

RANDOM-CLUSTER DYNAMICS IN \mathbb{Z}^2 : RAPID MIXING WITH GENERAL BOUNDARY CONDITIONS

BY ANTONIO BLANCA^{1,*}, REZA GHEISSARI² AND ERIC VIGODA^{1,**}

¹*School of Computer Science, Georgia Tech, * ablanca@cc.gatech.edu; ** vigoda@cc.gatech.edu*

²*Courant Institute, New York University, reza@cims.nyu.edu*

The random-cluster model with parameters (p, q) is a random graph model that generalizes bond percolation ($q = 1$) and the Ising and Potts models ($q \geq 2$). We study its Glauber dynamics on $n \times n$ boxes Λ_n of the integer lattice graph \mathbb{Z}^2 , where the model exhibits a sharp phase transition at $p = p_c(q)$. Unlike traditional spin systems like the Ising and Potts models, the random-cluster model has non-local interactions. Long-range interactions can be imposed as external connections in the boundary of Λ_n , known as *boundary conditions*. For select boundary conditions that do not carry long-range information (namely, wired and free), Blanca and Sinclair proved that when $q > 1$ and $p \neq p_c(q)$, the Glauber dynamics on Λ_n mixes in optimal $O(n^2 \log n)$ time. In this paper, we prove that this mixing time is polynomial in n for every boundary condition that is *realizable* as a configuration on $\mathbb{Z}^2 \setminus \Lambda_n$. We then use this to prove near-optimal $\tilde{O}(n^2)$ mixing time for “typical” boundary conditions. As a complementary result, we construct classes of nonrealizable (nonplanar) boundary conditions inducing slow (stretched-exponential) mixing at $p \ll p_c$.

1. Introduction. The *random-cluster model* [13] is a well studied generalization of independent bond percolation, with connections to the study of electrical networks and random spanning trees [21]. Perhaps most importantly, the random-cluster model is closely related to the *ferromagnetic Ising–Potts model*, which is the classical model for spin systems in statistical physics. In fact, the random-cluster model is often referred to as the FK-representation of this model and is a key tool in the analysis of its phase transition; see, for example, the recent breakthroughs on the infinite 2-dimensional integer lattice graph \mathbb{Z}^2 [2, 10, 11]. It also plays an indispensable role in the design of efficient Markov Chain Monte Carlo (MCMC) algorithms, like the Swendsen–Wang algorithm [34], for studying the Ising–Potts 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 $S \subseteq E$ given by

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

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

In addition to being interesting in its own right as a random graph and dependent percolation model, the random-cluster model provides a unifying framework for the study of several important probabilistic models. For example, for *integer* $q \geq 2$ the random-cluster model is dual, in a precise sense, to the q -state Ising–Potts model, where configurations are assignments of spin values $\{1, \dots, q\}$ to the vertices of G . Each configuration $\sigma \in [q]^V$ has probability $\mu_G(\sigma) \propto \exp(\beta H(\sigma))$, where $H(\sigma)$ is the number of edges connecting vertices with

Received July 2018; revised May 2019.

MSC2010 subject classifications. Primary 60K35; secondary 82B20, 82C20.

Key words and phrases. Random-cluster model, Glauber dynamics, mixing time, spatial mixing.

the same spin values, and $\beta > 0$ is a model parameter associated with the inverse temperature of the system. When $\beta = -\ln(1 - p) > 0$, it is straightforward to check via a probabilistic coupling [12] that correlations in the Ising–Potts model correspond to connectivities in the random-cluster setting; this has illuminated much of the study of these models and has made the random-cluster model an accepted generalization of the Ising–Potts model to noninteger values of q .

The random-cluster model, however, is not a spin system in the usual sense: the weight of a configuration S is not a function of local interactions between edges in G , but instead of global connectivity properties. This nonlocality is a crucial feature of the model but significantly complicates its analysis. The present paper studies the influence of this nonlocality on the speed of convergence of the Glauber dynamics (i.e., local Markov chains) on subsets of \mathbb{Z}^2 , where the model has been of particular interest.

On the infinite graph \mathbb{Z}^2 , both the random-cluster and the Ising–Potts model undergo phase transitions corresponding to the sudden emergence of long-range correlations as some parameter of the system is continuously varied (p and β in this case). A classical result of Onsager [31] established that the Ising model (the $q = 2$ case) undergoes a phase transition at a critical $\beta = \beta_c(2) = \ln(1 + \sqrt{2})$. Recently, it was shown in a celebrated result of Beffara and Duminil-Copin [2] that the Potts model ($q \geq 3$ integer), and more generally the random-cluster model for any $q > 1$, also undergo phase transitions at the critical points $\beta_c(q) = \ln(1 + \sqrt{q})$ and $p_c(q) = \sqrt{q}/(\sqrt{q} + 1)$, respectively. Note that $\beta_c(q) = -\ln(1 - p_c(q))$. These phase transitions can also be understood as transitions in the number of (unique vs. multiple) infinite-volume Gibbs measures on \mathbb{Z}^2 . The infinite-volume measures on \mathbb{Z}^2 are obtained by taking limits of the distributions on finite boxes with a sequence of “boundary conditions.”

We begin with the notion of boundary conditions in the Ising–Potts model. Let Λ_n be a $n \times n$ square region of \mathbb{Z}^2 with nearest-neighbor edges $E(\Lambda_n)$, and let $\partial\Lambda_n$ be its (inner) boundary (i.e., those vertices in Λ_n that are adjacent to vertices in $\mathbb{Z}^2 \setminus \Lambda_n$). An Ising–Potts boundary condition τ is a fixed assignment of spins to $\partial\Lambda_n$, and $\mu_{\Lambda_n}^\tau$ is the Gibbs distribution on Λ_n conditional on the assignment τ to $\partial\Lambda_n$. (Since the interactions are nearest-neighbor, this is the same as conditioning on a configuration on all of $\mathbb{Z}^2 \setminus \Lambda_n$.)

For the random-cluster model on Λ_n , a boundary condition ξ on $\partial\Lambda_n$ is a partition $\{\xi_1, \xi_2, \dots\}$ of the boundary vertices such that all vertices in ξ_i are always in the same connected component of a configuration S via “ghost” (or external) wirings; these connections are considered in the counting of $c(S)$ in (1) and can therefore impose highly nonlocal interactions. Of particular interest are boundary conditions for the random-cluster model corresponding to configurations on $\mathbb{Z}^2 \setminus \Lambda_n$: that is, where the boundary partition is induced by the connections 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.) We note that boundary conditions are fundamental to the study of static and dynamic properties of the random-cluster and Ising–Potts models, especially in finite subsets of \mathbb{Z}^2 .

In this paper, we consider the Glauber dynamics for the random-cluster model on the finite subgraph $(\Lambda_n, E(\Lambda_n))$ of \mathbb{Z}^2 , in the presence of boundary conditions. This Markov chain is of significant interest: it provides a simple Markov chain Monte Carlo (MCMC) algorithm for sampling configurations of the system; and is, in many cases, a plausible model for the evolution of the underlying system.

Specifically, we consider the following discrete-time variant of the Glauber dynamics chain, which we refer to as the *FK-dynamics*. For $t \in \mathbb{N}$, from $S_t \subseteq E(\Lambda_n)$, transition to $S_{t+1} \subseteq E(\Lambda_n)$ 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) , where S_t^ξ adds external wirings between all pairs of vertices in the same element of ξ . This Markov chain converges to $\pi_{\Lambda_n, p, q}$ and we study its speed of convergence.

A standard measure for quantifying the speed of convergence of a Markov chain is the *mixing time*: the time until the dynamics is close (in total variation distance) to its stationary distribution, starting from a worst-case initial state. 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 $|V|^\epsilon$ for some $\epsilon > 0$.

The corresponding Glauber dynamics for the Ising–Potts model (which updates spins one at a time according to the spins of their neighbors), is by now quite well understood on finite regions of \mathbb{Z}^2 . In the high-temperature region $\beta < \beta_c$ (corresponding to $p < p_c$), the Glauber dynamics has optimal mixing time $\Theta(n^2 \log n)$ on boxes Λ_n [1, 2, 7, 28]. These bounds rely on the exponential decay of correlations of the model at $\beta < \beta_c$, which holds even near the boundary for arbitrary Ising–Potts boundary conditions; this property is known as *strong spatial mixing*. Therefore, this $\Theta(n^2 \log n)$ mixing time bound holds for *every* boundary condition. In contrast, when $\beta > \beta_c$, the speed of convergence of the Glauber dynamics is expected to depend crucially on the boundary conditions and understanding its behavior for general boundaries is a long-standing open problem. At the moment, it is known that Glauber dynamics is torpidly mixing for free (no boundary) and periodic (toroidal) boundary conditions [6, 16, 35] and in the special case of the Ising model ($q = 2$), has a subexponential mixing time for uniform (e.g., all “1”) boundaries [25, 29].

The FK-dynamics is quite powerful since the self-duality of the model on \mathbb{Z}^2 implies that it is rapidly mixing in the low-temperature regime where the Ising–Potts Glauber dynamics is torpidly mixing. For the FK-dynamics on Λ_n , [4] showed that the mixing time is $\Theta(n^2 \log n)$ for all $q > 1$ whenever $p \neq p_c(q)$ for certain “nice” boundary conditions that have the relevant strong spatial mixing property. Specifically, the tight mixing time bound in [4] holds under boundary conditions that are free, wired or periodic. More recently, [15] examined the cutoff phenomenon in the FK-dynamics at $p \ll p_c(q)$; they also restricted attention to periodic boundaries. The behavior of the FK-dynamics under arbitrary random-cluster boundary conditions (not having strong spatial mixing) remained unclear. At the critical $p = p_c(q)$, the FK dynamics may exhibit torpid mixing depending on the “order” of the phase transition [16, 17]; notably when $q \gg 1$ and $p = p_c(q)$, the mixing time may be exponential or subexponential depending on the boundary conditions.

We prove that the FK-dynamics at $p \neq p_c(q)$ is rapidly mixing for all realizable boundary conditions.

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. As mentioned above, 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 the boundary $\partial \Lambda_n$) 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 in the “bulk” of Λ_n (i.e., at distance $\Theta(\log n)$ away from its boundary), 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 will be the key technical difficulty for us and is established in Theorem 4.1.

In the setting of spin systems, or boundary conditions that do not encode long-range interactions, a polynomial upper bound on $n \times \Theta(\log n)$ rectangles would follow from standard canonical paths arguments [23, 26, 27, 33]. However, even realizable boundary conditions can heavily distort the graph with external wirings, preventing this approach from succeeding. Instead, to prove Theorem 4.1 we devise a novel application of a recursive (block dynamics) scheme common in the analysis of spin systems. In lieu of splitting rectangles into two overlapping subrectangles, for example, the first two-thirds and the second two-thirds, as is done for spin systems, our choice of smaller subsets at every step of the recursion is delicately dictated by the boundary conditions. As a consequence, the subsets we recurse over are no longer restricted to rectangles, but can be arbitrary “groups of rectangles” that are “compatible” with the realizable boundary conditions. We point the reader to Section 4 for a more detailed proof overview.

Theorem 1.1 shows a polynomial upper bound on the mixing time, uniformly over all realizable boundary conditions. Utilizing this theorem, we prove near-optimal $O(|V|(\log |V|)^C)$ mixing time for “typical” boundaries as we detail now. The notion of typicality should be understood as with high probability under some probability distribution over realizable boundary conditions, with a natural choice being the marginal distribution of the infinite random-cluster measure $\pi_{\mathbb{Z}^2, p, q}$ on $\mathbb{Z}^2 \setminus \Lambda_n$.

As mentioned earlier, a key obstacle to proving mixing time upper bounds are long, distinct boundary connections that enforce long-range correlations. For any realizable boundary condition ξ corresponding to a partition $\{\xi_1, \xi_2, \dots\}$ of $\partial \Lambda_n$, let $L(\xi_i)$ be the smallest connected subgraph of $\partial \Lambda_n$ containing all vertices in ξ_i . The key class of boundary conditions we consider are those where all the $L(\xi_i)$ are small.

DEFINITION 1.2. We say a boundary condition ξ on $\partial \Lambda_n$ with corresponding partition $\{\xi_1, \xi_2, \dots\}$ is in \mathcal{C}_α if $\max_i |L(\xi_i)| \leq \alpha \log n$. We say that a realizable boundary condition ξ is in \mathcal{C}_α^* if its dual boundary condition ξ^* is in \mathcal{C}_α ; see Section 2 for the definition of dual configuration. We refer to the classes \mathcal{C}_α and \mathcal{C}_α^* as α -localized boundaries.

It is straightforward to see that if one samples a “random” boundary condition from the infinite-volume measure $\pi_{\mathbb{Z}^2, p, q}$, then with high probability, the induced boundary condition on $\partial \Lambda_n$ is α -localized for some $\alpha > 0$. Since there is a unique random-cluster measure on \mathbb{Z}^2 when $p \neq p_c(q)$, this is well defined.

THEOREM 1.3. For all $q > 1$, $p \neq p_c(q)$, let ω be a random-cluster configuration sampled from $\pi_{\mathbb{Z}^2, p, q}$. If ξ_ω is the boundary condition on $\partial \Lambda_n$ induced by the connections of ω in $E(\mathbb{Z}^2) \setminus E(\Lambda_n)$, then with probability $1 - o(1)$, ξ_ω is α -localized for $\alpha > 0$ sufficiently large.

There are other similar ways of defining typicality of realizable boundary conditions that may also be of interest. For example, if ω were sampled from the random-cluster measure in a large concentric box containing Λ_n , with probability $1 - o(1)$, ξ_ω would again be α -localized; see Remark 8.

We are able to prove nearly-optimal mixing time $\tilde{O}(n^2)$ for α -localized boundaries, and hence, with high probability, for a random boundary sampled from *any* off-critical random-cluster measure.

THEOREM 1.4. *For every $q > 1$, $p < p_c(q)$ (resp., $p > p_c(q)$), and every $\alpha > 0$, there exists a constant $C > 0$ such that for every realizable boundary condition $\xi \in \mathcal{C}_\alpha$ (resp., $\xi \in \mathcal{C}_\alpha^*$) on $\partial\Lambda_n$, the mixing time of the FK-dynamics on the $n \times n$ box Λ_n with boundary condition ξ is $O(n^2(\log n)^C)$.*

The proof of Theorem 1.4 uses Theorem 1.1 in a crucial way. Typical boundary conditions do not exhibit the strong spatial mixing property from [4]; however, for boundary conditions in C_α we are able to prove that correlations between edges near the boundary decay exponentially in their graph distance divided by $\alpha \log n$. Using this correlation bound, together with the aforementioned general framework in Section 3 to derive mixing time estimates from spatial mixing properties, 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 then 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 [16, 17] at $p = p_c(q)$; there, comparison methods were used to disregard the influence of long boundary connections at the expense of super-polynomial factors in the mixing time.

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 (nonrealizable) 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 1.5. *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 so the mixing time of the FK-dynamics on G completely governs the mixing time of FK-dynamics near the boundary $\partial\Lambda_n$. We can then 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 $q > 2$ [3, 14, 19, 20].

We remark about the condition $q > 2$ in Theorem 1.5. In [22], it was 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 $p \in (0, 1)$. It is believed that this rapid mixing holds for all $q \leq 2$; hence the requirement $q > 2$ appears to be sharp for Theorem 1.5. We believe that the above torpid mixing result may also 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 its critical point at which it exhibits slow mixing is $\Omega(1)$. Note that there are several examples of bounded degree graphs where torpid mixing is known [5, 6, 9, 16, 18].

Finally, we remark that mixing times of the FK-dynamics on a graph G are comparable, up to polynomial factors in $|E|$ to mixing times of the Chayes–Machta dynamics [8], as well as of the Swendsen–Wang algorithm that walks on the FK configurations of G [3, 36, 37]. By slight adaptations of the comparison results in [3, 36, 37], our theorems thus provide upper and lower bounds for these nonlocal dynamics in the presence of FK boundary conditions.

The rest of the paper is organized as follows. In Section 2, we formally define various preliminary notions that are used in our proofs. In Section 3, we introduce our 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. This is boosted to nearly-optimal mixing time for typical boundaries (Theorem 1.4) in Section 6. Finally, the torpid mixing result (Theorem 1.5) is proved in Section 7.

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

$$\Lambda_{n,l} = \{0, \dots, n\} \times \{0, \dots, l\} = \llbracket 0, n \rrbracket \times \llbracket 0, l \rrbracket.$$

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 $\mathbb{Z}^2 \setminus \Lambda_n$. The north, east, south, west, boundaries of Λ_n will be delineated $\partial_N\Lambda_n$, $\partial_E\Lambda_n$, $\partial_S\Lambda_n$ and $\partial_W\Lambda_n$, respectively.

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, for all $u, v \in \partial\Lambda_n$, u and v are in the same connected component of ω if and only if they are wired in ξ , then we say that the boundary condition ξ is *realizable*.

REMARK 1. Realizable boundary conditions are the most natural class of boundary conditions (see [21]) since they enforce the planarity of \mathbb{Z}^2 . However, nonrealizable boundary conditions are still relevant in some cases; for example, when considering the random-cluster model on nonlattice graphs such as trees. In Section 7, we consider the mixing time of the FK-dynamics under nonrealizable boundary conditions.

For $p \in (0, 1)$ and $q > 0$, the *random-cluster model* on Λ_n with a boundary condition ξ is the probability measure over the subsets $S \subseteq E(\Lambda_n)$ given by

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

where $c(S; \xi)$ corresponds to the number of connected components in the augmented graph (V, 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 subset $S \subseteq E(\Lambda_n)$, can be naturally identified with some edge

configuration $\omega : E(\Lambda_n) \rightarrow \{0, 1\}$ 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; for example, the random-cluster configuration on a set $R \subset \mathbb{Z}^2$ corresponds to the random-cluster configuration in the subgraph induced by R . We omit the subscripts p, q when understood from context.

Monotonicity. Define a partial order over boundary conditions by $\xi \leq \eta$ 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 $\xi \leq \eta$, then $\pi_{\Lambda_n}^\xi \leq \pi_{\Lambda_n}^\eta$, where \leq denotes stochastic domination.

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

$$p^* = \frac{q(1-p)}{q(1-p) + p}.$$

Under a realizable boundary condition ξ , this distributional equality becomes $\pi_{\Lambda_n, p, q}^\xi(S) = \pi_{\Lambda_n^*, p^*, q^*}^{\xi^*}(S)$, where ξ^* is the boundary condition induced by taking the dual configuration of the configuration on $\mathbb{Z}^2 \setminus \Lambda_n$ identified with ξ . Notice that \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*.

Infinite-volume measure and phase transition. A random-cluster measure $\pi_{\mathbb{Z}^2, p, q}$ can be defined on the infinite lattice \mathbb{Z}^2 as the limit as $n \rightarrow \infty$ of the sequence of random-cluster measures on $n \times n$ boxes with free boundary conditions. The measure $\pi_{\mathbb{Z}^2, p, q}$ exhibits a phase transition corresponding to the appearance of an infinite connected component. That is, there exists a critical value $p = p_c(q)$ such that if $p < p_c(q)$ (resp., $p > p_c(q)$), then, almost surely, all components are finite (resp., there exists an infinite component). For $q \geq 1$, the exact value of $p_c(q)$ for \mathbb{Z}^2 was recently settled in [2], proving

$$p_c(q) = p_{sd}(q) = \frac{\sqrt{q}}{\sqrt{q} + 1}.$$

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 $u, v \in \Lambda_n$,

$$(2) \quad \pi_{\Lambda_n, p, q}^\xi(u \overset{\Lambda_n}{\longleftrightarrow} v) \leq e^{-cd(u, v)},$$

where $d(u, v)$ is the graph distance between u, v in \mathbb{Z}^2 and $u \overset{\Lambda_n}{\longleftrightarrow} v$ 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).

2.1. *Random-cluster dynamics.* In this section, we overview some preliminaries related to Markov chain mixing times and the FK dynamics.

Mixing and coupling times. Consider an ergodic (i.e., irreducible and aperiodic) Markov chain \mathcal{M} with finite state space Ω , transition matrix P and stationary distribution μ . Let

$$t_{\text{MIX}}(\varepsilon) = \min\left\{t : \max_{X_0 \in \Omega} \|P^t(X_0, \cdot) - \mu\|_{\text{TV}} \leq \varepsilon\right\},$$

where $\|\cdot\|_{\text{TV}}$ is the total-variation distance. The *mixing time* of \mathcal{M} is given by $t_{\text{MIX}} := t_{\text{MIX}}(1/4)$, and for any positive $\varepsilon < 1/2$, by submultiplicativity, we have $t_{\text{MIX}}(\varepsilon) \leq \lceil \log_2 \varepsilon^{-1} \rceil t_{\text{MIX}}$. We use $t_{\text{MIX}}(\Lambda_n^\xi)$ to denote the mixing time of the FK-dynamics on $\Lambda_n \subset \mathbb{Z}^2$ with boundary condition ξ .

A *(one step) coupling* of the Markov chain \mathcal{M} specifies, for every pair of states $(X_t, Y_t) \in \Omega \times \Omega$, a probability distribution over (X_{t+1}, Y_{t+1}) such that the processes $\{X_t\}$ and $\{Y_t\}$, viewed in isolation, are faithful copies of \mathcal{M} , and if $X_t = Y_t$ then $X_{t+1} = Y_{t+1}$. The *coupling time*, denoted T_{coup} , is the minimum T such that $\Pr[X_T \neq Y_T] \leq 1/4$, starting from the worst possible pair of configurations X_0, Y_0 . The following inequality is standard: $t_{\text{MIX}} \leq T_{\text{coup}}$.

Spectral gap and conductance. If P is irreducible and reversible with respect to μ , then it has real eigenvalues $1 = \lambda_1 > \lambda_2 \geq \dots \geq \lambda_{|\Omega|} \geq -1$. The *absolute spectral gap* of P is defined by $\text{gap}(P) = 1 - \lambda_*$ where $\lambda_* = \max\{|\lambda_2|, |\lambda_{|\Omega|}|\}$. Let $\mu_{\min} = \min_{\omega \in \Omega} \mu(\omega)$; the following is then a standard inequality:

$$(3) \quad \text{gap}(P)^{-1} - 1 \leq t_{\text{MIX}} \leq \text{gap}(P)^{-1} \log(2e \cdot \mu_{\min}^{-1}).$$

For $A \subset \Omega$, the *conductance* of A is defined as

$$(4) \quad \Phi(A) = \frac{Q(A, A^c)}{\mu(A)} = \frac{\sum_{\omega \in A, \omega' \in A^c} \mu(\omega) P(\omega, \omega')}{\mu(A)}.$$

The *conductance* of the chain is given by $\Phi_\star = \min_{A: \mu(A) \leq 1/2} \Phi(A)$, and we have

$$(5) \quad \frac{\Phi_\star^2}{2} \leq \text{gap}(P) \leq 2\Phi_\star.$$

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. The edge-set of the dual graph Λ_n^* is not exactly in correspondence with the edge-set of a rectangle $\Lambda^* = \{-\frac{1}{2}, \dots, n + \frac{1}{2}\} \times \{-\frac{1}{2}, \dots, n + \frac{1}{2}\}$ 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 with at least one endpoint in $\Lambda_n \setminus \partial\Lambda_n$. The dual of this modified graph is then an $(n - 1) \times (n - 1)$ rectangle with all nearest-neighbor edges.

We note that the definition of realizability of the boundary condition is slightly different depending on the above choice of edge-set for a rectangle: while ξ is encoded in a configuration on $E(\mathbb{Z}^2) \setminus E(\Lambda_n)$, its dual ξ^* would be encoded in a configuration on $E((\mathbb{Z}^2)^*) \setminus (E(\Lambda^*) \setminus E(\partial\Lambda^*))$. However, it is straightforward to check that these two notions of realizability are equivalent: a partition of $\partial\Lambda_n$ can be encoded in a configuration on $E(\mathbb{Z}^2) \setminus E(\Lambda_n)$ if and only if it can be encoded in a configuration on $E(\mathbb{Z}^2) \setminus (E(\Lambda_n) \setminus E(\partial\Lambda_n))$. 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 $p < p_c(q)$.

Boundary condition modification. Finally, we will often appeal to a comparison inequality between mixing times of the FK-dynamics under different boundary conditions. This inequality is a consequence of a simple comparison between random-cluster measures.

DEFINITION 2.1. For two partitions ρ and ρ' of the vertex set of a graph $G = (V, E)$, we say that $\rho \leq \rho'$ if ρ is a finer partition than ρ' (so that the partition consisting only of singletons is the minimal element of this partial order). Then for two partitions $\rho \leq \rho'$ we define $D(\rho, \rho') = c(\rho) - c(\rho')$ where $c(\rho)$ is the number of components in ρ . For two partitions ρ, ρ' that are not comparable, let ρ'' be the smallest partition such that $\rho'' \geq \rho$ and $\rho'' \geq \rho'$ and set $D(\rho, \rho') = c(\rho) - c(\rho'') + c(\rho'') - c(\rho')$.

LEMMA 2.2. Let $G = (V, E)$ be an arbitrary graph, $p \in (0, 1)$ and $q > 0$. Let ρ and ρ' be two partitions of V encoding two distinct external wirings on the vertices of G . Let $\pi_G^\rho, \pi_G^{\rho'}$ be the resulting random-cluster measures. Then, for all FK configurations $\omega \in \{0, 1\}^E$, we have

$$q^{-2D(\rho, \rho')} \pi_G^{\rho'}(\omega) \leq \pi_G^\rho(\omega) \leq q^{2D(\rho, \rho')} \pi_G^{\rho'}(\omega).$$

With this in hand, the variational form of the spectral gap implies the following (see, e.g., [32]).

LEMMA 2.3. Let $G = (V, E)$ be an arbitrary graph, $p \in (0, 1)$ and $q > 0$. Consider the FK-dynamics on G with the external wirings ρ and ρ' , and let $t_{\text{MIX}}(G^\rho), t_{\text{MIX}}(G^{\rho'})$ denote their mixing times. Then

$$t_{\text{MIX}}(G^\rho) \leq q^{O(D(\rho, \rho'))} |E| t_{\text{MIX}}(G^{\rho'}).$$

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 subsets. In [4], 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, does not hold for most 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)*.

Notation. 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 ; that is, $E^c(R) := E(\Lambda_n) \setminus E(R)$. For a configuration $\omega : E(\Lambda_n) \rightarrow \{0, 1\}$, 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 ξ be a boundary condition for $\Lambda_n = (\Lambda_n, E(\Lambda_n))$ and let $\mathcal{B} = \{B_1, B_2, \dots, B_k\}$ be a collection of subsets of 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

$$(6) \quad \left| \pi_{\Lambda_n, p, q}^\xi(e = 1 \mid E^c(B_j) = 1) - \pi_{\Lambda_n, p, q}^\xi(e = 1 \mid E^c(B_j) = 0) \right| \leq \delta.$$

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

REMARK 3. SSM as defined in [4] 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. Namely, 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; see Theorem 1.5. However, if, for a collection $\mathcal{B} = \{B_1, B_2, \dots, B_k\}$, 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 $t_{\text{MIX}}(B^\tau)$ denote the mixing time of the FK-dynamics on the subset $B \subseteq \Lambda_n$ with boundary condition τ . (Recall that τ corresponds to a partition of ∂B and that ∂B consists of those vertices in B that are adjacent to vertices in $\mathbb{Z}^2 \setminus B$.)

DEFINITION 3.2. Let ξ be a boundary condition for $\Lambda_n = (\Lambda_n, E(\Lambda_n))$ and let $\mathcal{B} = \{B_1, B_2, \dots, B_k\}$ with $B_j \subset \Lambda_n$. We say that *local mixing (LM)* holds for \mathcal{B} and $T > 0$, if

$$t_{\text{MIX}}(B_j^{(1,\xi)}) \leq T \quad \text{and} \quad t_{\text{MIX}}(B_j^{(0,\xi)}) \leq T \quad \text{for all } j = 1, \dots, k,$$

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

REMARK 4. Observe that 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 $B_j \cap \partial\Lambda_n \neq \emptyset$, the connectivities from 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.

THEOREM 3.3. *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 $t_{\text{MIX}}(\Lambda_n^\xi) = O(Tn^2 \log n)$.*

PROOF. Consider two copies $\{X_t\}, \{Y_t\}$ of the FK-dynamics. We couple the evolution of $\{X_t\}$ and $\{Y_t\}$ by using the same random $e \in E(\Lambda_n)$ and the same uniform random number $r \in [0, 1]$ to decide whether to add or remove e from the configurations. This is a standard coupling of the steps of the FK-dynamics (see, e.g., [4, 21]); we call it the *identity coupling*. It is straightforward to verify that, when $q \geq 1$, the identity coupling is a *monotone coupling*, in the sense that if $X_t \subseteq Y_t$ then $X_{t+1} \subseteq Y_{t+1}$ with probability 1.

We bound the coupling time T_{coup} of the identity coupling for the FK-dynamics. The result then follows from the fact that $t_{\text{MIX}} \leq T_{\text{coup}}$. Since the identity coupling can be extended to a simultaneous coupling of *all* configurations that preserves the partial order \subseteq , the coupling time starting from any pair of configurations is bounded by the coupling time for initial configurations $Y_0 = \emptyset$ and $X_0 = E(\Lambda_n)$.

We prove that there exists $\hat{T} = O(Tn^2 \log n)$ such that the identity coupling \mathbb{P} satisfies

$$\mathbb{P}[X_{\hat{T}}(e) \neq Y_{\hat{T}}(e)] \leq \frac{1}{4|E(\Lambda_n)|}$$

for every $e \in E(\Lambda_n)$. A union bound over $e \in E(\Lambda_n)$ then implies that $T_{\text{coup}} \leq \hat{T}$.

Fix any edge $e \in E(\Lambda_n)$ and let $B_j \in \mathcal{B}$ be a subset for which (6) is satisfied (such a subset exists since moderate spatial mixing holds). To bound $\mathbb{P}[X_{\hat{T}}(e) \neq Y_{\hat{T}}(e)]$, we introduce two additional instances $\{Z_t^+\}, \{Z_t^-\}$ of the FK-dynamics. These two copies update only edges with both endpoints in B_j and reject all other updates. We set $Z_0^+ = E(\Lambda_n)$ and $Z_0^- = \emptyset$. The four Markov chains $\{X_t\}, \{Y_t\}, \{Z_t^+\}$ and $\{Z_t^-\}$ are coupled with the identity coupling, with updates outside B_j ignored by both $\{Z_t^+\}$ and $\{Z_t^-\}$. The monotonicity of this coupling, along with a triangle inequality, implies that for all $t \geq 0$,

$$(7) \quad \mathbb{P}[X_t(e) \neq Y_t(e)] \leq |\mathbb{P}[Z_t^+(e) = 1] - \pi_{\Lambda_n}^\xi(e = 1 | E^c(B_j) = 1)|$$

$$(8) \quad + |\pi_{\Lambda_n}^\xi(e = 1 | E^c(B_j) = 1) - \pi_{\Lambda_n}^\xi(e = 1 | E^c(B_j) = 0)|$$

$$(9) \quad + |\pi_{\Lambda_n}^\xi(e = 1 | E^c(B_j) = 0) - \mathbb{P}[Z_t^-(e) = 1]|.$$

Observe that the restrictions of the chains $\{Z_t^+\}$ and $\{Z_t^-\}$ to B_j are lazy versions of FK-dynamics on B_j with stationary measures $\pi_{\Lambda_n}^\xi(\cdot | E^c(B_j) = 1)$ and $\pi_{\Lambda_n}^\xi(\cdot | E^c(B_j) = 0)$, respectively. The laziness comes from the fact that they only accept updates that are in B_j , and by the local mixing assumption, once T updates have been done in B_j , the chains $\{Z_t^+\}$ and $\{Z_t^-\}$ will be mixed.

Now, after $\hat{T} = CT|E(\Lambda_n)| \log_2[24|E(\Lambda_n)|]$ steps, the expected number of updates in B_j is

$$CT|E(\Lambda_n)| \log_2[24|E(\Lambda_n)|] \frac{|E(B_j)|}{|E(\Lambda_n)|} \geq CT \log_2[24|E(\Lambda_n)|].$$

Let $\mathcal{A}_{\hat{T}}$ be the event that the number of updates in B_j after \hat{T} steps is at least $T \log_2[24 \times |E(\Lambda_n)|]$. A Chernoff bound then implies that, for a large enough constant $C > 0$,

$$\Pr[\mathcal{A}_{\hat{T}}^c] \leq \frac{1}{24|E(\Lambda_n)|}.$$

Therefore,

$$(10) \quad |\Pr[Z_{\hat{T}}^+(e) = 1 | \mathcal{A}_{\hat{T}}] - \Pr[Z_{\hat{T}}^+(e) = 1]| \leq \Pr[\mathcal{A}_{\hat{T}}^c] \leq \frac{1}{24|E(\Lambda_n)|}.$$

By the local mixing property, and the fact that $t_{\text{MIX}}(\varepsilon) \leq \lceil \log_2 \varepsilon^{-1} \rceil \cdot t_{\text{MIX}}$ for any positive $\varepsilon < 1/2$, we have

$$(11) \quad |\Pr[Z_{\hat{T}}^+(e) = 1 | \mathcal{A}_{\hat{T}}] - \pi_{\Lambda_n}^\xi(e = 1 | E^c(B_j) = 1)| \leq \frac{1}{24|E(\Lambda_n)|}.$$

It then follows from (10), (11) and the triangle inequality that when $t = \hat{T}$ the right-hand side in (7) is at most $\frac{1}{12|E(\Lambda_n)|}$. The same bound can be deduced for (9) in a similar manner.

Finally, since moderate spatial mixing holds for some $\delta < \frac{1}{12|E(\Lambda_n)|}$, we have

$$|\pi_{\Lambda_n}^\xi(e = 1 | E^c(B_j) = 1) - \pi_{\Lambda_n}^\xi(e = 1 | E^c(B_j) = 0)| \leq \frac{1}{12|E(\Lambda_n)|};$$

see Definition 3.1. Putting these together, we see that $\mathbb{P}[X_{\hat{T}}(e) \neq Y_{\hat{T}}(e)] \leq \frac{1}{4|E(\Lambda_n)|}$ as desired. \square

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 on thin rectangles of dimension $\Theta(n) \times \Theta(\log n)$ with realizable boundary conditions. To motivate this, we notice that since $p < p_c(q)$, the influence of the boundary is lost, with high probability, at a distance $\Theta(\log n)$. Thus the main difficulty will be to bound the mixing time of the FK-dynamics in the annulus of width $\Theta(\log n)$ with realizable boundary conditions on the outside. As such, the key ingredient in the proof of Theorem 1.1 will be the following mixing time bound on thin rectangles.

For an $n \times l$ rectangle $\Lambda_{n,l} = \llbracket 0, n \rrbracket \times \llbracket 0, l \rrbracket$, we use $\partial_N \Lambda_{n,l}$, $\partial_E \Lambda_{n,l}$, $\partial_S \Lambda_{n,l}$ and $\partial_W \Lambda_{n,l}$ for its north, east, south and west boundaries, respectively. Recall that for $a, b \in \mathbb{Z}$, we set $\llbracket a, b \rrbracket = \{a, a + 1, \dots, b\}$ and $\llbracket a, b \rrbracket^c = \llbracket 0, n \rrbracket \setminus \llbracket a, b \rrbracket$.

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 $\partial_E \Lambda_{n,l} \cup \partial_W \Lambda_{n,l} \cup \partial_S \Lambda_{n,l}$. 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 $\exp(O(l + \log n))$.*

Observe that 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 for us to prove Theorem 4.1 for the set of realizable boundary conditions ξ that are free on $\partial_E \Lambda_{n,l} \cup \partial_W \Lambda_{n,l} \cup \partial_S \Lambda_{n,l}$ and 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 $\partial_E \Lambda_{n,l} \cup \partial_W \Lambda_{n,l} \cup \partial_S \Lambda_{n,l}$; see Remark 2.

In Section 4.1, we give a overview of the main ideas in the proof of this theorem. In Sections 4.2–4.4, we introduce some crucial notions regarding groups of rectangular subsets of $\Lambda_{n,l}$ and their relations with the boundary conditions ξ . Sections 4.5–4.6 bound the mixing time of the block dynamics with respect to a suitably-chosen set of subsets of $\Lambda_{n,l}$. The recursive proof of Theorem 4.1 is then completed in Section 4.7.

4.1. Outline of proof of Theorem 4.1. We first mention some obstructions that boundary conditions present to proving Theorem 4.1 using approaches that are common in analogous problems for spin systems. A traditional approach to proving mixing time bounds for thin rectangles is the canonical paths method ([23, 26, 27, 33]), which gives an upper bound that is exponential in the shorter side length; however, boundary conditions can significantly distort the augmented graph with external wirings, preventing this approach from succeeding.

A sharper approach would be to use an inductive scheme [7, 17, 27], whereby, we bound the mixing time of the FK-dynamics on $n \times l$ rectangles by the mixing times in smaller rectangular blocks, for example,

$$(12) \quad B_W = \left[\left[0, \frac{2}{3}n \right] \right] \times \llbracket 0, l \rrbracket \quad \text{and} \quad B_E = \left[\left[\frac{1}{3}n, n \right] \right] \times \llbracket 0, l \rrbracket.$$

This method requires bounding by the mixing time of the so-called *block dynamics*.

DEFINITION 4.2. The *block dynamics* $\{X_t\}$ with blocks $B_W, B_E \subset \Lambda_{n,l}$ such that $E(B_W) \cup E(B_E) = E(\Lambda_{n,l})$ is the discrete-time Markov chain that, at each t , picks i uniformly at random from $\{W, E\}$ and updates the configuration in $E(B_i)$ with a sample from the stationary distribution of the chain conditional on the configuration of X_t on $E^c(B_i)$.

The spectral gap of the FK-dynamics on $\Lambda_{n,l}$ is bounded from below by the spectral gap of the block dynamics times the worst gap of the FK-dynamics in any B_i with worst-case configuration on $E^c(B_i)$; see Theorem 4.17 for a precise statement. With the choice of blocks



FIG. 1. (a) A boundary condition for which no configuration in $B_W \cap B_E$ isolates $B_W \setminus B_E$ from $B_E \setminus B_W$. (b) A boundary condition ξ where every pair of overlapping rectangles (as in (12)) must interact through ξ ; however, the two groups of rectangles $\mathcal{R}_1, \mathcal{R}_2$ do not interact through ξ .

in (12), applying this recursively, one would bound the spectral gap of the FK-dynamics on $\Lambda_{n,l}$ by the gap of the block dynamics raised to a $\Theta(\log n)$ power. Hence, establishing Theorem 4.1 would require an $\Omega(1)$ lower bound on the spectral gap of the block dynamics.

The mixing time and 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 $B_W \cap B_E$ will be such that it disconnects the influence of the configuration on $B_W \setminus B_E$ from $B_E \setminus B_W$ with probability $\Omega(1)$; this then allows a standard coupling argument to be used to bound the mixing time. In the presence of long-range boundary connections, however, it may be that no configuration on $B_W \cap B_E$ would disconnect the two sides from one another and facilitate coupling; see Figure 1(a) 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.

In particular, we will show that for every realizable boundary condition ξ on $\Lambda_{n,l}$, there exists a choice of two blocks, whose widths are at most $\frac{4}{5}n$, such that they are sufficiently isolated from one another in ξ . As Figure 1(b) demonstrates, there are realizable boundary conditions that would force these blocks to not be single rectangles, as in (12), but rather collections of rectangular subsets of $\Lambda_{n,l}$. Thus, our recursive argument will proceed instead on groups of rectangles,

$$\mathcal{R} = \bigcup R_i, \quad \text{where } R_i = [[a_i, b_i]] \times [[0, l]] \subset \Lambda_{n,l}$$

are disjoint, with boundary conditions induced by ξ and the configuration of the chain on $\Lambda_{n,l} \setminus \mathcal{R}$.

We formally define groups of rectangles (Definition 4.7), and their boundary conditions in Section 4.3. We consider next the notion of compatibility of a group of rectangles $\mathcal{R} \subset \Lambda_{n,l}$ with the boundary condition ξ . Roughly speaking, we say a group of rectangles \mathcal{R} is compatible with ξ if ξ limits the boundary interactions between the rectangles of \mathcal{R} for every possible configuration on $\Lambda_{n,l} \setminus \mathcal{R}$; see Definition 4.8. In Section 4.5, we provide an algorithm (see Lemma 4.9) that, for a group of rectangles \mathcal{R} compatible with ξ , finds two suitable subsets for the block dynamics: these subsets will each be group of rectangles compatible with ξ , of width between $1/5$ and $4/5$ of the width of \mathcal{R} . This will allow us to induct on groups of rectangles compatible with ξ , while ensuring that number of recursive steps is $O(\log n)$. Specifically, our splitting algorithm will find interior and exterior subsets \mathcal{A}_{INT} and \mathcal{A}_{EXT} with no boundary connections between the two; see Figure 3(a). These will be the cores of the two blocks for \mathcal{R} , but in order to bound the coupling time of the block dynamics, we want the two blocks to overlap and, therefore, we enlarge \mathcal{A}_{INT} and \mathcal{A}_{EXT} by $m = \Theta(\log l)$ to form the blocks \mathcal{R}_{INT} and \mathcal{R}_{EXT} for the block dynamics; see Figure 3(b).

Finally, in Section 4.6, we bound the coupling time of this block dynamics by some sufficiently large constant to conclude the proof. This follows by leveraging the fact that

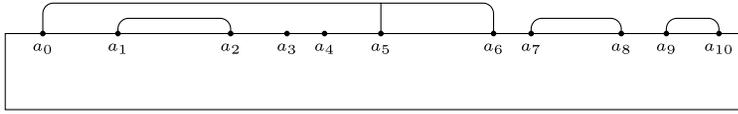


FIG. 2. The rectangle $\Lambda_{n,l}$ with a boundary condition ξ inducing disconnecting intervals. For example, $\llbracket a_1, a_4 \rrbracket$, $\llbracket a_3, a_4 \rrbracket$ and $\llbracket a_7, a_{10} \rrbracket$ are disconnecting intervals of free-type; $\llbracket a_1, a_2 \rrbracket$, $\llbracket a_0, a_6 \rrbracket$, $\llbracket a_7, a_8 \rrbracket$ and $\llbracket a_9, a_{10} \rrbracket$ are of free-wired-type; $\llbracket a_0, a_5 \rrbracket$ and $\llbracket a_5, a_6 \rrbracket$ are of wired-type.

$p < p_c$ ($p > p_c$) to condition on the existence of certain disconnecting dual (primal) paths in $\mathcal{R}_{\text{INT}} \cap \mathcal{R}_{\text{EXT}}$ that have $\Omega(1)$ probability.

We emphasize that in order to push through this recursion, it will be crucial that our notion of compatibility with ξ is strong enough to yield a uniform bound on this block dynamics coupling time, while being broad enough that the splitting algorithm always succeeds in finding subblocks that are themselves groups of rectangles compatible with ξ .

4.2. *Disconnecting intervals.* In this section, we introduce the notion of disconnecting intervals, one of the building blocks of our recursive proof of polynomial mixing. Recall that we use $\partial_N \Lambda_{n,l}$, $\partial_E \Lambda_{n,l}$, $\partial_S \Lambda_{n,l}$ and $\partial_W \Lambda_{n,l}$ for the north, east, south and west boundaries of the rectangle $\Lambda_{n,l}$, respectively.

DEFINITION 4.3. For a *realizable* boundary condition ξ on $\Lambda_{n,l}$ that is free on $\partial_E \Lambda_{n,l} \cup \partial_S \Lambda_{n,l} \cup \partial_W \Lambda_{n,l}$, an interval $\llbracket a, b \rrbracket \subset \llbracket 0, n \rrbracket$ is called *disconnecting* of:

1. *free-type*: if there are no boundary connections in ξ between $\llbracket a, b \rrbracket \times \{l\}$ and $\llbracket a, b \rrbracket^c \times \{l\}$;
2. *wired-type*: if there is a boundary component in ξ that contains the 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 $\llbracket a, b \rrbracket^c \times \llbracket 0, l \rrbracket$; in this case, we may refer to the interval as being of *free-wired-type*; see Figure 2 for several examples.

The following properties concerning the union and intersection of disconnecting intervals will be crucial to our proofs.

LEMMA 4.4. Let ξ be a *realizable* boundary condition on $\Lambda_{n,l}$ that is free on $\partial_S \Lambda_{n,l} \cup \partial_E \Lambda_{n,l} \cup \partial_W \Lambda_{n,l}$ and let $a < b < c$. If both $\llbracket a, b \rrbracket$ and $\llbracket b, c \rrbracket$ are disconnecting intervals of wired-type, then so is $\llbracket a, c \rrbracket$. If both $\llbracket a, b \rrbracket$ and $\llbracket b + 1, c \rrbracket$ are disconnecting intervals of free-type, then so is $\llbracket a, c \rrbracket$.

PROOF. If $\llbracket a, b \rrbracket$ and $\llbracket b, c \rrbracket$ are disconnecting intervals of wired-type, then by definition the vertices (a, l) , (b, l) and (c, l) are all in the same component of ξ ; hence $\llbracket a, c \rrbracket$ is a disconnecting interval of wired-type. If $\llbracket a, b \rrbracket$ and $\llbracket b + 1, c \rrbracket$ are disconnecting intervals of free-type, then by definition there are no connections in ξ between $\llbracket a, b \rrbracket$ and $\llbracket a, b \rrbracket^c \supset \llbracket a, c \rrbracket^c$ or between $\llbracket b + 1, c \rrbracket$ and $\llbracket b + 1, c \rrbracket^c \supset \llbracket a, c \rrbracket^c$. Consequently, there are not connections in ξ between $\llbracket a, c \rrbracket = \llbracket a, b \rrbracket \cup \llbracket b + 1, c \rrbracket$ and $\llbracket a, c \rrbracket^c$, and $\llbracket a, c \rrbracket$ is thus a disconnecting interval of free-type. \square

LEMMA 4.5. Let ξ be a *realizable* boundary condition on $\Lambda_{n,l}$ that is free on $\partial_S \Lambda_{n,l} \cup \partial_E \Lambda_{n,l} \cup \partial_W \Lambda_{n,l}$. Suppose there exist $a < b \leq c < d$ such that $\llbracket a, c \rrbracket$ and $\llbracket b, d \rrbracket$ are disconnecting intervals. Then either both $\llbracket a, c \rrbracket$ and $\llbracket b, d \rrbracket$ are of free-type or both are of wired-type.

PROOF. Suppose by way of contradiction and without loss of generality that $\llbracket a, c \rrbracket$ is only of free-type (i.e., free-type but not free-wired-type) and $\llbracket b, d \rrbracket$ is of wired-type. By definition, there exists a component of ξ that contains both (b, l) and (d, l) . Since $a < b \leq c < d$, there is therefore a connection in ξ between $\llbracket a, c \rrbracket \times \{l\}$ and $\llbracket a, c \rrbracket^c \times \{l\}$. Hence, $\llbracket a, c \rrbracket$ cannot be a disconnecting interval of free-type yielding the desired contradiction. \square

LEMMA 4.6. *Let ξ be a realizable boundary condition on $\Lambda_{n,l}$ that is free on $\partial_S \Lambda_{n,l} \cup \partial_E \Lambda_{n,l} \cup \partial_W \Lambda_{n,l}$. Suppose there exist $a < b < c < d$ such that $\llbracket a, c \rrbracket$ and $\llbracket b, d \rrbracket$ are disconnecting intervals.*

1. *If $\llbracket a, c \rrbracket$ and $\llbracket b, d \rrbracket$ are both of wired-type, then $\llbracket a, b \rrbracket$, $\llbracket b, c \rrbracket$, $\llbracket c, d \rrbracket$ and $\llbracket a, d \rrbracket$ are all disconnecting intervals of wired-type.*
2. *If $\llbracket a, c \rrbracket$ and $\llbracket b, d \rrbracket$ are both of free-type, then $\llbracket a, b - 1 \rrbracket$, $\llbracket b, c \rrbracket$, $\llbracket c + 1, d \rrbracket$ and $\llbracket a, d \rrbracket$ are all disconnecting intervals of free-type.*

PROOF. For the first part, suppose that both $\llbracket a, c \rrbracket$ and $\llbracket b, d \rrbracket$ are disconnecting intervals of wired-type. By definition, the vertices (a, l) and (c, l) are in the same component of ξ , as are (b, l) and (d, l) . By the planarity of realizable boundary conditions, it must be the case that these two components of ξ are indeed the same and, therefore, (a, l) , (b, l) , (c, l) , (d, l) are all in the same boundary component of ξ . As such, $\llbracket a, b \rrbracket$, $\llbracket b, c \rrbracket$, $\llbracket c, d \rrbracket$ and $\llbracket a, d \rrbracket$ are all disconnecting intervals of wired-type.

For the second part, suppose first by way of contradiction that $\llbracket a, b - 1 \rrbracket$ is not a disconnecting interval of free-type. Then there must be a boundary component in ξ with vertices in $\llbracket a, b - 1 \rrbracket \times \{l\}$ and $\llbracket a, b - 1 \rrbracket^c \times \{l\}$. If it is a component containing a vertex in $\llbracket a, c \rrbracket^c \times \{l\}$, it would violate the fact that $\llbracket a, c \rrbracket$ is disconnecting of free-type, while if it is a component containing a vertex in $\llbracket b, c \rrbracket \times \{l\}$, it would violate that $\llbracket b, d \rrbracket$ is disconnecting of free-type as $\llbracket a, b - 1 \rrbracket \subset \llbracket b, d \rrbracket^c$. By analogous reasoning $\llbracket c + 1, d \rrbracket$, $\llbracket b, c \rrbracket$ and $\llbracket a, d \rrbracket$ are all disconnecting intervals of free-type. \square

4.3. *Groups of rectangles.* In this section, we define *groups of rectangles* and their boundary conditions, which constitute the other building blocks of our recursive proof of polynomial mixing for the FK-dynamics on thin rectangles. As in the previous section, we consider an $n \times l$ rectangle $\Lambda_{n,l} = \llbracket 0, n \rrbracket \times \llbracket 0, l \rrbracket$ with a realizable boundary condition ξ that is free on $\partial_E \Lambda_{n,l} \cup \partial_S \Lambda_{n,l} \cup \partial_W \Lambda_{n,l}$.

Henceforth, we take

$$m = m(l) = C_\star \log l,$$

where C_\star is a large constant such that $C_\star > c^{-1}$ which we choose later, with c being the constant from (2).

A *rectangular subset* $R \subset \Lambda_{n,l}$ is a rectangle of the form $\llbracket a, b \rrbracket \times \llbracket 0, l \rrbracket$ for some $0 \leq a < b \leq n$. For such a rectangular subset, we denote by $W(R)$ its width; that is, $W(R) = b - a$. For the union of distinct and disjoint rectangular subsets $\mathcal{R} = \bigcup_{i=1}^{N(\mathcal{R})} R_i$, where $R_i = \llbracket a_i, b_i \rrbracket \times \llbracket 0, l \rrbracket$ and $a_1 < b_1 < a_2 < \dots < a_{N(\mathcal{R})} < b_{N(\mathcal{R})}$, its width is defined by $W(\mathcal{R}) = \sum_{i=1}^{N(\mathcal{R})} W(R_i)$ and its boundary is given by $\partial \mathcal{R} = \bigcup_{i=1}^{N(\mathcal{R})} \partial R_i$. Moreover, we let $\partial_N \mathcal{R} = \bigcup_{i=1}^{N(\mathcal{R})} \partial_N R_i$ and similarly define $\partial_S \mathcal{R}$; on the other hand, $\partial_E \mathcal{R}$ will be the eastern boundary of the right-most rectangle R_i and $\partial_W \mathcal{R}$, the western boundary of the left-most.

DEFINITION 4.7. *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 of $\Lambda_{n,l}$ such that $W(R_i) \geq 2m$ for every $i = 1, \dots, N(\mathcal{R})$.*

REMARK 5. The requirement that $W(R_i) \geq 2m$ for every i , which may seem arbitrary at the moment, is because in our recursive argument, we want our groups of rectangles \mathcal{R} to have interiors that are not influenced by the configuration on $E(\Lambda_{n,l}) \setminus E(\mathcal{R})$. When a group of rectangles has a thin constituent rectangle R_i , the influence of the outside configuration can permeate through all of R_i .

We will be considering blocks dynamics with blocks consisting of groups of rectangles. If we want to update the configuration on a group of rectangles $\mathcal{R} \subset \Lambda_{n,l}$, the boundary condition induced on $\partial\mathcal{R}$ will consist of the boundary condition on $\Lambda_{n,l}$ which will be fixed to be ξ , together with a fixed random-cluster configuration $\omega_{\mathcal{R}^c}$ on $E^c(\mathcal{R}) = E(\Lambda_{n,l}) \setminus E(\mathcal{R})$. Hence, our boundary conditions on \mathcal{R} will be of this form, and we denote them by the pair $\zeta = (\xi, \omega_{\mathcal{R}^c})$.

4.4. *Compatible boundary conditions.* We now define the notion of *compatibility* of groups of rectangles with boundary conditions ξ . This is crucial in our inductive argument; our algorithm for finding suitable blocks for the block dynamics will guarantee that if the starting group rectangles is compatible with respect to a given boundary condition, so will each of the blocks, enabling an inductive procedure.

Let ξ be a realizable boundary condition on $\partial\Lambda_{n,l}$ that is free on $\partial_S\Lambda_{n,l} \cup \partial_E\Lambda_{n,l} \cup \partial_W\Lambda_{n,l}$, and free in all vertices in $\partial_N\Lambda_{n,l}$ at distance at most m from $\partial_E\Lambda_{n,l} \cup \partial_W\Lambda_{n,l}$ (i.e., they appear as singletons in the corresponding boundary partition). This latter requirement for vertices near the corners of $\Lambda_{n,l}$ allows us to always choose blocks in our recursion whose boundaries are at least distance m from $\partial_E\Lambda_{n,l} \cup \partial_W\Lambda_{n,l}$; this simplifies our analysis.

The following will be the distinguishing property of our choice of blocks for the block dynamics.

DEFINITION 4.8. Let $\mathcal{R} = \bigcup_{i=1}^{N(\mathcal{R})} R_i$ be a group of rectangles with $R_i = \llbracket a_i, b_i \rrbracket \times \llbracket 0, l \rrbracket$ and $a_1 < b_1 < a_2 < b_2 < \dots < a_{N(\mathcal{R})} < b_{N(\mathcal{R})}$ (with $b_i - a_i \geq 2m$ for all i). We say \mathcal{R} is *compatible* with ξ , if

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

REMARK 6. It is clear from the definition that $\Lambda_{n,l}$ is compatible with ξ : the first condition is vacuous, while the second is satisfied since all vertices a distance at most m from $\partial_E\Lambda_{n,l} \cup \partial_W\Lambda_{n,l}$ are free. Observe also that $b_i - a_i \geq 2m$ for every i and so $b_{N(\mathcal{R})} - m \geq a_1 + m$; see Definition 4.7.

4.5. *Defining the blocks for the block dynamics.* Now that we have introduced disconnecting intervals, group of rectangles and the notion of compatibility, we describe our algorithm for picking two blocks $\mathcal{R}_{\text{INT}}, \mathcal{R}_{\text{EXT}}$ for the block dynamics, based on the boundary condition ξ . Recall the definitions of width $W(\cdot)$ of a rectangular subset and width of a collection of rectangles $W(\mathcal{R}) = \sum_{i=1}^{N(\mathcal{R})} W(R_i)$. It will also be convenient to have the following notation $\partial_{\parallel}\mathcal{R} = \bigcup_{i=1}^{N(\mathcal{R})} \partial_W R_i \cup \partial_E R_i$ for the vertical sides of the group of rectangles $\mathcal{R} = \bigcup_{i=1}^{N(\mathcal{R})} R_i$. The following lemma provides the basis for our splitting algorithm.

LEMMA 4.9. Let ξ be a realizable boundary condition on $\partial\Lambda_{n,l}$ that is free on $\partial_S\Lambda_{n,l} \cup \partial_E\Lambda_{n,l} \cup \partial_W\Lambda_{n,l}$ and free in all vertices in $\partial_N\Lambda_{n,l}$ at distance at most m from $\partial_E\Lambda_{n,l} \cup \partial_W\Lambda_{n,l}$. For every group of rectangles \mathcal{R} compatible with ξ , with $W(\mathcal{R}) \geq 100m$,

there exists a disconnecting interval $\llbracket c_\star, d_\star \rrbracket$ such that both (c_\star, l) and (d_\star, l) are in $\partial_N \mathcal{R}$, are distance at least m from $\partial_{\parallel} \mathcal{R}$, and

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

We pause to comment on why a disconnecting interval with such properties provides the desired blocks for the block dynamics. The interval $\llbracket c_\star, d_\star \rrbracket$ from the lemma will be used to define $\mathcal{A}_{\text{INT}} = \mathcal{R} \cap (\llbracket c_\star, d_\star \rrbracket \times \llbracket 0, l \rrbracket)$ and $\mathcal{A}_{\text{EXT}} = \mathcal{R} \setminus \mathcal{A}_{\text{INT}}$; their enlargements by m will form the blocks \mathcal{R}_{INT} and \mathcal{R}_{EXT} (see Figure 3). The requirement that $W(\mathcal{A}_{\text{INT}})$ be a fraction of $W(\mathcal{R})$ bounded away from 0 and 1 is so that we only recurse $O(\log n)$ times before reaching small enough widths. The requirement that the corners of $\llbracket c_\star, d_\star \rrbracket \times \llbracket 0, l \rrbracket$ are a distance at least m from $\partial_{\parallel} \mathcal{R}$ is so that when we enlarge the sets \mathcal{A}_{INT} , \mathcal{A}_{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 latter requirement will be guaranteed by the compatibility of \mathcal{R} with ξ .

PROOF OF LEMMA 4.9. We begin by finding a candidate disconnecting interval $\llbracket c, d \rrbracket$ with $(c, l), (d, l) \in \partial_N \mathcal{R}$ satisfying

$$(13) \quad \frac{1}{3} W(\mathcal{R}) \leq W(\mathcal{R} \cap (\llbracket c, d \rrbracket \times \llbracket 0, l \rrbracket)) \leq \frac{2}{3} W(\mathcal{R}).$$

In the second part of the proof, we show how to modify the interval $\llbracket c, d \rrbracket$ to obtain a disconnecting interval $\llbracket c_\star, d_\star \rrbracket$ with the added property that both (c_\star, l) and (d_\star, l) are distance at least m from $\partial_{\parallel} \mathcal{R}$.

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

$$(14) \quad \begin{aligned} W(\mathcal{R} \cap (\llbracket x, y \rrbracket \times \llbracket 0, l \rrbracket)) &< \frac{1}{3} W(\mathcal{R}), \quad \text{or} \\ W(\mathcal{R} \cap (\llbracket x, y \rrbracket \times \llbracket 0, l \rrbracket)) &> \frac{2}{3} W(\mathcal{R}). \end{aligned}$$

If the latter holds, then there is a pair, say $(x_0, l), (y_0, l) \in \partial_N \mathcal{R}$, for which the latter holds with a minimal width. The interval $\llbracket x_0, y_0 \rrbracket$ would be a disconnecting interval of wired-type and there is no other vertex $(z, l) \in \partial_N \mathcal{R}$ with $z \in \llbracket x_0 + 1, y_0 - 1 \rrbracket$ connected to (x_0, l) and (y_0, l) in ξ since if there were such a z it would violate the assumption that $\llbracket x_0, y_0 \rrbracket$ is of minimal width with $W(\mathcal{R} \cap (\llbracket x_0, y_0 \rrbracket \times \llbracket 0, l \rrbracket)) > \frac{2}{3} W(\mathcal{R})$ or that every pair of vertices $(x, l), (y, l) \in \partial_N \mathcal{R}$ connected in ξ satisfy (14). Consequently, all other connections through ξ between vertices $(x_1, l), (y_1, l) \in \partial_N \mathcal{R} \cap (\llbracket x_0 + 1, y_0 - 1 \rrbracket \times \llbracket 0, l \rrbracket)$ will be such that

$$W(\mathcal{R} \cap (\llbracket x_1, y_1 \rrbracket \times \llbracket 0, l \rrbracket)) < \frac{1}{3} W(\mathcal{R}).$$

We can then partition the vertices of $\partial_N \mathcal{R} \cap (\llbracket x_0 + 1, y_0 - 1 \rrbracket \times \{l\})$ into disjoint disconnecting intervals of free-wired-type using the following procedure:

1. Let $\rho = \{C_1, \dots, C_k\}$ be the partition of the vertices of $\partial_N \mathcal{R} \cap (\llbracket x_0 + 1, y_0 - 1 \rrbracket \times \{l\})$ induced by the boundary condition ξ ; every C_i corresponds to a distinct connected component of 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 $\partial_N \mathcal{R} \cap (\llbracket x_0 + 1, y_0 - 1 \rrbracket \times \{l\})$. Notice that some of the C_i 's may be singletons, which we view as disconnecting intervals of the free-wired-type;

3. Let $\{L_{i_1}, L_{i_2}, \dots, L_{i_\ell}\}$ be those disconnecting intervals which are maximal, in the sense that there does not exist j and k such that $L_{i_j} \subset L_{i_k}$.

The set of disconnecting intervals $\{L_{i_1}, L_{i_2}, \dots, L_{i_\ell}\}$ partitions $\llbracket x_0 + 1, y_0 - 1 \rrbracket$ into disjoint disconnecting intervals of free-wired-type with the property that $W(\mathcal{R} \cap (L_{i_j} \times \llbracket 0, l \rrbracket)) \leq \frac{1}{3}W(\mathcal{R})$ for every $j \in \{1, \dots, \ell\}$. We can then use Lemma 4.4 to merge adjacent disconnecting intervals until we obtain a candidate disconnecting interval $\llbracket c, d \rrbracket \subset \llbracket x_0, y_0 \rrbracket$ (of free-type), having width $W(\mathcal{R} \cap (\llbracket c, d \rrbracket \times \llbracket 0, l \rrbracket)) \in [\frac{1}{3}W(\mathcal{R}), \frac{2}{3}W(\mathcal{R})]$.

Now that we have found a candidate disconnecting interval $\llbracket c, d \rrbracket$ satisfying (13), we modify it to obtain a disconnecting interval $\llbracket c_\star, d_\star \rrbracket$ with the property that both (c_\star, l) and (d_\star, l) are distance at least m from $\partial_{\parallel} \mathcal{R}$.

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

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

note that the first case includes when \mathcal{I}_c is of free-wired-type, whereas the second case only applies when the interval is of free-type and not of wired-type. When (c, l) is instead at distance less than m from $\partial_E R_i$ for some i , then we simply set $c_\star = b_i - m$.

Symmetrically, if (d, l) is at distance less than m from $\partial_E R_i$ for some $R_i = \llbracket a_i, b_i \rrbracket \times \llbracket 0, l \rrbracket$, let $\mathcal{I}_d = \llbracket b_i - m, a_{i+1} + m \rrbracket$,

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

When (d, l) is at distance less than m from $\partial_W R_i$, let $d_\star = a_i + m$. To see that this process is well defined, notice that since $W(R_i) \geq 2m$ for every i , the points (c, l) , (d, l) cannot be both less than m away from $\partial_E R_i$ and less than m away from $\partial_W R_i$.

We claim that in all of these cases the interval $\llbracket c_\star, d_\star \rrbracket$ is a disconnecting interval. The fact that (c_\star, l) , $(d_\star, l) \in \partial_N \mathcal{R}$ are a distance at least m away from $\partial_{\parallel} \mathcal{R}$ follows directly from the construction.

First, observe that when (c, l) , (d, l) are both a distance at least m from $\partial_{\parallel} \mathcal{R}$, then we set $c_\star = c$ and $d_\star = d$; in this case $\llbracket c_\star, d_\star \rrbracket$ is disconnecting since $\llbracket c, d \rrbracket$ was chosen to be disconnecting.

Next, suppose that d was at least a distance m from $\partial_{\parallel} \mathcal{R}$ while c was at distance less than m from $\partial_W R_i$ for some i . In this setting, we establish that $\llbracket c_\star, d_\star \rrbracket$ is a disconnecting interval by considering the following four cases:

Case 1: $i = 1$. Note that $\llbracket a_1 + m, b_{N(\mathcal{R})} - m \rrbracket$ and $\llbracket c, d \rrbracket$ are disconnecting intervals by compatibility and construction, respectively. Also, $W(\mathcal{R} \cap (\llbracket c, d \rrbracket)) \geq \frac{100m}{3}$ since $W(\mathcal{R}) \geq 100m$. Hence, $c < a_1 + m < d < b_{N(\mathcal{R})} - m$ and Lemma 4.6 implies that $\llbracket c_\star, d_\star \rrbracket = \llbracket a_1 + m, d \rrbracket$ is disconnecting.

Case 2: $i > 1$ and $\llbracket b_{i-1} - m, a_i + m \rrbracket$ is of wired-type. Since $W(\mathcal{R} \cap (\llbracket c, d \rrbracket)) \geq \frac{100m}{3}$, we have $b_{i-1} - m < c < a_i + m < d$. Then, by Lemma 4.5, $\llbracket c, d \rrbracket$ is a disconnecting interval

of wired-type, and Lemma 4.6 implies that $\llbracket c_\star, d_\star \rrbracket = \llbracket a_i + m, d \rrbracket$ is a disconnecting interval of wired-type.

Case 3: $i > 1$, $\llbracket b_{i-1} - m, a_i + m \rrbracket$ is of free-type and $W(R_i) > 2m$. We again have $b_{i-1} - m < c < a_i + m < d$, and by Lemma 4.5 $\llbracket c, d \rrbracket$ is a disconnecting interval of free-type. Therefore, by Lemma 4.6, $\llbracket c_\star, d_\star \rrbracket = \llbracket a_i + m + 1, d \rrbracket$ is a disconnecting interval of free-type.

Case 4: $i > 1$, $\llbracket b_{i-1} - m, a_i + m \rrbracket$ is only of free-type and $W(R_i) = 2m$. In this case, it must be that $i < N(\mathcal{R})$. Therefore, by the compatibility of \mathcal{R} and ξ , $\llbracket b_i - m, a_{i+1} + m \rrbracket = \llbracket a_i + m, a_{i+1} + m \rrbracket$ is also a disconnecting interval. In fact, by Lemma 4.5, $\llbracket b_i - m, a_{i+1} + m \rrbracket$ is a disconnecting interval of free-type. Applying Lemma 4.6 with respect to $\llbracket a_i + m + 1, d \rrbracket$ and $\llbracket a_i + m, a_{i+1} + m \rrbracket$ implies $\llbracket c_\star, d_\star \rrbracket = \llbracket a_i + m, d \rrbracket$ is a disconnecting interval.

Suppose otherwise that c was at distance less than m from $\partial_E R_i$ for some i , and d was still at least a distance at least m from $\partial_{\parallel} \mathcal{R}$. In this case, $W(\mathcal{R} \cap (\llbracket c, d \rrbracket)) \geq \frac{100m}{3}$ implies $N(\mathcal{R}) > 1$ as well as $i < N(\mathcal{R})$. Moreover, $\llbracket b_i - m, a_{i+1} + m \rrbracket$ is a disconnecting interval by the compatibility of \mathcal{R} and ξ . Since $b_i - m < c < a_{i+1} + m < d$, Lemma 4.5 then implies that when $\llbracket b_i - m, a_{i+1} + m \rrbracket$ is of wired-type (resp., of free-type) then $\llbracket c, d \rrbracket$ is also of wired-type (resp., of free-type) and, therefore, by Lemma 4.6, $\llbracket c_\star, d_\star \rrbracket = \llbracket b_i - m, d \rrbracket$ is a disconnecting interval of wired-type (resp., of free-type).

The symmetric cases where (c, l) is a distance at least m from $\partial_{\parallel} \mathcal{R}$, and (d, l) is a distance less than m from $\partial_{\parallel} \mathcal{R}$ can be checked analogously. The remaining case in which both (c, l) and (d, l) are a distance less than m from $\partial_{\parallel} \mathcal{R}$ can also be checked similarly, by first modifying the candidate interval on one side in order to obtain a disconnecting interval $\llbracket c_\star, d \rrbracket$ having (c_\star, l) a distance at least m from $\partial_{\parallel} \mathcal{R}$, and then performing the modification on d to obtain the desired disconnecting interval $\llbracket c_\star, d_\star \rrbracket$.

Finally, we claim that in all such situations, $\llbracket c_\star, d_\star \rrbracket$ satisfies

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

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

We will now use the disconnecting interval given by Lemma 4.9 to define two subsets \mathcal{A}_{INT} and \mathcal{A}_{EXT} of \mathcal{R} , and set \mathcal{R}_{INT} and \mathcal{R}_{EXT} to be their enlargements by m .

DEFINITION 4.10. For a group of rectangles \mathcal{R} compatible with ξ , let $\llbracket c_\star, d_\star \rrbracket$ be the disconnecting interval given by Lemma 4.9. Then define the *interior and exterior cores* as

$$\mathcal{A}_{\text{INT}} = \mathcal{R} \cap (\llbracket c_\star, d_\star \rrbracket \times \llbracket 0, l \rrbracket) \quad \text{and} \quad \mathcal{A}_{\text{EXT}} = \mathcal{R} \cap (\llbracket c_\star, d_\star \rrbracket^c \times \llbracket 0, l \rrbracket);$$

see Figure 3(a). Using the disconnecting interval $\llbracket c_\star, d_\star \rrbracket$ define the *interior and exterior blocks* as follows:

1. Let $c_- = c_\star - m$ and $c_+ = c_\star + m$. Let $d_- = d_\star - m$ and $d_+ = d_\star + m$.
2. Define $\mathcal{R}_{\text{INT}} = \mathcal{R} \cap (\llbracket c_-, d_+ \rrbracket \times \llbracket 0, l \rrbracket)$ and $\mathcal{R}_{\text{EXT}} = \mathcal{R} \cap ((\llbracket 0, c_+ \rrbracket \cup \llbracket d_-, n \rrbracket) \times \llbracket 0, l \rrbracket)$; see Figure 3(b).

We will demonstrate that this choice of \mathcal{R}_{INT} and \mathcal{R}_{EXT} has certain fundamental properties that will facilitate the recursive argument in the proof of Theorem 4.1.

PROPOSITION 4.11. *If \mathcal{R} is a group of rectangles compatible with ξ , and moreover, $W(\mathcal{R}) \geq 100m$, then the sets \mathcal{R}_{INT} and \mathcal{R}_{EXT} are groups of rectangles satisfying the following properties:*

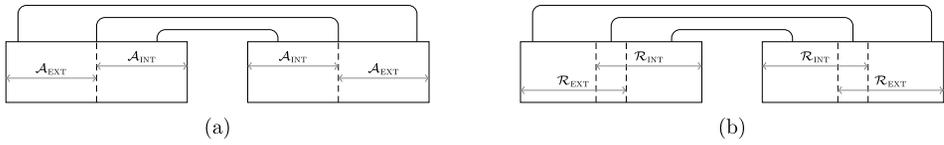


FIG. 3. (a) The cores \mathcal{A}_{INT} and \mathcal{A}_{EXT} and (b) the blocks \mathcal{R}_{INT} and \mathcal{R}_{EXT} . The blocks \mathcal{R}_{INT} and \mathcal{R}_{EXT} are the enlargements of \mathcal{A}_{INT} and \mathcal{A}_{EXT} by exactly m , and are thus, themselves, groups of rectangles.

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

PROOF. We show first that \mathcal{R}_{INT} and \mathcal{R}_{EXT} are groups of rectangles (see Definition 4.7). By construction \mathcal{R}_{INT} and \mathcal{R}_{EXT} are unions of rectangular subsets of $\Lambda_{n,l}$. Since \mathcal{R} is a group of rectangles, every constituent rectangular subset R_i of \mathcal{R} has $W(R_i) \geq 2m$ by definition. By Lemma 4.9, (c_\star, l) , (d_\star, l) were such that they are a distance at least m from $\partial_{\parallel}\mathcal{R}$; as a consequence $[[c_-, c_+] \times [[0, l]]$ and $[[d_-, d_+] \times [[0, l]]$ are subsets of \mathcal{R} . Moreover, every constituent rectangular subset of \mathcal{R}_{INT} and \mathcal{R}_{EXT} is either a constituent rectangle of \mathcal{R} , or contains one of $[[c_-, c_+] \times [[0, l]]$ and $[[d_-, d_+] \times [[0, l]]$, implying that every rectangular subset of \mathcal{R}_{INT} and \mathcal{R}_{EXT} has width at least $2m$. Together, these verify that \mathcal{R}_{INT} and \mathcal{R}_{EXT} are indeed groups of rectangles.

The first property follows from the facts that, by Lemma 4.9, $\frac{1}{4}W(\mathcal{R}) \leq W(\mathcal{A}_{\text{INT}}) \leq \frac{3}{4}W(\mathcal{R})$ and $\frac{1}{4}W(\mathcal{R}) \leq W(\mathcal{A}_{\text{EXT}}) \leq \frac{3}{4}W(\mathcal{R})$, while $W(\mathcal{R}) \geq 100m$, $W(\mathcal{R}_{\text{INT}}) - W(\mathcal{A}_{\text{INT}}) = 2m$ and $W(\mathcal{R}_{\text{EXT}}) - W(\mathcal{A}_{\text{EXT}}) = 2m$.

To verify the second property, consider \mathcal{R}_{EXT} first and label its rectangular subsets $R_1^{\text{EXT}}, \dots, R_{N(\mathcal{R}_{\text{EXT}})}^{\text{EXT}}$ (from left to right) with $R_j^{\text{EXT}} = [[a_j^{\text{EXT}}, b_j^{\text{EXT}}]]$. Let $i, i + 1$ be the indices of the two distinct rectangular subsets containing $\mathcal{R}_{\text{EXT}} \setminus \mathcal{A}_{\text{EXT}}$. As before, $R_1, \dots, R_{N(\mathcal{R})}$ are the constituent rectangular subsets of \mathcal{R} . Then, for every $j \in \{1, \dots, N(\mathcal{R}_{\text{EXT}}) - 1\} \setminus \{i\}$, there is a $k \in \{1, \dots, N(\mathcal{R}) - 1\}$ such that $b_j^{\text{EXT}} = b_k$ and $a_{j+1}^{\text{EXT}} = a_{k+1}$. Hence, the compatibility of \mathcal{R} with ξ guarantees that the interval $[[b_j^{\text{EXT}} - m, a_{j+1}^{\text{EXT}} + m]]$ is disconnecting for every $j \neq i$.

To see that $[[b_i^{\text{EXT}} - m, a_{i+1}^{\text{EXT}} + m]]$ is disconnecting, notice that by construction of \mathcal{R}_{EXT} , $[[b_i^{\text{EXT}} - m, a_{i+1}^{\text{EXT}} + m]] = [[c_\star, d_\star]]$. Finally, since (c_\star, l) , (d_\star, l) were at distance at least m from $\partial_{\parallel}\mathcal{R}$, the interval $[[a_1^{\text{EXT}} + m, b_{N(\mathcal{R}_{\text{EXT}})}^{\text{EXT}} - m]]$ matches the interval $[[a_1 + m, b_{N(\mathcal{R})} - m]]$, and thus by compatibility of \mathcal{R} with ξ implies the former is a disconnecting interval, altogether these imply the compatibility of \mathcal{R}_{EXT} with ξ .

Similarly, label the constituent rectangles of \mathcal{R}_{INT} as $R_1^{\text{INT}}, \dots, R_{N(\mathcal{R}_{\text{INT}})}^{\text{INT}}$ and notice that 1 and $N(\mathcal{R}_{\text{INT}})$ are the indices of the two rectangles containing $\mathcal{R}_{\text{INT}} \setminus \mathcal{A}_{\text{INT}}$. Every interval of the form $[[b_i^{\text{INT}} - m, a_{i+1}^{\text{INT}} + m]]$ corresponds (up to change of index) to such an interval for \mathcal{R} , so that by compatibility of \mathcal{R} with respect to ξ , these are all disconnecting intervals. The interval $[[a_1^{\text{INT}} + m, b_{N(\mathcal{R}_{\text{INT}})}^{\text{INT}} - m]]$ is exactly the interval $[[c_\star, d_\star]]$ given by Lemma 4.9 and, therefore, this is disconnecting by construction. Together, these imply the compatibility of \mathcal{R}_{INT} with ξ . \square

4.6. *Block dynamics coupling time.* Here, we consider the block dynamics on \mathcal{R} with blocks \mathcal{R}_{INT} and \mathcal{R}_{EXT} as defined in Section 4.5; see also Figure 3(b). We begin by defining the block dynamics for a group of rectangles.

DEFINITION 4.12. Let \mathcal{R} be a group of rectangles with boundary condition ζ . The block dynamics $\{X_t\}$ with blocks $\mathcal{B} = \{B_1, \dots, B_k\}$ such that $B_i \subset \mathcal{R}$ and $\bigcup_{i=1}^k E(B_i) = E(\mathcal{R})$ is the discrete-time Markov chain that at each t , selects i uniformly at random from $\{1, \dots, k\}$

and updates the configuration on $E(\mathcal{B}_i)$ from the stationary distribution conditional on the configuration $X_t(E^c(\mathcal{B}_i))$.

Fix $\mathcal{B} = \{\mathcal{R}_{\text{INT}}, \mathcal{R}_{\text{EXT}}\}$ and let $\text{gap}(\mathcal{R}^\xi; \mathcal{B})$ be the spectral gap of this block dynamics on \mathcal{R} with boundary condition $\zeta = (\xi, \omega_{\mathcal{R}^c})$, where ξ is a realizable boundary condition on $\Lambda_{n,l}$ and $\omega_{\mathcal{R}^c}$ is a configuration on $E^c(\mathcal{R})$.

LEMMA 4.13. *Let ξ be a realizable boundary condition on $\partial\Lambda_{n,l}$ that is free on $\partial_S\Lambda_{n,l} \cup \partial_E\Lambda_{n,l} \cup \partial_W\Lambda_{n,l}$ and free on vertices in $\partial_N\Lambda_{n,l}$ at distance at most m from $\partial_E\Lambda_{n,l} \cup \partial_W\Lambda_{n,l}$. 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 $\omega_{\mathcal{R}^c}$,*

$$\text{gap}(\mathcal{R}^{(\xi, \omega_{\mathcal{R}^c})}; \mathcal{B}) \geq K^{-1}.$$

PROOF. We consider the $p < p_c(q)$ case first. 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 . We design a coupling \mathbb{P} for the steps of $\{X_t\}$ and $\{Y_t\}$ and bound its coupling time. This yields upper bounds for both the mixing time and the inverse spectral gap of the block dynamics; see Section 2 for a brief overview of the coupling method. For this, we will show that for any two initial configurations X_0, Y_0

$$(15) \quad \mathbb{P}(X_2 = Y_2) = \Omega(1).$$

Since this bound will be uniform over X_0, Y_0 , we can make independent attempts at coupling the two chains every two steps. Hence, there would exist $T = O(1)$ such that

$$\max_{X_0, Y_0} \mathbb{P}(X_T \neq Y_T) \leq \frac{1}{4},$$

bounding the coupling time by $T = O(1)$ concluding the proof.

First, observe that with probability $1/4$ the first block to be updated is \mathcal{R}_{INT} and the second is \mathcal{R}_{EXT} . Suppose this is the case and let us consider the update on block \mathcal{R}_{INT} . Observe that $X_1(\mathcal{R}_{\text{INT}})$ is distributed according to the random-cluster measure π^{θ_X} on \mathcal{R}_{INT} , where θ_X is the boundary condition induced on $\partial\mathcal{R}_{\text{INT}}$ by the boundary condition ζ on \mathcal{R} and the configuration of X_0 in $E(\mathcal{R}) \setminus E(\mathcal{R}_{\text{INT}})$. Likewise, $Y_1(\mathcal{R}_{\text{INT}})$ has law π^{θ_Y} , with the boundary condition θ_Y defined analogously but considering the configuration of Y_0 in $E(\mathcal{R}) \setminus E(\mathcal{R}_{\text{INT}})$ instead. Therefore, any coupling for the random-cluster measures $\pi^{\theta_X}, \pi^{\theta_Y}$ yields a coupling for the first steps of $\{X_t\}$ and $\{Y_t\}$.

Let θ_1 be the boundary condition on $\partial\mathcal{R}_{\text{INT}}$ induced by ζ and the configuration that is all wired on $E(\mathcal{R}) \setminus E(\mathcal{R}_{\text{INT}})$; let π^{θ_1} be the corresponding random-cluster measure on \mathcal{R}_{INT} . Let $Q_W, Q_E \subset \mathcal{R}_{\text{INT}}$ be the two rectangles of width m that contain all the vertices in $\mathcal{R}_{\text{INT}} \setminus \mathcal{A}_{\text{INT}}$; that is, $Q_W \cup \mathcal{A}_{\text{INT}} \cup Q_E = \mathcal{R}_{\text{INT}}$, $Q_W \cap \mathcal{A}_{\text{INT}} = \emptyset$ and $Q_E \cap \mathcal{A}_{\text{INT}} = \emptyset$ (see Figure 4(a)). Let $\partial E(Q_W)$ be the set edges with one endpoint in Q_W and the other in \mathcal{A}_{INT} , and similarly define $\partial E(Q_E)$.

Let Γ_W be the set of configurations in \mathcal{R}_{INT} that have a dual-path in $E(Q_W) \cup \partial E(Q_W)$ connecting the top-most edge in $\partial E(Q_W)$ to an edge in $\partial_S Q_W$, and similarly define Γ_E as the set of configurations in \mathcal{R}_{INT} that have a dual-path in $E(Q_E) \cup \partial E(Q_E)$ from the top-most edge in $\partial E(Q_E)$ to an edge in $\partial_S Q_E$. (A dual-path is an open path in the dual configuration.) Let $\Gamma = \Gamma_E \cap \Gamma_W$; see Figure 4(b). The following lemma supplies the desired coupling.

LEMMA 4.14. *Let $q > 1$ and $p < p_c(q)$. There exists a coupling \mathbb{P}_1 of the distributions $\pi^{\theta_X}, \pi^{\theta_Y}, \pi^{\theta_1}$ such that if $(\omega^{\theta_X}, \omega^{\theta_Y}, \omega^{\theta_1})$ is sampled from \mathbb{P}_1 , then all of the following hold:*

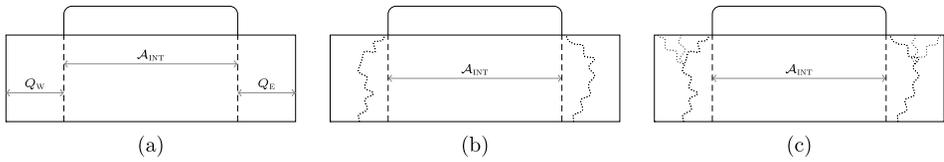


FIG. 4. (a) The block \mathcal{R}_{INT} with its subsets \mathcal{A}_{INT} , Q_W and Q_E . (b) The block \mathcal{R}_{INT} with the dual-paths (dotted) of a configuration in Γ . (c) The block \mathcal{R}_{INT} with the dual-paths (dotted) of a configuration in $\Gamma_W^{\text{NS}} \cap \Gamma_W^{\text{WE}} \cap \Gamma_E^{\text{NS}} \cap \Gamma_E^{\text{WE}} \subset \Gamma$.

1. $\mathbb{P}_1(\omega^{\theta_X}, \omega^{\theta_Y}, \omega^{\theta_1}) > 0$ only if $\omega^{\theta_X} \leq \omega^{\theta_1}$ and $\omega^{\theta_Y} \leq \omega^{\theta_1}$;
2. $\mathbb{P}_1(\omega^{\theta_X}(\mathcal{A}_{\text{INT}}) = \omega^{\theta_Y}(\mathcal{A}_{\text{INT}}) \mid \omega^{\theta_1} \in \Gamma) = 1$;
3. There exists a constant $\rho = \rho(p, q) > 0$ such that $\mathbb{P}_1(\omega^{\theta_1} \in \Gamma) \geq \rho$.

Hence, if we use the coupling \mathbb{P}_1 from Lemma 4.14 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 easily couple the update on \mathcal{R}_{EXT} in the second step so that $X_2 = Y_2$, since $X_1(E(\mathcal{A}_{\text{INT}})) = Y_1(E(\mathcal{A}_{\text{INT}}))$ 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}_{EXT} are identical. As a consequence, we obtain that for any X_0, Y_0 ,

$$\mathbb{P}(X_2 = Y_2) \geq \frac{\rho}{4}.$$

which gives (15) and thus concludes the proof for $p < p_c(q)$. \square

The case when $p > p_c(q)$ follows by an analogous dual argument. In this case, the set Γ has $o(1)$ probability and we therefore replace it by the set $\Gamma^* = \Gamma_E^* \cap \Gamma_W^*$, where Γ_W^* is the set of configurations in \mathcal{R}_{INT} that have an open path in $E(Q_W) \cup \partial E(Q_W)$ connecting the top-left corner (c_\star, l) to $\partial_S Q_W$; similarly, Γ_E^* is the set of configurations in \mathcal{R}_{INT} that have an open path in $E(Q_E) \cup \partial E(Q_E)$ connecting the top-right corner (d_\star, l) of \mathcal{A}_{INT} to $\partial_S Q_E$. Let θ_0 be the boundary condition on $\partial \mathcal{R}_{\text{INT}}$ induced by ζ and the all-free configuration on $E(\mathcal{R}) \setminus E(\mathcal{R}_{\text{INT}})$; let π^{θ_0} be the resulting random-cluster distribution on \mathcal{R}_{INT} . The constant bound on the coupling time of the block dynamics would then follow as above from the following dual analogue of Lemma 4.14.

LEMMA 4.15. *Let $q > 1$ and $p > p_c(q)$. There exists a coupling \mathbb{P}_0 of the distributions $\pi^{\theta_X}, \pi^{\theta_Y}, \pi^{\theta_0}$ such that if $(\omega^{\theta_X}, \omega^{\theta_Y}, \omega^{\theta_0})$ is sampled from \mathbb{P}_0 , then all of the following hold:*

1. $\mathbb{P}_0(\omega^{\theta_X}, \omega^{\theta_Y}, \omega^{\theta_0}) > 0$ only if $\omega^{\theta_X} \geq \omega^{\theta_0}$ and $\omega^{\theta_Y} \geq \omega^{\theta_0}$;
2. $\mathbb{P}_0(\omega^{\theta_X}(\mathcal{A}_{\text{INT}}) = \omega^{\theta_Y}(\mathcal{A}_{\text{INT}}) \mid \omega^{\theta_0} \in \Gamma^*) = 1$;
3. There exists a constant $\rho = \rho(p, q) > 0$ such that $\mathbb{P}_0(\omega^{\theta_0} \in \Gamma^*) \geq \rho$.

We proceed with the proof of the key Lemma 4.14. Parts 1 and 2 of the lemma follow from a fairly standard coupling technique; see, for example, [1, 4]. We shall also use this approach later in the proof of Lemma 5.4. Part 3 will be a consequence of the EDC property (2); see proof of Claim 4.16.

PROOF OF LEMMA 4.14. Let

$$L = \partial_W Q_W \cup \partial_N Q_W \cup \partial_E Q_E \cup \partial_N Q_E;$$

note that $L \cap \mathcal{A}_{\text{INT}} = \emptyset$. For a FK configuration ω on \mathcal{R}_{INT} , let

$$F(\omega) := \mathcal{R}_{\text{INT}} \setminus \bigcup_{v \in L} C(v, \omega),$$

where $C(v, \omega)$ is the vertex set of the connected component of v in ω , ignoring the boundary connections.

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}). We use \mathbb{P}_X and \mathbb{P}_Y to construct the coupling \mathbb{P}_1 as follows:

1. sample $(\omega^{\theta_X}, \omega^{\theta_1})$ from \mathbb{P}_X and sample ω^{θ_Y} from $\mathbb{P}_Y(\cdot \mid \omega^{\theta_1})$;
2. if $\mathcal{A}_{\text{INT}} \subseteq F(\omega^{\theta_1})$, sample ω_Δ from $\pi_\Delta^{\eta_1}$ —where $\Delta = \Delta(\omega^{\theta_1})$ is the subgraph induced by $F(\omega^{\theta_1})$ and η_1 is the boundary condition on $\partial F(\omega^{\theta_1})$ induced by θ_1 and the configuration of ω^{θ_1} on $E(\mathcal{R}_{\text{INT}}) \setminus E(F(\omega^{\theta_1}))$ —and make $\omega^{\theta_1}(F(\omega^{\theta_1})) = \omega^{\theta_X}(F(\omega^{\theta_1})) = \omega^{\theta_Y}(F(\omega^{\theta_1})) = \omega_\Delta$.

Let \mathbb{P}_1 be the resulting distribution of $(\omega^{\theta_X}, \omega^{\theta_Y}, \omega^{\theta_1})$. After step (i), \mathbb{P}_1 has the desired marginals. Moreover, we claim that replacing the configuration in $F(\omega^{\theta_1})$ with ω_Δ in step (ii) has no effect on the distribution, provided $\mathcal{A}_{\text{INT}} \subseteq F(\omega^{\theta_1})$. For this, we show that the three boundary conditions η_1, η_X, η_Y induced on $\partial F(\omega^{\theta_1})$ by the configurations of $\omega^{\theta_X}, \omega^{\theta_Y}, \omega^{\theta_1}$ on $E(\mathcal{R}_{\text{INT}}) \setminus E(F(\omega^{\theta_1}))$, respectively, and the corresponding boundary conditions $\theta_X, \theta_Y, \theta_1$ are all the same.

First, observe that the boundary condition on $\partial_S \mathcal{A}_{\text{INT}}$ is, in all three cases, free by assumption. Also, from the definition of $F(\omega^{\theta_1})$ it follows that every edge of $E(\mathcal{R}_{\text{INT}}) \setminus E(F(\omega^{\theta_1}))$ incident to $\partial F(\omega^{\theta_1})$ is closed in ω^{θ_1} ; hence 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 $(\partial \mathcal{A}_{\text{INT}} \cap \partial \mathcal{R}) \setminus \partial_S \mathcal{R}$. To show that η_1, η_X, η_Y also agree on $(\partial \mathcal{A}_{\text{INT}} \cap \partial \mathcal{R}) \setminus \partial_S \mathcal{R}$, we use the fact that top-left and top-right corners of \mathcal{A}_{INT} correspond to the endpoints of the disconnecting interval $[[c_\star, d_\star]]$. Indeed, for the boundary conditions η_1, η_X, η_Y to disagree on $(\partial \mathcal{A}_{\text{INT}} \cap \partial \mathcal{R}) \setminus \partial_S \mathcal{R}$ it must be the case that there are at least two distinct connected components of $\zeta = (\xi, \omega_{\mathcal{R}^c})$ connecting $(\partial \mathcal{A}_{\text{INT}} \cap \partial \mathcal{R}) \setminus \partial_S \mathcal{R}$ and $\partial \mathcal{R} \setminus \partial \mathcal{A}_{\text{INT}}$. Since $(c_\star, l), (d_\star, l) \notin \partial_{\parallel} \mathcal{R}$ and ξ is free on $\partial_S \mathcal{R}$, this would require at least two distinct connected components of ξ connecting vertices in $[[c_\star, d_\star]] \times \{l\}$ to vertices in $[[c_\star, d_\star]]^c \times \{l\}$. However, when $[[c_\star, d_\star]]$ is a disconnecting interval of free-type, there are no such connected components, and when it is disconnecting of wired-type, the planarity of realizable boundary conditions implies that there can be at most one such connected component, which is exactly the component of ξ containing both (c_\star, l) and (d_\star, l) .

Altogether, these together imply that when $\mathcal{A}_{\text{INT}} \subseteq F(\omega^{\theta_1})$, the three boundary conditions η_1, η_X, η_Y induced on $\partial F(\omega^{\theta_1})$ are the same. The *domain Markov property* of random-cluster measures (see, e.g., [21]) then guarantees that replacing the configuration in $F(\omega^{\theta_1})$ with ω_Δ had no effect on the distributions.

Finally, note that if $\omega^{\theta_1} \in \Gamma$, then the vertices in the boundary components of L , that is, $\bigcup_{v \in L} C(v, \omega^{\theta_1})$, will be confined to $Q_W \cup Q_E$, in which case $\mathcal{A}_{\text{INT}} \subseteq F(\omega^{\theta_1})$. This establishes parts 1 and 2, and part 3 follows directly from the following claim, which will conclude the proof. \square

CLAIM 4.16. *Let $q > 1$ and $p < p_c(q)$. There exists $\rho = \rho(p, q) > 0$ such that $\pi^{\theta_1}(\Gamma) \geq \rho$.*

PROOF. Let Γ_W^{NS} (resp., Γ_E^{NS}) be the set of configurations on \mathcal{R}_{INT} such that there is dual-path from $\partial_N Q_W$ to $\partial_S Q_W$ (resp., from $\partial_N Q_E$ to $\partial_S Q_E$) in $E(Q_W) \cup \partial E(Q_W)$ (resp., $E(Q_E) \cup \partial E(Q_E)$). Also, let Γ_W^{WE} (resp., Γ_E^{WE}) be the set of configurations on \mathcal{R}_{INT} such that there is dual-path in $E(Q_W) \cup \partial E(Q_W)$ (resp., in $E(Q_E) \cup \partial E(Q_E)$) between the left-most edge in $\partial_N Q_W$ (resp., the right-most edge in $\partial_N Q_E$) and the top-most edge of $\partial E(Q_W)$ (resp., $\partial E(Q_E)$); see Figure 4(c). Note that $\Gamma_W^{\text{NS}} \cap \Gamma_W^{\text{WE}} \cap \Gamma_E^{\text{NS}} \cap \Gamma_E^{\text{WE}} \subset \Gamma$.

The width of Q_W is $m = C_\star \log l$. Hence, the EDC property (2) implies that when C_\star is large enough there exists a constant $\rho_0 = \rho_0(p, q) > 0$ such that $\pi^{\theta_1}(\Gamma_W^{\text{NS}}) \geq \rho_0$ and similarly for Γ_E^{NS} . The EDC property (2) also implies that $\pi^{\theta_1}(\Gamma_W^{\text{WE}}) \geq \rho_1$ and $\pi^{\theta_1}(\Gamma_E^{\text{WE}}) \geq \rho_1$, for a suitable $\rho_1 = \rho_1(p, q) > 0$. We justify this as follows. By the EDC property (2), there exists some constant D such that with probability $\Omega(1)$, no pair of vertices whose distance is at least D , one of which is in $\partial_N Q_W \setminus \partial_W Q_W$ and the other in $\partial_E Q_W \cup \partial_S Q_W \cup \partial_E Q_W$, will be connected in Q_W . At the same time, with probability $\Omega(1)$, we can force $O(D)$ edges bordering those pairs of vertices that are distance less than D to be closed so that no pair of vertices that are closer than D are connected in Q_W either. The analogous reasoning holds for Γ_E^{WE} and Q_E . Since each of the events $\Gamma_W^{\text{NS}}, \Gamma_W^{\text{WE}}, \Gamma_E^{\text{NS}}, \Gamma_E^{\text{WE}}$ are decreasing events, by the FKG inequality (see, e.g., [21]) we get

$$\begin{aligned} \pi^{\theta_1}(\Gamma) &\geq \pi^{\theta_1}(\Gamma_W^{\text{NS}} \cap \Gamma_W^{\text{WE}} \cap \Gamma_E^{\text{NS}} \cap \Gamma_E^{\text{WE}}) \\ &\geq \pi^{\theta_1}(\Gamma_W^{\text{NS}}) \pi^{\theta_1}(\Gamma_W^{\text{WE}}) \pi^{\theta_1}(\Gamma_E^{\text{NS}}) \pi^{\theta_1}(\Gamma_E^{\text{WE}}) \geq \rho_0^2 \rho_1^2, \end{aligned}$$

and thus we can take $\rho = \rho_0^2 \rho_1^2$ to have the desired estimate. \square

PROOF OF LEMMA 4.15. The proof of Lemma 4.14 carries to Lemma 4.15 with certain natural modifications we describe next. As before, we let $L = \partial_W Q_W \cup \partial_N Q_W \cup \partial_E Q_E \cup \partial_N Q_E$ and let $(L^*, E(L^*))$ be the dual-graph induced by the set of dual-edges intersecting $E(L)$ in $(\mathbb{Z}^2)^*$; its vertex set consists of exactly $2E(L) - 1$ vertices, and we refer to those outside of \mathcal{R}_{INT} as $\partial_{\text{EXT}} L^*$. Similarly, let $(\mathcal{R}_{\text{INT}}^*, E(\mathcal{R}_{\text{INT}}^*))$ and $(\mathcal{A}_{\text{INT}}^*, E(\mathcal{A}_{\text{INT}}^*))$ be the dual-graphs induced by the dual-edges intersecting $E(\mathcal{R}_{\text{INT}})$ and $E(\mathcal{A}_{\text{INT}})$ in $(\mathbb{Z}^2)^*$, respectively.

For a FK configuration ω on $E(\mathcal{R}_{\text{INT}})$, we define a dual version of $F(\omega)$ as

$$F^*(\omega) := \mathcal{R}_{\text{INT}}^* \setminus \bigcup_{v^* \in \partial_{\text{EXT}} L^*} C^*(v^*, \omega),$$

where $C^*(v^*, \omega)$ is the dual-vertex set of the connected component of v^* in the dual-configuration ω^* (ignoring the boundary connections).

Using monotone couplings for π^{θ_0} and π^{θ_X} , and for π^{θ_0} and π^{θ_Y} , we can define \mathbb{P}_0 analogously to \mathbb{P}_1 with the configuration on $E(F^*(\omega^{\theta_0}))$ being resampled whenever $\mathcal{A}_{\text{INT}}^* \subseteq F^*(\omega^{\theta_0})$. Observe that updating the dual configuration on $E(F^*(\omega^{\theta_0}))$ is equivalent to updating the primal edges intersecting $E(F^*(\omega^{\theta_0}))$. The fact that resampling the configuration in $E(F^*(\omega^{\theta_0}))$ has no effect on the distribution when $\mathcal{A}_{\text{INT}}^* \subseteq F^*(\omega^{\theta_0})$ follows in similar fashion to the proof of Lemma 4.14. Indeed, from the definition of $F^*(\omega^{\theta_0})$, when $\mathcal{A}_{\text{INT}}^* \subseteq F^*(\omega^{\theta_0})$ there is a primal connection between (c_\star, l) and $\partial_S Q_W$ in $E(Q_W) \cup \partial E(Q_W)$ together with a primal connection between (d_\star, l) and $\partial_S Q_E$ in $E(Q_E) \cup \partial E(Q_E)$, so that $\mathcal{A}_{\text{INT}}^* \subseteq F^*(\omega^{\theta_0})$ implies the event Γ^* . Together with the fact that $\llbracket c_\star, d_\star \rrbracket$ is a disconnecting interval and the assumptions that ξ is free on $\partial_S \mathcal{A}_{\text{INT}}$ and $(c_\star, l), (d_\star, l) \notin \partial_{\parallel} \mathcal{R}$, this ensures that the three induced boundary conditions on $E(F^*(\omega^{\theta_0}))$ coincide.

Finally, part 3 of the lemma follows by an analogous argument to that in Claim 4.16, only replacing the EDC property by the matching exponential decay of dual-connectivities when $p > p_c(q)$. \square

4.7. Proof of Theorem 4.1. In this section, we put together the results from Sections 4.3–4.6 to prove Theorem 4.1. We first remind the reader of the following standard inequality concerning the spectral gaps of the FK and block dynamics.

THEOREM 4.17 ([27], Proposition 3.4). *Consider the FK-dynamics on a group of rectangles \mathcal{R} with boundary condition ζ . Let $\text{gap}(\mathcal{R}^\zeta)$ and $\text{gap}(\mathcal{R}^\zeta; \mathcal{B})$, respectively, be the spectral*

gaps of the FK-dynamics on \mathcal{R} and of the block dynamics with blocks $\mathcal{B} = \{\mathcal{B}_1, \dots, \mathcal{B}_k\}$ such that $\mathcal{B}_i \subset \mathcal{R}$ and $\bigcup_{i=1}^k E(\mathcal{B}_i) = E(\mathcal{R})$. For every p, q , there exists $\gamma = \gamma(p, q) \in (0, 1)$ such that

$$\text{gap}(\mathcal{R}^\zeta) \geq \gamma \cdot \left(\max_{e \in E(\mathcal{R})} \#\{i : E(\mathcal{B}_i) \ni e\} \right)^{-1} \cdot \text{gap}(\mathcal{R}^\zeta; \mathcal{B}) \cdot \min_{\substack{i=1, \dots, k \\ \eta \in \Omega(\mathcal{B}_i^\zeta)}} \text{gap}(\mathcal{B}_i^{(\zeta, \eta)}),$$

where $\Omega(\mathcal{B}_i^\zeta)$ denotes the set of FK configurations on $E(\mathcal{R}) \setminus E(\mathcal{B}_i)$.

The proposition in [27] is written in the spin system setting, but the proof follows mutatis mutandis for the random-cluster model and its proof is thus omitted. Also, we note that this theorem holds in more generality for arbitrary graphs with arbitrary boundary conditions, but for clarity we choose to state it here for groups of rectangles.

The final ingredient is the following spectral gap bound for the base case in our recursive proof.

LEMMA 4.18. *Consider a group of rectangles $\mathcal{R}_0 \subset \Lambda_{n,l}$ with $W(\mathcal{R}_0) \leq 100m$. For every $q > 1$ and $p \neq p_c(q)$, there exists $\kappa = \kappa(p, q) > 0$ such that for every boundary condition ζ on \mathcal{R}_0 ,*

$$\text{gap}(\mathcal{R}_0^\zeta) \geq \frac{1}{l(\log l)^2 \cdot q^{\kappa l}}.$$

PROOF. Note that $|\partial \mathcal{R}_0| = O(m + l)$. Hence, we can first modify the boundary conditions to be all free on all of $\partial \mathcal{R}_0$, incurring a cost of a $q^{O(l)}$ factor in the spectral gap by Lemma 2.3; recall that $m = O(\log l)$. Then we can use the fast mixing result of [4], for instance, to bound the mixing time on \mathcal{R}_0 with free boundary condition by $O(l(\log l)^2)$. This translates into a lower bound for the spectral gap and the result follows. \square

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 $\partial_E \Lambda_{n,l} \cup \partial_S \Lambda_{n,l} \cup \partial_W \Lambda_{n,l}$. By Lemma 2.3, we may modify ξ' to a boundary condition ξ that is also free on all vertices a distance at most $m = C_* \log l$ from $\partial_E \Lambda_{n,l} \cup \partial_W \Lambda_{n,l}$ at a cost of an exponential in m factor in the mixing time of the FK-dynamics. Let ξ be the resulting realizable boundary condition.

We wish to prove, by induction, that for every $100m \leq s \leq n$, every group of rectangles $\mathcal{R}_s \subset \Lambda_{n,l}$ that is compatible with ξ and has $W(\mathcal{R}_s) = s$ satisfies

$$(16) \quad \text{gap}(\mathcal{R}_s^{(\xi, \omega_{\mathcal{R}_s^c})}) \geq \frac{1}{l(\log l)^2 q^{\kappa l} \cdot b^{\log s}}$$

for some $b = b(p, q) > 0$ to be chosen, uniformly over all configurations $\omega_{\mathcal{R}_s^c}$ on $E^c(\mathcal{R}_s)$. Equation (16) concludes the proof since $\Lambda_{n,l}$ is a group of rectangles with $W(\Lambda_{n,l}) = n$ and is compatible with ξ .

The base case of this induction was shown in Lemma 4.18. Now suppose inductively that this holds for all $1 \leq k \leq s - 1$ for some $s \leq n$; we show that it also holds for s . Fix any \mathcal{R}_s that is compatible with ξ , and any configuration $\omega_{\mathcal{R}_s^c}$. Then, if we let $\mathcal{R}_{\text{INT}} = \mathcal{R}_{\text{INT}}(\mathcal{R}_s)$ and $\mathcal{R}_{\text{EXT}} = \mathcal{R}_{\text{EXT}}(\mathcal{R}_s)$ be the blocks given by Definition 4.10 and \mathcal{B}_s the block-dynamics with respect to these blocks, by Theorem 4.17

$$\begin{aligned} \text{gap}(\mathcal{R}_s^{(\xi, \omega_{\mathcal{R}_s^c})}) &\geq \frac{\gamma}{2} \cdot \text{gap}(\mathcal{R}_s^{(\xi, \omega_{\mathcal{R}_s^c})}; \mathcal{B}_s) \cdot \min_{i \in \{\text{INT}, \text{EXT}\}} \min_{\omega_{\mathcal{R}_i^c}} \text{gap}(\mathcal{R}_i^{(\xi, \omega_{\mathcal{R}_i^c})}) \\ &\geq \frac{\gamma}{2K} \cdot \min_{i \in \{\text{INT}, \text{EXT}\}} \min_{\omega_{\mathcal{R}_i^c}} \text{gap}(\mathcal{R}_i^{(\xi, \omega_{\mathcal{R}_i^c})}), \end{aligned}$$

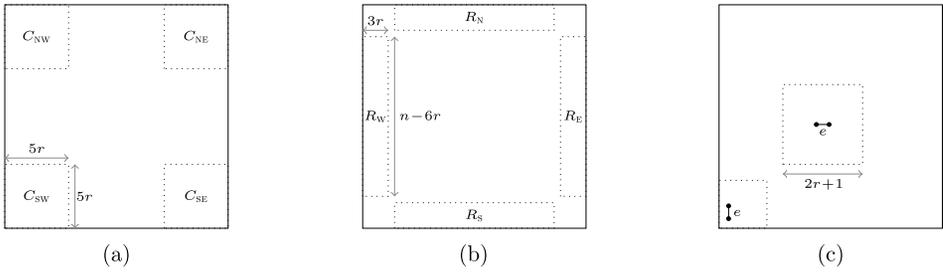


FIG. 5. (a) The subsets C_{NE} , C_{NW} , C_{SE} , and C_{SW} . (b) The subsets R_N , R_E , R_W and R_S . (c) $B(e, r)$ for two edges e of Λ_n .

where the second inequality follows from Lemma 4.13. By Proposition 4.11, $\max\{W(\mathcal{R}_{\text{INT}}), W(\mathcal{R}_{\text{EXT}})\} \leq \frac{4}{5}s$, and we can apply the inductive hypothesis to bound the second term on the right-hand side above. Combined with Lemma 4.18, we see that the choice of $b = (2\gamma^{-1}K)^{\frac{1}{\log(5/4)}}$ ensures that (16) holds also for \mathcal{R}_s . (Note that $2\gamma^{-1}K \geq 1$.)

This establishes the result for the case when the boundary condition is free on $\partial_E \Lambda_{n,l} \cup \partial_S \Lambda_{n,l} \cup \partial_W \Lambda_{n,l}$. As noted earlier (see Remark 2), this implies by duality the same bound for the class of realizable boundary conditions ξ that are wired on $\partial_E \Lambda_{n,l} \cup \partial_S \Lambda_{n,l} \cup \partial_W \Lambda_{n,l}$ for all $p \neq p_c(q)$. \square

5. Polynomial mixing time for realizable boundary conditions. In this section, we prove Theorem 1.1. This theorem is proved 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 3.1–3.2. 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 2.1.

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 straightforward 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 use Theorem 4.1.

We now define the collection of blocks for which we can establish both LM and MSM. Let $r \in \mathbb{N}$ and let $C_{NE}, C_{NW}, C_{SE}, C_{SW} \subset \Lambda_n$ be the four square boxes of side length $5r$ with a corner that coincides with a corner of Λ_n ; see Figure 5(a). Let $R_N \subset \Lambda_n$ be the $(n - 6r) \times 2r$ rectangle at distance $3r$ from both $\partial_W \Lambda_n$ and $\partial_E \Lambda_n$ whose top boundary is contained in $\partial_N \Lambda_n$ and let R_E, R_W, R_S be defined analogously; see Figure 5(b). Let $R = R_N \cup R_E \cup R_W \cup R_S$. 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 $d(\{e\}, \Lambda_n \setminus B(e, r)) \geq r$. Note that 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 5(c). Finally, let

$$(17) \quad \mathcal{B}_r = \{C_{NE}, C_{NW}, C_{SE}, C_{SW}, R\} \cup \{B(e, r) : e \in E(\Lambda_n), d(\{e\}, \partial \Lambda_n) > r\}.$$

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 $R = R_N \cup R_E \cup R_W \cup R_S$ 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. Theorem 1.1 then follows from Theorems 5.1, 5.2 and 3.3.

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 $\delta < 1/(12|E(\Lambda_n)|)$.*

We are now ready to prove Theorem 1.1 using the above.

PROOF OF THEOREM 1.1. As mentioned earlier, by duality of the dynamics and self-duality of the class of realizable boundary conditions, it suffices to prove the theorem for $p < p_c(q)$. Let \mathcal{P} be the set of realizable boundary conditions of $\Lambda_n = (\Lambda_n, E(\Lambda_n))$. For $\eta \in \mathcal{P}$, let $(\eta_1, \eta_2, \dots, \eta_k)$ denote the partition of $\partial\Lambda_n$ corresponding to η , and let $\eta(\ell)$ be the boundary condition obtained as follows: for each $v \in \partial\Lambda_n$, if $v \in \eta_i$ and v is at distance at most ℓ from a corner of Λ_n , remove v from η_i and add it as a singleton to the partition. Let \mathcal{P}_ℓ be the set of all boundary conditions obtained in this manner.

Consider Λ_n with arbitrary realizable boundary conditions $\xi \in \mathcal{P}$. By Lemma 2.3, we see that there exists $C > 0$ such that for every $\xi \in \mathcal{P}$, we have

$$t_{\text{MIX}}(\Lambda_n^\xi) \leq Cq^{8C\ell} \cdot n^2 \cdot t_{\text{MIX}}(\Lambda_n^{\xi(\ell)}).$$

It therefore suffices to prove the mixing time estimate uniformly over all modified boundary conditions $\eta \in \mathcal{P}_\ell$ for $\ell = 5r$ and $r = c_0(\log n)$ with c_0 taken to be sufficiently large, as $q^{8C\ell}$ would only be polynomial in n . By Theorem 5.2, for c_0 large enough, uniformly over all such boundary conditions we have moderate spatial mixing with respect to \mathcal{B}_r and $\eta \in \mathcal{P}_\ell$ with $\delta < 1/(12|E(\Lambda_n)|)$. Theorem 5.1 implies that LM holds for \mathcal{B}_r and $\eta \in \mathcal{P}_\ell$ with $T = O(n^c)$ where $c > 0$ constant. The result then follows from Theorem 3.3. \square

REMARK 7. We note that Theorem 1.1 also holds for the FK-dynamics on rectangles $\Lambda_{n,l} \subset \mathbb{Z}^2$ with $\ell \leq n$, provided these rectangles are not too thin. For example, when $l = \Omega((\log n)^2)$, our proofs would yield that the mixing time of the FK-dynamics is polynomial in n .

5.1. Local mixing for realizable boundary conditions. In this subsection, we prove Theorem 5.1. As mentioned earlier, this theorem may be viewed as a corollary of Theorem 4.1, which bounds the mixing time of the FK-dynamics on thin rectangles.

PROOF OF THEOREM 5.1. Let $r = c_0 \log n$. We wish to show that each of the subsets in \mathcal{B}_r has mixing time $O(n^c)$ under the boundary conditions $(1, \xi)$ and $(0, \xi)$. We begin by bounding the mixing time on the square boxes $C_{NE}, C_{NW}, C_{SE}, C_{SW}$ and $B(e, r)$ of \mathcal{B}_r . Since these have side length $O(\log n)$, a crude estimate on the mixing time is sufficient. For instance, by Lemma 2.3, at a cost of $\exp(O(r)) = n^{O(1)}$ factor, we can compare the mixing time in these boxes with boundary condition either $(1, \xi)$ or $(0, \xi)$ to the mixing time on equally sized boxes with free boundary conditions. In this setting, an upper bound of $O((\log n)^2 \log \log n)$ is known [4], and thus we obtain an $n^{O(1)}$ bound for their mixing times.

It remains to bound the mixing time of the FK-dynamics on the set $R = R_N \cup R_E \cup R_W \cup R_S$. For this, we use Theorem 4.1. We argue that the mixing time of the FK-dynamics on R is roughly equal to that of the FK-dynamics on a $[4(n - 6r) - 3] \times 2r$ rectangle Q with a suitably chosen boundary condition. We proceed to construct the rectangle Q and a boundary condition ξ' whose vertices, edges and wirings are identified with those of R and $(1, \xi)$. The case of R and $(0, \xi)$ is handled later in similar fashion.

We introduce some notation first. For a rectangle S , let S^α denote the rectangle that results from a clockwise rotation of S by an angle of amplitude α . Also, if S_1, \dots, S_k are rectangles of the same height, let $[S_1, \dots, S_k]$ denote the rectangle obtained by identifying the vertices in $\partial_E S_i$ with those in $\partial_W S_{i+1}$ for all $i = 1, \dots, k - 1$. When identifying the vertices, the double edges are removed. We take

$$Q = [R_W^{\pi/2}, R_N, R_E^{-\pi/2}, R_S^{-\pi}].$$

Observe that every vertex of Q , except those where the boundary overlaps occur, correspond to exactly one vertex in R ; vertices in the overlaps correspond to exactly two vertices in R . Conversely, every vertex in R corresponds to exactly one vertex of Q . The edges of Q and R are identified using this correspondence between the vertices. Observe also that $\partial_N Q = \partial R \cap \partial \Lambda_n$. We construct the boundary condition ξ' of Q as follows. If $u, v \in \partial R \cap \partial \Lambda_n$ are wired in ξ , the corresponding vertices are also wired in ξ' . The boundary condition ξ' is also wired along $\partial_W Q \cup \partial_N Q \cup \partial_E Q$.

CLAIM 5.3. *The boundary condition ξ' of Q is realizable. In particular, ξ' can be realized by a FK configuration in the half-plane of \mathbb{Z}^2 containing only vertices north of $\partial_N Q$, and a wiring of $\partial_E Q \cup \partial_S Q \cup \partial_W Q$.*

Finally, to completely capture the effect of the boundary condition $(1, \xi)$ on R , each of the three columns in Q that corresponds to overlaps of columns from R , are externally wired.

Now, by Theorem 4.1 and Lemma 2.3, we have

$$\text{gap}(Q^{\xi'}) = n^{-O(1)}.$$

We claim next that the FK-dynamics on R with boundary condition $(1, \xi)$ has roughly the same gap as the FK-dynamics on Q with boundary condition ξ' . To see this, we add a double edge to each edge of Q that corresponds to two edges in R . With this modification, there is now a one-to-one correspondence between the FK configurations in R and Q . Also, adding these edges has almost no effect on the mixing time of the FK-dynamics in Q , as their endpoints are wired, and so they only need to be updated once to mix. Moreover, by construction, the boundary condition ξ' for Q together with the wiring of the overlapping columns in Q encode exactly the same connectivities as the boundary condition $(1, \xi)$ for R . Hence, for every pair of FK configurations on Q , the FK-dynamics have the same transition probability as FK-dynamics on R between the corresponding configurations. Consequently, we can conclude that

$$\text{gap}(R^{(1, \xi)}) = n^{-O(1)}.$$

Finally, for the case of the FK-dynamics on R with boundary condition $(0, \xi)$, we can simply wire $\partial_W R_N$ to $\partial_N R_W$, $\partial_S R_W$ to $\partial_W R_S$, $\partial_E R_S$ to $\partial_S R_E$ and $\partial_N R_E$ to $\partial_E R_N$, which only incur a penalty of $n^{O(1)}$ by Lemma 2.3 and proceed as in the previous case. \square

PROOF OF CLAIM 5.3. First, note that $\partial_N Q$ corresponds to $\partial R \cap \partial \Lambda_n$. Let ω be a FK configuration on $\mathbb{Z}^2 \setminus \Lambda_n$ that realizes ξ . A path from $u \in \partial R \cap \partial \Lambda_n$ to $v \in \partial R \cap \partial \Lambda_n$ in ω splits $\partial R \cap \partial \Lambda_n$ into two parts R_1 and R_2 , one containing all the vertices from u to v in $\partial R \cap \partial \Lambda_n$ clockwise and the other all the vertices from u to v in $\partial R \cap \partial \Lambda_n$ counterclockwise. The planarity of \mathbb{Z}^2 implies that any other boundary component will be either completely contained in R_1 or R_2 . From this property, it follows that if $v_1, v_2 \in \partial_N Q$ are wired in ξ' , then $\llbracket v_1, v_2 \rrbracket$ is a disconnecting interval. This implies that the connectivities of ξ' in $\partial_N Q$ can be realized by a configuration on the half-plane of \mathbb{Z}^2 that contains all the vertices north of $\partial_N Q$. For example, every component $C = \{c_0, \dots, c_k\}$ of ξ , with c_i to the left of c_{i+1} , can be realized by the gadget consisting k paths of length h_C starting at c_0, \dots, c_k and going north, together with one path parallel to $\partial_N Q$ that joins the endpoints of all of these path. Since $\llbracket c_i, c_{i+1} \rrbracket$ is a disconnecting interval for all i and C , we can choose h_C for each C so that the resulting configuration is a valid configuration in the half-plane. \square

5.2. Moderate spatial mixing for realizable boundary conditions. In this section, we prove Theorem 5.2. We reduce the moderate spatial mixing condition (6) to bounding the probability of certain connectivities in a FK configuration. Specifically, if $e \in S \subset \Lambda_n$, the configuration on $E^c(S)$ affects the state of e when there are paths from e to the boundary of S ; the probability of such paths is maximized when we assume an all wired configuration on $E^c(S)$. Recall that for $S \subset \Lambda_n$, we let $S^c = \Lambda_n \setminus S$, and we use $E^c(S) = E(\Lambda_n) \setminus E(S)$.

LEMMA 5.4. *Consider the FK model on Λ_n with arbitrary boundary condition ξ on $\partial \Lambda_n$. For any $e \in E(\Lambda_n)$, any $S \subset \Lambda_n$ such that $e \in E(S)$, and any pair of configurations ω_1, ω_2 on $E^c(S)$:*

$$\begin{aligned} & \left| \pi_{\Lambda_n}^\xi(e = 1 \mid E^c(S) = \omega_1) - \pi_{\Lambda_n}^\xi(e = 1 \mid E^c(S) = \omega_2) \right| \\ & \leq \pi_{\Lambda_n}^\xi(\{e\} \overset{\xi}{\longleftrightarrow} \partial S \setminus \partial \Lambda_n \mid E^c(S) = 1), \end{aligned}$$

where $\{e\} \overset{\xi}{\longleftrightarrow} \partial S$ denotes the event that there is a path from e to ∂S taking into account the connections induced by ξ .

In the proof of Theorem 5.2, we use this lemma; its proof via machinery from [1] will be straightforward.

PROOF OF THEOREM 5.2. We need to show that as long as c_0 is large enough, for every $e \in E(\Lambda_n)$ there exists $B_e \in \mathcal{B}_r$ such that (6) holds for some $\delta < 1/(12|E(\Lambda_n)|)$. For each $e \in E(\Lambda_n)$, the subset B_e is chosen as follows:

1. If $d(\{e\}, \partial \Lambda_n) > r$, then $B_e = B(e, r)$;
2. Otherwise, if $e \in R$ and $d(\{e\}, \partial R \setminus \partial \Lambda) \geq r$, then $B_e = R$;
3. Otherwise, $e \in C_i$ for some $i \in \{\text{NE}, \text{NW}, \text{SW}, \text{SE}\}$, and we take $B_e = C_i$.

By Lemma 5.4, for every $e \in E(\Lambda_n)$,

$$\begin{aligned} & \left| \pi_{\Lambda_n}^\xi(e = 1 \mid E^c(B_e) = 1) - \pi_{\Lambda_n}^\xi(e = 1 \mid E^c(B_e) = 1) \right| \\ & \leq \pi_{\Lambda_n}^\xi(e \overset{\xi}{\longleftrightarrow} \partial B_e \setminus \partial \Lambda_n \mid E^c(B_e) = 1). \end{aligned}$$

In all three cases above, by construction $d(\{e\}, \partial B_e \setminus \partial \Lambda_n) \geq r$. This together with the fact that all vertices of $\partial \Lambda_n$ within distance $5r$ from the corners have no connections in ξ , implies that for e to be connected to $\partial B_e \setminus \partial \Lambda_n$ a path of open edges reaching a distance at least r is required in the configuration on B_e . The EDC property (see (2)) implies that for c_0 large enough

$$\pi_{\Lambda_n}^\xi(\{e\} \xleftrightarrow{\eta} \partial B_e \mid E^c(B_e) = 1) \leq \frac{1}{12|E(\Lambda_n)|},$$

and the result follows. \square

We conclude this section with the proof of Lemma 5.4.

PROOF OF LEMMA 5.4. Let $(1, \xi)$ be the boundary condition induced on S by ξ and the event $\{E^c(S) = 1\}$. Similarly, let θ_1 (resp., θ_2) be the boundary condition induced on S by configurations ω_1 (resp., ω_2) on $E^c(S)$ and ξ . For ease of notation, set $\pi^{\theta_1} = \pi_{\Lambda_n}^\xi(\cdot \mid E^c(S) = \omega_1)$, $\pi^{\theta_2} = \pi_{\Lambda_n}^\xi(\cdot \mid E^c(S) = \omega_2)$ and $\pi^{(1, \xi)} = \pi_{\Lambda_n}^\xi(\cdot \mid E^c(S) = 1)$. For a FK configuration ω on S , let

$$\Gamma^{(1, \xi)}(\omega) := S \setminus \bigcup_{v \in \partial S \setminus \partial \Lambda_n} C(v, \omega),$$

where $C(v, \omega)$ is the set of vertices in the connected component of v in ω , taking into account the connectivities induced by $(1, \xi)$. In words, $\Gamma^{(1, \xi)}(\omega)$ is the set of vertices of S not connected to $\partial S \setminus \partial \Lambda_n$ in ω using possibly the boundary connections.

We claim that there exists a coupling \mathbb{P} of the distributions π^{θ_1} , π^{θ_2} and $\pi^{(1, \xi)}$ such that $\mathbb{P}(\omega_1, \omega_2, \omega) > 0$ only if $\omega_1 \leq \omega$ and $\omega_2 \leq \omega$ on $E(S)$ and ω_1, ω_2 agree on all edges with both endpoints in $\Gamma^{(1, \xi)}(\omega)$. Given this coupling \mathbb{P} , we have

$$\begin{aligned} &|\pi_{\Lambda_n}^\xi(e = 1 \mid E^c(S) = \omega_1) - \pi_{\Lambda_n}^\xi(e = 1 \mid E^c(S) = \omega_2)| \\ &\leq \mathbb{P}(\omega_1(e) \neq \omega_2(e)) \\ &\leq \mathbb{P}(e \notin E(\Gamma^{(1, \xi)}(\omega))) \\ &= \pi_{\Lambda_n}^\xi(e \xleftrightarrow{\xi} \partial S \setminus \partial \Lambda_n \mid E^c(S) = 1), \end{aligned}$$

as claimed. The construction of the coupling \mathbb{P} is standard and is thus omitted; see, for example, [1, 4] and the proof Lemma 4.14 for similar constructions. \square

6. Near optimal mixing for typical boundary conditions. In this section, we provide the proof of Theorem 1.4, where we establish a sharper $\tilde{O}(n^2)$ mixing time upper bound for the FK-dynamics on $\Lambda_n = (\Lambda_n, E(\Lambda_n))$ for the class of boundary conditions we call *typical*.

DEFINITION 6.1. Let ω be a random-cluster configuration on \mathbb{Z}^2 , and let ξ_ω be the boundary condition on $\partial \Lambda_n$ induced by the edges of ω in $E(\mathbb{Z}^2) \setminus E(\Lambda_n)$. Suppose ω is sampled from $\pi_{\mathbb{Z}^2, p, q}$. A set \mathcal{C} of realizable boundary conditions for Λ_n is called *typical* (with respect to (p, q)) if $\xi_\omega \in \mathcal{C}$ with probability $1 - o(1)$.

Recall from Definition 1.2 the classes of boundary conditions \mathcal{C}_α and \mathcal{C}_α^* , consisting of realizable boundary conditions whose distinct boundary components consist only of vertices at most distance $\alpha \log n$ apart in $\partial \Lambda_n$. The following is a straightforward consequence of the EDC property (2) when $p < p_c(q)$.

LEMMA 6.2. *For every $q \geq 1$ and $p < p_c(q)$, the class of boundary conditions \mathcal{C}_α is typical with respect to (p, q) for sufficiently large $\alpha > 0$. Similarly, for every $q \geq 1$ and $p > p_c(q)$, the class \mathcal{C}_α^* is typical with respect to (p, q) for sufficiently large $\alpha > 0$.*

PROOF. By planar duality (namely the duality of the sets of boundary conditions \mathcal{C}_α and \mathcal{C}_α^* , it suffices to prove the case $p < p_c(q)$). For any $u, v \in \mathbb{Z}^2$, by the EDC property (2), we have that for $q \geq 1$ and $p < p_c(q)$ there exists $c = c(p, q) > 0$ such that $\pi_{\mathbb{Z}^2}(u \leftrightarrow v) \leq e^{-cd(u,v)}$. Let $u, v \in \partial\Lambda_n$ and suppose $d(u, v) \geq \alpha \log n$. Then there exists some $C(p, q) > 0$ such that for sufficiently large $\alpha > 0$,

$$\pi_{\mathbb{Z}^2 \setminus \Lambda_n}(u \overset{\mathbb{Z}^2 \setminus \Lambda_n}{\longleftrightarrow} v) = \pi_{\mathbb{Z}^2}(u \overset{\mathbb{Z}^2 \setminus \Lambda_n}{\longleftrightarrow} v) \leq \pi_{\mathbb{Z}^2}(u \leftrightarrow v) \leq Ce^{-c\alpha \log n} \leq \frac{1}{n^3},$$

where recall that $u \overset{\mathbb{Z}^2 \setminus \Lambda_n}{\longleftrightarrow} v$ denotes the event that there exists a path from u to v in $\mathbb{Z}^2 \setminus \Lambda_n$. A union bound over all pairs of vertices in $\partial\Lambda_n$ implies that if ω is sampled from $\pi_{\mathbb{Z}^2}$ and ξ_ω is the resulting boundary condition on $\partial\Lambda_n$, then $\xi_\omega \in \mathcal{C}_\alpha$ with probability $1 - o(1)$, and thus \mathcal{C}_α is typical. \square

REMARK 8. As mentioned in the Introduction, one may also be interested in the following notion of typicality, which sometimes comes up in recursive mixing time upper bounds. Let $q \geq 1$ and $p < p_c(q)$ (resp., $p > p_c(q)$) and consider a random-cluster sample from $\pi_{R_{2n}, p, q}^\xi$, where R_{2n} is the concentric box of side length $2n$ containing Λ_n with arbitrary boundary condition ξ . One could easily show that the boundary condition induced by the configuration on $R_{2n} \setminus \Lambda_n$ is in \mathcal{C}_α (resp., \mathcal{C}_α^*) with probability $1 - o(1)$. This follows by coupling this measure to the infinite-volume measure using the fact that \mathcal{C}_α is a decreasing event, and finding a dual circuit in the annulus $R_{2n} \setminus \Lambda_n$ (which exists with probability $1 - O(e^{-\Omega(n)})$).

We now show that when $p < p_c(q)$, the mixing time on Λ_n with boundary condition $\xi \in \mathcal{C}_\alpha$ satisfies

$$t_{\text{MIX}}(\Lambda_n^\xi) = O(n^2(\log n)^C),$$

where $C = C(p, q, \alpha) > 0$ is a constant independent of n and ξ . In particular, we prove Theorem 1.4 from the introduction in the regime $p < p_c(q)$ and $\xi \in \mathcal{C}_\alpha$ and this translates to a matching bound at $p > p_c(q)$ and $\xi \in \mathcal{C}_\alpha^*$ by duality. To prove this theorem, we again use the general framework from Theorem 3.3. Namely, we construct a collection of subsets of Λ_n for which we can establish MSM and LM; see Definitions 3.1 and 3.2. The fact that $\xi \in \mathcal{C}_\alpha$ will allow us to prove MSM with respect to $\Theta((\log n)^2) \times \Theta((\log n)^2)$ rectangles along the boundary, and Theorem 1.1 will provide the LM estimate on these rectangles.

Consider the collection $\mathcal{B}_r = \{B(e, r) : e \in E(\Lambda_n)\}$. Recall that for $r \geq 0$ and $e \in E(\Lambda_n)$, we set $B(e, r) \subset \Lambda_n$ to be the set of vertices in the minimal square box around e such that $d(\{e\}, \Lambda_n \setminus B(e, r)) \geq r$; see Figure 5(c). We first show that MSM holds for \mathcal{B}_r and $\xi \in \mathcal{C}_\alpha$ when $r = \Theta((\log n)^2)$ and $\delta < n^{-3}$.

LEMMA 6.3. *Let $q \geq 1$, $p < p_c(q)$, $\alpha > 0$, $\eta \in \mathcal{C}_\alpha$, $r = c_0(\log n)^2$ and $\mathcal{B} = \{B(e, r) : e \in E(\Lambda_n)\}$. For large enough $c_0 > 0$, MSM holds for η, \mathcal{B} for some $\delta < n^{-3}$.*

For this lemma, it is crucial that $r = \Theta((\log n)^2)$, as MSM does not hold for typical boundary conditions for \mathcal{B}_r when, for example, $r = \Theta(\log n)$. This is because in a typical configuration ω on $\mathbb{Z}^2 \setminus \Lambda_n$ it is likely that there exist pairs of vertices of $\partial\Lambda_n$ at distance $\gamma \log n$, for a suitably small constant $\gamma > 0$, that are connected in ω . Thus, for some $e \in E(\Lambda_n)$ close to

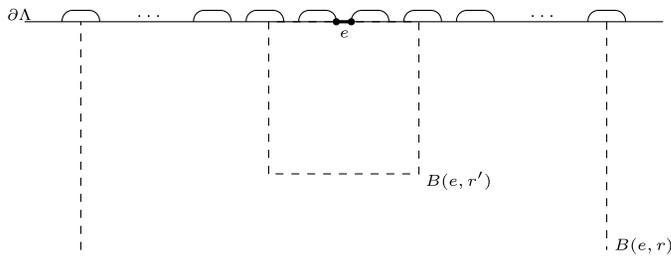


FIG. 6. If $r = \Theta((\log n)^2)$ and $r' = \Theta(\log n)$, influence from outside of $B(e, r')$ may be easily propagated to e through long boundary connections in $B(e, r')$; but to propagate influence from the exterior of $B(e, r)$, $\Omega(\log n)$ of them would have to be connected in Λ_n .

$\partial\Lambda_n$, it is possible for the configuration outside of $B(e, r)$ to exert a strong influence on the state of e when $r = \gamma' \log n$ with constant $\gamma' > 0$, even if $\gamma' \gg \gamma$; the presence of a constant number of open edges on (or near) $\partial\Lambda_n$ would propagate the influence from $\Lambda_n \setminus B(e, r)$ to e . Taking $r = \Omega((\log n)^2)$ avoids this issue, since, roughly speaking, $\Omega(\log n)$ open edges at specific points in $\partial\Lambda_n$ would now be required to propagate the influence from $\Lambda_n \setminus B(e, r)$ to e ; see Figure 6(b). The proof of Lemma 6.3 is provided in Section 6.1.

The final ingredient in the proof of Theorem 1.4 is a LM estimate for \mathcal{B}_r with $r = \Theta((\log n)^2)$. Such an estimate is readily provided by Theorem 1.1, with mixing time that is poly-logarithmic in n .

PROOF OF THEOREM 1.4. Let $\alpha > 0$ be sufficiently large and let $\eta \in C_\alpha$. By Lemma 6.3, MSM holds for η and \mathcal{B}_r with $r = \Theta((\log n)^2)$ for some $\delta < n^{-3}$. Observe also that every $B(e, r) \in \mathcal{B}_r$ with boundary condition $(1, \eta)$ or $(0, \eta)$ is a rectangle of side-length at most $r = O((\log n)^2)$ with a realizable boundary condition. Then by Theorem 1.1 (see also Remark 7), for every $e \in E(\Lambda_n)$ we have

$$\max\{t_{\text{MIX}}(B(e, r)^{0,\eta}), t_{\text{MIX}}(B(e, r)^{1,\eta})\} \leq (\log n)^C,$$

for a suitable $C > 0$, yielding the desired LM estimate. The result then follows from Theorem 3.3. \square

6.1. Moderate spatial mixing for C_α . We now prove Lemma 6.3. The proof involves showing that if $\xi \in C_\alpha$, when $p < p_c(q)$, the correlation between edges $e, e' \in E(\Lambda_n)$ near the boundary decays exponentially in $d(e, e')/(\alpha \log n)$ —whereas SSM would entail a decay rate that is exponential in just $d(e, e')$.

PROOF OF LEMMA 6.3. Fix an edge $e \in E(\Lambda_n)$ and for ease of notation let $B = B(e, r) \subset \Lambda_n$ and $\pi^\eta = \pi_{\Lambda_n, p, q}^\eta$. Let $(1, \eta)$ be the boundary condition induced on B by η and the event $\{E^c(B) = 1\}$.

Lemma 5.4 implies that for every pair of configurations ω_1, ω_2 on $E^c(B)$,

$$\begin{aligned} & |\pi^\eta(e = 1 \mid E^c(B) = \omega_1) - \pi^\eta(e = 1 \mid E^c(B) = \omega_2)| \\ (18) \quad & \leq \pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial B \setminus \partial \Lambda_n), \end{aligned}$$

where $\{e\} \xleftrightarrow{\eta} \partial B \setminus \partial \Lambda_n$ denotes the event that there is a path from e to $\partial B \setminus \partial \Lambda_n$ taking into account the connections induced by η . Thus, it is sufficient to bound the right-hand side of (18).

There are three cases corresponding to the location of e in Λ_n . First, if $d(\{e\}, \partial \Lambda_n) > r$, then $B \cap \partial \Lambda_n = \emptyset$ and $(1, \eta)$ is just the wired boundary condition on B . In this case the right-hand side of (18) is at most n^{-3} by the EDC property; see (2).

The second and third cases correspond to whether B intersects one or two sides of $\partial\Lambda_n$. For the second case, assume without loss of generality that B intersects $\partial_N\Lambda_n$ but not $\partial_W\Lambda_n$ or $\partial_E\Lambda_n$. That is, $d(\{e\}, \partial_N\Lambda_n) \leq r$, but $\{e\}$ is at distance at least r from $\partial_W\Lambda_n$ and $\partial_E\Lambda_n$. Let $\partial_W B, \partial_S B, \partial_E B$ be the west, south and east boundaries of B , all of which are wired in ω . By a union bound,

$$\begin{aligned} \pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial B \setminus \partial\Lambda_n) &\leq \pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial_W B) + \pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial_E B) \\ &\quad + \pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial_S B), \end{aligned}$$

where $\{e\} \xleftrightarrow{\eta} \partial_W B$ denotes the event that there is a path from e to $\partial_W B$ in B , taking into account those connections inherited from η (and ignoring the connections induced by the wired configuration on $E^c(B)$). Define $\{e\} \xleftrightarrow{\eta} \partial_E B$ and $\{e\} \xleftrightarrow{\eta} \partial_S B$ similarly.

The event $\{e\} \xleftrightarrow{\eta} \partial_S B$ implies that there exists a path of length at least r , either from $\{e\}$ or from $\partial\Lambda_n$ to $\partial_S B$. Therefore, the EDC property ((2)) implies that for large enough n ,

$$\pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial_S B) \leq \frac{1}{3n^3}.$$

We bound next $\pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial_W B)$. Let η_0, \dots, η_d be the boundary components of η . Since η is realizable, the planarity of \mathbb{Z}^2 implies that for every $i, j \in \{0, \dots, d\}$ there are only three possibilities: $L(\eta_i) \cap L(\eta_j) = \emptyset$, $L(\eta_i) \subset L(\eta_j)$ or $L(\eta_j) \subset L(\eta_i)$. (Recall that $L(\eta_i) \subset \partial\Lambda_n$ is the path of minimum length that contains all the vertices in η_i .) Call η_i a maximal boundary component if $\nexists j \in \{0, \dots, d\}$ such that $L(\eta_i) \subset L(\eta_j)$. The set of all maximal boundary components defines a partition for $\partial\Lambda_n$. Since also $\eta \in \mathcal{C}_\alpha$, we deduce that there exists a sequence of edges $e_0 = \{u_0, v_0\}, e_1, \dots, e_k = \{u_k, v_k\}$ in $B \cap \partial\Lambda_n$ such that 1) $(\gamma + \alpha) \log n \geq d(\{e_i\}, \{e_{i+1}\}) \geq \gamma \log n$ for all $i = 0, \dots, k - 1$, where γ is a large constant we choose later and $k \geq \frac{c_0}{2(\gamma + \alpha)} \log n$, and 2) the set $S_i = \llbracket v_i, u_{i+1} \rrbracket \subset B \cap \partial\Lambda_n$ is a disconnecting interval.

Let \mathcal{E}_i be the event that S_i is connected to S_{i+1} by a path of open edges in B . Let e_t be the closest edge in the sequence e_0, e_1, \dots, e_k to e and let $\hat{\mathcal{E}}_t = \bigcap_{i=0}^t \mathcal{E}_i$. Since $d(\{e\}, \partial_W B) \geq r$, we also have $t \geq \frac{c_0}{8(\gamma + \alpha)} \log n$. Then

$$(19) \quad \pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial_W B) \leq \pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial_W B \mid \hat{\mathcal{E}}_t^c) + \pi^{(1,\eta)}(\hat{\mathcal{E}}_t).$$

If the event $\hat{\mathcal{E}}_t^c$ occurs, then there exists $i < t$ such that S_i is not connected to S_{i+1} in B . This implies that there is a dual-path of length at least $\gamma \log n$ separating S_i from S_{i+1} . Consequently, a path from e to $\partial_W B$ would require two vertices at distance at least $\gamma \log n$ to be connected by a path of open edges in B . By the EDC property and a union bound, this has probability at most $1/(9n^3)$ for large enough γ . Thus,

$$(20) \quad \pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial_W B \mid \hat{\mathcal{E}}_t^c) \leq \frac{1}{9n^3}.$$

We bound next $\pi^{(1,\eta)}(\hat{\mathcal{E}}_t)$. Let $\{u_i, v_i\}$ denote the endpoints of the edge e_i , where u_i is to the left of v_i for all i . For $1 \leq i \leq t$, consider the rectangle $Q_i \subset B$ with corners at v_{i-1}, u_{i+3} and the other two corners on $\partial_S B$. Then

$$(21) \quad \pi^{(1,\eta)}(\hat{\mathcal{E}}_t) = \pi^{(1,\eta)}(\mathcal{E}_0, \dots, \mathcal{E}_t) \leq \pi^{(1,\eta)}(\mathcal{E}_1, \mathcal{E}_5, \mathcal{E}_9, \dots, \mathcal{E}_\ell),$$

where $t - 4 < \ell \leq t$. Now, let \mathcal{E}'_i be the event that S_i is connected to S_{i+1} by a path completely contained in Q_i . We have

$$\pi^{(1,\eta)}(\mathcal{E}_1, \mathcal{E}_5, \mathcal{E}_9, \dots, \mathcal{E}_\ell)$$

$$\begin{aligned} &= \pi^{(1,\eta)}(\mathcal{E}'_1, \mathcal{E}_5, \mathcal{E}_9, \dots, \mathcal{E}_\ell) + \pi^{(1,\eta)}(\mathcal{E}_1 \cap (\mathcal{E}'_1)^c, \mathcal{E}_5, \mathcal{E}_9, \dots, \mathcal{E}_\ell) \\ &\leq \pi^{(1,\eta)}(\mathcal{E}'_1, \mathcal{E}_5, \mathcal{E}_9, \dots, \mathcal{E}_\ell) + \frac{1}{n^4}, \end{aligned}$$

where the last inequality follows from the fact that for the event $\mathcal{E}_1 \cap (\mathcal{E}'_1)^c$ to occur there have to be two vertices at distance at least $\gamma \log n$ connected by a path in B ; by the EDC property this only occurs with probability at most n^{-4} for large γ . Iterating this procedure for $\mathcal{E}_5, \mathcal{E}_9, \dots$, we get

$$(22) \quad \pi^{(1,\eta)}(\mathcal{E}_1, \mathcal{E}_5, \mathcal{E}_9, \dots, \mathcal{E}_\ell) \leq \pi^{(1,\eta)}(\mathcal{E}'_1, \mathcal{E}'_5, \mathcal{E}'_9, \dots, \mathcal{E}'_\ell) + \frac{\ell}{n^4}.$$

Let $Q = \bigcup_{i=0}^{(\ell-1)/4} Q_{4i+1}$. Monotonicity implies that

$$\begin{aligned} \pi^{(1,\eta)}(\mathcal{E}'_1, \mathcal{E}'_5, \mathcal{E}'_9, \dots, \mathcal{E}'_\ell) &\leq \pi^1(\mathcal{E}'_1, \mathcal{E}'_5, \mathcal{E}'_9, \dots, \mathcal{E}'_\ell \mid E^c(Q) = 1) \\ &= \prod_{i=0}^{\frac{\ell-1}{4}} \pi^1(\mathcal{E}'_{4i+1} \mid E^c(Q) = 1), \end{aligned}$$

where for the last equality we use that the events $\mathcal{E}'_1, \mathcal{E}'_5, \dots, \mathcal{E}'_\ell$ are independent under the wired boundary condition when also conditioning on $\{E^c(Q) = 1\}$. We claim that there exists a constant $\rho \in (0, 1)$ (independent of n) such that for all $i = 0, \dots, \frac{\ell-1}{4}$,

$$(23) \quad \pi^1(\mathcal{E}'_{4i+1} \mid E^c(Q) = 1) \leq 1 - \rho.$$

To see this, let $j = 4i + 1$ and note that for large enough $D > 0$, by the EDC property and a union bound imply that there is no connection between any pair of vertices (u, v) with $u \in S_j$ and $v \in S_{j+1}$ and $d(u, v) \geq D$ with probability $\Omega(1)$. At the same time, by forcing an adjacent $O(D)$ edges to be closed at a cost of $e^{-O(D)}$, we see that with $\Omega(1)$ probability, in fact no other pairs (u, v) with $d(u, v) \leq D$ are connected either. Thus, (23) holds for a suitable $\rho \in (0, 1)$ and so

$$(24) \quad \pi^{(1,\eta)}(\mathcal{E}'_1, \mathcal{E}'_5, \mathcal{E}'_9, \dots, \mathcal{E}'_\ell) \leq (1 - \rho)^{\frac{\ell+3}{4}} \leq (1 - \rho)^{\frac{\ell-1}{4}} \leq \frac{1}{9n^3},$$

where the last inequality holds for sufficiently large c_0 since $t \geq \frac{c_0}{8(\gamma+\alpha)} \log n$.

Putting (24), (22), (20), (21) and (19) together, we get

$$\pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial_w B) \leq \frac{2}{9n^3} + \frac{\ell}{n^4} \leq \frac{1}{3n^3},$$

since $\ell = O(\log n)$. Analogously, we get $\pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial_E B) \leq \frac{1}{3n^3}$, and thus

$$\pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial B) \leq \frac{1}{n^3}.$$

Finally for the third case, suppose without loss of generality that B intersects $\partial_N \Lambda_n$ and $\partial_W \Lambda_n$, but not $\partial_E \Lambda_n$ or $\partial_S \Lambda_n$. A union bound implies that

$$(25) \quad \pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial B \setminus \partial \Lambda) \leq \pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial_S B) + \pi^{(1,\eta)}(\{e\} \xleftrightarrow{\eta} \partial_E B),$$

and each term in the right-hand side of (25) can be bounded in the same way as $\pi^{(1,\xi)}(\{e\} \xleftrightarrow{\eta} \partial_w B)$ in the second case; thus, the result follows. \square

7. Slow mixing under worst-case boundary conditions. In this section, we show that there are (nonrealizable) boundary conditions for the graph $(\Lambda_n, E(\Lambda_n))$ for which the FK-dynamics requires exponentially many steps to converge to stationarity. In particular, we prove Theorem 1.5 from the Introduction.

Theorem 1.5 is a corollary of a more general theorem we establish. This general theorem enables the transferring of mixing time lower bounds for the FK-dynamics on arbitrary graphs to mixing time lower bounds for the FK-dynamics on Λ_n , for suitably chosen boundary conditions. The high level idea is that any graph G with fewer than $\frac{n}{4}$ edges can be “embedded” into a subset L of the boundary $\partial_N \Lambda_n$ of Λ_n as a boundary condition we shall denote $\xi(G)$. When p is sufficiently small, the effect of the configuration on $\Lambda_n \setminus L$ becomes negligible, and so the mixing time of the FK-dynamics on Λ_n with boundary condition $\xi(G)$ is primarily dictated by its restriction to the embedded graph G .

We show first how to embed any graph $G = (V_G, E_G)$ into a subset $L \subset \partial_N \Lambda_n$. For $m \leq \lfloor n/4 \rfloor$, let

$$L = L(m) = \{ \llbracket 4i, 4i + 1 \rrbracket : i = 0, \dots, m - 1 \} \times \{n\} \subset \partial_N \Lambda_n$$

with edge set $E(L)$ consisting of all edges in $E(\Lambda_n)$ connecting vertices in L .

DEFINITION 7.1. Let $G = (V_G, E_G)$ be a graph with $|E_G| = m$ for $m \leq \lfloor n/4 \rfloor$ and let L be as above. We say a function $\phi : L \rightarrow V_G$ is an embedding of G into $(L, E(L))$ if for every $\{u, v\} \in E_G$ there exists a unique pair $x \in \phi^{-1}(u) \subseteq L$ and $y \in \phi^{-1}(v) \subseteq L$, where $\phi^{-1}(u)$ and $\phi^{-1}(v)$ denotes the preimage sets for u and v , respectively, such that $\{x, y\} \in E(L)$.

Notice that every graph G on $m \leq \lfloor n/4 \rfloor$ edges can be embedded into L by identifying each edge in E_G with an edge in $E(L)$.

FACT 7.2. For every graph $G = (V_G, E_G)$ with $m \leq \lfloor \frac{n}{4} \rfloor$ edges, there exists an embedding of G into $(L, E(L))$.

Now let $\xi(G)$ be the boundary condition on $\partial \Lambda_n$ defined by the partition:

$$\{ \{v\} : v \in \partial \Lambda_n \setminus L \} \cup \{ \phi^{-1}(v) : v \in V_G \}.$$

In words, $\xi(G)$ is the boundary condition that is free everywhere except in the vertices of L and where all the vertices in L that are mapped by ϕ to the same vertex of G are wired in $\xi(G)$. We are now ready to state our main comparison result from which Theorem 1.5 follows straightforwardly.

THEOREM 7.3. Let $G = (V_G, E_G)$ be a graph and suppose there exist $q > 2$ and $p = \lambda |E_G|^{-\alpha}$ with $\lambda > 0$, $\alpha > 1/3$ such that $\text{gap}(G) \leq \exp(-\Omega(|V_G|))$. Then, as long as $n \geq 4|E_G| \geq \varepsilon n$ for some $\varepsilon > 0$, with the same choice of p and q , we have

$$\text{gap}(\Lambda_n^{\xi(G)}) \leq e^{-\Omega(|V_G|)}.$$

PROOF OF THEOREM 1.5. It was established in [19] that for every $q > 2$ and every ℓ sufficiently large, there exists an interval $(\lambda_s(q), \lambda_S(q))$ such that if $p' = \lambda'/\ell$ with $\lambda' \in (\lambda_s(q), \lambda_S(q))$, then the spectral gap at parameters (p', q) satisfies

$$\text{gap}(K_\ell) \leq \exp(-\Omega(\ell)),$$

where K_ℓ denotes the complete graph on ℓ vertices. Therefore, for a fixed $p = \lambda n^{-\alpha}$ there exists a choice of $\ell = \Theta(n^\alpha)$ such that at parameters (p, q) , $\text{gap}(K_\ell) \leq \exp(-\Omega(\ell))$. Since the number of edges in K_ℓ is $\Theta(n^{2\alpha})$, the result follows from Theorem 7.3 and (3). \square

7.1. *Main comparison inequality: Proof of Theorem 7.3.* We now turn to the proof of Theorem 7.3. A standard tool for bounding spectral gaps is construction of bottleneck sets with small conductance; see Section 2.1. It will be easier to do so for the following *modified heat-bath* (MHB) dynamics, allowing us to isolate moves on $E(L)$, where we have embedded G , from those in $E^c(L) = E(\Lambda_n) \setminus E(L)$.

DEFINITION 7.4. Given a FK configuration X_t , one step of the MHB chain is given by:

1. Pick $e \in E(\Lambda_n)$ uniformly at random;
2. If both endpoints of e lie in L , then perform a heat-bath update on e , that is, replace the configuration in e with a sample from $\pi_{\Lambda_n, p, q}^\xi(\cdot \mid X_t(E(\Lambda_n) \setminus \{e\}))$;
3. Otherwise, replace the configuration in $E(\Lambda_n) \setminus E(L)$ with a sample from $\pi_{\Lambda_n, p, q}^\xi(\cdot \mid X_t(E(L)))$.

The MHB chain is clearly reversible with respect to $\pi_{\Lambda_n, p, q}^\xi$.

Let $\text{gap}_{\text{MHB}}(\Lambda_n^\xi)$ denote the spectral gap of the MHB dynamics on Λ_n with boundary condition ξ and parameters p and q . The following comparison inequality allows us to focus on finding upper bounds for the spectral gap of the MHB dynamics; its proof is deferred to Section 7.2.

LEMMA 7.5. For all $p \in (0, 1)$, $q > 0$, $n \in \mathbb{N}$ and boundary condition ξ for Λ_n , we have

$$\text{gap}(\Lambda_n^\xi) \leq \text{gap}_{\text{MHB}}(\Lambda_n^\xi).$$

With this in hand, we are now ready to prove Theorem 7.3.

PROOF OF THEOREM 7.3. Recall from (5) that since, by assumption, the FK-dynamics on G has $\text{gap}(G) \leq \exp(-\Omega(|V_G|))$, there must exist $S_\star \subset \Omega_G$ (the set of FK configurations on G) with $\pi_G(S_\star) \leq \frac{1}{2}$ such that

$$(26) \quad \Phi(S_\star) = \frac{Q_G(S_\star, S_\star^c)}{\pi_G(S_\star)} \leq e^{-\Omega(|V_G|)}.$$

Here, Q_G is the edge measure (4) of the FK-dynamics on G and $\pi_G = \pi_{G, p, q}$ denotes the random-cluster measure on G . We will construct from this set S_\star , a set $A_\star \subset \Omega$, such that

$$(27) \quad \begin{aligned} \Phi(A_\star) &= \frac{Q_{\text{MHB}}(A_\star, A_\star^c)}{\pi^{\xi(G)}(A_\star)} \leq e^{-\Omega(|V_G|)} \quad \text{and} \\ \Phi(A_\star^c) &= \frac{Q_{\text{MHB}}(A_\star, A_\star^c)}{\pi^{\xi(G)}(A_\star^c)} \leq e^{-\Omega(|V_G|)}, \end{aligned}$$

where Q_{MHB} denotes the edge measure (4) of the MHB dynamics on $\Lambda_n^{\xi(G)}$ and $\pi^{\xi(G)} = \pi_{\Lambda_n, p, q}^{\xi(G)}$. This implies Theorem 7.3 by combining it with (3) and (5).

Let $\{\xi_1, \dots, \xi_k\}$ be the partition of L induced by $\xi(G)$. For a FK configuration ω on $E^c(L)$, we say that $\xi_i \xleftrightarrow{\omega} \xi_j$ if there is an open path in ω from a vertex in ξ_i to a vertex in ξ_j . Let $\mathcal{S}^{\xi(G)}(\omega)$ be the set

$$\mathcal{S}^{\xi(G)}(\omega) = \{\xi_i \in \xi(G) : \xi_i \xleftrightarrow{\omega} \xi_j \text{ for some } j \neq i, j \in \{1, \dots, k\}\},$$

that is, those ξ_i that are connected to some ξ_j in ω . For $M \geq 0$, let

$$\mathcal{R}^{\xi(G)}(M) = \{\omega \in \{0, 1\}^{E^c(L)} : |\mathcal{S}^{\xi(G)}(\omega)| \leq M\}.$$

In words, $\mathcal{R}^{\xi(G)}(M)$ is the set of FK configurations on $E(\Lambda_n) \setminus E(L)$ that connect at most M elements of the partition $\{\xi_1, \dots, \xi_k\}$ of the vertex set L .

Observe that any configuration θ on $E(L)$ corresponds to a configuration on E_G . Namely, if $\theta(\{u, v\}) = 1$, then the edge $\{\phi(u), \phi(v)\}$ is open in the configuration on G , where ϕ is the embedding of G into L . With a slight abuse of notation, we may use θ also for the corresponding configuration on Ω_G . With this convention, let

$$(28) \quad A_M = \{\omega \in \Omega : \omega(E(L)) \in S_\star, \omega(E^c(L)) \in \mathcal{R}^{\xi(G)}(M)\}.$$

We show that if $M = \delta|V_G|$ for some $\delta > 0$ sufficiently small, and n is taken to be large enough, we can take $A_\star = A_M$. This will follow from the following two claims.

CLAIM 7.6. *The following are true of S_\star and A_M defined above:*

- (i) $\pi^{\xi(G)}(A_M) \geq q^{-M}(1 - e^{-\Omega(M)})\pi_G(S_\star)$;
- (ii) $\pi^{\xi(G)}(A_M^c) \leq e^{-O(M)}$.

CLAIM 7.7. *The modified heat-bath dynamics satisfies*

$$Q_{\text{MHB}}(A_M, A_M^c) \leq \pi^{\xi(G)}(A_M)e^{-\Omega(M \log n)} + \frac{q^{2M+1}}{p} Q_G(S_\star, S_\star^c).$$

Dividing the bound from Claim 7.7 by $\pi^{\xi(G)}(A_M)$ and using the bounds from Claim 7.6, we see that

$$\begin{aligned} \frac{Q_{\text{MHB}}(A_M, A_M^c)}{\pi^{\xi(G)}(A_M)} &\leq e^{-\Omega(M \log n)} + \frac{2q^{3M+1}}{p} \frac{Q_G(S_\star, S_\star^c)}{\pi_G(S_\star)} \\ &\leq e^{-\Omega(M \log n)} + e^{O(M)}e^{-\Omega(|V_G|)} \end{aligned}$$

for sufficiently large M , where the last inequality follows from (26) and the facts that $M = \delta|V_G|$ and $p \geq \lambda|V_G|^{-2\alpha}$. Similarly, we get

$$\frac{Q_{\text{MHB}}(A_M, A_M^c)}{\pi^{\xi(G)}(A_M^c)} \leq e^{-\Omega(M \log n)} + e^{O(M)}e^{-\Omega(|V_G|)}.$$

Then, since $M = \delta|V_G|$, for some $\delta > 0$ sufficiently small we obtain (27). \square

7.2. *Proof of auxiliary facts.* In this section, we provide the proofs of Lemma 7.5 and Claims 7.6 and 7.7.

PROOF OF LEMMA 7.5. For any $B \subset E(\Lambda_n)$, let P_B be the transition matrix corresponding to a heat-bath update on the entire set B . For $e \in E(\Lambda_n)$, we use P_e for $P_{\{e\}}$. Let P_{FK} and P_{MHB} be the transition matrices for the FK-dynamics and the MHB dynamics on Λ_n^ξ , respectively. Let $A = E(\Lambda_n) \setminus E(L)$. Then

$$P_{\text{MHB}} = \frac{1}{|E(\Lambda_n)|} \left(\sum_{e \in E(L)} P_e + \sum_{e \in A} P_A \right).$$

For ease of notation, set $\pi = \pi_{\Lambda_n, p, q}^\xi$. Then, for any $f, g \in \mathbb{R}^{|\Omega|}$, where Ω denotes the set of FK configurations on Λ_n , let

$$\langle f, g \rangle_\pi = \sum_{\omega \in \Omega} f(\omega)g(\omega)\pi(\omega).$$

If we endow $\mathbb{R}^{|\Omega|}$ with the inner product $\langle \cdot, \cdot \rangle_\pi$, we obtain a Hilbert space denoted $L_2(\pi) = (\mathbb{R}^{|\Omega|}, \langle \cdot, \cdot \rangle_\pi)$. Recall that if a matrix P is reversible with respect to π , it defines a *self-adjoint* operator from $L_2(\pi)$ to $L_2(\pi)$ via matrix vector multiplication, and thus $\langle f, Pg \rangle_\pi = \langle P^*f, g \rangle_\pi = \langle Pf, g \rangle_\pi$.

Now, the matrix P_{MHB} is positive semidefinite, since it is an average of positive semidefinite matrices. Hence, it is a standard fact (see, e.g., [24]) that if $\langle f, P_{\text{MHB}}f \rangle_\pi \leq \langle f, P_{\text{FK}}f \rangle_\pi$ for all $f \in \mathbb{R}^{|\Omega|}$, then $\text{gap}_{\text{MHB}}(\Lambda_n^\xi) \geq \text{gap}(\Lambda_n^\xi)$. To show this, note that for $e \in A$, we have $P_A = P_e P_A P_e$. Thus, for all $f \in \mathbb{R}^{|\Omega|}$,

$$\begin{aligned} \langle f, P_A f \rangle_\pi &= \langle f, P_e P_A P_e f \rangle_\pi = \langle P_e^* f, P_A P_e f \rangle_\pi = \langle P_e f, P_A P_e f \rangle_\pi \\ &\leq \langle P_e f, P_e f \rangle_\pi = \langle f, P_e^2 f \rangle_\pi = \langle f, P_e f \rangle_\pi, \end{aligned}$$

where we used that $P_e = P_e^*$, $P_e = P_e^2$ and that $\langle f, Pf \rangle_\pi \leq \langle f, f \rangle_\pi$ for every $f \in \mathbb{R}^{|\Omega|}$ and every matrix P reversible with respect to π . Then

$$\langle f, P_{\text{MHB}}f \rangle_\pi \leq \frac{1}{|E(\Lambda_n)|} \sum_{e \in E(\Lambda_n)} \langle f, P_e f \rangle_\pi = \langle f, P_{\text{FK}}f \rangle_\pi,$$

and the result follows. \square

Recall the notation of Lemma 7.3. In order to compare the marginal distribution in L to π_G , we need to bound the number of connections in $E^c(L)$ between different boundary components of $\xi(G)$ restricted to L : this will show that typical FK configurations on $E^c(L)$ do not have much influence on the connectivities among L . This bound follows from the fact that $p = O(n^{-\alpha})$ for some $\alpha > 1/3$ and the (approximate) independence of connections between ξ_i and ξ_j . For the remainder of this section, we set $\xi = \xi(G)$ for ease of notation. Claims 7.6–7.7 will then be seen as consequences of the following lemma.

LEMMA 7.8. *Let $q \geq 1$ and let $p = \lambda n^{-\alpha}$ for $\lambda > 0$ and $\alpha > 1/3$. Let ξ be any boundary condition on $\partial_N \Lambda_n$ and let η be any FK configuration on $E(L)$. For every $M \geq 1$,*

$$\pi_{\Lambda_n}^\xi(\mathcal{R}^\xi(M) \mid \eta) \geq 1 - \exp[-\Omega(M \log[Mn^{3\alpha-1}])].$$

PROOF. Let Y be the random variable for the number of vertices of L connected to at least one other vertex of L in an FK configuration on $E^c(L)$ sampled from the distribution $\pi^{\xi, \eta}(\cdot) = \pi_{\Lambda_n, p, q}^\xi(\cdot \mid \eta)$. It is sufficient to show that

$$\pi^{\xi, \eta}(Y \geq M) \leq \exp(-\Omega(M \log[Mn^{3\alpha-1}])).$$

By classical comparison inequalities (see, e.g., [21]), when $q \geq 1$, no matter the boundary conditions ξ, η , the random-cluster measure on $E^c(L)$ is stochastically dominated by the independent bond percolation distribution on $E^c(L)$ with the same parameter p , which we denote by $\nu = \nu_{\Lambda_n, p}$. (Recall that ν is the distribution on Ω that results from adding every edge in $E(\Lambda_n)$ independently with probability p .) Hence, if X is defined as Y but for ν on $E^c(L)$, we get

$$\pi^{\xi, \eta}(Y \geq M) \leq \nu(X \geq M).$$

Consider the subgraph $\hat{\Lambda} = (\Lambda_n, E^c(L))$. Let Z be the total number of vertices in L that are connected in a configuration sampled from ν to another vertex at distance 3 in $\hat{\Lambda}$. Then, since the distance between any two vertices in L in $\hat{\Lambda}$ is at least 3, we see that $X \leq Z$ and

$$\nu(X \geq M) \leq \nu(Z \geq M).$$

Thus, it suffices to establish a tail bound for Z . Enumerate the vertices of L as v_1, \dots, v_{2m} and let Z_j be the indicator random variable for the event that v_j is connected to another vertex at distance 3 in $\hat{\Lambda}$. For $r = 0, 1, 2, 3$, split up $Z = \sum_r \hat{Z}_r$, where

$$\hat{Z}_r = \sum_{i \geq 0} Z_{4i+r+1}.$$

We claim that there is some suitable $c > 0$ such that under ν , each \hat{Z}_r is stochastically dominated by the binomial random variable $S \sim \text{Bin}(\lceil n/4 \rceil, cn^{-3\alpha})$. This is because the random variables Z_{4i+r+1} are jointly dominated by independent Bernoulli random variables, $\text{Ber}(cn^{-3\alpha})$, for a suitable $c > 0$, since for every k the events $\{Z_k = 1\}$ and $\{Z_{k+4} = 1\}$ depend on disjoint sets of edges. The fact that the success probability of each one is at most $cn^{-3\alpha}$ follows from the fact that there are at most 16 choices of three adjacent edges from a vertex in $\partial_N \Lambda_n$, and $p = \lambda n^{-\alpha}$. Hence, by the Chernoff–Hoeffding inequality, for every $\delta > 0$,

$$\nu(\hat{Z}_r \geq \mathbb{E}_\nu[S] + \delta \lceil n/4 \rceil) \leq \exp\left[-\frac{n}{4} D(cn^{-3\alpha} + \delta \lceil n/4 \rceil, cn^{-3\alpha})\right],$$

where $D(a\|b)$ is the relative entropy between the Bernoulli random variables $\text{Ber}(a)$ and $\text{Ber}(b)$:

$$D(a\|b) = a \log(a/b) + (1 - a) \log((1 - a)/(1 - b)).$$

Since $\mathbb{E}_\nu[S] = O(n^{1-3\alpha})$, $\text{Var}(S) = O(n^{1-3\alpha})$, $\alpha > \frac{1}{3}$ and $M \geq 1$, it follows that for every $M \geq 1$,

$$\nu(\hat{Z}_r \geq M/4) \leq e^{-\Omega(M \log[Mn^{3\alpha-1}])}.$$

We have $Z = \hat{Z}_1 + \hat{Z}_2 + \hat{Z}_3 + \hat{Z}_4$, and so a union bound implies the matching bound for $\nu(Z \geq M)$. \square

Now recall the definitions of the set S_\star and A_M from (28).

PROOF OF CLAIM 7.6. For part (i), observe that if ω is sampled from π^ξ , then

$$\pi^\xi(A_M) = \pi^\xi(\omega(L) \in S_\star \mid \omega(E^c(L)) \in \mathcal{R}^\xi(M)) \pi^\xi(\omega(E^c(L)) \in \mathcal{R}^\xi(M)).$$

By Lemma 7.8,

$$\pi^\xi(\omega(E^c(L)) \in \mathcal{R}^\xi(M)) \geq 1 - e^{-\Omega(M)}.$$

Moreover, since

$$\pi^\xi(\omega(L) \in S_\star \mid \omega(E^c(L) = 0)) = \pi_G(S_\star),$$

it follows from Lemma 2.2 that

$$\pi^\xi(\omega(L) \in S_\star \mid \omega(E^c(L)) \in \mathcal{R}^\xi(M)) \geq q^{-M} \pi_G(S_\star),$$

and thus,

$$\pi^\xi(A_M) \geq q^{-M} (1 - e^{-\Omega(M)}) \pi_G(S_\star).$$

Similarly for part (ii), we have

$$\begin{aligned} \pi^\xi(A_M^c) &\geq \pi^\xi(\omega(L) \notin S_\star \mid \omega(E^c(L)) \in \mathcal{R}^\xi(M)) \pi^\xi(\omega(E^c(L)) \in \mathcal{R}^\xi(M)) \\ &\geq q^{-M} (1 - e^{-\Omega(M)}) \pi_G(S_\star^c) \end{aligned}$$

which is at least $e^{-O(M)}$ since $\pi_G(S_\star) \leq \frac{1}{2}$. \square

PROOF OF CLAIM 7.7. Let P_{MHB} be the transition matrix for the MHB dynamics and for ease of notation set $B = E^c(L)$. We have

$$(29) \quad \begin{aligned} Q_{\text{MHB}}(A_M, A_M^c) &\leq \sum_{\omega \in A_M} \sum_{\substack{\omega' \in \Omega: \\ \omega'(B) \notin \mathcal{R}^\xi(M)}} \pi^\xi(\omega) P_{\text{MHB}}(\omega, \omega') \\ &\quad + \sum_{\omega \in A_M} \sum_{\substack{\omega' \in \Omega: \\ \omega'(L) \notin S_\star}} \pi^\xi(\omega) P_{\text{MHB}}(\omega, \omega'). \end{aligned}$$

For the first term in (29), observe by definition of MHB dynamics, for every $\omega \in A_M$,

$$\sum_{\substack{\omega' \in \Omega: \\ \omega'(B) \notin \mathcal{R}^\xi(M)}} P_{\text{MHB}}(\omega, \omega') \leq \sum_{\substack{\omega' \in \Omega: \\ \omega'(B) \notin \mathcal{R}^\xi(M)}} \pi^\xi(\omega'(B) \mid \omega(L)) \leq e^{-\Omega(M \log n)},$$

where the last inequality follows from Lemma 7.8. Hence,

$$\sum_{\omega \in A_M} \sum_{\substack{\omega' \in \Omega: \\ \omega'(B) \notin \mathcal{R}^\xi(M)}} \pi^\xi(\omega) P_{\text{MHB}}(\omega, \omega') \leq \pi^\xi(A_M) e^{-\Omega(M \log n)}.$$

For the second term in (29), observe that $\omega \neq \omega'$ and that ω and ω' can differ in at most one edge e ; otherwise $P_{\text{MHB}}(\omega, \omega') = 0$. Thus, setting

$$\begin{aligned} p^+(q) &= \max \left\{ p, \frac{q(1-p)}{q(1-p) + p} \right\}, \\ p^-(q) &= \min \left\{ 1-p, \frac{p}{q(1-p) + p} \right\}, \end{aligned}$$

we obtain

$$\begin{aligned} P_{\text{MHB}}(\omega, \omega') &= \frac{1}{|E(\Lambda_n)|} \pi(\omega'(e) \mid \omega(E(\Lambda_n) \setminus \{e\})) \\ &\leq \frac{p^+(q)}{|E(\Lambda_n)|} \\ &\leq \frac{p^+(q) |E_G|}{p^-(q) |E(\Lambda_n)|} P_G(\omega(L), \omega'(L)). \end{aligned}$$

Then, since $|E_G| \leq |E(\Lambda_n)|$ and $p^+(q)/p^-(q) \leq q/p$,

$$\begin{aligned} &\sum_{\omega \in A_M} \sum_{\substack{\omega' \in \Omega: \\ \omega'(L) \notin S_\star}} \pi^\xi(\omega) P_{\text{MHB}}(\omega, \omega') \\ &\leq \frac{q}{p} \sum_{\omega \in A_M} \sum_{\substack{\omega' \in \Omega: \\ \omega'(L) \notin S_\star}} \pi^\xi(\omega) P_G(\omega(L), \omega'(L)) \\ &\leq \frac{q}{p} \sum_{\theta \in \mathcal{R}^\xi(M)} \pi^\xi(\theta) \sum_{\omega_1 \in S_\star} \sum_{\omega_2 \notin S_\star} \pi^\xi(\omega_1 \mid \theta) P_G(\omega_1, \omega_2) \\ &\leq \frac{q}{p} \pi^\xi(\mathcal{R}^\xi(M)) \sum_{\omega_1 \in S_\star} \sum_{\omega_2 \notin S_\star} \max_{\theta \in \mathcal{R}^\xi(M)} \pi^\xi(\omega_1 \mid \theta) P_G(\omega_1, \omega_2), \end{aligned}$$

where ω_1, ω_2 are FK configurations on $E(L)$ and θ is an FK configuration on B . Lemma 2.2 implies

$$\max_{\theta \in \mathcal{R}^\xi(M)} \pi^\xi(\omega_1 | \theta) \leq q^{2M} \pi_G(\omega_1),$$

and so

$$\begin{aligned} \sum_{\omega \in A_M} \sum_{\substack{\omega' \in \Omega: \\ \omega'(L) \notin S_\star}} \pi^\xi(\omega) P_{\text{MHB}}(\omega, \omega') &\leq \frac{q^{2M+1}}{p} \sum_{\omega_1 \in S_\star} \sum_{\omega_2 \notin S_\star} \pi_G(\omega_1) P_G(\omega_1, \omega_2) \\ &= \frac{q^{2M+1}}{p} Q_G(S_\star, S_\star^c). \end{aligned}$$

Combining these two bounds, we get

$$Q_{\text{MHB}}(A_M, A_M^c) \leq \pi^\xi(A_M) e^{-\Omega(M \log n)} + \frac{q^{2M+1}}{p} Q_G(S_\star, S_\star^c). \quad \square$$

Acknowledgments. The authors thank the anonymous referee for thorough and helpful suggestions. The authors also thank Insuk Seo for detailed comments.

The first and third authors were supported in part by NSF Grants CCF-1617306 and CCF-1563838.

REFERENCES

- [1] ALEXANDER, K. S. (2004). Mixing properties and exponential decay for lattice systems in finite volumes. *Ann. Probab.* **32** 441–487. MR2040789 <https://doi.org/10.1214/aop/1078415842>
- [2] BEFFARA, V. and DUMINIL-COPIN, H. (2012). The self-dual point of the two-dimensional random-cluster model is critical for $q \geq 1$. *Probab. Theory Related Fields* **153** 511–542. MR2948685 <https://doi.org/10.1007/s00440-011-0353-8>
- [3] BLANCA, A. and SINCLAIR, A. (2015). Dynamics for the mean-field random-cluster model. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques. LIPIcs. Leibniz Int. Proc. Inform.* **40** 528–543. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern. MR3441983
- [4] BLANCA, A. and SINCLAIR, A. (2017). Random-cluster dynamics in \mathbb{Z}^2 . *Probab. Theory Related Fields* **168** 821–847. MR3663632 <https://doi.org/10.1007/s00440-016-0725-1>
- [5] BORGS, C., CHAYES, J. T., FRIEZE, A., KIM, J. H., TETALI, P., VIGODA, E. and VU, V. H. (1999). Torpid mixing of some Monte Carlo Markov chain algorithms in statistical physics. In *40th Annual Symposium on Foundations of Computer Science (New York, 1999)* 218–229. IEEE Computer Soc., Los Alamitos, CA. MR1917562 <https://doi.org/10.1109/SFFCS.1999.814594>
- [6] BORGS, C., CHAYES, J. T. and TETALI, P. (2012). Tight bounds for mixing of the Swendsen–Wang algorithm at the Potts transition point. *Probab. Theory Related Fields* **152** 509–557. MR2892955 <https://doi.org/10.1007/s00440-010-0329-0>
- [7] CESI, F. (2001). Quasi-factorization of the entropy and logarithmic Sobolev inequalities for Gibbs random fields. *Probab. Theory Related Fields* **120** 569–584. MR1853483 <https://doi.org/10.1007/PL00008792>
- [8] CHAYES, L. and MACHTA, J. (1998). Graphical representations and cluster algorithms II. *Phys. A* **254** 477–516.
- [9] COOPER, C. and FRIEZE, A. M. (1999). Mixing properties of the Swendsen–Wang process on classes of graphs. *Random Structures Algorithms* **15** 242–261. MR1716764 [https://doi.org/10.1002/\(SICI\)1098-2418\(199910/12\)15:3/4<242::AID-RSA4>3.3.CO;2-3](https://doi.org/10.1002/(SICI)1098-2418(199910/12)15:3/4<242::AID-RSA4>3.3.CO;2-3)
- [10] DUMINIL-COPIN, H. and MANOLESCU, I. (2016). The phase transitions of the planar random-cluster and Potts models with $q \geq 1$ are sharp. *Probab. Theory Related Fields* **164** 865–892. MR3477782 <https://doi.org/10.1007/s00440-015-0621-0>
- [11] DUMINIL-COPIN, H., SIDORAVICIUS, V. and TASSION, V. (2017). Continuity of the phase transition for planar random-cluster and Potts models with $1 \leq q \leq 4$. *Comm. Math. Phys.* **349** 47–107. MR3592746 <https://doi.org/10.1007/s00220-016-2759-8>
- [12] EDWARDS, R. G. and SOKAL, A. D. (1988). Generalization of the Fortuin–Kasteleyn–Swendsen–Wang representation and Monte Carlo algorithm. *Phys. Rev. D* **38** 2009–2012. MR0965465 <https://doi.org/10.1103/PhysRevD.38.2009>

- [13] FORTUIN, C. M. and KASTELEYN, P. W. (1972). On the random-cluster model. I. Introduction and relation to other models. *Physica* **57** 536–564. [MR0359655](#)
- [14] GALANIS, A., ŠTEFANKOVIČ, D. and VIGODA, E. (2015). Swendsen–Wang algorithm on the mean-field Potts model. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques. LIPIcs. Leibniz Int. Proc. Inform.* **40** 815–828. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern. [MR3441999](#)
- [15] GANGULY, S. and SEO, I. (2018). Information percolation and cutoff for the random-cluster model. Preprint. Available at [arXiv:1812.01538](#).
- [16] GHEISSARI, R. and LUBETZKY, E. (2018). Mixing times of critical two-dimensional Potts models. *Comm. Pure Appl. Math.* **71** 994–1046. [MR3794520](#) <https://doi.org/10.1002/cpa.21718>
- [17] GHEISSARI, R. and LUBETZKY, E. (2019). Quasi-polynomial mixing of critical two-dimensional random cluster models. *Random Structures and Algorithms*. <https://doi.org/10.1002/rsa.20868>
- [18] GHEISSARI, R. and LUBETZKY, E. (2018). The effect of boundary conditions on mixing of 2D Potts models at discontinuous phase transitions. *Electron. J. Probab.* **23** Paper No. 57, 30. [MR3814251](#) <https://doi.org/10.1214/18-ejp180>
- [19] GHEISSARI, R., LUBETZKY, E. and PERES, Y. (2018). Exponentially slow mixing in the mean-field Swendsen–Wang dynamics. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms* 1981–1988. SIAM, Philadelphia, PA. [MR3775917](#) <https://doi.org/10.1137/1.9781611975031.129>
- [20] GORE, V. K. and JERRUM, M. R. (1999). The Swendsen–Wang process does not always mix rapidly. *J. Stat. Phys.* **97** 67–86. [MR1733467](#) <https://doi.org/10.1023/A:1004610900745>
- [21] GRIMMETT, G. (2006). *The Random-Cluster Model. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **333**. Springer, Berlin. [MR2243761](#) <https://doi.org/10.1007/978-3-540-32891-9>
- [22] GUO, H. and JERRUM, M. (2018). Random cluster dynamics for the Ising model is rapidly mixing. *Ann. Appl. Probab.* **28** 1292–1313. [MR3784500](#) <https://doi.org/10.1214/17-AAP1335>
- [23] JERRUM, M. and SINCLAIR, A. (1989). Approximating the permanent. *SIAM J. Comput.* **18** 1149–1178. [MR1025467](#) <https://doi.org/10.1137/0218077>
- [24] LEVIN, D. A. and PERES, Y. (2017). *Markov Chains and Mixing Times*, 2nd ed. Amer. Math. Soc., Providence, RI. [MR3726904](#)
- [25] LUBETZKY, E., MARTINELLI, F., SLY, A. and TONINELLI, F. L. (2013). Quasi-polynomial mixing of the 2D stochastic Ising model with “plus” boundary up to criticality. *J. Eur. Math. Soc. (JEMS)* **15** 339–386. [MR3017041](#) <https://doi.org/10.4171/JEMS/363>
- [26] MARTINELLI, F. (1994). On the two-dimensional dynamical Ising model in the phase coexistence region. *J. Stat. Phys.* **76** 1179–1246. [MR1298100](#) <https://doi.org/10.1007/BF02187060>
- [27] MARTINELLI, F. (1999). Lectures on Glauber dynamics for discrete spin models. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1997). Lecture Notes in Math.* **1717** 93–191. Springer, Berlin. [MR1746301](#) https://doi.org/10.1007/978-3-540-48115-7_2
- [28] MARTINELLI, F., OLIVIERI, E. and SCHONMANN, R. H. (1994). For 2-D lattice spin systems weak mixing implies strong mixing. *Comm. Math. Phys.* **165** 33–47. [MR1298940](#)
- [29] MARTINELLI, F. and TONINELLI, F. L. (2010). On the mixing time of the 2D stochastic Ising model with “plus” boundary conditions at low temperature. *Comm. Math. Phys.* **296** 175–213. [MR2606632](#) <https://doi.org/10.1007/s00220-009-0963-5>
- [30] MOSSEL, E. and SLY, A. (2013). Exact thresholds for Ising–Gibbs samplers on general graphs. *Ann. Probab.* **41** 294–328. [MR3059200](#) <https://doi.org/10.1214/11-AOP737>
- [31] ONSAGER, L. (1944). Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Phys. Rev.* **65** 117–149. [MR0010315](#)
- [32] SALOFF-COSTE, L. (1997). Lectures on finite Markov chains. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1996). Lecture Notes in Math.* **1665** 301–413. Springer, Berlin. [MR1490046](#) <https://doi.org/10.1007/BFb0092621>
- [33] SINCLAIR, A. (1992). Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combin. Probab. Comput.* **1** 351–370. [MR1211324](#) <https://doi.org/10.1017/S0963548300000390>
- [34] SWENDSEN, R. H. and WANG, J. S. (1987). Nonuniversal critical dynamics in Monte Carlo simulations. *Phys. Rev. Lett.* **58** 86–88.
- [35] THOMAS, L. E. (1989). Bound on the mass gap for finite volume stochastic Ising models at low temperature. *Comm. Math. Phys.* **126** 1–11. [MR1027910](#)
- [36] ULLRICH, M. (2013). Comparison of Swendsen–Wang and heat-bath dynamics. *Random Structures Algorithms* **42** 520–535. [MR3068036](#) <https://doi.org/10.1002/rsa.20431>
- [37] ULLRICH, M. (2014). Swendsen–Wang is faster than single-bond dynamics. *SIAM J. Discrete Math.* **28** 37–48. [MR3148642](#) <https://doi.org/10.1137/120864003>