathcal{X}} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) [\tilde{h}(\sigma) - \tilde{h}(\zeta)] = \frac{1 + o_\beta(1)}{q\mathbf{c}} e^{-\Gamma\beta}. \tag{10.14}$$

For simplicity, we define

$$\psi(\sigma) = \sum_{\zeta \in \mathcal{X}} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) [\tilde{h}(\sigma) - \tilde{h}(\zeta)], \tag{10.15}$$

so that we can rewrite our objective (10.14) as

$$\sum_{\sigma \in \mathcal{X}} h(\sigma) \psi(\sigma) = \frac{1 + o_\beta(1)}{q\mathbf{c}} e^{-\Gamma\beta}. \tag{10.16}$$

In summary, it suffices to prove (10.16) to prove that  $\tilde{h}$  satisfies part (1) of Proposition 3.2. The proof of (10.16) is divided into several lemmas. First, we demonstrate that  $\psi(\sigma)$  is negligible if  $\sigma$  is not a typical configuration.

**Lemma 10.5.** *For every  $\sigma \in \mathcal{X} \setminus (\mathcal{E}^{A,B} \cup \mathcal{B}^{A,B})$  (i.e.,  $\sigma \notin \widehat{\mathcal{N}}(\mathcal{S})$  by Proposition 9.6), it holds that  $\psi(\sigma) = o_\beta(e^{-\Gamma\beta})$ .*

*Proof.* Since  $\tilde{h} \equiv 1$  on  $\mathcal{X} \setminus (\mathcal{E}^{A,B} \cup \mathcal{B}^{A,B})$  by part (3) of Definition 10.2, it readily holds that

$$\psi(\sigma) = \sum_{\zeta \in \mathcal{E}^{A,B} \cup \mathcal{B}^{A,B}} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) [\tilde{h}(\sigma) - \tilde{h}(\zeta)].$$

Then, by (2.6) and part (2) of Proposition 9.6, if  $\zeta \in \mathcal{E}^{A,B} \cup \mathcal{B}^{A,B}$  with  $\sigma \sim \zeta$  then  $H(\sigma) \geq \Gamma + 1$ , and thus

$$\mu_\beta(\sigma) r_\beta(\sigma, \zeta) = \mu_\beta(\sigma) = O_\beta(e^{-(\Gamma+1)\beta}).$$

Along with the fact that  $0 \leq \tilde{h} \leq 1$ , we conclude that  $\psi(\sigma) = O_\beta(e^{-(\Gamma+1)\beta}) = o_\beta(e^{-\Gamma\beta})$ .  $\square$

We are left to consider  $\psi(\sigma)$  for  $\sigma \in \mathcal{E}^{A,B} \cup \mathcal{B}^{A,B} = \widehat{\mathcal{N}}(\mathcal{S})$ . To this end, we decompose as  $\psi = \psi_1 + \psi_2$  where

$$\psi_1(\sigma) = \sum_{\zeta \in \mathcal{E}^{A,B} \cup \mathcal{B}^{A,B}} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) [\tilde{h}(\sigma) - \tilde{h}(\zeta)], \tag{10.17}$$

$$\psi_2(\sigma) = \sum_{\zeta \notin \mathcal{E}^{A,B} \cup \mathcal{B}^{A,B}} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) [\tilde{h}(\sigma) - \tilde{h}(\zeta)].$$

In fact, we can show that  $\psi_2(\sigma)$  is negligible.

**Lemma 10.6.** *For  $\sigma \in \mathcal{E}^{A,B} \cup \mathcal{B}^{A,B}$ , we have  $\psi_2(\sigma) = o_\beta(e^{-\Gamma\beta})$ .*

*Proof.* This follows directly by the same argument presented in the proof of Lemma 10.5. □

Now, to estimate  $\psi_1(\sigma)$ , let us first look at the bulk typical configurations that are not the edge typical configurations.

**Lemma 10.7.** *We have that  $\psi_1(\sigma) = o_\beta(e^{-\Gamma\beta})$  for all  $\sigma \in \mathcal{G}_{[\mathfrak{m}_K, M-\mathfrak{m}_K-1]}^{A, B}$ .*

*Proof.* For  $\sigma \in \mathcal{G}_{[\mathfrak{m}_K, M-\mathfrak{m}_K-1]}^{A, B}$ , by definition we can write

$$\begin{aligned} \psi_1(\sigma) &= \sum_{\zeta \in \mathcal{E}^{A, B} \cup \mathcal{B}^{A, B}} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) [\tilde{h}(\sigma) - \tilde{h}(\zeta)] \\ &= \frac{\mathfrak{b}}{\mathfrak{c}(M - 2\mathfrak{m}_K)} \sum_{\zeta \in \mathcal{E}^{A, B} \cup \mathcal{B}^{A, B}} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) [\tilde{h}^{2D}(\sigma^{(m)}) - \tilde{h}^{2D}(\zeta^{(m)})] \end{aligned}$$

for some  $m \in Q \setminus P$  with  $P \prec Q$  (cf. Definition 10.2), where  $\sigma^{(m)}$  and  $\zeta^{(m)}$  are considered as 2D Ising configurations. Then, by Theorem 2.2 and (5.2), the last display equals

$$\frac{2\mathfrak{b}(1 + o_\beta(1))}{q\mathfrak{c}(M - 2\mathfrak{m}_K)} e^{-2KL\beta} \times \sum_{\zeta \in \mathcal{E}^{A, B} \cup \mathcal{B}^{A, B}} \mu_\beta^{2D}(\sigma) r_\beta^{2D}(\sigma, \zeta) [\tilde{h}^{2D}(\sigma^{(m)}) - \tilde{h}^{2D}(\zeta^{(m)})].$$

Since  $\sigma^{(m)}$  is a 2D gateway configuration, by part (1) of Proposition 5.13, the last summation equals  $o_\beta(e^{-\Gamma^{2D}\beta})$ . Therefore, we conclude that

$$\psi_1(\sigma) = \frac{2\mathfrak{b}(1 + o_\beta(1))}{q\mathfrak{c}(M - 2\mathfrak{m}_K)} e^{-2KL\beta} \times o_\beta(e^{-\Gamma^{2D}\beta}) = o_\beta(e^{-\Gamma\beta}).$$

□

**Lemma 10.8.** *For all  $i \in \llbracket \mathfrak{m}_K + 1, M - \mathfrak{m}_K - 1 \rrbracket$ , we have that*

$$\sum_{\sigma \in \widehat{\mathcal{R}}_i^{A, B}} \psi_1(\sigma) = 0.$$

Moreover,  $|\psi_1(\sigma)| \leq C e^{-\beta\Gamma}$  for all  $\sigma \in \widehat{\mathcal{R}}_i^{A, B}$ , for some fixed constant  $C > 0$ .

*Proof.* Recall from the definition that  $\tilde{h}$  is defined as constant on each  $\widehat{\mathcal{R}}_i^{A, B}$ . Thus,  $\psi_1(\sigma) = 0$  for all  $\sigma \in \widehat{\mathcal{R}}_i^{A, B} \setminus \mathcal{N}(\mathcal{R}_i^{A, B})$  and it suffices to show that

$$\sum_{\sigma \in \mathcal{N}(\mathcal{R}_i^{A, B})} \psi_1(\sigma) = o_\beta(e^{-\beta\Gamma}).$$

It remains to prove that for all  $a \in A$ ,  $b \in B$ , and  $P \in \mathfrak{S}_M$  such that  $|P| \in \llbracket \mathfrak{m}_K + 1, M - \mathfrak{m}_K - 1 \rrbracket$ ,

$$\sum_{\sigma \in \mathcal{N}(\sigma_P^{a, b})} \sum_{\zeta \in \mathcal{E}^{A, B} \cup \mathcal{B}^{A, B}} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) [\tilde{h}(\sigma) - \tilde{h}(\zeta)] = o_\beta(e^{-\beta\Gamma}). \tag{10.18}$$

Indeed, the left-hand side can be written as

$$\begin{aligned} &\sum_{Q \in \mathfrak{S}_M: P \prec Q} \sum_{\sigma \in \mathcal{C}_{P, Q}^{a, b} \cap \mathcal{N}(\sigma_P^{a, b}), \zeta \in \mathcal{C}_{P, Q}^{a, b}: \zeta \sim \sigma} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) [\tilde{h}(\sigma) - \tilde{h}(\zeta)] \\ &+ \sum_{Q' \in \mathfrak{S}_M: Q' \prec P} \sum_{\sigma \in \mathcal{C}_{Q', P}^{a, b} \cap \mathcal{N}(\sigma_P^{a, b}), \zeta \in \mathcal{C}_{Q', P}^{a, b}: \zeta \sim \sigma} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) [\tilde{h}(\sigma) - \tilde{h}(\zeta)]. \end{aligned}$$

Since we constructed the test function  $\tilde{h}$  between  $\sigma_P^{a,b}$  and  $\sigma_Q^{a,b}$  ( $P \prec Q$ ) and between  $\sigma_{Q'}^{a,b}$  and  $\sigma_{P'}^{a,b}$  ( $Q' \prec P'$ ) in the same manner, the two summations above cancel out with each other, and thus we obtain (10.18).

Finally, for the last statement of the lemma, it suffices to see that if  $\sigma \in \mathcal{N}(\mathcal{R}_i^{A,B})$  and  $\zeta \notin \mathcal{N}(\mathcal{R}_i^{A,B})$  with  $\sigma \sim \zeta$  then  $H(\zeta) \geq \Gamma$ , and thus

$$\mu_\beta(\sigma) r_\beta(\sigma, \zeta) = \mu_\beta(\zeta) \leq \frac{1}{Z_\beta} e^{-\beta\Gamma} = O_\beta(e^{-\beta\Gamma}),$$

where the last equality holds by Theorem 2.2. This proves the last statement of the lemma since the number of summands in (10.17) does not depend on  $\beta$ , and since we have  $0 \leq \tilde{h} \leq 1$  (cf. Remark 10.3).  $\square$

Next, we turn to the edge typical configurations.

**Lemma 10.9.** *The following statements hold.*

1. If  $\sigma \in \mathcal{O}^A \setminus (\widehat{\mathcal{R}}_{m_K}^{A,B} \cup \mathcal{N}(\mathcal{S}(A)))$ , we have  $\psi_1(\sigma) = o_\beta(e^{-\Gamma\beta})$ .
2. If  $\sigma \in \bar{\mathcal{I}}^A \setminus (\widehat{\mathcal{R}}_{m_K}^{A,B} \cup \mathcal{N}(\mathcal{S}(A)))$ , it holds that

$$\sum_{\zeta \in \mathcal{N}(\sigma)} \psi_1(\zeta) = 0, \tag{10.19}$$

and  $|\psi_1(\zeta)| \leq C e^{-\Gamma\beta}$  for all  $\zeta \in \mathcal{N}(\sigma)$  where  $C$  is a constant independent of  $\beta$ .

*Proof.* (1) By part (1) of Proposition 9.12 and the definition of  $\tilde{h}$ , we calculate

$$\begin{aligned} \psi_1(\sigma) &= \sum_{\zeta \in \mathcal{E}^A} \frac{\epsilon_A}{q\mathbf{c}} e^{-\Gamma\beta} r^A(\sigma, \Pi^A(\zeta)) [\mathfrak{h}^A(\sigma) - \mathfrak{h}^A(\Pi^A(\zeta))] + O_\beta(e^{-(\Gamma+1)\beta}) \\ &= \frac{\epsilon_A}{q\mathbf{c}} e^{-\Gamma\beta} \cdot |\mathcal{Y}^A| \cdot (-L^A \mathfrak{h}^A)(\sigma) + O_\beta(e^{-(\Gamma+1)\beta}). \end{aligned}$$

Since  $L^A \mathfrak{h}^A = 0$  on  $\mathcal{O}^A \setminus (\widehat{\mathcal{R}}_{m_K}^{A,B} \cup \mathcal{N}(\mathcal{S}(A)))$  by the elementary property of equilibrium potentials (cf. (3.6)), we may conclude that  $\psi_1(\sigma) = O_\beta(e^{-(\Gamma+1)\beta}) = o_\beta(e^{-\Gamma\beta})$ .

(2) First, we prove (10.19). Note that  $\tilde{h}$  is constant on  $\mathcal{N}(\sigma)$ . Thus,

$$\sum_{\zeta \in \mathcal{N}(\sigma)} \psi_1(\zeta) = \sum_{\zeta \in \mathcal{N}(\sigma)} \sum_{\zeta' \in \mathcal{O}^A} \mu_\beta(\zeta) r_\beta(\zeta, \zeta') [\tilde{h}(\zeta) - \tilde{h}(\zeta')].$$

By part (2) of Proposition 9.12 and the definition of  $\tilde{h}$ , this is equal to

$$\begin{aligned} &\sum_{\zeta' \in \mathcal{O}^A} \frac{\epsilon_A}{q\mathbf{c}} e^{-\Gamma\beta} \cdot r^A(\sigma, \zeta') [\mathfrak{h}^A(\sigma) - \mathfrak{h}^A(\zeta')] + O_\beta(e^{-(\Gamma+1)\beta}) \\ &= \frac{\epsilon_A}{q\mathbf{c}} e^{-\Gamma\beta} \cdot |\mathcal{Y}^A| \cdot (-L^A \mathfrak{h}^A)(\sigma) + o_\beta(e^{-\Gamma\beta}). \end{aligned}$$

Since  $L^A \mathfrak{h}^A = 0$  on  $\bar{\mathcal{I}}^A \setminus (\widehat{\mathcal{R}}_{m_K}^{A,B} \cup \mathcal{N}(\mathcal{S}(A)))$ , we conclude that  $\sum_{\zeta \in \mathcal{N}(\sigma)} \psi_1(\zeta) = o_\beta(e^{-\Gamma\beta})$  and (10.19) is now proved.

Finally, for the last statement, the last display implies that for all  $\zeta \in \mathcal{N}(\sigma)$ ,

$$\begin{aligned} |\psi_1(\zeta)| &= \left| \sum_{\zeta' \in \mathcal{O}^A} \frac{\epsilon_A}{q\mathbf{c}} e^{-\Gamma\beta} \cdot r^A(\sigma, \zeta') [\mathfrak{h}^A(\sigma) - \mathfrak{h}^A(\zeta')] \right| + O_\beta(e^{-(\Gamma+1)\beta}) \\ &\leq \sum_{\zeta' \in \mathcal{O}^A} \frac{\epsilon_A}{q\mathbf{c}} r^A(\sigma, \zeta') e^{-\Gamma\beta} + O_\beta(e^{-(\Gamma+1)\beta}) \leq C e^{-\Gamma\beta}, \end{aligned}$$

where the first inequality holds since  $0 \leq \mathfrak{h}^A \leq 1$  and the second inequality holds by Proposition 9.9, (9.10), and the fact that the number of such  $\zeta' \in \mathcal{O}^A$  with  $\sigma \sim \zeta'$  does not depend on  $\beta$ . This concludes the proof.  $\square$

**Lemma 10.10.** *It holds that*

$$\sum_{\sigma \in \widehat{\mathcal{R}}_{\mathfrak{m}_K}^{A,B}} \psi_1(\sigma) = o_\beta(e^{-\Gamma\beta}),$$

and that  $|\psi_1(\sigma)| \leq Ce^{-\Gamma\beta}$  for all  $\sigma \in \widehat{\mathcal{R}}_{\mathfrak{m}_K}^{A,B}$  where  $C$  is a constant independent of  $\beta$ .

*Proof.* First, we consider the first statement. Proposition 9.10 and the definition of  $\tilde{h}$  on  $\widehat{\mathcal{R}}_{\mathfrak{m}_K}^{A,B}$  imply that  $\psi_1(\sigma) = 0$  for all  $\sigma \in \widehat{\mathcal{R}}_{\mathfrak{m}_K}^{A,B} \setminus \mathcal{N}(\mathcal{R}_{\mathfrak{m}_K}^{A,B})$ . Hence, it suffices to prove that

$$\sum_{\sigma \in \mathcal{N}(\mathcal{R}_{\mathfrak{m}_K}^{A,B})} \psi_1(\sigma) = o_\beta(e^{-\Gamma\beta}). \tag{10.20}$$

Since  $\tilde{h}$  is constant on  $\mathcal{N}(\mathcal{R}_{\mathfrak{m}_K}^{A,B})$ , the left-hand side can be decomposed into

$$\left[ \sum_{\sigma \in \mathcal{N}(\mathcal{R}_{\mathfrak{m}_K}^{A,B}), \zeta \in \mathcal{E}^A} + \sum_{\sigma \in \mathcal{N}(\mathcal{R}_{\mathfrak{m}_K}^{A,B}), \zeta \in \mathcal{B}^{A,B}} \right] \mu_\beta(\sigma) r_\beta(\sigma, \zeta) [\tilde{h}(\sigma) - \tilde{h}(\zeta)]. \tag{10.21}$$

Let us analyze the first summation of (10.21). By part (2) of Proposition 9.12, this equals

$$\sum_{\sigma \in \mathcal{R}_{\mathfrak{m}_K}^{A,B}} \sum_{\zeta \in \mathcal{O}^A} \frac{\mathfrak{c}^A}{q\mathfrak{c}} e^{-\Gamma\beta} \times r^A(\sigma, \zeta) [\mathfrak{h}^A(\sigma) - \mathfrak{h}^A(\zeta)] + O_\beta(e^{-(\Gamma+1)\beta}).$$

By the property of capacities (e.g., [13, (7.1.39)]) and Proposition 9.10, we have

$$\mathfrak{c}_A^{-1} = |\mathcal{V}^A| \text{cap}^A(\mathcal{S}(A), \mathcal{R}_{\mathfrak{m}_K}^{A,B}) = - \sum_{\sigma \in \mathcal{R}_{\mathfrak{m}_K}^{A,B}} \sum_{\zeta: \{\sigma, \zeta\} \in \mathcal{E}^A} r^A(\sigma, \zeta) \{\mathfrak{h}^A(\sigma) - \mathfrak{h}^A(\zeta)\}. \tag{10.22}$$

Summing up, we obtain

$$\sum_{\sigma \in \mathcal{N}(\mathcal{R}_{\mathfrak{m}_K}^{A,B}), \zeta \in \mathcal{E}^A} \psi_1(\sigma) = -\frac{1}{q\mathfrak{c}} e^{-\Gamma\beta} + o_\beta(e^{-\Gamma\beta}). \tag{10.23}$$

Next, we analyze the second summation of (10.21).

$$\sum_{a \in A, b \in B} \sum_{P, Q \in \mathfrak{S}_M: P \prec Q, |P| = \mathfrak{m}_K} \sum_{\sigma \in \mathcal{N}(\sigma_{P,Q}^{a,b}), \zeta \in \mathcal{B}^{a,b}} \mu_\beta(\sigma) r_\beta(\sigma, \zeta) [\tilde{h}(\sigma) - \tilde{h}(\zeta)]. \tag{10.24}$$

By Theorem 2.2, (5.2), and part (2) of Proposition 5.13, this becomes (recall the 2D constant  $\kappa^{2D}$  from (5.11))

$$|A||B| \times 2M \times \frac{2\mathfrak{b}(1 + o_\beta(1))}{q\mathfrak{c}(M - 2\mathfrak{m}_K)} e^{-2KL\beta} \times \frac{1}{2\kappa^{2D}} e^{-\Gamma^{2D}\beta} = \frac{1 + o_\beta(1)}{q\mathfrak{c}} e^{-\Gamma\beta}, \tag{10.25}$$

where the identity follows from the definition of  $\mathfrak{b}$  in (3.8). Combining this with (10.21) and (10.23), we can prove the first statement of the lemma.

For the second statement, from the discussion before (10.20) it is inferred that we only need to prove for  $\sigma \in \mathcal{N}(\mathcal{R}_{\mathfrak{m}_K}^{A,B})$ . For such  $\sigma \in \mathcal{N}(\mathcal{R}_{\mathfrak{m}_K}^{A,B})$ , the previous proof implies that

$$\psi_1(\sigma) = \left[ \sum_{\zeta \in \mathcal{E}^A} + \sum_{\zeta \in \mathcal{B}^{A,B}} \right] \mu_\beta(\sigma) r_\beta(\sigma, \zeta) [\mathfrak{h}^A(\sigma) - \mathfrak{h}^A(\zeta)] + O_\beta(e^{-(\Gamma+1)\beta}),$$

where we used the fact that  $0 \leq \tilde{h} \leq 1$ . By (10.22) and Proposition 9.9, the first summation in the right-hand side is bounded by  $Ce^{-\Gamma\beta}$ . By (10.24) and (10.25), the second summation in the right-hand side is also bounded by  $Ce^{-\Gamma\beta}$ . Therefore, we conclude the proof of the second statement.  $\square$

**Lemma 10.11.** *It holds that*

$$\sum_{\sigma \in \mathcal{N}(\mathcal{S}(A))} \psi_1(\sigma) = \frac{1 + o_\beta(1)}{qc} e^{-\Gamma\beta} \quad \text{and} \quad \sum_{\sigma \in \mathcal{N}(\mathcal{S}(B))} \psi_1(\sigma) = o_\beta(e^{-\Gamma\beta}).$$

Moreover, it holds that  $|\psi_1(\sigma)| \leq Ce^{-\Gamma\beta}$  for all  $\sigma \in \mathcal{N}(\mathcal{S}(A)) \cup \mathcal{N}(\mathcal{S}(B))$  where  $C$  is a constant independent of  $\beta$ .

*Proof.* We concentrate on the claim for  $\mathcal{N}(\mathcal{S}(A))$ , since the corresponding claim for  $\mathcal{N}(\mathcal{S}(B))$  can be proved in the exact same way.

By the property of capacities (e.g., [13, (7.1.39)]) as above, we can write that

$$\epsilon_A^{-1} = |\mathcal{Y}^A| \text{cap}^A(\mathcal{S}(A), \mathcal{R}_{m_K}^{A,B}) = \sum_{\sigma \in \mathcal{S}(A)} \sum_{\zeta: \{\sigma, \zeta\} \in \mathcal{E}^A} r^A(\sigma, \zeta) \{\mathfrak{h}^A(\sigma) - \mathfrak{h}^A(\zeta)\}.$$

Therefore, by the definition of  $\tilde{h}$  and part (2) of Proposition 9.12,  $\sum_{\sigma \in \mathcal{N}(\mathcal{S}(A))} \psi_1(\sigma)$  equals

$$\sum_{\sigma \in \mathcal{S}(A)} \sum_{\zeta: \{\sigma, \zeta\} \in \mathcal{E}^A} \frac{\epsilon_A}{qc} e^{-\Gamma\beta} \cdot r^A(\sigma, \zeta) [\mathfrak{h}^A(\sigma) - \mathfrak{h}^A(\zeta)] + O_\beta(e^{-(\Gamma+1)\beta}) = \frac{1}{qc} e^{-\Gamma\beta} + o_\beta(e^{-\Gamma\beta}).$$

This proves the first statement. As before, the fact that  $|\psi_1| \leq Ce^{-\Gamma\beta}$  on  $\mathcal{N}(\mathcal{S}(A))$  is straightforward from the observations made in the proof.  $\square$

Finally, we present a proof of Proposition 3.2 by combining all computations above.

*Proof of Proposition 3.2.* It remains to prove that  $\tilde{h}$  satisfies part (1) since we already verified in the previous subsection that it satisfies part (2). By the discussion at the beginning of the subsection, it suffices to prove (10.16). By the definition of  $\psi$  given in (10.15) and the series of Lemmas 10.5-10.11, and the fact that  $0 \leq h \leq 1$ , we have

$$\sum_{\sigma \in \mathcal{X}} h(\sigma) \psi(\sigma) = \sum_{\sigma \in \mathcal{N}(\mathcal{S}(A))} h(\sigma) \psi_1(\sigma) + o_\beta(e^{-\Gamma\beta}) = \sum_{\sigma \in \mathcal{N}(\mathcal{S}(A))} \psi_1(\sigma) + o_\beta(e^{-\Gamma\beta}),$$

where the second identity follows from Lemmas 10.4. Thus, by applying Lemma 10.11, we can complete the proof of (10.16).  $\square$

## 11 Remarks on open boundary condition

Thus far, we have only considered the models under periodic boundary conditions. In this section, we consider the same models under open boundary conditions. The proofs for the open boundary case differ slightly to those of the periodic case; however, the fundamentals of the proofs are essentially identical. Hence, we do not repeat the detail but focus solely on the technical points producing the different forms of the main results.

### Energy Barrier

We start by explaining that for the open boundary case, the energy barrier is given by

$$\Gamma = KL + K + 1. \tag{11.1}$$

One can observe that the canonical path explained in Figure 6 becomes an optimal path (note that we should start from a corner of box in this case) with height  $KL + K + 1$  between ground states. This proves that the energy barrier  $\Gamma$  is at most  $KL + K + 1$ . Hence, it remains to prove the corresponding lower bound, i.e., of the fact that  $\Gamma \geq KL + K + 1$ . Rigorous proof of this has been developed in [42] for the 2D model, and the same argument also applies to the 3D model as well using the arguments given in Section 8.

**Sub-exponential prefactor**

As mentioned earlier, the large deviation-type results (Theorems 2.7 and 2.9) hold under open boundary conditions without modification, except for the value of  $\Gamma$ . On the other hand, for the precise estimates (Theorems 2.10 and 2.12), the prefactor  $\kappa$  must be appropriately modified.

For simplicity, we assume that  $q = 2$  and analyze the transition from  $s_1$  to  $s_2$ . To heuristically investigate the speed of this transition in the open boundary case via a comparison to the periodic one, it suffices to check the bulk part of the transition, because the edge part is negligible (as  $K \rightarrow \infty$ ) as in the periodic boundary case. The bulk transition must start from a configuration filled with  $m_K$  lines of spin 2 at either the bottom or top of the lattice box  $\Lambda$ . In the periodic case, there are  $M$  choices for these starting clusters (of spins 2) of size  $KL \times m_K$ ; thus, we can observe that the speed of the transition is slowed by a factor of  $M/2$  under this restriction. Now, let us suppose that we are at a configuration such that several floors of spin 2 are located at the bottom of the lattice, as in Figure 6. When we expand this cluster of spin 2 in the periodic case, there are 2 (namely, up and down) possible choices for the next floor to be filled; on the other hand, there is only one (namely, up) possible choice in the open boundary case. This further slows down the transition by a factor of 2. Next, when we expand the floor at the top of the cluster of spin 2, we may again look at the bulk part of the spin updates (cf. Definition 5.8). Thus, we suppose that there are two lines filled with spin 2 on that floor. There are  $L$  possible choices of the location in the periodic case, but just two possible choices in the open case. Thus, this gives us a factor of  $L/2$ . Moreover, we may choose one of two directions of growth of lines in the periodic case, which gives us additional factor of 2. Finally, there are  $K$  possible ways to form a protuberance in the periodic case; however, we now have only two (at the corners) possible choices. This further slows down the transition by a factor of  $K/2$ . Once the protuberance has been formed, we have only one direction in which to expand it, whereas we have two directions in the periodic case. This slows down the transition by a factor of 2. Summing up, the transition on the bulk is slowed by a factor of

$$\frac{M}{2} \times 2 \times \frac{L}{2} \times 2 \times \frac{K}{2} \times 2 = KLM .$$

Turning this into a rigorous argument (via the same logic applied to the periodic case), we obtain the following Eyring–Kramers law with a modified (compared to the periodic case) prefactor. Recall that we assumed  $K \leq L \leq M$ .

**Theorem 11.1.** *Suppose that we impose open boundary conditions on the model. Then, there exists a constant  $\kappa' = \kappa'(K, L, M) > 0$  such that, for all  $s, s' \in S$ ,*

$$\mathbb{E}_s^\beta [\tau_{\bar{s}}] = (1 + o_\beta(1)) \frac{\kappa'}{q - 1} e^{\Gamma\beta} \quad \text{and} \quad \mathbb{E}_s^\beta [\tau_{s'}] = (1 + o_\beta(1)) \kappa' e^{\Gamma\beta} .$$

Moreover, the constant  $\kappa'$  satisfies

$$\lim_{K \rightarrow \infty} KLM \cdot \kappa'(K, L, M) = \begin{cases} 1/8 & \text{if } K < L < M , \\ 1/16 & \text{if } K = L < M \text{ or } K < L = M , \\ 1/48 & \text{if } K = L = M . \end{cases} \quad (11.2)$$

The constant  $\kappa'$  can be defined in terms of new bulk and edge constants  $\mathfrak{b}'(n)$  and  $\mathfrak{e}'(n)$ , in the exact same manner as done in Section 3.1.

Then, Theorem 2.12 also holds for open boundary conditions with modified limiting Markov chain  $X'(\cdot)$  with rate  $r_{X'}(s, s') = (\kappa')^{-1}$  for all  $s, s' \in S$ .

## References

- [1] L. Alonso and R. Cerf. The three dimensional polyominoes of minimal area. *Electronic Journal of Combinatorics*, 3(1):1–39, 1996. MR3445787
- [2] J. Beltrán and C. Landim. Tunneling and metastability of continuous time Markov chains. *Journal of Statistical Physics*, 140(6):1065–1114, 2010. MR2218873
- [3] J. Beltrán and C. Landim. Metastability of reversible condensed zero range processes on a finite set. *Probability Theory and Related Fields*, 152(3–4):781–807, 2012. MR2642889
- [4] J. Beltrán and C. Landim. Tunneling and metastability of continuous time Markov chains II, the nonreversible case. *Journal of Statistical Physics*, 149(4):598–618, 2012. MR2094397
- [5] J. Beltrán and C. Landim. A Martingale approach to metastability. *Probability Theory and Related Fields*, 161(1–2):267–307, 2015. MR2120991
- [6] G. Ben Arous and R. Cerf. Metastability of the three dimensional Ising model on a torus at very low temperatures. *Electronic Journal of Probability*, 1:1–55, 1996. MR1898856
- [7] G. Bet, A. Gallo, and F. R. Nardi. Critical configurations and tube of typical trajectories for the Potts and Ising models with zero external field. *Journal of Statistical Physics*, 184(3):30, 2021. MR0749840
- [8] G. Bet, A. Gallo, and F. R. Nardi. Metastability for the degenerate Potts model with negative external magnetic field under Glauber dynamics. *Journal of Mathematical Physics*, 63(12):123303, 2022. MR3401022
- [9] G. Bet, A. Gallo, and F. R. Nardi. Metastability for the degenerate Potts model with positive external magnetic field under Glauber dynamics. *arXiv:2108.04011*, 2021. MR1386350
- [10] G. Bet, V. Jacquier, and F. R. Nardi. Effect of energy degeneracy on the transition time for a series of metastable states: application to probabilistic cellular automata. *Journal of Statistical Physics*, 184(1):8, 2021. MR1961341
- [11] A. Bianchi, A. Bovier, and D. Ioffe. Sharp asymptotics for metastability in the random field Curie-Weiss model. *Electronic Journal of Probability*, 14:1541–1603, 2009. MR0267221
- [12] A. Bianchi, S. Dommers, and C. Giardinà. Metastability in the reversible inclusion process. *Electronic Journal of Probability*, 22:1–34, 2017. MR4288339
- [13] A. Bovier and F. den Hollander. *Metastability: a potential-theoretic approach*. Grundlehren der Mathematischen Wissenschaften 351, Springer, 2015. MR4313074
- [14] A. Bovier, F. den Hollander, and F. R. Nardi. Sharp asymptotics for Kawasaki dynamics on a finite box with open boundary. *Probability Theory and Related Fields*, 135(2):265–310, 2006. MR4227174
- [15] A. Bovier, F. den Hollander, and C. Spitoni. Homogeneous nucleation for Glauber and Kawasaki dynamics in large volumes and low temperature. *Annals of Probability*, 38(2):661–713, 2010.
- [16] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes I: Sharp asymptotics for capacities and exit times. *Journal of the European Mathematical Society*, 6(4):399–424, 2004. MR3960293
- [17] A. Bovier, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes II: Precise asymptotics for small eigenvalues. *Journal of the European Mathematical Society*, 7(1):69–99, 2005. MR3513256
- [18] A. Bovier and F. Manzo. Metastability in Glauber dynamics in the low-temperature limit: Beyond exponential asymptotics. *Journal of Statistical Physics*, 107(3–4):757–779, 2002. MR3968863
- [19] M. Cassandro, A. Galves, E. Olivieri, and M. E. Vares. Metastable behavior of stochastic dynamics: A pathwise approach. *Journal of Statistical Physics*, 35(5–6):603–634, 1984. MR4515693
- [20] E. N. M. Cirillo, F. R. Nardi, and J. Sohler. Metastability for general dynamics with rare transitions: Escape time and critical configurations. *Journal of Statistical Physics*, 161(2):365–403, 2015.

- [21] E. N. M. Cirillo and E. Olivieri. Metastability and nucleation for the Blume-Capel model. Different mechanisms of transition. *Journal of Statistical Physics*, 83(3–4):473–554, 1996. MR3900816
- [22] F. den Hollander, F. R. Nardi, E. Olivieri, and E. Scoppola. Droplet growth for three-dimensional Kawasaki dynamics. *Probability Theory and Related Fields*, 125(2):153–194, 2003. MR3568163
- [23] M. I. Freidlin and A. D. Wentzell. On small random perturbations of dynamical systems. *Uspekhi Matematicheskikh Nauk*, 25(1):3–55, 1970. English translation, *Russian Mathematical Surveys*, 25(1):1–56, 1970. MR4029142
- [24] S. Kim. Second time scale of the metastability of reversible inclusion processes. *Probability Theory and Related Fields*, 180(3–4):1135–1187, 2021. MR4408002
- [25] S. Kim. Metastability of Blume–Capel model with zero chemical potential and zero external field. *Journal of Statistical Physics*, 184(3):33, 2021. MR4382580
- [26] S. Kim and I. Seo. Condensation and metastable behavior of non-reversible inclusion processes. *Communications in Mathematical Physics*, 382(2):1343–1401, 2021. MR4408505
- [27] S. Kim and I. Seo. Approximation method to metastability: an application to non-reversible, two-dimensional Ising and Potts models without external fields. *arXiv:2212.13746*, 2022. MR4482068
- [28] C. Landim. Metastable Markov chains. *Probability Surveys*, 16:143–227, 2019. MR3726904
- [29] C. Landim and P. Lemire. Metastability of the two-dimensional Blume–Capel model with zero chemical potential and small magnetic field. *Journal of Statistical Physics*, 164(2):346–376, 2016. MR2070109
- [30] C. Landim, P. Lemire, and M. Mourragui. Metastability of the two-dimensional Blume–Capel model with zero chemical potential and small magnetic field on a large torus. *Journal of Statistical Physics*, 175(2):456–494, 2019. MR4013872
- [31] C. Landim, D. Marcondes, and I. Seo. Metastable behavior of weakly mixing Markov chains: The case of reversible, critical zero-range processes. *Annals of Probability*, 51(1):157–227, 2023. MR3441372
- [32] C. Landim, D. Marcondes, and I. Seo. A resolvent approach to metastability. *Journal of the European Mathematical Society (to appear)*, *arXiv:2102.00998*, 2021. MR1101685
- [33] C. Landim, M. Mariani, and I. Seo. Dirichlet’s and Thomson’s principles for non-selfadjoint elliptic operators with application to non-reversible metastable diffusion processes. *Archive for Rational Mechanics and Analysis*, 231(2):887–938, 2019. MR1151800
- [34] C. Landim and I. Seo. Metastability of non-reversible, mean-field Potts model with three spins. *Journal of Statistical Physics*, 165(4):693–726, 2016. MR2123364
- [35] C. Landim and I. Seo. Metastability of one-dimensional, non-reversible diffusions with periodic boundary conditions. *Annales de l’Institut Henri Poincaré Probabilités et Statistiques*, 55(4):1850–1889, 2019.
- [36] Le Peutrec, D.; Michel, L.: Sharp spectral asymptotics for nonreversible metastable diffusion processes. *Probability and Mathematical Physics*. **1**: 3-53. (2020) MR4575020
- [37] J. Lee. Energy landscape and metastability of Curie–Weiss–Potts model. *Journal of Statistical Physics*, 187(1):2, 2022. MR3922538
- [38] J. Lee and I. Seo. Non-reversible metastable diffusions with Gibbs invariant measure I: Eyring–Kramers formula. *Probability Theory and Related Fields*, 182(3–4):849–903, 2022.
- [39] J. Lee and I. Seo. Non-reversible metastable diffusions with Gibbs invariant measure II: Markov chain convergence. *Journal of Statistical Physics*, 189(2):25, 2022.
- [40] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2017.
- [41] F. Manzo, F. R. Nardi, E. Olivieri, and E. Scoppola. On the essential features of metastability: Tunnelling time and critical configurations. *Journal of Statistical Physics*, 115(1–2):591–642, 2004.
- [42] F. R. Nardi and A. Zocca. Tunneling behavior of Ising and Potts models in the low-temperature regime. *Stochastic Processes and their Applications*, 129(11):4556–4575, 2019.

- [43] F. R. Nardi, A. Zocca, and S. C. Borst. Hitting time asymptotics for hard-core interactions on grids. *Journal of Statistical Physics*, 162(2):522–576, 2016.
- [44] E. J. Neves and R. H. Schonmann. Critical droplets and metastability for a Glauber dynamics at very low temperatures. *Communications in Mathematical Physics*, 137(2):209–230, 1991.
- [45] E. J. Neves and R. H. Schonmann. Behavior of droplets for a class of Glauber dynamics at very low temperature. *Probability Theory and Related Fields*, 91(3–4):331–354, 1992.
- [46] E. Olivieri and M. E. Vares. *Large deviations and metastability*. Encyclopedia of Mathematics and Its Applications 100, Cambridge University Press, Cambridge, 2005.
- [47] C. Oh and F. Rezakhanlou. Metastability of zero range processes via Poisson equations. *Unpublished manuscript*.
- [48] F. Rezakhanlou and I. Seo. Scaling limit of small random perturbation of dynamical systems. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 59(2):867–903, 2023.
- [49] I. Seo. Condensation of non-reversible zero-range processes. *Communications in Mathematical Physics*, 366(2):781–839, 2019.

**Acknowledgments.** The authors thank the anonymous referee for his/her valuable comments that helped us strengthen the presentation of the article.