INFORMATION PERCOLATION AND CUTOFF FOR THE RANDOM-CLUSTER MODEL

SHIRSHENDU GANGULY AND INSUK SEO

ABSTRACT. We consider the Random-Cluster model on $(\mathbb{Z}/n\mathbb{Z})^d$ with parameters $p \in (0,1)$ and $q \geq 1$. This is a generalization of the standard bond percolation (with edges open independently with probability p) which is biased by a factor q raised to the number of connected components. We study the well known FK-dynamics on this model where the update at an edge depends on the global geometry of the system unlike the Glauber Heat-Bath dynamics for spin systems, and prove that for all small enough p (depending on the dimension) and any $q > 1$, the FK-dynamics exhibits the cutoff phenomenon at $\lambda_\infty^{-1} \log n$ with a window size $O(\log \log n)$, where λ_∞ is the large n limit of the spectral gap of the process. Our proof extends the *Information Percolation* framework of Lubetzky and Sly [23] to the Random-Cluster model and also relies on the arguments of Blanca and Sinclair [5] who proved a sharp $O(\log n)$ mixing time bound for the planar version. A key aspect of our proof is the analysis of the effect of a sequence of dependent (across time) Bernoulli percolations extracted from the graphical construction of the dynamics, on how information propagates.

CONTENTS

1. INTRODUCTION AND MAIN RESULT

The random-cluster (Fortuin-Kasteleyn/FK) model is an extensively studied model in statistical physics, generalizing electrical networks, percolation, and spin systems like the Ising and Potts models, under a single framework. In this work, we study the so called heat-bath Glauber dynamics or FK-dynamics for the model on the d-dimensional torus. The main result of this paper establishes a sharp convergence to equilibrium for this Markov chain also known as the *cutoff phenomenon*.

1.1. **Random-cluster model (RCM).** For $d \geq 2$, denote by $\Lambda_n = \mathbb{Z}_n^d$, the d-dimensional discrete torus and by $E_n = E(\Lambda_n)$, the set of edges in Λ_n. We will fix the dimension to be d throughout the entire paper. The random-cluster measure $\mu_{p,q}^n$ on the graph (Λ_n, E_n) with parameters $p \in (0,1)$ and $q > 0$ is a probability measure on the space of subsets of E_n defined by

$$\mu_{p,q}^n(S) = \frac{1}{Z_{p,q}^n} p^{|S|}(1-p)^{|E_n \setminus S|} q^{c(S)} \ ; \ S \subset E_n \ ,$$

where $Z_{p,q}^n$ is the partition function turning $\mu_{p,q}^n$ into a probability measure, and $c(S)$ is the number of connected components of the graph (Λ_n, S). Clearly the measure $\mu_{p,q}^n$ can be regarded as a

probability measure on $\Omega_n = \{0, 1\}^{E_n}$, i.e., we will identify $X = (X(e))_{e \in E_N} \in \Omega_n$ with a subset A of E_n where $e \in A$ if and only if $X(e) = 1$. Hence, by slight abuse of notation, we can always regard $X \in \Omega_n$ as a subset of E_n. The random-cluster model was introduced by Fortuin and Kasteleyn (see [12, 13]) and unifies the study of various objects in statistical mechanics such as random graphs, spin systems and electrical networks (see [17]). When $q = 1$, this model corresponds to the standard bond percolation but when $q > 1$ (resp., $q < 1$), the probability measure biases subgraphs with more (resp., fewer) connected components. For the special case of integer $q \geq 2$ the random-cluster model is a dual to the classical ferromagnetic q-state Potts model, via the so called Edward-Sokal coupling of the models (see, e.g., [11]). However, note that unlike spin systems, the probability that an edge e belongs to A does not depend only on the dispositions of its neighboring edges but on the entire configuration A, since connectivity is a global property (see Figure 1.1 for an illustration).

1.2. **FK-dynamics (Glauber/Heat-bath dynamics).** The FK-dynamics is a reversible Markov process $X_t = \{X_t(e)\}_{e \in E_n}$ on Ω_n, whose invariant measure is given by $\mu_{p,q}^n$. Informally, at rate one, the state of every edge $X(e)$ is resampled conditionally on the state of the remaining edges i.e.,

$$X(e) = \begin{cases} 1 & \text{w.p. } p \text{ if } e \text{ is not a cut-edge,} \\ 1 & \text{w.p. } \frac{p}{p+(1-p)q} \text{ if } e \text{ is a cut-edge,} \\ 0 & \text{otherwise,} \end{cases}$$

where we use the standard terminology *cut-edge* to denote an edge whose removal increases the number of connected components by one. A more formal treatment appears in Definition 2.1.

Note that unlike Glauber dynamics on spin systems like Ising or Potts models, the FK-dynamics has long range dependencies (see Figure 1.1). The FK-dynamics has been an object of significant interest and has played a key role in several recent works. A by no means complete, but nonetheless representative list includes [].

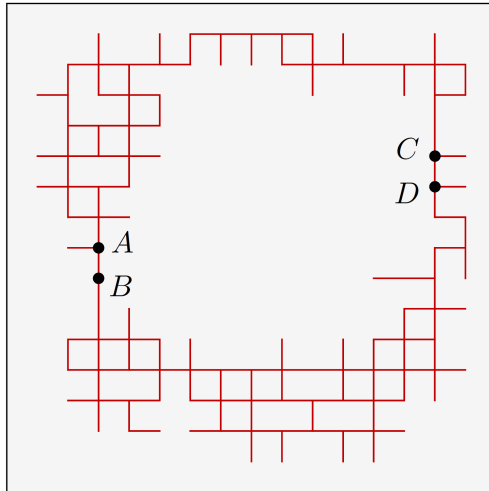The key statistic we will consider is the time taken by the above dynamics to converge to equilibrium.

FIGURE 1.1. Illustrating long range dependencies in FK-dynamics. Consider a configuration where the red edges are open while everything else is closed. The probability of the edge AB to be open then depends on whether the edge CD is open or not.

1.3. **Mixing of Markov chains and cutoff phenomenon.** We review in brief the set up of interest for us from the theory of reversible Markov chains with finite state spaces. For an extensive account of all the details and recent progress in various directions see [19]. For two probability measures μ_1 and μ_2 on S we will be interested in the L^1-distance or the so-called *total variation distance* between them denoted by $\|\mu_1 - \mu_2\|_{\mathrm{TV}}$[1]:

$$\sup_{A \subset S} (\mu_1(A) - \mu_2(A)) = \frac{1}{2} \sum_{x \in S} |\mu_1(x) - \mu_2(x)| = \frac{1}{2} \sum_{x \in S} \left| \frac{\mu_1(x)}{\mu_2(x)} - 1 \right| \mu_2(x) . \qquad (1.1)$$

For concreteness consider a continuous time reversible Markov chain Y_t with a finite state space S and equilibrium measure π.

Denote by $\mathbb{P}_x$, $x \in S$, the law of Markov chain starting from x.

We will be primarily interested in the total variation mixing time defined by

$$t_{\mathrm{mix}}(\varepsilon) = \inf \left\{ t : \sup_{y \in S} \|\mathbb{P}_y[Y_t \in \cdot] - \pi\|_{\mathrm{TV}} \leq \varepsilon \right\} \; ; \; \varepsilon \in (0, 1) .$$

For notational brevity, we will denote by $d(t)$, the worst case total variation distance to stationarity for the FK-dynamics, i.e.,

$$d(t) = d_n(t) := \sup_{x \in \Omega_n} \|\mathbb{P}_x[X_t \in \cdot] - \mu_{p,q}^n\|_{\mathrm{TV}} \qquad (1.2)$$

from now on. Many naturally occurring Markov chains are expected to exhibit a sharp transition in convergence, in the sense that the total variation distance to equilibrium drops from one to zero in a rather short time window. This is formalized by the notion of cutoff formulated by Aldous and Diaconis [1] (see also [6]). Formally a sequence of Markov chain $Y_t^{(1)}, Y_t^{(2)}, \ldots$ with mixing times given by $t_{\mathrm{mix}}^{(1)}(\varepsilon), t_{\mathrm{mix}}^{(2)}(\varepsilon), \ldots$ is said to exhibit the **_Cutoff Phenomenon_** if for any $\varepsilon \leq 1/2$,

$$\lim_{i \to \infty} \frac{t_{\mathrm{mix}}^{(i)}(\varepsilon)}{t_{\mathrm{mix}}^{(i)}(1 - \varepsilon)} = 1 .$$

Moreover cutoff is said to occur with window size w_i if for any $\varepsilon \leq 1/2$ one has

$$t_{\mathrm{mix}}^{(i)}(\varepsilon) - t_{\mathrm{mix}}^{(i)}(1 - \varepsilon) = O_\varepsilon(w_i) ,$$

where $w_i = o(t_{\mathrm{mix}}^{(i)}(\frac{1}{4}))$.

1.4. **Main result.** Given the above definitions, our main result establishes cutoff for the FK-dynamics for a range of sub-critical values of the parameters p, q.

Theorem 1.1. *For any $d \geq 2$, there exists $p_0 = p_0(d) > 0$ such that, for all $p \in (0, p_0)$ and $q > 1$, there exists a constant $\lambda_\infty = \lambda_\infty(p, q)$ such that the FK-dynamics on Ω_n exhibits cutoff at $\frac{d}{2\lambda_\infty} \log n$ with order $O(\log \log n)$ window size.*

Some remarks are in order. Note that the case $q = 1$ is the well known example of random walk on a hypercube where cutoff occurs for all values of p see [19, Theorem 18.3]. Similarly in the case $d = 1$, one notices that each edge is a cut edge unless the configuration is completely full. Thus the process in this case can also be coupled with a random walk on a hypercube, implying cutoff for all values of p and q.

The value of the threshold p_0 in the statement above, only depends on the dimension through the value of the critical bond percolation probability and does not depend on q. We shall assume

---

[1]$\|\mu_1 - \mu_2\|_2$ will be used to denote the L^2-distance where the 1-norm in (1.1) is replaced by the 2-norm, i.e.,

$$\|\mu_1 - \mu_2\|_2 = \sum_{x \in S} \left| \frac{\mu_1(x)}{\mu_2(x)} - 1 \right|^2 \mu_2(x) .$$

that $q > 1$ is fixed from now on. Notice that by a duality argument as in [5, Section 7], in the planar case (i.e., $d = 2$) it follows that Theorem 1.1 holds also when p is close enough to 1. We will also elaborate on a description of λ_∞ in terms of the spectral gap of the Glauber dynamics for the infinite volume RCM in Section 7.2.

1.5. **Background and related work.** There has been much activity over the past two decades in analyzing Glauber dynamics for spin systems in both statistical physics and computer science leading to deep connections between the mixing time and the phase structure of the physical model. In contrast, the Glauber dynamics for the RCM remains less understood. The main reason for this is that connectivity is a global property. Ullrich in a series of important papers [32, 31, 33] established comparison estimates between the FK-dynamics and the well known non-local Swendsen-Wang (SW) dynamics ([30]) using functional analytic arguments. Although initially the arguments appeared only for integer values of q exploring connections with the Ising/Potts models, the analysis extends to all $q > 1$, which appeared in [4]. Until recently, all existing bounds on the FK-dynamics were via transferring results for the SW or related dynamics [30] using comparison estimates as above. However these methods typically yield highly sub-optimal bounds and does not provide any insight into the behavior of RCM. Recently the authors of [5] established a fast mixing time of order $O(n^2 \log n)$ bound for the discrete time FK-dynamics on RCM in a box of size n in $\mathbb{Z}^2$ with a special class of boundary conditions. The proof works for all $q \geq 1$ and $p \neq p_c(q)$. Furthermore, although not explicitly mentioned, the arguments extend to periodic boundary conditions as well. The key ingredients used were planar duality, tools developed for mixing of spin systems in [26] and most importantly the exponential decay of connectivity below $p_c(q)$ established in the breakthrough work [2]. More recently [3] extends the results to a more general class of boundary conditions with weaker bounds. Among various things, the latter work in particular also shows that boundary conditions can have a drastic effect on the mixing time.

A general conjecture of Peres [28] indicates that one should expect cutoff to occur in the regime of fast mixing for many natural chains as above. In the breakthrough papers, [20, 21], Lubetzky and Sly verified the above conjecture for Glauber dynamics for Ising and Potts models, putting forward a host of new methods using ideas similar to the Propp-Wilson *coupling from the past* [29] as well as relating L^1-mixing to L^2-mixing using powerful log-Sobolev inequalities [7]. Subsequently in [23, 24], the results of the above papers were refined by inventing the general **Information percolation** machinery. Furthermore in very recent work, [27] extended the above framework to prove cutoff results for the non-local SW dynamics for Potts models on the torus in any dimension for suitably high temperatures.

However as indicated above, the FK-dynamics has significant differences with the above described spin models and whether cutoff occurs in the fast mixing regime in this case was left open. The main theorem of this paper answers this question in the affirmative as long as p is small enough and $q > 1$. In the process, we extend the Information Percolation framework to the RCM setting as well. An elaborate description of the various geometric difficulties and how to encounter them is presented in the next section. We end this section by also mentioning the recent work of Lubetzky and Gheissari on proving quasi-polynomial bounds for the mixing time at criticality for FK-dynamics in two dimensions and related bounds for critical spin systems in [14, 15, 16] based on recent breakthroughs in [9, 10] .

2. Idea of the proof and organization of the article

We first develop a graphical construction (grand coupling of FK-dynamics) which will be quite useful in constructing coupling arguments. We then discuss the key issues that one faces towards proving the main result and what new ideas one needs beyond the existing literature to address them.

2.1. **Graphical construction/Monotone coupling.** We will define the FK-dynamics formally through the following graphical construction by creating what is now popularly called in the literature as the ***Update sequence*** (see [20, 27]). For $e \in E_n$, define the sequence of updates as

$$\text{Upd}(e) = \{(t_1, U_1), (t_2, U_2), \cdots\} , \qquad (2.1)$$

where $t_1 < t_2 < \cdots$ is a sequence of update times obtained from an independent Poisson process with rate 1 attached at e, and for each i, U_i is a uniform random variable in $[0, 1]$ independent of all other randomness. The sequence $\text{Upd}(e)$ is the update sequence corresponding to e. Then, we define the full update sequence as

$$\text{Upd} = \bigcup_{e \in E_n} \text{Upd}(e) . \qquad (2.2)$$

Note that $t_i(e) \neq t_j(e')$ for all $i, j \in \mathbb{N}$ and $e, e' \in E_n$ almost surely. It would also be useful to define for $0 < t_1 < t_2$, the update sequence of e in the time interval $(t_1, t_2]$ as

$$\text{Upd}[t_1, t_2](e) = \{(s, U) : (s, U) \in \text{Upd}(e), s \in (t_1, t_2]\} , \qquad (2.3)$$

and

$$\text{Upd}[t_1, t_2] = \bigcup_{e \in E_n} \text{Upd}[t_1, t_2](e) .$$

For $X \in \Omega_n$, we say that $e \in E_n$ is a *cut-edge* if $c(X \setminus \{e\}) \neq c(X \cup \{e\})$, (recall that $c(\cdot)$ denotes the number of connected components). Furthermore, from now on, we shall assume $q > 1$ and write

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

for convenience.

We now introduce a construction of the FK-dynamics suitable for our purposes. This is the standard grand coupling for the FK-dynamics (see [18])

Definition 2.1 (FK-dynamics/Monotone Coupling). For each $(t, U) \in \text{Upd}(e)$ for some $e \in E_n$,
(1) (a) If $U < 1 - p + p^*$, we let

$$X_t(e) = \begin{cases} 0 & \text{if } U \in [0, 1 - p), \\ 1 & \text{if } U \in [1 - p, 1 - p + p^*). \end{cases}$$

 (b) If $U \geq 1 - p + p^*$, we let $X_t(e) = 0$ if e is a cut-edge in (Λ_n, X_{t-}), and $X_t(e) = 1$ if e is not a cut-edge in (Λ_n, X_{t-}).
(2) We set $X_t(e') = X_{t-}(e')$ for all $e' \neq e$.

We will denote by $\mathbb{P}_{x_0} = \mathbb{P}_{x_0}^{p, q, n}$ the law of the FK-dynamics starting from $x_0 \in \Omega_n$. Similarly, for a probability measure ν on Ω_n, denote by $\mathbb{P}_\nu$ the law of FK-dynamics starting from the initial distribution ν. Note that the FK-dynamics is reversible with respect to its invariant measure $\mu_{p,q}^n$. Naturally the update sequence allows a grand coupling of (X_t) started from all possible configurations x_0. A well known fact is the monotonicity of FK-dynamics i.e., if (X_t) and (Y_t) are two copies of the Markov chain started from x_0 and y_0 with $x_0 \leq y_0$ in the usual partial order on Ω_n, then under the grand coupling for all later times t one has $X_t \leq Y_t$. Thus often this coupling is called the *monotone coupling* and the corresponding law is denoted by $\mathbb{P}_{x_0, y_0}$. Note that another perhaps more canonical way to define the dynamics would be to first check if e is a cut-edge (resp. not) and then accordingly set it to 0 or 1 depending on whether $U < 1 - p^*$ or not (resp. $U < 1 - p$ or not). However the above alternative formulation has the nice property that if $U < 1 - p + p^*$, we do not need to check whether e is a cut-edge or not, and the randomness at e only depends on U, not the entire configuration of X_t. This will be used throughout the paper in various coupling arguments.

2.2. **The key ideas of the proof.** In the work of Lubetzky and Sly [20] on the Ising model, the key idea was to break the dependencies in the Markov chain to reduce the analysis to the study of a product chain of Glauber dynamics on small boxes. The proof then relied on the relation between the L^1-mixing time of the product chain to L^2-mixing time of the individual coordinates and sharp estimates on the latter obtained via Log-Sobolev inequalities (LSI). Unfortunately such functional analytic tools are not available for the RCM. Although it is perhaps natural to predict that such estimates hold at least in some part of the parameter space, it is important to point out that the standard arguments which work for nearest neighbor spin systems fail owing to long range effects. Whether the LSI indeed holds for the RCM thus remains an important open problem.

Furthermore, to improve the size of the cutoff window to $O(1)$, in [23, 24], the powerful machinery of information percolation was invented to bypass the use of log-Sobolev inequalities to estimate the L^2-mixing time. The proof however still relied heavily on the local nature of Glauber dynamics for spin systems. On the other hand a non-local Markov chain admitting global changes is the well known Swendsen-Wang (SW) dynamics for Potts model. In SW dynamics for the Potts model, one proceeds by sampling an independent bond percolation on each of the mono-chromatic components (connected component of vertices with the same spins) and then for each connected component of the percolation sampled, a uniformly random spin is assigned. This is done at every time step independently of the past and hence the interaction of the spin at every vertex at every time step in only limited to spins within its percolation cluster.

Very recently in [27] the strategy was extended to SW dynamics. The latter work is based on the observation made above that while in Glauber dynamics, in one step the spin at a vertex can only depend on its immediate neighbors, the state of a vertex in SW by definition depends on all the vertices inside an independent percolation cluster sampled at each time step. Thus in the subcritical regime, since the cluster diameters have exponential tails, one can expect the same approach to go through and indeed this is what is made rigorous in [27]. The arguments in this article draw inspiration mostly from this last article.

As indicated before, at a very high level, one of the main contributions of our approach is extending the Information Percolation framework to the setting of FK-dynamics. However in RCM, in one step the update of an edge can depend on the status of an arbitrarily far located edge (see Figure 1.1). To bypass this, we first run the process for an $O(1)$ burn-in time which allows the process to be dominated by a subcritical Bernoulli percolation.

At this point we try to analyze the information percolation clusters. Very informally (see Section 5 for precise definitions) this approach involves keeping track of the interactions between various edges as they are updated, *backwards* in time. For e.g.,: if an edge e is updated using an element $(t, U) \in \mathrm{Upd}(e)$ one of two things could happen (recall Definition 2.1):

- $U < 1-p+p^*$, in which case the updated value of the edge is a Bernoulli variable independent of the state of the system. In this case we call the edge to become ***Oblivious***.
- However if $U > 1 - p + p^*$ one needs to check whether e is a cut-edge or not and in the process interacts (shares information) with several edges.

Formally one considers a space-time slab (see Figure 5.1) and evolves backward in time by branching out to all possible edges an update shares information with, or gets killed in case of an oblivious update. The key usefulness of this approach as exploited in [20, 21, 23, 24, 27] is that if the backward branching process (called the ***History diagram***) is subcritical then, the process will be killed before reaching the initial configuration in this backward evolution causing the final configuration to be independent of the initial one implying coupling of all starting states under the grand coupling. However this is an overkill since for cutoff to occur one can tolerate some mild dependence on the initial condition as long as that is hidden inside the natural fluctuation of the system.

To bound the growth rate of the history diagram we first discretize time with interval length $\Delta = \frac{1}{\sqrt{p}}$ (as the reader will notice, this choice of Δ is not special and a host of other choices will

work too) and consider the interval $[\tau_i, \tau_{i+1}]$ where $\tau_i = i\Delta$ and define the history diagram only at times τ_i. We first extract several auxiliary percolation models based on the update sequence (see Table 1), and one of which denoted by $\Xi_i(\cdot)$ captures the following: For every i, $\Xi_i(e)$ is 1 iff e has not been updated in the interval $[\tau_i, \tau_{i+1}]$ or e is open at least once in $[\tau_i, \tau_{i+1}]$ for the Glauber dynamics for the standard Bernoulli percolation with parameter p, (random walk on the hypercube) using the same update sequence and starting from the empty configuration. Now given the history diagram up to time τ_{i+1} for any edge we first check if it has been updated or not in an interval $[\tau_i, \tau_{i+1}]$ (recall the history diagram flows backwards). If not, the edge continues to be a part of the history diagram, if it is updated using an oblivious update it gets killed, otherwise we bound the spreading of information by the connected component of e in the percolation $\Xi_i \cup \Xi_{i-1}$.

Note that to ensure that the state of the edge e throughout the interval $[\tau_i, \tau_{i+1}]$ does not depend on any edge not included in the history diagram we need the boundary of the latter to be closed throughout the entire interval i.e., we must consider its connected component 'forward in time' which a priori depends on the entire time interval $[0, \tau_i]$. However this is the point at which we use the smallness of p crucially, which creates an environment which is subcritical and hence the connected component can be bounded by the connected component of $\Xi_i \cup \Xi_{i-1}$, i.e., instead of the entire interval $[0, \tau_{i+1}]$ we can get by, just using the information on $[\tau_{i-1}, \tau_{i+1}]$.

Given the above, the situation is similar to the definition of the SW dynamics considered in [27], except that the percolation sampled at every discrete time step is now 1-dependent across time. This creates the need for a refined and delicate analysis of the information percolation clusters to yield L^2-mixing bounds. This is stated as Theorem 5.1 and Proposition 5.2. The proof of the latter is the core of this work. The above approach adopted in the paper of extracting dependent percolation models that can be analyzed could be of independent interest and useful in other general contexts in bounding how passage of information occurs in such dynamical settings.

Assuming these results, the arguments used to show cutoff are quite similar to the ones already appearing in [27] based on the methods in [20]. An additional ingredient needed to prove Theorem 5.1 from Proposition 5.2 is that the spectral gap of the FK-dynamics is positive uniformly in the system size. In SW the lower bound on the spectral gap follows by path coupling by establishing a one step contraction which unfortunately is absent in our setting; instead we rely on the a priori mixing time bounds obtained in [5]. c.

Finally, we mention that for the Ising model, [23] exploited monotonicity of the system, to prove an $O(1)$ bound on the cutoff window without resorting to the methods of [20]. Such sharp bounds are missing in [27] which deals with the general Potts model. However the RCM is monotone and whether this can be used to prove a similar improvement of Theorem 1.1 is not pursued in this paper and is left for further research. Furthermore, another possible direction to investigate is the effect of boundary conditions. While the current paper only deals with periodic boundary conditions, for local dynamics on Ising and Potts models [21] proved sharp mixing time results for general boundary conditions. Recall that typically in addressing such questions, there are two goals. One is to control the cutoff window size and the other is to pin down the location. Under certain special cases, in [21], the location of mixing was related to infinite volume objects. Moreover, to bound the window size, [21] relied on certain worst case Log-Sobolev constants. Since these are not available in our setting and boundary conditions can lead to delicate global dependencies, the current arguments in the paper do not directly go through. Nonetheless, this is an important project to be taken up in the future.

2.3. **Organization of the article.** We prove and collect results about a priori bounds on the mixing time and the spectral gap in Section 3 to be used throughout the rest of the article. As mentioned above, we need to define several auxiliary percolation models based on the update sequence. This is done in Section 4. **Section 5 is the core of this work and the main contribution in**

this paper which bounds the L^2-mixing time by defining suitable information percolation clusters. This section is rather long and has several new constructions and delicate geometric arguments. However assuming the main result of this section, the proof of Theorem 1.1 is quite similar to the arguments appearing in [20, 23, 27]. The reader not familiar with the latter papers can choose to first assume the results of Section 5 to see how they are used in the subsequent sections to then come back to the proofs of Section 5.

The proof of the main result Theorem 1.1 spans Section 6 where certain modifications of arguments of [20] and Section 7 where the final proof appears. The outstanding proofs of some of the stated claims are collected in the Appendix (Section 8).

3. A PRIORI BOUNDS ON MIXING TIME AND SPECTRAL GAP

We start by recalling the following standard result.

Proposition 3.1. [19, Theorems 12.3 and 12.4] *Let (Z_t) be a discrete time ergodic reversible Markov chain on a finite state space S with the equilibrium measure π, let $\mathbb{Q}_z$ be the law of Markov chain (Z_t) starting from $z \in S$, and let γ be the spectral gap of the Markov chain (Z_t). Then,*

$$(1-\gamma)^t \leq 2 \sup_{z \in S} \|\mathbb{Q}_z(Z_t \in \cdot) - \pi\|_{TV} \leq \frac{1}{\pi_{\min}} (1-\gamma)^t \,,$$

where $\pi_{min} = \min_{x \in S} \pi(x)$.

In [5], a discrete version of FK-dynamics is considered where at every discrete time step, an uniformly chosen edge is updated. Denote by $(\widehat{X}_k)_{k \geq 0}$ the discrete FK-dynamics in Ω_n, and by $\widehat{\mathbb{P}}_{x_0, y_0}$, the law of the monotone coupling (Definition 2.1) of two copies of discrete FK-dynamics $\widehat{X}_k$ and $\widehat{Y}_k$ starting from two initial conditions $x_0, y_0 \in \Omega_n$ respectively. Moreover, let $\hat{\lambda}(n) = \hat{\lambda}(n, p, q)$ denote the spectral gap of the above process. Furthermore let $\hat{t}_{\mathrm{mix}} = \hat{t}_{\mathrm{mix}}(1/4)$ and $\hat{d}(t)$ be the mixing time and the worst-case distance to stationarity respectively in the sense of (1.2) for the discrete time dynamics. Then, the following sharp mixing time results were either obtained or are consequences of the results in [5]. In the latter, only the two dimensional case was treated but one can easily verify that the arguments extend to general dimensions under exponential decay of connectivity. We provide brief sketches of the proofs of these results with pinpoint references to the relevant literature for the remaining details.

Theorem 3.2. *For any dimension d, there exists $p_0 = p_0(d)$ such that for all $q \geq 1$ and $p < p_0$, there exists $C = C(p) > 0$ and $\lambda = \lambda(p) > 0$ such that:*

(1) *For all $x_0, y_0 \in \Omega_n$, $k \leq o(n^{1/(d+2)})$ and $e \in E_n$, it holds that*

$$\widehat{\mathbb{P}}_{x_0, y_0} \left[\widehat{X}_{kn^d}(e) \neq \widehat{Y}_{kn^d}(e) \right] \leq e^{-Ck} \,.$$

(2) *The mixing time $\hat{t}_{\mathrm{mix}}$ of discrete process $\widehat{X}_k$ is $\Theta(n^d \log n)$.*

(3) *For all $n \in \mathbb{N}$, $\hat{\lambda}(n) \geq \lambda n^{-d}$.*

Remark 3.3. Indeed, one can take p_0 to be the critical Bernoulli bond percolation probability on $\mathbb{Z}^d$. For $d = 2$, thanks to the complete knowledge about exponential decay of connectivity up to the critical point established in [2], the results of Theorem 3.2 were shown to hold for all subcritical p, for each $q \geq 1$ in [5].

Proof. (1) and (2) appear as [5, Display (13)], and [5, Theorem 6.1] respectively. Note that (1) proves the upper bound in (2) by taking $k = C \log n$. The proof of the lower bound of mixing time appears in [5, Theorem 6.1]. Although (3) does not quite appear in [5] it is a consequence of (1). To see this, we will use the well known lower bound of total variation distance in terms of spectral gap recalled in Proposition 3.1. Namely, using the above and union bounding over all elements in E_n, we get that $\hat{d}(kn^d)$, the worst-case total variation distance at time kn^d is $e^{-\Omega(k)+d \log n}$, and hence

$$(1 - \hat{\lambda})^{kn^d} \leq e^{-\Omega(k)+d \log n}$$

for all $k \leq o(n^{1/(d+2)})$. Now taking logs we get $-kn^d \hat{\lambda} \leq -\Omega(k) + d \log n$, and therefore for some $C > 0$,

$$\frac{1}{n^d} \left(C - \frac{\log n}{k} \right) \leq \hat{\lambda} \,.$$

Thus by choosing a large enough $k = o(n^{1/(d+2)})$ the result follows. $\qquad\square$

However for our purposes, we will need a translation of the result for the continuous time setting. Denote by $\lambda(n) = \lambda(n, p, q)$, the spectral gap of the continuous time FK-dynamics defined in Definition 2.1.

Corollary 3.4. *For any dimension d, there exists $p_0 = p_0(d)$ such that for all $q \geq 1$ and $p < p_0$, there exists $C = C(p) > 0$ and $\lambda = \lambda(p) > 0$ such that:*

(1) *For all $x_0, y_0 \in \Omega_n$ and $k \leq o(n^{1/(d+2)})$, it holds that,*

$$\mathbb{P}_{x_0, y_0} [X_t(e) \neq Y_t(e)] \leq e^{-Ct} \,.$$

(2) *The FK-dynamics in Λ has mixing time of order $\Omega(\log n)$.*
(3) *For all $n \in \mathbb{N}$, it holds that $\lambda(n) \geq \lambda$.*

Proof. All these results are immediate from Theorem 3.2 since the continuous dynamics is n^d times faster than the discrete counterpart. In particular, to show part (3), see [19, Lemmas 20.5 and 20.11]. $\qquad\square$

4. Auxiliary percolation models, and disagreement propagation bounds

Given the randomness defined by the update sequence in (2.1), we will need to define several auxiliary percolation models extracted from the graphical construction, which though simple will be useful in various comparison arguments appearing throughout the paper. We will also state useful bounds on the speed of propagation of disagreements. We start with the percolation models. Before providing precise definitions, for the reader's benefit we give short descriptions off what each of these models capture. Furthermore, for ease of reference throughout the article, all the definitions are collected in Table 1 at the end of this section and the reader can choose to skip the precise definitions at first read referring to the table whenever needed.

(1) Standard Percolation dynamics ($q = 1$)/Random walk on the hypercube, i.e., edges are randomly refreshed at rate one with a Bernoulli(p) variable independently. This will dominate the FK-dynamics in the regime of our interest.
(2) Enlarged percolation: An edge is said to be open if it was open at least once in the Standard Percolation dynamics in a given (to be specified) time interval.
(3) Update/Non-update percolation: An edge is open if it has not been updated at least once in a given interval of time.

4.1. **Standard percolation dynamics (SPD).** It will be useful to discretize time as we will see in later applications. Throughout the article we will fix $\Delta := \Delta(p) = p^{-1/2}$, to be the basic unit of discretization and let $\tau_i := i\Delta$. (The choice of Δ is not special as long as it satisfies the properties discussed in this section.) Also let $\mathbb{Z}_+$ be the set of non-negative integers.

Definition 4.1 (SPD associated to the update sequence Upd)**.** For each $i \in \mathbb{Z}_+$, we construct a SPD $(\mathscr{F}_t^i)_{t \geq \tau_i}$ in Ω_n as follows:

  (1) $\mathscr{F}_{\tau_i}^i = E_n$.
  (2) For each $t > \tau_i$ and $e \in E_n$,
      (a) If $\mathrm{Upd}[\tau_i, t](e) = \emptyset$, we let $\mathscr{F}_t^i(e) = \mathscr{F}_{\tau_i}^i(e)(= 1)$.
      (b) Otherwise, let (t^*, U^*) be the last update in $\mathrm{Upd}[\tau_i, t](e)$.
          (i) We let $\mathscr{F}_t^i(e) = 1$ if $U^* > 1 - p$,
          (ii) else let $\mathscr{F}_t^i(e) = 0$ if $U^* \leq 1 - p$.

We define the dynamics $(\mathscr{E}_t^i)_{t \geq \tau_i}$ in an identical manner by replacing step (1) with $\mathscr{E}_{\tau_i}^i = \emptyset$. In other words, $(\mathscr{F}_t^i)$ and $(\mathscr{E}_t^i)$ are the Glauber dynamics of the percolation measure with open probability p on Ω_n starting at $t = \tau_i$ from the full and empty configurations, respectively.

Since $(\mathscr{F}_t^i)$ and $(\mathscr{E}_t^i)$, for $i \in \mathbb{Z}_+$, and the FK-dynamics (X_t), share the same update sequence, we can couple all of them in the time window $[\tau_i, \infty)$ in a natural manner calling this as the *canonical coupling*. We record some simple but useful lemmas below.

Lemma 4.2. *Under the canonical coupling, for all $i \in \mathbb{Z}_+$, it holds that*

$$X_t \leq \mathscr{F}_t^i \text{ for all } t \geq \tau_i .$$

Proof. Denote by X_t^{full} the FK-dynamics on Ω_n with $X_0 = E_n$, the full configuration. Via the monotone coupling, we have $X_t \leq X_t^{\mathrm{full}}$ for all $t \geq 0$. Now the inclusion $X_t^{\mathrm{full}} \leq \mathscr{F}_t^0$ for all $t \geq 0$ comes directly from the definitions of FK-dynamics and percolation dynamics. Since we have $\mathscr{F}_t^0 \leq \mathscr{F}_t^i$ for all $t \geq \tau_i$ for all $i \in \mathbb{Z}_+$ under the canonical coupling, we are done. $\square$

For $s \in [0, 1]$, denote by $\mathrm{Perc}_n(s)$ the standard bond percolation on E_n where an edge e is open with probability s. Denote by $\preceq$ the usual stochastic domination.

Lemma 4.3. *For all $i \in \mathbb{Z}_+$ and $t \geq 0$, the law of $\mathscr{F}_{t+\tau_i}^i$ is given by $\mathrm{Perc}_n(e^{-t} + p[1 - e^{-t}])$. Therefore, for all $x_0 \in \Omega_n$, it holds that*

$$\mathbb{P}_{x_0}[X_t \in \cdot] \preceq \mathrm{Perc}_n(e^{-t} + p[1 - e^{-t}]) .$$

Proof. By definition, $\mathscr{F}_{t+\tau_i}^i(e) = 1$ if $\mathrm{Upd}[\tau_i, \tau_i + t](e) = \emptyset$. Otherwise, i.e., if $\mathrm{Upd}[\tau_i, \tau_i + t](e) \neq \emptyset$,

$$\mathscr{F}_{t+\tau_i}^i(e) = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } 1 - p \end{cases}$$

since the status of $\mathscr{F}_{t+\tau_i}^i(e)$ depends only on the last update for this edge before $t + \tau_i$. Since

$$\mathbb{P}[\mathrm{Upd}[\tau_i, \tau_i + t](e) = \emptyset] = e^{-t} ,$$

it follows that

$$\mathbb{P}[\mathscr{F}_{t+\tau_i}^i(e) = 1] = e^{-t} + p[1 - e^{-t}] .$$

The proof of the first assertion is completed since the status of edges are independent under SPD. The second assertion follows from Lemma 4.2 and choosing $i = 0$. $\square$

As indicated in Section 2, we will allow ourselves an $O(1)$ burn-in time which will be enough by the above domination results for the configuration to look like a sample of a subcritical percolation. This then creates a situation where no connected component is large and hence the interactions between various edges are still rather local. To make this formal, denote by $p_{\mathrm{perc}}(d) \in (0, 1)$ the critical

probability of the edge percolation in $\mathbb{Z}^d$. **From now on we will assume that $p \in (0, p_{\mathbf{perc}}(d))$ and further arguments would put additional smallness conditions on p.** Define

$$p_{\mathrm{init}} = p_{\mathrm{init}}(p) := \frac{1}{2}(p + p_{\mathrm{perc}}(d)) \in (p,\, p_{\mathrm{perc}}(d))\,,$$

and let $t_{\mathrm{init}} = t_{\mathrm{init}}(p)$ be the solution of the following equation:

$$p(1 - e^{-t_{\mathrm{init}}}) + e^{-t_{\mathrm{init}}} = p_{\mathrm{init}}\,. \tag{4.1}$$

As the next lemma will show, we can restrict our initial conditions to the class of measures ν satisfying $\nu \preceq \mathrm{Perc}_n(p_{\mathrm{init}})$. More precisely, define

$$\widehat{d}(t) = \sup_{\nu:\nu\preceq\mathrm{Perc}_n(p_{\mathrm{init}})} \left\| \mathbb{P}_\nu\left[X_t \in \cdot\,\right] - \mu_{p,q}^n \right\|_{\mathrm{TV}}\,,$$

and

$$\widehat{t}_{\mathrm{mix}}(\epsilon) = \inf\left\{t : \widehat{d}(t) < \epsilon\right\}\,.$$

Then, we obtain the following comparison result between $t_{\mathrm{mix}}(\epsilon)$ and $\widehat{t}_{\mathrm{mix}}(\epsilon)$.

Lemma 4.4. *For all $p < p_{perc}(d)$ and $t > t_{init}$, we have*

$$\sup_{x_0 \in \Omega_n} \left\| \mathbb{P}_{x_0}\left[X_t \in \cdot\,\right] - \mu_{p,q}^n \right\|_{TV} \leq \sup_{\nu:\nu\preceq\mathrm{Perc}_n(p_{init})} \left\| \mathbb{P}_\nu\left[X_{t-t_{init}} \in \cdot\,\right] - \mu_{p,q}^n \right\|_{TV}\,. \tag{4.2}$$

Therefore, we have

$$\widehat{t}_{mix}(\epsilon) \leq t_{mix}(\epsilon) \leq \widehat{t}_{mix(\epsilon)} + t_{init}\,. \tag{4.3}$$

Proof. By Lemma 4.3 and definition of t_{init} and p_{init}, we have that the distribution of $X_{t_{\mathrm{init}}}$ given any initial configuration is stochastically bounded by $\mathrm{Perc}_n(p_{\mathrm{init}})$. Hence, the first assertion of proposition follows. The inequalities in (4.3) follow since $\widehat{d}(t) \leq d(t) \leq \widehat{d}(t - t_{\mathrm{init}})$ by (4.2). $\qquad\square$

Thus we will take t_{init} to be our burn-in time.

4.2. **Enlarged and non-update percolations.** In this section we define the second and the third models indicated at the beginning of the section.

Definition 4.5. We define two sequences of random configurations $(\overline{\mathscr{F}}_i)_{i\in\mathbb{N}}$ and $(\overline{\mathscr{E}}_i)_{i\in\mathbb{N}}$ in Ω_n based on the definitions $(\mathscr{F}_t^i)_{i\in\mathbb{N}}$ and $(\mathscr{E}_t^i)_{i\in\mathbb{N}}$ as follows:

  (1) For $i \in \mathbb{N}$, define $\overline{\mathscr{F}}_i \in \Omega_n$ as

$$\overline{\mathscr{F}}_i(e) = 1 \text{ iff } \mathscr{F}_t^{i-1}(e) = 1 \text{ for some } t \in [\tau_i,\, \tau_{i+1}]\,.$$

   Note that here we consider $\mathscr{F}_t^{i-1}$ instead of $\mathscr{F}_t^i$ since otherwise $\overline{\mathscr{F}}_i(e)$ would be deterministically 1.
  (2) For $i \in \mathbb{Z}_+$, define $\overline{\mathscr{E}}_i \in \Omega_n$ as

$$\overline{\mathscr{E}}_i(e) = 1 \text{ iff } \mathscr{E}_t^i(e) = 1 \text{ for some } t \in [\tau_i,\, \tau_{i+1}]\,.$$

The following result is a static version of Lemma 4.2.

Lemma 4.6. *Under the canonical coupling, for all $i \in \mathbb{N}$, we have*

$$X_t \leq \overline{\mathscr{F}}_i \text{ for all } t \in [\tau_i,\, \tau_{i+1}]\,.$$

Proof. Since $X_t \leq \mathscr{F}_t^{i-1}$ for all $t \in [\tau_i,\, \tau_{i+1}]$ by Lemma 4.2, the proof is immediate from the definition of $\overline{\mathscr{F}}_i$. $\qquad\square$

Now we investigate the distributions of $\overline{\mathscr{E}}_i$ and $\overline{\mathscr{F}}_i$. To this end we introduce the non-update percolation $\mathscr{N}_i \in \Omega_n$, for $i \in \mathbb{Z}_+$, as the following:

$$\mathscr{N}_i(e) = \begin{cases} 1 & \text{if } \mathrm{Upd}[\tau_i, \tau_{i+1}](e) = \emptyset \,, \\ 0 & \text{if } \mathrm{Upd}[\tau_i, \tau_{i+1}](e) \neq \emptyset \,. \end{cases} \tag{4.4}$$

In order words, $\mathscr{N}_i(e) = 0$ if and only if there is an update $(t_1, U_1) \in \mathrm{Upd}(e)$ such that $t_1 \in (\tau_i, \tau_{i+1}]$. Given the above definitions, we have the following comparison results.

Lemma 4.7. *The following holds:*
(1) *For all $i \in \mathbb{Z}_+$, we have $\overline{\mathscr{E}}_i \preceq \mathrm{Perc}_n(p^{1/2})$.*
(2) *For all $i \in \mathbb{Z}_+$, we have $\mathscr{N}_i \preceq \mathrm{Perc}_n(p^{1/2})$.*
(3) *For all $i \in \mathbb{N}$, we have $\overline{\mathscr{F}}_i \preceq \mathrm{Perc}_n(3p^{1/2})$.*

Proof. We start by observing that $\overline{\mathscr{E}}_i(e) = 1$ if and only if

$$\{U > 1 - p \text{ for some } (t, U) \in \mathrm{Upd}[\tau_i, \tau_{i+1}](e)\} \,. \tag{4.5}$$

To compute the probability of the latter notice that given the event $|\mathrm{Upd}[\tau_i, \tau_{i+1}](e)| = k$, the event (4.5) happens with probability $1 - (1 - p)^k$. Hence, the probability of the event (4.5) can be written as

$$\sum_{k=0}^{\infty} e^{-\Delta} \frac{\Delta^k}{k!} (1 - (1 - p)^k) = 1 - e^{-p\Delta} \leq p\Delta = p^{1/2} \,.$$

This finishes the proof of (1). Part (2) can be readily obtained from the observation that

$$\mathscr{N}_i \sim \mathrm{Perc}_n(e^{-\Delta}) \preceq \mathrm{Perc}_n(p^{1/2}) \,.$$

For part (3), we claim that

$$\overline{\mathscr{F}}_i \leq \mathscr{N}_{i-1} \cup \overline{\mathscr{E}}_{i-1} \cup \overline{\mathscr{E}}_i \,. \tag{4.6}$$

This claim along with parts (1) and (2) will finish the proof. To prove the claim, first suppose that $\overline{\mathscr{F}}_i(e) = 1$ and $\mathscr{N}_{i-1}(e) = 0$. Then, $\mathrm{Upd}[\tau_{i-1}, \tau_i](e) \neq \emptyset$ and hence we can take the last update (t_1, U_1) in $\mathrm{Upd}[\tau_{i-1}, \tau_i](e)$. Since $\overline{\mathscr{F}}_i(e) = 1$, at least one update (t, U) in $\{(t_1, U_1)\} \cup \mathrm{Upd}[\tau_i, \tau_{i+1}](e)$ satisfies $U > 1 - p$. It implies either $\overline{\mathscr{E}}_{i-1}(e) = 1$ or $\overline{\mathscr{E}}_i(e) = 1$. This finishes the proof. $\square$

We end this section with a final definition. For $i \in \mathbb{Z}_+$, let

$$\Xi_i := \overline{\mathscr{E}}_i \cup \mathscr{N}_i \in \Omega_n \,. \tag{4.7}$$

We record a key fact in the next lemma. In short the lemma says that the FK-dynamics across time can be dominated by a sequence of Bernoulli percolations which are one dependent across time. This will be crucially used in the analysis of how information spreads in the FK-dynamics.

Proposition 4.8. *The following hold:*
(1) *For all $i \in \mathbb{Z}_+$, the distribution of Ξ_i is stochastically dominated by $\mathrm{Perc}_n(2p^{1/2})$.*
(2) *For all $i \in \mathbb{N}$, under the canonical coupling, we have that*

$$X_t \leq \Xi_{i-1} \cup \Xi_i \text{ for all } t \in [\tau_i, \tau_{i+1}] \,.$$

Proof. The proof of part (1) is immediate from (1) and (2) of Lemma 4.7, while the proof of part (2) is an immediate consequence of Lemma 4.6 and (4.6). $\square$

For purpose of easy reference throughout the article we record all the percolation models defined so far in Table 1.

Percolation	Description	Defined in
$\mathscr{E}_t^{\circ i}$	Percolation on $[\tau_i, \infty)$, starting at τ_i from empty	Def. 4.1
$\mathscr{F}_t^i$	Percolation on $[\tau_i, \infty)$, starting at τ_i from full	Def. 4.1
$\overline{\mathscr{E}}_i$	Open some time in $[\tau_i, \tau_{i+1}]$ starting with empty at τ_i	Def. 4.5
$\overline{\mathscr{F}}_i$	Open some time in $[\tau_i, \tau_{i+1}]$ starting with full at τ_{i-1}	Def. 4.5
$\mathscr{N}_i$	Non-update in $[\tau_i, \tau_{i+1}]$ implies open	(4.4)
Ξ_i	$\overline{\mathscr{E}}_i \cup \mathscr{N}_i$	(4.7)

TABLE 1. Different kinds of percolation.

4.3. **Decay of connectivity.** We now record some useful exponential decay of connectivity results for a non-equilibrium RCM. It is well-known that for a sub-critical bond percolation or RCM, one observes an exponential decay of connectivity, i.e., the probability that two sites u and v belong to the same cluster decays exponentially in the graph distance $d(u, v)$, (cf. [2, Theorem 2]). We would need a dynamical version for our purposes and start with some definitions. Note that $\overline{\mathscr{F}}_i$ had so far been defined for $i \geq 1$ only. We now define $\overline{\mathscr{F}}_0$ as

$$\overline{\mathscr{F}}_0 = X_0 \cup \overline{\mathscr{E}}_0 .$$

Then, by definition

$$X_t \leq \overline{\mathscr{F}}_0 \text{ for all } t \in [0, \tau_1] . \tag{4.8}$$

Proposition 4.9. *For all small enough p, there exists $\gamma = \gamma(p) > 0$ such that,*

$$\sup_{\nu: \nu \preceq \mathrm{Perc}_n(p_{init})} \mathbb{P}_\nu \left[u \xleftrightarrow{\overline{\mathscr{F}}_i} v \right] \leq e^{-\gamma d(u, v)}$$

for all $i \in \mathbb{Z}_+$, $n \in \mathbb{N}$, and $u, v \in \Lambda_n$.

Proof. By Lemma 4.7, the distribution of $\overline{\mathscr{F}}_i$ is dominated by $\mathrm{Perc}_n(3p^{1/2})$ for $i \geq 1$. For $i = 0$, we notice from the definition of $\overline{\mathscr{F}}_0$ that the distribution of the latter is dominated by $\mathrm{Perc}_n(p_{\mathrm{init}} + p^{1/2})$.

In conclusion, for all small enough p, the distribution of $\overline{\mathscr{F}}_i$, $i \geq 0$, is dominated by $\mathrm{Perc}_n(s)$ for some $s < p_{\mathrm{perc}}(d)$ and hence we are done by decay of connectivity for subcritical percolation [8, Theorem 3.7]. $\square$

From now on, all the statements are asymptotic in n, so that they hold only when n is large enough. In addition, we write C or c for positive constants whose different occurrences might denote different values. We shall not repeat stating these explicitly.

The next result follows from similar arguments as in the proof of the previous proposition.

Lemma 4.10. *Suppose that two disjoint subsets A and B of E_n satisfy $d(A, B) \geq c \log^2 n$ for some $c > 0$. Denote by $\mu_{B^c}^+$ the random-cluster measure on $B^c = E_n \setminus B$ under the full boundary condition on B. Denote by $X \in \{0, 1\}^{B^c}$ a random-cluster configuration sampled according to $\mu_{B^c}^+$, and denote by $\mathrm{Conn}(B; X)$ the set of edges in B^c connected to an edge of B via an open path in X. Then, for all small enough p, we have*

$$\mu_{B^c}^+ [\mathrm{Conn}(B; X) \cap A = \emptyset] \geq 1 - \frac{1}{n^{2d}} .$$

Proof. One can readily observe that the decay of connectivity established in the previous result holds for any connected domain with any boundary condition. Hence, we get

$$\mu^+_{B^c}\left[\mathrm{Conn}(B;X)\cap A \neq \emptyset\right] \leq \mu^+_{B^c}\left[u \xleftrightarrow{X} v \text{ for some } u, v \in B^c \text{ such that } d(u, v) \geq c\log^2 n\right]$$

$$\leq e^{-c\log^2 n}\binom{n^d}{2} < \frac{1}{n^{2d}},$$

where the second inequality follows by the union bound. □

The final result of this section records a statement about how fast disagreement percolates in FK-dynamics.

4.4. **Estimates on the propagation of disagreements.** We fix a subset $A \subset E_n$ this section. Define an enlargement A^+ of A as

$$A^+ = \{e \in E_n : d(e, A) \leq \log^4 n\}. \tag{4.9}$$

The main objective in this section is to show that, under monotone coupling, FK-dynamics started from two configurations that agree on A^+ and are reasonably sparse, continue to agree on A for all $t \in [0, t_{\max}]$ where

$$t_{\max} = \log^2 n. \tag{4.10}$$

Consider two censored dynamics (Z^+_t) (resp. (Z^-_t)) as FK-dynamics on $\{0, 1\}^{A^+}$ conditioned on full (resp. empty) configuration on $E_n \setminus A^+$. Let $\mathrm{Perc}_n^{A^+}(p_{\mathrm{init}})$ denote the percolation measure on A^+ with open probability p_{init}.

Lemma 4.11. *Consider two copies of FK-dynamics (Z^+_t) and (Z^-_t) on $\{0, 1\}^{A^+}$ coupled via the monotone coupling. Suppose that the law of the initial condition Z^-_0 follows a law ν on $\{0, 1\}^{A^+}$ satisfying $\nu \preceq \mathrm{Perc}_n^{A^+}(p_{init})$, and suppose further that $Z^+_0 = Z^-_0$. Then, for all sufficiently small p, we have that*

$$\mathbb{P}\left[Z^+_t(A) = Z^-_t(A) \text{ for all } t \in [0, t_{max}]\right] \geq 1 - \frac{1}{n^{3d}}. \tag{4.11}$$

Remark 4.12. Note that the probability in (4.11) is with respect to both the FK-dynamics and also the initial measure ν. In other words, this is an annealed probability.

Remark 4.13. Even though we considered the two worst boundary conditions, namely, full and empty, a simple monotonicity consideration allows us to conclude that

$$\mathbb{P}\left[Z^+_t(A) = Z^-_t(A) = Z_t(A) \text{ for all } t \in [0, t_{\max}]\right] \geq 1 - \frac{1}{n^{3d}},$$

where Z_t is one of the following processes on $\{0, 1\}^{A^+}$:

- Censored FK-dynamics on A^+ conditioned on any configuration on $E_n \setminus A^+$, i.e., one that only updates sites in A^+.
- If A, and hence A^+, are square boxes, the FK-dynamics on $\{0, 1\}^{A^+}$ with periodic boundary conditions.
- The FK-dynamics on E_n projected to $A^+ \subset E_n$, i.e., $X_t(A^+)$.

Remark 4.14. In the above theorem, the size of the ambient space Λ_n (which is n) is not important. Taking the ambient space to be Λ_m which contains A^+ suffices. Moreover, we can replace $\log^4 n$ in the statement of lemma with $\log^{3+\delta}$ for any $\delta > 0$ with $t_{\max} = \log^{1+\delta} n$.

The proof follows the arguments in [5, 26] and is postponed to the Appendix (Section 8).
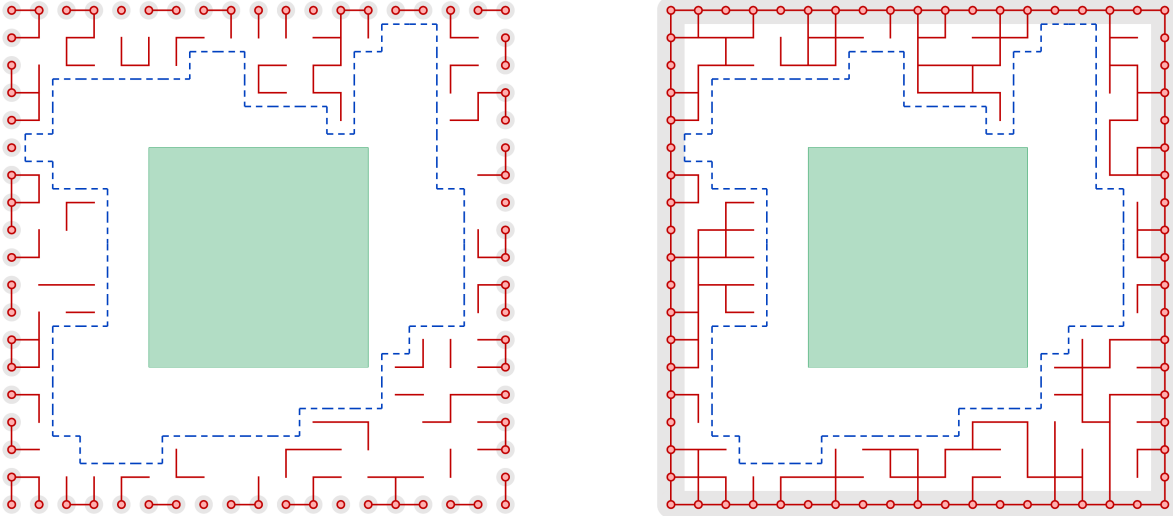
FIGURE 4.1. Figure illustrating the weak spatial mixing property of the subcritical RCM. Here we consider the equilibrium measures with free (LHS) and wired boundary conditions (RHS). By monotonicity of the equilibrium measures with respect to their boundary conditions, there exists a coupling such that the LHS is dominated by the RHS. However under this coupling by the exponential decay of connectivity the RHS (and hence the LHS) has a closed surface (contour in the planar case) within $O(\log n)$ distance from the boundary and they agree in the interior of the surface in particular on the green region.

5. Information percolation clusters and time dependent Bernoulli percolations

As emphasized before, this is the section which contains all the new ideas in the paper. The main result is the following bound on L^2-mixing. Recall the spectral gap $\lambda(n)$ from Corollary 3.4.

Theorem 5.1. *For all small enough $p > 0$, there exists $C = C(p) > 0$ such that the following L^2-bound holds for all large enough n:*

$$\max_{x_0 \in \Omega_n} \left\| \mathbb{P}_{x_0} \left[X_t \in \cdot \right] - \mu_{p,q}^n \right\|_{L^2(\mu_{p,q}^n)} \leq 2 \exp \left\{ -\lambda(n)(t - C \log n) \right\}$$

for all $t \geq C \log n$.

Recall that the spectral gap governs the rate of decay of L^2 norm. More precisely for any $s \leq t$ and any starting state $x_0 \in \Omega_n$ we have (see for example, [19, Lemma 20.5]),

$$\left\| \mathbb{P}_{x_0} \left[X_t \in \cdot \right] - \mu_{p,q}^n \right\|_{L^2(\mu_{p,q}^n)} \leq e^{-\lambda(n)(t-s)} \left\| \mathbb{P}_{x_0} \left[X_s \in \cdot \right] - \mu_{p,q}^n \right\|_{L^2(\mu_{p,q}^n)} . \tag{5.1}$$

By Corollary 3.4, it suffices to prove the following proposition.

Proposition 5.2. *For all small enough $p > 0$, there exists $C = C(p) > 0$ such that for $t_\star = C \log n$,*

$$\max_{x_0 \in \Omega_n} \left\| \mathbb{P}_{x_0} \left[X_{t_\star} \in \cdot \right] - \mu_{p,q}^n \right\|_{L^2(\mu_{p,q}^n)} \leq 2 . \tag{5.2}$$

The proof of Proposition 5.2 is the heart of this work and is rather long, intricate and involves several percolation arguments based on the models introduced in Section 4. As mentioned earlier, using the results of this section as inputs, the arguments of the following sections are quite similar to the ones appearing in [20, 27]. Readers not familiar with these papers, at first read, to get a sense of the overall flow of arguments, could choose to assume Theorem 5.1 and read the subsequent easier sections first, before coming back to this section.

We provide a roadmap for this section for the ease of reading.

- The construction of information percolation is done in Section 5.1 relying on the definitions in Section 4, particularly the percolation models listed in Table 1. At a very high level it amounts to classifying vertices into green, red and blue where the state of the red vertices depend on the initial configuration, the blue vertices are independent Bernoulli variables independent of everything else, whereas the green vertices have a complicated dependency on each other but are still independent of the initial configuration (Theorem 5.5).

- Using the above, the proof of Proposition 5.2 occupies Sections 5.2 and 5.3. The key steps are the following:

  (1) To bound the L^2-distance it suffices to condition on the green clusters. Then the strategy is to compute the L^2-distance of the conditional distribution to a product Bernoulli Measure instead of the equilibrium measure (Lemma 5.13). The Bernoulli measure is exactly the one which describes the law of the blue vertices. Thus this distance would be zero if there does not exist any red cluster.

  (2) We then establish the key estimate showing exponential unlikeliness of red vertices with time in Proposition 5.11 which makes the above step sufficient. The proof of this proposition uses a comparison with a subcritical branching process and is presented in Section 5.3. In particular, the proof involves delicate geometric arguments relying on several properties of the auxiliary percolation models defined in Table 1.

5.1. **Information percolation (IP).** As mentioned before (Section 4.1), we will discretize time using τ_i and will define IP on the space-time slab $E_n \times [\tau_1, \tau_m]$ for some $m \in \mathbb{N}$. We shall take $m = \Omega(\log n)$ later, but for the moment we think of m as a fixed integer. We also recall the various percolations defined in Table 1.

For $\Xi \in \Omega_n$, and $e = (u, v) \in E_n$ where $u, v \in \Lambda_n$, if $\Xi(e) = 1$, define $\mathrm{Conn}(e; \Xi)$ as the connected component of Ξ containing (u, v). On the other hand, we define $\mathrm{Conn}(e; \Xi) = \emptyset$ if $\Xi(e) = 0$.

Furthermore define $\partial \mathrm{Conn}(e; \Xi)$ as the edge boundary of $\mathrm{Conn}(e; \Xi)$ i.e., as the set of edges in $E_n \setminus \mathrm{Conn}(e; \Xi)$, which are adjacent to an edge in $\mathrm{Conn}(e; \Xi)$ and define

$$\overline{\mathrm{Conn}}(e; \Xi) = \mathrm{Conn}(e; \Xi) \cup \partial \mathrm{Conn}(e; \Xi) . \tag{5.3}$$

We set $\overline{\mathrm{Conn}}(e; \Xi) = \emptyset$ if $\Xi(e) = 0$. Given the above notations, we now define IP for the FK-dynamics. It would be notationally convenient to define $\tau_{i+1/2} = (i + 1/2)\Delta$, for $i \in \mathbb{N}$. Furthermore to distinguish between edges (elements of E_n) and connections across time, we will call the former 'space edges' as just edges and the latter as 'time edges' (see Figure 5.1 for an illustration).

Definition 5.3 (Information percolation)**.** The information percolation cluster is defined on the space-time slab $E_n \times [\tau_1, \tau_m]$ for some fixed $m \geq 2$. For an edge $e \in E_n$, we define the history $\mathscr{H}_e = (\mathscr{H}_e(t))_{t \in [\tau_1, \tau_m]}$ associated to the edge e backward in time recursively as follows: Start by setting $\mathscr{H}_e(\tau_m) = \{e\}$.

  (1) For each $t = \tau_{i+1}$ with $i \in [\![1, m-1]\!]$, suppose that $\mathscr{H}_e(\tau_{i+1})$ is given by a subset of E_n. Then we let $\mathscr{H}_e(\tau_{i+1/2})$ be the same as $\mathscr{H}_e(\tau_{i+1})$, as well as for any $w \in \mathscr{H}_e(\tau_{i+1})$ we connect the two edges (w, τ_{i+1}) and $(w, \tau_{i+1/2})$, by a 'time edge' in the time direction (see Figure 5.1.)

  (2) For each $w \in \mathscr{H}_e(\tau_{i+1/2})$, we check if it has been updated in the time interval (τ_i, τ_{i+1}) (recall the various notations from Table 1).

    (a) If $\mathscr{N}_i(e) = 1$, then introduce the 'space edge' (w, τ_i) and connect $(w, \tau_{i+1/2})$ and (w, τ_i) by a time edge.

    (b) If $\mathscr{N}_i(e) = 0$, we take the last update (t_0, U_e) for e in $(\tau_i, \tau_{i+1}]$.

      (i) If $U_e < 1 - p + p^*$, this update is called **oblivious** and we do not take any action on the edge $(w, \tau_{i+1/2})$.

(ii) If $U_e > 1 - p + p^*$, then w is open in $\overline{\mathscr{e}}_i$, and hence is open in Ξ_i as well (cf. (4.7)). In this case, we include all the edges in $\overline{\mathrm{Conn}}(w; \Xi_{i-1} \cup \Xi_i)$ in $\mathscr{H}_e(\tau_{i+1/2})$ and $\mathscr{H}_e(\tau_i)$. Finally we connect the space edges $(w', \tau_{i+1/2})$ and (w', τ_i) for all the edges w' in $\overline{\mathrm{Conn}}(w; \Xi_{i-1} \cup \Xi_i)$ using time edges.

(3) Steps (1) and (2) above define $\mathscr{H}_e(\tau_i)$ as a subset of E_n. Now return to the first step if $i \geq 2$ to use the above construction recursively.

For $A \subset E_n$, define $\mathscr{H}_A = (\mathscr{H}_A(t))_{t \in [\tau_1, \tau_m]}$ as $\mathscr{H}_A = \bigcup_{e \in A} \mathscr{H}_e$. Two histories $\mathscr{H}_e$ and $\mathscr{H}_{e'}$ are connected if they share an edge.

Some remarks are in order. First, we emphasize that two histories $\mathscr{H}_e$ and $\mathscr{H}_{e'}$ are regarded as two disconnected pieces if they share vertices only. Second, by the construction rule, one can observe that:

$$\mathscr{H}_e(\tau_{i+1/2}) = \mathscr{H}_e(\tau_{i+1}) \cup \mathscr{H}_e(\tau_i) . \tag{5.4}$$

Using terminology from existing literature we will often refer to the collection $\mathscr{H} := \{\mathscr{H}_e\}_{e \in E_n}$ as the history diagram. This induces a new graph structure on E_n. i.e. e and e' are connected if $\mathscr{H}_e$ and $\mathscr{H}_{e'}$ are connected. Note that the vertex set for this graph is E_n.

With the above conventions, each connected component of this new graph is called an **_information percolation cluster._** We shall simply refer to them as *clusters*. Let them be indexed by the set $\mathcal{C}$.

Definition 5.4 (IP clusters and their colors)**.** Each cluster $C \in \mathcal{C}$ is colored red, blue or green according to the following rule:

- Colored **red** if $\mathscr{H}_C(\tau_1) \neq \emptyset$.
- Colored **blue** if $\mathscr{H}_C(\tau_1) = \emptyset$ and $|C| = 1$.
- Colored **green** if $\mathscr{H}_C(\tau_1) = \emptyset$ and $|C| \geq 2$.

Denote by $\mathcal{C}_\mathcal{R}$, $\mathcal{C}_\mathcal{B}$ and $\mathcal{C}_\mathcal{G}$ the collection of red, blue and green clusters, respectively. Define

$$E_\mathcal{R} = \{e : e \in C \text{ for some } C \in \mathcal{C}_\mathcal{R}\} \tag{5.5}$$

and define $E_\mathcal{B}$ and $E_\mathcal{G}$ similarly. We use the following simplified notations to denote the history diagrams emanating from the various colored edges:

$$\mathscr{H}_\mathcal{R} := \mathscr{H}_{E_\mathcal{R}} , \quad \mathscr{H}_\mathcal{B} := \mathscr{H}_{E_\mathcal{B}} , \text{ and } \mathscr{H}_\mathcal{G} := \mathscr{H}_{E_\mathcal{G}} .$$

The following theorem justifies the above definitions. In short, it says that to reconstruct the state of the edges in $\mathscr{H}_A(\tau_{i+1})$, all one needs is the update sequence and the state of the edges $\mathscr{H}_A(\tau_i)$ at time τ_i provided that A is a cluster.

Theorem 5.5. *Given a history diagram $\mathscr{H}$, suppose that a set $A \subset E_n$ is a cluster. Then, for each $i \in [\![1, m-1]\!]$, the configuration $X_{\tau_{i+1}}(\mathscr{H}_A(\tau_{i+1}))$ is a deterministic function of*

$$X_{\tau_i}(\mathscr{H}_A(\tau_i)) \quad and \quad \bigcup_{e \in \mathscr{H}_A(\tau_{i+1/2})} \mathrm{Upd}[\tau_i, \tau_{i+1}](e) . \tag{5.6}$$

In particular, if $\mathscr{H}_A(\tau_i) = \emptyset$, for some $i \geq 1$, then $X_{\tau_{i+1}}$ is independent of X_{τ_i} and therefore of X_{τ_1}.

Remark 5.6. Note that not all update sequences are compatible with the diagram $\mathscr{H}$. In particular, the inner boundary of Green cluster is always closed and hence any update sequence for which the diagram occurs with positive probability must respect such constraints.

The proof of the above theorem is provided below after introducing some notations and observing some consequences of the already stated definitions. We momentarily fix $A \subset E_n$ and suppressing the dependence on A, define

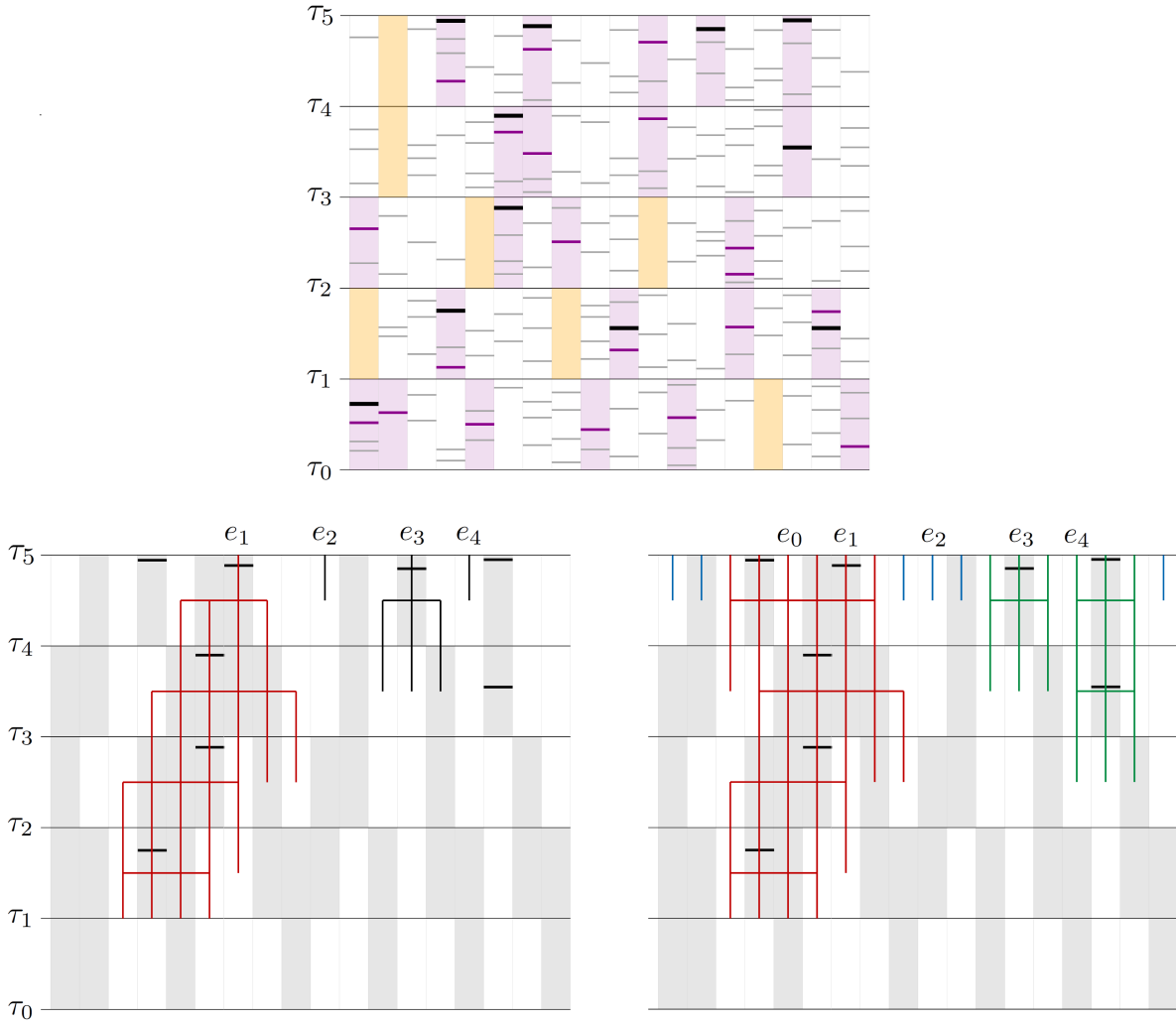$$W_j = \mathscr{H}_A(\tau_j) ; \ j \in [\![1, m]\!] . \tag{5.7}$$

FIGURE 5.1. **(Up)** The various colors indicate the values of the uniform variables for each update: gray $\leftrightarrow \{U < 1 - p\}$, purple $\leftrightarrow \{1 - p \leq U \leq 1 - p + p^*\}$, black $\leftrightarrow \{U > 1 - p + p^*\}$. The purple region denotes $\overline{\mathscr{E}}_i(e) = 1$, while the yellow region implies that $\mathscr{N}_i(e) = 1$. **(Down)** In the two graphs, the gray region indicates whether $\Xi_i \cup \Xi_{i-1}(e)$ is 1 or 0 with gray indicating the former. **(Down-left)** History diagrams for e_1, e_2, e_3, e_4. We can assert that $\mathscr{H}_{e_1}$ is red, but not able to say anything about the remaining ones; **(Down-right)** History diagram for e_0 is combined with that of e_1. e_4 belongs to green cluster although its last update is oblivious. The vertical edges acting as connections across time are referred to as 'time edges' in the article.

In the proof of the main result of this section (i.e., Theorem 5.1), the key ingredient is the analysis of the evolution of $|W_j|$ backwards in time. This is formulated in Proposition 5.19. A crucial role is played by the following two decompositions of W_j. The first decomposition is according to the type of evolution that occurs in the time interval $[\tau_{j-1}, \tau_j]$:

$$W_j = W_j^{\mathrm{NU}} \cup W_j^{\mathrm{Ob}} \cup W_j^{\mathrm{NOb}} , \tag{5.8}$$

where,

$$W_j^{\mathrm{NU}} = \{e \in W_j : \mathscr{N}_{j-1}(e) = 1\} \text{ i.e., the edges that have not been updated in } [\tau_{j-1}, \tau_j] \,,$$

$$W_j^{\mathrm{Ob}} = \{e \in W_j : \mathscr{N}_{j-1}(e) = 0 \text{ and the last update for } e \text{ in } [\tau_{j-1}, \tau_j] \text{ is oblivious}\} \,,$$

$$W_j^{\mathrm{NOb}} = \{e \in W_j : \mathscr{N}_{j-1}(e) = 0 \text{ and the last update for } e \text{ in } [\tau_{j-1}, \tau_j] \text{ is non-oblivious}\} \,.$$

Now, we consider the second decomposition of W_j. For this, we classify each edge according to the origin of its evolution in $[\tau_j, \tau_{j+1}]$. For each $j \in [\![1, m-1]\!]$, we write

$$C_j = \bigcup_{e \in W_{j+1}^{\mathrm{NOb}}} \overline{\mathrm{Conn}}(e; \Xi_{j-1} \cup \Xi_j) \,. \tag{5.9}$$

Hence, the set C_j represents the collection of edges in W_j that arise from non-oblivious expansions (i.e., step (2)-(b)-(ii) of Definition 5.3). Since each edge in W_j is either due to such an expansion or is inherited from W_{j+1} owing to no update at the corresponding edge in the time interval $[\tau_j, \tau_{j+1}]$, we obtain that

$$W_j = C_j \cup W_{j+1}^{\mathrm{NU}} \,. \tag{5.10}$$

Therefore, by writing

$$N_j = W_j \setminus C_j \,, \tag{5.11}$$

we obtain another decomposition of W_j given by

$$W_j = C_j \cup N_j \,. \tag{5.12}$$

We next record some basic properties of these decompositions.

Lemma 5.7. *For all $j \in [\![1, m-1]\!]$, it holds that*

$$W_{j+1}^{NOb} \subset C_j \ \text{ and } \ N_j \subset W_{j+1}^{NU} \,.$$

Proof. For the first inclusion, we note that $e \in W_{j+1}^{\mathrm{NOb}}$ implies that $\overline{\mathscr{E}}_j(e) = 1$ and thus $\Xi_j(e) = 1$. Hence, the definition (5.9) indicates that $e \in C_j$ as well and thus the first inclusion trivially holds. For the latter one, it suffices to recall (5.10) and the definition (5.11) of N_j. □

For $S \subset E_n$, define $\partial^- S$ as the set of edges in S which are adjacent to at least one edge in S^c, i.e., $\partial^- S = \partial(E \setminus S)$. We record the following simple fact.

Lemma 5.8. *For all $j \in [\![1, m-1]\!]$, all the edges in $\partial^- C_j$ are closed in $\Xi_{j-1} \cup \Xi_j$. In particular, there is no open path in $\Xi_{j-1} \cup \Xi_j$ connecting an open edge in C_j and an edge in N_j.*

Proof. The proof is direct from the definition of C_j where we included the closed (outer) boundary of $\mathrm{Conn}(e; \Xi_{j-1} \cup \Xi_j)$. □

Lemma 5.9. *For all $j \in [\![1, m-1]\!]$, for each $e \in C_j$, and for all $t \in [\tau_j, \tau_{j+1}]$, the process $X_t(e)$ is a deterministic function of*

$$X_{\tau_j}(\mathscr{H}_{C_j}(\tau_j)) \ \text{ and } \ \bigcup_{e' \in C_j} \mathrm{Upd}[\tau_j, t](e') \,.$$

Proof. Let

$$\mathcal{U}_t = \bigcup_{e' \in C_j} \mathrm{Upd}[\tau_j, t](e') \ ; \ t \in [\tau_j, \tau_{j+1}] \,.$$

We fix $e \in C_j$ and $t \in [\tau_j, \tau_{j+1}]$ and denote by (t_0, U_0) the last update for e in $[\tau_j, t]$. If $U_0 < 1-p+p^*$ then, in view of Definition 2.1, the configuration $X_t(e)$ is 1 if $U_0 < 1-p$, and 0 if $U_0 \geq 1-p$. Thus, we can determine $X_t(e)$ solely in terms of $(t_0, U_0) \in \mathrm{Upd}[\tau_j, t](e) \subset \mathcal{U}_t$. Now we consider the case $U_0 > 1-p+p^*$. In this case, the configuration $X_t(e) = X_{t_0}(e)$ is determined by checking whether e is a cut-edge or not in the configuration X_{t_0-}. In order to check this, one has to investigate

$\mathrm{Conn}(e; X_{t_0-} \cup \{e\})$ to determine whether removing e disconnects some component of $X_{t_0-} \cup \{e\}$ or not. Note that e is open in $\overline{\mathscr{E}}_i$ (and hence in Ξ_i) since $U_0 > 1 - p + p^* > 1 - p$. Thus, by Proposition 4.8, we have

$$\mathrm{Conn}(e; X_{t_0-} \cup \{e\}) \subset \mathrm{Conn}(e; \Xi_j \cup \Xi_{j+1}) \subset C_j \setminus \partial^- C_j . \tag{5.13}$$

Therefore, we can determine $X_t(e)$ in terms of $X_{t_0-}(C_j)$ and $(t_0, U_0) \in \mathcal{U}_t$.

If $\mathcal{U}_{t_0} = \{(t_0, U_0)\}$, we have $X_{t_0-}(C_j) = X_{\tau_j}(C_j)$, so we can conclude the proof. Otherwise, we take the last update (t_1, U_1) in $\mathcal{U}_{t_0}$ other than (t_0, U_0). Then, we have,

$$X_{t_0-}(C_j) = X_{t_1}(C_j) .$$

Since there are finitely many updates in $[\tau_j, \tau_{j+1}]$ almost surely, we can repeat this procedure to finish the proof. An important fact implicitly used above is that in repeating the argument all the edges $\tilde{e}$ that we encounter with an update time $\tilde{t} \in [\tau_j, t_0]$ has the property that the connected component of

$$X_{\tilde{t}-}(\tilde{e}) \subset \overline{\mathrm{Conn}}(e; \Xi_{j-1} \cup \Xi_j) \subset C_j .$$

since the edge boundary of $\mathrm{Conn}(e; \Xi_{j-1} \cup \Xi_j)$ remains closed throughout the interval $[\tau_j, \tau_{j+1}]$. $\quad\square$

The proof of Theorem 5.5 now follows.

Proof of Theorem 5.5. In view of (5.8) and the first inclusion of Lemma 5.7, it suffices to consider the following three cases separately.

- Case 1: $e \in W_{i+1}^{\mathrm{NU}}$. By (2)-(a) of Definition 5.3, we have $X_{\tau_{i+1}}(e) = X_{\tau_i}(e)$ and thus configuration of $X_{\tau_{i+1}}(e)$ is determined by $X_{\tau_i}(W_{i+1}^{\mathrm{NU}})$. Since $W_{i+1}^{\mathrm{NU}} \subset W_i$ (cf. (5.10)), the proposition holds for this case.
- Case 2: $e \in W_{i+1}^{\mathrm{Ob}} \setminus C_i$. By (2)-(b)-(i) of Definition 5.3, the configuration $X_{\tau_{i+1}}(e)$ is solely determined by the last update for e in $(\tau_i, \tau_{i+1}]$ and therefore the proposition holds as well.
- Case 3: $e \in W_{i+1} \cap C_i$. This case is immediate from Lemma 5.9.

$\quad\square$

The following corollary is an immediate consequence of the previous theorem.

Corollary 5.10. *Given a history diagram $\mathscr{H}$, the following holds.*
1. *The configurations $X_{\tau_m}(E_{\mathcal{G}})$ and $X_{\tau_m}(E_n \setminus E_{\mathcal{G}})$ are independent.*
2. *The configuration $X_{\tau_m}(E_{\mathcal{G}})$ is independent of X_{τ_1}.*
3. *For $e \in E_{\mathcal{B}}$, the distribution of $X_{\tau_m}(e)$ is a Bernoulli random variable with parameter $\frac{p^*}{1-p+p^*}$, and is independent of all other randomness.*

Proof. Parts (1) and (2) are direct consequences of Theorem 5.5 and the definition of a green cluster. We now consider part (3). For $e \in E_{\mathcal{B}}$, the configuration $X_{\tau_m}(e)$ is determined by the last update (t, U) for e in $[\tau_1, \tau_m]$. Furthermore, since $e \in E_{\mathcal{B}}$, this last update is oblivious and therefore we know that $U < 1 - p + p^*$. Given this condition, we have $X_{\tau_m}(e) = 1$ if $U < 1 - p$ and $X_{\tau_m}(e) = 0$ if $U \in [1 - p, 1 - p + p^*]$ otherwise. This finishes the proof of part (3). $\quad\square$

For each $A \subset E_n$, define

$$\mathscr{H}_A^- = \mathscr{H}_{E_n \setminus A} .$$

As in [27, 24], it would be crucial to estimate the probability of A being a red cluster or a collection of singleton blue clusters i.e.,

$$\{A \in \mathcal{C}_{\mathcal{R}}\} \cup \{A \subset E_{\mathcal{B}}\} . \tag{5.14}$$

Furthermore, technical aspects make it important to estimate the above probabilities conditioned on the history diagram of the complement of A. For this conditional probability to be non-zero a necessary condition is that,

$$\mathscr{H}_A^- \cap \{A \times \{t = \tau_{m-1/2}\}\} = \emptyset; , \tag{5.15}$$

for the following reason. Suppose that $e \in A$ satisfies $(e, \tau_{m-1/2}) \in \mathscr{H}_{e'}$ for some $e' \in E_n \setminus A$. Then, by the definition of the information percolation cluster, the cluster containing e must contain e' as well.

Thus this is a compatibility condition to guarantee that $\{A \in \mathcal{C}_{\mathcal{R}}\} \cup \{A \subset E_{\mathcal{B}}\}$ is a non-empty event which we denote by $\mathscr{H}_A^- \in \mathscr{H}_{\text{com}}(A)$. Given this, we define

$$\mathcal{P}_A = \sup_{\mathscr{H}_A^- \in \mathscr{H}_{\text{com}}(A)} \mathbb{P}\left[A \in \mathcal{C}_{\mathcal{R}} \,\middle|\, \mathscr{H}_A^-, \{A \in \mathcal{C}_{\mathcal{R}}\} \cup \{A \subset E_{\mathcal{B}}\}\right], \tag{5.16}$$

i.e., the maximum probability of A being a red cluster conditioned on a compatible $\mathscr{H}_A^-$. Given the above preparation, the following proposition is the main estimate (similar to [27, Lemma 4.8]) needed. For $A \subset E_n$, we denote by $|\text{Conn}(A)|$, the smallest number of edges in any connected subgraph of (Λ_n, E_n) containing A.

Proposition 5.11. *For any $\theta > 0$, we can find two constants $C = C(\theta) > 0$ and $p_0 = p_0(\theta) > 0$ such that, for any $p \in (0, p_0)$, there exists a constant $\alpha = \alpha(p) > 0$ satisfying*

$$\mathcal{P}_A \leq Ce^{-(\theta|\text{Conn}(A)|+\alpha\tau_m)} \quad \textit{for all } A \subset E_n \, .$$

A notable feature of this proposition is the fact that α is independent of θ. In the remaining part of the current section, α always refers to the constant above. The proof of this proposition is postponed to Section 5.3. A corollary of this proposition is the following lemma which lower bounds the probability that there are no red clusters.

Lemma 5.12. *For all small enough p, there exists a constant $C = C(p) > 0$ satisfying*

$$\sup_{\mathscr{H}_{\mathcal{G}}} \mathbb{P}\left[\mathscr{H}_{\mathcal{R}} = \emptyset | \mathscr{H}_{\mathcal{G}}\right] \geq 1 - Cn^2 e^{-\alpha\tau_m} \, .$$

Proof. By the union bound and the definition of $\mathcal{P}_A$,

$$1 - \mathbb{P}\left[\mathscr{H}_{\mathcal{R}} = \emptyset | \mathscr{H}_{\mathcal{G}}\right] \leq \sum_{A \subset E_n, \, A \neq \emptyset} \mathbb{P}\left[A \in \mathcal{C}_{\mathcal{R}} | \mathscr{H}_{\mathcal{G}}\right] \leq \sum_{A \subset E_n, \, A \neq \emptyset} \mathcal{P}_A \, .$$

Now, by Proposition 5.11 and the translation invariance of the periodic lattice,

$$\sum_{A \subset E_n, \, A \neq \emptyset} \mathcal{P}_A \leq \sum_{e \in E_n} \sum_{A : A \ni e} \mathcal{P}_A \leq Cn^2 e^{-\alpha\tau_m} \sum_{k=1}^{\infty} \sum_{A : A \ni e, \, |\text{Conn}(A)|=k} e^{-\theta k} \, . \tag{5.17}$$

For a fixed $e \in E_n$, we have that

$$|\{A \subset E_n : A \ni e, \, |\text{Conn}(A)| = k\}| \leq (k+1)(8d^2)^k \, .$$

The verification is elementary and we leave the proof to the reader. Finally, we can combine the last two displays to deduce

$$\mathbb{P}\left[\mathscr{H}_{\mathcal{R}} = \emptyset | \mathscr{H}_{\mathcal{G}}\right] \geq 1 - Cn^2 e^{-\alpha\tau_m} \sum_{k=1}^{\infty} (k+1)(8d^2 e^{-\theta})^k \, . \tag{5.18}$$

Now by taking θ large enough so that $8d^2 e^{-\theta} < 1/2$, the proof of the lemma is complete. $\qquad \square$

The remainder of the section is now devoted to proving (5.2).

5.2. **Proof of Proposition 5.2.** For $A \subset E_n$, define ν_A as a Bernoulli percolation measure on A with open probability $\overline{p}$ where

$$\overline{p} = \frac{p^*}{1 - p + p^*} \,. \tag{5.19}$$

In the remaining part of the section, we will simply write $\mu := \mu_{p,q}^n$, $E := E_n$ and denote by μ_A, $A \subset E$, the projection of μ on A. We first prove the following lemma which shows that the L^2-distance to μ can be controlled by the L^2-distance of the measure on the complement of the green clusters to the measure ν.

Lemma 5.13. *For all small enough p, we can find $C = C(p) > 0$ such that for $m \geq C \log n$ we have*

$$\left\| \mathbb{P}_{x_0} \left[X_t \in \cdot \right] - \mu \right\|_{L^2(\mu)} \leq 2 \sup_{\mathscr{H}_{\mathcal{G}}} \left\| \mathbb{P}_{x_0} \left[X_t(E \setminus E_{\mathcal{G}}) \in \cdot \,|\, \mathscr{H}_{\mathcal{G}} \right] - \nu_{E \setminus E_{\mathcal{G}}} \right\|_{L^2(\nu_{E \setminus E_{\mathcal{G}}})} + 1$$

for all $x_0 \in \Omega_n$,

Proof. Consider two copies of FK-dynamics (X_t) and (Y_t) where $X_0 = x_0$ and Y_0 is distributed according to μ. We couple them via the monotone coupling introduced in Definition 2.1. Now by Jensen's inequality (for details see [27, Lemma 4.13]) we obtain

$$\left\| \mathbb{P}_{x_0} \left[X_t \in \cdot \right] - \mu \right\|_{L^2(\mu)}^2 = \left\| \mathbb{P}_{x_0} \left[X_t \in \cdot \right] - \mathbb{P}_{\mu} \left[Y_t \in \cdot \right] \right\|_{L^2(\mu)}^2$$

$$\leq \int \left\| \mathbb{P}_{x_0} \left[X_t \in \cdot \,|\, \mathscr{H}_{\mathcal{G}} \right] - \mathbb{P}_{\mu} \left[Y_t \in \cdot \,|\, \mathscr{H}_{\mathcal{G}} \right] \right\|_{L^2(\mu_{p,q}^n(\cdot \,|\, \mathscr{H}_{\mathcal{G}}))}^2 d\mathbb{P}(\mathscr{H}_{\mathcal{G}})$$

$$\leq \sup_{\mathscr{H}_{\mathcal{G}}} \left\| \mathbb{P}_{x_0} \left[X_t \in \cdot \,|\, \mathscr{H}_{\mathcal{G}} \right] - \mathbb{P}_{\mu} \left[Y_t \in \cdot \,|\, \mathscr{H}_{\mathcal{G}} \right] \right\|_{L^2(\mu_{p,q}^n(\cdot \,|\, \mathscr{H}_{\mathcal{G}}))}^2 \,. \tag{5.20}$$

Given $\mathscr{H}_{\mathcal{G}}$, the diagram $\mathscr{H}_{E \setminus E_{\mathcal{G}}}$ is disjoint from $\mathscr{H}_{\mathcal{G}} = \mathscr{H}_{E_{\mathcal{G}}}$, and as we noticed in Corollary 5.10 configurations $X_t(E_{\mathcal{G}})$ (resp. $Y_t(E_{\mathcal{G}})$) and $X_t(E \setminus E_{\mathcal{G}})$ (resp. $Y_t(E \setminus E_{\mathcal{G}})$) are independent. Moreover, $Y_t(E_{\mathcal{G}})$ and $X_t(E_{\mathcal{G}})$ are identical by Theorem 5.5. Thus, the projection onto $E \setminus E_{\mathcal{G}}$ does not change the L^2-norm. Combining this observation with (5.20), we obtain

$$\left\| \mathbb{P}_{x_0} \left[X_{\tau_m} \in \cdot \right] - \mu \right\|_{L^2(\mu)}$$
$$\leq \sup_{\mathscr{H}_{\mathcal{G}}} \left\| \mathbb{P}_{x_0} \left[X_{\tau_m}(E \setminus E_{\mathcal{G}}) \in \cdot \,|\, \mathscr{H}_{\mathcal{G}} \right] - \mathbb{P}_{\mu} \left[Y_{\tau_m}(E \setminus E_{\mathcal{G}}) \in \cdot \,|\, \mathscr{H}_{\mathcal{G}} \right] \right\|_{L^2(\mu_{E \setminus E_{\mathcal{G}}}(\cdot \,|\, \mathscr{H}_{\mathcal{G}}))} \,. \tag{5.21}$$

Now by Lemma 5.12, for $m \geq C \log n$ where $C = C(p)$ is large enough,

$$\mathbb{P} \left[\mathscr{H}_{\mathcal{R}} = \emptyset \,|\, \mathscr{H}_{\mathcal{G}} \right] \geq \frac{1}{2} \,.$$

Then, for all $Z \subset \{0, 1\}^{E \setminus E_{\mathcal{G}}}$, we can deduce that,

$$\mathbb{P}_{\mu} \left[Y_{\tau_m}(E \setminus E_{\mathcal{G}}) = Z \,|\, \mathscr{H}_{\mathcal{G}} \right] \geq \mathbb{P} \left[\mathscr{H}_{\mathcal{R}} = \emptyset \,|\, \mathscr{H}_{\mathcal{G}} \right] \nu_{E \setminus E_{\mathcal{G}}}(Z) \geq \frac{1}{2} \nu_{E \setminus E_{\mathcal{G}}}(Z) \,. \tag{5.22}$$

Note that the first inequality follows from the fact that the distribution on $E_{\mathcal{B}}$ is $\nu_{E_{\mathcal{B}}}$, and that under $\mathscr{H}_{\mathcal{R}} = \emptyset$, we have $E \setminus E_{\mathcal{G}} = E_{\mathcal{B}}$. We are now able to complete the proof of the lemma by combining (5.21), (5.22), and the definition of L^2-norm. $\quad\square$

Thus the task has now been reduced to measuring the L^2-distance of certain measures to the product measure ν. The Miller-Peres inequality establishes a simple yet extremely useful bound for such cases. It first appeared in [25] where the product measure was given by independent $Ber(1/2)$ variables. This was extended later in [23, Lemma 4.3] which is the version we will use.

Lemma 5.14. *Let $\Omega = \{0, 1\}^S$ for a finite set S, and let η be a probability measure on the space of subsets of S. For each $R \subset S$, suppose that a probability measure φ_R on $\{0, 1\}^R$ is given. For $p \in (0, 1/2)$, denote by ν_p the measure on $\{0, 1\}^S$ given by the product of independent $Ber(p)$ variables. Let μ_p be a measure on Ω obtained first by sampling a subset R of S according to η,*

and then sampling an element of $\{0,1\}^R$ according to φ_R, and sampling an element of $\{0,1\}^{S \setminus R}$ according to the restriction of ν_p on $\{0,1\}^{S \setminus R}$. Then, we have

$$\|\mu_p - \nu_p\|_{L^2(\nu_p)}^2 \leq \mathbb{E}\left[p^{-|R \cap R'|}\right] - 1 \, ,$$

where R, $R' \subset S$ are two independent samples of η.

In view of Lemmas 5.13 and 5.14, we obtain that

$$\|\mathbb{P}_{x_0}[X_{\tau_m} \in \cdot] - \mu\|_{L^2(\mu)} \leq 2 \sup_{\mathscr{H}_{\mathcal{G}}} \mathbb{E}\left[\frac{1}{\overline{p}^{|E_{\mathcal{R}} \cap E_{\mathcal{R}'}|}} \,\Big|\, \mathscr{H}_{\mathcal{G}}\right] + 1 \, , \tag{5.23}$$

provided that p is small enough so that $\overline{p} < 1/2$, for all $m > C_1 \log n$ where C_1 is the constant in Lemma 5.13 and $E_{\mathcal{R}}$ and $E_{\mathcal{R}'}$, are two independent samples of the set $E_{\mathcal{R}}$ of red clusters (see (5.5)) conditioned on $\mathscr{H}_{\mathcal{G}}$. To analyze the right-hand side of (5.23), we recall the following domination results from [22, 23]. Let $\{J_A : A \subset E\}$ be a family of independent indicators such that $\mathbb{P}(J_A = 1) = \mathcal{P}_A$ for all $A \subset E$ and similarly let $\{J_{A,A'} : A, A' \subset E\}$ be a family of independent indicators such that $\mathbb{P}(J_{A,A'} = 1) = \mathcal{P}_A \mathcal{P}_{A'}$ for all $A, A' \subset E$.

Lemma 5.15 ([22], Lemma 2.3, Corollary 2.4). *Then following coupling results hold.*

(1) *The conditional distribution of red clusters given $\mathscr{H}_{\mathcal{G}}$ can be coupled to J_A such that*

$$\{A : A \in \mathcal{C}_{\mathcal{R}}\} \subset \{A : J_A = 1\} \, .$$

(2) *Similarly, the conditional distribution of $(E_{\mathcal{R}}, E_{\mathcal{R}}')$ given $\mathscr{H}_{\mathcal{G}}$ can be coupled such that*

$$|E_{\mathcal{R}} \cap E_{\mathcal{R}'}| \leq \sum_{A \cap A' \neq \emptyset} |A \cup A'| J_{A,A'} \, .$$

We are now ready to prove Proposition 5.2.

Proof of Proposition 5.2. It suffices to prove that the right-hand side of (5.23) is bounded by 2 for $m = C \log n$ with large enough C. Write $\kappa := \log(1/\overline{p}) > 0$. By part (2) of Lemma 5.15, we have

$$\sup_{\mathscr{H}_{\mathcal{G}}} \mathbb{E}\left[\overline{p}^{-|E_{\mathcal{R}} \cap E_{\mathcal{R}'}|} \,\big|\, \mathscr{H}_{\mathcal{G}}\right] \leq \mathbb{E}\exp\left\{\kappa \sum_{A \cap A' \neq \emptyset} |A \cup A'| J_{A,A'}\right\}$$

$$= \prod_{A \cap A' \neq \emptyset} \mathbb{E}\exp\left\{\kappa |A \cup A'| J_{A,A'}\right\}$$

$$\leq \prod_{e \in E_n} \prod_{(A,A'):e \in A, e \in A'} \left[(e^{\kappa(|A|+|A'|)} - 1)\mathcal{P}_A \mathcal{P}_{A'} + 1\right]$$

$$\leq \exp\left\{|E|\Big[\sum_{A:e \in A} e^{\kappa|A|} \mathcal{P}_A\Big]^2\right\} \, ,$$

where e in the last line is an arbitrary edge in E. The last inequality follows from $x + 1 \leq e^x$ and the translation invariance of the underlying graph. Hence, it suffice to show that

$$\sum_{A:e \in A} e^{\kappa|A|} \mathcal{P}_A \leq \frac{1}{n^3}$$

for $m = C \log n$ with sufficiently large C. To this end, we recall Proposition 5.11 so that

$$\sum_{A:e \in A} e^{\kappa|A|} \mathcal{P}_A \leq C e^{-\alpha \tau_m} \sum_{A:e \in A} e^{\kappa|A| - \theta|\mathrm{Conn}(A)|} \leq C e^{-\alpha \tau_m} \sum_{A:e \in A} e^{(\kappa-\theta)|\mathrm{Conn}(A)|} \, .$$

Thus, we can proceed as in (5.17) and (5.18) to deduce that the last summation bounded by is bounded by 1, provided that θ is large enough. This finishes the proof. $\qquad\square$

5.3. **Proof of Proposition 5.11: domination by subcritical branching processes.** We now prove Proposition 5.11 to complete our discussion on Theorem 5.1. For $S \subset E$, define $\mathcal{C}^*_{\mathcal{R}(S)}$ to be the collection of red clusters that arises when exposing the joint histories of elements of S i.e., $\mathscr{H}_S$ only. Similarly define $\mathcal{C}^*_{\mathcal{B}(S)}$ for blue clusters.

Lemma 5.16. *There exists $c = c(p) > 0$ such that, for all $A \subset E$ we have*

$$\mathcal{P}_A \leq e^{c|\mathrm{Conn}(A)|}\, \mathbb{P}\left[A \in \mathcal{C}^*_{\mathcal{R}(A)} \right]\,. \tag{5.24}$$

To prove the above we will first attempt to understand the effect of conditioning on the event $\mathscr{H}^-_A = \mathcal{X} \in \mathscr{H}_{\mathrm{com}}(A)$. We will determine a subset of $\mathrm{Upd}[0, \tau_m]$ that is enough to determine the event $\{\mathscr{H}^-_A = \mathcal{X}\}$. We write $\mathcal{X}_i = \mathcal{X} \cap \{t = \tau_i\}$ for $i \in \{n/2 : n \in \mathbb{Z}\}$. Recall from (5.4) that the event $\{\mathscr{H}^-_A = \mathcal{X}\}$ is non-empty only when $\mathcal{X}$ satisfies the consistency condition

$$\mathcal{X}_{i+1/2} = \mathcal{X}_i \cup \mathcal{X}_{i+1} \text{ for all } i \in [\![0,\, m-1]\!]\,. \tag{5.25}$$

For each $i \in [\![1,\, m-1]\!]$ and $e \in E_n$, we define

$$\mathcal{U}_i(e) = \begin{cases} \mathrm{Upd}[\tau_{i-1},\, \tau_{i+1}](e) & \text{if } e \in \mathcal{X}_i\,, \\ \mathrm{Upd}[\tau_i,\, \tau_{i+1}](e) & \text{if } e \in \mathcal{X}_{i+1} \setminus \mathcal{X}_i\,, \\ \emptyset & \text{otherwise.} \end{cases} \tag{5.26}$$

Then, we define

$$\mathcal{U}_i = \bigcup_{e \in E_n} \mathcal{U}_i(e) \quad \text{and} \quad \mathcal{U} = \bigcup_{i=1}^{m-1} \mathcal{U}_i\,.$$

Note that $\mathcal{U}$ depends on $\mathcal{X}$.

Lemma 5.17. *The event $\{\mathscr{H}^-_A = \mathcal{X}\}$ is independent of the update variables not in $\mathcal{U}$.*

Proof. Write $\mathcal{Y}_i = \mathscr{H}^-_A \cap \{t = \tau_i\}$ and define the event $\mathcal{E}_i$ by

$$\mathcal{E}_i = \{\mathcal{Y}_i = \mathcal{X}_i\}\,.$$

If $\mathcal{X}$ satisfies the condition (5.25), we can write

$$\{\mathscr{H}^-_A = \mathcal{X}\} = \bigcap_{i=1}^{m} \mathcal{E}_i\,.$$

We claim that given $\mathcal{E}_{i+1}$, the event $\mathcal{E}_i$ depends only on the events in $\mathcal{U}_i$. Given $\mathcal{E}_{i+1}$, we decompose $\mathcal{Y}_{i+1} = \mathcal{X}_{i+1}$ as following (similar to those in Theorem 5.5):

$$\mathcal{Y}^{\mathrm{NU}}_{i+1} = \{e \in \mathcal{Y}_{i+1} : \mathcal{N}_i(e) = 1\}\,,$$
$$\mathcal{Y}^{\mathrm{Ob}}_{i+1} = \{e \in \mathcal{Y}_{i+1} : \mathcal{N}_i(e) = 0 \text{ and the last update for } e \text{ in } [\tau_i,\, \tau_{i+1}] \text{ is oblivious}\}\,,$$
$$\mathcal{Y}^{\mathrm{NOb}}_{i+1} = \{e \in \mathcal{Y}_{i+1} : \mathcal{N}_i(e) = 1 \text{ and the last update for } e \text{ in } [\tau_i,\, \tau_{i+1}] \text{ is non-oblivious}\}\,.$$

This classification can be carried out if we only know

$$\bigcup_{e \in \mathcal{Y}_{i+1}} \mathrm{Upd}[\tau_i,\, \tau_{i+1}](e) \subset \mathcal{U}_i\,.$$

Now we suppose that this classification is given. Then, we have

$$\mathcal{Y}_i = \mathcal{Y}^{\mathrm{NU}}_{i+1} \cup \bigcup_{e \in \mathcal{Y}^{\mathrm{NOb}}_{i+1}} \overline{\mathrm{Conn}}(e; \Xi_{i-1} \cup \Xi_i)\,,$$

and therefore $\mathcal{X}_i = \mathcal{Y}_i$ holds if

$$\mathcal{X}_i \setminus \mathcal{Y}_{i+1}^{\mathrm{NU}} \subset \bigcup_{e \in \mathcal{Y}_{i+1}^{\mathrm{NOb}}} \overline{\mathrm{Conn}}(e; \Xi_{i-1} \cup \Xi_i) \subset \mathcal{X}_i \ .$$

This event can be determined by knowing $\bigcup_{e \in \mathcal{X}_i} \mathcal{U}_i(e)$. This finishes the proof. $\square$

Lemma 5.18. *For all $\mathcal{X}$ satisfying* (5.25), *it holds that*

$$\mathbb{P}\left[A \in \mathcal{C}_{\mathcal{R}(A)}^*, \ \mathscr{H}_A \cap \mathcal{X} = \emptyset \ \middle| \ \mathscr{H}_A^- = \mathcal{X}\right] = \mathbb{P}\left[A \in \mathcal{C}_{\mathcal{R}(A)}^*, \ \mathscr{H}_A \cap \mathcal{X} = \emptyset\right] \ .$$

Proof. We keep the notation from the previous lemma. In view of the previous lemma, it suffices to demonstrate that the event

$$\mathcal{E} = \{A \in \mathcal{C}_{\mathcal{R}(A)}^*\} \cap \{\mathscr{H}_A \cap \mathcal{X} = \emptyset\}$$

does not depend on the updates in $\mathcal{U}$. Note that this event is the same as saying $\mathscr{H}_A$ reaches $t = \tau_1$ without touching $\mathcal{X}$. We prove this by induction (see Figure 5.2 for an illustration). Write $W_i(A) = \mathscr{H}_A \cap \{t = \tau_i\}$ and for each $i \in [\![2, m]\!]$, define the event $\mathcal{E}_i$ as

$$\mathcal{E}_i = \{W_{i-1}(A) \neq \emptyset \ \text{and} \ W_{i-1}(A) \cap (\mathcal{X}_{i-3/2} \cup \mathcal{X}_{i-1/2}) = \emptyset\} \ .$$
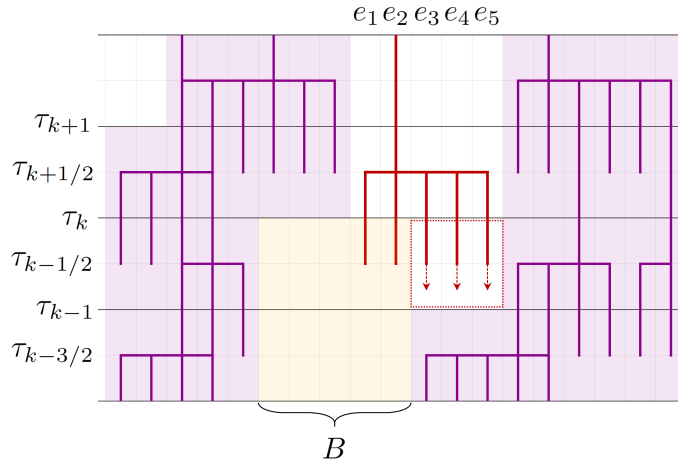

FIGURE 5.2. Illustrating the proofs of Lemmas 5.17 and 5.18. The purple graph is $\mathcal{X}$ and $\mathcal{U}$ is the set of updates in the purple region. The red graph is the history diagram $\mathscr{H}_A$. At time $t = \tau_k$, the occurrence of the event $\mathcal{E}_k$ does not depend on the updates in the purple region. Note that for the latter event to occur e_3, e_4, e_5 cannot hit the purple region and hence the last updates for each of them in $(\tau_{k-1}, \tau_k]$ should be oblivious. This depends on the updates in the red box. For e_1 and e_2, they can be expanded and one of them must be to ensure that they all together form a red cluster. However this expansion should be confined to B. This can be determined by the updates in yellow region and therefore also independent of updates in the purple region.

We suppose that $\mathcal{X}$ satisfies $\mathcal{X}_m \cap A = \emptyset$ and $\mathcal{X}_{m-1/2} \cap A = \emptyset$ since otherwise the event $\{\mathscr{H}_A \cap \mathcal{X} = \emptyset\}$ (and hence $\mathcal{E}$) cannot happen. Under this minimal consistency assumption, we can write $\mathcal{E} = \bigcap_{i=2}^m \mathcal{E}_i$.

We now claim that, for each $k \in [\![2, m-1]\!]$, given $\bigcap_{i=k+1}^m \mathcal{E}_i$, the event $\mathcal{E}_k$ does not depend on updates in $\mathcal{U}$. For each $e \in W_k(A)$, we consider two cases:

(1) $e \in \mathcal{X}_{k-3/2}$: The last update in $(\tau_{k-1}, \tau_k]$ must be oblivious to have $W_{k-1}(A)$ disjoint from $\mathcal{X}_{k-3/2}$. This update belongs to $\mathcal{U}$ only if $e \in \mathcal{X}_{k-1}$ or $e \in \mathcal{X}_k$, which cannot happen under $\mathcal{E}_{k+1}$, since under $\mathcal{E}_{k+1}$ the set $W_k(A)$ is disjoint to $\mathcal{X}_{k-1/2} = \mathcal{X}_{k-1} \cup \mathcal{X}_k$ (cf. (5.25)).

(2) $e \notin \mathcal{X}_{k-3/2}$: We still have two cases: either there is no update in $(\tau_{k-1}, \tau_k]$ for the edge e, or the last update in $(\tau_{k-1}, \tau_k]$ for the edge e is oblivious and

$$\text{Conn}(e; \Xi_{k-1} \cup \Xi_k) \bigcap (\mathcal{X}_{k-3/2} \cup \mathcal{X}_{k-1/2}) = \emptyset .$$

Determining whether this holds or not can be performed by looking only at

$$\bigcup_{e' \notin \mathcal{X}_{k-3/2} \cup \mathcal{X}_{k-1/2}} \text{Upd}[\tau_{k-2}, \tau_k](e') .$$

These updates are disjoint to $\mathcal{U}$ since $e' \notin \mathcal{X}_{k-3/2} \cup \mathcal{X}_{k-1/2}$ implies $\text{Upd}[\tau_{k-2}, \tau_k](e') \cap \mathcal{U} = \emptyset$ Furthermore, the non-emptiness of $W_{k-1}(A)$ implies that at least one of the last updates of $e \in W_k$ is non-oblivious or there is an edge $e \in W_k$ such that there is no update in $(\tau_{k-1}, \tau_k]$. By the same reasoning as (1), this is independent of the updates in $\mathcal{U}$. Summing up, for the event $\mathcal{E}_i$ to occur, all the events described above must occur simultaneously and the probability of this is independent of the conditioning on the randomness in $\mathcal{U}$. □

Proof of Lemma 5.16. Given the above preparation, the remaining steps of the proof already appears in [21, 27]. Note first that, conditioned on $\mathscr{H}_A^- = \mathcal{X} \in \mathscr{H}_{\text{com}}(A)$, one has $\{A \in \mathcal{C}_\mathcal{R}\} = \{A \in \mathcal{C}_{\mathcal{R}(A)}^*\} \cap \{\mathscr{H}_A \cap \mathcal{X} = \emptyset\}$ and similarly $\{A \subset \mathcal{C}_\mathcal{B}\} = \{A \subset \mathcal{C}_{\mathcal{B}(A)}^*\} \cap \{\mathscr{H}_A \cap \mathcal{X} = \emptyset\}$. Therefore, we can deduce

$$\mathbb{P}\left[A \in \mathcal{C}_\mathcal{R} \,\middle|\, \mathscr{H}_A^- = \mathcal{X}, \, \{A \in \mathcal{C}_\mathcal{R}\} \cup \{A \subset E_\mathcal{B}\}\right]$$

$$= \mathbb{P}\left[A \in \mathcal{C}_{\mathcal{R}(A)}^*, \, \mathscr{H}_A \cap \mathcal{X} = \emptyset \,\middle|\, \mathscr{H}_A^- = \mathcal{X}, \, \{A \in \mathcal{C}_\mathcal{R}\} \cup \{A \subset E_\mathcal{B}\}\right]$$

$$= \frac{\mathbb{P}\left[A \in \mathcal{C}_{\mathcal{R}(A)}^*, \, \mathscr{H}_A \cap \mathcal{X} = \emptyset \,\middle|\, \mathscr{H}_A^- = \mathcal{X}\right]}{\mathbb{P}\left[\{A \in \mathcal{C}_{\mathcal{R}(A)}^*\} \cup \{A \subset \mathcal{C}_{\mathcal{B}(A)}^*\}, \, \mathscr{H}_A \cap \mathcal{X} = \emptyset \,\middle|\, \mathscr{H}_A^- = \mathcal{X}\right]}$$

$$\leq \frac{\mathbb{P}\left[A \in \mathcal{C}_{\mathcal{R}(A)}^*, \, \mathscr{H}_A \cap \mathcal{X} = \emptyset\right]}{\mathbb{P}\left[A \subset \mathcal{C}_{\mathcal{B}(A)}^*, \, \mathscr{H}_A \cap \mathcal{X} = \emptyset \,\middle|\, \mathscr{H}_A^- = \mathcal{X}\right]} \quad \text{(by Lemma 5.18)}$$

$$\leq \frac{\mathbb{P}\left[A \in \mathcal{C}_{\mathcal{R}(A)}^*\right]}{\mathbb{P}\left[A \subset \mathcal{C}_{\mathcal{B}(A)}^*, \, \mathscr{H}_A \cap \mathcal{X} = \emptyset \,\middle|\, \mathscr{H}_A^- = \mathcal{X}\right]} . \tag{5.27}$$

Now we bound the denominator of (5.27) from below. Since $\mathcal{X}$ satisfies the compatibility condition (5.15) by hypothesis, an event which implies the event in the denominator is the following: all the edges in A are updated in the time interval $[\tau_{m-1/2}, \tau_m]$ with oblivious updates. Note that this implies that $\mathscr{H}_A$, the history diagram of A, will only intersect $E \times \{\tau_{m-1/2}, \tau_m\}$ and hence will not intersect $\mathcal{X}$. Now the probability of an edge being updated in $[\tau_{m-1/2}, \tau_m]$ is $1 - e^{-\frac{\Delta}{2}}$ where Δ appeared in the definition of the τ_i's. Moreover the probability of an update being oblivious is $1 - p + p^*$. Putting the above together, we get that the denominator of (5.27) is bounded below by $e^{-c(p)|A|}$ for some $c(p) > 0$. This completes the proof of (5.24). □

Thereby, it only remains to prove the following proposition.

Proposition 5.19. *For any $\theta > 0$, we can find two constants $C = C(\theta) > 0$ and $p_0 = p_0(\theta) > 0$ such that, for any $p \in (0, p_0)$ there exists a constant $\alpha = \alpha(p) > 0$ satisfying*

$$\mathbb{P}\left[A \in \mathcal{C}_{\mathcal{R}(A)}^*\right] \leq C e^{-(\theta|\text{Conn}(A)| + \alpha \tau_m)} \quad \text{for all } A \subset E.$$

5.3.1. *Domination by sub-critical branching process.* To estimate the probability $\mathbb{P}[A \in \mathcal{C}^*_{\mathcal{R}(A)}]$, we fix A and m in the remaining part of the current section. Recall the notation W_i from (5.7). The main idea of the proof is that for sufficiently small p, the sequence W_m, W_{m-1}, ..., W_1 is dominated by a subcritical branching process in a suitable sense that will be explained below. Note that the event $\{A \in \mathcal{C}^*_{\mathcal{R}(A)}\}$ requires that

(1) $\mathscr{H}_e$ for some $e \in A$ starting at time $t = \tau_m$ survives to time $t = \tau_1$.
(2) All the history diagrams $\mathscr{H}_e$, $e \in A$, are connected together before arriving at $t = \tau_1$.

Comparing them with sub-critical branching processes will allow us to bound the probabilities of the above events. As Lemma 5.21 and the discussion following that will show, the analysis has to take into account that the 1-dependence across time of the Bernoulli percolation clusters used to define the information percolation history diagrams prevents a contraction every time step. Nonetheless this is sufficient to yield subcritical behavior once every two steps which is enough for our purposes.

We start with a general lemma. For $r \in (0, p_{\mathrm{perc}}(d))$, let ω_r be an i.i.d. standard bond percolation configuration on the lattice $(\mathbb{Z}^d, E(\mathbb{Z}^d))$ where each edge is open with probability r. Denote by $\overline{\mathrm{Conn}}(e; \omega_r)$ the closure of the open cluster containing an edge e as in (5.3), and let m_r be the distribution of $|\overline{\mathrm{Conn}}(e; \omega_r)|$, i.e.,

$$m_r(k) = \mathbb{P}\left[|\overline{\mathrm{Conn}}(e; \omega_r)| = k\right] \quad; \; k \in \mathbb{Z}_+ . \tag{5.28}$$

It is well-known (see [8, 17]) that there exists a constant $\rho(r) > 0$ such that, for all $e \in E(\mathbb{Z}^d)$,

$$m_r(k) \leq e^{-\rho(r)k} \quad \text{for all } k \geq 1 . \tag{5.29}$$

Lemma 5.20. *Fix a non-empty set $A \subset E$ and consider a random configuration $X \in \Omega_n$ whose distribution is stochastically dominated by $\mathrm{Perc}_n(r)$ for some $r \in (0, p_{perc}(d))$.*

Given X, we define

$$A(X) = \bigcup_{e \in A} \overline{\mathrm{Conn}}(e; X) .$$

Let $(y_i)_{i=1}^\infty$ be a sequence of i.i.d. random variables in $\mathbb{Z}_+$ distributed according to m_r. Then, $|A(X)|$ is stochastically dominated by $y_1 + y_2 + \cdots + y_{|A|}$.

Proof. Take an arbitrary enumeration $A = \{e_1, e_2, \ldots, e_{|A|}\}$ and define disjoint sets $G_1, G_2, \ldots, G_{|A|}$ as $G_1 = \overline{\mathrm{Conn}}(e_1; X)$ and

$$G_k = \overline{\mathrm{Conn}}(e_k; X) \setminus \left[\bigcup_{i=1}^{k-1} \overline{\mathrm{Conn}}(e_i; X)), \right] \quad; \; k \in [\![2, |A|]\!] .$$

Then, the set $A(X)$ can be represented as the disjoint union of $G_1, G_2, \ldots, G_{|A|}$ and thus

$$|A(X)| = \sum_{i=1}^{|A|} |G_i| .$$

We now claim that

$$\sum_{i=1}^k |G_i| \preceq \sum_{i=1}^k y_i \;\; \text{for all } k \in [\![1, |A|]\!] .$$

Clearly this is true for $k = 1$. To finish the proof by the induction, it suffices to prove that the distribution of $|G_{i+1}|$ given $G_1, \ldots, G_i$, is stochastically dominated by y_{i+1}. This follows by the spatial independence of bond percolation. More precisely, given $G_1, \ldots, G_i$, the edge configuration on $(G_1 \cup \cdots \cup G_i)^c$ is a Bernoulli percolation with the same parameter, and thus the distribution of G_{i+1} is dominated by that of $\overline{\mathrm{Conn}}(e_{i+1}; X)$. By (5.28) and the fact that X is dominated by $\mathrm{Perc}_n(r)$, the size of the latter is dominated by the distribution $m_r(\cdot)$ and the proof is completed. $\square$

Recalling (5.7), let

$$a_i := |W_i| \; ; \; i \in [\![0, \, m-1]\!] \; ,$$

and let

$$p_1 := 2p^{1/2} \; . \tag{5.30}$$

A direct application of Lemma 5.20 is the following bound on a_{m-1} which along with the fact that $\mathbb{P}(y_i = 0) = 1 - o_p(1)$ (by (5.29)), shows that the information percolation history diagram exhibits a contraction from $W_m = A$ to W_{m-1} similar to a subcritical branching process.

Lemma 5.21. *Suppose that $p_1 < p_{perc}(d)$ and let $(y_i)_{i=1}^{\infty}$ be a sequence of i.i.d. random variables with distribution m_{2p_1}. Then, we have*

$$a_{m-1} \preceq y_1 + \cdots + y_{|A|} \; .$$

Proof. We apply Lemma 5.20 with $X = \Xi_{m-1} \cup \Xi_{m-2}$. By Proposition 4.7 and union bound, the distribution of X is stochastically dominated by $\mathrm{Perc}_n(2p_1)$. We now claim that in the construction of the evolution of the history diagram of an edge e over the time interval $[\tau_{m-1}, \tau_m]$, e gets expanded to a subset of $\overline{\mathrm{Conn}}(e; X)$. To see this, we first note that there are only three possible cases: the oblivious update, the non-oblivious update, and the non-update and the corresponding expansions being the empty set, $\overline{\mathrm{Conn}}(e; X)$, and e, respectively. Thus, to prove the claim, it suffices to just check the non-update case for which the expansion set is merely e. However in this case, the claim is verified by observing that e is open in $\mathscr{N}_{m-1} \preceq \Xi_{m-1} \preceq X$ where all the inclusions are by definition. Hence, we can conclude that $W_{m-1} \subset A(X)$. The assertion of the lemma is now immediate from Lemma 5.20. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

We now state the main result regarding the domination by branching process. For $i \in [\![1, \, m]\!]$, denote by $\mathcal{F}_i$ the σ-algebra on Ω_n generated by update sequence $\mathrm{Upd}[\tau_i, \, \tau_m]$ (Hence $\mathcal{F}_m = \{\emptyset, \, \Omega_n\}$).

Proposition 5.22. *Suppose that p is small enough so that $3p_1 < p_{perc}(d)$. Let $(y_i)_{i=1}^{\infty}$ be a sequence of i.i.d. random variables with distribution m_{3p_1} defined in (5.28). For all $i \in [\![1, \, m-2]\!]$, the distribution of a_i given $(\mathcal{F}_{i+2}, W_{i+2})$ is stochastically dominated by*

$$y_1 + y_2 + \cdots + y_{a_{i+2}} \; .$$

One might expect that the proof of Proposition 5.22 can be carried out similarly as that of Lemma 5.21. However this does not work, roughly because of the following: Assume that we condition on W_{i+1} and try to control $a_i = |W_i|$. Then, W_i is determined by W_{i+1}, the environment $\Xi_i \cup \Xi_{i-1}$, and the update sequence in $[\tau_i, \tau_{i+1}]$. However, by the same reasoning, W_{i+1} is determined from W_{i+2}, $\Xi_{i+1} \cup \Xi_i$ and the update sequence in $[\tau_{i+1}, \tau_{i+2}]$, and thus W_{i+1} already contains some information on Ξ_i. Therefore, the distribution of Ξ_i given W_i is hard to analyze. In particular, in the *worst case*, if all the edges in W_{i+1} belong to Ξ_i, one cannot expect a contraction estimate of a_i in terms of a_{i+1} described in the previous lemma. However, at this point one notices that Ξ_{i-1} and Ξ_{i-2} are still independent of W_{i+1} and hence one can possibly obtain a bound for a_{i-1} instead. In other words, if we conditioned on W_{i+2} and all the relevant information prior to it, the distribution of a_i, instead of a_{i+1}, can be dominated in an appropriate manner.

This is done through the next result whose proof crucially uses the definitions listed in Table 1. Recall the notations C_j and N_j from (5.9).

Proposition 5.23. *For $i \in [\![1, \, m-2]\!]$, define $\theta_i \in \Omega_n$ as following:*

$$\theta_i(e) = \begin{cases} (\Xi_{i-1} \cup \Xi_i)(e) & \text{if } e \in C_{i+2} \; , \\ (\Xi_{i-1} \cup \Xi_i \cup \Xi_{i+1})(e) & \text{if } e \in E \setminus C_{i+2} \; . \end{cases}$$

Define

$$Z_i = \bigcup_{e \in W_{i+2}} \overline{\text{Conn}}(e; \theta_i) . \tag{5.31}$$

Then, it holds that $W_i \subset Z_i$.

The proof of this proposition is based on two geometric lemmas (Lemmas 5.24 and 5.25). We refer to Figure 5.3 for the illustration of the proofs of these two lemmas and Proposition 5.23. However before proving the latter we first finish the proof of Proposition 5.22.
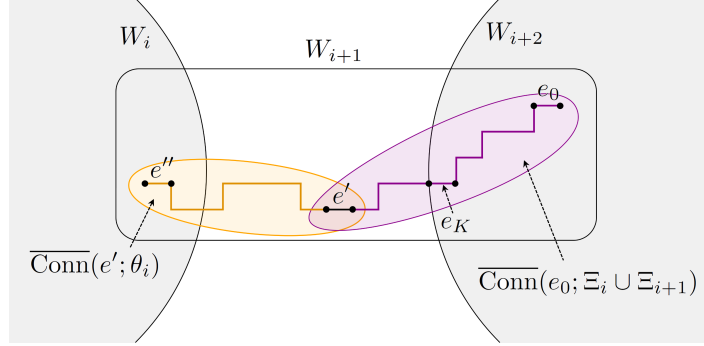

FIGURE 5.3. Figure illustrating the proof of Lemmas 5.24, 5.25, and Proposition 5.23. The crucial fact is that $e' \in \text{Conn}(e_0; \theta_i)$ (without bar).

Note that W_i is determined by $\text{Upd}[\tau_{i-1}, \tau_m]$, and hence a_i is a random variable measurable with respect to $\mathcal{F}_{i-1}$.

Proof of Proposition 5.22. We consider the following configuration

$$\Xi_{i+1}^o(e) = \begin{cases} 0 & \text{if } e \in C_{i+2} , \\ \Xi_{i+1}(e) & \text{if } e \in E \setminus C_{i+2} . \end{cases}$$

We first make the following claim.

Claim. *Given $(\mathcal{F}_{i+2}, W_{i+2})$, the distribution of Ξ_{i+1}^o is dominated by $\text{Perc}_n(p_1)$.*

Assuming this claim, since $\theta_i = \Xi_{i-1} \cup \Xi_i \cup \Xi_{i+1}^o$, by Proposition 4.8, it follows that the distribution of θ_i given $(\mathcal{F}_{i+2}, W_{i+2})$ is stochastically dominated by $\text{Perc}_n(3p_1)$. Hence, by Lemma 5.20 and the definition (5.31) of Z_i, we can conclude that $|Z_i|$ is stochastically bounded above by $y_1 + \cdots + y_{a_{i+2}}$. Thus we are done by Proposition 5.23.

It remains to prove the claim. We start by noting that C_{i+2} is not a deterministic function of $\mathcal{F}_{i+2}$ and W_{i+2}. However, by (5.10), W_{i+3}^{NU} and W_{i+3}^{NOb} which are subsets of W_{i+2} are indeed measurable with respect to $\mathcal{F}_{i+2}$. Next recalling how C_{i+2} is constructed from (5.9), note that given $\mathcal{F}_{i+2}$ and W_{i+2}, by standard exploration of $\overline{\text{Conn}}(e; \Xi_{i+1} \cup \Xi_{i+2})$ for $e \in W_{i+3}^{\text{NOb}}$, further conditioning on C_{i+2}, does not affect the distribution of the updates in

$$\bigcup_{e \in E \setminus C_{i+2}} \text{Upd}[\tau_{i+1}, \tau_{i+2}](e), \tag{5.32}$$

(note that here we are crucially using the fact that $\overline{\text{Conn}}(e; \Xi_{j-1} \cup \Xi_j)$ includes the closed boundary edges since otherwise conditioning on C_{i+2} would yield information about its boundary edges which would then have been members of $E \setminus C_{i+2}$). Thus, from now we assume that C_{i+2} is given, and suppose that $e \notin C_{i+2}$. Since the configuration $\Xi_{i+1}(E \setminus C_{i+2})$ is determined by the updates in (5.32), we can conclude that the distribution of $\Xi_{i+1}(E \setminus C_{i+2})$ given $(\mathcal{F}_{i+2}, W_{i+2}, C_{i+2})$ is stochastically bounded by percolation on $E \setminus C_{i+2}$ with open probability p_1, by Proposition 4.7 and the definition

of p_1 in (5.30). Since $\Xi_{i+1}^o(e) = 0$ for $e \in C_{i+2}$, the claim holds conditionally on $(\mathcal{F}_{i+2}, W_{i+2}, C_{i+2})$ and hence by averaging over C_{i+2}, conditionally on $(\mathcal{F}_{i+2}, W_{i+2})$. $\qquad\square$

Lemma 5.24. *For $i \in [\![1, m-1]\!]$, we have that*

$$W_i \subset \bigcup_{e \in W_{i+1}} \overline{\mathrm{Conn}}(e; \Xi_{i-1} \cup \Xi_i) . \tag{5.33}$$

Proof. In view of the definition (5.9), the decomposition (5.12), and the fact that $W_{i+1}^{\mathrm{NU}} \subset W_{i+1}$, it suffice to check that

$$N_i \subset \bigcup_{e \in W_{i+1}^{\mathrm{NU}}} \overline{\mathrm{Conn}}(e; \Xi_{i-1} \cup \Xi_i) . \tag{5.34}$$

If $e \in W_{i+1}^{\mathrm{NU}}$, we have $\mathcal{N}_i(e) = 1$ by the definition of $\mathcal{N}_i$, and thus $(\Xi_i \cup \Xi_{i-1})(e) = 1$ since $\mathcal{N}_i \leq \Xi_i$. Therefore, we have $e \in \overline{\mathrm{Conn}}(e; \Xi_i \cup \Xi_{i-1})$. Hence, the right-hand side of (5.34) contains W_{i+1}^{NU}, and hence contains N_i by the second inclusion of Lemma 5.7. $\qquad\square$

Lemma 5.25. *For $i \in [\![1, m-2]\!]$, define $\xi_i \in \Omega_n$ as follows:*

$$\xi_i(e) = \begin{cases} \Xi_i(e) & \text{if } e \in C_{i+2} , \\ (\Xi_i \cup \Xi_{i+1})(e) & \text{if } e \in E \setminus C_{i+2} . \end{cases}$$

Then, we have

$$W_{i+1} \setminus W_{i+2} \subset \bigcup_{e \in W_{i+2}} \overline{\mathrm{Conn}}(e; \xi_i) .$$

Proof. For $e' \in W_{i+1} \setminus W_{i+2}$, we know from Lemma 5.24 that there exists $e_0 \in W_{i+2}$ and a path

$$e_0, e_1, \cdots, e_k(= e')$$

in E such that $(\Xi_i \cup \Xi_{i+1})(e_l) = 1$ for all $l \in [\![0, k-1]\!]$. If none of $e_0, e_1, \cdots, e_k$ belongs to C_{i+2} then the assertion of lemma is immediate since $\Xi_i \cup \Xi_{i+1} = \xi_i$ along this path. Otherwise, let

$$K = \max\{h : e_h \in C_{i+2}\} .$$

Since $e' \notin W_{i+2}$, we have $K < k$. Then, since $e_{K+1} \notin C_{i+2}$, we have $e_K \in \partial^- C_{i+2}$ and thus $(\Xi_{i+1} \cup \Xi_{i+2})(e_K) = 0$ by Lemma 5.8. Since $(\Xi_i \cup \Xi_{i+1})(e_K) = 1$, we can conclude that $\Xi_i(e_K) = 1$. This implies that $e' = e_k \in \overline{\mathrm{Conn}}(e_K; \xi_i)$, where $e_K \in C_{i+2} \subset W_{i+2}$. This completes the proof. $\qquad\square$

Now we are ready to prove Proposition 5.23.

Proof of Proposition 5.23. Fix arbitrary $e'' \in W_i$. It suffices to verify that $e'' \in Z_i$. Since $\Xi_{i-1} \cup \Xi_i \leq \theta_i$, by Lemma 5.24, there exists $e' \in W_{i+1}$ such that

$$e'' \in \overline{\mathrm{Conn}}(e'; \Xi_{i-1} \cup \Xi_i) \subset \overline{\mathrm{Conn}}(e'; \theta_i) . \tag{5.35}$$

If $e' \in W_{i+2}$, we can immediately assert that $e' \in Z_i$ by the definition of Z_i.

On the other hand, if $e' \in W_{i+1} \setminus W_{i+2}$, then by Lemma 5.25 and by the fact that $\xi_i \leq \theta_i$, there exists $e_0 \in W_{i+2}$ such that

$$e' \in \overline{\mathrm{Conn}}(e_0; \xi_i) \subset \overline{\mathrm{Conn}}(e_0; \theta_i) . \tag{5.36}$$

We remark that (5.35) implies that $\theta_i(e') = 1$ since otherwise $\overline{\mathrm{Conn}}(e'; \theta_i) = \emptyset$. Therefore we can replace $e' \in \overline{\mathrm{Conn}}(e_0; \theta_i)$ in (5.36) with $e' \in \mathrm{Conn}(e_0; \theta_i)$. Combining this with (5.35) ensures that $e'' \in \overline{\mathrm{Conn}}(e_0, \theta_i) \subset Z_i$. This completes the proof. $\qquad\square$

5.3.2. *Bounds on a_i based on domination by branching processes.* Now we present two consequences of the previous branching process type estimate. These will play a fundamental role in the proof of Proposition 5.19.

For a random variable y in $\mathbb{Z}_+$ following the law m_{3p_1} defined in (5.28), we define $M = M(p)$ as the solution of

$$e^{-2M} = \mathbb{E}(y) \ .$$

It readily follows that

$$\lim_{p \to 0} M = \infty \ . \tag{5.37}$$

Lemma 5.26. *For $k \in [\![1, m]\!]$, select $\mathfrak{r} \in \{1, 2\}$ so that $(k - \mathfrak{r}) \bmod 2 = 0$. Then, for some constant $C = C(p) > 0$, it holds that*

$$\mathbb{E}\left[a_{\mathfrak{r}} | a_k\right] \leq C e^{-Mk} a_k \ .$$

Proof. It follows from Proposition 5.22 that, for all $k \in [\![3, m]\!]$,

$$\mathbb{E}\left[a_{k-2} | a_k, \mathcal{F}_k\right] \leq e^{-2M} a_k \ . \tag{5.38}$$

Then, the proof of lemma is completed by the induction. $\square$

Lemma 5.27. *For all sufficiently small p, there exists $c_0 = c_0(p) > 0$ such that,*

$$\mathbb{E} \exp\left\{c_0 \sum_{i=1}^{m-1} a_i\right\} \leq e^{|A|} \ .$$

Furthermore, $\lim_{p \to 0} c_0(p) = +\infty$.

Proof. By the Cauchy-Schwarz inequality,

$$\mathbb{E}\left[\exp\left\{c \sum_{i=1}^{m-1} a_i\right\}\right]^2 \leq \mathbb{E} \exp\left\{2c \sum_{i:2i \in [\![1, m-1]\!]} a_{2i}\right\} \cdot \mathbb{E} \exp\left\{2c \sum_{i:2i+1 \in [\![1, m-1]\!]} a_{2i+1}\right\} \ . \tag{5.39}$$

Denote by y the random variable with distribution m_{3p_1} defined in (5.28). Note that the following equation on c

$$\mathbb{E} e^{(2c+1)y} = e \tag{5.40}$$

has a positive solution $c_0 = c_0(p)$ and we can readily check that $\lim_{p \to 0} c_0 = +\infty$. Now it suffices to prove that, for all ℓ,

$$\mathbb{E} \exp\left\{2c_0 \sum_{i=1}^{\ell} a_{m-2i}\right\} \leq e^{|A|} \quad \text{and} \quad \mathbb{E} \exp\left\{2c_0 \sum_{i=0}^{\ell} a_{m-2i-1}\right\} \leq e^{|A|} \ . \tag{5.41}$$

By Proposition 5.22 and (5.40), for all $i \in [\![3, m]\!]$, we have

$$\mathbb{E}\left[e^{(2c_0+1)a_{i-2}} \middle| a_i, \mathcal{F}_i\right] \leq \mathbb{E}\left[e^{(2c_0+1)y}\right]^{a_i} \leq e^{a_i} \ .$$

Consequently, for all $\ell \geq 1$,

$$\mathbb{E}\left[e^{a_{m-2\ell}} \cdot \exp\left\{2c_0 \sum_{i=1}^{\ell} a_{m-2i}\right\} \middle| a_{m-2\ell+2}, \mathcal{F}_{m-2\ell+2}\right] \leq e^{a_{m-2\ell+2}} \exp\left\{2c_0 \sum_{i=1}^{\ell-1} a_{m-2i}\right\} \ .$$

Repeating this procedure, we obtain

$$\mathbb{E}\exp\left\{2c_0\sum_{i=1}^{\ell}a_{m-2i}\right\} \leq \mathbb{E}\left[e^{a_{2\ell}}\exp\left\{2c_0\sum_{i=1}^{\ell}a_{m-2i}\right\}\right]$$

$$\leq \mathbb{E}\left[e^{a_{2\ell+2}}\exp\left\{2c_0\sum_{i=1}^{\ell-1}a_{m-2i}\right\}\right] \quad (5.42)$$

$$\leq \cdots \leq \mathbb{E}\left[e^{(2c_0+1)a_{m-2}}\right] \leq e^{a_m} = e^{|A|}\,.$$

This proves the first inequality in (5.41). By a similar argument as above, one can show that

$$\mathbb{E}\exp\left\{2c_0\sum_{i=0}^{\ell}a_{m-2i-1}\right\} \leq \mathbb{E}e^{(2c_0+1)a_{m-1}}\,. \quad (5.43)$$

By Lemma 5.21 and the fact that m_{2p_1} is dominated by m_{3p_1}, we have

$$\mathbb{E}e^{(2c_0+1)a_{m-1}} \leq \mathbb{E}\left[e^{(2c_0+1)y}\right]^{|A|} \leq e^{|A|}\,, \quad (5.44)$$

where the last inequality follows from (5.40). Now, (5.41) is proven by combining (5.42), (5.43), and (5.44). $\square$

We now proceed to proving Proposition 5.19.

5.3.3. *Proof of Proposition 5.19.* Recall that we want to bound $\mathbb{P}\left[A \in \mathcal{C}^*_{\mathcal{R}(A)}\right]$.

We start by defining

$$\sigma := \begin{cases} \max\{i : a_i = 1\} & \text{if } a_i = 1 \text{ for some } i \in [\![1,\,m]\!], \\ 0 & \text{otherwise.} \end{cases}$$

We further define events $\mathcal{A}$ and $\mathcal{B}$ as

$$\mathcal{A} = \Big(\{\sigma > 0\} \cap \{\mathscr{H}_A \text{ merges to one point in } [\tau_\sigma, \tau_m]\}\Big) \bigcup \Big(\{\sigma = 0\} \cap \{A \in \mathcal{C}^*_{\mathcal{R}(A)}\}\Big),$$

$$\mathcal{B} = \Big(\{\sigma > 0\} \cap \{\text{The history diagram starting from } W_\sigma \text{ at } t = \tau_\sigma \text{ survives until } t = \tau_1\}\Big)$$

$$\bigcup\{\sigma = 0\}.$$

Note that in the definition of $\mathcal{A}$, on the event $\{\sigma > 0\}$, we put the additional constraint that all the history diagrams in $\mathcal{H}_A$ merge to a point in $[\tau_\sigma, \tau_m]$. Note that this is not guaranteed just by assuming $\sigma > 0$, since it may happen that all the history diagrams have been killed except for one edge which survives on its own up to τ_σ. Clearly, the event $\{A \in \mathcal{C}^*_{\mathcal{R}(A)}\}$ is a subset of $\mathcal{A} \cap \mathcal{B}$ since if $\sigma > 0$, the only way $\{A \in \mathcal{C}^*_{\mathcal{R}(A)}\}$ can occur is if $\mathcal{A} \cap \mathcal{B}$ occurs since otherwise $\mathcal{C}_{\mathcal{R}(A)}$ has multiple connected components. Thus we have

$$\mathbb{P}\left[A \in \mathcal{C}^*_{\mathcal{R}(A)}\right] \leq \mathbb{P}\left[\mathcal{A} \cap \mathcal{B}\right]\,. \quad (5.45)$$

We next make and prove the following claim.

Claim. *Conditioned on the event $\{\sigma = k\}$ with $k \geq 0$, two events $\mathcal{A}$ and $\mathcal{B}$ are independent.*

Proof. For $k = 0$, the claim is immediate from the definitions of $\mathcal{A}$ and $\mathcal{B}$. For $k \geq 1$, let us write $W_k = \{e\}$. Then, conditioned on the event $\{\sigma = k\}$, it suffices prove that the event $\mathcal{A}$ is independent of $\mathrm{Upd}[0, \tau_k]$ since $\mathcal{B}$ depends only on $\mathrm{Upd}[0, \tau_k]$. Clearly the behavior of the history diagram starting from e in $(\tau_{k+1}, \tau_m]$ is independent of $\mathrm{Upd}[0, \tau_k]$. Hence, it only suffices to check the interval $(\tau_k, \tau_{k+1}]$. The event $\mathcal{A}$ imposes that all the edges in $W_{k+1} \setminus \{e\}$ exhibit the oblivious

update in $(\tau_k, \tau_{k+1}]$, while $e \in W_{k+1}$ survives to τ_k without expanding to $\overline{\mathrm{Conn}}(e; \Xi_{k-1} \cup \Xi_k)$ which by definition includes e as well as its adjacent edges. The first event is determined by $\mathrm{Upd}[\tau_k, \tau_{k+1}]$ and hence is independent of $\mathrm{Upd}[0, \tau_k]$. The second event occurs only when there is no update at e in $\mathrm{Upd}[\tau_k, \tau_{k+1}]$, and hence this event is also independent of $\mathrm{Upd}[0, \tau_k]$ as well. This completes the proof. $\qquad\square$

By this claim and (5.45), we deduce that

$$\mathbb{P}\left[A \in \mathcal{C}^*_{\mathcal{R}(A)}\right] \leq \mathbb{E}\left[\mathbb{P}[\mathcal{A} \mid \sigma]\, \mathbb{P}[\mathcal{B} \mid \sigma]\right] . \tag{5.46}$$

We now claim that there exists $C, M > 0$ such that for all $k \geq 0$,

$$\mathbb{P}[\mathcal{B} \mid \sigma = k] \leq C e^{-Mk} . \tag{5.47}$$

Since this bound trivially holds for $k = 0$ if we take $C > 1$, it suffice to consider the case $k \geq 1$. Select $\mathfrak{r} \in \{1, 2\}$ so that $(\sigma - \mathfrak{r}) \bmod 2 = 0$. Then, by Lemma 5.26, we get

$$\mathbb{P}[\mathcal{B} \mid \sigma = k] \leq \mathbb{P}[a_{\mathfrak{r}} > 0 \mid a_k = 1] \leq \mathbb{E}[a_{\mathfrak{r}} \mid a_k = 1] \leq C e^{-Mk} .$$

This proves the bound (5.47). Now by (5.46) and (5.47), we have

$$\mathbb{P}\left[A \in \mathcal{C}^*_{\mathcal{R}(A)}\right] \leq \mathbb{E}\left[C e^{-M\sigma} \mathbf{1}_{\mathcal{A}}\right] . \tag{5.48}$$

If $|A| = 1$ so that $\sigma = m$, this inequality proves the assertion of the proposition. Now we assume that $|A| \geq 2$, so that $\sigma \leq m - 2$. Note that σ cannot be $m - 1$ since $W_{m-1} \supset A$ under $\mathcal{A}$ since otherwise for $e \in A \setminus W_{m-1}$, the set $\mathscr{H}_e$ is a singleton and hence $\mathscr{H}_e \cap \mathscr{H}_{A \setminus \{e\}} = \emptyset$. Thus conditioned on the event $\{\sigma = k\}$ with $k \geq 1$, the event $\mathcal{A}$ implies that

$$\{a_{k+1} + \cdots + a_{m-1} \geq |\mathrm{Conn}(A)| - 1\} ,$$

where, recall that $|\mathrm{Conn}(A)|$ is the smallest possible number of edges of the connected subgraph of (Λ_n, E_n) containing A. We can neglect a_m since $W_m = A \subset W_{m-1}$ under $\mathcal{A}$, and we used the fact that $a_k = 1$. On the other hand, for $k = 0$, the event $\mathcal{A}$ implies that

$$\{a_1 + \cdots + a_{m-1} \geq |\mathrm{Conn}(A)|\} .$$

Recall from Definition 5.3, that our definition of the information percolation clusters only extended up to τ_1 and not τ_0. This is reflected in the fact that the above sum starts from a_1 instead of a_0. Therefore, we can bound the right-hand side of (5.48) from above by

$$\sum_{k=1}^{m-2} C e^{-Mk} \mathbb{E}\left[\mathbf{1}\left\{a_{k+1} + \cdots + a_{m-1} \geq |\mathrm{Conn}(A)| - 1\right\} \mathbf{1}\{\sigma = k\}\right]$$

$$+ C \mathbb{E}\left[\mathbf{1}\left\{a_1 + \cdots + a_{m-1} \geq |\mathrm{Conn}(A)|\right\} \mathbf{1}\{\sigma = 0\}\right] .$$

By applying $\mathbf{1}\{\sigma = k\} \leq \mathbf{1}\{a_{k+1}, \cdots, a_{m-1} \geq 2\}$ here, we obtain

$$\mathbb{E}\left[C e^{-M\sigma} \mathbf{1}_{\mathcal{A}}\right] \leq C \sum_{k=0}^{m+1} H_k , \tag{5.49}$$

where

$$H_k = e^{-Mk} \mathbb{E}\left[\mathbf{1}\left\{a_{k+1} + \cdots + a_{m-1} \geq |\mathrm{Conn}(A)| - 1\right\} \mathbf{1}\left\{a_{k+1}, \cdots, a_{m-1} \geq 2\right\}\right] .$$

For any $C_1, C_2 > 0$, by the Chebyshev inequality,

$$H_k \leq e^{-Mk} e^{-C_1(|\mathrm{Conn}(A)| - 1)} e^{-2C_2(m-k)} \mathbb{E} e^{(C_1 + C_2)(a_{k+1} + \cdots + a_{m-1})} .$$

Now we take C_1 and C_2 such that $C_1 + C_2 < c_0$ where the constant c_0 is the one appeared in Lemma 5.27. Then, by Lemma 5.27 and the fact that $|A| \leq |\text{Conn}(A)|$, we can further obtain

$$H_k \leq e^{-Mk} e^{-C_1(|\text{Conn}(A)|-1)} e^{-2C_2(m-k)} e^{|\text{Conn}(A)|} . \tag{5.50}$$

For given $\theta > 0$, we first take p small enough so that $c_0 > \theta + 2$ and $M > 1$. This is possible since

$$\lim_{p \to 0} c_0 = \lim_{p \to 0} M = +\infty$$

by (5.37) and Lemma 5.27. Take $C_1 = \theta + 1$ and $C_2 = 1/2$. With this selection, the bound (5.50) becomes

$$H_k \leq e^{\theta+1} e^{-(m+\theta|\text{Conn}(A)|)} e^{-(M-1)k} .$$

Combining this with (5.49) yields

$$\mathbb{E}\left[Ce^{-M\sigma} \mathbf{1}_{\mathcal{A}}\right] \leq C(\theta) e^{-(m+\theta|\text{Conn}(A)|)} .$$

Thus the statement of the proposition follows by recalling that $\tau_m = m\Delta$. $\qquad\square$

6. REDUCTION TO A PRODUCT CHAIN

From now on, we define

$$r = r(n) = 3 \log^5 n .$$

Moreover recall $t_{\max}$ from (4.10). Denote by $(X_t^\dagger)_{t \geq 0}$ the FK-dynamics defined on the periodic lattice $\mathbb{Z}_r^d$. Let $\Omega_r = \{0, 1\}^{E_r}$ where $E_r = E(\mathbb{Z}_r^d)$, and denote by $\pi^\dagger := \mu_{p,q}^r$ the random-cluster measure on $\Omega_r = \{0, 1\}^{E_r}$. Let $\Lambda \subset E_r$ be a box of size $2 \log^5 n$. Then, define

$$\mathbf{d}_t = \mathbf{d}_{t,n} = \max_{x_0^\dagger \in \Omega_r} \left\| \mathbb{P}_{x_0} \left[X_t^\dagger(\Lambda) \in \cdot \right] - \pi_\Lambda^\dagger \right\|_{L^2(\pi_\Lambda^\dagger)} , \tag{6.1}$$

where $X_t^\dagger(\Lambda)$ represents the configuration of $X_t^\dagger$ on Λ, and $\pi_\Lambda^\dagger$ stands for the projection of $\pi^\dagger$ onto the set Λ. The main result of this section is the following theorem.

Theorem 6.1. *For all sufficiently small p, there exists a constant $C_1 = C_1(p)$ such that the following hold.*

(1) *For $s \in [C_1 \log \log n, t_{max}]$ and $t \in [0, t_{max}]$, it holds that*

$$\max_{\nu:\nu \preceq \text{Perc}_n(t_{init})} \left\| \mathbb{P}_\nu \left[X_{t+s} \in \cdot \right] - \mu_{p,q}^n \right\|_{TV} \leq \frac{1}{2} \left[\exp \left\{ \frac{n^d}{\log^{12d} n} \mathbf{d}_t^2 \right\} - 1 \right]^{1/2} + \frac{4}{n^{2d}} .$$

(2) *If $t \geq C_1 \log \log n$ and*

$$\lim_{n \to \infty} \left(\frac{n}{\log^{10} n} \right)^d \mathbf{d}_t^2 = +\infty ,$$

then we have

$$\liminf_{n \to \infty} \max_{\nu:\nu \preceq \text{Perc}_n(t_{init})} \left\| \mathbb{P}_\nu \left[X_t \in \cdot \right] - \mu_{p,q}^n \right\|_{TV} = 1 .$$

Remark 6.2. As the proof will reveal, part (1) of the theorem holds even when X_t is the FK-dynamics on $\mathbb{Z}_m^d$, instead of $\mathbb{Z}_n^d$, where $m \in [\![\log^5 n, n]\!]$. The inequality in this case is

$$\max_{\nu:\nu \preceq \text{Perc}_m(t_{\text{init}})} \left\| \mathbb{P}_\nu \left[X_{t+s} \in \cdot \right] - \mu_{p,q}^m \right\|_{TV} \leq \frac{1}{2} \left[\exp \left\{ m^d \mathbf{d}_t^2 \right\} - 1 \right]^{1/2} + \frac{4}{n^{2d}} ,$$

where $\mathbf{d}_t$ is as in (6.1).

Henceforth, the constant $C_1 > 0$ will always refer to the constant appeared in this theorem. The proof of this theorem will be presented in the remaining part of the current section. We shall assume that p is small enough so that all the results established in Sections 4 (including 4.4) and 5 are valid. As indicated in Section 2, following the strategy in [20] where a similar statement as Theorem 6.1 appears, we will reduce the chain to an approximate product chain. The only major difference in the statement of Theorem 6.1, as compared to statements appearing in previous articles is that owing to the non-locality of the dynamics, we initialize from a sparse initial condition dominated by a sub-critical percolation which can be obtained by evolving the initial configuration for a burning time (see Lemma 4.4).

We start by giving a short roadmap of what the various subsections achieve.

- The first part (Section 6.1) constructs the so called Barrier dynamics where the FK-dynamics on $(\mathbb{Z}/n\mathbb{Z})^d$ gets compared to FK-dynamics on a disjoint collection of $(\mathbb{Z}/r\mathbb{Z})^d$ where $r = \log^5 n$.
- We define the notion of Update support in Section 6.2. To get an upper bound on the mixing time, we bound the total variation distance at $d(t + s)$ where $t = t_{\text{mix}} = O(\log n)$ and $s = \log\log(n)$. The natural strategy is to couple the configurations at time $t + s$ starting from any two arbitrary initial configurations. At this point the key observation is that irrespective of the configuration at time t, all but a sparse set of small boxes couple at time $t + s$. The remainder is called the 'Update support' for reasons which will be clear later and hence the remaining task is to ensure that the time interval $[0, t]$ is sufficient for the FK-dynamics starting from two arbitrary configurations to couple on the 'Update support'.
- We prove Theorem 6.1 in Section 6.3.

6.1. **Coupling with barrier-dynamics.** Divide E_n into disjoint squares of size $\log^5 n$ as follows. Let us write $K = n/\log^5 n$ and assume that K and $\log^5 n$ are integers for the simplification of notation. Define

$$V_n = \{0, \log^5 n, 2\log^5 n, \ldots, (K-1)\log^5 n\}^d \subset \Lambda_n .$$

For each $v \in V_n$, we define an **edge box** B_v by

$$B_v = \left\{ (u, u + e_j) : u \in v + [\![0, \log^5 n - 1]\!]^d \text{ and } j \in [\![1, d]\!] \right\} , \tag{6.2}$$

where e_j represents the jth standard normal vector in $\mathbb{R}^d$. One can think of B_v as a box of size $\log^5 n$ with some boundary edges are removed. Note that $(B_v)_{v \in V_n}$ is a decomposition of E_n. Furthermore, we mention that all the boxes below of various sizes, are edge boxes and hence for brevity we will refer to them as boxes.
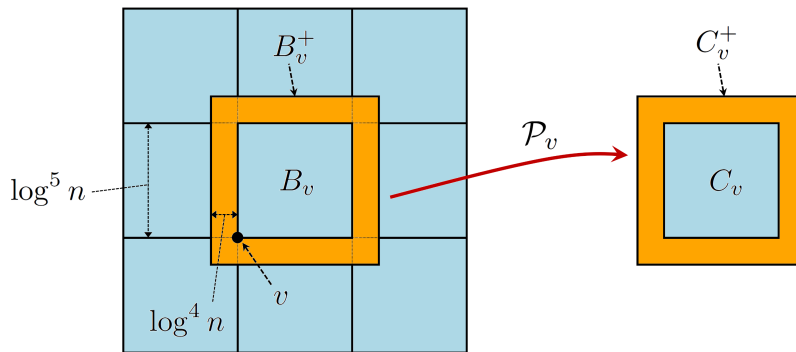

FIGURE 6.1. Figure illustrating the maps $\mathcal{P}_v$.

Then, for each $v \in V_n$, consider the expanded box $B_v^+ \subset \Lambda_n$ of B_v in the sense of (4.9). Then, B_v^+ is a box of size $\log^5 n + 2\log^4 n$ which is concentric with B_v. Let C_v^+ be another square lattice

of size $\log^5 n + 2\log^4 n$ and define a natural identification map $\mathcal{P}_v : B_v^+ \to C_v^+$. Define $C_v = \mathcal{P}_v(B_v)$ so that (C_v, C_v^+) is a copy of (B_v, B_v^+) (see Figure 6.1). We define

$$\widehat{E}_n = \bigsqcup_{v \in V_n} C_v^+ \text{ and } \widehat{\Omega}_n = \{0, 1\}^{\widehat{E}_n} .$$

Note that the last union is a disjoint union.

Definition 6.3 (Barrier-dynamics). For each $v \in V_n$, the barrier-dynamics is a FK-dynamics X_t^v on C_v^+ coupled with X_t by sharing the same update sequence via the following rules:

(1) (Initial condition) The initial edge configuration on C_v^+ is identical to that of B_v^+ through $\mathcal{P}_v$. In other words, $X_0^v(\mathcal{P}_v(e)) = X_0(e)$ for all $e \in B_v$.
(2) (Dynamics) We define the FK-dynamics $(X_t^v)_{t\geq0}$ on C_v^+ with periodic boundary condition by using the update sequence of B_v^+. Formally stating, we perform updates for each $e \in C_v^+$ by using the update sequence $\mathrm{Upd}(\mathcal{P}_v^{-1}(e))$ of the edge $\mathcal{P}_v^{-1}(e) \in B_v^+$.

For $t \geq 0$, we define a random map $\mathcal{G}_t : \Omega_n \to \Omega_n$ such that, for all $X_0 \in \Omega_n$,

$$[\mathcal{G}_t(X_0)](e) = X_t^v(\mathcal{P}_v(e)) ,$$

where $v \in V_n$ is the unique index such that $e \in B_v$. The next lemma now says that the actual dynamics and the barrier dynamics stay coupled for a significant amount of time provided the initial condition is sparse enough (note that for a spin system the latter condition is not needed since each update only depends on its immediate neighbors).

Lemma 6.4. *Suppose that p is small enough and the law of the initial condition X_0 follows the law ν such that $\nu \preceq \mathrm{Perc}_n(p_{init})$. Then, we have*

$$\mathbb{P}\big[X_t = \mathcal{G}_t(X_0) \text{ for all } t \in [0, t_{max}]\big] \geq 1 - n^{-2d} .$$

Proof. By Lemma 4.11 (cf. Remark 4.13), it holds that

$$\mathbb{P}\big[X_t(B_i) = X_t^v(C_i) \text{ for all } t \in [0, t_{\max}]\big] \geq 1 - n^{-3d} .$$

Thus, the conclusion of the lemma follows from the union bound since $|V_n| < n^d$. $\square$

6.2. Sparsity of update support.

Definition 6.5 (Update support). For each $s > 0$, denote by $\mathrm{U}_s = \mathrm{Upd}[0, s]$ the update sequence between time $[0, s]$. Then, the random map $\mathcal{G}_s$ is completely determined by U_s and hence we can write $\mathcal{G}_s = g_{\mathrm{U}_s}$ for some function $g_{\mathrm{U}_s} : \Omega_n \to \Omega_n$. The update support of U_s is the minimum subset $\Gamma_{\mathrm{U}_s} \subset E_n$ such that $\mathcal{G}_s$ is a function of $X(\Gamma_{\mathrm{U}_s})$ for all $X \in \Omega_n$, i.e.,

$$g_{\mathrm{U}_s}(X) = f_{\mathrm{U}_s}(X(\Gamma_{\mathrm{U}_s}))$$

for some $f_{\mathrm{U}_s} : \{0, 1\}^{\Gamma_{\mathrm{U}_s}} \to \Omega_n$.

Lemma 6.6. [20, Lemma 3.8] *Fix $t \geq 0$ and let U_s represent the update sequence for the time interval $[t, t+s]$ for $s \leq t_{max}$ where t_{max} was defined in (4.10). Suppose that p is small enough and a probability measure ν in Ω_n satisfies $\nu \preceq \mathrm{Perc}_n(t_{init})$. Then, we have*

$$\big\|\mathbb{P}_\nu[X_{t+s} \in \cdot\,] - \mu_{p,q}^n\big\|_{TV} \leq \int \big\|\mathbb{P}_\nu[X_t(\Gamma_{\mathrm{U}_s}) \in \cdot\,] - \mu_{\Gamma_{\mathrm{U}_s}}\big\|_{TV} d\mathbb{P}(\mathrm{U}_s) + 2n^{-3d} ,$$

where $\mu_{\Gamma_{\mathrm{U}_s}}$ represents the projection of $\mu_{p,q}^n$ on Γ_{U_s}.

Proof. The proof in the above reference relies only on the coupling of X_t and $\mathcal{G}_t(X_0)$ for $t \in [0, t_{\max}]$. For our model this has been established in Lemma 6.4 based on the bound on disagreement percolation using the sparse initial conditions. $\square$

Now we establish the sparsity of the update support Γ_{U_s}.

Definition 6.7 (Sparse set). A $\mathcal{S} \in \Omega_n$ is called *sparse* if for some $K \leq n^d (\log n)^{-12d}$, the graph induced by $\mathcal{S}$ can be decomposed into disjoint components $A_1, A_2, \cdots, A_K$ such that

    (1) For all distinct $i, j \in [\![1, K]\!]$, there is no open path in S connecting A_i and A_j.
    (2) Every A_i, $i \in [\![1, K]\!]$, has diameter at most $\log^5 n$. In particular, there is a box of size $2 \log^5 n$ containing A_i.
    (3) The distance between any distinct A_i and A_j is at least $4 \log^4 n$.

We write Spa_n to denote the set of sparse configurations in Ω_n.
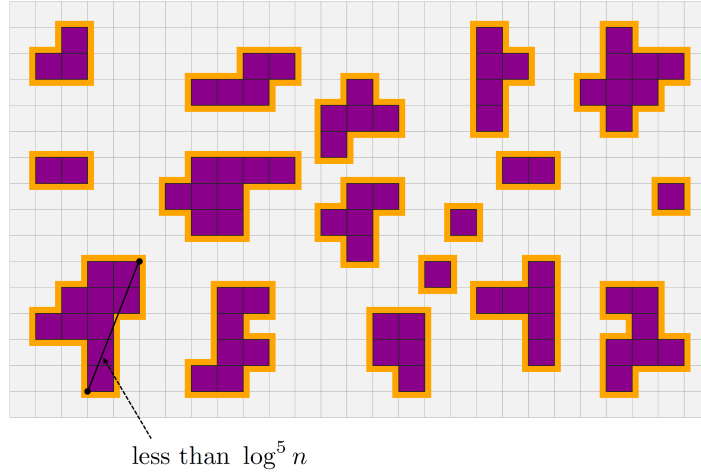

less than $\log^5 n$

FIGURE 6.2. The figure illustrates the sparse update support. The purple regions (including the orange buffer around them) denote the sets A_i (and A_i^+) as in Definitions 6.7 and 6.10.

Lemma 6.8. [20, Lemma 3.9] *There exists $C_2 = C_2(p) > 0$ such that, for all $s \geq C_2 \log \log n$,*

$$\mathbb{P}[\Gamma_{U_s} \in Spa_n] \geq 1 - n^{-3d} . \tag{6.3}$$

Proof. The only model-dependent part is the proof of the following fact: For $t \geq C_2 \log \log n$ with a large enough C_2,

$$\sum_{e \in C_v^+} \mathbb{P}\left[X_t^{v,\text{full}}(e) \neq X_t^{v,\text{empty}}(e)\right] \leq \log^{-10d} n , \tag{6.4}$$

where $(X_t^{v,\text{full}})_{t \geq 0}$ (resp. $(X_t^{v,\text{empty}})_{t \geq 0}$) is the FK-dynamics on periodic lattice C_v^+ with full (resp. empty) initial condition. The proof of this fact in our setting follows from Corollary 3.4 which indicates that, for some $C > 0$,

$$\sum_{e \in C_v^+} \mathbb{P}\left[X_t^{v,\text{full}}(e) \neq X_t^{v,\text{empty}}(e)\right] \leq |C_v^+| e^{-C \log \log n} .$$

Hence, the bound (6.4) follows if we take C large enough. The remaining part is identical to cited proofs and will not be repeated here. $\square$

By Lemmas 6.6 and 6.8, we obtain the following result.

Proposition 6.9. *Suppose that p is small enough and ν is a probability distribution on Ω_n satisfying $\nu \preceq \text{Perc}_n(p_{init})$. Then, for all $s \in [C_0 \log \log n, t_{max}]$ where C_0 is the constant appearing in Lemma 6.8, there exists a measure $\mathbb{Q}$ on Spa_n such that,*

$$\left\| \mathbb{P}_\nu(X_{t+s} \in \cdot) - \mu_{p,q}^n \right\|_{TV} \leq \int_{Spa_n} \left\| \mathbb{P}_\nu(X_t(\Gamma \in \cdot) - \mu_\Gamma \right\|_{TV} d\mathbb{Q}(\Gamma) + 3n^{-3d} .$$

6.3. **Proof of Theorem 6.1.** Before jumping into the proof we will need some technical preparation. The next few results use coupling arguments to compare the actual chain to a product chain. We start by defining a notion of good sets, and then introduce a generalized version of barrier dynamics.

Definition 6.10. A collection of disjoint subsets $A_1, A_2, \cdots, A_K$ of Ω_n are m-good for some $m \in \left[\log^4 n, (1/2)\log^5 n\right]$ if, each A_i is contained in a box of size $2\log^5 n$, and the expanded sets A_i^+, $i \in [\![1, K]\!]$, are disjoint where

$$A_i^+ = \{e \in E_n : d(e, A_i) \leq m\} .$$

As a consequence of Lemma 6.8, the sets A_i in the update support are $\log^4 n-$good (see Figure 6.2).

Let us take a box of size $r = 3\log^5 n$ containing A_i^+ and denote this box by $\widehat{A}_i^+$ (this is possible since $m \leq (1/2)\log^5 n$). Take K copies $\mathcal{L}_1, \mathcal{L}_2, \cdots, \mathcal{L}_K$ of the periodic lattice $\mathbb{Z}_r^d$, and embed each $\widehat{A}_i^+$ to $\mathcal{L}_i$ by a identification map $\mathcal{P}_i : \widehat{A}_i^+ \to \mathcal{L}_i$.

For $i \in [\![1, K]\!]$, denote by $(Y_t^{(i)})_{t \geq 0}$, the FK-dynamics on $\mathcal{L}_i$, whose update sequence and initial condition are inherited from that of $\widehat{A}_i^+ = \mathcal{P}_i^{-1}(\mathcal{L}_i)$ of $(X_t(\widehat{A}_i^+))_{t \geq 0}$. Let $\pi^{(i)}$ be the random-cluster measure on $\mathcal{L}_i$ so that $\pi^{(i)}$ is the invariant measure of $Y_t^{(i)}$. Define the product spaces:

$$\mathcal{L} = \prod_{i=1}^{K} \mathcal{L}_i \ , \quad \pi = \prod_{i=1}^{K} \pi^{(i)}, \ \text{and} \ Y_t^* = \prod_{i=1}^{K} Y_t^{(i)} \ ,$$

and let

$$\Gamma = \bigcup_{i=1}^{K} A_i \subset \Omega_n \ \text{and} \ \Gamma^* = \bigcup_{i=1}^{K} \mathcal{P}(A_i) \subset \mathcal{L} \ .$$

By slight abuse of notations, we identify A_i and $\mathcal{P}(A_i)$, for $i \in [\![1, K]\!]$ or Γ and Γ^* and simply write $\mathcal{P}(A_i) = A_i$ and $\Gamma^* = \Gamma$. With this identification, we can regard $X_t(A_i)$ and $Y_t^*(A_i)$ or $X_t(\Gamma)$ and $Y_t^*(\Gamma)$ as processes defined on the same space.

We first recall from Remark 4.13, that we can couple (X_t) and (Y_t^*). In the lemmas below where we record various coupling statements, we assume that the collection $A_1, A_2, \cdots, A_K$ of subsets of Ω_n is m-good for some $m \in \left[\log^4 n, (1/2)\log^5 n\right]$.

Lemma 6.11. *Suppose that p is small enough and the law of the initial condition X_0 follows the law ν such that $\nu \preceq \mathrm{Perc}_n(p_{init})$. Then, we have*

$$\mathbb{P}\left[X_t(\Gamma) = Y_t^*(\Gamma) \text{ for all } t \in [0, t_{max}]\right] \geq 1 - \frac{1}{n^{2d}} \ .$$

Proof. Since $K \leq n^d$, it suffices to show that, for all $i \in [\![1, K]\!]$,

$$\mathbb{P}\left[X_t(A_i) = Y_t^*(A_i) \text{ for all } t \in [0, t_{\max}]\right] \geq 1 - \frac{1}{n^{3d}} \ .$$

This follows directly from Lemma 4.11. $\square$

Now we obtain upper and lower bounds for the total-variation distance of (Y_t^*) in the two lemmas below. Combined with the previous coupling result, they yield bounds on the total-variation distance for (X_t).

Lemma 6.12. *For all sufficiently small p, we have that*

$$\sup_{x_0 \in \Omega_n} \|\mathbb{P}_{x_0}[Y_t^*(\Gamma) \in \cdot] - \pi_\Gamma\|_{TV} \leq \frac{1}{2}\left[e^{K\mathbf{d}_t^2} - 1\right]^{\frac{1}{2}},$$

where π_Γ represents the projection of π onto Γ.

Notation 6.13. In the statement of lemma, $\mathbb{P}_{x_0}[Y_t^*(\Gamma) \in \cdot]$ means that the starting configuration of Y_t^* is inherited from $x_0 \in \Omega_n$ by the collection map $\prod_{i=1}^K \mathcal{P}^{(i)} : \prod_{i=1}^K \widehat{A}_i^+ \to \prod_{i=1}^K \mathcal{L}_i$. We define $\mathbb{P}_\nu[Y_t^*(\Gamma) \in \cdot]$ for a probability distribution ν on Ω_n in the same manner.

Proof. By the L^1-L^2 inequality we have

$$\|\mathbb{P}_{x_0}[Y_t^*(\Gamma) \in \cdot] - \pi_\Gamma\|_{\mathrm{TV}} \leq \frac{1}{2} \|\mathbb{P}_{x_0}[Y_t^*(\Gamma) \in \cdot] - \pi_\Gamma\|_{L^2(\pi_\Gamma)} . \tag{6.5}$$

Denote by $\pi_{A_i}^{(i)}$ the projection of $\pi^{(i)}$ onto A_i. Then, since $\pi_\Gamma = \prod_{i=1}^K \pi_{A_i}^{(i)}$, by the bound of L^2-norm for product space (cf. [20, Section 3.2]), we obtain that

$$\|\mathbb{P}_{x_0}[Y_t^*(\Gamma) \in \cdot] - \pi_\Gamma\|_{L^2(\pi_\Gamma)} \leq \left[\exp\left\{\sum_{i=1}^K \left\|\mathbb{P}_{x_0}[Y_t^{(i)}(A_i) \in \cdot] - \pi_{A_i}^{(i)}\right\|_{L^2(\pi_i^*)}^2\right\} - 1\right]^{1/2} . \tag{6.6}$$

By the definition of $\mathbf{d}_t$ (see (6.1)) and by the fact that A_i is a subset of box of size $2\log^5 n$, we can deduce from the definition of $\mathbf{d}_t$ that

$$\left\|\mathbb{P}_{x_0}[Y_t^{(i)}(A_i) \in \cdot] - \pi_{A_i}^{(i)}\right\|_{L^2(\pi_i^*)} \leq \mathbf{d}_t . \tag{6.7}$$

We now conclude using (6.5), (6.6) and (6.7).   □

Recall that μ_Γ represents the projection of $\mu_{p,q}^n$ to Γ. Using spatial mixing properties, we conclude now that μ_Γ is close to π_Γ. This follows from Lemma 4.10 which implies that the effect of the boundary condition does not reach beyond the buffer region $A_i^+ \setminus A_i$ for each i (see Figure 6.2). Using this we prove that the total-variation distance between μ_Γ and π_Γ is small.

Lemma 6.14. *It holds that*

$$\|\mu_\Gamma - \pi_\Gamma\|_{TV} \leq \frac{1}{n^{2d}} .$$

Proof. We apply Lemma 4.10 with $A = \Gamma$ and $B = E_n \setminus \bigcup_{i=1}^K A_i^+$. Recall the measure $\mu_{B^c}^+$ and the configuration X from Lemma 4.10. The latter implies that, with probability more than $1 - n^{-2d}$, there exists a closed surface in $X(A_i^+ \setminus A_i)$ enclosing A_i for all $i \in [\![1, K]\!]$. This implies the Lemma by the domain Markov property of random cluster measure. For details about this argument, see [5, Proof of Claim 4.2].   □

Lemma 6.15. *Suppose that p is small enough, $t \in [0, t_{max}]$ and the collection $A_1, A_2, \cdots, A_K$ of subsets of Ω_n is m-good for some $m \in \left[\log^4 n, (1/2)\log^5 n\right]$. Then under the notations of Definition 6.10, we have*

$$d(t) + \frac{2}{n^{2d}} \geq \sup_{x_0 \in \Omega_n} \|\mathbb{P}_{x_0}[Y_t^*(\Gamma) \in \cdot] - \pi_\Gamma\|_{TV} ,$$

where $d(t)$ the total-variation distance at time t was defined in Section 1.3.

Proof. Since projection does not increase total-variation norm, we have

$$d(t) \geq \sup_{x_0 \in \Omega_n} \|\mathbb{P}_{x_0}[X_t(\Gamma) \in \cdot] - \mu_\Gamma\|_{\mathrm{TV}} . \tag{6.8}$$

By Lemma 6.11, we have

$$\sup_{x_0 \in \Omega_n} \|\mathbb{P}_{x_0}[X_t(\Gamma) \in \cdot] - \mathbb{P}_{x_0}[Y_t^*(\Gamma) \in \cdot]\|_{\mathrm{TV}} \leq \frac{1}{n^{2d}} . \tag{6.9}$$

By combining (6.8), (6.9), and Lemma 6.14, the proof is completed.   □

We are finally ready to finish the proof of Theorem 6.1

6.3.1. *Proof of part (1): upper bound.* In view of Proposition 6.9, it suffices to prove the following proposition.

Proposition 6.16. *Suppose that p is sufficiently small, $\Gamma \in Spa_n$, and $t \in [0, t_{max}]$. Then, we have*

$$\|\mathbb{P}_\nu[X_t(\Gamma) \in \cdot\,] - \mu_\Gamma\|_{TV} \leq \frac{1}{2}\left[\exp\left\{\frac{n^d}{\log^{12d} n}\mathbf{d}_t^2\right\} - 1\right]^{1/2} + \frac{2}{n^{2d}} . \tag{6.10}$$

Proof. Denote by $A_1, A_2, \dots, A_K$ the connected components of Γ in the sense of Definition 6.7. Then, then $A_1, A_2, \dots, A_K$ are m-good with $m = \log^4 n$. Now we recall the notations from Definition 6.10 and Lemma 6.12. We bound the total-variation norm at the left-hand side of (6.10) by

$$\|\mathbb{P}_\nu[X_t(\Gamma) \in \cdot\,] - \mathbb{P}_\nu[Y_t^*(\Gamma) \in \cdot\,]\|_{\mathrm{TV}} + \|\mathbb{P}_\nu[Y_t^*(\Gamma) \in \cdot\,] - \pi_\Gamma\|_{\mathrm{TV}} + \|\pi_\Gamma - \mu_\Gamma\|_{\mathrm{TV}} . \tag{6.11}$$

We recall Notation 6.13 for the notation $\mathbb{P}_\nu[Y_t^*(\Gamma) \in \cdot\,]$. We now bound these three terms separately to complete the proof. For the first term, by Lemma 6.11 we have

$$\|\mathbb{P}_\nu[X_t(\Gamma) \in \cdot\,] - \mathbb{P}_\nu[Y_t^*(\Gamma) \in \cdot\,]\|_{\mathrm{TV}} \leq \frac{1}{n^{2d}} . \tag{6.12}$$

By the Lemma 6.12, and the fact $K \leq n^d/\log^{12d} n$, the second term is bounded by

$$\|\mathbb{P}_\nu[Y_t^*(\Gamma) \in \cdot\,] - \pi_\Gamma\|_{\mathrm{TV}} \leq \frac{1}{2}\left[\exp\left\{\frac{n^d}{\log^{12d} n}\mathbf{d}_t^2\right\} - 1\right]^{1/2} . \tag{6.13}$$

Finally, the last term at (6.11) is at most $1/n^{2d}$ by Lemma 6.14. Combining this with (6.11), (6.12), and (6.13), we can finish the proof. $\square$

6.3.2. *Proof of Part (2): lower bound.* Given the above ingredients the proof of the lower bound is almost verbatim from [20, Section 3.3] but nonetheless we include the proof in the appendix for completeness.

In the following section we finish the proof of Theorem 1.1.

7. Proof of main result

We keep the notation $r = 3\log^5 n$. The following lemma provides a sharp bound on $\mathbf{d}_t$.

Lemma 7.1. [20, Lemma 4.1] *For all small enough p, there exists a constant $C_3 = C_3(p) > 0$ such that*

$$e^{-\lambda(r)(t+C_3 \log\log n)} - n^{-2d} \leq \mathbf{d}_t \leq e^{-\lambda(r)(t-C_3 \log\log n)} \tag{7.1}$$

for all $t \in [C_3 \log\log n, t_{max}]$.

Proof. Since

$$\mathbf{d}_t \leq \max_{x_0^\dagger \in \Omega_r} \left\|\mathbb{P}_{x_0^\dagger}[X_t^\dagger \in \cdot\,] - \pi^\dagger\right\|_{L^2(\pi^\dagger)} ,$$

the upper bound part of (7.1) is immediate from Theorem 5.1. We note from this bound that

$$r^{d/2}\mathbf{d}_t = o(1) \text{ for } t = C \log\log n \tag{7.2}$$

with sufficiently large C. Here we implicitly used Corollary 3.4. For the lower bound part, we first recall the bound

$$e^{-\lambda(r)t} \leq 2 \max_{x_0^\dagger \in \Omega_r} \left\|\mathbb{P}_{x_0^\dagger}[X_t^\dagger \in \cdot\,] - \pi^\dagger\right\|_{\mathrm{TV}} \quad \text{for all } t \geq 0 ,$$

which is the continuous time version of Proposition 3.1 (see [19, Lemma 20.11]). By Lemma 4.4 (in particular, (4.2)), we have that

$$e^{-\lambda(r)(t+t_{\text{init}})} \leq 2 \max_{\nu:\nu \in \mathrm{Perc}_r} \left\|\mathbb{P}_{x_0^\dagger}[X_t^\dagger \in \cdot\,] - \pi^\dagger\right\|_{\mathrm{TV}} \quad \text{for all } t \geq 0 .$$

We now take $s = C_1 \log \log n$ and $t \in [C \log \log n, t_{\max}]$, where C_1 and C are the constants appeared in Theorem 6.1 and in (7.2), respectively. Then, by the previous inequality and part (1) of Theorem 6.1 for the lattice $\mathbb{Z}_r^d$ (cf. Remark 6.2 with $m = r = 3 \log^5 n$), we have that

$$e^{-\lambda(r)(t+t_{\text{init}}+C_1 \log \log n)} \leq 2 \max_{\nu:\nu \in \text{Perc}_r} \left\| \mathbb{P}_{x_0^\dagger}[X_{t+s}^\dagger \in \cdot] - \pi^\dagger \right\|_{\text{TV}}$$

$$\leq \left[\exp\{r^d \mathbf{d}_t^2\} - 1\right]^{\frac{1}{2}} + 8n^{-2d} \leq 2r^{d/2}\mathbf{d}_t + 8n^{-2d} ,$$

where the last inequality follows from (7.2) and the elementary inequality $e^x - 1 \leq 4x$ for $x \in [0, 1]$. We can deduce the lower bound from this computation. $\qquad\square$

Given the above, the proof of Theorem 1.1 involves two steps:

- Prove a version (Proposition 7.2) with λ_∞ replaced by $\lambda(r)$ where r was chosen above.
- Show that $\lambda(r)$ converges to λ_∞ and have bounds on the convergence rate (Proposition 7.3).

Define

$$t(n) = \frac{d}{2\lambda(r)} \log n \quad \text{and} \quad w(n) = \log \log n .$$

Proposition 7.2. *For all small enough p, there exist two constants $c_1 = c_1(p)$, $c_2 = c_2(p)$ such that*

$$\lim_{n\to\infty} \max_{x_0 \in \Omega_n} \left\| \mathbb{P}_{x_0} \left[X_{t(n)-c_1 w(n)} \in \cdot\right] - \mu_{p,q}^n \right\|_{TV} = 1 , \tag{7.3}$$

$$\lim_{n\to\infty} \max_{x_0 \in \Omega_n} \left\| \mathbb{P}_{x_0} \left[X_{t(n)+c_2 w(n)} \in \cdot\right] - \mu_{p,q}^n \right\|_{TV} = 0 . \tag{7.4}$$

Proof. By Lemma 4.4, it suffices to consider the initial condition ν satisfying $\nu \preceq \text{Perc}_n(p_{\text{init}})$. We recall the constant C_3 from the statement of Lemma 7.1. First, by the lower bound in Lemma 7.1 and by part (3) of Corollary 3.4,

$$\left(\frac{n}{\log^{10} n}\right)^d \mathbf{d}_{t(n)-c_1 w(n)}^2 \geq \left(\frac{n}{\log^{10} n}\right)^d e^{-2\lambda(r)(t(n)-c_1 w(n)+C_3 \log \log n)} - n^{-d}$$

$$\geq (\log n)^{2\lambda(c_1-C_3)-10d} - n^{-d} .$$

Therefore, for $c_1 > C_3 + \frac{11d}{2\lambda}$, we have

$$\lim_{n\to\infty} \left(\frac{n}{\log^{10} n}\right)^d \mathbf{d}_{t(n)-c_1 w(n)}^2 = +\infty ,$$

and thus by part (2) of Theorem 6.1 we obtain (7.3). Now we turn to (7.4). For $c \in (0, C_3)$, by the upper bound of Lemma 7.1,

$$\frac{n^d}{\log^{12d} n} \mathbf{d}_{t(n)+cw(n)}^2 \leq \frac{n^d}{\log^{12d} n} e^{-2\lambda(r)(t(n)+cw(n)-C_2 \log \log n)} \leq (\log n)^{\lambda(C_2-c)-12d} .$$

By taking c close enough to C_2 we obtain

$$\frac{n^d}{\log^{12d} n} \mathbf{d}_{t(n)+cw(n)}^2 \leq \frac{1}{\log^{11d} n} . \tag{7.5}$$

Let $c_2 = C_1 + c$ where C_1 is the constant appeared in Theorem 6.1. Then, by part (1) of Theorem 6.1 (note that this is where the sparsity assumption on ν is used) and (7.5),

$$\max_{\nu:\nu \preceq \text{Perc}_n(t_{\text{init}})} \left\| \mathbb{P}_\nu \left[X_{t(n)+c_2 w(n)} \in \cdot\right] - \mu_{p,q}^n \right\|_{\text{TV}} \leq \frac{1}{2} \left[\exp\left\{\frac{n^d}{\log^{12d} n} \mathbf{d}_{t(n)+cw(n)}^2\right\} - 1\right]^{\frac{1}{2}} + \frac{4}{n^{2d}}$$

$$\leq \frac{1}{2} \left[\exp \frac{1}{\log^{11d} n} - 1\right]^{\frac{1}{2}} + \frac{4}{n^{2d}} .$$

This completes the proof of (7.4). $\qquad\square$

Notice that since $\lambda(r) = \Theta(1)$ (Corollary 3.4), we have $w(n) \ll t(n)$, and therefore the previous proposition already demonstrates the cutoff phenomenon provided that p is small enough. The next result shows that the sequence $(\lambda(r))_{r \geq 1}$ is a convergent sequence.

Proposition 7.3. [20, Lemma 4.3] *There exists $\lambda_\infty = \lambda_\infty(p) > 0$ such that*

$$|\lambda(r) - \lambda_\infty| \leq r^{-1/4 + o(1)} .$$

Proof. We only provide the modified choice of parameters needed for our purpose. A careful reading of the proof shows that entire arguments presented above are still in force if we replace $r = 3\log^5 n$ with $r = \log^{4+\delta}$ for any δ. Of course the constants that we obtained above must be modified to depend on δ, and the time $t_{\max}$ should be defined as $\log^{1+\delta} n$ (cf. Remark 4.14). Taking $r_1 = \log^{4+\delta}$ and $r_2 \in [r_1, r_1^2]$ and applying Proposition 7.2 with $r = r_1$ and $r = r_2$, respectively, yields

$$\frac{d}{2\lambda(r_1)}\log n - Cw(n) \leq \frac{d}{2\lambda(r_2)}\log n + Cw(n)$$

for some constant $C = C(p, \delta)$. Since $\lambda(\cdot)$ is bounded below, we obtain

$$\lambda(r_1) - \lambda(r_2) \leq C\frac{\log\log n}{\log n} \leq r_1^{-1/4+\delta}$$

for all sufficiently large n. The rest of the arguments are exactly the same as [20, Lemma 4.3] and are omitted.

$\square$

7.1. **Proof of Theorem 1.1.** As mentioned before we can combine Propositions 7.2 and 7.3 to deduce Theorem 1.1. Define

$$t^*(n) = \frac{d}{2\lambda_\infty}\log n .$$

Thus we need to show that for all small enough p, there exist two constants $c_1 = c_1(p)$, $c_2 = c_2(p)$ such that

$$\lim_{n \to \infty} \max_{x_0 \in \Omega_n} \left\| \mathbb{P}_{x_0}\left[X_{t^*(n)-c_1 w(n)} \in \cdot \right] - \mu_{p,q}^n \right\|_{\mathrm{TV}} = 1 , \tag{7.6}$$

$$\lim_{n \to \infty} \max_{x_0 \in \Omega_n} \left\| \mathbb{P}_{x_0}\left[X_{t^*(n)+c_2 w(n)} \in \cdot \right] - \mu_{p,q}^n \right\|_{\mathrm{TV}} = 0 . \tag{7.7}$$

The proof is now immediate from

$$|t^*(n) - t(n)| \leq C|\lambda(r) - \lambda_\infty|\log n \leq C\log^{-1/4} n .$$

$\square$

7.2. **Comparison to infinite volume dynamics.** It is quite natural to predict that λ_∞ is in fact the spectral gap of the infinite volume FK-dynamics with the same parameters p and q. Defining the latter is not trivial but this has been carried out in [17, Chapter 8]. For the Ising model a similar result was shown in [20] using the monotonicity of the underlying dynamics as well as Log-Sobolev inequalities. Even though the lack of monotonicity of the Potts model prevented the authors in [21] to prove a similar conclusion, this was settled in [27, Section 6.2] using the Information Percolation machinery which also implies the same for SW dynamics. Furthermore in [27], the authors remark that the argument relies on bounds on disagreement propagation and an infinite version of the exponential L^2-mixing rate and hence holds in more generality for spin systems.

Thus in our context of the FK-dynamics to prove a similar result, given the disagreement propagation bounds, stated in Section 4.4, the only remaining step is to establish an analog of Theorem 5.1 for the infinite system by proving an analog of Proposition 5.2 in the same setting. The argument in [27] proceeds by defining Information Percolation clusters for the infinite process. We believe that this can be carried out in our setting as well, by suitable extensions of the arguments

for finite systems presented in Section 5. However, we do not pursue verifying the precise details in the paper.

8. APPENDIX

We provide the proofs that were omitted from the main article.

Proof of Lemma 4.11. We recall notations from Section 4.1 and in particular $t_{\max}$ from (4.10) and write $t_{\max} = L\Delta$ so that
$$0 = \tau_0 < \tau_1 < \cdots < \tau_L = \log^2 n .$$
We regard $(Z_t^-)_{t\geq 0}$ as a Markov chain on Ω_n such that the configuration outside of A^+ is empty. Also consider $(X_t)_{t\geq 0}$, the FK-dynamics on Ω_n starting from an initial condition which agrees with (Z_t^-) on A^+. One can observe that under the monotone coupling,
$$Z_t^- \leq X_t \leq Z_t^+ \text{ for all } t \geq 0 . \tag{8.1}$$
We recall the enlarged percolation $\overline{\mathscr{F}}_i$ from Table 1, and denote by $\mathcal{E}_i$ the event that there is no open path of length $\beta \log n$ in $\overline{\mathscr{F}}_i$ for some $\beta = \beta(p) > 0$. Then, by Proposition 4.9 and the union bound,
$$\mathbb{P}[\mathcal{E}_i] \leq \binom{|E_n|}{2} \exp\{-\gamma\beta \log n\} < \frac{1}{n^{4d}}$$
provided that β is large enough. Define $\mathcal{E} = \bigcup_{i=1}^{L} \mathcal{E}_i$. Since $L = t_{\max}/\Delta = \Omega(\log^2 n)$, the union bound implies that
$$\mathbb{P}[\mathcal{E}] \leq L\frac{1}{n^{4d}} < \frac{1}{n^{3d}} . \tag{8.2}$$

We now claim that $\mathcal{E}^c$ implies that $Z_t^+(A) = Z_t^-(A)$ for all $t \in [0, t_{\max}]$. Thus the claim along with (8.2), finishes the proof of the lemma. To prove the claim, define A_i, $i \in [\![0, L]\!]$, inductively as $A_L = A$ and
$$A_{i-1} = \{e \in E_n : d(e, A_i) \leq t_{\max}\Delta\} ,$$
so that
$$A^+ = A_0 \supset A_1 \supset \cdots \supset A_L = A .$$
For all $i \in [\![0, L-1]\!]$, we shall prove that $Z_{\tau_i}^+(A_i) = Z_{\tau_i}^-(A_i)$ implies $Z_t^+(A_{i+1}) = Z_t^-(A_{i+1})$ for all $t \in (\tau_i, \tau_{i+1}]$. since then the proof of the claim is completed by the induction. Now it suffices to observe that there exists a closed surface of $\overline{\mathscr{F}}_i$ in $A_i\backslash A_{i+1}$ under $\mathcal{E}_i$ since the set $\bigcup_{e\in\partial A_i} \mathrm{Conn}(e; \overline{\mathscr{F}}_i)$ is disjoint to A_{i+1} as there is no connected path of length $\Omega(\log^2 n)$ in $\overline{\mathscr{F}}_i$ (call this surface as V_i). The proof now follows by noticing that the FK-dynamics for both Z_t^+ and Z_t^- agree on the component of $E_n \setminus V_i$ (say $\tilde{A}_i$) containing A_{i+1} (and hence on A_{i+1}) throughout $[\tau_i, \tau_{i+1}]$, since the starting configurations for both the chains agree on $\tilde{A}_i$ by induction and the dynamics has zero boundary condition throughout $[\tau_i, \tau_{i+1}]$. $\qquad\square$

Proof of Theorem 6.1, Part (2): lower bound. Recall $r = 3\log^5 n$, and let us divide $\mathbb{Z}_n^d$ by $K = \lfloor n/r \rfloor^d$ square boxes $A_1^+, A_2^+, \cdots, A_K^+$ of size r as we did in Section 6.1. Then, let A_i be the box of size $2r/3$ which is concentric with A_i^+. Then, the collection $A_1, A_2, \cdots, A_K$ is m-good with $m = (1/2)\log^5 n$. We recall the notations from Definition 6.10. By definition (6.1) of $\mathbf{d}_t$, we can find $x_0^* = x_0^*(t) \in \Omega_r$ satisfying $\mathbf{d}_t = \|\mathbb{P}_{x_0^*}[X_t^\dagger(\Lambda) \in \cdot] - \pi_\Lambda^\dagger\|_{L^2(\pi_\Lambda^\dagger)}$. Let U_i be a configurations on B_i distributed according to $\pi_{A_i}^{(i)}$, where $\{U_i, 1 \leq i \leq K\}$ is a collection of independent random variables. Define a sequence of i.i.d. random variable u_i as
$$u_i = \frac{\mathbb{P}[Y_t^{(i)}(A_i) = U_i \,|\, Y_0^{(i)} = x_0^*]}{\pi_{B_i}^{(i)}(U_i)} \;\; ; \; i \in [\![1, K]\!] .$$

The condition $Y_0^{(i)} = x_0^*$ means $X_0(\widehat{A}_i^+) = x_0^*$. By the definition of u_i and x_0^*, one can readily check that

$$\mathbb{E}u_i = 1 \ \text{ and } \ \operatorname{Var} u_i = \mathbf{d}_t^2 \ . \tag{8.3}$$

By the L^∞-L^2 reduction for reversible Markov chains, we have

$$\|u_i - 1\|_\infty = \left\| \mathbb{P}\left[Y_t^{(i)}(A_i) \in \cdot \,|\, Y_0^{(i)} = x_0^* \right] - \pi_{A_i}^{(i)} \right\|_{L^\infty(\pi_{A_i}^{(i)})}$$

$$\leq \left\| \mathbb{P}\left[Y_{t/2}^{(i)}(A_i) \in \cdot \,|\, Y_0^{(i)} = x_0^* \right] - \pi_{A_i}^{(i)} \right\|_{L^2(\pi_{A_i}^{(i)})} \leq \mathbf{d}_{t/2} \ .$$

Hence, by Theorem 5.1 we obtain

$$\|u_i - 1\|_\infty \leq e^{-c \log \log n} \tag{8.4}$$

for some $c > 0$. Then, by (8.3) and (8.4), we have

$$\mathbb{E}|u_i - 1|^3 \leq e^{-c \log \log n} \, \mathbf{d}_t^2 = o(1) \, \mathbf{d}_t^2 \ . \tag{8.5}$$

Given the above inputs, the rest of the proof follows by arguments identical to [20, Section 3.3] and is omitted. $\qquad\square$

References

[1] David Aldous and Persi Diaconis. Shuffling cards and stopping times. *The American Mathematical Monthly*, 93(5):333–348, 1986.

[2] Vincent Beffara and Hugo Duminil-Copin. The self-dual point of the two-dimensional random-cluster model is critical for $q > 1$. *Probability Theory and Related Fields*, 153(3-4):511–542, 2012.

[3] Antonio Blanca, Reza Gheissari, and Eric Vigoda. Random-cluster dynamics in $\mathbb{Z}^2$: rapid mixing with general boundary conditions. *arXiv preprint arXiv:1807.08722*, 2018.

[4] Antonio Blanca and Alistair Sinclair. Dynamics for the mean-field random-cluster model. *arXiv preprint arXiv:1412.6180*, 2014.

[5] Antonio Blanca and Alistair Sinclair. Random-cluster dynamics in $\mathbb{Z}^2$. *Probability Theory and Related Fields*, 168(3-4):821–847, 2017.

[6] Persi Diaconis. The cutoff phenomenon in finite Markov chains. *Proceedings of the National Academy of Sciences*, 93(4):1659–1664, 1996.

[7] Persi Diaconis and Laurent Saloff-Coste. Logarithmic sobolev inequalities for finite Markov chains. *The Annals of Applied Probability*, 6(3):695–750, 1996.

[8] Hugo Duminil-Copin. Introduction to Bernoulli percolation. *www.ihes.fr/duminil/publi/2017percolation.pdf*, 2018.

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

[10] Hugo Duminil-Copin, Vladas Sidoravicius, and Vincent Tassion. Continuity of the phase transition for Planar Random-Cluster and Potts models with $1 \leq q \leq 4$. *Communications in Mathematical Physics*, 349(1):47–107, 2017.

[11] Robert G Edwards and Alan D Sokal. Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm. *Physical review D*, 38(6):2009, 1988.

[12] Cornelis Marius Fortuin. On the random-cluster model: III. the simple random-cluster model. *Physica*, 59(4):545–570, 1972.

[13] Cornelius Marius Fortuin and Piet W Kasteleyn. On the random-cluster model: I. introduction and relation to other models. *Physica*, 57(4):536–564, 1972.

[14] Reza Gheissari and Eyal Lubetzky. Quasi-polynomial mixing of critical 2d random cluster models. *arXiv preprint arXiv:1611.01147*, 2016.

[15] Reza Gheissari and Eyal Lubetzky. The effect of boundary conditions on mixing of 2d Potts models at discontinuous phase transitions. *Electronic Journal of Probability*, 23, 2018.

[16] Reza Gheissari and Eyal Lubetzky. Mixing times of critical two-dimensional Potts models. *Communications on Pure and Applied Mathematics*, 71(5):994–1046, 2018.

[17] Geoffrey Grimmett. Percolation, 1989.

[18] Geoffrey Grimmett. The random-cluster model. In *Probability on discrete structures*, pages 73–123. Springer, 2004.

[19] David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer. *Markov chains and mixing times*. American Mathematical Soc., 2009.

[20] Eyal Lubetzky and Allan Sly. Cutoff for the Ising model on the lattice. *Inventiones mathematicae*, 191(3):719–755, 2013.

[21] Eyal Lubetzky and Allan Sly. Cutoff for general spin systems with arbitrary boundary conditions. *Communications on Pure and Applied Mathematics*, 67(6):982–1027, 2014.

[22] Eyal Lubetzky and Allan Sly. An exposition to information percolation for the Ising model. *Annales de la Faculté des Sciences de Toulouse*, 29(4):745–761, 2015.

[23] Eyal Lubetzky and Allan Sly. Information percolation and cutoff for the stochastic Ising model. *Journal of the American Mathematical Society*, 29(3):729–774, 2016.

[24] Eyal Lubetzky and Allan Sly. Universality of cutoff for the Ising model. *The Annals of Probability*, 45(6A):3664–3696, 2017.

[25] Jason Miller and Yuval Peres. Uniformity of the uncovered set of random walk and cutoff for lamplighter chains. *The Annals of Probability*, 40(2):535–577, 2012.

[26] Elchanan Mossel and Allan Sly. Exact thresholds for Ising–Gibbs samplers on general graphs. *The Annals of Probability*, 41(1):294–328, 2013.

[27] Danny Nam and Allan Sly. Cutoff for the Swendsen-Wang dynamics on the lattice. *arXiv preprint arXiv:1805.04227*, 2018.

[28] Y Peres. American institute of mathematics (AIM) research workshop 'sharp thresholds for mixing times'(Palo Alto, december 2004). *Summary available at http://www. aimath.org/WWN/mixingtimes*, 12, 2004.

[29] James Propp and David Wilson. Coupling from the past: a user's guide. *Microsurveys in Discrete Probability*, 41:181–192, 1998.

[30] Robert H Swendsen and Jian-Sheng Wang. Nonuniversal critical dynamics in Monte Carlo simulations. *Physical review letters*, 58(2):86, 1987.

[31] Mario Ullrich. Rapid mixing of Swendsen-Wang dynamics in two dimensions. *arXiv preprint arXiv:1212.4908*, 2012.

[32] Mario Ullrich. Comparison of Swendsen-Wang and heat-bath dynamics. *Random Structures & Algorithms*, 42(4):520–535, 2013.

[33] Mario Ullrich. Swendsen–wang is faster than single-bond dynamics. *SIAM Journal on Discrete Mathematics*, 28(1):37–48, 2014.

S. Ganguly

Department of Statistics, UC Berkeley, Berkeley, California, CA 94720, USA.
E-mail address: sganguly@berkeley.edu

I. Seo

Department of Mathematical Science and RIMS, Seoul National University, Seoul, South Korea.
E-mail address: insuk.seo@snu.ac.kr