Mixing of the square plaquette model on a critical length scale*

Paul Chleboun†   Aaron Smith‡

Abstract

Plaquette models are short range ferromagnetic spin models that play a key role in the dynamic facilitation approach to the liquid glass transition. In this paper we study the dynamics of the square plaquette model at the smallest of the three critical length scales discovered in [7]. Our main results are estimates of the spectral gap and mixing time for two natural boundary conditions. As a consequence, we observe that these time scales depend heavily on the boundary condition in this scaling regime.

Contents

1 Introduction

In this paper we consider the dynamics of the square plaquette model (SPM) at low temperature. Spin plaquette models were originally associated with glassy behaviour in [13, 20], where it was argued that they have more physically-realistic dynamics and thermodynamic properties than kinetically-constrained models, while remaining mathematically tractable. These models have also recently been generalised to quantum systems, called fracton models, which show extremely similar relaxation behaviour [3, 19]. Plaquette models are defined over an integer lattice $\Lambda \subseteq \mathbb{Z}^d$, and configurations of the plaquette model correspond to ± 1-valued labellings of the lattice Λ. Every configuration of the SPM has an associated energy given by certain short range ferromagnetic interactions that are defined in terms of the *plaquettes*, a collection of subsets of $\mathbb{Z}^2$ denoted by $\mathcal{P}$.

More formally, plaquette models are families of probability distributions with state space $\{-1, +1\}^\Lambda$. Every configuration $\sigma \in \{-1, +1\}^\Lambda$ has an associated energy value, given by the Hamiltonian

$$H(\sigma) = -\frac{1}{2} \sum_{P \in \mathcal{P}} \prod_{x \in P} \sigma_x \,.$$

For any fixed inverse-temperature $\beta > 0$, we then associate to this Hamiltonian the probability distribution π_Λ^β on $\{-1, 1\}^\Lambda$ given by $\pi_\Lambda^\beta(\sigma) \propto \exp(-\beta H_\Lambda(\sigma))$. In the case of the SPM, we always take $\Lambda \subset \mathbb{Z}^2$ and the plaquettes are exactly the collection of unit squares contained in Λ. There are several natural Markov chains associated with these probability distributions, many of which have very similar behaviour. In this paper we study the most popular of these Markov chains, the *continuous-time single-spin Glauber dynamics* (also known as the *Gibbs sampler*). Roughly speaking, this Markov chain evolves from a configuration σ by choosing a random site $x \in \Lambda$ at unit rate and then updating the value $\sigma(x)$ conditional on $\{\sigma(y)\}_{y \in \Lambda \setminus \{x\}}$ according to the measure π_Λ^β.

Despite the relatively simple form of the Hamiltonian, the thermodynamics of these measures is non trivial [7]. Although there is no phase transition in the SPM (i.e. there is a unique infinite volume Gibbs measure for each β), static correlation lengths grow extremely quickly as the temperature tends to zero ($\beta \to \infty$). Furthermore, the ground states (configurations σ that minimise $H(\sigma)$) depend on the boundary conditions, and are often highly degenerate. It turns out that this plays an important role for the dynamics of the process. In this paper we consider the dynamics of the SPM in the low temperature regime, i.e. as $\beta \to \infty$, on the smallest of the critical length scales in [7].

Plaquette spin models, under single spin-flip Glauber dynamics, have recently attracted a great deal of attention in the physics literature in the context of glassy materials (see e.g. [17] and references therein). Understanding the liquid-glass transition and the dynamics of amorphous materials remains a significant challenge in condensed matter physics (for a review see [1]). One particularly successful approach to studying such systems, known as *dynamic facilitation*, supposes that local relaxation events facilitate further relaxation events in neighbouring regions, but in the absence of such events the system is locally unable to change state (transitions are *blocked*). This idea led to the introduction of a class of interacting particle systems called kinetically constrained models (KCMs), which feature trivial stationary measures but complicated dynamics. These models display many of the key features of glassy systems, such as aging [10] and dynamical heterogeneity [4], and have been extremely well studied in both the mathematical and physics literature (see [14] for a review). There are two crucial difficulties in justifying KCMs as models for the liquid glass transition: it is not clear how kinetic constraints emerge from the microscopic dynamics of many-body systems, and, since KCMs have trivial thermodynamics, KCMs cannot account for the growth of static

correlations. It turns out that plaquette spin models address both these issues, and in particular their dynamics are effectively constrained [13, 20].

The dynamics of plaquette spin models has been the focus of several works in the physics literature. Initial simulations clearly indicate the occurrence of glassy dynamics (extremely slow relaxation) at low temperature [16, 20]. The products $\prod_{x \in P} \sigma_x$ of the spins over the individual plaquettes $P \in \mathcal{P}$ play an important role in studying these dynamics. These products are called the *plaquette variables*, and a plaquette variable equal to -1 is said to be a *defect*. In the SPM, we note that the plaquettes can be naturally identified with their lower-left corner, and so the plaquette variables also form a spin system on an integer lattice.

At low temperature, it is natural to describe the dynamics of the SPM via the locations of the defects, which are effectively constrained. Indeed, flipping a single spin will flip the value of all the plaquette variables whose plaquettes contain the corresponding site – in the SPM, this means flipping the four adjoining plaquette variables. If the plaquette variables associated to the four plaquettes containing a given spin are currently all $+1$ (i.e. there are no defects) then the spin flips at an extremely slow rate of $e^{-4\beta}$; if there is currently one defect associated with the spin and the other plaquette variables are positive then the spin flips with rate $e^{-2\beta}$; finally, if there are two or more defects associated with a spin, then it flips with rate 1 (see Figure 1). For this reason, defects are infrequently created, and isolated defects move extremely slowly, while paired defects can move quickly.

Simulations, and heuristic analysis based on this observation, suggest that for the SPM the relaxation time scales like $e^{c\beta}$ (Arrhenius scaling), and that the dynamics are closely related to those of the Frederickson-Andersen KCM for which mixing properties have been well studied (see [23, 24, 2] and references therein). On the other hand, in a related model called the triangular plaquette model the relaxation time is expected to scale like $e^{c\beta^2}$ (super-Arrhenius scaling) [13]. These dynamics are closely related to a particularly KCM known as the East model which has been widely studied (see [5, 11, 12] and references therein). This difference between Arrhenius and super-Arrhenius scaling is fundamental due to the nature of the energy barriers that should be overcome to bring isolated defects together and annihilate them. Despite their importance, as far as we are aware this work represents the first rigorous results related to the dynamics of plaquette spin models.

Our main results are on the dynamics of the SPM in boxes $\Lambda = \{1, 2, \ldots, L\}^2$ with side length L given by what physicists call the *critical length scale* – the correlation length for the product of spin variables in the infinite volume Gibbs measure. In [7], this critical length scale was shown to satisfy $L \approx e^{\frac{\beta}{2}}$ as $\beta \to 0$. Our results show that the relaxation time (inverse spectral gap) and total variation mixing times indeed have Arrhenius scaling on this critical length scale. We also show that, on this length scale, the relaxation time has a dramatic dependence on the boundary conditions, a phenomena that has been previously observed for certain kinetically constrained models [6].

Our first result is that, for the SPM in boxes of side length $L \approx e^{\frac{\beta}{2}}$ with all plus boundary conditions, the relaxation time scales as $e^{3.5\beta}$ up to polynomial factors in β as $\beta \to \infty$. Furthermore, the total variation mixing time in this case is between $e^{3.5\beta}$ and $e^{4\beta}$ (See Theorem 3.1). Our second main result is on the same length scale for periodic boundary conditions. We find that with periodic boundary conditions on the critical length scale the relaxation time and total variation mixing time both scale as $e^{4\beta}$ up to polynomial corrections in β (see Theorem 3.2).

It turns out that these different time scales are caused by the structure of the ground states. In particular, with all plus boundary conditions there is a unique ground state (all sites have spin $+1$), and the low temperature dynamics are dominated by the time

to reach the ground state. On the other hand with periodic boundary conditions there are 2^{2L-1} ground states, where L is the side length of the box, which correspond to flipping all the spins in any set of rows and columns with respect to the all plus state. In this case the dynamics are dominated by an induced random walk on the ground states. It is possible to construct boundary conditions such that there are a small number of very well separated ground states (see Remark 3.5). In this case we conjecture that the relaxation time is at least $e^{4.5\beta}$. In a companion paper [8], we give Arrhenius bounds on the mixing time and spectral gap for all boundary conditions for length $L \gg e^{\frac{\beta}{2}}$.

The main tools used in the proof of the upper bounds will be detailed canonical path bounds using multi-commodity flows [25], combined with the spectral profile method introduced in [15]. Rather surprisingly, it turns out that a lot of effort is required to construct flows which do not have very bad congestion on edges connecting certain low-probability configurations with many defects. The main difficulty, informally sketched in Section 2.5 and near the beginning of Section 5.3, is common to plaquette models and other models for which mixing is greatly facilitated by the exchange of small, short-lived, and rapidly-moving configurations of interacting particles (see Figure 1). In particular we believe that much of the effort in constructing the multi-commodity flows will be useful also for the study of other plaquette models. To prove the upper bounds in the case of periodic boundary conditions we compare the trace of the process on the ground states with the simple random walk on the hypercube, and use several times the result that mixing times are related to the hitting times of large sets [21, 22].

1.1 Guide to paper

We begin by setting basic notation for the paper in Section 2. This section also includes a heuristic description of the dynamics of the SPM near stationarity. This heuristic guides our proof strategy and also suggests the final results, and we suggest that readers fully digest this heuristic before reading the more precise proofs. Section 3 includes a precise statement of our main results. Section 4 gives some important basic results that will be used frequently throughout the paper, including a description of all possible plaquette configurations and a concentration result for the number of defects under the stationary measure. Section 5 is the bulk of the paper. It describes the canonical path method, derives some specific forms of these bounds that will be used in this paper, constructs two families of canonical paths that will be used to analyze the SPM, and gives detailed bounds on the properties of these paths. The start of this section includes a more detailed guide to its contents. Finally, Sections 6 and 7 contain the proofs of our main results: bounds on the mixing and relaxation time for the all-plus and periodic boundary conditions respectively. Both sections rely heavily on the canonical path arguments for their upper bounds, and use ad-hoc constructions of special test functions for their lower bounds.

2 Notation and background

2.1 Basic conventions

We denote the two canonical basis vectors of $\mathbb{Z}^2$ by $e_1 = (1, 0)$ and $e_2 = (0, 1)$. For $x \in \mathbb{Z}^2$ we denote its projection on e_1 and e_2 by x_1 and x_2 respectively. We define the shorthand $[a:b] = \{a, a+1, \ldots, b\}$ when $b - a \in \mathbb{N}$.

Given $\Lambda \subseteq \mathbb{Z}^2$ we will denote by Ω_Λ the state space of the plaquette model, given by $\{-1, 1\}^\Lambda$, endowed with the product topology. We let $\Omega = \Omega_{\mathbb{Z}^2}$. Given $A \subset \Lambda \subseteq \mathbb{Z}^2$ and a configuration $\sigma \in \Omega_\Lambda$ we define $\sigma|_A$ as the restriction of σ to A. We define the *plaquette* at site $x \in \mathbb{Z}^2$ by the set of four sites $B_x = \{x, x+e_1, x+e_2, x+e_1+e_2\}$. For short we also write $\overline{B}_x = B_{x-e_1-e_2} = \{x, x-e_1, x-e_2, x-e_1-e_2\}$.

For functions $f, g : \mathbb{R}^+ \mapsto \mathbb{R}^+$ we write $f = O(g)$ if there exists $0 < C, X < \infty$ so that $f(x) \leqslant C g(x)$ for all $x > X$. We also write $f = o(g)$ if $\lim_{x \to \infty} \frac{f(x)}{g(x)} = 0$, and we write $f = \Omega(g)$ if $g = O(f)$. Finally, we write $f = \Theta(g)$ if both $f = O(g)$ and $g = O(f)$. To save space, we also write $f \lesssim g$ for $f = O(g)$, we write $f \gtrsim g$ for $f = \Omega(g)$, and we write $f \asymp g$ for $f = \Theta(g)$. Similarly, to save space, all inequalities should be understood to hold only for all $\beta > \beta_0$ sufficiently large. For example, we may write $e^\beta \geqslant \beta + 25$ without additional comment. Since all of our results are asymptotic as β goes to infinity, this convention will not cause any difficulties. For any function $f : A \to B$, we denote the image of f by $f(A)$.

Finally, we define two orders on $\mathbb{Z}^2$. For $x \neq y \in \mathbb{Z}^2$, we say that x is less than y in *lexicographic order*[1] if and only if one of the two following conditions hold:

1. $x_2 > y_2$, or
2. $x_2 = y_2$ and $x_1 < y_1$.

Similarly, we say that x is less than y in *anti-lexicographic order* if and only if one of the two following conditions hold:

1. $x_2 > y_2$, or
2. $x_2 = y_2$ and $x_1 > y_1$.

By a small abuse of notation, we say that a set S_1 is less than a set S_2 in lexicographic order if *every* element of S_1 is less than *every* element of S_2.

2.2 Equilibrium Gibbs measures

We will define the finite volume Gibbs measures on $\Lambda \subset \mathbb{Z}^2$ with fixed and periodic boundary conditions. Let $\mathcal{B}(\Lambda) = \{x \in \mathbb{Z}^2 : B_x \cap \Lambda \neq \emptyset\}$ be the set of plaquettes which intersect Λ, indexed by their bottom left site, and let $\mathcal{B}_-(\Lambda) = \{x \in \mathbb{Z}^2 : \overline{B}_x \subset \Lambda\}$ (if Λ is a rectangle, then $\mathcal{B}_-(\Lambda)$ is just Λ without the left most column and bottom row). For a boundary condition $\tau \in \Omega$ we will denote by $\Omega_\Lambda^\tau = \{\sigma \in \Omega : \sigma|_{\Lambda^c} \equiv \tau|_{\Lambda^c}\}$. Finally we denote the external boundary of Λ by $\partial(\Lambda) = \cup_{x \in \mathcal{B}(\Lambda)} B_x \setminus \Lambda$.

For fixed boundary conditions τ, the plaquette variables associated with a spin configuration are defined by the map $p^\tau : \Omega_\Lambda^\tau \to \Omega_{\mathcal{B}(\Lambda)}$ which is given by the formula

$$p_x^\tau(\sigma) = \prod_{y \in B_x} \sigma_y = \sigma_{(x_1,x_2)} \sigma_{(x_1+1,x_2)} \sigma_{(x_1,x_2+1)} \sigma_{(x_1+1,x_2+1)}, \quad \text{for } x \in \mathcal{B}(\Lambda). \qquad (2.1)$$

Similarly, for periodic boundary conditions on a box $\Lambda = [0\!:\!L_1 - 1] \times [0\!:\!L_2 - 1]$, define $p^{\mathrm{per}} : \Omega_\Lambda \to \Omega_\Lambda$ by

$$p_x^{\mathrm{per}}(\sigma) = \prod_{y \in B_x} \sigma_y = \sigma_{(x_1,x_2)} \sigma_{(x_1+1,x_2)} \sigma_{(x_1,x_2+1)} \sigma_{(x_1+1,x_2+1)}, \quad \text{for } x \in \Lambda, \qquad (2.2)$$

where the sums $x_1 + 1$ and $x_2 + 1$ above are taken modulo L_1 and L_2 respectively. We say there is a *defect* in σ at $x \in \mathcal{B}(\Lambda)$ if $p_x^\tau(\sigma) = -1$ (similarly for periodic boundary conditions). By a small abuse of notation, we consider "per" to be a boundary condition.

For $\sigma \in \Omega_\Lambda$ or Ω_Λ^τ let $|\sigma| = |\{x \in \Lambda : \sigma_x = -1\}|$ denote the number of minus spins and $|p^\tau(\sigma)| = |\{x \in \mathcal{B}(\Lambda) : p_x^\tau(\sigma) = -1\}|$ the number of defects (similarly for $|p^{\mathrm{per}}(\sigma)|$). We define a partial order on plaquette variables with respect to defects by

$$p^\tau(\sigma) \leqslant p^\tau(\eta) \iff \{x \in \mathcal{B}(\Lambda) : p_x^\tau(\sigma) = -1\} \subset \{x \in \mathcal{B}(\Lambda) : p_x^\tau(\eta) = -1\}, \qquad (2.3)$$

---

[1] There are several different "lexicographic" orders in the literature. The order in this paper corresponds to the order in which words are read in English if the Cartesian plane is drawn in the usual way.

similarly for periodic boundary conditions.

We define the Hamiltonian $H_\Lambda^\tau : \Omega_\Lambda^\tau \to \mathbb{R}$ with boundary condition τ by

$$H_\Lambda^\tau(\sigma) = -\frac{1}{2} \sum_{x \in \mathcal{B}(\Lambda)} p_x^\tau(\sigma) \,, \tag{2.4}$$

and similarly for periodic boundary condition. The finite volume Gibbs measure on Λ with boundary condition τ is then denoted by π_Λ^τ and given by

$$\pi_\Lambda^\tau(\sigma) = \frac{e^{-\beta H_\Lambda^\tau(\sigma)}}{Z_\Lambda^\tau(\beta)} \,, \tag{2.5}$$

where $Z_\Lambda^\tau(\beta) = \sum_{\sigma \in \Omega_\Lambda^\tau} e^{-\beta H_\Lambda^\tau(\sigma)}$ is the partition function. The analogous formula gives the finite volume Gibbs measure $\pi_\Lambda^{\mathrm{per}}$ for periodic boundary conditions. For brevity, if $\tau \equiv \pm 1$ we will replace τ with $\pm$; for example we write H_Λ^+, π_Λ^+ for plus boundary conditions. Also, where there is no confusion, we denote by $+ \in \Omega$ the configuration of all $+1$ spins. When the boundary conditions and lattice are clear from the context, we may drop the boundary condition superscript and the lattice subscript.

It follows from (2.4), (2.5) and Lemmas 4.1 and 4.2 in Section 4, that the measure induced by π_Λ^τ on the plaquette variables satisfy

$$\pi_\Lambda^\tau(p^\tau(\sigma) = p) \propto e^{-\beta|p|} \mathbb{1}_{p^\tau(\Omega_\Lambda^\tau)}(p), \quad \text{for } p \in \Omega_{\mathcal{B}(\Lambda)} \,. \tag{2.6}$$

Also, since $|p^+(+)| = |p^{\mathrm{per}}(+)| = 0$, we will frequently use

$$\pi_\Lambda^\tau(\sigma) = \pi_\Lambda^\tau(+) e^{-\beta|p^\tau(\sigma)|} \,, \quad \text{for } \sigma \in \Omega_\Lambda^\tau, \quad \text{and } \tau \in \{+, \mathrm{per}\} \,. \tag{2.7}$$

2.3 Finite volume Glauber dynamics

For a set $S \subset \Lambda$ and $\sigma \in \Omega_\Lambda$, denote by σ^S the configuration obtained by flipping all the spins of σ that lie in S. With slight abuse of notation, we define $\sigma^x = \sigma^{\{x\}}$ for $x \in \Lambda$.

Given a finite region Λ and boundary condition τ, we consider the continuous time Markov process determined by the generator

$$\mathcal{L}_\Lambda^\tau f(\sigma) = \sum_{x \in \Lambda} c_\Lambda^\tau(x, \sigma)(f(\sigma^x) - f(\sigma)) = \sum_{x \in \Lambda} c_\Lambda^\tau(x, \sigma) \nabla_x f(\sigma) \,, \tag{2.8}$$

where we define $\nabla_x f(\sigma) = (f(\sigma^x) - f(\sigma))$, and where the Metropolis spin-flip rates $c_\Lambda^\tau(x, \sigma)$ are given by the formula

$$c_\Lambda^\tau(x, \sigma) = \begin{cases} e^{-\beta(H_\Lambda^\tau(\sigma^x) - H_\Lambda^\tau(\sigma))} & \text{if } H_\Lambda^\tau(\sigma^x) > H_\Lambda^\tau(\sigma) \,, \\ 1 & \text{otherwise.} \end{cases} \tag{2.9}$$

With a slight abuse of notation, we denote the elements of the associated transition rate matrix by by $\mathcal{L}_\Lambda^\tau(\sigma, \eta)$, for $\sigma, \eta \in \Omega_\Lambda^\tau$. The process is reversible with respect to the finite volume equilibrium measure π_Λ^τ.

Remark 2.1. All our results hold equally well for the standard heat-bath dynamics, since $\left(1 + e^{\beta(H_\Lambda^\tau(\sigma^x) - H_\Lambda^\tau(\sigma))}\right)^{-1} \asymp \min\{e^{-\beta(H_\Lambda^\tau(\sigma^x) - H_\Lambda^\tau(\sigma))}, 1\}$ for large β.

Since $H_\Lambda^\tau(\sigma)$ only depends on the plaquette variables, the spin dynamics also induce a dynamics on these "defect" variables which is Markov. The generator of the defect dynamics is given by

$$\mathcal{Q}_\Lambda^\tau f(p) = \sum_{x \in \Lambda} k_\Lambda^\tau(x, p) \left(f(p^{\overline{B}_x}) - f(p) \right) \,, \tag{2.10}$$

where we recall

$$p_z^{\overline{B}_x} = \begin{cases} -p_z & \text{if } z \in \overline{B}_x = \{x - e_1 - e_2, x - e_1, x - e_2, x\}, \\ p_z & \text{otherwise.} \end{cases}$$

From (2.9), the transition rates for this process are given by

$$k_\Lambda^\tau(x, p) = \min\left\{\exp\left[\beta\left(|p| - |p^{\overline{B}_x}|\right)\right], 1\right\}. \tag{2.11}$$

2.4 Measures of mixing rate

We set notation for some common notions related to mixing rates. The Dirichlet form associated with $\mathcal{L}_\Lambda^\tau$ is denoted by $\mathcal{D}_\Lambda^\tau(f) = -\pi_\Lambda^\tau(f\mathcal{L}_\Lambda^\tau f)$, and it satisfies the formula

$$\mathcal{D}_\Lambda^\tau(f) = \frac{1}{2}\sum_{\eta \in \Omega}\sum_{x \in \Lambda}\pi_\Lambda^\tau(\eta)c_\Lambda^\tau(x, \eta)\left(\nabla_x f(\eta)\right)^2. \tag{2.12}$$

Define $\mathrm{Var}_\Lambda^\tau(f)$ to be the variance of f with respect to π_Λ^τ.

Definition 2.2 (Relaxation time). *The smallest positive eigenvalue of* $-\mathcal{L}_\Lambda^\tau$ *is called the spectral gap and it is denoted by* $\mathrm{gap}(\mathcal{L}_\Lambda^\tau)$. *It satisfies the Rayleigh-Ritz variational principle*

$$\mathrm{gap}(\mathcal{L}_\Lambda^\tau) := \inf_{\substack{f:\Omega_\Lambda \mapsto \mathbb{R} \\ f \text{ non constant}}}\frac{\mathcal{D}_\Lambda^\tau(f)}{\mathrm{Var}_\Lambda^\tau(f)}. \tag{2.13}$$

The relaxation time $T_{\mathrm{rel}}^\tau(\Lambda)$ *is defined as the inverse of the spectral gap:*

$$T_{\mathrm{rel}}^\tau(\Lambda) = \frac{1}{\mathrm{gap}(\mathcal{L}_\Lambda^\tau)}. \tag{2.14}$$

If $\Lambda = [1:L]^2$ *we simply write* $T_{\mathrm{rel}}^\tau(L)$.

Definition 2.3 (Total Variation Distance). *We denote the total variation distance between two measures* μ, ν *on a common* σ-*algebra* $\mathcal{F}$ *by:*

$$\|\mu - \nu\|_{\mathrm{TV}} = \sup_{A \in \mathcal{F}}|\mu(A) - \nu(A)|.$$

Definition 2.4 (Mixing time). *We denote by*

$$T_{\mathrm{mix}}^\tau(\Lambda) = \inf\{s > 0 \,:\, \max_{\sigma \in \Omega_\Lambda}\|e^{s\,\mathcal{L}_\Lambda^\tau}(\sigma, \cdot) - \pi_\Lambda^\tau(\cdot)\|_{\mathrm{TV}} < 1/4\}$$

the mixing time of $\mathcal{L}_\Lambda^\tau$. *If* $\Lambda = [1, L]^2$ *we simply write* $T_{\mathrm{mix}}^\tau(L)$.

We define the critical scale as the correlation length for the product of spin variables in the infinite volume Gibbs measure, see [7] for further details and other important length scales.

Definition 2.5 (The critical scale). *We define the critical length scale by* $L_c = \lfloor e^{\frac{\beta}{2}}\rfloor$.

2.5 Heuristics for mixing with + boundary

Call a configuration (other than the ground state) *metastable* if no plaquette has more than one defect, and *unstable* otherwise. As the unstable states are short-lived, we concentrate on the transitions between "nearby" low-energy metastable states.

Due to parity constraints (see Lemma 4.1), the lowest-energy metastable configurations have exactly four defects, placed at the vertices of a rectangle. Starting the SPM process in such a "rectangular" configuration, with height and width of the associated
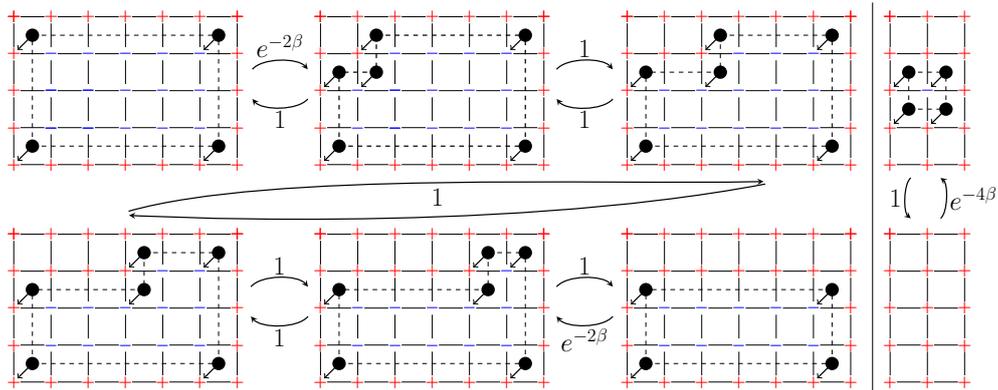
Figure 1: The spin configuration is represented by $+$ (red) and $-$ (blue) on the lattice. The black circles represent defects (-1 plaquette variables), which are associated with the site at the bottom left of the corresponding plaquette. Dashed lines separate regions of $+$ and $-$ spins. The labeled arrows indicate the associated transition rates. An isolated defect creates two defects at rate $e^{-2\beta}$, subsequently a pair of defects is emitted and moves along an edge of the rectangle according to a simple random walk, and is then annihilated upon colliding with another defect (possibly the defect that emitted the pair). Right: the only type of transition not included on the left is to add or remove four neighbouring defects.

rectangle at least four, it is overwhelmingly likely that the next metastable configuration will be another "rectangular" configuration, with either height or width changed by exactly one. The typical intermediate dynamics between such metastable states are shown in Figure 1.

Note that the "rectangular" configurations are entirely determined by the upper-left and lower-right vertices. Following the heuristic of Figure 1, these corner defects perform nearly-independent simple random walks as shown in Figure 2. A routine SRW calculation says that this rectangle will collapse to the unique ground state in a time that scales like $\Theta\left(e^{3.5\beta}\right)$. Since typical configurations have few defects, it is natural to guess that the relaxation time of these simple configurations will be very close to the relaxation time of the full process. This turns out to be correct, and will guide our proof.
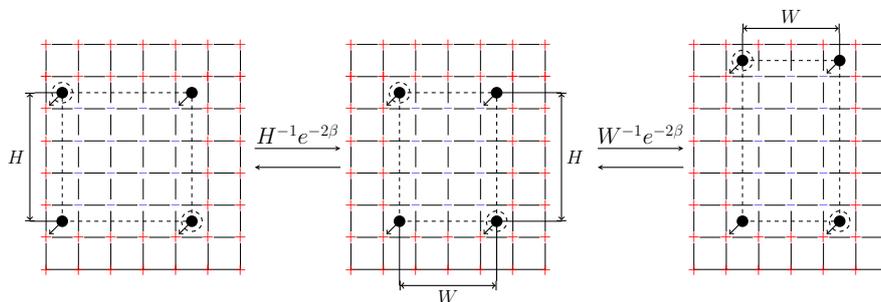

Figure 2: The defects at the corners of a box perform random walks at different rates: they move left-right at rate $H^{-1}e^{-2\beta}$ and up-down at rate $W^{-1}e^{-2\beta}$, where H is the height of the box and W is its width. These random walks are close to independent for β large.

The canonical path method (see Section 5.1) can be used to extend this heuristic argument to configurations with many defects. If we restrict our attention to configurations whose defects form non-overlapping rectangles, the following canonical path construction would again give an $O(e^{3.5\beta})$ bound on the relaxation time:

1. Pick a rectangle uniformly at random.
2. Follow a minimal-length path that collapses this rectangle without ever adding more than 2 defects to the initial configuration, as given by the heuristic in Figures 1 and 2.
3. If any defects remain, go back to step **(1)**.

In other words, if the rectangles don't overlap, we can essentially treat them as evolving independently. Unfortunately, configurations with overlapping rectangles are possible. Allowing for overlaps, this simple path has very high congestion and so gives very poor bounds (see Figure 5). This problem occurs essentially becase the pairs of short-lived particles in Figure 2 can "interact" in this situation, and a similar problem is expected to occur in other plaquette models. In Section 5.3 we will show that it is possible to salvage our heuristic calculations by choosing rectangles according to a much more complicated rule. This rule is governed by a "soft" selection process that focuses on the most "prominent" rectangles. We use this to show that, on average, these interactions between short-lived defect pairs don't contribute significantly to the congestion. We suspect that this basic approach can be used to analyze similar interacting-particle systems (in particular other plaquette models).

3 Main results

We give bounds on the mixing and relaxation time of the square plaquette model on the critical scale:

Theorem 3.1 (Mixing and Relaxation Times for Plus Boundary Conditions)**.** *The square plaquette process with all plus boundary conditions, on the critical scale, satisfies*

$$e^{3.5\beta} \lesssim T_{\mathrm{rel}}^{+}(L_c) \lesssim \beta^6 e^{3.5\beta}. \tag{3.1}$$

Furthermore,

$$e^{3.5\beta} \lesssim T_{\mathrm{mix}}^{+}(L_c) \lesssim \beta^9 e^{4\beta}. \tag{3.2}$$

Theorem 3.2 (Mixing and Relaxation Times for Periodic Boundary Conditions)**.** *The square plaquette process with periodic boundary conditions, on the critical scale, satisfies*

$$e^{4\beta} \lesssim T_{\mathrm{rel}}^{\mathrm{per}}(L_c) \lesssim \beta^4 e^{4\beta}. \tag{3.3}$$

Furthermore,

$$e^{4\beta} \lesssim T_{\mathrm{mix}}^{\mathrm{per}}(L_c) \lesssim \beta^9 e^{4\beta}. \tag{3.4}$$

Remark 3.3. With some extra technical effort in the canonical paths argument, the upper bound on the mixing time with plus boundary conditions can be improved (see [9] where it is shown that $T_{\mathrm{mix}}^{+}(L_c) \lesssim \beta^7 e^{3.75\beta}$). We conjecture that the lower bound given here is correct up to polynomial factors in β.

Remark 3.4. We recall that the fundamental reason for the different scaling of the relaxation and mixing time with all plus and periodic boundary conditions is the structure of the ground states. With all plus boundary conditions there is a unique ground state (in
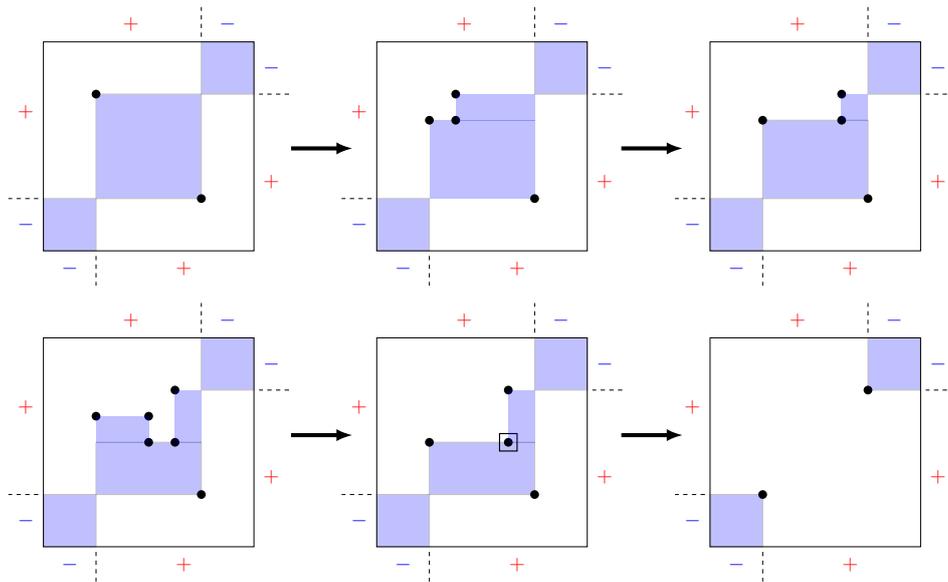
Figure 3: Two well separated ground states (top left and bottom right), with a typical path between them, as described in Remark 3.5. The shaded blue region indicates the $-$ spins, and the empty regions are $+$ spins.

which all sites have spin $+1$), and the low temperature dynamics are dominated by the time to reach the ground state. On the other hand, with periodic boundary conditions, there are 2^{2L-1} ground states, and the low temperature dynamics are dominated by the behaviour of the induced random walk among these ground states. There is a natural bijection between these ground states and the vertices of the $(2L-1)$-dimensional hypercube (see Definition 7.1), and it turns out that the induced random walk is quite closely related to the usual simple random walk on the hypercube $\{-1, +1\}^{2L-1}$ with edge set

$$E = \{(x, y) \in \{-1, +1\}^{2L-1} \times \{-1, +1\}^{2L-1} : \sum_i \mathbf{1}_{x[i] \neq y[i]} = 1 \text{ or } x = -y\}.$$

Remark 3.5. For certain special choices of boundary conditions, there exist a small number of ground states that are well separated, in the sense that it takes an extremely long time to switch between them. In the case of the boundary conditions given in Figure 3 we conjecture

$$e^{4.5\beta} \lesssim T_{\text{rel}}^{\tau_{\text{bad}}}(L_c).$$

The heuristic behind this bound is as follows. Typical paths between the two ground states are of the form described in the figure. Initially two new defects are created, at one of the existing defects, which occurs with rate $e^{-2\beta}$. Subsequently, with probability $1/L$, a pair of defects is not immediately re-absorbed, and travels a distance of order L. At some point during this excursion (top right frame of Figure 3) the process may generate two new defects, with probability $\approx e^{-2\beta}L^2$. Subsequently, with probability $1/L$ again, these defects reach one of the isolated defects before being re-absorbed, leading to the stable configuration in the bottom middle Figure 3. Finally the defect marked with a square box makes a two dimensional random walk following the same mechanism as given in Figure 2. The projection of the position of this defect onto the diagonal (bottom left to top right) is a martingale, and so with probability $O(1/L)$ this defect will travel a

distance $\Omega(L)$ before being re-absorbed in the previous ground state. Putting all these factors together gives the heuristic bound.

4 Preliminary calculations and notation

In this section, we make some simple observations about which configurations are possible and likely at stationarity. We first observe that only defect configurations satisfying certain parity constraints are possible, and these depend on the boundary conditions (Lemmas 4.1, and 4.2). These parity conditions immediately imply some rough bounds on the number of configurations with a given number of defects (Lemma 4.3), which in turn implies that the ground state has high probability under the stationary measure (Lemma 4.4).

In the case $\Lambda = [1 : \ell_1] \times [1 : \ell_2]$, we have $\mathcal{B}(\Lambda) = [0 : \ell_1] \times [0 : \ell_2]$. In this case we denote all the sites belonging to the ith column of $\mathcal{B}(\Lambda)$ by $C_i = \{i\} \times [0 : \ell_2]$, and the jth row by $R_j = [0 : \ell_1] \times \{j\}$. For $x, y \in \mathcal{B}(\Lambda)$ we write $y \ll x$ if $y_1 < x_1$ and $y_2 < x_2$.

Lemma 4.1 (τ-b.c. parity constraints)**.** *If* $\Lambda = [1 : \ell_1] \times [1 : \ell_2]$ *then for boundary condition* $\tau \in \Omega$

$$p^\tau(\Omega_\Lambda^\tau) = \{p \in \Omega_{\mathcal{B}(\Lambda)} \,:\, p \text{ satisfies } (4.1) \text{ and } (4.2) \text{ below}\},$$

where the parity constraints are given by

$$\prod_{x \in C_i} p_x = \tau_{(i,0)} \tau_{(i+1,0)} \tau_{(i,\ell_2+1)} \tau_{(i+1,\ell_2+1)}, \quad \forall i \in [0 : \ell_1] \quad \text{and,} \tag{4.1}$$

$$\prod_{x \in R_j} p_x = \tau_{(0,j)} \tau_{(0,j+1)} \tau_{(\ell_1+1,j)} \tau_{(\ell_1+1,j+1)}, \quad \forall j \in [0 : \ell_2]. \tag{4.2}$$

Furthermore, $p^\tau : \Omega_\Lambda^\tau \to p^\tau(\Omega_\Lambda^\tau)$ *is bijective, and the inverse may be written as*

$$(p^\tau)^{-1}(p)_x = \sigma^\tau(p)_x := \begin{cases} \tau_{(0,0)} \tau_{(x_1,0)} \tau_{(0,x_2)} \prod_{y \ll x} p_y & \text{if } x \in \Lambda, \\ \tau_x & \text{otherwise,} \end{cases} \tag{4.3}$$

for $p \in p^\tau(\Omega_\Lambda^\tau)$.

Lemma 4.2 (Periodic b.c. parity constraints)**.** *If* $\Lambda = [0 : \ell_1] \times [0 : \ell_2]$ *then for periodic boundary conditions*

$$p^{\text{per}}(\Omega_\Lambda) = \{p \in \Omega_\Lambda \,:\, p \text{ satisfies } (4.4)\},$$

where the parity constraints are given by

$$\prod_{x \in C_i} p_x = \prod_{x \in R_j} p_x = 1, \quad \forall i \in [0 : \ell_1 - 1] \quad \text{and} \quad \forall j \in [0 : \ell_2 - 1]. \tag{4.4}$$

and for any fixed $\tau \in \{-1, 1\}^{C_0 \cup R_0}$, *we have* $p^{\text{per}} : \{\sigma \in \Omega_\Lambda \,:\, \sigma_{C_0 \cup R_0} \equiv \tau\} \to p^{\text{per}}(\Omega_\Lambda)$ *is bijective.*

Proof. The proof relies on the following observation; for $\sigma \in \{-1, 1\}^{\Lambda'}$, for some Λ' finite, and $A, B \subset \Lambda'$ we have

$$\prod_{x \in A} \sigma_x \prod_{y \in B} \sigma_y = \prod_{x \in A \triangle B} \sigma_x, \tag{4.5}$$

where $A \triangle B$ denotes the symmetric difference of A and B. It follows that for $\sigma \in \Omega_\Lambda^\tau$ and $i \in [0 : \ell_1]$

$$\prod_{x \in C_i} p_x^\tau(\sigma) = \prod_{x \in C_i} \prod_{y \in B_x} \sigma_y = \tau_{(i,0)} \tau_{(i+1,0)} \tau_{(i,\ell_2+1)} \tau_{(i+1,\ell_2+1)},$$

and similarly for rows, i.e. $p^\tau(\Omega_\Lambda^\tau) \subset \{p \in \Omega_{\mathcal{B}(\Lambda)} : p \text{ satisfies (4.1) and (4.2)}\}$.

To prove the opposite inclusion, fix $\zeta \in \{p \in \Omega_{\mathcal{B}(\Lambda)} : p \text{ satisfies (4.1) and (4.2)}\}$, then $\sigma^\tau(\zeta) \in \Omega_\Lambda^\tau$ by the definition given in (4.3). Furthermore, using (2.1) and (4.5),

$$p^\tau(\sigma^\tau(\zeta))_x = \prod_{y \in B_x} \sigma^\tau(\zeta)_y = \prod_{y \ll x} \zeta_y \prod_{z \ll x+e_1} \zeta_z \prod_{y' \ll x+e_2} \zeta_{y'} \prod_{z' \ll x+e_1+e_2} \zeta_{z'} = \zeta_x$$

This completes the proof of Lemma 4.1, since $|\Omega_\Lambda^\tau| = |\{p \in \Omega_{\mathcal{B}(\Lambda)} : p \text{ satisfies (4.1) and } (4.2)\}| = 2^{|\Lambda|}$. The proof of Lemma 4.2 follows the same arguments. $\square$

These parity considerations imply the following bounds on the number of configurations with a given number of defects:

Lemma 4.3. *Fix $k \in \mathbb{N}$. If $\Lambda = [1 : L]^2$, then with all plus boundary conditions*

$$|\{\sigma \in \Omega_\Lambda^+ : |p^+(\sigma)| = 2k\}| \leqslant \min\{(ek)^{2k}L^{2k}, L^{3k}\}. \tag{4.6}$$

If $\Lambda = [0 : L]^2$, then with periodic boundary conditions

$$|\{\sigma \in \Omega_\Lambda : |p^{\text{per}}(\sigma)| = 2k\}| \leqslant 2^{2L+1}\min\{(ek)^{2k}L^{2k}, L^{3k}\}. \tag{4.7}$$

Proof. We begin with the $+$ boundary conditions. By Lemma 4.1 we have, for each $k \in \mathbb{N}$,

$$|\{\sigma \in \Omega_\Lambda^+ : |p^+(\sigma)| = 2k\}| = |\{p \in p^+(\Omega_\Lambda^+) : |p| = 2k\}|$$

and $|\{p \in p^+(\Omega_\Lambda^+) : |p| = 2k\}|$ is exactly the number of configurations in $\{0,1\}^\Lambda$, where we identify defects with 1's, which have even row and column sums and the total number of defects is $2k$. Since any row which contains at least one defect must also contain at least two, the total number of rows which have at least one defect is clearly bounded above by k (similarly for columns). It follows that the number of configurations which satisfy the row and column parity constraint is at most the number of ways of firstly choosing k rows and k columns (with replacement), and then arranging the $2k$ defects anywhere on the sub-lattice defined by these rows and columns. The number of ways of arranging the former is L^{2k}, and the number of ways of arranging the latter is clearly bounded above by $\binom{k^2}{2k}$. Using $\binom{n}{k} < (e\,n)^k/k^k$ and taking the product gives the first upper bound. The second upper bound follows by using only the row parity, which implies that the defects may be grouped into k disjoint pairs, such that defects belonging to the same pair occupy the same row. There are then at most L^3 ways to arrange each pair on Λ and the upper bound L^{3k} follows immediately.

We next consider periodic boundary conditions. By the above argument, for every choice of $\tau \in \{-1, 1\}^{C_0 \cup R_0}$ we have

$$|\{\sigma \in \Omega_\Lambda : |p^{\text{per}}(\sigma)| = 2k, \sigma|_{C_0 \cup R_0} = \tau\}| \leqslant \min\{(ek)^{2k}L^{2k}, L^{3k}\}.$$

Applying the bijection in Lemma 4.2 and noting that $|\{-1, 1\}^{C_0 \cup R_0}| = 2^{2L+1}$ completes the proof. $\square$

For a given boundary condition τ, define the collection of *ground states* by:

$$\mathcal{G} = \mathcal{G}^\tau = \{\sigma \in \Omega_\Lambda^\tau : H_\Lambda^\tau(\sigma) = \min_{\eta \in \Omega_\Lambda^\tau} H_\Lambda^\tau(\eta)\}. \tag{4.8}$$

Lemma 4.4 (Domination of Ground States). *Let $\tau \in \{+, \text{per}\}$ and $L = L_c$, then*

$$\pi(\mathcal{G}) \geqslant 1 - O(e^{-2\beta}). \tag{4.9}$$

Proof. By Lemmas 4.1 and 4.2, we note that all possible configurations have an even number of defects, and also that no configurations can have exactly 2 defects. When $\tau = +$, there is a unique ground state $+ \equiv (+1, \ldots, +1)$. In this situation, Lemma 4.3 gives

$$1 - \pi(+) = \sum_{k \geqslant 2} \pi(+)|\{\sigma \in \Omega_\Lambda^+ \ : \ |p^+(\sigma)| = 2k\}|e^{-2k\beta} \lesssim \pi(+)e^{-2\beta}.$$

Thus $\pi(+) = 1 - O(e^{-2\beta})$, completing the proof of the lemma for $\tau = +$. For the case of periodic boundary conditions, observe that there are exactly 2^{2L+1} ground states, so $\pi(\mathcal{G}) = 2^{2L+1}\pi(+)$ and the same argument holds. $\qquad \square$

5 Construction and analysis of canonical paths

Our arguments for the upper bounds in Theorems 3.1 and 3.2 will be based on the method of "canonical paths." The idea in this method is to construct a family of (possibly random) paths between any pairs of configurations σ, η, such that the paths do not use any one edge too much. In this section, we construct the canonical paths that will be used for those proofs, and also give some initial analysis of their properties. As a guide to the remainder of this section:

- In Section 5.1, we state generic bounds on relaxation and mixing times of a Markov chain in terms of canonical paths. These are small variants of well-known results, but may be useful in the study of other Markov chains.
- In Section 5.2, we give preliminary notation related to canonical paths.
- In Section 5.3, we construct and analyze the main "building block" of our canonical path construction. This building block determines the entire path for any initial configuration σ with $O(\beta L)$ defects.
- In Section 5.4, we construct a "building block" that may be used when σ has many more than L defects. We also combine our two building blocks to construct the main canonical path used throughout this paper.
- In Section 5.5, we give simple bounds on various quantities related to combining the building blocks into a longer canonical path.
- In Section 5.6, we combine the results in this section to obtain final bounds on the properties of the canonical paths we have defined.

5.1 Canonical path bounds

In this section we describe the use of canonical path arguments to bound the spectral gap and spectral profile of a Markov process on finite state space. We first set some standard notation for this subsection only. For a general Markov chain with transition rate matrix K on state space Ω, denote by $\mathbf{E} = \{(\eta, \sigma) \in \Omega^2 \ : \ K(\eta, \sigma) > 0\}$ the collection of transitions allowed by K. We recall the notion of *paths* in such a Markov chain:

Definition 5.1 (Path). *A sequence $\gamma = (\eta^{(0)}, \eta^{(1)}, \ldots, \eta^{(m)}) \in \Omega$ is called a path from $\eta^{(0)}$ to $\eta^{(m)}$ if $(\eta^{(i-1)}, \eta^{(i)}) \in \mathbf{E}$ for all $1 \leqslant i \leqslant m$. We say that this path has length $|\gamma| \equiv m$. For $1 \leqslant i \leqslant m$, we call the pair $(\eta^{(i-1)}, \eta^{(i)})$ the i'th edge of γ. For $\eta, \sigma \in \Omega$, we denote by $\Gamma_{\eta,\sigma}$ the collection of all paths from η to σ. Similarly we let $\Gamma_\eta = \cup_\sigma \Gamma_{\eta,\sigma}$ be the collection of all paths starting at η and $\Gamma = \cup_{\eta,\sigma \in \Omega_\Lambda} \Gamma_{\eta,\sigma}$ be the collection of all paths.*

For a path $\gamma = (\eta^{(0)}, \ldots, \eta^{(m)})$, we denote by $\gamma^{(\text{init})} = \eta^{(0)}$ and $\gamma^{(\text{fin})} = \eta^{(m)}$ the initial and final elements of γ. If γ_1, γ_2 are two paths with $\gamma_1^{(\text{fin})} = \gamma_2^{(\text{init})}$, we denote by $\gamma_1 \bullet \gamma_2$ the concatenation of γ_1 and γ_2 with the repeated element $\gamma_1^{(fin)}, \gamma_2^{(init)}$ removed. With some abuse of notation, we define $\emptyset \bullet \gamma = \gamma \bullet \emptyset = \gamma$ for any path γ.

When applying the results in this Section, we will always use the process given by (2.8).

For the remainder of this section, assume that K has a unique reversible measure μ and let $\mu_* = \min_{\eta \in \Omega} \mu(\eta)$. For $S \subset \Omega$, define $c_0(S)$ to be the non-constant functions on Ω with support contained in S. For $S \subset \Omega$, also define

$$\lambda(S) = \inf_{f \in c_0(S)} \frac{\mathcal{D}_K(f)}{\operatorname{Var}_\mu(f)}, \tag{5.1}$$

where $\mathcal{D}_K$ is the Dirichlet form associated with K,

$$\mathcal{D}_K = -\mu(f\, Kf) = \frac{1}{2} \sum_{\eta, \sigma \in \Omega} \mu(\eta) K(\eta, \sigma)(f(\eta) - f(\sigma))^2. \tag{5.2}$$

Lemma 5.2 (Multicommodity Flows for Bounding the Spectral Profile)**.** *Let $S \subsetneq \Omega$. For $\eta \in S$, let F_η be a probability measure on paths $\gamma = (\gamma^0, \ldots, \gamma^m)$ in Ω that have starting point $\gamma^0 = \eta$ and final point $\gamma^m \in S^c$. Then*

$$\lambda(S) \geqslant \mathcal{A}^{-1},$$

where the constant

$$\mathcal{A} \equiv 2 \max_{e \in \mathbf{E}} \sum_{\eta \in S} \sum_{\gamma \ni e} F_\eta(\gamma) \, |\gamma| \, \frac{\mu(\eta)}{\mu(e_-) K(e_-, e_+)} \tag{5.3}$$

is the cost of the flow $\{F_\eta\}_{\eta \in S}$ for the set S.

Remark 5.3. We will use this with $S = \Omega_\Lambda \backslash \{+\}$ when bounding the spectral gap of the SPM with + boundary conditions. The more general bound is useful for bounding the spectral profile, which we will need to obtain a more refined mixing time bound.

Proof. We begin by reducing to the case of paths with no repeated edges. Consider a path γ and configuration $\eta \in S$ with $F_\eta(\gamma) > 0$. We note that excising all cycles from γ can only reduce the associated term $|\gamma| \frac{\mu(\eta)}{\mu(e_-) K(e_-, e_+)}$ appearing in (5.3), possibly to zero if the path without cycles does not use the edge e. Thus, we can assume without loss of generality that F_η is supported on paths with no cycles, for all $\eta \in S$.

Next, fix $f \in c_0(S)$ with $\mathbb{E}_\mu[f] = 0$. By the Cauchy-Schwarz inequality and the fact that $f(\eta) = 0$ for all $\eta \in S^c$,

$$\operatorname{Var}_\mu(f) = \sum_{\eta \in \Omega} \mu(\eta) f(\eta)^2 = \sum_{\eta \in S} \mu(\eta) \sum_\gamma F_\eta(\gamma) \Big(\sum_{i=1}^{|\gamma|} (f(\gamma^{i-1}) - f(\gamma^i)) \Big)^2$$

$$\leqslant \sum_{\eta \in S} \sum_\gamma \mu(\eta) F_\eta(\gamma) \, |\gamma| \sum_{i=1}^{|\gamma|} (f(\gamma^{i-1}) - f(\gamma^i))^2$$

$$= \sum_{\eta \in S} \sum_\gamma \mu(\eta) F_\eta(\gamma) \, |\gamma| \sum_{i=1}^{|\gamma|} \frac{\mu(\gamma^{i-1}) K(\gamma^{i-1}, \gamma^i)}{\mu(\gamma^{i-1}) K(\gamma^{i-1}, \gamma^i)} (f(\gamma^{i-1}) - f(\gamma^i))^2.$$

Comparing this to the formula for the Dirichlet form, this gives

$$\frac{\operatorname{Var}_\mu(f)}{\mathcal{D}_K(f)} \leqslant 2 \max_{e \in \mathbf{E}} \sum_{\eta \in S} \sum_{\gamma \ni e} F_\eta(\gamma) \, |\gamma| \, \frac{\mu(\eta)}{\mu(e_-) K(e_-, e_+)},$$

completing the proof of the lemma. $\qquad \square$

This bound is all that we will require to bound the spectral gap of the SPM model with + boundary conditions. In order to bound the mixing time of both chains, as well as the spectral gap of the SPM with periodic boundary conditions, we will use the spectral profile method introduced in [15]. We quickly recall the main results of that paper:

Definition 5.4 (Spectral Profile Bounds.). *For $r > 0$, define*

$$\Lambda(r) = \inf_{\mu_* \leqslant \mu(S) \leqslant r} \lambda(S).$$

Theorem 1.1 of [15] states that the mixing time T_{mix} of K satisfies

$$T_{\mathrm{mix}} \leqslant \int_{4\mu_*}^{16} \frac{2}{x\,\Lambda(x)} dx. \tag{5.4}$$

In our application, the spectral profile $\Lambda(r)$ is defined as an infimum over a very large collection of subsets $S \subset \Omega_\Lambda^\tau$; we will find it easier to work with a much-reduced collection of subsets. For fixed $r > 0$, let

$$k(r) = \min\{k \,:\, \pi(+)\,e^{-\beta k} \leqslant r\}. \tag{5.5}$$

For fixed $k \in \mathbb{N}$, let

$$S_k = \{\sigma \in \Omega_\Lambda^\tau \,:\, |p^\tau(\sigma)| \geqslant k\}. \tag{5.6}$$

We observe that, for fixed $r > 0$ and $S \subset \Omega_\Lambda$ with $\pi(S) \leqslant r$, we have

$$S \subset S_{k(r)}. \tag{5.7}$$

Thus, we have

$$\Lambda(r) = \inf_{\mu_* < \mu(S) < r} \inf_{f \in c_0(S)} \frac{\mathcal{D}_\mathcal{L}(f)}{\mathrm{Var}_\pi(f)} \tag{5.8}$$

$$\geqslant \inf_{f \in c_0(S_{k(r)})} \frac{\mathcal{D}_\mathcal{L}(f)}{\mathrm{Var}_\pi(f)} = \lambda(S_{k(r)}).$$

Thus, to bound the mixing time, it is enough to bound $\inf_{f \in c_0(S)} \frac{\mathcal{D}_\mathcal{L}(f)}{\mathrm{Var}_\pi(f)}$ below, for sets of the form $S = S_k$ and some $k \in \mathbb{N}$.

5.2 Preliminary notation

We will use the following notation for defects frequently:

Definition 5.5 (Neighbouring Defects). *Let $\Lambda = [\ell_1 : \ell_2] \times [\ell_3 : \ell_4]$. Fix $p \in \Omega_{\mathcal{B}(\Lambda)}$ and define*

$$D(p) = \{x \in \Omega_{\mathcal{B}(\Lambda)} \,:\, p_x = -1\}.$$

For $p \in \Omega_{\mathcal{B}(\Lambda)}$ and $x = (x_1, x_2) \in D(p)$, define the left row neighbour, right row neighbour, down column neighbour and up column neighbour of the defect at x by

$$\ell(x) = (\max\{x_1' < x_1 \,:\, (x_1', x_2) \in D(p)\}, x_2) \tag{5.9}$$
$$r(x) = (\min\{x_1' > x_1 \,:\, (x_1', x_2) \in D(p)\}, x_2)$$
$$d(x) = (x_1, \max\{x_2' < x_2 \,:\, (x_1, x_2') \in D(p)\})$$
$$u(x) = (x_1, \min\{x_2' > x_2 \,:\, (x_1, x_2') \in D(p)\})$$

when they exist; otherwise by $\emptyset$ as appropriate. Finally, by a small abuse of notation we define $D(\sigma) = D(p^+(\sigma))$ for $\sigma \in \Omega_\Lambda^+$.
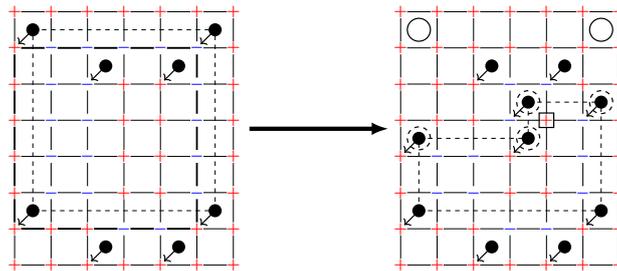
Figure 4: Left: an initial configuration σ, with defects $D(\sigma)$ indicated by black circles. The thick-black box shows the rectangle R, the sites inside the dashed box correspond to $\mathcal{B}_-(R)$ which are flipped. Right: An example of a configuration $\eta \in \gamma_{\sigma,R}$, the next spin to be flipped in the path is marked by a solid box. Open circles denote *absent* defects (with respect to σ), dashed circles denote defect variables which are also flipped with respect to σ. In this example the top row of the rectangle R, bottom row, and the two rows containing defect variables which differ from $D(\sigma)$, are all distinct; it is also possible for the rows to intersect (see Equation (5.20) for an explicit list of these cases - the example above shows a **mid-type edge).**

We now define a simple "base path" that flips all the spins in a rectangle $R \subset \mathbb{Z}^2$ in a sensible order. Typically in application, the rectangle R in the following definition will have at least three defects at its vertices in the initial configuration σ, and this "base path" removes two or four of them without ever adding more than two in intermediate stages:

Definition 5.6 (Rectangle-Removal Path)**.** *Let* $\Lambda = [\ell_1 : \ell_2] \times [\ell_3 : \ell_4]$. *Given* $\sigma \in \Omega_\Lambda$ *and a rectangle* $R = [x_1 : x_2] \times [y_1 : y_2] \subset \mathcal{B}(\Lambda)$ *we define a path* $\gamma_{\sigma,R}$, *starting at* σ, *that flips all the spins at all the sites in* $\mathcal{B}_-(R)$. *We construct the path according to the following cases:*

1. *If* $(x_1, y_1), (x_1, y_2), (x_2, y_2) \in D(\sigma)$, *or fewer than 3 of the four corners of the rectangle are in* $D(\sigma)$, *we define* $\gamma_{\sigma,R}$ *to be the path that flips all spins in* $\mathcal{B}_-(R)$ *in lexicographic order.*

2. $(x_2, y_2), (x_1, y_2), (x_2, y_1) \in D(\sigma)$ *but* $(x_1, y_1) \notin D(\sigma)$, *we define* $\gamma_{\sigma,R}$ *to be the path that that flips all spins in* $\mathcal{B}_-(R)$ *in anti-lexicographic order.*

3. *Otherwise, denote by* $M : \Omega_\Lambda \mapsto \Omega_\Lambda$ *the "mirror reflection" map*

$$\mathrm{M}(\sigma)_{\mathrm{i,j}} = \sigma_{\mathrm{i},\ell_4+\ell_3-\mathrm{j}} \tag{5.10}$$

which flips the lattice in the x-axis. Note that $M(R)$ *is now in one of the two previous cases, and we define* $\gamma_{\sigma,R} = M^{-1}(\gamma_{M(\sigma),M(R)})$.

Although a *spin configuration* $\eta \in \gamma_{\sigma,R}$ can be very different from σ, their associated *defect configurations* $D(\eta), D(\sigma)$ can only differ in a few locations, as illustrated in Figure 4.

Denote by $\mathbf{R}_\Lambda$ the collection of rectangles in the lattice Λ. For $R = [x_1 : x_2] \times [y_1 : y_2] \in \mathbf{R}_\Lambda$, we informally call $y_2 - y_1$ the "height" and $x_2 - x_1$ the "width" of R. Denote by $\mathbf{E} = \{(\sigma, \eta) \in \Omega_\Lambda^2 : \mathcal{L}_\Lambda^+(\sigma, \eta) > 0\}$ the collection of edges associated with the generator $\mathcal{L}_\Lambda^+$. Finally, For any set S, denote by $\mathcal{P}(S)$ the power set of S.

5.3 Path construction: initial segment

In this section, we give a description of the main building block of our canonical path. The basic heuristic for these paths was presented in Section 2.5: pick a random

rectangle with at least three defects in its corners, then "squish" the rectangle using the path in Definition 5.6. For reasons discussed shortly, picking a rectangle uniformly at random leads to very poor bounds. In this section we show that the heuristic can be saved with a small tweak: rather than choosing uniformly at random, we make a very careful choice of "good" rectangles at each stage.

Throughout, we will denote by F a function which maps a configuration to a collection of "good" rectangles, and $F_\sigma^{(\text{pt})}$ a measure on paths associated with this function F. In order to upper-bound the congestion in Equation (5.3) (i.e. the total weight of paths using a given edge), we define a recovery function G:

$$G(e) = \{\sigma \in \Omega_\Lambda \ : \ \exists \gamma \ni e, \ \ s.t. \ \ F_\sigma^{(\text{pt})}(\gamma) > 0\},$$

(see Equation (5.19) at the start of Definition 5.19 for an equivalent definition). Our goal in the following definitions is to make $|G(e)|$ small (without too badly increasing the length of the associated paths or decreasing the minimum probability of any configuration along them). Given both an observed edge e contained in a path *and* the rectangle R that is being "squished" by a path, we can uniquely identify the initial configuration in $G(e)$. So, in order to bound $|G(e)|$, it is enough to bound the number of rectangles we *might* be flipping when we observe the edge e.

The details of the functions F and G are a bit complicated, and so for the reader's convenience we do not make explicit reference to them outside of this section. The only results required from this section are the notation for the path measure $F_\sigma^{(\text{pt})}$ (see Definition 5.16), and our final estimates in Proposition 5.20.

We now give a rough description of our path construction, followed by the full details. The main problem with the heuristic of Section 2.5 is illustrated in Figure 5. In this figure, we see an edge $e = (e_-, e_-^x)$ in a path $\gamma_{\sigma, R}$ for which the row containing the spin x that is being flipped by e has very many defects. In this case, the edge could appear in an enormous number of paths of the form in Definition 5.6. As a result, the congestion of our path is large and the canonical path method gives a very poor bound on the relaxation time.
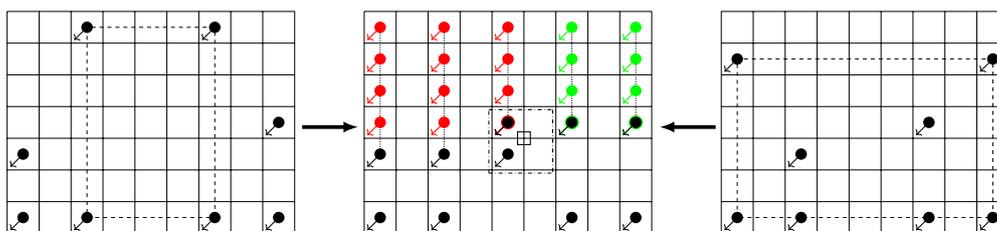

Figure 5: The middle display shows an edge (e_-, e_+); the defects in e_- are marked by black circles, and the spin that is flipped in e_+ is marked with a square (the corresponding defect variables that are flipped are contained in the dot-dashed box). If this edge occurs along the path in Definition 5.6, the top-left and top-right defects in the initial configuration could have appeared anywhere marked by red and green defects respectively, such that they are both in the same row. Two representative possible initial configurations which both go through this same edge are shown left and right. Only part of the lattice Λ is shown; the configurations do not satisfy the parity constraint, Equation 4.2, on this part of the lattice.

There is a conceptually-simple fix to the problem illustrated in Figure 5: if no row has "too many" defects, select rectangles *uniformly at random among all rectangles* with at least three defects at vertices; otherwise, choose *uniformly at random among rectangles*

with top on the high-density row. This definition doesn't quite work, but leads to the following more careful strategy:

1. If no row contains more than $\beta^{\Theta(1)}$ defects, choose any rectangle with defects at three or more corners.

2. Otherwise, let K be the maximal number of defects in any row and, for $1 \leqslant k \leqslant K$, let $\#(k)$ be the collection of rows with k defects. The "fix" above tells us to choose any rectangle that has defects at three or more corners *and* whose top edge is in $\#(K)$. However, in this case $|G(e)|$ may still be large.

   To see this, imagine we observe an edge in which the maximal number of defects in any row is K. Then it is possible that the top edge of the rectangle being flipped is in a row in $\#(K)$, but it is also possible that the top edge is in a row in $\#(K-2)$. In particular, if $|\#(K-2)|$ is much larger than $|\#(K)|$, the set $G(e)$ may be too large in comparison to the number of rectangles with top edges in $\#(K)$.

   Fortunately, this turns out to be almost the only obstacle: roughly, we can choose k to be the largest number that satisfies $|\#(k)| \gtrsim \beta^{-1}|\#(j)|$ for each j in some bounded interval near k. We then choose any rectangle that has defects at three or more corners *and* whose top edge is in $\#(k)$. It turns out this allows us to make a sufficiently good estimate of which row the top of the rectangle was in, having observed the edge e.

If we were only interested in bounding the relaxation time for $+$ boundary conditions, it would suffice to choose a rectangle uniformly at random from the subset constructed above. However, in order to obtain a bound on the mixing time, and on the relaxation time for the periodic boundary conditions, we need to get a much tighter bound for initial configurations that have very many defects. This tighter bound gives us better control on the constant appearing in (5.3) when the initial set S appearing in that lemma is small.

It turns out that we need a tighter bound exactly when the initial configuration, σ, has many more than L defects. To obtain a better bound in this case, we will "split" the lattice Λ into roughly $n(\sigma) \approx \frac{|D(\sigma)|}{L}$ pieces, each of which contains $O(L)$ defects. It will turn out that our main congestion bounds will apply to each "piece" of the split, and so this splitting analysis will give us a "free" improvement to our final bound on the order of $n(\sigma)$. This improvement is critical for obtaining a good bound on the mixing time of the process, since it will imply that defects can disappear much more quickly when there are very many of them. We don't know of any simple way to obtain this improvement by a factor of $n(\sigma)$ by other small modifications of our main argument.

The picture presented by this construction is quite simple, but we require quite a large amount of notation to describe it precisely. The main difficulties are that (i) we must also be able to reconstruct the part $i \in [1:L]$ of the split by observing an edge in a sample path, and (ii) the strategy used to define the collection of good rectangles in each split depends on comparing certain rough statistics of σ to certain threshold values, and it is not possible to use a single threshold value for all $\sigma \in \Omega_\Lambda$. We resolve this by partitioning the state space Ω_Λ into a small number of parts indexed by the choice of threshold. This type of threshold argument, and associated partitioning of the configuration space, seems to be useful for similar path arguments.

We now give some notation which is necessary to construct the functions F discussed above. We begin by introducing some notation for the rectangles that will be chosen by these functions. When three sites x, y, z in a lattice are corners of a rectangle, we denote this rectangle by $\mathbf{R}(x, y, z)$. We then introduce some special collections of rectangles:

Definition 5.7 (Marked Rectangles). *We follow the notation of Definition 5.5, assuming again that Λ is of the form $\Lambda = [\ell_1 : \ell_2] \times [\ell_3 : \ell_4]$. For $x, y \in \Lambda$, define $\mathbf{Top}(x, y)$ and $\mathbf{Bottom}(x, y)$ to be the maximum and minimum of x and y with respect to lexicographical order (respectively).*

For a defect configuration $p \in \Omega_{\mathcal{B}(\Lambda)}$ and $x = (x_1, x_2), y = (y_1, y_2) \in D(p)$, we define the associated extended "down-" and "up-" rectangles of x and y by setting

$$T_d(x, y) = \mathbf{R}(x, y, \mathbf{Top}(d(x), d(y))) \quad \textit{if } x_2 = y_2 \textit{ and at least one of } d(x), d(y) \textit{ exist,}$$
$$T_u(x, y) = \mathbf{R}(x, y, \mathbf{Bottom}(u(x), u(y))) \quad \textit{if } x_2 = y_2 \textit{ and at least one of } u(x), u(y) \textit{ exist,}$$

otherwise, say $T_d(x, y) = \emptyset$ or $T_u(x, y) = \emptyset$ respectively.

We denote by

$$\mathfrak{T}_p = \{T_a(x, y) : x, y \in D(p), \text{ and } a \in \{u, d\}\} \tag{5.11}$$

the collection of extended rectangles of p. For $i \in \mathbb{N}$, define

$$\mathfrak{T}_p(i) = \{T_a(x, y) : x, y \in D(p), i = y_2 = x_2, a \in \{u, d\}\} \tag{5.12}$$

to be the collection of extended rectangles with two defects in row i. When the lattice Λ is not clear from the context, we use the notation $\mathfrak{T}_{p,\Lambda}$ and $\mathfrak{T}_{p,\Lambda}(i)$ for Equations (5.11) and (5.12).

Finally, by a small abuse of notation we define $\mathfrak{T}_\sigma = \mathfrak{T}_{p^+(\sigma)}$ and $\mathfrak{T}_\sigma(i) = \mathfrak{T}_{p^+(\sigma)}$ for $\sigma \in \Omega_\Lambda^+$.

We will split the lattice Λ into roughly $\max(1, \frac{|D(\sigma)|}{L})$ pieces, each of which is a union of adjacent columns. Recall, if we consider spin configurations $\sigma \in \Omega_\Lambda$ for some box $\Lambda = [1\colon L]^2$ then the associated defect configurations, $p(\sigma) \in \Omega_{\mathcal{B}(\Lambda)}$, are on the slightly larger box $\mathcal{B}(\Lambda) = [0\colon L]^2$.

Definition 5.8 (Splits). *Let $\bar{\Lambda} = \mathcal{B}([1 : L]^2)$ and*

$$\bar{\Lambda}(i, j) = \{i, i + 1, \ldots, j\} \times [0 : L], \quad \textit{for } 0 \leqslant i < j \leqslant L.$$

For $\sigma \in \Omega_\Lambda$ set $s_1(\sigma) = 0$, then inductively define

$$s_{i+1}(\sigma) = \min\{j > s_i(\sigma) : |D(\sigma) \cap \bar{\Lambda}(s_i(\sigma), j - 1)| \geqslant 100\, L\}, \quad \textit{if this is finite,} \tag{5.13}$$
$$s_{i+1}(\sigma) = L + 1, \qquad \textit{otherwise,}$$

for $1 \leqslant i \leqslant m(\sigma) \equiv \min\{j : s_{j+1}(\sigma) = L + 1\}$. When σ is clear from context we write s_i for $s_i(\sigma)$. Finally, define $\mathcal{S}(\sigma) = \{s_i\}_{i=1}^{m+1}$ to be this splitting of $[0 : L]$, and $\mathcal{S}_i(\sigma) = \Lambda(s_i, s_{i+1} - 1)$.

The following definition gives some notation for counting the number of defects in a row of a sublattice:

Definition 5.9 (Occupancy). *Fix a subset $\Lambda = [\ell_1\colon\ell_2] \times [1\colon L] \subset [1\colon L]^2$, and fix $p \in \Omega_{\mathcal{B}(\Lambda)}$. For $0 \leqslant i \leqslant L$, define the occupancy of row i by*

$$\mathfrak{N}_i(p) = |D(p) \cap ([\ell_1 - 1 : \ell_2] \times \{i\})|, \tag{5.14}$$

the number of defects in row i. When the lattice is not clear from the context, we use the extended notation $\mathfrak{N}_{i,\Lambda}(p)$ for $\mathfrak{N}_i(p)$. Again, by a slight abuse of notation, for $\sigma \in \Omega_\Lambda^+$ we let $\mathfrak{N}_i(\sigma) = \mathfrak{N}_i(p^+(\sigma))$.

We give some basic facts about splits. Recall that we view p as a defect configuration, even if it is not in the image of Equation (2.1) for any configuration σ. In particular p may not satisfy the parity condition, for example the parity condition in a part of the split are satisfied in each column but not necessarily in each row. We now show that any piece of a split must have many extended rectangles, relative to the number of defects which are in the piece:

Lemma 5.10 (Splits Contain Many Rectangles). *Let* $\Lambda = [1:L]^2$, *for any* $\sigma \in \Omega_\Lambda^+$,

$$|\mathfrak{T}_\sigma| \;\geqslant\; \frac{1}{4}|D(\sigma)|\,, \tag{5.15}$$

and for $1 \;\leqslant\; i \;\leqslant\; m-1$, *letting* $p = p^+(\sigma)|_{\mathcal{S}_i(\sigma)}$, *then*

$$|\mathfrak{T}_{\sigma,\mathcal{S}_i(\sigma)}| \;\geqslant\; \frac{1}{4}(|D(p)| - L - 1)\,. \tag{5.16}$$

Furthermore, for any j *with* $\mathfrak{N}_{j,\mathcal{S}_i(\sigma)}(\sigma) \;\geqslant\; 1$,

$$|\mathfrak{T}_{\sigma,\mathcal{S}_i(\sigma)}(j)| \;\geqslant\; \frac{1}{2}(\mathfrak{N}_{j,\mathcal{S}_i(\sigma)}(\sigma) - 1)^2. \tag{5.17}$$

Proof. For $\sigma \in \Omega_\Lambda^+$, for each defect in $p^+(\sigma)$ there is at least one defect in the same row and at least one in the same column by the parity constraint. It follows that each defect belongs to at least one rectangle in $\mathfrak{T}_\sigma$, and can be located at one of at most four vertices of any such rectangle. Inequality (5.15) follows directly.

Note that, for any fixed $j \in [0:L]$, the set

$$\{x = (x_1, x_2) \in D(p) \,:\, x_2 = j, \, r(x) = \emptyset \text{ and } \ell(x) = \emptyset\}$$

of defects in row j, of part $\mathcal{S}_i(\sigma)$, with no left or right-partners cannot have more than one element, and hence the total number of defects without left or right-partners is bounded by $L + 1$. Every other defect also has at least one up or down-partner by the parity constraint, and the same argument as above implies Inequality (5.16). Essentially the same considerations also imply Inequality (5.17): any pair of defects in the same row must have at least one associated extended rectangle. $\square$

We observe that each piece of a split has at least $100L$ defects (except possibly the last piece), and splits associated with a high-defect configuration σ will have $m(\sigma) \gtrsim \frac{|D(\sigma)|}{L}$ pieces:

Lemma 5.11 (Counting Split Components). *Fix* $\sigma \in \Omega_{\mathcal{B}(\Lambda)}$, *and let* $\mathcal{S}(\sigma) = \{s_i\}_{i=1}^{m+1}$ *be as in Definition 5.8. Then*

$$\left\lfloor \frac{|D(\sigma)|}{101\,L} \right\rfloor \;\leqslant\; m(\sigma) \;\leqslant\; \frac{|D(\sigma)|}{100\,L} + 1.$$

Proof. By inspection,

$$100\,L \;\leqslant\; |D(\sigma) \cap \mathcal{S}_i(\sigma)| \;\leqslant\; 101\,L$$

for all $1 \;\leqslant\; i \;\leqslant\; m-1$, where the upper bound also holds for $i = m$. The result follows immediately. $\square$

To reduce notation in the rest of this section we define the occupancy vector of the splits.

Definition 5.12 (Occupancy vector). *Let* $\Lambda = [1 : L]^2$*, we define the row occupation vector function,* $r : \Omega_\Lambda \times [1 : L] \rightarrow [0 : L]^{L+1}$*, by*

$$r_k(\sigma, i) = \mathfrak{N}_{k, \mathcal{S}_i(\sigma)}(\sigma), \quad \text{for } k \in [0 : L],$$

so that $r_k(\sigma, i)$ *is the number of defects in row* k *of part* i *of the split.*
For any vector $v \in [0 : L]^{L+1}$ *we write* $\#_v(k) = |\{i \in [0 : L] : v_i = k\}|$ *and, with a slight abuse of notation,* $v_{\max} = \max_{k \in [0:L]} v_k$*, for short.*

We now partition the set of possible occupation vectors in order to define our choice of good rectangles. The parameter θ will allow us to approximate which row we started from, having observed an edge, without loosing more than a factor of $O(\beta)$.

Definition 5.13 (Good Partition). *Let* $\Theta = [0 : \beta]$*, define* $\mathcal{V}_0 = \{v \in [0 : L]^L : v_{\max} \leqslant \beta^2\}$ *and for* $\theta \in \Theta \setminus \{0\}$

$$\bar{\mathcal{V}}_\theta = \{v \in [0 : L]^L : v_{\max} > \beta^2, \ \forall k \in [-32, 32] \ \beta \#_v(v_{\max} - \theta) \geqslant \#_v(v_{\max} - \theta - k)\}.$$

For $\theta \geqslant 1$*, define recursively* $\mathcal{V}_\theta = \bar{\mathcal{V}}_\theta \setminus \bigcup_{i=0}^{\theta-1} \mathcal{V}_i$*.*

Lemma 5.14. *For all* β *sufficiently large, the sets* $\mathcal{V}_0, \mathcal{V}_1, \ldots, \mathcal{V}_\beta$*, partition* $[0 : L_c]^{L_c}$*.*

Proof. The sets $\mathcal{V}_0, \mathcal{V}_1, \ldots, \mathcal{V}_\beta$ are mutually disjoint by construction. Fix $v \in [0 : L_c]^{L_c}$. If $v_{\max} \leqslant \beta^2$ then $v \in \mathcal{V}_0$. Otherwise, when $v_{\max} > \beta^2$, we claim that $v \in \bigcup_{i=1}^{\beta} \mathcal{V}_i$. We prove this by contradiction: If not, there exists $n \in [0 : v_{\max}]$ such that $\#_v(n) > \beta^{\beta/32}$, which is greater than L for all sufficiently large β. $\qquad\square$

The configurations with row occupancy vectors in $\mathcal{V}_0$ are called *sparse* and will be treated differently from the other configurations. We note that the following definition is stated for general L, but makes use of the partition which is guaranteed to exist by Lemma 5.14. This is not a problem, since any L is equal to the critical length L_c for *some* value of β.

Definition 5.15 (Reconstructable Rectangles). *Fix* $\Lambda = [1 : L]^2$*, for* $\sigma \in \Omega_\Lambda$ *let* $\mathcal{S}(\sigma) = \{s_i\}_{i=1}^{m+1}$ *be the split associated with* σ*. For* $i \notin [1 : m(\sigma)]$*, set* $F(\sigma, i) = \emptyset$*. For* $i \in [1 : m(\sigma)]$*, let* $v = r(\sigma, i) \in [0 : L]^L$*, and define the set of good rectangles by*

$$F(\sigma, i) = \begin{cases} \mathfrak{T}_{\sigma, \mathcal{S}_i(\sigma)} & \text{if } v \in \mathcal{V}_0, \\ \bigcup_{j : v_j = v_{\max} - \underline{\theta}(v)} \mathfrak{T}_{\sigma, \mathcal{S}_i(\sigma)}(j) & \text{otherwise,} \end{cases}$$

where we define $\underline{\theta}(v)$ *to be the unique value of* θ *such that* $v \in \mathcal{V}_\theta$*.*

In the definition above, if the max row occupation in split i is small then just take all rectangles in split i with defects in at least three vertices. Otherwise, select *only* rectangles with two vertices containing a defect in a row with occupation $v_{\max} - \underline{\theta}(v)$ and at least one other vertex containing a defect.

We note that if $|D(\sigma)|$ is sufficiently large, so that the associated split has more than one part, then the last part of the split (the $m(\sigma)^{\text{th}}$) may contain no good rectangles (i.e. $F(\sigma, m(\sigma))$ may be empty). For this reason we avoid choosing this part of the split in the next definition.

We will say that $\sigma \in \Omega_\Lambda$ is *sparse* in $\mathcal{S}_i(\sigma)$ if $v \in \mathcal{V}_0$ and θ-*dense* if $v \in \mathcal{V}_\theta$ for $\theta > 0$. We now define the short paths which are used to construct the initial part of the full random paths. In words, we pick uniformly a part i of the split from 1 to $n(\sigma) := 1 \vee (m(\sigma) - 1)$, and then independently and uniformly choose a good rectangle to flip from split i.

Definition 5.16 (Partial Random Path). *For $\sigma \in \Omega_{[1:L]^2}$, define the "partial path" probability measure $F_\sigma^{(\mathrm{pt})}$ by*

$$F_\sigma^{(\mathrm{pt})}(\gamma) = \frac{1}{n(\sigma)\,|F(\sigma, i)|}\,,$$

if $\gamma = \gamma_{\sigma,R}$ for some $R \in F(\sigma, i)$ with $i \in [1 : n(\sigma)]$, and 0 otherwise.[2] Recall, here $n(\sigma) := 1 \vee (m(\sigma) - 1)$.

It is necessary to identify the index of the split initially chosen given an edge.

Definition 5.17 (Identifying the split index). *For $e = (e_-, e_+) \in \mathbf{E}$ let $x(e) = (x_1, x_2) \in \Lambda$ be the single site that is flipped by e, so that $e_+ = e_-^{x(e)}$, then define $\mathbf{I} : \mathbf{E} \mapsto [1 : L]$ by*

$$\mathbf{I}(e) = i \quad \text{s.t.} \quad (x_1 - 1, x_2) \in \mathcal{S}_i(e_-)\,, \tag{5.18}$$

which is unique since $\{\mathcal{S}_j(e_-)\}_{j=1}^m$ partitions $\mathcal{B}(\Lambda)$.

The function $\mathbf{I}$ always correctly identifies the index of the initial split chosen, made precise in the following sense.

Lemma 5.18. *For each edge $e \in \mathbf{E}$ and path $\gamma \ni e$ such that $F_\sigma^{(\mathrm{pt})}(\gamma) > 0$ for some $\sigma \in \Omega_\Lambda$, we have $\gamma = \gamma_{\sigma,R}$ for some $R \in F(\sigma, \mathbf{I}(e))$.*

Proof. Fix $e \in \mathbf{E}$, let $x = (x_1, x_2) \in \Lambda$ be the single site that is flipped by e, and fix σ and a path γ such that $e \in \gamma$ and $F_\sigma^{(\mathrm{pt})}(\gamma) > 0$. By Definition 5.16 there exists an $i \in [1, n(\sigma)]$ and an $R \in F(\sigma, i)$ such that $\gamma = \gamma_{\sigma,R}$. By Definition 5.20 we have $R \subset \mathcal{S}_i(\sigma)$. Without loss of generality suppose that the top two corners of R are in $D(\sigma)$ (otherwise we flip the configurations according to the mirror map M). Label the columns of Λ containing the left edge of R, the site x and the right edge of R by ℓ, x_1, r respectively. By construction of $\gamma_{\sigma,R}$ the defect configurations $p(\sigma)$ and $p(e_-)$ are equal outside of the columns $\ell, x_1 - 1, r$ (See Figure 4). It follows that $\mathcal{S}_j(\sigma) = \mathcal{S}_j(e_-)$ for $j < i$. Furthermore, the number of defects in columns ℓ and r in $p(e_-)$ are less than or equal to the number of defects in the same columns of $p(\sigma)$, see Definition 5.6 and Figure 4. In particular $|D(e_-) \cap \bar{\Lambda}(s_i(e_-), x_1 - 2)| \leqslant |D(\sigma) \cap \bar{\Lambda}(s_i(\sigma), x_1 - 2)|$, and hence $s_{i+1}(e_-) - 1 \geqslant x_1 - 1$ which implies $(x_1 - 1, x_2) \in \mathcal{S}_i(e_-)$ as required. $\square$

For bounding the congestion of out paths it will be convenient to introduce the initial configurations which are compatible with a given edge.

Definition 5.19 (Compatible Initial Configurations). *For each edge $e \in \mathbf{E}$ we define the set of all initial configurations $\sigma \in \Omega_\Lambda$ such that $F_\sigma^{(\mathrm{pt})}$ gives positive weight to a path that use the edge e*

$$G(e) = \Big\{\sigma \in \Omega_\Lambda \;:\; \exists R \in \bigcup_{i=1}^{n(\sigma)} F(\sigma, i) \;\; \text{s.t.}\;\; e \in \gamma_{\sigma,R}\Big\}. \tag{5.19}$$

We partition $G(e)$ in terms of the edge type and parameter θ. Consider $\sigma \in \Omega_\Lambda$ and a rectangle R with at least three of its four corners in $D(\sigma)$. Assume for now that both of the top two corners of R are in $D(\sigma)$. In this case, set

$$\mathbf{TYPE}(e, R, \sigma) = \begin{cases} \textbf{none}, & e \notin \gamma_{\sigma,R}\,, \\ \textbf{init}, & e \in \gamma_{\sigma,R} \text{ and } x \text{ is in the top row of } \mathcal{B}_-(R) \text{ and} \\ & \qquad \mathcal{B}_-(R) \text{ has more than 1 row}\,, \\ \textbf{fin}, & e \in \gamma_{\sigma,R} \text{ and } x \text{ is in the last row of } \mathcal{B}_-(R)\,, \\ \textbf{mid}, & \textit{otherwise.} \end{cases} \tag{5.20}$$

---

[2]Note that $F(\sigma, i) \cap F(\sigma, j) = \emptyset$ for $i \neq j$ since $F(\sigma, i)$ only contains rectangles for which all vertices are inside the ith part of the split, so this is a probability measure.

When the top two corners of R are not both in $D(\sigma)$, define the type by flipping all elements according to the mirror reflection map M given in Definition 5.6. Then

$$G_\theta^{(\mathrm{U})}(e) = \{\eta \in G(e) \,:\, \exists R \in F(\eta, \mathbf{I}(e)) \text{ s.t. } \mathbf{TYPE}(e, R, \eta) = U \text{ and } \underline{\theta}\big(v(\eta, \mathbf{I}(e))\big) = \theta\}\,.$$

We now summarize the required congestion and energy bounds that will be used in the final analysis for the complete paths.

Proposition 5.20. *Consider $\beta, L > 0$ with $L = L_c(\beta)$ and set $\Lambda = [1 : L]^2$. Then:*

1. ***Covering Property:*** *For all $e \in \mathbf{E}$*

$$G(e) = \bigcup_{\theta \in \Theta} \big(G_\theta^{(\mathrm{init})}(e) \cup G_\theta^{(\mathrm{mid})}(e) \cup G_\theta^{(\mathrm{fin})}(e) \big)\,. \tag{5.21}$$

2. ***Small Congestion:*** *For all $e = (e_-, e_+) \in \mathbf{E}$ and $\theta \in \Theta$,*

$$\sum_{\sigma \in G_\theta^{(\mathrm{U})}(e)} \sum_{\gamma \ni e} F_\sigma^{(\mathrm{pt})}(\gamma) \lesssim \frac{\beta^4 L}{|D(e_-)|}, \quad U \in \{\mathrm{init}, \mathrm{mid}\}, \tag{5.22}$$

$$\sum_{\sigma \in G_\theta^{(\mathrm{fin})}(e)} \sum_{\gamma \ni e} F_\sigma^{(\mathrm{pt})}(\gamma) \lesssim \frac{\beta^2 L^2}{|D(e_-)|}\,.$$

3. ***Many Options:*** *We have*

$$F(\sigma, i) \neq \emptyset\,, \quad \text{for all } \sigma \in \Omega_\Lambda \backslash \mathcal{G}, i \in [1 : n(\sigma)] \text{ and} \tag{5.23}$$

$$n(\sigma) \gtrsim \frac{|D(\sigma)|}{L}\,, \quad \text{for all } \sigma \in \Omega_\Lambda\,.$$

4. ***Energy Bound:*** *For all $\sigma \in \Omega_\Lambda$, $\theta \in \Theta$ and $e \in \mathbf{E}$,*

$$\pi(e_-)\mathcal{L}(e_-, e_+) \geqslant e^{-2\beta}\pi(\sigma)\,, \qquad \sigma \in G_\theta^{(\mathrm{init})}(e) \cup G_\theta^{(\mathrm{mid})}(e) \tag{5.24}$$

$$\pi(e_-)\mathcal{L}(e_-, e_+) \geqslant \pi(\sigma)\,, \qquad \sigma \in G_\theta^{(\mathrm{fin})}(e)\,.$$

To give the congestion bounds above we will use the following preliminary results related to the row occupancy vectors.

Lemma 5.21. *Fix $e \in \mathbf{E}$, and let $I = \mathbf{I}(e_-)$ as in Definition 5.17, define $u \in [0 : L]^2$ by $u = r(e_-, I)$, then*

$$\sup_{\sigma \in G(e)} \|r(\sigma, I) - u\|_\infty \leqslant 8\,,$$

where $r(\sigma, I)$ is given in Definition 5.12.

Proof. Fix $e \in E$ and $\sigma \in G(e)$, then there exists a rectangle R such that $e \in \gamma_{\sigma, R}$. Any Rectangle-Removal Path (see Definition 5.6) can change the number of defects in any row of the lattice by at most two, i.e. $|\mathfrak{N}_k(e_-) - \mathfrak{N}_k(\sigma)| \leqslant 2$ for $k \in [0 : L]$. Furthermore the total number of defects can differ by at most six, $\big||D(e_-)| - |D(\sigma)|\big| \leqslant 6$. By the same argument as in the proof of Lemma 5.18 we have $s_I(e_-) = s_I(\sigma)$, and since $\big||D(e_-)| - |D(\sigma)|\big| \leqslant 6$ we know that $S_I(e_-) \triangle S_I(\sigma)$ contains at most six columns which contain defects in e_-. Therefore, the difference in the number of defects in any row of the Ith split is bounded above by $2 + 6 = 8$. $\qquad\square$

The following lemma will be used, in the θ-dense case, to bound the number of rows a rectangle removal path could have started from.

Lemma 5.22. *Consider* $\beta, L > 0$ *with* $L = L_c(\beta)$. *For each* $u \in [0 : L]^L$ *and* $\theta \in \Theta \setminus \{0\}$, *let*

$$\Delta_u^\theta = \{v \in \mathcal{V}_\theta \, : \, \|v - u\|_\infty \leqslant 8\},$$

and

$$H_\theta(u) := \{i \in [1 : L] \, : \, \exists v \in \Delta_u^\theta \ \text{s.t.} \ v_i = v_{\max} - \theta\}, \tag{5.25}$$

then

$$|H_\theta(u)| \lesssim \beta \inf_{v \in \Delta_u^\theta} \#_v(v_{\max} - \theta). \tag{5.26}$$

Proof. Fix $u \in [0 : L]^L$ and $\theta \in \Theta \setminus \{0\}$. We first show that

$$H_\theta(u) \subseteq \{i \in [1 : L] \, : \, \exists \delta \in [-16 : 16] \ \text{s.t.} \ u_i = u_{\max} - \theta + \delta\}.$$

Fix $j \in [1 : L]$ and suppose there exists a $v \in \Delta_u^\theta$ such that $v_j = v_{\max} - \theta$, then $u_j = v_{\max} - \theta + \delta'$ for some $\delta' \in [-8 : 8]$ by definition of Δ_u^θ. Similarly $u_{\max} = v_{\max} + \delta''$ for some $\delta'' \in [-8 : 8]$, and putting these together $u_j = u_{\max} - \theta + \delta$ for some $\delta \in [-16 : 16]$, proving the claim. We now show that the set on the right hand side of the display above is not too large. Observe that

$$|\{i \in [1 : L] \, : \, \exists \delta \in [-16 : 16] \ \text{s.t.} \ u_i = u_{\max} - \theta + \delta\}| = \sum_{\delta \in [-16:16]} \#_u(u_{\max} - \theta + \delta),$$

and by the same argument as above, for any $v \in \Delta_u^\theta$ we have

$$\#_u(u_{\max} - \theta + \delta) \leqslant \sum_{\delta' \in [-16:16]} \#_v(v_{\max} - \theta + \delta + \delta').$$

Hence

$$\left|\{i \in [1 : L] \, : \, \exists v \in \Delta_u^\theta \ \text{s.t.} \ v_i = v_{\max} - \theta\}\right| \leqslant 256 \sup_{k \in [-32:32]} \#_v(v_{\max} - \theta + k)$$
$$\lesssim \beta \, \#_v(v_{\max} - \theta),$$

where the last inequality follows from the definition of $\mathcal{V}_\theta$. $\qquad\square$

We now give the main bounds on the sizes of $G_\theta^{(\text{init})}$, $G_\theta^{(\text{mid})}$ and $G_\theta^{(\text{fin})}$ used for bounding the congestion in Proposition 5.20. We give bounds in two cases, corresponding to initially sparse configurations and initially dense configurations.

Lemma 5.23 (Number of Sparse Preimages). *Consider* $\beta, L > 0$ *with* $L = L_c(\beta)$ *and the original lattice* $[1 : L]^2$. *Fix an edge* $e = (e_-, e_+) \in \mathbf{E}$. *Let* $x = (x_1, x_2)$ *be the spin that is flipped by edge* e, *so that* $e_+ = e_-^x$. *Let* $I = \mathbf{I}(e)$. *We have:*

$$|\{(\sigma, R) \in G_0^{(\text{fin})}(e) \times \mathbf{R}_{[1:L]^2} \, : \, e \in \gamma_{\sigma,R} \ and \ F_\sigma^{(\text{pt})}(\gamma_{\sigma,R}) > 0\}| \lesssim \beta^2 L^2 \tag{5.27}$$

and

$$|\{(\sigma, R) \in G_0^{(\text{init})}(e) \times \mathbf{R}_{[1:L]^2} \, : \, e \in \gamma_{\sigma,R} \ and \ F_\sigma^{(\text{pt})}(\gamma_{\sigma,R}) > 0\}| \lesssim \beta^4 L \tag{5.28}$$
$$|\{(\sigma, R) \in G_0^{(\text{mid})}(e) \times \mathbf{R}_{[1:L]^2} \, : \, e \in \gamma_{\sigma,R} \ and \ F_\sigma^{(\text{pt})}(\gamma_{\sigma,R}) > 0\}| \lesssim \beta^4 L.$$

Proof. Before giving a proof, we note that our reconstruction possibilities are illustrated in Figure 6. The following proof is essentially a detailed explanation of the figure, and the heuristic appearing in the figure's caption does not hide any important details.
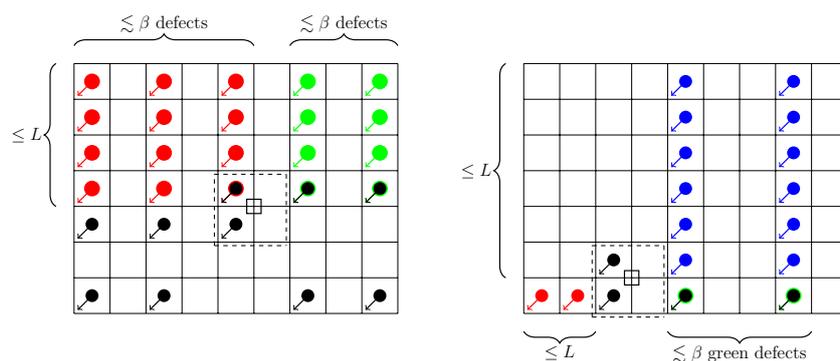
Figure 6: The defects in $D(e_-)$ are indicated by black circles, the site x which is flipped in the edge e is marked by a solid box, the dashed box shows the associated defects which are flipped. Left: As in Figure 5, the red and green circles indicate where the initial top defects of the rectangle R in $D(\sigma)$ are allowed to be, for **mid** or **int** type. Right: Type **fin**; the top right corner of R must be above one of the defects of e_-, in row $x_2 - 1$, to the right of x (marked by green), the bottom left corner must be to the left of x, in row $x_2 - 1$ (marked by red circles), and the top right corner must be in one of at most L sites above the bottom left corner.

We begin by proving Inequality (5.27). Let

$$P = \left\{ (\sigma, R) \in G_0^{(\text{fin})}(e) \times \mathbf{R}_{[1:L]^2} \; : \; e \in \gamma_{\sigma, R} \text{ and } F_\sigma^{(\text{pt})}(\gamma_{\sigma, R}) > 0 \right\}. \qquad (5.29)$$

It is clear from Definition 5.6 that, for any fixed edge e and rectangle R, there is at most one σ satisfying $e \in \gamma_{\sigma, R}$; denote it $\sigma = \sigma(e, R)$. Based on this observation, we will find an upper bound on $|P|$ by reconstructing all candidate rectangles R such that $(\sigma(e, R), R) \in P$.

Set $\mathcal{S}_I'(e) = \Lambda(s_I(e_-), s_{I+1}'(e_-))$, where

$$s_{I+1}'(e_-) = \min\{ j > s_I(\sigma) \; : \; |D(e_-) \cap \bar{\Lambda}(s_I(e_-), j - 1)| \; \geqslant \; 100\,L + 6 \}. \qquad (5.30)$$

Then for any $\sigma \in G(e)$ with $e \in \gamma_{\sigma, R}$, we have $\mathcal{S}_I(\sigma) \subset \mathcal{S}_I'(e)$ and $\mathcal{B}_-(R) \subset \mathcal{S}_I'(e)$. To reduce notation let $\Lambda = \mathcal{S}_I'(e)$ for the remainder of this proof.

For a rectangle R, denote by $(R^{(u,\ell)}, R^{(u,r)}, R^{(d,r)}, R^{(d,\ell)})$ the 4 corners of R in the usual clockwise order, starting at the upper-left hand corner. Note that it is possible to reconstruct all of R if you know the location of the opposite corners $R^{(u,\ell)}$ and $R^{(d,r)}$.

We next introduce notation that is meant to mimic the usual notation for conditional probability. Denote $R^{(a,b)} = (R_1^{(a,b)}, R_2^{(a,b)})$ where $a \in \{u, d\}$ and $b \in \{\ell, r\}$. For

$$A, B \subset \{(a, b, c)\}_{a \in \{u,d\}, b \in \{\ell, r\}, c \in \{1, 2\}}$$

some subsets of the 8 parameters specifying the coordinates of vertices of R, denote by $N[(R_c^{(a,b)})_{(a,b,c) \in A} \mid (R_c^{(a,b)})_{(a,b,c) \in B}]$ the total number of possible values of the collection $\{R_c^{(a,b)}\}_{(a,b,c) \in A}$ such that $(\sigma(e, R), R) \in P$, given that $\{R_c^{(a,b)}\}_{(a,b,c) \in B}$ are fixed. Note that $N[\cdot|\cdot]$ depends on the edge e and the type, but we do not emphasize this in the notation. By a small abuse of notation, set $N[\cdot] = N[\cdot \mid \emptyset]$.

Note that the configuration σ determined by e and R must have defects in at least three corners of R; we assume for the remainder of the proof that it has defects in the upper-left, upper-right, and bottom-left corners. Note that this assumption under-counts the number of valid rectangles by at most a factor of 4, and so it will not affect the order of the final estimate of $|P|$.

We begin with the trivial bound

$$N[R^{(u,\ell)}] \leqslant L^2, \tag{5.31}$$

since there are at most L^2 elements of Λ. Given $R^{(u,\ell)}$, we know that $R_1^{(d,\ell)} = R_1^{(u,\ell)}$ and $R_2^{(d,\ell)} = (x_2 - 1)$, since $\mathbf{TYPE}(e, R, \sigma) = \mathbf{fin}$, so

$$N[R^{(d,\ell)}|R^{(u,\ell)}] \leqslant 1. \tag{5.32}$$

Since $R_2^{(d,r)} = R_2^{(d,\ell)}$ and $R^{(d,r)} \in D(e_-)$,

$$N[R^{(d,r)}|R^{(d,\ell)}] \leqslant \max_{(\sigma,R)\in P} \max_{j\in\mathbb{N}} \mathfrak{N}_{j,\mathcal{S}_I(\sigma)}(\sigma) + 2 \lesssim \beta^2, \tag{5.33}$$

where the last inequality follows from the fact that $\underline{\theta}(v(\sigma, I)) = 0$ for $(\sigma, R) \in P$. Combining Inequalities (5.31), (5.32) and (5.33), we conclude in this case that

$$|P| \leqslant N[(R^{(u,\ell)}, R^{(d,r)})] \lesssim \beta^2 L^2.$$

We now prove Inequality (5.28) following the same strategy. We now define P analogously to Equation (5.29) for type **init** or **mid**. We begin with the trivial bound $N[R_2^{(u,\ell)}] \leqslant L$, since the height of Λ is L. Next, note that the column containing $R^{(u,\ell)}$ must intersect the row $(x_2 - 1)$ at either a defect of e_- or at $(x_1 - 1, x_2 - 1)$. Since there are at most $\max_{j\in\mathbb{N}} \mathfrak{N}_{j,\mathcal{S}_I(\sigma)}(\sigma) + 2 \lesssim \beta^2$ defects in any row of e_- for any $(\sigma, R) \in P$, this observation implies $N[R_1^{(u,\ell)}|R_2^{(u,\ell)}] \lesssim \beta^2$, and so

$$N[R^{(u,\ell)}] \lesssim \beta^2 L. \tag{5.34}$$

Inequality (5.32) also holds in this case, though the argument is slightly different, as follows. Inspecting Equation (5.11) from Definition 5.7, we see that if $\hat{p} \in \Omega_{[\ell_1:\ell_2]\times[\ell_3\times\ell_4]}$ is a defect configuration and $\hat{R}$ is a rectangle with three corners in $\mathfrak{T}_{\hat{p}}$, then $\hat{p}$ has no defects in the line between $\hat{R}^{(u,\ell)}$ and $\hat{R}^{(d,\ell)}$. Therefore, $R^{(d,\ell)}$ must be the highest defect of p that is in the same column as $R^{(u,\ell)}$ and below the row $(x_2 - 1)$, and so Inequality (5.32) holds. Inequality (5.33) holds in this case as well, thus combining Inequalities (5.34), (5.32) and (5.33) we have

$$|P| \leqslant N[R^{(u,\ell)}, R^{(d,r)}] \lesssim \beta^4 L.$$

This completes the proof of Inequality (5.28). $\qquad\square$

We now prove the analogous result for the remaining configurations:

Lemma 5.24 (Number of Dense Preimages). *Consider $\beta, L > 0$ with $L = L_c(\beta)$ and the original lattice $[1 : L]^2$. Fix $\theta \in \Theta \setminus \{0\}$ and an edge $e = (e_-, e_+) \in \mathbf{E}$, let $x = (x_1, x_2)$ be the spin that is flipped by edge e, and let $u = r(e_-, \mathbf{I}(e_-))$ with $u_{\max} = \max_{k\in[0:k]} u_k$. Then we have:*

$$|\{(\sigma, R) \in G_\theta^{(\text{init})}(e) \times \mathbf{R}_{[1:L]^2} : e \in \gamma_{\sigma,R} \text{ and } F_\sigma^{(\text{pt})}(\gamma_{\sigma,R}) > 0\}| \lesssim u_{\max}^2 |H_\theta(u)|, \tag{5.35}$$

$$|\{(\sigma, R) \in G_\theta^{(\text{mid})}(e) \times \mathbf{R}_{[1:L]^2} : e \in \gamma_{\sigma,R} \text{ and } F_\sigma^{(\text{pt})}(\gamma_{\sigma,R}) > 0\}| \lesssim u_{\max}^2 |H_\theta(u)|,$$

and

$$|\{(\sigma, R) \in G_\theta^{(\text{fin})}(e) \times \mathbf{R}_{[1:L]^2} : e \in \gamma_{\sigma,R} \text{ and } F_\sigma^{(\text{pt})}(\gamma_{\sigma,R}) > 0\}| \lesssim L\, u_{\max} |H_\theta(u)|. \tag{5.36}$$

Proof. The proof will be very similar to the proof of Lemma 5.23, and the overall picture is somewhat similar to Figure 6. The main difference is that, in the previous lemma and figure, we bounded the number of locations for the top row of the rectangle R using the trivial upper bound of L. Here, we use the much better bound $|H_\theta(u)|$. Since any nontrivial dense configuration is quite complicated, we do not include a new figure in this case.

We begin by proving Inequality (5.35). We set notation, fixing $U \in \{\mathrm{mid}, \mathrm{init}\}$ and then:

$$P = \{(\sigma, R) \in G_\theta^{(U)}(e) \times \mathbf{R}_{[1:L]^2} \; : \; e \in \gamma_{\sigma,R} \text{ and } F_\sigma^{(\mathrm{pt})}(\gamma_{\sigma,R}) > 0\}\,.$$

Again, we bound the number of elements $(\sigma, R) \in P$ by counting the number of ways to reconstructing a suitable rectangle R from our knowledge of e.

Set $\Lambda = \mathcal{S}_I'(e)$ as defined in the proof of Lemma 5.23. Also, define $R, N[\cdot|\cdot]$ as in the proof of Lemma 5.23. By the same argument as in Lemma 5.23, leading to Inequality (5.34), we know that $R^{(u,\ell)}$ has to be in a column that either

1. is $x_1 - 1$, or
2. contains some $y = (y_1, y_2) \in D(e_-)$ with $y_2 = x_2 - 1$.

Since no row of Λ contains more than $u_{\max} + 8$ defects in e_-, this implies

$$N[R_1^{(u,\ell)}] \lesssim u_{\max}. \tag{5.37}$$

On the other hand, by the definition of $F(\sigma, I)$ together with Lemmas 5.21 and 5.22, the row containing $R^{(u,\ell)}$ must be in $H_\theta(u)$. Thus,

$$N[R_2^{(u,\ell)}|R_1^{(u,\ell)}] \lesssim |H_\theta(u)|. \tag{5.38}$$

Combining Inequalities (5.37) and (5.38), we have

$$N[R^{(u,\ell)}] \lesssim u_{\max}|H_\theta(u)|. \tag{5.39}$$

Next, $R^{(u,r)}$ must be in the same row as $R^{(u,\ell)}$, and must be in either

1. column x_1, or
2. a column containing an element $y = (y_1, y_2) \in D(e_-)$ with $y_2 = x_2$.[3]

Thus,

$$N[R^{(u,r)}|R^{(u,\ell)}] \lesssim u_{\max}\,. \tag{5.40}$$

Finally, Inequality (5.32) holds by the same arguments as in the previous lemma (recall that this inequality had two proofs – one if $U = \mathrm{fin}$, another if $U \in \{\mathrm{mid}, \mathrm{init}\}$). Thus, combining Inequalities (5.39), (5.40) and (5.32),

$$|P| \;\leqslant\; N[R^{(u,\ell)}, R^{(d,r)}] \lesssim u_{\max}^2|H_\theta(u)|.$$

This completes the proof of Inequality (5.35).

Next, we prove Inequality (5.36) using a similar argument, again we redefine P accordingly. Inequality (5.37) does not hold in this case, since we have already removed

---

[3]It is not an accident that the value of y_2 is off-by-one from our analysis of the location of $R^{(u,\ell)}$; see Figure 6.

the defect that was originally in the bottom left corner of R, so we use instead the weaker trivial bound

$$N[R_1^{(u,\ell)}] \leqslant L,$$

since there are at most L columns in Λ. Next, noting that Inequality (5.38) holds with the same argument as above, this implies

$$N[R^{(u,\ell)}] \lesssim L\,|H_\theta(u)|. \tag{5.41}$$

Inequalities (5.40) and (5.32) hold by the same arguments given above. Thus, combining Inequalities (5.41), (5.40) and (5.32), we conclude

$$|P| \leqslant N[R^{(u,\ell)}, R^{(d,r)}] \lesssim L\,u_{\max}\,|H_\theta(u)|.$$

This completes the proof of the lemma. $\qquad\square$

As our final preliminary congestion estimate, we bound the number of choices given by the function F. Recall $n(\sigma) = 1 \vee (m(\sigma) - 1)$ is the number of parts of the split which contain at least $100L$ defects if there are two or more parts.

Lemma 5.25. *Consider* $\beta, L > 0$ *with* $L = L_c(\beta)$, *let* $\Lambda = [1:L]^2$, *and fix* $\sigma \in \Omega_\Lambda$. *Let* $p = p^+(\sigma)$, *fix* $i \in [1:n(\sigma)]$ *and define* $v = r(\sigma, i)$. *Then we have the following bounds on* $|F(\sigma, i)|$:

1. *If* $\underline{\theta}(v) = 0$, *then*

$$|F(\sigma, i)| \gtrsim |D(p|_{\mathcal{S}_i(\sigma)})|. \tag{5.42}$$

2. *If* $\underline{\theta}(v) \in \Theta \setminus \{0\}$, *then*

$$|F(\sigma, i)| \gtrsim (v_{\max} - \underline{\theta}(v))^2\,\#_v(v_{\max} - \underline{\theta}(v)). \tag{5.43}$$

Proof. In the case $n(\sigma) = 1$, Inequality (5.42) follows immediately from Inequality (5.15) in Lemma 5.10. When $n(\sigma) > 1$, then $|D(p|_{\mathcal{S}_i(\sigma)})| \geqslant 100L$ for each $i \in [1:n(\sigma)]$, so Inequality (5.42) follows from (5.16) in Lemma 5.10. Inequality (5.43) follows immediately from (5.17) in Lemma 5.10. $\qquad\square$

Having completed the preliminary congestion bounds we now prove Proposition 5.20:

Proof of Proposition 5.20. Properties **(1)**, and **(3)** are almost immediate. Property **(1)**, given in (5.21), follows immediately from Definition 5.19. Property **(3)**, given in Equation (5.23), follows from the definition of F in Definition 5.15 and Lemma 5.11. Thus, it only remains to prove Properties **(2)** and **(4)**.

We prove the last line of Inequality (5.22). Fix $e \in \mathbf{E}$ and let $x = (x_1, x_2)$ be the spin that is flipped by edge e (so that $e_+ = e_-^x$) and fix $\theta \in \Theta \setminus \{0\}$ such that $G_\theta^{(\text{fin})}$ is not empty, finally let $i = \mathbf{I}(e)$ and $u = r(e_-, i)$. Recalling (5.27) of Lemma 5.23 and Inequality (5.36) of Lemma 5.24,

$$|\{(\sigma, R) \in G_0^{(\text{fin})}(e) \times \mathbf{R}_{[1:L]^2} : e \in \gamma_{\sigma,R} \text{ and } F_\sigma^{(\text{pt})}(\gamma_{\sigma,R}) > 0\}| \lesssim \beta^2 L^2, \tag{5.44}$$

$$|\{(\sigma, R) \in G_\theta^{(\text{fin})}(e) \times \mathbf{R}_{[1:L]^2} : e \in \gamma_{\sigma,R} \text{ and } F_\sigma^{(\text{pt})}(\gamma_{\sigma,R}) > 0\}| \lesssim L\,u_{\max}\,|H_\theta(u)|.$$

On the other hand, consider $\sigma \in G_0^{(\text{fin})}(e)$. By Lemma 5.25, the fact that first $m(\sigma) - 1$ splits each contain at least $100L$ defects, and the fact that the only split contains $|D(\sigma)|$ defects if $m(\sigma) = 1$, we have

$$|F(\sigma, i)| \gtrsim \min(L, |D(e_-)|), \qquad \sigma \in G_0^{(\text{fin})}(e); \tag{5.45}$$

we have used the observation $|D(\sigma)| \asymp |D(e_-)|$, since any rectangle-removal path creates or removes at most six defects and $|D(e_-)| \geqslant 4$. When $\sigma \in G_\theta^{(\mathrm{fin})}(e)$ with $\theta \in [1:\beta]$, we have

$$|F(\sigma, i)| \overset{\text{Lemma } 5.25}{\gtrsim} (v_{\max} - \theta)^2 \, \#_v(v_{\max} - \theta) \gtrsim u_{\max}^2 \, \#_v(v_{\max} - \theta) \qquad (5.46)$$
$$\overset{\text{Ineq. } (5.26)}{\gtrsim} u_{\max}^2 \beta^{-1} |H_\theta(u)|,$$

where in the second inequality we used $v_{\max} \geqslant \beta^2$, $v_{\max} > u_{\max} - 16$ and $\theta \leqslant \beta$.

We consider the initially sparse and θ-dense cases separately:

1. **Sparse initial configurations**. Fix $\tau \in G_0^{(\mathrm{fin})}(e)$, in this case, Inequalities (5.44) and (5.45) give

$$\frac{|\{(\sigma, R) \in G_0^{(\mathrm{fin})}(e) \times \mathbf{R}_{[1:L]^2} \; : \; e \in \gamma_{\sigma, R} \text{ and } F_\sigma^{(\mathrm{pt})}(\gamma_{\sigma, R}) > 0\}|}{F(\tau, i)} \leqslant \frac{\beta^2 L^2}{\min(L, |D(e_-)|)}.$$

2. **θ-dense initial configurations**. Fix $\tau \in G_\theta^{(\mathrm{fin})}(e)$, in this case, Inequalities (5.44) and (5.46) give

$$\frac{|\{(\sigma, R) \in G_\theta^{(\mathrm{fin})}(e) \times \mathbf{R}_{[1:L]^2} \; : \; e \in \gamma_{\sigma, R} \text{ and } F_\sigma^{(\mathrm{pt})}(\gamma_{\sigma, R}) > 0\}|}{|F(\tau, i)|} \lesssim \frac{L \, u_{\max} \, |H_\theta(u)|}{\beta^{-1} u_{\max}^2 \, |H_\theta(u)|}$$
$$\lesssim \beta^{-1} L$$
$$\leqslant \frac{\beta^2 L^2}{\min(L, |D(e_-)|)}.$$

Combining these two cases,

$$\sum_{\sigma \in G_\theta^{(\mathrm{fin})}(e)} \sum_{R \, : \, e \in \gamma_{\sigma, R}} \frac{1}{|F(\sigma, i)|} \leqslant \sum_{\sigma \in G_0^{(\mathrm{fin})}(e)} \sum_{R \, : \, e \in \gamma_{\sigma, R}} \frac{1}{|F(\sigma, i)|} + \sum_{\sigma \in G_\theta^{(\mathrm{fin})}(e)} \sum_{R \, : \, e \in \gamma_{\sigma, R}} \frac{1}{|F(\sigma, i)|}$$
$$\lesssim \frac{\beta^2 L^2}{\min(L, |D(e_-)|)}.$$

If we consider initial configurations σ such that $|D(\sigma)| \lesssim L$ and apply the trivial bound $n(\sigma) \geqslant 1$ we get the desired result by the definition of $F_\sigma^{(\mathrm{pt})}$. Finally considering the initial configurations σ such that $|D(\sigma)| \gtrsim L$, in this case we apply **(3)** proved above to get the desired bound. This completes the proof of the last line of Inequality (5.22).

The proof of the other two inequalities in (5.22) is essentially identical, with the following simple changes to the two cases:

1. Replace references to Inequality (5.27) of Lemma 5.23 with references to Inequality (5.28) of Lemma 5.23 (the upper bound is better by a factor of $\beta^2 L^{-1}$).
2. Replace references to Inequality (5.36) of Lemma 5.24 with references to Inequality (5.35) of Lemma 5.24 (in this case the upper bound is better by a factor of $L^{-1} u_{\max}$).

Finally, we prove Property **(4)**, the energy bound. Inequality (5.24) can be observed directly from the definition of the rectangle-removal paths in Definition 5.6. We find it easiest to see this by simply following along in Figure 4; we now give a short description of what can occur. Begin by assuming there are no defects in the interior of the rectangle. In this case the first spin that is "flipped" in any row besides the last one will create a new pair of vertically adjacent defects that travel across the row, and the last spin that is "flipped" in the row will delete both elements of this pair. This proves the first line of

Inequality (5.24) in this case. The second line of Inequality (5.24) follows in this case from noticing that the last row of the rectangle will *already* have a pair of vertically aligned defects, and so no new defects are created. Finally, it is straightforward to see that any defects in the *interior* of the rectangle can only decrease the number of excess defects created during a rectangle-removal path. This completes the proof of Proposition 5.20. $\qquad\square$

5.4 Path construction: high density

The first path in this section will be appended to previously-defined paths after roughly L defects have been removed, as it allows us to give a good bound on the total path length (at the cost of an enormous penalty to both congestion and energy). The path simply flips all -1 spins to $+1$ spins, in order:

Definition 5.26 (Naive Paths). *For fixed $\sigma \in \Omega_{[1:L]^2}$, we define a ($\{0,1\}$-valued) probability mass $F_\sigma^{(na)}$ on $\Gamma_{\sigma,+}$ by giving an explicit algorithm for sampling from the measure. Let $z^{(1)}, z^{(2)}, \ldots$ be all points of $[1:L]^2$ in lexicographic order.*

1. *Initialize by setting $\gamma = \{\sigma\}$ and $i = 0$.*

2. *While $\gamma^{(fin)} \neq +$, do the following:*

   (a) *Set $i = i + 1$.*
   (b) *If $\gamma_{z^{(i)}}^{(fin)} = -1$, do the following:*

      i. *Set $\sigma = \gamma^{(fin)}$.*
      ii. *Set $\gamma = \gamma \bullet (\gamma^{(fin)}, \sigma^{z^{(i)}})$.*

3. *Return the path γ.*

We then define the complete measure on paths that will be analyzed in the remainder of this paper:

Definition 5.27 (Full Random Path). *For fixed $\sigma \in \Omega_{[1:L]^2}$, we define a probability measure F_σ on $\Gamma_{\sigma,+}$ by giving an explicit algorithm for sampling from this measure:*

1. *Initialize by setting $\gamma = \{\sigma\}$ and $i = 1$.*

2. *While $D(\gamma^{(fin)}) \neq \emptyset$, iteratively sample subpaths $\gamma_1, \gamma_2, \ldots$ according to the following loop.*

   (a) *If $i \leqslant \beta L$, sample $\gamma_i \sim F_{\gamma^{(fin)}}^{(pt)}$.*
   (b) *If $i > \beta L$, sample $\gamma_i \sim F_{\gamma^{(fin)}}^{(na)}$, according to Definition 5.26.*
   (c) *In both of these cases, set $\gamma = \gamma \bullet \gamma_i$ and then $i = i + 1$.*

3. *Once γ satisfies $D(\gamma^{(fin)}) = \emptyset$, return the path γ.*

When bounding the mixing time, it is useful to consider truncated paths as follows. Fix a truncation level $0 \leqslant k \leqslant L^2$. We define the measure $F_\sigma^{(k)}$ on Γ by the following algorithm:

1. *Sample the path $(\sigma^{(0)}, \ldots, \sigma^{(m)}) \sim F_\sigma$.*

2. *Let $i_{min} = \min\{i : |D(\sigma^{(i)})| \leqslant k\}$.*

3. *Return the path $\gamma = (\sigma^{(0)}, \ldots, \sigma^{(i_{min})})$.*

5.5 Path analysis: bounds on congestion and path length

We analyze the random path constructed in Definition 5.27. We begin by bounding the congestion associated with the subpath given in Definition 5.26. We will then bound the congestion associated to entire paths.

Definition 5.28 (Compatible Naive Paths). *Fix an edge $e = (e_-, e_+) \in \mathbf{E}$. We now define the set $G^{(\mathrm{na})}(e) \subset \Gamma$ of compatible naive paths associated with edge e as follows. Say $\gamma \in G^{(\mathrm{na})}(e)$ if it satisfies both of the following:*

1. *$e \in \gamma$, and*
2. *there exists a configuration $\sigma \in \Omega_{[1:L]^2}$ so that $F_\sigma^{(\mathrm{na})}(\gamma) > 0$.*

We now bound the sizes of these sets of compatible paths. Recall the naive paths are deterministic, so $F_\sigma^{(\mathrm{na})}(\gamma) > 0$ for exactly one path γ.

Lemma 5.29 (Number of Compatible Naive Paths). *Fix an edge $e \in \mathbf{E}$ and integer $k \in \mathbb{N}$. With notation as above,*

$$|\{\gamma \in G^{(\mathrm{na})}(e) \,:\, |D(\gamma^{(\mathrm{init})})| = |D(e_-)| + k\}| \;\leqslant\; 2^{2L} \sum_{m=(k-2L)^+}^{k+2L} \binom{L^2}{m}, \qquad (5.47)$$

where $(k - 2L)^+ = \max\{k - 2L, 0\}$.

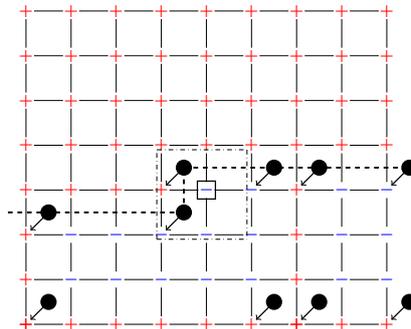*Proof.* Figure 7 and its caption give an informal proof; a formal proof follows.

Figure 7: Reconstruction of σ from an observed edge e. Defects in e_- are shown by black circles. The small black square indicates the site x where the spin is flipped in e. The four defects that are flipped by the edge e are contained in the dot-dashed box. The region above the thick dashed line is all $+$ under e_- (and in particular has no defects), and could have had any values in the initial configuration σ (so that the total number of defects is $|D(e_-)| + k$). The e_- and σ have the same spin configuration below the dashed line, and in particular have the same number of defects in this region excluding those on the dashed line, so there are no choices here. The dashed line intersects at most $(L + 2)$ plaquettes, and in principle any of these could contain defects in e_- that do not appear in σ (in particular there may be up to $(L + 2)$ extra defects here and thus 2^{L+2} choices).

More formally, let $z^{(1)}, z^{(2)}, \ldots$ be the elements of $\Lambda = [1 : L]^2$ in lexicographic order. Recall that $e_- = e_+^{z^{(i)}}$ for some $z^{(i)} = (z_1^{(i)}, z_2^{(i)}) \in \Lambda$ and $1 \leqslant i \leqslant |\Lambda|$. Let $\gamma \in G^{(\mathrm{na})}(e)$, and let $\sigma = \gamma^{(\mathrm{init})}$. We note that γ is determined entirely by σ (since the measure $F_\sigma^{(\mathrm{na})}$ puts all mass on a single path). Thus, it is enough to bound the possible choices of σ. We note the following about the possible choices for σ:

1. We must have $\sigma(z^{(j)}) = e_-(z^{(j)})$ for all $j \geqslant i$.
2. Knowledge of e gives no obvious restrictions on $\sigma(z^{(j)})$ for $j < i$.

This gives rise to the following restrictions on the locations of the defects of σ,

1. They must agree with the defects of e_- for all sites $\{z^{(k)} : k > i+L+1\}$. These are the sites below the boundary "strip" in Figure 7.
2. The defects in the set $\{z^{(k)} : k < i-1\}$ of sites above the boundary "strip" can be in any position.

There are also some complicated restrictions on the strip $\{z^{(k)} : i-1 \leqslant k \leqslant i+L+1\}$ itself, but we will ignore these restrictions (at the cost of over-counting the true number of possible configurations σ). Thus, to create a candidate compatible path starting at ancestor configuration σ with $|D(\sigma)| = |D(e_-)| + k$, we can:

1. Place defects in the set $\{z^{(k)} : k > i+L+1\}$ below the horizontal strip so that they agree with e_- (there is one way to do this).
2. Place defects on the horizontal strip $\{z^{(k)} : i-1 \leqslant k \leqslant i+L+1\}$ in any way (there are $(L+2)$ sites on the strip, so at most $2^{L+2} \leqslant 2^{2L}$ ways to do this).
3. Place the remaining defects anywhere in the set $\{z^{(k)} : k < i-1\}$ above the strip (there are between $(k-2L)^+$ and $k+2L$ defects to place after choosing the configuration on the strip, and so at most $\sum_{m=(k-2L)^+}^{k+2L} \binom{L^2}{m}$ ways to do this).

Not all candidate ancestors are in fact possible ancestors, but this gives the desired upper bound on the number of possible ancestors (and thus the same upper bound on the number of possible compatible paths). □

Next, we bound the number of compatible paths associated with a full path segment $\gamma_{\sigma,R}$ obtained from the function F in Proposition 5.20. We note that any such allowed path $\gamma_{\sigma,R}$ will certainly have $|D(\gamma_{\sigma,R}^{(\text{fin})})| = |D(\gamma_{\sigma,R}^{(\text{init})})| - k$ for $k \in \{2,4\}$, motivating the definition:

Definition 5.30 (Long Compatible Non-Naive Paths). *Fix $\eta \in \Omega_{[1:L]^2}$ and $k \in \{2,4\}$. We now define the set $G^{(\text{long})}(k,\eta) \subset \Gamma$ of long compatible typical paths associated with η as follows. Say $\gamma \in G^{(\text{long})}(k,\eta)$ if it satisfies all of the following:*

1. *$\gamma^{(\text{fin})} = \eta$, and*
2. *$|D(\gamma^{(\text{init})})| = |D(\eta)| + k$, and*
3. *There exists $\sigma \in \Omega_{[1:L]^2}$ such that $F_\sigma^{(\text{pt})}(\gamma) > 0$.*

We now bound the number of compatible paths for a given configuration:

Lemma 5.31 (Number of Long Compatible Non-Naive Paths). *For fixed $\eta \in \Omega_{[1:L]^2}$, we have the inequalities*

$$|G^{(\text{long})}(2,\eta)| \leqslant |D(\eta)|\,(L^2 - |D(\eta)|) \leqslant \frac{L^4}{4} \qquad (5.48)$$

$$|G^{(\text{long})}(4,\eta)| \leqslant (L^2 - |D(\eta)|)^2 \leqslant L^4$$

Proof. Let $\gamma \in G^{(\text{long})}(2,\eta)$. Note that, to reconstruct γ from η, it is enough to reconstruct the rectangle $R \subset [1:L]^2$ that satisfies $\gamma = \gamma_{\sigma,R}$ for some $\sigma \in \Omega_{[1:L]^2}$ (note that knowledge of R and η allows you to reconstruct σ uniquely). Let $r^{(\text{init})}$, $r^{(\text{fin})}$ be, respectively, the first and last elements of R that are flipped while traversing γ. Since only two defects are removed over the course of γ, we must have that $r^{(\text{fin})} \in D(\eta)$ and $r^{(\text{init})} \notin D(\eta)$. Thus, there are at most $(L^2 - |D(\eta)|)$ choices for $r^{(\text{init})}$ and at most $|D(\eta)|$ choices for $r^{(\text{fin})}$.

Since the choice of two opposite corners of R determines R completely, this completes the proof of the first line of Inequality (5.48). The proof of the second line is essentially identical. □

Having bounded the congestion of the path types, we bound the path lengths.

Lemma 5.32 (Path Length). *Fix* $\sigma \in \Omega_{[1:L]^2}$ *and let* $\gamma \in \Gamma_{\sigma,+}$ *satisfy* $F_\sigma(\gamma) > 0$. *Then the path length is bounded by*

$$|\gamma| \leqslant L^2 \min\left(\beta L + 1, \frac{|D(\sigma)|}{2}\right).$$

Proof. Recall that the (non-truncated) paths constructed in Definition 5.27 are given in terms of a decomposition of each path γ into subpaths:

$$\gamma = \gamma_1 \bullet \ldots \bullet \gamma_\ell$$

for some ℓ. Note that for all $i < \ell$, we have γ_i is of the form $\gamma_{\gamma_{i-1}^{(\mathrm{fin})}, R_i}$ for some rectangle R_i (with the convention $\gamma_0^{(\mathrm{fin})} = \sigma$). Recalling that $|D(\gamma_i^{(\mathrm{fin})})| \leqslant |D(\gamma_i^{(\mathrm{init})})| - 2$, this implies $\ell \leqslant \frac{|D(\sigma)|}{2}$. Recalling from Definition 5.27 that we start a naive path of the form in Definition 5.26 after at most βL rectangle-removal paths of the form Definition 5.6, this implies $\ell \leqslant \min(\beta L + 1, \frac{|D(\sigma)|}{2})$.

Recalling that each path component γ_i consists of flipping all the spins in a single rectangle $R_i \subset [1:L]^2$ at most once, we have $|\gamma_i| \leqslant |R_i| \leqslant |[1:L]^2| = L^2$. Thus,

$$|\gamma| \leqslant \sum_{i=1}^m |\gamma_i| \leqslant m L^2 \leqslant L^2 \min\left(\beta L + 1, \frac{|D(\sigma)|}{2}\right). \qquad \square$$

5.6 Final bounds on canonical paths

In this section, we put together our main bounds on the canonical paths studied in this paper. The calculation is somewhat lengthy, so we remind the reader of some conventions that are used throughout this section:

1. Throughout this section, we consider $\beta, L > 0$ with $L = L_c(\beta)$ and let $\Lambda = [1:L]^2$.

2. In this section only, we fix the constant $M = \lfloor \beta L \rfloor + 1$, one more than the number of path components taken before using the "naive" path, i.e. an upper bound on the total number of sub-paths used to construct each path to the ground state.

3. All paths that have positive measure under Definition 5.27 will consist of a sequence of paths

$$\gamma_1 \bullet \ldots \bullet \gamma_m, \tag{5.49}$$

   where each element γ_i has positive measure under Definition 5.16, possibly followed by a path of the form given in Definition 5.26. We will always write our paths in terms of this decomposition. When a single path has several decompositions of this form, we sum over all of these representations.

4. The measures $\{F_\sigma^{(\mathrm{pt})}\}$ are as in Definition 5.16, the measures $\{F_\sigma\}$, $\{F_\sigma^{(k)}\}$ are as in Definition 5.27, and the measures $\{F_\sigma^{(\mathrm{na})}\}$ are as in Definition 5.26. We note that $F_\sigma^{(\mathrm{na})}$ assigns full mass to a single path, which we denote $\gamma_{\sigma,\mathrm{na}}$.

5. For $\sigma \in \Omega$ and $e = (e_-, e_+) \in \mathbf{E}$, we define the energy associated with a path started at σ on edge e as $Q(\sigma, e) = \pi(\sigma)/\big(\pi(e_-)\mathcal{L}(e_-, e_+)\big)$.

We fix an edge $e \in \mathbf{E}$ and calculate a bound on the sum that appears in Lemma 5.2, with sets $S = S_k$ of the form (5.6) for some $k \in [0 : L^2]$. Although in principle we could take advantage of the fact that paths from S_k that are truncated when they enter S_k^c are shorter than paths which go all the way from S_k to $\{+\}$, we will not do so. However, we *will* use the fact that the initial state of any path from S_k to S_k^c will have at least k defects.

Recalling $M \lesssim \beta L$, we have:

$$\sum_{\sigma \in S_k} \sum_{\gamma \ni e} F_\sigma^{(k)}(\gamma) |\gamma| Q(\sigma, e)$$

$$\overset{\text{Lemma 5.32}}{\lesssim} \sum_{\sigma \in S_k} \sum_{\gamma \ni e} F_\sigma^{(k)}(\gamma) L^2 \, \min(|D(\sigma)|, \beta L) Q(\sigma, e)$$

$$\overset{\text{Eq. (5.49)}}{=} \sum_{m=1}^{M} \sum_{j=1}^{m} \sum_{\substack{\gamma = \gamma_1 \bullet \ldots \bullet \gamma_m \\ \gamma^{(\text{init})} \in S_k, \, \gamma_j \ni e}} F_{\gamma^{(\text{init})}}^{(k)}(\gamma) L^2 \, \min(|D(\gamma^{(\text{init})})|, \beta L) Q(\gamma^{(\text{init})}, e)$$

$$= L^2 \sum_{m=1}^{M} \sum_{j=1}^{\min(m, M-1)} \sum_{\substack{\gamma = \gamma_1 \bullet \ldots \bullet \gamma_m \\ \gamma^{(\text{init})} \in S_k, \, \gamma_j \ni e}} F_{\gamma^{(\text{init})}}^{(k)}(\gamma) \, \min(|D(\gamma^{(\text{init})})|, \beta L) Q(\gamma^{(\text{init})}, e)$$

$$+ L^2 \sum_{\substack{\gamma = \gamma_1 \bullet \ldots \bullet \gamma_M \\ \gamma^{(\text{init})} \in S_k, \, \gamma_M \ni e}} F_{\gamma^{(\text{init})}}^{(k)}(\gamma) \, \min(|D(\gamma^{(\text{init})})|, \beta L) Q(\gamma^{(\text{init})}, e)$$

$$\equiv L^2 S_{\text{init}} + L^2 S_{\text{naive}}. \tag{5.50}$$

We bound these last two terms separately, starting with S_{init}. As the subscripts in the following sums will become somewhat complicated, we introduce some short-term notation to reduce the visual clutter. In the following definitions, we always think of η as the configuration in γ that appears "just before" the part of the path containing the edge e, and $2u$ as the number of defects removed in going from $\gamma^{(\text{init})}$ to η in excess of the minimal number.

$$\Gamma^{(i)}(m) = \{\gamma_1 \bullet \ldots \bullet \gamma_m \in \Gamma \, : \, \gamma_i \ni e, \, \gamma_1^{(\text{init})}, \ldots, \gamma_i^{(\text{init})} \in S_k\} \quad \text{for } i \leqslant m,$$

$$\Gamma(m) = \Gamma^{(m)}(m) = \{\gamma_1 \bullet \ldots \bullet \gamma_m \in \Gamma \, : \, \gamma_m \ni e, \, \gamma_1^{(\text{init})}, \ldots, \gamma_m^{(\text{init})} \in S_k\},$$

$$\Gamma(m, u) = \{\gamma_1 \bullet \ldots \bullet \gamma_m \in \Gamma(m) \, : |D(\gamma^{(\text{init})})| = |D(\gamma_m^{(\text{init})})| + 2(m - 1 + u)\},$$

$$\Delta(\eta) = \{\gamma \in \Gamma_\eta \, : \, \gamma \ni e\},$$

$$h(m) = \min(|D(e_-)| + 4m, \beta L).$$

In this notation, we have omitted the dependence on many variables (e.g. the edge e and starting level k) that do not change in the calculation. We compute:

$$S_{\text{init}} = \sum_{m=1}^{M} \sum_{i=1}^{\min(m, M-1)} \sum_{\gamma \in \Gamma^{(i)}(m)} F_{\gamma^{(\text{init})}}^{(k)}(\gamma) \, \min(|D(\gamma^{(\text{init})})|, \beta L) \, Q(\gamma^{(\text{init})}, e) \tag{5.51}$$

$$= \sum_{m=1}^{M-1} \sum_{\gamma \in \Gamma(m)} \left(\prod_{j=1}^{m} F_{\gamma_j^{(\text{init})}}^{(\text{pt})}(\gamma_j) \right) \min(|D(\gamma^{(\text{init})})|, \beta L) \, Q(\gamma^{(\text{init})}, e),$$

where in the second line we took the marginal distribution of the first i segments of the path γ that is being summed over (integrating out the remaining segments from index $i + 1$ to m, which are not relevant to the other terms in the sum). Denote the inner sum

by:

$$S(m,e) \equiv \sum_{\gamma \in \Gamma(m)} \left(\prod_{j=1}^{m} F^{(\text{pt})}_{\gamma_j^{(\text{init})}}(\gamma_j) \right) \min(|D(\gamma^{(\text{init})})|, \beta L) \, Q(\gamma^{(\text{init})}, e) \,. \qquad (5.52)$$

We now decompose $S(m,e)$ over the possible values of u and factor out the term which is due to the final part of the path, and observe that $\min(|D(\gamma^{(\text{init})})|, \beta L) \leqslant h(m)$, to get

$$S(m,e) = \sum_{u=0}^{m-1} \sum_{\gamma \in \Gamma(m,u)} \left(\prod_{j=1}^{m} F^{(\text{pt})}_{\gamma_j^{(\text{init})}}(\gamma_j) \right) \min(|D(\gamma^{(\text{init})})|, \beta L) \, Q(\gamma_m^{(\text{init})}, e) e^{-2(m+u-1)\beta}$$

$$\leqslant \sum_{u=0}^{m-1} \sum_{\eta \in S_k} \sum_{\gamma \in \Delta(\eta)} F^{(\text{pt})}_{\eta}(\gamma) h(m) Q(\eta, e) e^{-2(m-1+u)\beta} \cdot \sum_{\gamma \in \Phi(m,u,\eta)} \left(\prod_{i=1}^{m-1} F^{(\text{pt})}_{\gamma_i^{(\text{init})}}(\gamma_i) \right),$$
$$(5.53)$$

where

$$\Phi(m,u,\eta) = \{\gamma_1 \bullet \dots \bullet \gamma_{m-1} : F^{(\text{pt})}_{\gamma_i^{(\text{init})}}(\gamma_i) > 0 \;\forall i \in [1:(m-1)], \; \gamma_1^{(\text{init})} \in S_k, \; \gamma_{m-1}^{(\text{fin})} = \eta,$$

$$|D(\gamma^{(\text{init})})| = |D(\eta)| + 2(m-1+u)\} \,.$$

To continue, we bound the inner part of the sum, $\sum_{\gamma \in \Phi(m,u,\eta)} \left(\prod_{i=1}^{m-1} F^{(\text{pt})}_{\gamma_i^{(\text{init})}}(\gamma_i) \right)$, in two cases:

1. $|D(\eta)| \leqslant \beta^{-1} L^2$. In this case we begin with the trivial bound $\prod_{i=1}^{m-1} F^{(\text{pt})}_{\gamma_i^{(\text{init})}}(\gamma_i) \leqslant 1$. Thus, it is enough to bound the size of the set $\Phi(m,u,\eta)$. Let $\mathbf{Bin}(u) = \{s \in \{0,1\}^{m-1} : \sum_{i=1}^{m-1} s_i = u\}$; for $s \in \mathbf{Bin}(u)$ define

   $$\Phi_s(m,u,\eta) = \{\gamma_1 \bullet \dots \bullet \gamma_{m-1} \in \Phi(m,u,\eta) :$$
   $$\forall i \in [1:(M-1)], \; |D(\gamma_i^{(\text{init})})| = |D(\gamma_i^{(\text{fin})})| + 2(1+s_i)\} .$$

   Since each component γ_i of a path in $\Phi(m,u,\eta)$ satisfies $|D(\gamma_i^{(\text{init})})| - |D(\gamma_i^{(\text{fin})})| \in \{2,4\}$, the sets $\{\Phi(u,\eta,s)\}_{s \in \mathbf{Bin}(u)}$ in fact partitions $\Phi(u,\eta)$:

   $$\Phi(m,u,\eta) = \sqcup_{s \in \mathbf{Bin}(u)} \Phi_s(m,u,\eta). \qquad (5.54)$$

   Since $m \leqslant \beta L$ and $|D(\eta)| \leqslant \beta^{-1} L^2$ we have $|D(\gamma_i^{(\text{fin})})| \lesssim \beta^{-1} L^2$ for each $i \in [1:m-1]$. So applying the first bounds in Lemma 5.31 we have

   $$|\Phi_s(m,u,\eta)| \lesssim L^{4(m-1)} \beta^{-m+1+u} \qquad (5.55)$$

   for all $s \in \mathbf{Bin}(u)$. Combining these bounds,

   $$\sum_{\gamma \in \Phi(m,u,\eta)} \left(\prod_{i=1}^{M-1} F^{(\text{pt})}_{\gamma_i^{(\text{init})}}(\gamma_i) \right) \leqslant |\Phi(m,u,\eta)| \overset{(5.54)}{\leqslant} \sum_{s \in \mathbf{Bin}(u)} |\Phi_s(m,u,\eta)|$$

   $$\overset{(5.55)}{\lesssim} L^{4(m-1)} \beta^{-m+1+u} \binom{m-1}{u}.$$

2. $|D(\eta)| > \beta^{-1} L^2$. The bound in this case is essentially the same. Since $|D(\eta)| > \beta^{-1} L^2$ is large, we can improve the trivial bound $\prod_{i=1}^{M-1} F^{(\text{pt})}_{\gamma_i^{(\text{init})}}(\gamma_i) \leqslant 1$, by applying

(5.23), to the (still very weak) bound $\prod_{i=1}^{m-1} F^{(\mathrm{pt})}_{\gamma_i^{(\mathrm{init})}}(\gamma_i) \leqslant \beta^{-m+1}$ (in fact the upper bound is closer to $(L/\beta)^{-m+1}$ but we don't use this additional strength). On the other hand, since $|D(\eta)|$ is large, applying Lemma 5.31 gives only the bound

$$|\Phi(u, \eta, s)| \leqslant L^{4(m-1)}, \tag{5.56}$$

so our upper bound is larger by a factor of $\beta^{m-(u+1)}$ relative to Inequality (5.55). Since we have improved by β^{-m+1} and worsened by a smaller factor of $\beta^{m-(u+1)}$, the conclusion

$$\sum_{\gamma \in \Phi(m,u,\eta)} \left(\prod_{i=1}^{m-1} F^{(\mathrm{pt})}_{\gamma_i^{(\mathrm{init})}}(\gamma_i) \right) \leqslant L^{4(m-1)} \beta^{-m+1+u} \binom{m-1}{u} \tag{5.57}$$

holds in this case as well.

Thus, Inequality (5.57) holds in both cases. Substituting Inequality (5.57) into (5.53) and continuing:

$$S(m, e) \lesssim \sum_{u=0}^{m-1} \binom{m-1}{u} \beta^{-(m-1-u)} L^{4(m-1)} \sum_{\eta \in S_k} \sum_{\gamma \in \Delta(\eta)} F^{(\mathrm{pt})}_\eta(\gamma) h(m) Q(\eta, e) e^{-2(m-1+u)\beta} \tag{5.58}$$

$$\lesssim \beta^{-(m-1)} \sum_{\eta \in S_k} \sum_{\gamma \in \Delta(\eta)} F^{(\mathrm{pt})}_\eta(\gamma) h(m) Q(\eta, e) \sum_{u=0}^{m-1} \binom{m-1}{u} (\beta e^{-2\beta})^u$$

$$\lesssim \beta^{-(m-1)} h(m) \sum_{\eta \in S_k \cap G(e)} \sum_{\gamma \in \Delta(\eta)} F^{(\mathrm{pt})}_\eta(\gamma) h(m) Q(\eta, e),$$

where on the second line we used $L^{4(m-1)} e^{-2(m-1)\beta} \lesssim 1$.

Next, we apply the congestion and energy bounds in Proposition 5.20. We note there are effectively two cases: $\eta \in \cup_{\theta \in \Theta} G^{(\mathrm{init})}_\theta(e) \cup G^{(\mathrm{mid})}_\theta(e)$ or $\eta \in \cup_{\theta \in \Theta} G^{(\mathrm{fin})}_\theta(e)$. In the first case, the congestion bound is slightly stronger while the energy bound is slightly weaker; in the second case, the congestion bound is slightly weaker while the energy bound is much stronger. Thus, in both cases, we continue from Inequality (5.58) to find:

$$S(m, e) \lesssim |\Theta| \beta^4 e^{2.5\beta} \beta^{-(m-1)} \frac{h(m)}{|D(e_-)|}. \tag{5.59}$$

Combining Inequalities (5.51) and (5.59), and noting that $h(m) \leqslant \min(|D(e_-)|, \beta L) + 4m$,

$$S_{\mathrm{init}} \lesssim \sum_{m=1}^{M-1} S(m, e) \lesssim \beta^6 e^{2.5\beta} \min(1, \frac{\beta L}{|D(e_-)|}). \tag{5.60}$$

We now bound the term S_{naive} by a similar factorisation as we did for S_{init}. We replace the previous simplifying notation with the following,

$$\Gamma_{\mathrm{init}}(\eta) = \{\gamma_1 \bullet \ldots \bullet \gamma_{M-1} \in \Gamma : \gamma_1^{(\mathrm{init})}, \ldots, \gamma_{M-1}^{(\mathrm{init})} \in S_k, \ \gamma_{M-1}^{(\mathrm{fin})} = \eta\}, \tag{5.61}$$

$$\Omega_{\mathrm{na}} = \{\eta \in S_k : \gamma_{\eta,\mathrm{na}} \ni e\},$$

$$\Omega(v) = \{\eta \in S_k : |D(\eta)| = |D(e_-)| + 2v, \gamma_{\eta,\mathrm{na}} \ni e\}.$$

In the following we now think of η as the configuration in γ that appears at the beginning of the naive part of the path (which contains e), and $2v$ as the number of defects removed from η to e_-, which could be negative.

$$S_{\mathrm{naive}} = \sum_{\substack{\gamma = \gamma_1 \bullet \ldots \bullet \gamma_M \\ \gamma^{(\mathrm{init})} \in S_k, \gamma_M \ni e}} F^{(k)}_{\gamma^{(\mathrm{init})}}(\gamma) \ \min(|D(\gamma^{(\mathrm{init})})|, \beta L) Q(\gamma^{(\mathrm{init})}, e) \tag{5.62}$$

$$= \sum_{\eta \in \Omega_{\mathrm{na}}} \sum_{\gamma \in \Gamma_{\mathrm{init}}(\eta)} \left(\prod_{i=1}^{M-1} F^{(\mathrm{pt})}_{\gamma_i^{(\mathrm{init})}}(\gamma_i) \right) \min(|D(\gamma^{(\mathrm{init})})|, \beta L) Q(\gamma^{(\mathrm{init})}, e)$$

Now, defining $\Gamma_{\mathrm{init}}(\eta, x) = \{\gamma_1 \bullet \ldots \bullet \gamma_{m-1} \in \Gamma : \gamma_1^{(\mathrm{init})} \in S_k, \ \gamma_{m-1} \ni (\eta^x, \eta)\}$, we observe that $\Gamma_{\mathrm{init}}(\eta) \subset \cup_{x \in \Lambda} \Gamma_{\mathrm{init}}(\eta, x)$. Furthermore, using reversibility and that the rates are bounded by 1, we have

$$Q(\gamma^{(\mathrm{init})}, e)/Q(\gamma^{(\mathrm{init})}, (\eta^x, \eta)) \ \leqslant \ Q(\eta, e) \, ,$$

it follows that

$$S_{\mathrm{naive}} \leqslant \sum_{\eta \in \Omega_{\mathrm{na}}} Q(\eta, e) \sum_{x \in \Lambda} \underbrace{\sum_{\gamma \in \Gamma_{\mathrm{init}}(\eta, x)} \left(\prod_{i=1}^{M-1} F^{(\mathrm{pt})}_{\gamma_i^{(\mathrm{init})}}(\gamma_i) \right) \min(|D(\gamma^{(\mathrm{init})})|, \beta L) Q(\gamma^{(\mathrm{init})}, (\eta^x, \eta))} \, ,$$

where the under-braced term is exactly $S(M, (\eta^x, \eta))$ defined in Eq. (5.52). Applying Ineq. (5.59) and summing over the possible number of defects removed in going from η to e_-,

$$S_{\mathrm{naive}} \lesssim \beta^6 e^{8\beta} \beta^{-(M-1)} \sum_{v \, \geqslant \, -L-2} \sum_{\eta \in \Omega(v)} e^{-2v\beta} \, ,$$

where we applied the bounds $h(m)/|D(e_-)| \ \leqslant \ \beta L$ and $\mathcal{L}(e_-, e_+) \ \geqslant \ e^{-4\beta}$. Finally, applying Lemma 5.29

$$S_{\mathrm{naive}} \lesssim \beta^6 e^{8\beta} \beta^{-(M-1)} \sum_{2v \, \geqslant \, -2L} 2^{2L} \sum_{k=(2v-2L)^+}^{2v+2L} \binom{L^2}{k} e^{-2v\beta}$$

$$\leqslant \beta^6 e^{8\beta} \beta^{-(M-1)} \sum_{2v \, \geqslant \, -2L} 2^{2L} (4L+1) \frac{L^{2(2v+2L)}}{(2v-2L)^+!} e^{-2v\beta}$$

$$\lesssim \beta^6 e^{8.5\beta} \beta^{-(M-1)} 2^{2L} e^{2L\beta} \sum_{n \, \geqslant \, 0} \frac{(L^2 e^{-\beta})^n}{(n-4L)^+!}$$

$$\overset{L^2 e^{-\beta} \approx 1}{\lesssim} \beta^6 e^{9\beta} \beta^{-(M-1)} 2^{2L} e^{2L\beta} \, . \tag{5.63}$$

Since $M = \Theta(\beta L)$ and $L = \Theta(e^{0.5\beta})$, we observe that the term $L^2 S_{\mathrm{naive}}$ is negligible compared to $L^2 S_{\mathrm{init}}$.

Summarizing the calculations in this section, we have by Inequalities (5.50), (5.63) and (5.60)

$$\sum_{\sigma \in S_k} \sum_{\gamma \ni e} F^{(k)}_{\sigma}(\gamma) |\gamma| \frac{\pi(\sigma)}{\pi(e_-) \mathcal{L}(e_-, e_+)} \lesssim L^2 (S_{\mathrm{init}} + S_{\mathrm{naive}}) \tag{5.64}$$

$$\lesssim L^2 \left(\beta^6 e^{2.5\beta} \min(1, \frac{\beta L}{|D(e_-)|}) + L^{4M+1} e^{(9-2M)\beta} 2^{2L} \beta^{-M} \right)$$

$$\lesssim \beta^6 e^{3.5\beta} \min(1, \frac{\beta L}{|D(e_-)|}) \, ,$$

for any edge e with $e_- \in S_k$.

6 Analysis of all-plus boundary condition: proof of Theorem 3.1

6.1 Upper bounds

Applying Lemma 5.2 with $S = \{\sigma \in \Omega_\Lambda : |D(\sigma)| = 0\}$ and the usual variational characterization of the spectral gap given in Equation (2.13),

$$T^+_{\text{rel}}(L_c) \leqslant \mathcal{A},$$

where $\mathcal{A}$ is defined as in Equation (5.3). By Inequality (5.64),

$$\mathcal{A} \lesssim \beta^6 \, e^{3.5\beta}.$$

Combining these two bounds completes the proof of the upper bound in Inequality (3.1).

Next, we prove the upper bound on the mixing time. Define the function $k : [0,1] \mapsto \mathbb{N}$ as in Equality (5.5), and for $r > 0$ let $S_{k(r)}$ be as in Equation (5.6). Applying Inequality (5.8), we have by Lemma 5.2 and the bound on $\mathcal{A}$ given in Inequality (5.64) that

$$
\begin{aligned}
\lambda^{-1}(S_{k(r)}) &\lesssim \beta^6 \, e^{3.5\beta}, & e^{-\beta L_c} &< r, \\
\lambda^{-1}(S_{k(r)}) &\lesssim \beta^8 \, e^{4\beta} \frac{-1}{\log(r)}, & 0 &< r < e^{-\beta L_c},
\end{aligned}
\tag{6.1}
$$

where we recall from Lemma 4.4 that $\pi(+) \approx 1$, and $|D(e_-)| \geqslant k(r)$ for any edge e with $e_- \in S_{k(r)}$. Applying Inequality (5.4) with the bound from Inequality (6.1), we conclude

$$
\begin{aligned}
T^+_{\text{mix}}(L_c) &\overset{\text{Ineq (5.4)}}{\leqslant} \int_{4 \min_\sigma \pi(\sigma)}^{16} \frac{2}{x \, \Lambda_{\mathcal{L}^+_\Lambda}(x)} dx \\
&\overset{\text{Ineqs (5.7),(6.1)}}{\lesssim} 2\beta^8 \, e^{4\beta} \int_{4e^{-L_c^2\beta}}^{e^{-L_c\beta}} \frac{1}{x \log(1/x)} dx + 2\beta^6 e^{3.5\beta} \int_{e^{-L_c\beta}}^{16} \frac{1}{x} dx \lesssim \beta^9 \, e^{4\beta},
\end{aligned}
$$

completing the proof of the upper bound in Inequality (3.2).

6.2 Lower bound on relaxation time

We give a lower bound on the relaxation time by choosing an explicit test function that relaxes to equilibrium quite slowly, and explicitly bounding the two terms in the ratio (2.13). We choose a test function motivated by the heuristic that the relaxation time is dominated by the time it takes for 4 initially well seperated defects, at the corners of a rectangle, to annihilate. As illustrated in Figure 2, these opposite corners perform nearly-independent symmetric random walk until they are in adjacent columns or rows, and so our test function will be quite similar to the optimal test function for symmetric random walk in $[0 : L_c]^2$.

Throughout the proof, to reduce notation we use $\pi = \pi^+_\Lambda$, $c = c^+_\Lambda$, $p = p^+$, $L = L_c$ and $\Lambda = [1 : L]^2$. Let $f : \Omega^+_\Lambda \to \mathbb{R}_+$ be given by

$$f(\sigma) = g\left(\frac{|\sigma|}{L^2}\right),\tag{6.2}$$

where $g : [0,1] \to [0,1]$ is defined for $x \in [0,1/2]$ by

$$
\begin{cases}
g(x) = 0 & \text{if } x \in [0, 1/4], \\
g(x) = 12x - 3 & \text{if } x \in (1/4, 1/3) \\
g(x) = 1 & \text{if } x \in [1/3, 1/2],
\end{cases}
\tag{6.3}
$$

and $g(x) = g(1-x)$ for $x \in [1/2, 1]$.

To estimate $\mathrm{Var}_\Lambda^+(f)$ from below, it will turn out to be enough to compare the $+$ configuration to configurations consisting of exactly 4 defects at the corners of a rectangle (i.e. configurations with a single rectangular region of -1 spins, and all $+1$ spins outside). Informally, there are $\Theta(L^4) = \Theta(e^{2\beta})$ rectangles of area $\Omega(L^2)$, and by Lemma 4.4 and Equality (2.7) the probability of any particular such configuration is $\Theta(e^{-4\beta})$. It will follow from these estimates on the number and probability of the "large-rectangle" configurations that the $\mathrm{Var}_\Lambda^+(f) \gtrsim e^{-2\beta}$, as shown below.

More formally, let $\mathbf{B}$ be the collection of rectangles $R \subset [1:L]^2$ with areas $\frac{L^2}{3} \leqslant |R| \leqslant \frac{2L^2}{3}$, and define

$$\mathcal{C} = \{\sigma \in \Omega_\Lambda^+ \,:\, \exists R \in \mathbf{B} \text{ s.t. } \sigma_x \equiv 1 - 2\mathbb{1}_{x \in R}\}$$

to be the rectangles with spin (-1) exactly on a rectangle in $\mathbf{B}$.

Observe that $|\mathcal{C}| \geqslant cL^4$ for some $c > 0$. Since $f(\sigma) \geqslant 1$ for all $\sigma \in \mathcal{C}$ and $f(+) = 0$, our bound on $|\mathcal{C}| \geqslant cL^4$ gives the following bound on $\mathrm{Var}_\Lambda^+(f)$:

$$\mathrm{Var}_\Lambda^+(f) = \frac{1}{2} \sum_{\sigma, \eta \in \Omega_\Lambda} \pi(\sigma)\pi(\eta)\left(f(\sigma) - f(\eta)\right)^2$$

$$\geqslant \frac{1}{2}\pi(+) \sum_{\sigma \in \mathcal{C}} \pi(\sigma)f^2(\sigma) \geqslant \frac{1}{2}\pi(+)^2 \sum_{\sigma \in \mathcal{C}} e^{-4\beta} \gtrsim e^{-2\beta}.$$

To estimate the Dirichlet form from above, we consider various cases defined by the number of defects present in the initial and final configuration. Let

$$\mathcal{A}_n = \{\eta \in \Omega_\Lambda \,:\, |D(\eta)| = n\},$$
$$\mathcal{B}_n = \mathcal{A}_n \cup \mathcal{A}_{n-2} \cup \mathcal{A}_{n-4},$$

where $\mathcal{A}_n$ is non-empty only for $n = 0$ or $n \geqslant 4$ and n even. By reversibility, we may consider transitions which only remove defects (reducing the number by 0, 2 or 4), or leave the number of defects the same, so

$$\mathcal{D}_\Lambda^+(f) \leqslant \sum_{\sigma \in \Omega} \sum_{\substack{x \in \Omega_\Lambda \,: \\ |D(\sigma)| \geqslant |D(\sigma^x)|}} \pi(\sigma)c(x, \sigma)\left(\nabla_x f(\sigma)\right)^2$$

$$= \sum_{k=2}^{L^2/2} \sum_{\sigma \in \mathcal{A}_{2k}} \sum_{\substack{x \,: \\ \sigma^x \in \mathcal{B}_{2k}}} \pi(\sigma)c(x, \sigma)\left(\nabla_x f(\sigma)\right)^2. \tag{6.4}$$

It remains to bound this sum from above. We divide the contributions to the sum in (6.4) into 5 cases. In the Cases $1, 4$ and 5 we will use the following general bound on $\nabla_x f(\sigma)$. For all $\sigma \in \Omega_\Lambda$, we have [4]

$$\left(\nabla_x g(\sigma)\right)^2 = \left(g\left(\frac{|\sigma|}{L^2}\right) - g\left(\frac{|\sigma| \pm 1}{L^2}\right)\right)^2$$

$$= g'\left(\frac{|\sigma|}{L^2}\right)^2 \frac{1}{L^4} + o\left(\frac{1}{L^4}\right) \lesssim e^{-2\beta}.$$

Case 1 ($\sigma \in \bigcup_{k=4}^{L^2/2} \mathcal{A}_{2k}$): We will show that the contribution to the sum in (6.4) due to configurations σ with 8 or more defects are negligible. For $\sigma \in \mathcal{A}_{2k}$ we have $\pi(\sigma) =$

---

[4]Note that $g'(x)$ does not exist at some isolated points x. To ensure that the following calculation still holds, we define $g'(x) = \lim_{y \uparrow x} g'(y)$ at these points. Note that essentially any sensible choice for these isolated points will not change the final estimate.

$\pi(+)e^{-2k\beta}$, also $c(x,\sigma) \leqslant 1$ and $|\{x \in \Lambda \,:\, \sigma^x \in \mathcal{B}_{2k}\}| \leqslant 8k$. Combining these bounds with Lemma 4.3, it follows that

$$\sum_{k=4}^{\infty} \sum_{\sigma \in \mathcal{A}_{2k}} \sum_{\sigma^x \in \mathcal{B}_{2k}}^{x\,:} \pi(\sigma)c(x,\sigma)\left(\nabla_x f(\sigma)\right)^2 \lesssim \left(L^8 e^{-8\beta} + L^{12} e^{-12\beta} + \sum_{k=8}^{\infty} k L^{3k} e^{-2k\beta} \right) e^{-2\beta}$$

$$\lesssim \left(e^{-4\beta} + e^{-4\beta}\left(8 - 7e^{-\beta/2}\right) \right) e^{-2\beta} \lesssim e^{-6\beta},$$

where we use the bound $|\mathcal{A}_{2k}| \leqslant (ek)^{2k} L^{2k}$ from Lemma 4.3 to bound the two terms corresponding to $k=4$ and $k=6$, and we use the bound $|\mathcal{A}_{2k}| \leqslant L^{3k}$ from the same lemma to bound the remaining terms with $k \geqslant 8$.

Case 2 ($\sigma \in \mathcal{A}_4$ and $\sigma^x \in \mathcal{A}_0$): If a single spin flip at x removes all defects, then $\sigma_x = -1$ and $\sigma_y = 1$ for $y \in \Lambda \setminus \{x\}$. Therefore $\left(\nabla_x f(\sigma)\right)^2 \leqslant \left(g(1/L^2)\right)^2 = 0$ for $L > 2$ by (6.3). It follows that there is no contribution to the sum in (6.4) from this case.

Case 3 ($\sigma \in \mathcal{A}_4$ and $\sigma^x \in \mathcal{A}_4$): By a similar argument as in Case 2, these transitions do not contribute to the sum for $L > 4$. Observe that, if σ and σ^x have the same number of defects, this means that there must be exactly two defects in the plaquette associated with x – that is, $|D(\sigma) \cap \{x, x-e_1, x-e_2, x-e_1-e_2\}| = 2$. Since $\sigma \in \mathcal{A}_4$, this fact (and the parity Lemma 4.1) implies that the collection $\{y \in [1:L]^2 \,:\, \sigma_y = -1\}$ of sites with spin -1 must be a rectangle of width or height 1. In particular, $|\sigma| \leqslant L$ and so $\left(\nabla_x f(\sigma)\right)^2 \leqslant 4\left(g(1/L)\right)^2 = 0$ for $L > 4$ by (6.3).

Case 4 ($\sigma \in \mathcal{A}_6$ and $\sigma^x \in \mathcal{A}_4$): The only spin flips which remove 2 defects are those in which the initial state has 3 neighboring defects which are replaced by a single defect (see the first transition in Figure 1, which is a generic transition in this sense). It follows that $|\{\sigma \in \mathcal{A}_6 \,:\, \exists x \in \Lambda \text{ with } \sigma^x \in \mathcal{A}_4\}| \lesssim |\mathcal{A}_4|$. By Lemma 4.3, $|\mathcal{A}_4| \lesssim L^4$. Also for any $\sigma \in \mathcal{A}_6$ the number of sites at which a spin flip will reduce the number of defects is trivially bounded from above by 24 (the total number of sites that can contain a defect in any of their associated plaquettes). It follows that

$$\sum_{\sigma \in \mathcal{A}_6} \sum_{\sigma^x \in \mathcal{A}_4}^{x\,:} \pi(\sigma)c(x,\sigma)\left(\nabla_x f(\sigma)\right)^2 \lesssim e^{2\beta} e^{-8\beta} \pi(+) \lesssim e^{-6\beta}. \tag{6.5}$$

The following case is the only one which gives a non-negligible contribution to our upper bound on the Dirichlet form.

Case 5 ($\sigma \in \mathcal{A}_6$ and $\sigma^x \in \mathcal{A}_6$): As in Case 3, σ has at least two neighbouring defects. By the same argument as in Lemma 4.3, we have $|\{\sigma \in \mathcal{A}_6 \,:\, \exists x \in \Lambda \text{ with } \sigma^x \in \mathcal{A}_6\}| \lesssim L^5$. Furthermore, any spin flip at a site which is not a neighbour of a defect must change the number of defects, so $|\{x \in \Lambda \,:\, |p(\sigma)| = |p(\sigma^x)|\}| \leqslant 24$. Finally for $\sigma \in \mathcal{A}_6$ and $\sigma^x \in \mathcal{A}_6$ we have $\pi(\sigma) = \pi(+)e^{-6\beta}$ and $c(x,\sigma) = 1$, it follows that

$$\sum_{\sigma \in \mathcal{A}_6} \sum_{\sigma^x \in \mathcal{A}_6}^{x\,:} \pi(\sigma)c(x,\sigma)\left(\nabla_x f(\sigma)\right)^2 \lesssim e^{-2\beta} e^{(5/2)\beta} e^{-6\beta} \pi(+) \approx e^{-2\beta}\, e^{-(7/2)\beta}. \tag{6.6}$$

Combining the bounds from Cases 1–5 above with Inequality (6.4) we have,

$$T_{\text{rel}}^+(L_c) \geqslant \frac{\text{Var}_\pi^+(f)}{\mathcal{D}_\Lambda^+(f)} \gtrsim e^{(7/2)\beta}. \tag{6.7}$$

7 Analysis of periodic boundary condition: proof of Theorem 3.2

Throughout this section, we often reserve subscripts for a time index and use the "bracket" notation $x[i] \equiv x_i$ to indicate an element of a vector $x = (x_1, \ldots, x_k)$. In this section, we have many explicit probabilistic calculations related to Markov processes

$\{X_t\}_{t \geqslant 0}$ running according to some generator and with starting points $X_0 = x$. When we wish to emphasize the starting point of a process in such a calculation, we use subscripts, as in *e.g.* $\mathbb{P}_x[X_t \in S]$, $\mathbb{E}_x[f(X_t)]$.

7.1 Lower bound

We prove the lower bound in Theorem 3.2 by constructing and analyzing a test function.

We begin by giving the heuristic that guides our proof of this inequality. Roughly speaking, we expect the relaxation time of the dynamics of the spin model $\{X_t\}_{t \geqslant 0}$ to be bounded below by the relaxation time of the dynamics of the trace $\{\hat{X}_t\}_{t \geqslant 0}$ of $\{X_t\}_{t \geqslant 0}$ on the set $\mathcal{G} = \{X \in \Omega_\Lambda : H_\Lambda^{\mathrm{per}}(X) = \min_{\eta \in \Omega_\Lambda} H_\Lambda^{\mathrm{per}}(\eta)\}$ of ground states. Furthermore, this "trace walk" $\{\hat{X}_t\}_{t \geqslant 0}$ behaves very much like simple random walk on the hypercube $\{-1, +1\}^{2L-1}$. This can be made precise in the following way: there is a natural bijection $w : \mathcal{G} \mapsto \{-1, +1\}^{2L-1}$ (see Definition 7.1 below) under which the dynamics of the trace walk $\{w(\hat{X}_t)\}_{t \geqslant 0}$ are a very small perturbation of the usual simple random walk on the hypercube $\{-1, +1\}^{2L-1}$, with one added transition that flips all signs. Thus, we expect that a test function based on an eigenfunction for the kernel of simple random walk on the hypercube with largest eigenvalue, the usual "Hamming weight" function, should be able to pick out the slowest mixing behaviour of $\{X_t\}_{t \geqslant 0}$.

We now make this heuristic precise. We begin by defining the bijection $w : \mathcal{G} \mapsto \{-1, +1\}^{2L-1}$, used to identify the ground states, as well as a collection of functions that will allow us to extend a test function on $\mathcal{G}$ to a test function on all of Ω_Λ. Note, that to avoid excessive subscripts we use the notation $\sigma[i, j]$ to denote the $(i, j)^{\mathrm{th}}$ component of σ, previously denoted $\sigma_{(i,j)}$. In words, w assigns a -1 to all the rows and columns of a ground state that are flipped with respect to the $+$ configuration, and $+1$ to all the others, where the columns are labeled 1 to L (left to right), and the rows $L + 1$ to $2L - 1$ (bottom to top). The top row of spins are fully specified by all the other rows and columns.

We give the full notation required to define our test function in the following definition; the informal description of this notation in Remark 7.2 is sufficient to follow the remainder of the argument:

Definition 7.1 (Preliminary Notation for Test Function). *We define a bijection $w : \mathcal{G} \mapsto \{-1, +1\}^{2L-1}$ between the collection of ground states and the $(2L - 1)$-hypercube by the following formula for its inverse:*

$$w^{-1}(v)[i, j] = v[i]v[L + j], \qquad i \in [1 : L], \, j \in [1 : (L - 1)] \tag{7.1}$$
$$w^{-1}(v)[i, L] = v[i], \qquad i \in [1 : L].$$

For $v \in \{-1, 1\}^{2L-1}$, define $|v| = \sum_{i=1}^{2L-1} \mathbb{1}_{v[i]=1}$ to be the Hamming weight of v. We also denote by $\prec$ some fixed total order on $\{-1, +1\}^{2L-1}$ that extends the usual Hamming partial order.

For $m \in [1 : L]$, $k \in [1 : L]$, and $\sigma, \eta \in \mathcal{G}$ with $w(\sigma), w(\eta)$ differing at the single index $\ell \in [1 : L]$ and with $w(\sigma) \prec w(\eta)$, define the configuration $R(\sigma, \eta, m, k)$ by

$$R(\sigma, \eta, m, k)[\ell, j] = \eta[\ell, j], \qquad j \in [m : (m + k - 1)]$$
$$R(\sigma, \eta, m, k)[i, j] = \sigma[i, j], \qquad \text{all other entries,}$$

where in this case the "interval" $[m : (m + k - 1)]$ is defined modulo L (so that, if $L = 9$, we have $[7 : 2] = \{7, 8, 9, 1, 2\}$).

Similarly, for σ, η with $w(\sigma), w(\eta)$ differing at the single index $\ell \in [(L + 1) : (2L - 1)]$ and with $w(\sigma) \prec w(\eta)$, define the configuration $R(\sigma, \eta, k, m)$ by

$$R(\sigma, \eta, m, k)[i, \ell] = \eta[i, \ell - L], \qquad i \in [m : (m + k - 1)]$$

$$R(\sigma, \eta, m, k)[i, j] = \sigma[i, j], \qquad \textit{all other entries.}$$

Finally, for σ, η with $w(\sigma) = -w(\eta)$ and $w(\sigma) \prec w(\eta)$, define the configuration $R(\sigma, \eta, m, k)$ by

$$R(\sigma, \eta, m, k)[i, L] = \eta[i, L], \qquad i \in [m : (m + k - 1)]$$
$$R(\sigma, \eta, m, k)[i, j] = \sigma[i, j], \qquad \textit{all other entries.}$$

For convenience, when $w(\sigma) \prec w(\eta)$ we let $R(\sigma, \eta, m, 0) = \sigma$, and when $w(\eta) \prec w(\sigma)$ we define $R(\sigma, \eta, m, k) \equiv \emptyset$ for each $k \in [0 : L]$.

We note that, if you restrict the range of k to the set $[1 : L]$, then R is an injective map[5] – that is, if $R(\sigma, \eta, m, k) = R(\sigma', \eta', m', k')$ for $k, k' \in [2 : L-1]$, then $\sigma = \sigma'$, $\eta = \eta'$, $m = m'$ and $k = k'$. For $k \in [1 : L]$, let

$$\mathcal{R}_k = \{\zeta \in \Omega_{[1:L]^2} : \exists \sigma, \eta, m \ \ \text{s.t.} \ \ \zeta = R(\sigma, \eta, m, k)\}, \tag{7.2}$$

so that $\mathcal{R} \equiv \cup_{k=1}^{L} \mathcal{R}_k$ is the collection of points along these minimal-length and -energy paths between ground states. Note that, for $k, k' \in [2 : (L-2)]$, $\mathcal{R}_k \cap \mathcal{R}_{k'} = \emptyset$. Thus, for $\zeta \in \cup_{k=2}^{L-1} \mathcal{R}_k$, let $\sigma(\zeta), \eta(\zeta), m(\zeta), k(\zeta)$ be the unique elements that satisfy

$$\zeta = R(\sigma(\zeta), \eta(\zeta), m(\zeta), k(\zeta)). \tag{7.3}$$

When $\zeta \in \mathcal{R}_1 \cup \mathcal{R}_{L-1}$, there may be several choices that satisfy Equation (7.3), since these configurations have a single spin flipped relative to an element of the ground state; this single spin might be the first or last spin to be flipped in a path that is flipping a row or column; in total there can be up to four choices. Denote by $\phi(\zeta)$ the number of these choices, and denote the choices themselves by $(\sigma_1(\zeta), \ldots, k_1(\zeta)), \ldots, (\sigma_{\phi(\zeta)}(\zeta), \ldots, k_{\phi(\zeta)}(\zeta))$, in any arbitrary but fixed order.

Remark 7.2. For fixed $m \in L$, the sequence $\{R(\sigma, \eta, m, k)\}_{k=0}^{L}$ gives one of many minimal-length paths from σ to η: that is, $R(\sigma, \eta, m, 0) = \sigma$, $R(\sigma, \eta, m, L) = \eta$, and $R(\sigma, \eta, m, k)$ is obtained by flipping k adjacent spins of the column on which σ, η disagree. Roughly speaking: m identifies where in the column we start "flipping" from σ to η, and k tells us "how far" we are in the path from σ to η. See Figure 8 for a sample pair of ground states and an element along the path between them.

We will see that the paths of the form $(R(\sigma, \eta, m, k))_{k=0}^{L}$ describe the only high-probability paths between pairs of "neighboring" ground states (differ by one coordinate of w). This allows us to lift test functions from $\{-1, +1\}^{2L-1}$ to similar test functions on $\Omega_{[1:L]^2}$. We can now define a simple test function:

Definition 7.3 (Test Function). *Define $\hat{g} : \{-1, +1\}^{2L-1} \mapsto \mathbb{R}$ by*

$$\hat{g}(u) = \left| |u| - \frac{2L-1}{2} \right|.$$

Define $g : \Omega_\Lambda \mapsto \mathbb{R}$ by

$$g(\zeta) = \begin{cases} \hat{g}(w(\zeta)), & \zeta \in \mathcal{G}, \\ \frac{k(\zeta)}{L} \hat{g}(w(\eta(\zeta))) + (1 - \frac{k(\zeta)}{L}) \hat{g}(w(\sigma(\zeta))), & \zeta \in \cup_{k=2}^{L-2} \mathcal{R}_k, \\ \frac{k_1(\zeta)}{L} \hat{g}(w(\eta_1(\zeta))) + (1 - \frac{k_1(\zeta)}{L}) \hat{g}(w(\sigma_1(\zeta))), & \zeta \in \mathcal{R}_1 \cup \mathcal{R}_{L-1}, \\ 0, & \zeta \notin \mathcal{R}. \end{cases}$$

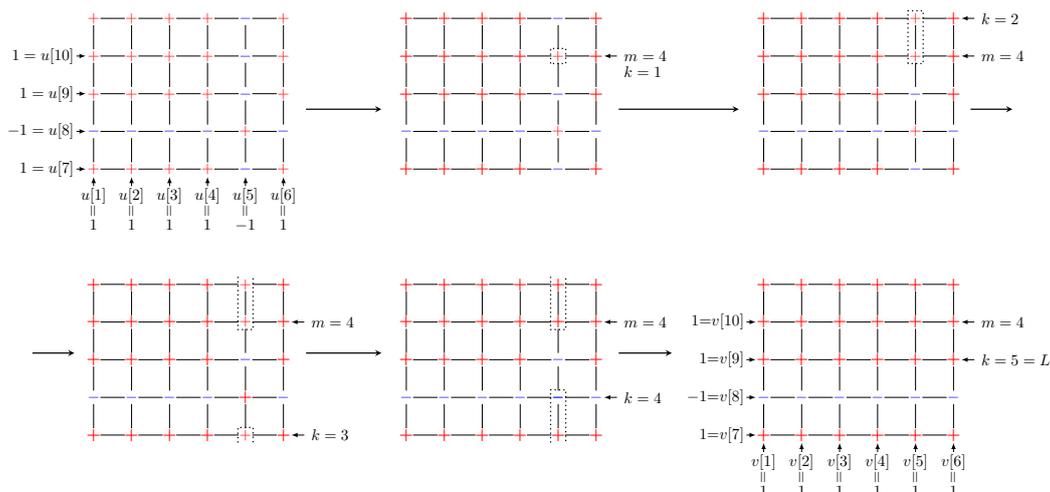[5]Note that obtaining this injectivity is the reason we insist that $w(\sigma) \prec w(\eta)$.

Figure 8: A minimal-length path $\{R(\sigma, \eta, m, k)\}_{k=0}^{L}$ between two ground states $\sigma = w^{-1}(u)$ and $\eta = w^{-1}(v)$, where u and v are given in the first and final frames.

This is well-defined for $L > 4$, since $\mathcal{R}_1 \cup \mathcal{R}_{L-1}$ is disjoint from $\cup_{k=2}^{L-2} \mathcal{R}_k$ (the former are obtained by flipping exactly 1 spin from a ground state; the latter all differ from any ground state by at least 2 spins). Note that this test function essentially ignores the behaviour of $\{X_t\}_{t \geqslant 0}$ except on the collection of ground states and minimal-length paths between ground states. On the minimal-length paths between ground states, g linearly interpolates the values of $\hat{g}$, with the possibility of an error of $\frac{2}{L}$ for configurations that are the first or last configuration along a minimal-length path.

We are primarily interested in the following property of g: if $\zeta_1, \zeta_2 \in \mathcal{R}$ and $\mathcal{L}_\Lambda^{\mathrm{per}}(\zeta_1, \zeta_2) \neq 0$, then

$$|g(\zeta_1) - g(\zeta_2)| \leqslant \frac{3}{L}; \tag{7.4}$$

note that we require a factor of 3 (rather than 1) in this inequality to deal with the possibility that ζ_1 or ζ_2 are in the set $\mathcal{R}_1 \cup \mathcal{R}_{L-1}$.

We now compare the Dirichlet form and variance of the test function g given in Definition 7.3, restricting our attention to pairs β, L with $L = L_c(\beta)$. We begin with a lower bound on the variance. If $X \sim \pi_\Lambda^{\mathrm{per}}(\cdot)$, and $Y \sim \pi_\Lambda^{\mathrm{per}}(\cdot \mid Y \in \mathcal{G})$ is distributed uniformly on the ground states, then

$$\begin{aligned} \mathrm{Var}_\Lambda^{\mathrm{per}}(g) &= \mathbb{E}[g(X)^2] - \mathbb{E}[g(X)]^2 \\ &\geqslant \pi(\mathcal{G}) \times (\mathbb{E}[g(Y)^2] - \mathbb{E}[g(Y)]^2) \\ &\gtrsim (\mathbb{E}[g(Y)^2] - \mathbb{E}[g(Y)]^2) \gtrsim L, \end{aligned} \tag{7.5}$$

where the second-last inequality uses the fact that $\pi(\mathcal{G}) \gtrsim 1$ from Lemma 4.4 and the last inequality uses the fact that $g(Y) \sim \mathrm{Bin}(2L - 1, 0.5)$.

We next calculate an upper bound on the Dirichlet form, still restricting our attention to the case $L = L_c$. We observe that if σ, η are two ground states with $w(\sigma) = -w(\eta)$, then $g(\sigma) = g(\eta)$, and so these minimal-length transitions do not contribute to the following calculation. Using reversibility of $\mathcal{L}_\Lambda^{\mathrm{per}}$ and the fact that $\mathcal{L}_\Lambda^{\mathrm{per}}(\sigma, \eta) = 0$ for all pairs $\sigma \in \mathcal{G}$,

$\eta \notin \mathcal{R}$, we have

$$2\mathcal{D}_\Lambda^{\mathrm{per}}(g) = \sum_{\sigma,\eta \in \Omega_\Lambda} (g(\sigma) - g(\eta))^2 \pi_\Lambda^{\mathrm{per}}(\sigma)\mathcal{L}(\sigma,\eta) \tag{7.6}$$

$$\leqslant \sum_{\sigma \in \mathcal{G}} \sum_{\eta \in \mathcal{R}} (g(\sigma) - g(\eta))^2 \pi_\Lambda^{\mathrm{per}}(\sigma)\mathcal{L}_\Lambda^{\mathrm{per}}(\sigma,\eta) + \sum_{k=1}^{L-1} \sum_{\sigma \in \mathcal{R}_k} \sum_{\eta \in \mathcal{R}_{k\pm1}} (g(\sigma) - g(\eta))^2 \pi_\Lambda^{\mathrm{per}}(\sigma)\mathcal{L}(\sigma,\eta)$$

$$+ \sum_{\sigma \in \mathcal{R}} \sum_{\eta \notin \mathcal{R}} (g(\sigma) - g(\eta))^2 \pi_\Lambda^{\mathrm{per}}(\sigma)\mathcal{L}(\sigma,\eta)$$

$$\equiv S_1 + S_2 + S_3,$$

where the inequality is there because the sets $\{\mathcal{R}_k\}_{k=1}^{L-1}$ are not quite disjoint. We calculate these three terms separately. Using the estimate $\pi_\Lambda^{\mathrm{per}}(\mathcal{G}) = 1 - o(1)$ from Lemma 4.4, the bound (7.4), and the fact that $\mathcal{L}(\sigma,\eta) \in \{e^{-4\beta}, 0\}$ for $\sigma \in \mathcal{G}, \eta \notin \mathcal{G}$, we have

$$S_1 = \sum_{\sigma \in \mathcal{G}} \sum_{\eta \in \mathcal{R}} (g(\sigma) - g(\eta))^2 \pi_\Lambda^{\mathrm{per}}(\sigma)\mathcal{L}(\sigma,\eta) \tag{7.7}$$

$$\lesssim \sum_{\sigma \in \mathcal{G}} |\{\eta \in \mathcal{R} \,:\, \mathcal{L}(\sigma,\eta) > 0\}|\, L^{-2}\, |\mathcal{G}|^{-1}\, e^{-4\beta}$$

$$\lesssim |\mathcal{G}| \times L^2 \times L^{-2} \times |\mathcal{G}|^{-1} \times e^{-4\beta}$$

$$= e^{-4\beta}.$$

Next, note that an element $\zeta \in \mathcal{R}_k$ is uniquely determined by a choice of $\sigma \in \mathcal{G}, \eta \in \mathcal{G}$ s.t. $|w(\sigma) - w(\eta)| = 1$, and $m \in [1:L]$. Multiplying these three factors,

$$|\mathcal{R}_k| \leqslant |\mathcal{G}| \times (2L) \times (L).$$

Using this bound, the trivial bound $\mathcal{L}(\sigma,\eta) \leqslant 1$ for any configurations σ,η, and applying Lemma 4.4 and Inequality (7.4) as above, we have:

$$S_2 = \sum_{k=1}^{L-1} \sum_{\sigma \in \mathcal{R}_k} \sum_{\eta \in \mathcal{R}_{k\pm1}\,:\,\mathcal{L}(\sigma,\eta)>0} (g(\sigma) - g(\eta))^2 \pi_\Lambda^{\mathrm{per}}(\sigma)\mathcal{L}(\sigma,\eta) \tag{7.8}$$

$$\lesssim L \times (|\mathcal{G}| \times L^2) \times (1) \times L^{-2} \times (e^{-4\beta} \times \frac{1}{|\mathcal{G}|}) \times (1)$$

$$= L\, e^{-4\beta} \lesssim e^{-\frac{7}{2}\beta}.$$

Finally, we note that for $\sigma \in \mathcal{R}$ and η with $\mathcal{L}(\sigma,\eta) > 0$, we must have $|D(\eta)| \leqslant |D(\sigma)| + 4 \leqslant 8$. Furthermore, if $|D(\eta)| = 6$, $|D(\sigma)| = 4$, and $\eta = \sigma^x$ for some $x \in [1:L]^2$, then the single flipped spin x must share a plaquette with one of the $O(1)$ defects in $D(\sigma)$. In particular, for each fixed $\sigma \in \mathcal{R}$,

$$|\{\eta \notin \mathcal{R} \,:\, \mathcal{L}(\sigma,\eta) > 0, |D(\eta)| = 6\}| = O(1).$$

Finally, we use the trivial bound $g(\sigma) \leqslant L$. Then, applying Lemma 4.4 and Inequality (7.4) as above, we have:

$$S_3 = \sum_{\sigma \in \mathcal{R}} \sum_{\eta \notin \mathcal{R}} (g(\sigma) - g(\eta))^2 \pi_\Lambda^{\mathrm{per}}(\sigma)\mathcal{L}(\sigma,\eta) \tag{7.9}$$

$$= \sum_{\sigma \in \mathcal{R}} \sum_{\eta \notin \mathcal{R},\, |D(\eta)|=6} (g(\sigma) - g(\eta))^2 \pi_\Lambda^{\mathrm{per}}(\sigma)\mathcal{L}(\sigma,\eta)$$

$$+ \sum_{\sigma \in \mathcal{R}} \sum_{\eta \notin \mathcal{R}, \, |D(\eta)|=8} (g(\sigma) - g(\eta))^2 \pi_\Lambda^{\mathrm{per}}(\sigma) \mathcal{L}(\sigma, \eta)$$

$$\lesssim (L^2 \times |\mathcal{G}|) \times (1) \times L^2 \times (\frac{1}{|\mathcal{G}|} e^{-4\beta}) \times e^{-2\beta} + (L^2 \times |\mathcal{G}|) \times L^2 \times L^2 \times (\frac{1}{|\mathcal{G}|} e^{-4\beta}) \times e^{-4\beta}$$

$$= L^4 e^{-6\beta} + L^6 e^{-8\beta} \lesssim e^{-4\beta}.$$

Combining Inequality (7.6) with Inequalities (7.7), (7.8) and (7.9), we have

$$\mathcal{D}_\Lambda^{\mathrm{per}}(g) \lesssim e^{-\frac{7}{2}\beta}. \tag{7.10}$$

Combining Inequalities (7.5) and (7.10), the variational characterization of the spectrum in (2.13) gives

$$\begin{aligned} T_{\mathrm{rel}}(L_c, +) &\geqslant \sup_{\mathrm{Var}_\pi(f) \neq 0} \frac{\mathrm{Var}_\Lambda^{\mathrm{per}}(f)}{\mathcal{D}_\Lambda^{\mathrm{per}}(f)} \\ &\geqslant \frac{\mathrm{Var}_\Lambda^{\mathrm{per}}(g)}{\mathcal{D}_\Lambda^{\mathrm{per}}(g)} \\ &\gtrsim L \times e^{\frac{7}{2}\beta} = e^{4\beta}. \end{aligned}$$

This completes the proof of the lower bound of Theorem 3.2.

7.2 Upper bound

We now prove the upper bound in Theorem 3.2. Throughout this section, we restrict our attention to pairs β, L with $L = L_c(\beta)$.

As in the lower bound, the main heuristics are that the trace $\{\hat{X}_t\}_{t \geqslant 0}$ of $\{X_t\}_{t \geqslant 0}$ on $\mathcal{G}$ should determine the mixing time, and that this trace is very similar to simple random walk on the hypercube. Roughly speaking, our argument has the following steps:

1. We show that the dynamics of the trace process can be compared to simple random walk on the hypercube, and conclude that it has a similar mixing time of order $O(L \log(L)^2)$ (Lemma 7.5 and Corollary 7.7).
2. Using the characterization of mixing times in terms of hitting times of large sets (see Theorem 1 of [22]), we conclude from step 1 that the hitting time of any large set (for the trace process) is bounded by $O(L \log(L)^2)$.
3. We show, for the original process, that the expected hitting time of the set $\mathcal{G}$ is $O\left(\beta^9 e^{4\beta}\right)$. (Lemma 7.8).
4. We show, for the original process, that the expected transition time between elements of $\mathcal{G}$ is $O\left(\beta^8 e^{3.5\beta}\right)$ (Lemma 7.9).
5. Putting these bounds together, and recalling that $\pi_\Lambda^{\mathrm{per}}(\mathcal{G}) = 1 - o(1)$, we conclude that the hitting time of any large set (for the original process) cannot be much more than $O(L \times \beta^8 e^{3.5\beta} + \beta^9 e^{4\beta}) = O(\beta^9 e^{4\beta})$. We conclude the argument by again using the characterization of [22].

The main ingredient in Step 3 is relating the defect dynamics in the case of periodic boundary conditions to the defect dynamics in the case of all-plus boundary conditions. This relation is a consequence of the following three simple observations:

1. By Lemma 4.1, the sets $p^+(\Omega_{[1:L]^2})$ and $p^{\mathrm{per}}(\Omega_{[0:L]^2})$ are the same.
2. Let $\{X_t^{(+)}\}_{t \geqslant 0}$, $\{X_t^{(\mathrm{per})}\}_{t \geqslant 0}$ be Markov chains evolving according to the generators $\mathcal{L}_{[1:L]^2}^+$ and $\mathcal{L}_{[0:L]^2}^{(\mathrm{per})}$ respectively, as defined in Equation (2.8). Then $\{p^+(X_t^+)\}_{t \geqslant 0}$ and $\{p^{\mathrm{per}}(X_t^{(\mathrm{per})})\}_{t \geqslant 0}$ are both themselves Markov chains, on the common space

$p^+(\Omega_{[1:L]^2}) = p^{\mathrm{per}}(\Omega_{[0:L]^2})$, and they have the same stationary distribution. We denote their generators by $\mathcal{Q}^+_{[1:L]^2}$ and $\mathcal{Q}^{(\mathrm{per})}_{[0:L]^2}$ respectively.

3. For all $x, y \in p^+(\Omega_{[1:L]^2})$,

$$\mathcal{Q}^{(\mathrm{per})}_{[0:L]^2}(x, y) \;\geqslant\; \mathcal{Q}^+_{[1:L]^2}(x, y). \tag{7.11}$$

Note that this inequality is really not an equality, as there are transitions in the periodic boundary condition case that are not possible in the all-plus boundary condition case.

There is an important sequence of observations that makes most of the following calculation quite straightforward, even if some of the details are messy: First, almost all "excursions" between elements of the ground state $\mathcal{G}$ remain within the low-energy set $\mathcal{R}$. Furthermore, each excursion from $\mathcal{G}$ can be decomposed into its first jump into $\mathcal{R}_1 \cup \mathcal{R}_{k-1}$, and the remaining excursions into sets of the form $\{R(\sigma, \eta, m, k)\}_{k=2}^{L-2}$ for some fixed $\sigma, \eta \in \mathcal{G}$ and starting point $m \in [1 : L]$. As long as these longer excursions remain within a set $\{R(\sigma, \eta, m, k)\}_{k=2}^{L-2}$, they follow *exactly* the law of a simple random walk on $[2 : (L-2)]$, with the obvious bijection between $\{R(\sigma, \eta, m, k)\}_{k=2}^{L-2}$ and $[2 : (L-2)]$ (see Figure 8). In particular, essentially all of the following calculations will be written in terms of simple random walk with killing, combined with simple 1-step analyses of what occurs between excursions.

Continuing more formally, we introduce the following definition for the discrete-time trace of a continuous-time Markov chain:

Definition 7.4 (Non-Lazy Trace Chain). *Let $\{X_t\}_{t \geqslant 0}$ be a càdlàg Markov process on finite set Ω, and let $S \subset \Omega$. Define two sequences of times $\{t_i\}_{i \in \mathbb{N}}$, $\{s_i\}_{i \in \mathbb{N}}$ by the following recursions:*

$$\begin{aligned}
t_1 &= \inf\{t > 0 \,:\, X_t \in S\} \\
s_i &= \inf\{t > t_i \,:\, X_t \notin S\} \\
t_{i+1} &= \inf\{t > s_i \,:\, X_t \in S, \, X_t \neq X_{t_i}\}.
\end{aligned}$$

Then define the trace process $\{\hat{X}_k\}_{k \in \mathbb{N}}$, on S, by the equation

$$\hat{X}_k \equiv X_{t_k}.$$

We denote by $Q_{\mathcal{G}}^{\mathrm{per}}$ the transition kernel associated with the trace of $\{X_t\}_{t \geqslant 0}$ on the set $\mathcal{G}$.

Denote by Q_{hyp} the following discrete-time transition kernel on $\{-1, +1\}^{2L-1}$:

$$\begin{aligned}
Q_{\mathrm{hyp}}(u, v) &= \frac{1}{2L-1}, &&\text{if } |u - v| = 1, \\
Q_{\mathrm{hyp}}(u, v) &= 0, &&\text{otherwise.}
\end{aligned}$$

Recall $|\cdot|$ is the Hamming distance, i.e. $|u - v| = 1$ iff u and v differ at exactly one coordinate.

The following estimate allows us to make very strong conclusions about the mixing of $Q_{\mathcal{G}}^{\mathrm{per}}$ in terms of Q_{hyp}:

Lemma 7.5. *For $Q_{\mathcal{G}}^{\mathrm{per}}$, Q_{hyp} as above, and the function w as in Definition 7.1, we have*

$$Q_{\mathcal{G}}^{\mathrm{per}}(w^{-1}(u), w^{-1}(v)) \gtrsim Q_{\mathrm{hyp}}(u, v) \tag{7.12}$$

uniformly in $u, v \in \{-1, +1\}^{2L-1}$. Furthermore, for $\{X_t\}_{t \geqslant 0}$ a copy of the plaquette dynamics started at $X_0 = \sigma \in \mathcal{G}$, let

$$\tau_{\text{start}} = \inf\{t > 0 : X_t \neq \sigma\} \tag{7.13}$$
$$\tau_{\text{end}} = \inf\{t > \tau_{\text{start}} : X_t \in \mathcal{G} \text{ or } |D(X_t)| > 4\}$$

be the times at which the first excursion from $\mathcal{G}$ starts and ends. Then

$$\mathbb{P}_\sigma[|D(X_{\tau_{\text{end}}})| > 4] \lesssim L^2 e^{-2\beta} \tag{7.14}$$
$$1 - \mathbb{P}_\sigma[X_{\tau_{\text{end}}} = \sigma] \approx L^{-1}.$$

Remark 7.6 (Coupling to Simple Random Walks)**.** The proof of Lemma 7.5 (and the rest of this section) relies heavily on the following simple observation: every configuration $\zeta \in \mathcal{R} \backslash (\mathcal{G} \cup \mathcal{R}_1 \cup \mathcal{R}_{L-1})$ consists of two pairs of adjacent defects, and these two pairs each *exactly* undergo independent random walk until the first time they are adjacent or new defects are added (again, see Figure 8).

More formally, fix $m \in [1 : L]$ and $k \in [2 : (L-2)]$, and let $\{L_t, U_t\}_{t \geqslant 0}$ be two independent simple random walks on $\mathbb{Z}$ with rate 1 and starting points $L_0 = m$, $U_0 = m + k$. Let $\{X_t\}_{t \geqslant 0}$ be the usual defects process, started at a point $X_0 = \mathcal{R}(\sigma, \eta, m, k)$. Finally, let

$$\tau_1 = \inf\{s > 0 : X_s \notin \cup_{m \in [1:L]} \cup_{k \in [2:(L-1)]} \mathcal{R}(\sigma, \eta, m, k)\}$$
$$\tau_2 = \inf\{s > 0 : U_s = L_s + 1 \text{ or } U_s = L_s + L - 1\}.$$

Then, inspecting the generators of these processes, it is possible to couple $\{(X_t, L_t, U_t)\}_{t \geqslant 0}$ so that

$$X_t = \mathcal{R}(\sigma, \eta, L_t, U_t - L_t)$$

for all $0 \leqslant t \leqslant \tau_1 \leqslant \tau_2$.

This coupling gives a *complete* description of the excursions of $\{X_t\}_{t \geqslant 0}$ in sets of the form $\{R(\sigma, \eta, m, k)\}_{k=2}^{L-2}$ for some fixed $\sigma, \eta \in \mathcal{G}$ and starting point $m \in [1 : L]$. We use this coupling, and its immediate consequences, extensively in the remainder of this section to "translate" well-known facts about simple random walk to our setting. Most of the proof simply consists of applying this coupling to excursions on the set $\mathcal{R} \backslash (\mathcal{G} \cup \mathcal{R}_1 \cup \mathcal{R}_{L-1})$.

The behaviour of the process outside of these excursions on $\mathcal{R} \backslash (\mathcal{G} \cup \mathcal{R}_1 \cup \mathcal{R}_{L-1})$ is fairly simple. Some excursions end by exiting $\mathcal{R}$, and we will see that this is quite rare. The remaining excursions will end at $\mathcal{R}_1 \cup \mathcal{R}_{L-1} \cup \mathcal{G}$; since each element $\sigma \in \mathcal{R}_1 \cup \mathcal{R}_{L-1}$ has at most 4 neighbours η with high transition probability $\mathcal{L}(\sigma, \eta) = 1$ (again, see Figure 8), it is straightforward to analyze all transitions from within the set $\mathcal{R}_1 \cup \mathcal{R}_{L-1} \cup \mathcal{G}$ that have non-negligible probability.

Proof. Fix $u, v \in \{-1, +1\}^{2L-1}$ that differ at a single index ℓ and let $\sigma = w^{-1}(u)$, $\eta = w^{-1}(v)$. Assume for now that $1 \leqslant \ell \leqslant L$; the other cases will be essentially identical. Recalling notation from Definition 7.1, which will be used heavily throughout this proof, we assume without loss of generality that $u \prec v$.

We study $\{X_t\}_{t \geqslant 0}$ by charting its movement from a ground state, to an adjacent non-ground state with exactly 4 adjacent defects, to moving along minimal-energy paths between ground states; we stop keeping track of $\{X_t\}_{t \geqslant 0}$ as soon as it either returns to a ground state *or* enters a state with 6 or more defects.

Let $\{X_t\}_{t \geqslant 0}$ be a copy of the plaquette dynamics started at $X_0 = \sigma \in \mathcal{G}$ and recall $\tau_{\text{start}} = \inf\{t > 0 : X_t \neq \sigma\}$. By symmetry we have

$$\mathbb{P}_\sigma[X_{\tau_{\text{start}}} \in \cup_{m=1}^L (R(\sigma, \xi, m, 1) \cup R(\xi, \sigma, m, L-1))] = \frac{1}{L} \tag{7.15}$$

for all $\xi \in \mathcal{G}$ such that $|w(\sigma) - w(\xi)| = 1$ or $w(\sigma) = -w(\xi)$. Note that this means all other transition probabilities are 0.

Define the event $\mathcal{E}_{\text{start}} = \{X_{\tau_{\text{start}}} \in \cup_{m=1}^{L} R(\sigma, \eta, m, 1)\}$. Next, let $\tau_{\text{exc}}^{(1)} = \inf\{t > \tau_{\text{start}} : X_t \neq X_{\tau_{\text{start}}}\}$. Then define a sequence of successive return and excursion times by the following formulae, with $i \in \mathbb{N}$:

$$\tau_{\text{ret}}^{(i)} = \inf\{t > \tau_{\text{exc}}^{(i)} : X_t \notin \cup_{m=1}^{L} \cup_{k=2}^{L-2} R(\sigma, \eta, m, k)\}$$
$$\tau_{\text{exc}}^{(i+1)} = \inf\{t > \tau_{\text{ret}}^{(i)} : X_t \neq X_{\tau_{\text{ret}}^{(i)}}\}.$$

The reader may note that this formula looks slightly different from *e.g.* (7.15) in that it contains only unions over $R(\sigma, \eta, m, k)$, with none over $R(\eta, \sigma, m, k)$; this is because we have assumed that $\eta \in \Omega_{[1:L]^2}$ satisfies $w(\sigma) \prec w(\eta)$, so that *e.g.* $R(\eta, \sigma, m, k) = \emptyset$ for all m, k.

We keep track of the behaviour of $\{X_t\}_{t \geqslant 0}$ in the excursions from $\cup_{m=1}^{L}(R(\sigma, \eta, m, 1) \cup \mathcal{R}(\sigma, \eta, m, L-1))$ by using the following families of indicator functions:

$$\phi_{\mathcal{G}}^{(i)} = \mathbb{1}\{X_{\tau_{\text{exc}}^{(i)}} \in \mathcal{G}\}, \quad \phi_{\mathcal{R}}^{(i)} = \mathbb{1}\{|D(X_{\tau_{\text{exc}}^{(i)}})| = 4\}, \quad \phi_{\Delta}^{(i)} = 1 - \phi_{\mathcal{G}}^{(i)} - \phi_{\mathcal{R}}^{(i)}$$

and

$$\psi_1^{(i)} = \mathbb{1}\{X_{\tau_{\text{ret}}^{(i)}} = X_{\tau_{\text{exc}}^{(i)}}\}, \quad \psi_{\Delta}^{(i)} = \mathbb{1}\{|D(X_{\tau_{\text{ret}}^{(i)}})| > 4\}, \quad \psi_{L-1}^{(i)} = 1 - \psi_0^{(i)} - \psi_{\Delta}^{(i)}.$$

In the family associated with the letter ϕ, the first indicates that the i'th excursion begins with $\{X_t\}_{t \geqslant 0}$ immediately returning to a ground state; the second indicates that the excursion begins by moving along a minimal-energy path between neighbouring ground states; the third indicates that something else has happened. In the family associated with the letter ψ, the first indicates that the i'th excursion ended where it started; the second indicates that the i'th excursion involves moving off of a minimal-energy path; the third indicates that the i'th excursion involves travelling along a minimal-energy path until a point that is adjacent to a different ground state.

Let K_{SRW} be the transition kernel on state space $[1 : (L-1)] \cup \{\Delta\}$ with holding at $\{1, L-1, \Delta\}$ and other nonzero transitions rates:

$$K_{\text{SRW}}(i, i \pm 1) = 2$$
$$K_{\text{SRW}}(i, \Delta) = 8e^{-2\beta} + (L^2 - 12)e^{-4\beta}.$$

Let $\{Z_t\}_{t \geqslant 0}$ be a Markov process with this transition kernel started at 2, and let $\hat{\tau} = \inf\{t > 0 : Z_t \in \{1, L-1, \Delta\}\}$. By the discussion in Remark 7.6 and simple counting of the states adjacent to $\cup_{k \in [2:(L-2)]} \cup_{m \in [1:L]} \mathcal{R}(\sigma, \eta, m, k)$ with 6 or more defects, on the event $\{\phi_{\mathcal{R}}^{(i)} = 1\}$, we can couple $\{Z_t\}_{t \geqslant 0}$ to $\{X_t\}_{\tau_{\text{exc}}^{(i)}}^{\tau_{\text{ret}}^{(i)}}$ so that the following hold:

$$\hat{\tau} = \tau_{\text{ret}}^{(i)} - \tau_{\text{exc}}^{(i)} \tag{7.16}$$
$$\mathbb{1}\{Z_{\hat{\tau}} = z\} = \psi_z^{(i)}, \qquad z \in \{1, L-1, \Delta\}.$$

This correspondence allows us to analyze excursions using well-known properties of K_{SRW}, which is simple random walk on the path with killing at constant rate. Recall the definition of τ_{end} from Equality (7.13), and let $I \in \mathbb{N}$ be the unique (random) integer satisfying

$$\tau_{\text{end}} \in \{\tau_{\text{exc}}^{(I)}, \tau_{\text{ret}}^{(I)}\};$$

that is, the first time that the walk either enters a ground state *or* a state with 6 or more defects.

Combining Equalities (7.15) and (7.16), and using standard facts about hitting probabilities for simple random walk on the path, we have

$$
\begin{aligned}
\mathbb{P}_\sigma[X_{\tau_{\mathrm{end}}} = \eta] &\gtrsim \mathbb{P}_\sigma[\mathcal{E}_{\mathrm{start}}]\mathbb{P}_\sigma[X_{\tau_{\mathrm{end}}} = \eta | \mathcal{E}_{\mathrm{start}}] \\
&\gtrsim L^{-1} \sum_{i \in \mathbb{N}} \mathbb{P}_\sigma[X_{\tau_{\mathrm{end}}} = \eta | \mathcal{E}_{\mathrm{start}}, I = i]\mathbb{P}_\sigma[I = i | \mathcal{E}_{\mathrm{start}}] \\
&\gtrsim L^{-1} \sum_{i \in \mathbb{N}} (L^{-1} - Le^{-2\beta})\mathbb{P}_\sigma[I = i | \mathcal{E}_{\mathrm{start}}] \approx L^{-2}.
\end{aligned}
$$

By essentially the same calculation, we have

$$
\begin{aligned}
\mathbb{P}_\sigma[|D(X_{\tau_{\mathrm{end}}})| > 4] &\lesssim L^2 e^{-2\beta} \\
1 - \mathbb{P}_\sigma[X_{\tau_{\mathrm{end}}} = \sigma] &\approx L^{-1}.
\end{aligned}
$$

This completes the proof of (7.14) in the case that $1 \leqslant \ell \leqslant L$ and $w(\sigma) \prec w(\eta)$. Combining these two inequalities, and using symmetry,

$$
\begin{aligned}
\mathbb{P}_\sigma[X_{\tau_{\mathrm{end}}} = \eta | X_{\tau_{\mathrm{end}}} \neq \sigma] &\gtrsim L^{-1}\frac{\mathbb{P}_\sigma[X_{\tau_{\mathrm{end}}} = \eta]}{\mathbb{P}_\sigma[X_{\tau_{\mathrm{end}}} \in \mathcal{G}\backslash\{\sigma\}] + \mathbb{P}[|D(X_{\tau_{\mathrm{end}}})| > 4]} \\
&\gtrsim \frac{L^{-2}}{L^{-1} + L^2 e^{-2\beta}} \approx L^{-1}.
\end{aligned}
$$

This completes the proof of Inequality (7.12), and thus the lemma, in the case that $1 \leqslant \ell \leqslant L$ and $w(\sigma) \prec w(\eta)$. The other cases are essentially identical by symmetry. $\square$

This comparison implies the mixing bound:

Corollary 7.7. *The mixing time of the transition kernel* $\frac{1}{2}(\mathrm{Id} + Q_{\mathcal{G}}^{\mathrm{per}})$ *is* $O(L \log(L)^2)$.

Proof. This follows immediately from the well-known fact that the mixing time of $\frac{1}{2}$-lazy simple random walk on $\{-1, +1\}^{2L-1}$ is $O(L \log(L))$, the comparison of $Q_{\mathcal{G}}^{\mathrm{per}}$ to the transition kernel for simple random walk on $\{-1, +1\}^{2L-1}$ given in Lemma 7.5, and Theorem 1 of [18]. $\square$

Fix $\sigma \in \Omega_{[1:L]^2}$ and let $\{X_t\}_{t \geqslant 0}$ be a copy of the original Markov process started at point $X_0 = \sigma$. For $A \subset \Omega_{[1:L]^2}$, denote by

$$
\tau_A = \inf\{t > 0 : X_t \in A\}
$$

the hitting time of the set A. Using the upper bound on the mixing time in Theorem 3.1, we have:

Lemma 7.8. *The hitting time* $\tau_{\mathcal{G}}$ *satisfies*

$$
\mathbb{E}_\sigma[\tau_{\mathcal{G}}] \lesssim \beta^9 e^{4\beta},
$$

uniformly in $\sigma \in \Omega_\Lambda$.

Proof. Denote by $\tau_{\mathrm{mix},d}^{\mathrm{per}}$ and $\tau_{\mathrm{mix},d}^+$ the mixing times of the defect dynamics given by the generators $\mathcal{Q}_{[0:L]^2}^{(\mathrm{per})}$ and $\mathcal{Q}_{[1:L]^2}^+$ respectively. Since the defect dynamics are a deterministic function of the spin dynamics, it is clear that

$$
\tau_{\mathrm{mix},d}^+ \leqslant \tau_{\mathrm{mix}}^+, \quad \tau_{\mathrm{mix},d}^{\mathrm{per}} \leqslant \tau_{\mathrm{mix}}^{\mathrm{per}}.
$$

By the tightness of the spectral profile (as stated in Theorem 1 of [18]) and Inequality (7.11), we also have

$$
\tau_{\mathrm{mix},d}^{\mathrm{per}} \lesssim \beta \tau_{\mathrm{mix},d}^+.
$$

Combining these two bounds with Theorem 3.1,

$$\tau_{\mathrm{mix},d}^{\mathrm{per}} \lesssim \beta\tau_{\mathrm{mix},d}^{+} \lesssim \beta^9 e^{4\beta}. \tag{7.17}$$

Recalling that $\pi(\mathcal{G}) = 1 - o(1)$ (by Lemma 4.4), Inequality (7.17) and Theorem 1 of [22] immediately imply the result. $\qquad\square$

For $\sigma \in \mathcal{G}$, we have the improved bound on the expected hitting time:

Lemma 7.9. *Let $\sigma \in \mathcal{G}$. Then*

$$\mathbb{E}_\sigma[\tau_{\mathcal{G}\setminus\{\sigma\}}] \lesssim \beta^9 \, e^{3.5\beta}.$$

Proof. Let $\{X_t\}_{t \geqslant 0}$ be a copy of the original Markov process with starting point $X_0 = \sigma$. Define inductively the sequence of times

$$t_1 = \inf\{t > 0 \,:\, X_t \neq \sigma\}$$
$$s_i = \inf\{t > t_i \,:\, X_t \in \mathcal{G}\}$$
$$t_{i+1} = \inf\{t > s_i \,:\, X_t \neq X_{s_i}\}.$$

Denote by

$$\mathcal{B}_i = \{\max_{t_i \leqslant t \leqslant s_i} |D(X_t)| > 4\}$$

the event that $\{X_t\}_{t \geqslant 0}$ escapes from the collection of minimal-energy paths during the i'th excursion. Define

$$J = \min\{j \in \mathbb{N} \,:\, X_{s_j} \neq \sigma\},$$

so that $\tau_{\mathcal{G}\setminus\{\sigma\}} = s_J$. Thus, we have

$$\mathbb{E}_\sigma[\tau_{\mathcal{G}\setminus\{\sigma\}}] = \mathbb{E}_\sigma[s_J] \tag{7.18}$$
$$= \mathbb{E}_\sigma\Big[\sum_{i=1}^{J}(s_i - t_i) + \sum_{i=2}^{J}(t_i - s_{i-1}) + t_1\Big]$$
$$= \sum_{j=1}^{\infty}\Big(\sum_{i=1}^{j}\mathbb{E}_\sigma[(s_i - t_i)|J = j] + \sum_{i=2}^{j}\mathbb{E}_\sigma[(t_i - s_{i-1})|J = j] + \mathbb{E}_\sigma[t_1|J = j]\Big)\mathbb{P}_\sigma[J = j].$$

We now bound these three types of terms, and the probability $\mathbb{P}_\sigma[J = j]$. By direct inspection of the transition probabilities, and the Markov property,

$$\mathbb{E}_\sigma[t_1|J = j] \lesssim L^{-2}e^{4\beta} \approx e^{3\beta}.$$

By the same calculation,

$$\mathbb{E}_\sigma[t_i - s_{i-1}|J = j] = \mathbb{E}_\sigma[t_1|J = j] \lesssim L^{-2}e^{4\beta} \approx e^{3\beta} \tag{7.19}$$

for $1 \leqslant i \leqslant j$. By Inequality (7.14),

$$\mathbb{P}_\sigma[J = 1] \gtrsim L^{-1},$$

and so by the Markov property

$$\mathbb{P}_\sigma[J > j] \leqslant (1 - C\,L^{-1})^j \tag{7.20}$$

for some fixed $0 < C < \infty$ and all $j \in \mathbb{N}$. Finally, we calculate the term $\mathbb{E}[s_i - t_i | J = j]$. By the coupling Equality (7.16) from the proof of Lemma 7.5 and the usual bound of $O(L^2)$ on the first hitting time of $\{1, L-1\}$ for simple random walk on $[1 : (L-1)]$,

$$\mathbb{E}_\sigma[(s_i - t_i)\mathbb{1}_{\mathcal{B}_i^c} | J = j] \lesssim L^2. \tag{7.21}$$

By Inequality (7.14),

$$\mathbb{P}_\sigma[\mathcal{B}_i | J = j] \lesssim L^2 e^{-2\beta} \approx e^{-\beta}.$$

Thus, also applying Lemma 7.8 and then again Inequality (7.14),

$$\begin{aligned}
\mathbb{E}_\sigma[(s_i - t_i)\mathbb{1}_{\mathcal{B}_i} | J = j] &= \mathbb{E}_\sigma[(s_i - t_i) | J = j, \mathcal{B}_i]\mathbb{P}_\sigma[\mathcal{B}_i | J = j] \\
&\overset{\text{Lemma 7.8}}{\lesssim} \beta^9 e^{4\beta}\mathbb{P}_\sigma[\mathcal{B}_i | J = j] \\
&\overset{\text{Inequality (7.14)}}{\lesssim} \beta^9 e^{4\beta} L^2 e^{-2\beta} \\
&\approx \beta^9 e^{3\beta}.
\end{aligned}$$

Combining this with Inequality (7.21), we conclude

$$\mathbb{E}_\sigma[(s_i - t_i) | J = j] \lesssim \beta^9 e^{3\beta}. \tag{7.22}$$

Applying Inequalities (7.19), (7.20), and (7.22) to Inequality (7.18), we have:

$$\begin{aligned}
\mathbb{E}_\sigma[\tau_{\mathcal{G}\setminus\{\sigma\}}] &= \sum_{j=1}^\infty \left(\sum_{i=1}^j \mathbb{E}_\sigma[(s_i - t_i) | J = j] + \sum_{i=2}^j \mathbb{E}_\sigma[(t_i - s_{i-1}) | J = j] + \mathbb{E}_\sigma[t_1 | J = j]\right)\mathbb{P}[J = j] \\
&\lesssim \sum_{j=1}^\infty j\,\beta^9 e^{3\beta}\mathbb{P}_\sigma[J = j] \\
&\lesssim \beta^9 e^{3.5\beta}.
\end{aligned}$$

This completes the proof. $\qquad\square$

We now put these results together:

Proof of Upper Bound of Theorem 3.2. Fix a point $\sigma \in \Omega_{[1:L]^2}$ and a subset $A \subset \Omega_{[1:L]^2}$ with $\pi_\Lambda^{\text{per}}(A) \geqslant \frac{3}{4}$. We wish to bound the expected hitting time $\mathbb{E}_\sigma[\tau_A]$.

Assume for now that $\sigma \in \mathcal{G}$. Let $\{\hat{X}_i\}_{i\in\mathbb{N}}$ be the trace of $\{X_t\}_{t \geqslant 0}$ on $\mathcal{G}$, as in Definition 7.4. Let

$$\hat{\tau}_A = \min\{i \,:\, \hat{X}_i \in A\}.$$

By Lemma 4.4, we know that $\pi_\Lambda^{\text{per}}(A \cap \mathcal{G}) \geqslant \frac{2}{3}$ for all $\beta > \beta_0$ sufficiently large. Thus, by Theorem 1 of [22] and Corollary 7.7,

$$\mathbb{E}_\sigma[\hat{\tau}_A] \lesssim L \log(L)^2,$$

uniformly in the particular choice of $\sigma \in \mathcal{G}$ and $A \subset \Omega_\Lambda$ with large measure. In particular, there exists a constant $0 < C_1 < \infty$ that satisfies

$$\mathbb{P}_\sigma[\hat{\tau}_A > C_1 L \log(L)^2] \leqslant 0.01 \tag{7.23}$$

uniformly in these choices.

Next, let $\{t_i\}_{i \in \mathbb{N}}$ be the random sequence of times given in Definition 7.4 when constructing the trace $\{\hat{X}_i\}_{i \in \mathbb{N}}$. Define also the timescale $I = \lceil C_1 L \log(L)^2 + 1 \rceil$. By Lemma 7.9,

$$\mathbb{E}[t_{i+1} - t_i | X_{t_i} = \eta] \lesssim \beta^9 e^{3.5\beta}, \tag{7.24}$$

uniformly in the choice of $\eta \in \mathcal{G}$. In particular, there exists a constant $0 < C_2 < \infty$ that satisfies

$$\mathbb{P}_\sigma[t_I > C_2 \beta^9 e^{3.5\beta} L \log(L)^2] \leqslant 0.01$$

uniformly in the initial point $\sigma \in \mathcal{G}$. Combining Inequalities (7.23) and (7.24), we have uniformly in $\sigma \in \mathcal{G}$ and $A \subset \Omega_{[1:L]^2}$ with sufficiently large measure

$$\mathbb{P}_\sigma[\tau_A > C_2 \beta^9 e^{3.5\beta} L \log(L)^2] \leqslant \mathbb{P}_\sigma[\hat{\tau}_A > I] + \mathbb{P}_\sigma[t_I > C_2 \beta^9 e^{3.5\beta} L \log(L)^2] \tag{7.25}$$
$$\leqslant 0.02.$$

Next we extend this to general starting positions. By Lemma 7.8, there exists $0 < C_3 < \infty$ so that, for all $\sigma \in \Omega_{[1:L]^2}$,

$$\mathbb{P}_\sigma[\tau_\mathcal{G} > C_3 \beta^9 e^{4\beta}] \leqslant 0.01.$$

Combining this with Inequality (7.25), we conclude that for all $\sigma \in \Omega_{[1:L]^2}$,

$$\mathbb{P}_\sigma[\tau_A > C_3 \beta^9 e^{4\beta} + C_2 \beta^9 e^{3.5\beta} L \log(L)^2] \leqslant \mathbb{P}_\sigma[\tau_\mathcal{G} > C_3 \beta^9 e^{4\beta}] +$$
$$+ \sup_{\eta \in \mathcal{G}} \mathbb{P}_\eta[\tau_A > C_2 \beta^9 e^{3.5\beta} L \log(L)^2]$$
$$\leqslant 0.01 + 0.02 = 0.03.$$

This immediately implies

$$\max_{\sigma \in \Omega_{[1:L]^2}} \mathbb{E}_\sigma[\tau_A] \lesssim \beta^9 e^{4\beta}.$$

By Theorem 1 of [22], this implies the desired bound on the mixing time. $\qquad\square$

References

[1] L. Berthier and G. Biroli, *Theoretical perspective on the glass transition and amorphous materials*, Reviews of Modern Physics **83** (2011), no. 2, 587–645.

[2] O. Blondel, N. Cancrini, F. Martinelli, C. Roberto, and C. Toninelli, *Fredrickson-Andersen one spin facilitated model out of equilibrium*, Markov Process. Relat. Fields **19** (2013), 383–406. MR-3156958

[3] C. Chamon, *Quantum glassiness in strongly correlated clean systems: An example of topological overprotection*, Phys. Rev. Lett. **94** (2005), no. 4, 1–4.

[4] P. Chleboun, A. Faggionato, and F. Martinelli, *Time Scale Separation and Dynamic Heterogeneity in the Low Temperature East Model*, Communications in Mathematical Physics **328** (2014), no. 3, 955–993. MR-3201217

[5] P. Chleboun, A. Faggionato, and F. Martinelli, *Mixing time and local exponential ergodicity of the East-like process in $\mathbb{Z}^d$*, Annales de la faculté des sciences de Toulouse Mathématiques **XXIV** (2015), no. 4, 717–743. MR-3434253

[6] P. Chleboun, A. Faggionato, and F. Martinelli, *Relaxation to equilibrium of generalized East processes on $\mathbb{Z}^d$: Renormalization group analysis and energy-entropy competition*, Ann. Probab. **44** (2016), no. 3, 1817–1863. MR-3502595

[7] P. Chleboun, A. Faggionato, F. Martinelli, and C. Toninelli, *Mixing Length Scales of Low Temperature Spin Plaquettes Models*, Journal of Statistical Physics **169** (2017), no. 3, 441–471. MR-3711606

[8] P. Chleboun, A. Faggionato, F. Martinelli, A. Smith, and C. Toninelli, *Relaxation time of plaquette spin models*, Still being written (2019).

[9] P. Chleboun and A. Smith, *Cutoff for the square plaquette model on a critical length scale*, arXiv preprint arXiv:1910.14439 [math.PR] (2019).

[10] A. Faggionato, F. Martinelli, C. Roberto, and C. Toninelli, *Aging Through Hierarchical Coalescence in the East Model*, Communications in Mathematical Physics **309** (2011), no. 2, 459–495 (en). MR-2864800

[11] A. Faggionato, F. Martinelli, C. Roberto, and C. Toninelli, *The East model: recent results and new progresses*, Markov Process. Related Fields **3** (2013), no. 19, 407–452. MR-3156959

[12] S. Ganguly, E. Lubetzky, and F. Martinelli, *Cutoff for the East Process*, Commun. Math. Phys. **335** (2015), no. 3, 1287–1322. MR-3320314

[13] J. P. Garrahan, *Glassiness through the emergence of effective dynamical constraints in interacting systems*, Journal of Physics: Condensed Matter **14** (2002), no. 7, 1571–1579.

[14] J. P. Garrahan, P. Sollich, and C. Toninelli, *Kinetically Constrained Models*, Dynamical heterogeneities in glasses, colloids, and granular media, OUP, 2011.

[15] S. Goel, R. Montenegro, and P. Tetali, *Mixing time bounds via the spectral profile*, Electronic Journal of Probability **11** (2006), no. 2000, 1–26. MR-2199053

[16] R. L. Jack, L. Berthier, and J. P. Garrahan, *Static and dynamic length scales in a simple glassy plaquette model*, Physical Review E **72** (2005), no. 1, 016103.

[17] R. L. Jack and J. P. Garrahan, *Phase transition for quenched coupled replicas in a plaquette spin model of glasses*, Phys. Rev. Lett. **116** (2016), 055702.

[18] G. Kozma, *On the precision of the spectral profile*, Latin American Journal of Probability and Mathematical Statistics **3** (2007), 321–329. MR-2372888

[19] R. M. Nandkishore and M. Hermele, *Fractons*, arXiv preprint arXiv:1803.11196 (2018), 1–30.

[20] M. E. Newman and C. Moore, *Glassy dynamics and aging in an exactly solvable spin model.*, Phys. Rev. E **60** (1999), no. 5, 5068–5072.

[21] R. I. Oliveira, *Mixing and hitting times for finite Markov chains*, Electronic Journal of Probability **17** (2012), no. May 2010, 1–12. MR-2968677

[22] Y. Peres and P. Sousi, *Mixing times are hitting times of large sets*, Journal of Theoretical Probability **28** (2015), no. 2, 488–519. MR-3370663

[23] N. S. Pillai and A. Smith, *Mixing times for a constrained Ising process on the torus at low density*, Annals of Probability **45** (2017), no. 2, 1003–1070. MR-3630292

[24] N. S. Pillai and A. Smith, *Mixing times for a constrained ising process on the two-dimensional torus at low density*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques **55** (2019), no. 3, 1649–1678. MR-4010947

[25] A. Sinclair, *Improved Bounds For Mixing Rates Of Markov Chains And Multicommodity Flow*, Combinatorics, Probability & Computing **1** (1992), no. April 1992, 351–370. MR-1211324

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm