Random-Cluster Dynamics in $\mathbb{Z}^2$: Rapid Mixing with General Boundary Conditions

Antonio Blanca
Department of Computer Science and Engineering, Pennsylvania State University, USA
ablanca@cse.psu.edu

Reza Gheissari
Courant Institute of Mathematical Sciences, New York University, USA
reza@cims.nyu.edu

Eric Vigoda
School of Computer Science, Georgia Institute of Technology, USA
vigoda@cc.gatech.edu

Abstract

The random-cluster (FK) model is a key tool for the study of phase transitions and for the design of efficient Markov chain Monte Carlo (MCMC) sampling algorithms for the Ising/Potts model. It is well-known that in the high-temperature region $\beta < \beta_c(q)$ of the q-state Ising/Potts model on an $n \times n$ box Λ_n of the integer lattice $\mathbb{Z}^2$, spin correlations decay exponentially fast; this property holds even arbitrarily close to the boundary of Λ_n and uniformly over all boundary conditions. A direct consequence of this property is that the corresponding single-site update Markov chain, known as the Glauber dynamics, mixes in optimal $O(n^2 \log n)$ steps on Λ_n for all choices of boundary conditions. We study the effect of boundary conditions on the FK-dynamics, the analogous Glauber dynamics for the random-cluster model.

On Λ_n the random-cluster model with parameters (p, q) has a sharp phase transition at $p = p_c(q)$. Unlike the Ising/Potts model, the random-cluster model has non-local interactions which can be forced by boundary conditions: external wirings of boundary vertices of Λ_n. We consider the broad and natural class of boundary conditions that are *realizable* as a configuration on $\mathbb{Z}^2 \setminus \Lambda_n$. Such boundary conditions can have many macroscopic wirings and impose long-range correlations even at very high temperatures ($p \ll p_c(q)$). In this paper, we prove that when $q > 1$ and $p \neq p_c(q)$ the mixing time of the FK-dynamics is polynomial in n for *every* realizable boundary condition. Previously, for boundary conditions that do not carry long-range information (namely wired and free), Blanca and Sinclair (2017) had proved that the FK-dynamics in the same setting mixes in optimal $O(n^2 \log n)$ time. To illustrate the difficulties introduced by general boundary conditions, we also construct a class of non-realizable boundary conditions that induce slow (stretched-exponential) convergence at high temperatures.

2012 ACM Subject Classification Mathematics of computing → Markov-chain Monte Carlo convergence measures; Theory of computation → Random walks and Markov chains

Keywords and phrases Markov chain, mixing time, random-cluster model, Glauber dynamics, spatial mixing

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2019.67

Category RANDOM

Related Version A full version of the paper is available at `https://arxiv.org/abs/1807.08722`.

Funding *Antonio Blanca*: Research supported in part by NSF grants CCF-1617306 and CCF-1563838.
Eric Vigoda: Research supported in part by NSF grants CCF-1617306 and CCF-1563838.

Acknowledgements The authors thank the anonymous referees for their helpful suggestions.

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019).
Editors: Dimitris Achlioptas and László A. Végh; Article No. 67; pp. 67:1–67:19
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

Statistical physics models are designed to study physical phase transitions where a small change in a parameter which controls the local interactions, such as temperature, causes abrupt changes in the macroscopic behavior of the system. Phase transitions are captured by the onset of long-range correlations between vertices in the underlying graph, the infinite 2-dimensional integer lattice graph $\mathbb{Z}^2$ being a widely considered example. These long-range correlations manifest in the asymptotic effect of "boundary conditions" in large finite volumes. For example, if we take an $n \times n$ box Λ_n of $\mathbb{Z}^2$ and fix a configuration on the boundary of this box, as we formalize momentarily, this fixed boundary condition may affect the static (equilibrium state) and dynamic (approach to equilibrium) properties of the system.

The most notable and well-studied statistical physics model is the Ising/Potts model of ferromagnetism. The (ferromagnetic) Ising/Potts model on a (finite) graph, say the $n \times n$ box $\Lambda_n \subset \mathbb{Z}^2$ with nearest-neighbor edges $E(\Lambda_n)$, is defined on the set of spin assignments $\{1, \dots, q\}^{\Lambda_n}$. The probability of a configuration $\sigma \in \{1, \dots, q\}^{\Lambda_n}$ in the associated Gibbs distribution μ_{Λ_n} is proportional to $\exp(\beta H(\sigma))$, where $H(\sigma)$ is the number of edges of Λ_n whose endpoints are assigned the same spin in σ; the parameter $\beta > 0$ corresponds to the inverse temperature and controls the strength of the nearest-neighbor interactions.

An Ising/Potts boundary condition τ is a fixed assignment of spins to $\partial\Lambda_n$, the (inner) boundary of Λ_n; i.e., those vertices in Λ_n that are adjacent to vertices in $\mathbb{Z}^2 \setminus \Lambda_n$. The Gibbs distribution on Λ_n conditioned on the fixed assignment τ to $\partial\Lambda_n$, denoted $\mu^\tau_{\Lambda_n}$, is used for example to define the infinite Ising/Potts Gibbs measures on $\mathbb{Z}^2$. These are obtained as the limits of the distributions on finite boxes for distinct boundary conditions τ; i.e, $\lim_{n\to\infty} \mu^\tau_{\Lambda_n}$ for different τ.

On $\mathbb{Z}^2$, it is known that the Ising/Potts model undergoes a sharp phase transition at a critical point $\beta = \beta_c(q) = \ln(1 + \sqrt{q})$ [31, 2]. This phase transition marks the onset of long-range correlations and can also be understood as a transition in the number of (unique vs. multiple) infinite-volume Gibbs measures. In finite regions of $\mathbb{Z}^2$ such as Λ_n, this phase transition corresponds to whether an arbitrary boundary condition τ on $\partial\Lambda_n$ may have macroscopic effects on the Gibbs distribution. For instance, in the low-temperature region $\beta > \beta_c(q)$, if τ is the all "1" configuration on $\partial\Lambda_n$, the spins of all vertices, even those near the center of Λ_n will prefer the spin "1" and thus align with the boundary. In contrast, in the high-temperature region $\beta < \beta_c(q)$, there is exponential decay (with distance) of spin correlations: crucially this holds uniformly over all boundary conditions and over all vertices (i.e., even for those near the boundary); this property is known as *strong spatial mixing (SSM)*.

This phase transition also exhibits itself in the dynamic properties of the system, e.g., through the speed of convergence to stationarity of natural Markov chains for the Ising/Potts model. The classical Glauber dynamics, for example, which in each step updates the spin of a random vertex according to the spins of its neighbors, is known to converge in $\Theta(n^2 \log n)$ steps when $\beta < \beta_c(q)$ [28, 8, 2, 1]; this bound relies on the SSM property described above and, as such, it holds for *every* fixed boundary condition. In contrast, when $\beta > \beta_c(q)$ the speed of convergence of the Glauber dynamics is expected to depend crucially on the boundary condition and understanding its behavior for general boundaries is a long-standing open problem. At the moment, it is known that Glauber dynamics requires exponentially (in n) many steps to converge for free (no boundary) and periodic (toroidal) boundary conditions [34, 7, 16] and, in the special case of the Ising model ($q = 2$), sub-exponentially many steps for uniform (e.g., all "1") boundaries [25, 29].

Our focus here is the random-cluster (FK) model [13], which is a random graph model intimately connected to the Ising/Potts model. Indeed, it has been central to the study of the Ising/Potts phase transition (see, e.g., the recent breakthroughs on $\mathbb{Z}^2$ [2, 12, 11]) and plays an indispensable role in the design of efficient Markov Chain Monte Carlo (MCMC) algorithms for the Ising/Potts model (e.g., the Swendsen-Wang dynamics [33, 22]). We study the effects of boundary conditions on Λ_n on the speed of convergence of the FK-dynamics, the analog of the Ising/Potts Glauber dynamics for the random-cluster model. Despite the close connection between these models, the boundary effects are fundamentally different. Whereas the SSM property of the Ising/Potts model at $\beta < \beta_c(q)$ is uniform over the choice of boundary condition, in the random-cluster setting, SSM is limited to only a few select choices of boundary conditions.

We seek to understand the dynamics in situations where spatial mixing is destroyed near the boundary by the boundary condition. First we establish that for all *realizable* FK boundary conditions (those which are consistent with the planarity of $\mathbb{Z}^2$), the FK-dynamics converges in polynomially many (in n) steps, both at high and low temperatures. To illustrate the difficulties introduced by general boundary conditions, we also construct a class of non-realizable boundary conditions that induce slow (stretched-exponential) convergence at high temperatures.

The random-cluster model. For a graph $G = (V, E)$ and parameters $p \in (0, 1)$ and $q > 0$, random-cluster configurations are subsets of edges in $\Omega = \{S \subseteq E\}$, with the probability of $S \subset E$ given by

$$\pi_{G,p,q}(S) = \frac{1}{Z} p^{|S|}(1-p)^{|E \setminus S|} q^{c(S)} , \tag{1}$$

where $c(S)$ is the number of connected components (including isolated vertices) in the subgraph (V, S), and $Z = Z_{G,p,q}$ is the normalizing constant that makes $\pi_{G,p,q}$ a probability measure.

For *integer* $q \geq 2$ connectivities in the random-cluster model correspond to spin correlations in the Ising/Potts setting, and it is consequently viewed as a generalization of the ferromagnetic Ising/Potts model to non-integer values of q. The random-cluster model, however, is not a spin system in the usual sense, as the weight of a configuration S is not a function of local interactions between edges in G, but instead of global connectivity properties. This non-local structure is a crucial feature of the model but significantly complicates its analysis; for example, it allows boundary conditions to induce long-range connections in G.

We consider the random-cluster model on the $n \times n$ box Λ_n of $\mathbb{Z}^2$, where, for $q \geq 1$, the model is also known to exhibit a phase transition corresponding to the emergence of long-range correlations in the form of large connected components [2]. That is, there exists a critical value $p = p_c(q) = \sqrt{q}/(\sqrt{q}+1)$ such that, with high probability, when $p < p_c(q)$ all connected components are of size $O(\log n)$ whereas when $p > p_c(q)$ there exists a "giant" component of size $\Theta(n^2)$ [2].

A random-cluster boundary condition ξ on $\partial\Lambda_n$ is a partition $\{\xi_1, \xi_2, ...\}$ of the boundary vertices such that all vertices in ξ_i are connected via "ghost" (or external) wirings; these connections are considered in the counting of $c(S)$ in (1) and can therefore impose highly non-local interactions. Of particular interest are boundary conditions where the partition is induced by the connectivity components of a random-cluster configuration on $E(\mathbb{Z}^2) \setminus E(\Lambda_n)$. We call such boundary conditions *realizable*. In fact, many works, including the standard text [21], often restrict attention to realizable boundary conditions, but non-realizable boundary conditions are still relevant in some cases.

The FK-dynamics. In this paper we study the *mixing time* of the FK-dynamics in the presence of boundary conditions. (The mixing time is the number of steps until a Markov chain is close to its stationary distribution in total variation distance, starting from the worst possible initial configuration.) For a configuration $S_t \subseteq E(\Lambda_n)$, a transition $S_t \to S_{t+1} \subseteq E(\Lambda_n)$ of the FK-dynamics is defined as follows:

1. Choose an edge $e \in E(\Lambda_n)$ uniformly at random;
2. let $S_{t+1} = S_t \cup \{e\}$ with probability

$$\frac{\pi_{\Lambda_n,p,q}(S_t \cup \{e\})}{\pi_{\Lambda_n,p,q}(S_t \cup \{e\}) + \pi_{\Lambda_n,p,q}(S_t \setminus \{e\})} = \begin{cases} \frac{p}{q(1-p)+p} & \text{if } e \text{ is a "cut-edge" in } (\Lambda_n, S_t); \\ p & \text{otherwise;} \end{cases}$$

3. else let $S_{t+1} = S_t \setminus \{e\}$.

We say e is a *cut-edge* in (Λ_n, S_t) if the number of connected components in $S_t \cup \{e\}$ and $S_t \setminus \{e\}$ differ. Under a boundary condition ξ, the property of e being a cut-edge is defined with respect to the augmented graph (Λ_n, S_t^ξ). The FK-dynamics converges to (1) by construction, and we study its mixing time. We say the dynamics is *rapidly mixing* if the mixing time is polynomial in $|V|$, and *torpidly mixing* when the mixing time is exponential in $|V|^\varepsilon$ for some $\varepsilon > 0$.

Results. The FK-dynamics is quite powerful since it is known to mix in $\Theta(n^2 \log n)$ steps for all $q > 1$ both at high and low temperatures (i.e., for all $p \neq p_c(q)$) for certain "nice" boundary conditions that do not carry information about random-cluster connectivities in non-local ways: namely, configurations in different regions of Λ_n do not interact through these boundaries [5]. (In comparison, the Ising/Potts Glauber dynamics is torpidly mixing in the low-temperature regime.) Specifically, the tight mixing time bound in [5] holds under boundary conditions that are free (no boundary condition), wired (all boundary vertices are connected to one another) or periodic (the torus). More recently, [15] examined the cutoff phenomenon in the FK-dynamics at $p \ll p_c(q)$; they also restricted attention to periodic boundaries. At the critical $p = p_c(q)$ the FK-dynamics may exhibit torpid mixing depending on the "order" (i.e., the continuity) of the phase transition [16, 18]; notably, when $q \gg 1$ and $p = p_c(q)$, the mixing time may be exponential or sub-exponential depending on the choice of boundary conditions [17].

The stability of the FK-dynamics to the choice of boundary conditions remained unclear at $p \neq p_c(q)$; we show that the FK-dynamics is in fact rapidly mixing for *all* realizable boundary conditions at $p \neq p_c(q)$.

▶ **Theorem 1.1.** *For every $q > 1$, $p \neq p_c(q)$, there exists a constant $C > 0$ such that the mixing time of the FK-dynamics on the $n \times n$ box $\Lambda_n \subset \mathbb{Z}^2$ with any realizable boundary condition is $O(n^C)$.*

We pause to comment on the proof of Theorem 1.1. The proofs of fast mixing when $p \neq p_c(q)$ have relied crucially on a strong spatial mixing property, which in the random-cluster model would say that correlations between edges (even near $\partial\Lambda_n$) decay exponentially in the graph distance between them. It is easy to construct examples of realizable boundary conditions where this correlation does not decay at all, even if $p \ll p_c(q)$, as the boundary can enforce long-range interactions. Since the exponential decay of correlations does hold for edges at distance $\Theta(\log n)$ away from $\partial\Lambda_n$, we are able to reduce the proof of Theorem 1.1 to proving a polynomial upper bound for the mixing time of the FK-dynamics on thin rectangles of dimension $n \times \Theta(\log n)$ with realizable boundary conditions. This reduction is a byproduct of a more general framework we describe in Section 3 for deriving mixing time estimates from spatial mixing properties. The analysis of the FK-dynamics on thin rectangles is then the key technical challenge for us; see Theorem 4.1 and Section 4.1 for a detailed outline of its proof and the novelties therein.

Theorem 1.1 shows a polynomial upper bound on the mixing time, uniformly over all realizable boundary conditions. Utilizing this theorem we can prove near-optimal $\tilde{O}(n^2)$ mixing time for "typical" boundaries. The notion of typicality should be understood as with high probability under some distribution over realizable boundary conditions, with a natural choice being the marginal of the infinite-volume random-cluster measure $\pi_{\mathbb{Z}^2,p,q}$ on $E(\mathbb{Z}^2) \setminus E(\Lambda_n)$ (when $p \neq p_c(q)$ this measure is unique: see [21]).

▶ **Theorem 1.2.** *Let $q > 1$, $p \neq p_c(q)$ and suppose ω is a random-cluster configuration sampled from $\pi_{\mathbb{Z}^2,p,q}$. Let ξ_ω be the boundary condition on $\partial\Lambda_n$ induced by the edges of ω in $\mathbb{Z}^2 \setminus \Lambda_n$. Then, there exists a constant $C > 0$ such that with probability $1 - o(1)$ the mixing time of the FK-dynamics on the $n \times n$ box Λ_n with boundary condition ξ_ω is $O(n^2(\log n)^C)$.*

The proof of Theorem 1.2 uses Theorem 1.1 in a crucial way. Typical boundary conditions do not exhibit the strong spatial mixing property from [5]; however, for such boundary conditions we are able to prove that correlations between edges near the boundary decay exponentially in their graph distance divided by a $\Theta(\log n)$ factor. Using this spatial mixing bound, together with the aforementioned general framework in Section 3, we reduce bounding the mixing time on Λ_n with typical boundaries to bounding the mixing time on $\Theta((\log n)^2) \times \Theta((\log n)^2)$ rectangles with arbitrary realizable boundary conditions. Theorem 1.1 implies that the mixing time of the FK-dynamics in these smaller rectangles is at most poly-logarithmic in n. Similar classes of typical boundary conditions were considered in [18] at $p = p_c(q)$.

Given that our rapid mixing result for realizable boundaries relies heavily on the planarity of the boundary connections in $\mathbb{Z}^2 \setminus \Lambda_n$, one may wonder whether rapid mixing holds for all possible FK boundary conditions (including those not realizable as configurations on $\mathbb{Z}^2 \setminus \Lambda_n$). We answer this in the negative, showing that there exist (non-realizable) boundaries for which the FK-dynamics is torpidly mixing even while $p \neq p_c(q)$. In fact, this torpid mixing holds at $p \ll p_c(q)$, which may sound especially surprising as correlations in the Gibbs measure $\pi_{\Lambda_n,p,q}$ die off faster as p decreases.

▶ **Theorem 1.3.** *Let $q > 2$. For every $\alpha \in (0, \frac{1}{2}]$ and $\lambda > 0$ there exists a boundary condition ξ, such that when $p = \lambda n^{-\alpha}$ the mixing time of the FK-dynamics on the $n \times n$ box Λ_n with boundary condition ξ is $\exp(\Omega(n^\alpha))$.*

Our proof of this theorem is constructive: we take any graph G on m edges for which torpid mixing of the FK-dynamics is known at some value of $p(m) < p_c(q)$ and show how to embed G into the boundary of Λ_n. We then develop a procedure to transfer mixing time bounds from G to Λ_n. The high-level idea is that for sufficiently small $p(m)$ the effect of the configuration away from the boundary is negligible, and the mixing time of the FK-dynamics on G completely governs the mixing time of FK-dynamics near the boundary $\partial\Lambda_n$. Therefore, we can use known torpid mixing results for the mean-field random-cluster model (the case where G is the complete graph) in its critical window at $q > 2$ [20, 4, 14, 19].

We remark that the requirement $q > 2$ appears to be sharp for Theorem 1.3, since it was recently shown that the mixing time of FK-dynamics when $q = 2$ is at most polynomial in the number of vertices on *any* graph and at every $p \in (0, 1)$ [22]. It is expected that this rapid mixing holds for all $q \leq 2$. We believe that our torpid mixing result may extend to small, but $\Omega(1)$ values of $p < p_c(q)$, though our current proof does not allow for this. In principle, one would want to embed a bounded degree graph into $\partial\Lambda_n$, so that the value of p at which it exhibits slow mixing is $\Omega(1)$. There are already several examples of bounded degree graphs where torpid mixing is known [10, 6, 7, 16, 17].

Finally, we remark that by slight adaptations of the comparison results in [35, 36, 4], our theorems translate (up to polynomial factors in n) to bounds for the mixing times of popular non-local dynamics like the Chayes-Machta dynamics [9] and the Swendsen-Wang dynamics on FK configurations [35, 36, 4].

The paper is organized as follows. In Section 2, we define various preliminary notions that are used in our proofs. In Section 3, we introduce a general framework to deduce mixing time estimates on Λ_n from spatial and local mixing properties. We then present our key rapid mixing result for thin rectangles (Theorem 4.1) in Section 4, before completing the proof of Theorem 1.1 in Section 5. The proofs from Section 4 are deferred to Section 6, and those of Theorems 1.2 and 1.3 are included in the full manuscript [3].

2 Preliminaries: the random-cluster model in $\mathbb{Z}^2$

In this section we introduce a number of definitions, notation, and background results that we will refer to repeatedly. More details and proofs can be found in the books [21, 24]. We will be considering the random-cluster model on rectangular subsets of $\mathbb{Z}^2$ of the form $\Lambda_{n,l} = \{0, ..., n\} \times \{0, ..., l\} = [\![0, n]\!] \times [\![0, l]\!]$. When $n = l$, we use Λ_n for $\Lambda_{n,n}$. For simplicity, in this preliminary section we shall focus on the $n = l$ case, but everything stated here holds more generally for rectangular subsets with $n \neq l$.

Abusing notation, we will also use Λ_n for the graph $(\Lambda_n, E(\Lambda_n))$ where $E(\Lambda_n)$ consists of all nearest neighbor pairs of vertices in Λ_n. We denote by $\partial\Lambda_n$ the (inner) boundary of Λ_n; that is the vertex set consisting of all vertices in Λ_n adjacent to vertices in $\mathbb{Z}^2 \setminus \Lambda_n$. A *boundary condition* ξ of Λ_n is a partition of the vertices in $\partial\Lambda_n$. When $u, v \in \partial\Lambda_n$ are in the same element of ξ, we say that they are *wired* in ξ. If there exists a random-cluster configuration ω on $E(\mathbb{Z}^2) \setminus E(\Lambda_n)$ such that $u, v \in \partial\Lambda_n$ are connected in ω if and only if they are wired in ξ, we say that the boundary condition ξ is *realizable*.

For $p \in (0, 1)$ and $q > 0$, the *random-cluster model* on Λ_n with boundary conditions ξ is the probability measure $\pi^{\xi}_{\Lambda_n,p,q}$ over the subsets $S \subseteq E(\Lambda_n)$ given by (1) with the $q^{c(S)}$ term replaced by $q^{c(S;\xi)}$, where $c(S; \xi)$ corresponds to the number of connected components in the augmented graph (Λ_n, S^{ξ}) and S^{ξ} adds auxiliary edges between all pairs of vertices in $\partial\Lambda_n$ that are in the same element of ξ. Every random-cluster configuration $S \subseteq E(\Lambda_n)$, can be identified with some $\omega : E(\Lambda_n) \to \{0, 1\}$ via $\omega(e) = 1$ if $e \in S$ (e is *open*) and $\omega(e) = 0$ if $e \notin S$ (e is *closed*). We sometimes interchange vertex sets with the subgraph they induce; e.g., the random-cluster configuration on a set $B \subset \mathbb{Z}^2$ corresponds to the configuration in the subgraph induced by B. We omit the subscripts p, q when understood from context.

Exponential decay of connectivities (EDC). A consequence of the results in [1, 2] is that for every $q > 1$ and $p < p_c(q)$, there is a $c = c(p, q) > 0$ such that for every boundary condition ξ and all $u, v \in \Lambda_n$,

$$\pi^{\xi}_{\Lambda_n,p,q}(u \stackrel{\Lambda_n}{\longleftrightarrow} v) \le \mathrm{e}^{-cd(u,v)}\,, \tag{2}$$

where $d(u, v)$ is the graph distance between u, v in $\mathbb{Z}^2$ and $u \stackrel{\Lambda_n}{\longleftrightarrow} v$ denotes that there is an open path between u and v in the FK configuration on $E(\Lambda_n)$ (not using the connections of ξ).

Monotonicity. Define a partial order over boundary conditions by $\xi \le \eta$ if the partition corresponding to ξ is *finer* than that of η. The extremal boundary conditions then, are the *free* boundary where $\xi = \{\{v\} : v \in \partial\Lambda_n\}$, which we denote by $\xi = 0$, and the *wired* boundary where $\xi = \{\partial\Lambda_n\}$, denoted by $\xi = 1$. When $q > 1$, the random-cluster model satisfies the following monotonicity in boundary conditions: if ξ, η are two boundary conditions on $\partial\Lambda_n$ with $\xi \le \eta$, then $\pi^{\xi}_{\Lambda_n} \preceq \pi^{\eta}_{\Lambda_n}$, which is to say that $\pi^{\eta}_{\Lambda_n}$ stochastically dominates $\pi^{\xi}_{\Lambda_n}$ with respect to the natural partial order on FK configurations.

Planar duality. Let $\Lambda_n^* = (\Lambda_n^*, E(\Lambda_n^*))$ denote the planar dual of Λ_n. That is, Λ_n^* corresponds to the set of faces of Λ_n, and for each $e \in E(\Lambda_n)$, there is a dual edge $e^* \in E(\Lambda_n^*)$ connecting the two faces bordering e. The random-cluster distribution satisfies $\pi_{\Lambda_n,p,q}(S) = \pi_{\Lambda_n^*,p^*,q}(S^*)$, where S^* is the dual configuration to $S \subseteq E$ (i.e., $e^* \in S^*$ iff $e \notin S$), and $p^* = q(1-p)/(q(1-p)+p)$. Notice that the infinite graph $\mathbb{Z}^2$ is isomorphic to its dual. The unique value of p satisfying $p = p^*$, denoted $p_{sd}(q)$, is called the *self-dual point* and [2] established that $p_c(q) = p_{sd}(q)$; recall that $p_c(q)$ is the critical point for the phase transition in $\mathbb{Z}^2$.

Mixing time and spectral gap. Consider an ergodic (i.e., irreducible and aperiodic) Markov chain $\mathcal{M}$ with finite state space Ω, transition matrix P and stationary distribution μ. The *mixing time* of $\mathcal{M}$ is given by $t_{\rm MIX} := t_{\rm MIX}(1/4)$, where $t_{\rm MIX}(\varepsilon) = \min\{t : \max_{X_0\in\Omega} \|P^t(X_0,\cdot) - \mu\|_{\rm TV} \le \varepsilon\}$ and $\|\cdot\|_{\rm TV}$ denotes total-variation distance. For any positive $\varepsilon < 1/2$, we have $t_{\rm MIX}(\varepsilon) \le \lceil \log_2 \varepsilon^{-1} \rceil \, t_{\rm MIX}$. We use $t_{\rm MIX}(\Lambda_n^{\xi})$ to denote the mixing time of the FK-dynamics on $\Lambda_n \subset \mathbb{Z}^2$ with boundary condition ξ. If P is irreducible and reversible with respect to μ, then it has real eigenvalues $1 = \lambda_1 > \lambda_2 \ge \cdots \ge \lambda_{|\Omega|} \ge -1$. The *spectral gap* of P is defined by $\text{gap}(P) = 1 - \max\{|\lambda_2|, |\lambda_{|\Omega|}|\}$, and the inverse of the spectral gap captures the mixing time up to a $O(\log(\mu_{\min}^{-1}))$ factor, where $\mu_{\min} := \min_{\omega\in\Omega} \mu(\omega)$. For the various dynamics consider in this paper this factor is $\text{poly}(n)$.

FK-dynamics and duality. Each run of the FK-dynamics on Λ_n, with realizable boundary conditions ξ and parameters p, q, determines a valid run of the FK-dynamics on the dual graph Λ_n^* with boundary conditions ξ^* and parameters p^*, q. (Simply identify the FK configuration in each step with its dual configuration; it can be straightforwardly verified that the transitions of the FK-dynamics on the dual graph occur with the correct probabilities.) Hence, the two dynamics have the same mixing times.

▶ Remark 2.1. The edge-set of the dual graph Λ_n^* is not exactly in correspondence with the edge-set of a rectangle $\Lambda^* = \{-\frac{1}{2}, ..., n+\frac{1}{2}\} \times \{-\frac{1}{2}, ..., n+\frac{1}{2}\}$ as it does not include any edges that are between boundary vertices of Λ^*. All the proofs in the paper carry through, only with the natural minor geometric modifications, to the case of rectangles Λ_n with modified edge-set that only contains edges edges with at least one endpoint in $\Lambda_n \setminus \partial\Lambda_n$. The dual of this modified graph is then a $(n-1) \times (n-1)$ rectangle with all nearest-neighbor edges. With these considerations, it often suffices for us to prove our theorems for $p < p_c(q)$. For example, it is sufficient to prove Theorem 1.1 for $p < p_c(q)$.

3 Mixing time upper bounds: a general framework

In this section we introduce a general framework for bounding the mixing time of the FK-dynamics on $\Lambda_n = (\Lambda_n, E(\Lambda_n))$ by its mixing times on certain smaller subsets. In [5] it was shown that a strong form of spatial mixing (encoding exponential decay of correlations uniformly over subsets of Λ_n) implies optimal mixing of the FK-dynamics. However, this notion, known as *strong spatial mixing (SSM)* and described in Remark 3.2, does not hold for many boundary conditions for which fast mixing of the FK-dynamics is still expected. To circumvent this, we introduce a weaker notion, which we call *moderate spatial mixing (MSM)*.

We introduce some notation first. For a set $R \subseteq \Lambda_n$, let $E(R) \subseteq E_n$ be the set of edges of $E(\Lambda_n)$ with both endpoints in R. We will denote by R^c the vertex set $\Lambda_n \setminus R$ and by $E^c(R)$ the edge-complement of R; i.e., $E^c(R) := E(\Lambda_n) \setminus E(R)$. For a configuration $\omega : E(\Lambda_n) \to \{0,1\}$, we will use $\omega(R)$, or alternatively $\omega(E(R))$, for the configuration of ω on $E(R)$. With a slight abuse of notation, for an edge set $F \subseteq E(\Lambda_n)$, we use $\{F = \omega\}$ for the event that the configuration on F is given by ω; when ω is the all free or the all wired configuration, we simply use $\{F = 0\}$ and $\{F = 1\}$, respectively.

▶ **Definition 3.1.** *Let $\mathcal{B} = \{B_1, B_2, \dots, B_k\}$ be a collection of subsets of $\Lambda_n = (\Lambda_n, E(\Lambda_n))$ and let ξ be a boundary condition on Λ_n. We say that* moderate spatial mixing (MSM) *holds on Λ_n for ξ, $\mathcal{B}$ and $\delta > 0$ if for all $e \in E(\Lambda_n)$, there exists $B_j \in \mathcal{B}$ such that*

$$\left|\pi^{\xi}_{\Lambda_n,p,q}(\, e = 1 \mid E(B_j^c) = 1\,) - \pi^{\xi}_{\Lambda_n,p,q}(\, e = 1 \mid E(B_j^c) = 0\,)\right| \;\leq\; \delta\,. \tag{3}$$

In words, MSM holds for $\mathcal{B}$ if for every edge $e \in E(\Lambda_n)$ we can find B_j such that $e \in B_j$ and the "influence" of the configuration on $\Lambda_n \setminus B_j$ on the state of e is bounded by δ.

▶ Remark 3.2. SSM as defined in [5] holds when

MSM holds for a specific sequence of collections of subsets: if $\mathcal{B}_r$ is the set of subsets containing all the square boxes of side length $2r$ centered at each $e \in E(\Lambda_n)$ (intersected with $E(\Lambda_n)$), then SSM holds if MSM holds for $\mathcal{B}_r$ for every $r \geq 1$ with $\delta = \exp(-\Omega(r))$.

MSM does not capture the fast mixing of the FK-dynamics the way SSM does; it is easy to find collections of subsets for which MSM holds for *all* boundary conditions, including those boundary conditions for which we later prove slow mixing (Theorem 1.3). However, if, for a collection $\mathcal{B} = \{B_1, \dots, B_k\}$, we also bound the mixing time of the FK-dynamics on every B_j, we can deduce a mixing time bound for the FK-dynamics on Λ_n. Let $t_{\rm MIX}(B^\tau)$ denote the mixing time of the FK-dynamics on $B \subseteq \Lambda_n$ with boundary condition τ.

▶ **Definition 3.3.** *Let $\mathcal{B} = \{B_1, B_2, \dots, B_k\}$ be a collection of subsets of $\Lambda_n = (\Lambda_n, E(\Lambda_n))$ and let ξ be a boundary condition on Λ_n. We say that* local mixing (LM) *holds for $\mathcal{B}$ and $T > 0$, if*

$$t_{\rm MIX}\big(B_j^{(1,\xi)}\big) \leq T \quad \textit{and} \quad t_{\rm MIX}\big(B_j^{(0,\xi)}\big) \leq T \qquad \textit{for all } j = 1, ..., k$$

where $(1,\xi)$ (resp., $(0,\xi)$) denotes the boundary condition on B_j induced by the event $\{E(B_j^c) = 1\}$ (resp. $\{E(B_j^c) = 0\}$) and the boundary condition ξ.

▶ Remark 3.4. When $B_j \cap \partial\Lambda_n = \emptyset$, $(1,\xi)$ and $(0,\xi)$ are simply the wired and free boundary condition on B_j, respectively. When $B_j \cap \partial\Lambda_n \neq \emptyset$, ξ could also induce some connections in $(1,\xi)$ and $(0,\xi)$.

Our next theorem, roughly speaking, establishes the following implication:

$$\text{MSM} + \text{LM} \implies \text{upper bound for mixing time of FK-dynamics},$$

with the quality of the bound depending on the T for which LM holds. A similar (and inspiring) implication for the Glauber dynamics of the Ising model in graphs of bounded degree was established by Mossel and Sly in [30]; there, the notion of MSM is replaced by a form of spatial mixing which is stronger than SSM. The proof of this theorem is provided in the full version of this paper [3].

▶ **Theorem 3.5.** *Let ξ be a boundary condition on $\Lambda_n = (\Lambda_n, E(\Lambda_n))$ and let $\mathcal{B} = \{B_1, B_2, \dots, B_k\}$ with $B_j \subset \Lambda_n$ for all $j = 1, \dots, k$. If for ξ and $\mathcal{B}$, moderate spatial mixing holds for some $\delta \leq 1/(12|E(\Lambda_n)|)$ and local mixing holds for some $T > 0$, then $t_{\rm MIX}(\Lambda_n^\xi) = O(Tn^2 \log n)$.*

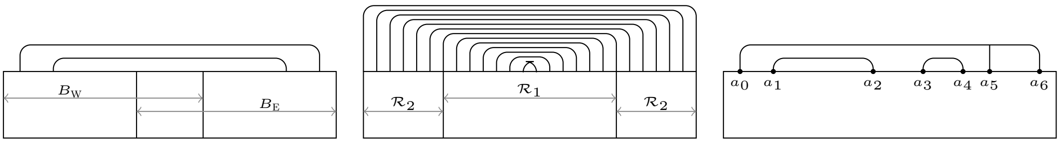

Figure 1 (a) A boundary condition for which no configuration in $B_{\rm W} \cap B_{\rm E}$ isolates $B_{\rm W} \setminus B_{\rm E}$ from $B_{\rm E} \setminus B_{\rm W}$. **(b)** A boundary condition ξ where every pair of overlapping rectangles must interact through ξ; the two *groups of rectangles* $\mathcal{R}_1$, $\mathcal{R}_2$ do not interact through ξ. **(c)** A boundary condition with disconnecting intervals: $[\![a_1, a_4]\!]$ of free-type; $[\![a_1, a_2]\!]$, $[\![a_3, a_4]\!]$, $[\![a_0, a_6]\!]$ of free-wired-type; and $[\![a_0, a_5]\!]$, $[\![a_5, a_6]\!]$ of wired-type.

4 Fast mixing on thin rectangles

The main difficulty in proving Theorem 1.1 using the general framework from Section 3 is obtaining mixing time estimates for the FK-dynamics on thin rectangles of dimension $\Theta(n) \times \Theta(\log n)$ with realizable boundary conditions. To motivate this we notice that in Λ_n, when $p \neq p_c(q)$, the influence of the boundary condition is lost with high probability at a distance $\Theta(\log n)$ from $\partial \Lambda_n$. (This is a consequence of the EDC property when $p < p_c(q)$, or the corresponding dual property when $p > p_c(q)$.) Consequently, the main challenge will be to establish the mixing time of the FK-dynamics in the annulus of width $\Theta(\log n)$ with realizable boundary conditions on the outside. The key ingredient in the proof of Theorem 1.1 will be the following mixing time estimate on thin rectangles. For an $n \times l$ rectangle $\Lambda_{n,l} = [\![0, n]\!] \times [\![0, l]\!]$, let $\partial_{\rm N} \Lambda_{n,l}$, $\partial_{\rm E} \Lambda_{n,l}$, $\partial_{\rm S} \Lambda_{n,l}$ and $\partial_{\rm W} \Lambda_{n,l}$ denote its north, east, south and west boundaries respectively.

▶ **Theorem 4.1.** *Consider $\Lambda_{n,l} = (\Lambda_{n,l}, E(\Lambda_{n,l}))$ for $l \leq n$ with an arbitrary* realizable *boundary condition ξ that is either free or wired on $\partial_{\rm E} \Lambda_{n,l} \cup \partial_{\rm W} \Lambda_{n,l} \cup \partial_{\rm S} \Lambda_{n,l}$. Then, for every $q > 1$ and $p \neq p_c(q)$, the mixing time of the FK-dynamics on $\Lambda_{n,l}$ is at most* $\exp(O(l + \log n))$.

When $l = O(\log n)$, this implies the mixing time is $n^{O(1)}$, which will be the setting of interest in our proofs. Moreover, we note that it suffices to prove Theorem 4.1 for the set of realizable boundary conditions ξ that are free on $\partial_{\rm E} \Lambda_{n,l} \cup \partial_{\rm W} \Lambda_{n,l} \cup \partial_{\rm S} \Lambda_{n,l}$ for all $p \neq p_c(q)$, as the set of boundary conditions dual to these are exactly the set of realizable boundary conditions that are wired on $\partial_{\rm E} \Lambda_{n,l} \cup \partial_{\rm W} \Lambda_{n,l} \cup \partial_{\rm S} \Lambda_{n,l}$; see Remark 2.1.

4.1 Proof of Theorem 4.1

Remarks about previous methods. To motivate our proof approach, we first mention some obstructions that FK boundary conditions present if we tried to adapt methods for the analogous problems in the context of spin systems. A traditional technique to proving mixing time bounds for thin rectangles is the canonical paths method ([26, 27, 23, 32]), which gives an upper bound that is exponential in the shorter side length; however, this approach relies on bounding the *cut-width* of $\Lambda_{n,l}$ which can be significantly distorted in the augmented graph $\Lambda_{n,l}^{\xi}$ by the FK boundary conditions ξ.

A sharper technique is an inductive scheme [8, 27, 18], whereby, the mixing time of the FK-dynamics on the rectangle $\Lambda_{n \times l}$ is bounded by the mixing times in two smaller (overlapping) rectangular subsets, e.g., the left two-thirds $B_{\rm W} = [\![0, \frac{2}{3}n]\!] \times [\![0, l]\!]$ and the right two-thirds $B_{\rm E} = [\![\frac{1}{3}n, n]\!] \times [\![0, l]\!]$; see Figure 1(a). This approach requires bounding the spectral gap of a *block dynamics*, whose updates consist of resampling the configuration on a random block $B_i \in \{B_{\rm W}, B_{\rm E}\}$ from $\pi^{\xi}_{\Lambda_{n,l}}$ conditional on the configuration on $E^c(B_i)$.

It follows from classical results that the spectral gap of the FK-dynamics on $\Lambda_{n,l}$ is bounded from below by the spectral of the block dynamics times the worst spectral gap of the FK-dynamics in any block B_i, assuming a worst-case configuration on $E^c(B_i)$; see, e.g., Proposition 3.4 in [27]. With the choice of blocks $B_{\text{W}}, B_{\text{E}}$, applying this recursively, the spectral gap of the FK-dynamics on $\Lambda_{n,l}$ is bounded from below by the gap of the block dynamics raised to a $\Theta(\log n)$ power. Therefore, establishing Theorem 4.1 requires an $\Omega(1)$ lower bound on the spectral gap of the block dynamics.

The spectral gap of the block dynamics is typically bounded by showing that after the first block update in either B_{W} or B_{E}, the configuration in $B_{\text{W}} \cap B_{\text{E}}$ disconnects the influence of the configuration on $B_{\text{W}} \setminus B_{\text{E}}$ from $B_{\text{E}} \setminus B_{\text{W}}$ with probability $\Omega(1)$. This would then allow a standard coupling argument to lower bound the spectral gap by $\Omega(1)$. In the presence of long-range boundary connections, however, it could be that no configuration on $B_{\text{W}} \cap B_{\text{E}}$ would disconnect the two sides from one another and facilitate coupling; see Figure 1(b) for such an example. As such, our choices of blocks will depend on the boundary conditions and will be chosen to allow for the block dynamics to couple in $O(1)$ time, while ensuring that the blocks are still at most a fraction of the size of the original rectangle, so that after $O(\log n)$ recursive steps we arrive at a sufficiently small base scale.

Definitions and main results for thin rectangles. As Figure 1(b) demonstrates, there are realizable boundary conditions that would force the blocks for the block dynamics to not be single rectangles, but rather unions of rectangular subsets of $\Lambda_{n,l}$ of the form $R = [\![a, b]\!] \times [\![0, l]\!]$ with $0 \le a < b \le n$; for ease of notation, let $[\![a, b]\!]^c = [\![0, n]\!] \setminus [\![a, b]\!]$. Our recursive argument will proceed instead on *groups of rectangles*.

► **Definition 4.2.** *Let $m = C_\star \log l$ where $C_\star$ is a suitably large constant. A* group of rectangles *$\mathcal{R} = \bigcup_{i=1}^{N(\mathcal{R})} R_i$ is the union of $N(\mathcal{R})$ disjoint rectangular subsets $R_i = [\![a_i, b_i]\!] \times [\![0, l]\!]$ of $\Lambda_{n,l}$ such that $W(R_i) := b_i - a_i \ge 2m$ for every $i = 1, ..., N(\mathcal{R})$.*

The requirement that the width $W(R_i)$ of every constituent rectangle R_i is at least $2m$, is so that the interior of the R_i's are can be isolated from the configuration on $E(\Lambda_{n,l}) \setminus E(\mathcal{R})$. Indeed, the constant $C_\star$ is chosen so that $C_\star > c^{-1}$ with c being the constant from the EDC property (2).

We show that for every group of rectangle $\mathcal{R}$ there is a choice of two suitable blocks, which in turn will be group of rectangles, for the block dynamics. By suitable we mean two group of rectangles whose width are a constant fraction of that of $\mathcal{R}$ and that are sufficiently isolated from one another in ξ; see Proposition 4.6. (The width of a group of rectangles $\mathcal{R} = \bigcup_{i=1}^{N(\mathcal{R})} R_i$, denoted $W(\mathcal{R})$, is simply the sum of the width of its constituent rectangles; that is $W(\mathcal{R}) = \sum_{i=1}^{N(\mathcal{R})} W(R_i)$.)

For this, we introduce the key notions of *disconnecting intervals* of a boundary condition ξ and *compatibility* of a group of rectangles $\mathcal{R} \subset \Lambda_{n,l}$ with ξ. These allow us to manage the unwieldy interactions that may be induced by the realizable boundary condition ξ. Roughly speaking, a disconnecting interval is a segment $[\![a, b]\!] \times \ell$ of $\partial_{\text{N}}\Lambda_{n,l}$ that has no interaction through ξ with the remaining vertices in $\partial_{\text{N}}\Lambda_{n,l}$.

► **Definition 4.3.** *For a* realizable *boundary condition ξ on $\Lambda_{n,l}$ that is free on $\partial_{\text{E}}\Lambda_{n,l} \cup \partial_{\text{S}}\Lambda_{n,l} \cup \partial_{\text{W}}\Lambda_{n,l}$, an interval $[\![a, b]\!] \subset [\![0, n]\!]$ is called* disconnecting *of*

1. free-type*: if there are no boundary connections in ξ between $[\![a, b]\!] \times \{l\}$ and $[\![a, b]\!]^c \times \{l\}$.*
2. wired-type*: if there is a boundary component in ξ that contains both vertices (a, l) and (b, l).*

Observe that an interval can be both of free-type and of wired-type if (a, l) and (b, l) are connected through ξ but are not connected to any boundary vertex in $[\![a, b]\!]^c \times [\![0, l]\!]$; in this case, we may refer to the interval as being of *free-wired-type*; see Figure 1(c) for several examples.

We say a group of rectangles $\mathcal{R}$ is compatible with ξ if the boundary interactions between the rectangular subsets of $\mathcal{R}$ are limited in the following way.

▶ **Definition 4.4.** *Let $\mathcal{R} = \bigcup_{i=1}^{N(\mathcal{R})} R_i$ be a group of rectangles with $R_i = [\![a_i, b_i]\!] \times [\![0, l]\!]$ and $a_1 < b_1 < \ldots < a_{N(\mathcal{R})} < b_{N(\mathcal{R})}$. Let ξ be a realizable boundary condition on $\Lambda_{n,l}$ that is free on $\partial_{\rm S}\Lambda_{n,l} \cup \partial_{\rm E}\Lambda_{n,l} \cup \partial_{\rm W}\Lambda_{n,l}$, and free in all vertices in $\partial_{\rm N}\Lambda_{n,l}$ at distance at most m from $\partial_{\rm E}\Lambda_{n,l} \cup \partial_{\rm W}\Lambda_{n,l}$.*

We say $\mathcal{R}$ is compatible *with ξ, if*

1. *Between every two consecutive rectangles $R_i = [\![a_i, b_i]\!] \times [\![0, l]\!]$ and $R_{i+1} = [\![a_{i+1}, b_{i+1}]\!] \times [\![0, l]\!]$ the interval $[\![b_i - m, a_{i+1} + m]\!]$ is a disconnecting interval; and*
2. *The interval $[\![a_1 + m, b_{N(\mathcal{R})} - m]\!]$ is also a disconnecting interval.*

It is clear from the definition that $\Lambda_{n,l}$ is compatible with ξ: the first condition is vacuous, while the second is satisfied by the additional assumption that all vertices a distance at most m from $\partial_{\rm E}\Lambda_{n,l} \cup \partial_{\rm W}\Lambda_{n,l}$ are free (i.e., they appear as singletons in the corresponding boundary partition)

With the definition of group of rectangles, disconnecting intervals and compatibility in hand, we can now design a "splitting" algorithm for picking two blocks $\mathcal{R}_{\rm INT}, \mathcal{R}_{\rm EXT}$ for the block dynamics with the desired properties. The following lemma, proved in Section 6, provides the basis of such an algorithm.

▶ **Lemma 4.5.** *Let ξ be a realizable boundary condition on $\partial\Lambda_{n,l}$ that is free on $\partial_{\rm S}\Lambda_{n,l} \cup \partial_{\rm E}\Lambda_{n,l} \cup \partial_{\rm W}\Lambda_{n,l}$ and free in all vertices in $\partial_{\rm N}\Lambda_{n,l}$ at distance at most m from $\partial_{\rm E}\Lambda_{n,l} \cup \partial_{\rm W}\Lambda_{n,l}$. For every group of rectangles $\mathcal{R} = \bigcup_{i=1}^{N(\mathcal{R})} R_i$ compatible with ξ, with $W(\mathcal{R}) \geq 100m$, there exists a disconnecting interval $[\![c_\star, d_\star]\!]$ such that $(c_\star, l), (d_\star, l) \in \partial_{\rm N}\mathcal{R}$, are distance at least m from the vertical sides $\bigcup_{i=1}^{N(\mathcal{R})} \partial_{\rm W}R_i \cup \partial_{\rm E}R_i$ of $\mathcal{R}$, and*

$$\frac{1}{4}W(\mathcal{R}) \leq W(\mathcal{R} \cap ([\![c_\star, d_\star]\!] \times [\![0, l]\!])) \leq \frac{3}{4}W(\mathcal{R})\,.$$

With the disconnecting interval $[\![c_\star, d_\star]\!]$ from Lemma 4.5, we define $\mathcal{A}_{\rm INT} = \mathcal{R} \cap ([\![c_\star, d_\star]\!] \times [\![0, l]\!])$ and $\mathcal{A}_{\rm EXT} = \mathcal{R} \cap ([\![c_\star, d_\star]\!]^c \times [\![0, l]\!])$. Their enlargements by m will form the blocks $\mathcal{R}_{\rm INT}$ and $\mathcal{R}_{\rm EXT}$:

$$\mathcal{R}_{\rm INT} = \mathcal{R} \cap ([\![c_\star - m, d_\star + m]\!] \times [\![0, l]\!]) \quad \text{and} \quad \mathcal{R}_{\rm EXT} = \mathcal{R} \cap (([\![0, c_\star + m]\!] \cup [\![d_\star - m, n]\!]) \times [\![0, l]\!])\,;$$

These sets are depicted in Figure 2(a)–(b). The requirement that the corners of $[\![c_\star, d_\star]\!] \times [\![0, l]\!]$ are a distance at least m from the vertical sides of $\mathcal{R}$ is so that when we enlarge the sets $\mathcal{A}_{\rm INT}, \mathcal{A}_{\rm EXT}$ by m, we do not overflow beyond the rectangles containing $(c_\star, l)$ and $(d_\star, l)$. Crucially, our ability to pick disconnecting segments that satisfy this requirement is guaranteed by the compatibility of $\mathcal{R}$ with ξ.

It follows from Lemma 4.5, and the definitions of disconnecting interval and compatibility, that $\mathcal{R}_{\rm INT}$ and $\mathcal{R}_{\rm EXT}$ have the following properties, which will facilitate our recursive argument to prove Theorem 4.1.

▶ **Proposition 4.6.** *If $\mathcal{R}$ is a group of rectangles compatible with ξ, and moreover, $W(\mathcal{R}) \geq 100m$, then the sets $\mathcal{R}_{\rm INT}$ and $\mathcal{R}_{\rm EXT}$ are groups of rectangles satisfying the following properties:*

1. *$\frac{1}{5}W(\mathcal{R}) \leq W(\mathcal{R}_{\rm INT}) \leq \frac{4}{5}W(\mathcal{R})$ and likewise $\frac{1}{5}W(\mathcal{R}) \leq W(\mathcal{R}_{\rm EXT}) \leq \frac{4}{5}W(\mathcal{R})$;*
2. *Both $\mathcal{R}_{\rm INT}$ and $\mathcal{R}_{\rm EXT}$ are compatible with ξ.*

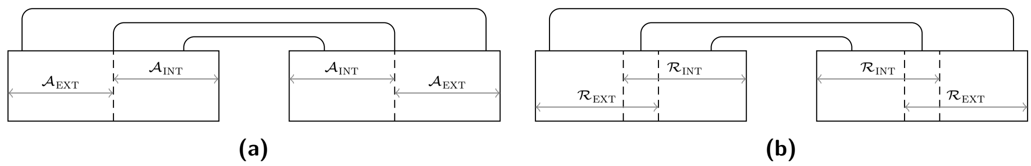

Figure 2 **(a)** The cores $\mathcal{A}_{\text{INT}}$ and $\mathcal{A}_{\text{EXT}}$ and **(b)** the blocks $\mathcal{R}_{\text{INT}}$ and $\mathcal{R}_{\text{EXT}}$. The blocks $\mathcal{R}_{\text{INT}}$ and $\mathcal{R}_{\text{EXT}}$ are the enlargements of $\mathcal{A}_{\text{INT}}$ and $\mathcal{A}_{\text{EXT}}$ by exactly m, and are thus, themselves, groups of rectangles.

Finally, we consider the spectral gap of the block dynamics $\{X_t\}$ on $\mathcal{R}$ with blocks $\mathscr{B} = \{\mathcal{R}_{\text{INT}}, \mathcal{R}_{\text{EXT}}\}$. In this case, $\{X_t\}$ is the discrete-time Markov chain that in each step picks i uniformly at random from $\{\text{INT}, \text{EXT}\}$ and updates the configuration in $E(\mathcal{R}_i)$ with a sample from the stationary distribution of the chain conditional on the configuration on $E^c(\mathcal{R}_i)$. Let $\text{gap}(\mathcal{R}^\zeta; \mathscr{B})$ be the spectral gap of this block dynamics on the group of rectangle $\mathcal{R}$ with boundary condition ζ induced on $\mathcal{R}$ by ξ and a fixed random-cluster configuration $\omega_{\mathcal{R}^c}$ on $E^c(\mathcal{R}) = E(\Lambda_{n,l}) \setminus E(\mathcal{R})$; hence, we may identify ζ with the pair $(\xi, \omega_{\mathcal{R}^c})$.

▶ **Lemma 4.7.** *Let ξ be a realizable boundary condition on $\partial\Lambda_{n,l}$ that is free on $\partial_{\text{S}}\Lambda_{n,l} \cup \partial_{\text{E}}\Lambda_{n,l} \cup \partial_{\text{W}}\Lambda_{n,l}$ and free on vertices in $\partial_{\text{N}}\Lambda_{n,l}$ at distance at most m from $\partial_{\text{E}}\Lambda_{n,l} \cup \partial_{\text{W}}\Lambda_{n,l}$. For every $q > 1$ and $p \neq p_c(q)$, there exists $K = K(p,q) \geq 1$ such that for every group of rectangles $\mathcal{R}$ compatible with ξ, and every configuration $\omega_{\mathcal{R}^c}$ on $E^c(\mathcal{R})$, we have* $\text{gap}(\mathcal{R}^{(\xi, \omega(\mathcal{R}^c))}; \mathscr{B}) \geq K^{-1}$.

The proof of Lemma 4.7 is deferred to Section 6. We are now ready to prove Theorem 4.1.

Proof of Theorem 4.1. Fix $q > 1$, $p \neq p_c(q)$ and $\Lambda_{n,l}$ with a realizable boundary condition ξ' that is *free* on $\partial_{\text{E}}\Lambda_{n,l} \cup \partial_{\text{S}}\Lambda_{n,l} \cup \partial_{\text{W}}\Lambda_{n,l}$. We modify ξ' to a boundary condition ξ that is also free on all vertices a distance at most $m = C_\star \log l$ from $\partial_{\text{E}}\Lambda_{n,l} \cup \partial_{\text{W}}\Lambda_{n,l}$ at a cost of an exponential in m factor in the mixing time of the FK-dynamics (see Lemma 2.3 in [3]). Let ξ be the resulting realizable boundary condition.

Let $\mathcal{R} \subset \Lambda_{n,l}$ be a group of rectangles that is compatible with ξ and has $W(\mathcal{R}) = s$ for $100m \leq s \leq n$. Let $\mathcal{R}_{\text{INT}}$ and $\mathcal{R}_{\text{EXT}}$ be the group rectangles defined earlier and consider the block dynamics with respect to these blocks. Recall that we use $\text{gap}(\mathcal{R}^\zeta)$ and $\text{gap}(\mathcal{R}^\zeta; \mathscr{B})$ for the spectral gaps of the FK-dynamics and the blocks dynamics with respect to $\mathscr{B} = \{\mathcal{R}_{\text{INT}}, \mathcal{R}_{\text{EXT}}\}$ respectively. As discussed earlier, for any boundary condition $\zeta = (\xi, \omega_{\mathcal{R}}^c)$, Proposition 3.4 from [27] implies that for a suitable constant $\gamma \in (0,1)$

$$\text{gap}\left(\mathcal{R}^\zeta\right) \geq \gamma \cdot \text{gap}\left(\mathcal{R}^\zeta; \mathscr{B}\right) \cdot \min_{\substack{i \in \{\text{INT}, \text{EXT}\} \\ \omega \in \Omega(R_i^c)}} \text{gap}\left(\mathcal{R}_i^{(\xi,\omega)}\right) \geq \frac{\gamma}{K} \cdot \min_{\substack{i \in \{\text{INT}, \text{EXT}\} \\ \omega \in \Omega(R_i^c)}} \text{gap}\left(\mathcal{R}_i^{(\xi,\omega)}\right), \quad (4)$$

where the second inequality follows from Lemma 4.7 and $\Omega(R_i^c)$ is the set of FK configurations on $E(R_i^c)$. Observe that Proposition 4.6 implies that $\max\{W(\mathcal{R}_{\text{INT}}), W(\mathcal{R}_{\text{EXT}})\} \leq 4s/5$. Therefore, applying (4) $O(\log n)$ times, we deduce that $\text{gap}(\mathcal{R}^\zeta) \geq \exp(\Omega(-\log n)) \cdot \text{gap}(\mathcal{R}_0^{\zeta_0})$, where $\mathcal{R}_0$ is a group of rectangles with $W(\mathcal{R}_0) \leq 100m$ and $\zeta_0 = (\xi, \omega_0)$ is an arbitrary boundary condition for $\mathcal{R}_0$.

Finally, since $|\partial\mathcal{R}_0| = O(m + l) = O(l)$, the lower bound for $\text{gap}(\mathcal{R}_0^{\zeta_0})$ follows from the following crude argument. Observe that we can first modify the boundary condition ζ_0 to be all free on all of $\partial\mathcal{R}_0$, incurring a cost of a $q^{-\Omega(l)}$ factor in the spectral gap; see Lemma 2.3 in [3]. The fast mixing result from [5] for the free boundary condition then implies that $\text{gap}(\mathcal{R}_0^{\zeta_0}) \geq \exp\left(-\Omega(l)\right)$ and so the result follows. ◀

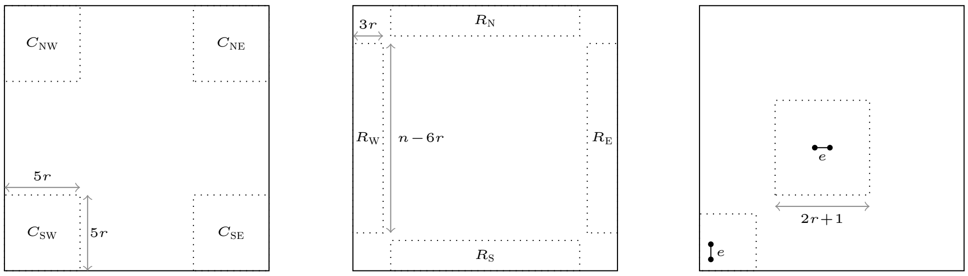

Figure 3 Subsets (a) C_{NE}, C_{NW}, C_{SE}, and C_{SW} and (b) R_{N}, R_{E}, R_{W} and R_{S}. **(c)** $B(e,r)$ for two edges e of Λ_n.

5 Polynomial mixing time for realizable boundary conditions

In this section we finalize the proof of Theorem 1.1 for $p < p_c(q)$ using the technology introduced in Section 3; namely, we construct a collection of subsets $\mathcal{B}$ for which we can establish LM and MSM; see Definitions 3.1–3.3. To establish LM we crucially use Theorem 4.1. The results for $p > p_c(q)$ follow from the self-duality of the model and of realizable boundary conditions, as explained in Section 2.

For general realizable boundary conditions, proving LM for a collection of subsets $\mathcal{B}$ for which MSM holds is the main challenge. This is because, for MSM to hold for a collection $\mathcal{B}$ for all realizable boundary conditions, a subset in $\mathcal{B}$ needs to contain $\Omega(n)$ edges. In particular, some element of $\mathcal{B}$ must include most (or all) edges near $\partial\Lambda_n$, as otherwise it is easy to construct examples of realizable boundary conditions for which MSM does not hold. Thus, a trivial (exponential in the perimeter) upper bound for the mixing time on those subsets with $\Omega(n)$ edges would be unhelpful, and we ought to use Theorem 4.1 instead.

We define a collection of blocks for which we can establish both LM and MSM. Let $r \in \mathbb{N}$ and let $C_{\text{NE}}, C_{\text{NW}}, C_{\text{SE}}, C_{\text{SW}} \subset \Lambda_n$ be the four square boxes of side length $5r$ with a corner that coincides with a corner of Λ_n; see Figure 3(a). Let $R_{\text{N}} \subset \Lambda_n$ be the $(n-6r) \times 2r$ rectangle at distance $3r$ from both $\partial_{\text{W}}\Lambda_n$, $\partial_{\text{E}}\Lambda_n$ whose top boundary is contained in $\partial_{\text{N}}\Lambda_n$ and let $R_{\text{E}}, R_{\text{W}}, R_{\text{S}}$ be defined analogously; see Figure 3(b). Let $R = R_{\text{N}} \cup R_{\text{E}} \cup R_{\text{W}} \cup R_{\text{S}}$. Now, for $e \in E(\Lambda_n)$, let $B(e,r) \subset \Lambda_n$ be the set of vertices in the minimal square box around e such that $d(\{e\}, \Lambda_n \setminus B(e,r)) \geq r$. If $d(\{e\}, \partial\Lambda_n) > r$, then $B(e,r)$ is just a square box of side length $2r+1$ centered at e; otherwise $B(e,r)$ intersects $\partial\Lambda_n$; see Figure 3(c). Finally, let

$$\mathcal{B}_r = \{C_{\text{NE}}, C_{\text{NW}}, C_{\text{SE}}, C_{\text{SW}}, R\} \cup \{B(e,r) : e \in E(\Lambda_n), d(\{e\}, \partial\Lambda_n) > r\}. \tag{5}$$

We claim that LM holds for $\mathcal{B}_r$ with $r = \Theta(\log n)$ and $T = O(n^C)$ for some constant $C > 0$.

▶ **Theorem 5.1.** *Let $q \geq 1$, $p < p_c(q)$ and $r = c_0 \log n$ with $c_0 > 0$ independent of n. There exists a constant $C > 0$ such that LM holds for every realizable boundary condition ξ and $\mathcal{B}_r$ with $T = O(n^C)$.*

The subsets $B(e,r)$ in $\mathcal{B}_r$ and the corner boxes C_{NE}, C_{NW}, C_{SE} and C_{SW} are small enough that crude bounds for their mixing times are sufficient. As mentioned earlier, the main challenge for proving local mixing for $\mathcal{B}_r$ is to derive a mixing time bound for $R = R_{\text{N}} \cup R_{\text{E}} \cup R_{\text{W}} \cup R_{\text{S}}$ as it intersects the boundary of Λ_n and contains $\Omega(n)$ vertices. To establish such a bound we rely on Theorem 4.1. In particular, we relate the mixing time of the FK-dynamics on R

to that of the FK-dynamics on a single thin rectangle by concatenating the four rectangles constituting R, one after another, such that the union of their outer boundaries make up the northern boundary of the new rectangle.

The final ingredient of the proof is establishing MSM for the collection $\mathcal{B}_r$. We show that MSM holds for $\mathcal{B}_r$ with $r = \Theta(\log n)$ for all realizable boundary conditions ξ where the vertices in $\partial\Lambda_n$ at distance $5r$ from the corners of Λ_n are free in ξ. This is sufficient since any realizable boundary condition can be turned into a realizable boundary condition with this property by simply removing all connections in ξ involving vertices near the corners of Λ_n; this modification can change the mixing time of the FK-dynamics by a factor of at most $\exp(O(r))$; see Lemma 2.3 in [3].

▶ **Theorem 5.2.** *Let $q \geq 1$, $p < p_c(q)$ and $r = c_0 \log n$ with $c_0 > 0$ independent of n. Let ξ be a realizable boundary condition with the property that every vertex $v \in \partial\Lambda_n$ at distance at most $5r$ from a corner of Λ_n is free in ξ. Then, for all sufficiently large $c_0 > 0$, MSM holds for ξ and $\mathcal{B}_r$ with $\delta < 1/(12|E(\Lambda_n)|)$.*

Theorem 1.1 follows from Theorems 3.5 and 5.1–5.2. Their proofs are found in the full version [3].

6 Proofs from Section 4

We prove here the two key results from Section 4: Lemmas 4.5 and 4.7. We begin with the proof of Lemma 4.5 which describes how to find an appropriate disconnecting interval for the block dynamics in a group of rectangles $\mathcal{R}$. The proof uses the following important geometric observations regarding disconnecting intervals. For more details we refer to [3].

▶ **Lemma 6.1.** *Let ξ be a realizable boundary condition on $\Lambda_{n,l}$ that is free on $\partial_{\mathrm{S}}\Lambda_{n,l} \cup \partial_{\mathrm{E}}\Lambda_{n,l} \cup \partial_{\mathrm{W}}\Lambda_{n,l}$ and let $a < b < c$. If both $[\![a,b]\!]$ and $[\![b,c]\!]$ are disconnecting intervals of wired-type, then so is $[\![a,c]\!]$. If both $[\![a,b]\!]$ and $[\![b+1,c]\!]$ are disconnecting intervals of free-type, then so is $[\![a,c]\!]$.*

▶ **Lemma 6.2.** *Let ξ be a realizable boundary condition on $\Lambda_{n,l}$ that is free on $\partial_{\mathrm{S}}\Lambda_{n,l} \cup \partial_{\mathrm{E}}\Lambda_{n,l} \cup \partial_{\mathrm{W}}\Lambda_{n,l}$. Suppose there exist $a < b < c < d$ such that $[\![a,c]\!]$ and $[\![b,d]\!]$ are disconnecting intervals; then either both are of free-type or both are of wired-type. Moreover, if both are*

1. *of* wired-type*: then $[\![a,b]\!], [\![b,c]\!], [\![c,d]\!]$ and $[\![a,d]\!]$ are all disconnecting intervals of wired-type.*
2. *of* free-type*: then $[\![a,b-1]\!], [\![b,c]\!], [\![c+1,d]\!]$ and $[\![a,d]\!]$ are all disconnecting intervals of free-type.*

Proof of Lemma 4.5. We find a candidate disconnecting interval $[\![c,d]\!]$ with $(c,l),(d,l) \in \partial_{\mathrm{N}}\mathcal{R}$ satisfying:

$$\tfrac{1}{3}W(\mathcal{R}) \leq W(\mathcal{R} \cap ([\![c,d]\!] \times [\![0,l]\!])) \leq \tfrac{2}{3}W(\mathcal{R})\,. \tag{6}$$

In the second part of the proof we show how to modify $[\![c,d]\!]$ to obtain a disconnecting interval $[\![c_\star,d_\star]\!]$ with the added property that both $(c_\star,l)$ and $(d_\star,l)$ are distance at least m from $\partial_{\|}\mathcal{R} := \bigcup_{i=1}^{N(\mathcal{R})} \partial_{\mathrm{W}}R_i \cup \partial_{\mathrm{E}}R_i$.

If there exist vertices $(x,l),(y,l) \in \partial_{\mathrm{N}}\mathcal{R}$ such that $\frac{1}{3}W(\mathcal{R}) \leq W(\mathcal{R} \cap ([\![x,y]\!] \times [\![0,l]\!])) \leq \frac{2}{3}W(\mathcal{R})$ with (x,l) connected to (y,l) through ξ, then we take $c = x$, $d = y$ and use $[\![c,d]\!] = [\![x,y]\!]$ as our candidate disconnecting interval. Suppose otherwise that there does not exist any such boundary connection: then every pair $(x,l),(y,l) \in \partial_{\mathrm{N}}\mathcal{R}$ connected through ξ is such that

$$W(\mathcal{R} \cap ([\![x,y]\!] \times [\![0,l]\!])) < \tfrac{1}{3}W(\mathcal{R})\,, \qquad \text{or} \qquad W(\mathcal{R} \cap ([\![x,y]\!] \times [\![0,l]\!])) > \tfrac{2}{3}W(\mathcal{R})\,. \tag{7}$$

If the latter holds, then there is a pair, say $(x_0,l),(y_0,l) \in \partial_{\mathrm{N}}\mathcal{R}$, for which the latter holds with a minimal width. Consequently, all other connections through ξ between vertices $(x_1,l),(y_1,l) \in \partial_{\mathrm{N}}\mathcal{R} \cap ([\![x_0+1,y_0-1]\!] \times [\![0,l]\!])$ will be such that $W(\mathcal{R} \cap ([\![x_1,y_1]\!] \times [\![0,l]\!])) < \frac{1}{3}W(\mathcal{R})$. We can then partition the vertices of $\partial_{\mathrm{N}}\mathcal{R} \cap ([\![x_0+1,y_0-1]\!] \times \{l\})$ into disjoint disconnecting intervals of free-wired-type as follows:

1. Let $\rho = \{C_1,\dots,C_k\}$ be the partition of $\partial_{\mathrm{N}}\mathcal{R} \cap ([\![x_0+1,y_0-1]\!] \times \{l\})$ induced by ξ;
2. For each C_i, consider the disconnecting interval L_i of free-wired-type determined by the left-most and right-most vertices of C_i in $\partial_{\mathrm{N}}\mathcal{R} \cap ([\![x_0+1,y_0-1]\!] \times \{l\})$ (n.b. these may be singletons);
3. Let $\{L_{i_1},...,L_{i_\ell}\}$ be those which are maximal, i.e., there does not exist j and k such that $L_{i_j} \subset L_k$.

The set of disconnecting intervals $\{L_{i_1},\dots,L_{i_\ell}\}$ partitions $[\![x_0+1,y_0-1]\!]$ into disjoint disconnecting intervals of free-wired-type with the property that $W(\mathcal{R} \cap (L_{i_j} \times [\![0,l]\!])) \le \frac{1}{3}W(\mathcal{R})$ for every $j \in \{1,\dots,\ell\}$. We can then use Lemma 6.1 to merge adjacent disconnecting intervals until we obtain a candidate disconnecting interval $[\![c,d]\!] \subset [\![x_0,y_0]\!]$ (of free-type), having width $W(\mathcal{R} \cap ([\![c,d]\!] \times [\![0,l]\!]) \in [\frac{1}{3}W(\mathcal{R}), \frac{2}{3}W(\mathcal{R})]$.

Now that we have found a candidate disconnecting interval $[\![c,d]\!]$ satisfying (6), we modify it to obtain a disconnecting interval $[\![c_\star,d_\star]\!]$ with the property that both $(c_\star,l)$, $(d_\star,l)$ are distance at least m from $\partial_{\|}\mathcal{R}$.

If (c,l) is at distance at least m from $\partial_{\|}\mathcal{R}$, set $c_\star = c$, and similarly if (d,l) is at distance at least m from $\partial_{\|}\mathcal{R}$, then set $d_\star = d$. Otherwise, suppose (c,l) is at distance less than m from $\partial_{\mathrm{W}}R_i$ for some constituent rectangular subset $R_i = [\![a_i,b_i]\!] \times [\![0,l]\!]$ of $\mathcal{R}$. Since $\mathcal{R}$ is compatible with ξ, the interval $\mathcal{I}_c = [\![b_{i-1}-m, a_i+m]\!]$ is a disconnecting interval, and we set

$$c_\star = \begin{cases} a_i + m\,, & \text{if } \mathcal{I}_c \text{ is of wired-type, or } i=1, \text{ or } W(R_i)=2m; \\ a_i+m+1\,, & \text{if } \mathcal{I}_c \text{ is only of free-type, and } W(R_i) > 2m; \end{cases} \,.$$

When (c,l) is at distance less than m from $\partial_{\mathrm{E}}R_i$ for some i, then we simply set $c_\star = b_i - m$. Symmetrically, if (d,l) is at distance less than m from $\partial_{\mathrm{E}}R_i$ for $R_i = [\![a_i,b_i]\!] \times [\![0,l]\!]$, let $\mathcal{I}_d = [\![b_i - m, a_{i+1}+m]\!]$

$$d_\star = \begin{cases} b_i - m\,, & \text{if } \mathcal{I}_d \text{ is of wired-type, or } i=N(\mathcal{R}), \text{ or } W(R_i)=2m; \\ b_i-m-1\,, & \text{if } \mathcal{I}_d \text{ is only of free-type, and } W(R_i) > 2m. \end{cases}$$

When (d,l) is at distance less than m from $\partial_{\mathrm{W}}R_i$, let $d_\star = a_i + m$. Since $W(R_i) \ge 2m$ for every i, the points $(c,l),(d,l)$ cannot be both less than m away from $\partial_{\mathrm{E}}R_i$ and less than m away from $\partial_{\mathrm{W}}R_i$.

One can check via a case analysis, exploiting the compatibility of $\mathcal{R}$ with ξ and using Lemma 6.2, that in all of these cases the interval $[\![c_\star,d_\star]\!]$ is a disconnecting interval; we defer these details to the full manuscript [3]. The fact that $(c_\star,l),(d_\star,l) \in \partial_{\mathrm{N}}\mathcal{R}$ are a distance at least m away from $\partial_{\|}\mathcal{R}$ follows directly from the construction. Finally, we claim that in all such situations, $[\![c_\star,d_\star]\!]$ satisfies

$$\tfrac{1}{4}W(\mathcal{R}) \le W(\mathcal{R} \cap ([\![c_\star,d_\star]\!] \times [\![0,l]\!])) \le \tfrac{3}{4}W(\mathcal{R})\,.$$

This follows from the facts that $W(\mathcal{R}) \ge 100m$, $|c - c_\star| \le m$ and $|d - d_\star| \le m$. ◀

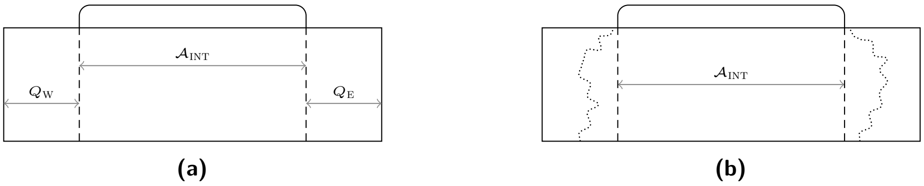

Figure 4 **(a)** The block $\mathcal{R}_{\text{INT}}$ with its subsets $\mathcal{A}_{\text{INT}}$, Q_{W} and Q_{E}. **(b)** The block $\mathcal{R}_{\text{INT}}$ with the dual-paths (dotted) of a configuration in Γ allowing coupling inside $\mathcal{A}_{\text{INT}}$.

We proceed with the proof of Lemma 4.7, where we establish a lower bound for the spectral gap of the block dynamics in the group of rectangles $\mathcal{R}$ with blocks $\mathcal{R}_{\text{INT}}$ and $\mathcal{R}_{\text{EXT}}$ as defined in Section 4.1.

Proof of Lemma 4.7. We consider the $p < p_c(q)$ case; the case of $p > p_c(q)$ follows from a similar (dual) argument which we defer to the full manuscript [3]. Let $\{X_t\}, \{Y_t\}$ be two instances of the block dynamics on $\mathcal{R}$ with boundary condition $\zeta = (\xi, \omega_{\mathcal{R}^c})$ started from initial configurations X_0, Y_0. It suffices to construct a coupling $\mathbb{P}$ of the steps of $\{X_t\}, \{Y_t\}$ such that $\min_{X_0,Y_0} \mathbb{P}(X_2 = Y_2) = \Omega(1)$; see [24].

With probability 1/4 the first block to be updated is $\mathcal{R}_{\text{INT}}$ and the second is $\mathcal{R}_{\text{EXT}}$. Suppose this is the case and let us consider the update on $\mathcal{R}_{\text{INT}}$. Let θ_ω be the boundary condition on $\partial\mathcal{R}_{\text{INT}}$ induced by ζ and the restriction of a configuration ω to $E(\mathcal{R}) \setminus E(\mathcal{R}_{\text{INT}})$, and let π^{θ_ω} be the FK distribution on $\mathcal{R}_{\text{INT}}$ with boundary conditions θ_ω. As $X_1(\mathcal{R}_{\text{INT}}), Y_1(\mathcal{R}_{\text{INT}})$ have laws $\pi^{\theta_{X_0}}$, $\pi^{\theta_{Y_0}}$, respectively, a coupling for $\pi^{\theta_{X_0}}$ and $\pi^{\theta_{Y_0}}$ yields a coupling for X_1 and Y_1.

We describe next such a coupling for $\pi^{\theta_{X_0}}$ and $\pi^{\theta_{Y_0}}$. Let $Q_{\text{W}}, Q_{\text{E}} \subset \mathcal{R}_{\text{INT}}$ be the two rectangles of width m that contain all the vertices in $\mathcal{R}_{\text{INT}} \setminus \mathcal{A}_{\text{INT}}$; i.e., $Q_{\text{W}} \cup \mathcal{A}_{\text{INT}} \cup Q_{\text{E}} = \mathcal{R}_{\text{INT}}$, $Q_{\text{W}} \cap \mathcal{A}_{\text{INT}} = \emptyset$ and $Q_{\text{E}} \cap \mathcal{A}_{\text{INT}} = \emptyset$ (see Figure 4(a)). Let $\partial E(Q_{\text{W}})$ be the set of edges with one endpoint in Q_{W} and the other in $\mathcal{A}_{\text{INT}}$, and similarly define $\partial E(Q_{\text{E}})$. Let Γ_{W} be the set of configurations in $\mathcal{R}_{\text{INT}}$ that have a *dual-path* in $E(Q_{\text{W}}) \cup \partial E(Q_{\text{W}})$ connecting the top-most edge in $\partial E(Q_{\text{W}})$ to an edge in $\partial_{\text{S}} Q_{\text{W}}$, and similarly define Γ_{E} as the set of configurations in $\mathcal{R}_{\text{INT}}$ that have a dual-path in $E(Q_{\text{E}}) \cup \partial E(Q_{\text{E}})$ from the top-most edge in $\partial E(Q_{\text{E}})$ to an edge in $\partial_{\text{S}} Q_{\text{E}}$. (A dual-path is an open path in the dual configuration.) Let $\Gamma = \Gamma_{\text{E}} \cap \Gamma_{\text{W}}$; see Figure 4(b). Let θ_1 be the boundary condition on $\partial\mathcal{R}_{\text{INT}}$ induced by ζ and the wired configuration on $E(\mathcal{R}) \setminus E(\mathcal{R}_{\text{INT}})$. The following lemma supplies the desired coupling.

▶ **Lemma 6.3.** *Let $q > 1$ and $p < p_c(q)$. There exists a coupling $\mathbb{P}_1$ of the distributions $\pi^{\theta_{X_0}}$, $\pi^{\theta_{Y_0}}$, π^{θ_1} such that if $(\omega^{\theta_X}, \omega^{\theta_Y}, \omega^{\theta_1})$ is sampled from $\mathbb{P}_1$, the following hold:*
1. $\mathbb{P}_1(\omega^{\theta_X}(\mathcal{A}_{\text{INT}}) = \omega^{\theta_Y}(\mathcal{A}_{\text{INT}}) \mid \omega^{\theta_1} \in \Gamma) = 1$*;*
2. *There exists a constant $\rho = \rho(p,q) > 0$ such that $\mathbb{P}_1(\omega^{\theta_1} \in \Gamma) \geq \rho$.*

If we use the coupling $\mathbb{P}_1$ from Lemma 6.3 to couple the first step of the chains, then X_1 and Y_1 will agree on $E(\mathcal{A}_{\text{INT}})$ with probability at least $\rho > 0$. If this occurs, then we can couple the update on $\mathcal{R}_{\text{EXT}}$ in the second step so that $X_2 = Y_2$ with probability one. This is because $X_1(E(\mathcal{A}_{\text{INT}})) = Y_1(E(\mathcal{A}_{\text{INT}}))$ implies $X_1(E(\mathcal{R}) \setminus E(\mathcal{R}_{\text{EXT}})) = Y_1(E(\mathcal{R}) \setminus E(\mathcal{R}_{\text{EXT}}))$, and thus the boundary conditions induced by the two instances of the chain on $\mathcal{R}_{\text{EXT}}$ are identical. As a consequence, we obtain that for any X_0, Y_0, $\mathbb{P}(X_2 = Y_2) \geq \rho/4$, which concludes the proof for $p < p_c(q)$. ◀

We conclude this section with the proof of Lemma 6.3.

Proof of Lemma 6.3. Let $L = \partial_{\mathrm{W}} Q_{\mathrm{W}} \cup \partial_{\mathrm{N}} Q_{\mathrm{W}} \cup \partial_{\mathrm{E}} Q_{\mathrm{E}} \cup \partial_{\mathrm{N}} Q_{\mathrm{E}}$. For a configuration ω on $\mathcal{R}_{\mathrm{INT}}$ let $F(\omega) := \mathcal{R}_{\mathrm{INT}} \setminus \bigcup_{v \in L} C(v,\omega)$, where $C(v,\omega)$ is the vertex set of the connected component of v in ω, ignoring the boundary connections. Note that $\omega \in \Gamma$ if and only if the vertices in the boundary components of L, i.e., $\bigcup_{v \in L} C(v,\omega)$, are confined to $Q_{\mathrm{W}} \cup Q_{\mathrm{E}}$, in which case $\mathcal{A}_{\mathrm{INT}} \subseteq F(\omega)$.

Clearly, $\pi^{\theta_1} \succeq \pi^{\theta_X}$ and $\pi^{\theta_1} \succeq \pi^{\theta_Y}$ and thus there exist monotone couplings $\mathbb{P}_X$ (resp., $\mathbb{P}_Y$) for π^{θ_X} and π^{θ_1} (resp., π^{θ_Y} and π^{θ_1}). The coupling $\mathbb{P}_1$ is defined as follows. First sample $(\omega^{\theta_X}, \omega^{\theta_1})$ from $\mathbb{P}_X$ and ω^{θ_Y} from $\mathbb{P}_Y(\,\cdot \mid \omega^{\theta_1}\,)$. If $\mathcal{A}_{\mathrm{INT}} \subseteq F(\omega^{\theta_1})$, then re-sample the configuration on $E(F(\omega^{\theta_1}))$ in ω^{θ_1} and update $\omega^{\theta_X}(F(\omega^{\theta_1}))$ and $\omega^{\theta_Y}(F(\omega^{\theta_1}))$ such that $\omega^{\theta_1}(F(\omega^{\theta_1})) = \omega^{\theta_X}(F(\omega^{\theta_1})) = \omega^{\theta_Y}(F(\omega^{\theta_1}))$.

To deduce part 1, it now suffices to show that if $\mathcal{A}_{\mathrm{INT}} \subseteq F(\omega^{\theta_1})$ the three boundary conditions η_1, η_X, η_Y induced on $\partial F(\omega^{\theta_1})$ by the configurations of ω^{θ_X}, ω^{θ_Y}, ω^{θ_1} on $E(\mathcal{R}_{\mathrm{INT}}) \setminus E(F(\omega^{\theta_1}))$, respectively, and the corresponding boundary conditions θ_X, θ_Y, θ_1 are identical; if this is the case part 1 follows from the domain Markov property (see [21]).

First observe that the boundary condition on $\partial_{\mathrm{S}} \mathcal{A}_{\mathrm{INT}}$ is always free by assumption. Also from the definition of $F(\omega^{\theta_1})$ every edge of $E(\mathcal{R}_{\mathrm{INT}}) \setminus E(F(\omega^{\theta_1}))$ incident to $\partial F(\omega^{\theta_1})$ is closed in ω^{θ_1}, so by the monotonicity of the coupling, the same holds for ω^{θ_X} and ω^{θ_Y}. The remaining portion of $\partial F(\omega^{\theta_1})$ is precisely the set of vertices $(\partial \mathcal{A}_{\mathrm{INT}} \cap \partial \mathcal{R}) \setminus \partial_{\mathrm{S}} \mathcal{R}$. In order for the boundary conditions η_1, η_x, η_Y to differ on this set, there must be at least two distinct boundary components in $\zeta = (\xi, \omega_{\mathcal{R}^c})$ between $(\partial \mathcal{A}_{\mathrm{INT}} \cap \partial \mathcal{R}) \setminus \partial_{\mathrm{S}} \mathcal{R}$ and $[\![c_\star, d_\star]\!]^c \times \{l\}$; this cannot happen because $[\![c_\star, d_\star]\!]$ is disconnecting.

Part 2 of the lemma is a straightforward consequence of the EDC property of the random-cluster model at $p < p_c(q)$; see (2). Namely, since the width of Q_{W} is $m = C_\star \log l$, (2) implies that when $C_\star$ is large enough there is a constant $\rho_{\mathrm{W}}(p,q) > 0$ such that $\pi^{\theta_1}(\Gamma_{\mathrm{W}}^c) \le \pi_\Delta^1(L \overset{\Delta}{\longleftrightarrow} \partial_{\mathrm{W}} \mathcal{A}_{\mathrm{INT}}) \le 1 - \rho_{\mathrm{W}}$, where Δ is the subgraph induced by the edges in $E(\mathcal{R}_{\mathrm{INT}}) \setminus E(\mathcal{A}_{\mathrm{INT}})$. A matching bound holds for Γ_{E}^c. Since Γ_{E}, Γ_{W} are both decreasing events, by the FKG inequality (see [21]), $\pi^{\theta_1}(\Gamma) \ge \rho_{\mathrm{W}} \rho_{\mathrm{E}} =: \rho$, concluding the proof. ◀

References

1. Kenneth S. Alexander. On weak mixing in lattice models. *Probab. Theory Related Fields*, 110(4):441–471, 1998. `doi:10.1007/s004400050155`.
2. Vincent Beffara and Hugo Duminil-Copin. The self-dual point of the two-dimensional random-cluster model is critical for $q \ge 1$. *Probab. Theory Related Fields*, 153(3-4):511–542, 2012. `doi:10.1007/s00440-011-0353-8`.
3. Antonio Blanca, Reza Gheissari, and Eric Vigoda. Random-cluster dynamics in $\mathbb{Z}^2$: rapid mixing with general boundary conditions. *preprint available at* `arXiv:1807.08722.`, 2018.
4. Antonio Blanca and Alistair Sinclair. Dynamics for the Mean-field Random-cluster Model. In *Proc. of the 19th International Workshop on Randomization and Computation (RANDOM 2015)*, pages 528–543, 2015. `doi:10.4230/LIPIcs.APPROX-RANDOM.2015.528`.
5. Antonio Blanca and Alistair Sinclair. Random-Cluster Dynamics in $\mathbb{Z}^2$. *Probab. Theory Related Fields*, 2016. Extended abstract appeared in Proc. of the 27th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2016), pp. 498–513. `doi:10.1007/s00440-016-0725-1`.
6. Christian Borgs, Jennifer T. Chayes, Alan Frieze, Jeong H. Kim, Prasad Tetali, Eric Vigoda, and Van H. Vu. Torpid mixing of some Monte Carlo Markov chain algorithms in statistical physics. In *Proc. of the 40th Annual Symposium on Foundations of Computer Science (FOCS 1999)*, pages 218–229, 1999. `doi:10.1109/SFFCS.1999.814594`.
7. Christian Borgs, Jennifer T. Chayes, and Prasad Tetali. Tight bounds for mixing of the Swendsen-Wang algorithm at the Potts transition point. *Probab. Theory Related Fields*, 152(3-4):509–557, 2012. `doi:10.1007/s00440-010-0329-0`.

8 Filippo Cesi. Quasi-factorization of the entropy and logarithmic Sobolev inequalities for Gibbs random fields. *Probability Theory and Related Fields*, 120(4):569–584, August 2001. doi:10.1007/PL00008792.

9 Lincoln Chayes and Jon Machta. Graphical representations and cluster algorithms I. Discrete spin systems. *Physica A: Statistical Mechanics and its Applications*, 239(4):542–601, 1997.

10 Colin Cooper and Alan M. Frieze. Mixing properties of the Swendsen–Wang process on classes of graphs. *Random Structures and Algorithms*, 15:242–261, 1999.

11 Hugo Duminil Copin, Maxim Gagnebin, Matan Harel, Ioan Manolescu, and Vincent Tassion. Discontinuity of the phase transition for the planar random-cluster and Potts models with $q > 4$. *CoRR*, 2016. arXiv:1611.09877.

12 Hugo Duminil-Copin, Vladas Sidoravicius, and Vincent Tassion. Continuity of the Phase Transition for Planar Random-Cluster and Potts Models with $1 \leq q \leq 4$. *Communications in Mathematical Physics*, 349(1):47–107, January 2017. doi:10.1007/s00220-016-2759-8.

13 Cornelius M. Fortuin and Pieter W. Kasteleyn. On the random-cluster model. I. Introduction and relation to other models. *Physica*, 57:536–564, 1972.

14 Andreas Galanis, Daniel Štefankovic, and Eric Vigoda. Swendsen-Wang Algorithm on the Mean-Field Potts Model. In *Proc. of the 19th International Workshop on Randomization and Computation (RANDOM 2015)*, pages 815–828, 2015. doi:10.4230/LIPIcs.APPROX-RANDOM.2015.815.

15 Shirshendu Ganguly and Insuk Seo. Information Percolation and Cutoff for the Random-Cluster Model. *CoRR*, 2018. arXiv:1812.01538.

16 Reza Gheissari and Eyal Lubetzky. Mixing Times of Critical Two-Dimensional Potts Models. *Comm. Pure Appl. Math*, 71(5):994–1046, 2018.

17 Reza Gheissari and Eyal Lubetzky. The effect of boundary conditions on mixing of 2D Potts models at discontinuous phase transitions. *Electron. J. Probab.*, 23:30 pp., 2018. doi:10.1214/18-EJP180.

18 Reza Gheissari and Eyal Lubetzky. Quasi-polynomial mixing of critical two-dimensional random cluster models. *Random Structures and Algorithms*, 2019. doi:10.1002/rsa.20868.

19 Reza Gheissari, Eyal Lubetzky, and Yuval Peres. Exponentially slow mixing in the mean-field Swendsen–Wang dynamics. *Annales de l'Institut Henri Poincare (B)*, 2019. to appear. Extended abstract appeared in Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2018), pp. 1981–1988.

20 Vivek K. Gore and Mark R. Jerrum. The Swendsen-Wang process does not always mix rapidly. *J. Statist. Phys.*, 97(1-2):67–86, 1999. doi:10.1023/A:1004610900745.

21 Geoffrey Grimmett. The random-cluster model. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 73–123. Springer, Berlin, 2004. doi:10.1007/978-3-662-09444-0_2.

22 Heng Guo and Mark Jerrum. Random cluster dynamics for the Ising model is rapidly mixing. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017, Barcelona, Spain, Hotel Porta Fira, January 16-19*, pages 1818–1827, 2017. doi:10.1137/1.9781611974782.118.

23 Mark Jerrum and Alistair Sinclair. Approximating the permanent. *SIAM J. Comput.*, 18(6):1149–1178, 1989. doi:10.1137/0218077.

24 David A. Levin, Malwina J. Luczak, and Yuval Peres. Glauber dynamics for the mean-field Ising model: cut-off, critical power law, and metastability. *Probab. Theory Related Fields*, 146(1-2):223–265, 2010. doi:10.1007/s00440-008-0189-z.

25 Eyal Lubetzky, Fabio Martinelli, Allan Sly, and Fabio Lucio Toninelli. Quasi-polynomial mixing of the 2D stochastic Ising model with "plus" boundary up to criticality. *J. Eur. Math. Soc. (JEMS)*, 15(2):339–386, 2013. doi:10.4171/JEMS/363.

26 Fabio Martinelli. On the two-dimensional dynamical Ising model in the phase coexistence region. *J. Statist. Phys.*, 76(5-6):1179–1246, 1994. doi:10.1007/BF02187060.

27 Fabio Martinelli. Lectures on Glauber dynamics for discrete spin models. In *Lectures on probability theory and statistics (Saint-Flour, 1997)*, volume 1717 of *Lecture Notes in Math.*, pages 93–191. Springer, Berlin, 1999. `doi:10.1007/978-3-540-48115-7_2`.

28 Fabio Martinelli, Enzo Olivieri, and Roberto H. Schonmann. For 2-D lattice spin systems weak mixing implies strong mixing. *Comm. Math. Phys.*, 165(1):33–47, 1994. URL: `http://projecteuclid.org/euclid.cmp/1104271032`.

29 Fabio Martinelli and Fabio Lucio Toninelli. On the mixing time of the 2D stochastic Ising model with "plus" boundary conditions at low temperature. *Comm. Math. Phys.*, 296(1):175–213, 2010. `doi:10.1007/s00220-009-0963-5`.

30 Elchanan Mossel and Allan Sly. Exact thresholds for Ising–Gibbs samplers on general graphs. *Ann. Probab.*, 41(1):294–328, January 2013. `doi:10.1214/11-AOP737`.

31 Lars Onsager. Crystal Statistics. I. A Two-Dimensional Model with an Order-Disorder Transition. *Phys. Rev.*, 65:117–149, February 1944. `doi:10.1103/PhysRev.65.117`.

32 Alistair Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combin. Probab. Comput.*, 1(4):351–370, 1992. `doi:10.1017/S0963548300000390`.

33 Robert H. Swendsen and Jian-Sheng Wang. Nonuniversal critical dynamics in Monte Carlo simulations. *Phys. Rev. Lett.*, 58:86–88, January 1987. `doi:10.1103/PhysRevLett.58.86`.

34 Lawrence E. Thomas. Bound on the mass gap for finite volume stochastic Ising models at low temperature. *Comm. Math. Phys.*, 126(1):1–11, 1989. URL: `http://projecteuclid.org/euclid.cmp/1104179720`.

35 Mario Ullrich. Comparison of Swendsen-Wang and heat-bath dynamics. *Random Structures and Algorithms*, 42(4):520–535, 2013. `doi:10.1002/rsa.20431`.

36 Mario Ullrich. Rapid mixing of Swendsen-Wang dynamics in two dimensions. *Dissertationes Math. (Rozprawy Mat.)*, 502:64, 2014. `doi:10.4064/dm502-0-1`.