RANDOM-CLUSTER DYNAMICS ON RANDOM REGULAR GRAPHS IN TREE UNIQUENESS

ANTONIO BLANCA AND REZA GHEISSARI

ABSTRACT. We establish rapid mixing of the random-cluster Glauber dynamics on random Δ-regular graphs for all $q \geq 1$ and $p < p_u(q, \Delta)$, where the threshold $p_u(q, \Delta)$ corresponds to a uniqueness/non-uniqueness phase transition for the random-cluster model on the (infinite) Δ-regular tree. It is expected that this threshold is sharp, and for $q > 2$ the Glauber dynamics on random Δ-regular graphs undergoes an exponential slowdown at $p_u(q, \Delta)$.

More precisely, we show that for every $q \geq 1$, $\Delta \geq 3$, and $p < p_u(q, \Delta)$, with probability $1 - o(1)$ over the choice of a random Δ-regular graph on n vertices, the Glauber dynamics for the random-cluster model has $\Theta(n \log n)$ mixing time. As a corollary, we deduce fast mixing of the Swendsen–Wang dynamics for the Potts model on random Δ-regular graphs for every $q \geq 2$, in the tree uniqueness region. Our proof relies on a sharp bound on the "shattering time", i.e., the number of steps required to break up any configuration into $O(\log n)$ sized clusters. This is established by analyzing a delicate and novel iterative scheme to simultaneously reveal the underlying random graph with clusters of the Glauber dynamics configuration on it, at a given time.

1. INTRODUCTION

The random-cluster model is a random graph model, unifying the study of electrical networks, independent bond percolation, and the ferromagnetic Ising/Potts model from statistical physics [21,31]. It is defined on a graph $G = (V, E)$ and parametrized by an edge probability $p \in (0, 1)$ and cluster weight $q > 0$. Each configuration consists of a subset of edges $\omega \subseteq E$ (equivalently $\omega \in \{0, 1\}^E$) and is assigned probability

$$\pi_{G,p,q}(\omega) = \frac{1}{Z_{G,p,q}} p^{|\omega|}(1-p)^{|E|-|\omega|} q^{c(\omega)}\,, \tag{1.1}$$

where $c(\omega)$ is the number of connected components in (V, ω) and $Z_{G,p,q}$ is a normalizing constant.

Aside from its inherent interest as a model of random networks, the random-cluster model provides an elegant class of Markov Chain Monte Carlo (MCMC) algorithms for sampling from the Ising/Potts model. For integer $q \geq 2$, a sample ω from (1.1) can be transformed into one for the q-state ferromagnetic Potts model by independently assigning a random spin from $\{1, \ldots, q\}$ to each connected component of (V, ω); see, e.g., [18, 31]. Random-cluster based sampling algorithms, which include the popular Swendsen–Wang algorithm [50], are a widely-used alternative to the standard Ising/Potts Markov chains since the former are often efficient at "low-temperatures" (large p) where the latter suffer exponential slowdowns (see [8, 33]).

Our focus here is on the Glauber dynamics of the random-cluster model. Specifically, we consider the following discrete-time Glauber dynamics chain, which we refer to as the *FK-dynamics*. From a configuration $\omega_t \subseteq E$, one step of the FK-dynamics transitions to a new configuration $\omega_{t+1} \subseteq E$ as follows:

(1) Choose an edge $e_t \in E$ uniformly at random;

(2) Set $\omega_{t+1} = \omega_t \cup \{e_t\}$ with probability $\begin{cases} \hat{p} := \frac{p}{q(1-p)+p} & \text{if } e_t \text{ is a "cut-edge" in } (V, \omega_t); \\ p & \text{otherwise;} \end{cases}$

(3) Otherwise set $\omega_{t+1} = \omega_t \setminus \{e_t\}$.

We say e is a *cut-edge* in (V, ω_t) if changing the state of e_t changes the number of connected components $c(\omega_t)$ in (V, ω_t). This chain is, by design, reversible with respect to $\pi_{G,p,q}$.

A central question in the study of Markov chains is how the *mixing time*—defined as the number of steps until the Markov chain is close to stationarity starting from the worst possible initial configuration—grows as the size of the graph G increases. Of particular interest in the context of random-cluster and Ising/Potts dynamics is the relation of mixing times to the rich equilibrium phase transitions of the model.

We consider this question when G is a random Δ-regular graph on n vertices. The study of spin systems and their dynamics on random graphs is quite active [12, 14–16, 19, 20, 25, 44, 45]. Random Δ-regular graphs

are a canonical example of graphs having exponential volume growth, with a non-trivial geometry, making them an attractive alternative to lattices or trees. More generally, the study of spin systems on random graphs yields insight into hard instances of the classical computational problems of sampling, counting, learning and testing [2, 23, 48, 49] and features in the study of random constraint satisfaction problems [32, 54].

The phase transition of the random-cluster model on random Δ-regular graphs is expected to involve three critical points [11, 34, 39, 41]. Most relevant to us would be the critical threshold $p_u(q,\Delta)$ corresponding to a *uniqueness/non-uniqueness* phase transition for the random-cluster model on the infinite Δ-regular wired tree (in which the leaves are externally wired to be in the same connected component). Roughly speaking, the uniqueness/non-uniqueness phase transition captures whether the wired boundary has an effect or not on the configuration near the root of the tree (in the limit as the height of the tree grows). It is believed that the mixing time slows down at $p_u(q,\Delta)$, either polynomially or exponentially depending on $q \le 2$ or $q > 2$.

In this paper, we establish optimal mixing for the FK-dynamics on random Δ-regular graphs throughout the uniqueness region $p < p_u(q,\Delta)$ for all real $q \ge 1$ and all $\Delta \ge 3$.

Theorem 1.1. *Fix any $q \ge 1, \Delta \ge 3$, and $p < p_u(q,\Delta)$. Consider the FK-dynamics on a uniformly random Δ-regular graph on n vertices. With probability $1 - o(1)$ over the choice of the random graph $\mathcal{G}$, the mixing time of the FK-dynamics on $\mathcal{G}$ is $\Theta\big(n \log n\big)$.*

The FK-dynamics are known to be resistant to sharp analysis with the known techniques for Markov chains for spin systems. This is due, in part, to the fact that the random-cluster model presents highly non-local interactions: an update on an edge e_t depends on the entire configuration $\omega_t(E \setminus \{e_t\})$. Indeed, the only other setting where the speed of convergence of FK-dynamics is well-understood via direct analysis is in square subsets of $\mathbb{Z}^2$ [6, 8, 26–29]. Other bounds to date have been obtained either indirectly, via comparison with global Markov chains using the results of [52, 53] (and as a result, these bounds are off by polynomial factors), or by taking either p very small (e.g., under a Dobrushin-type condition) or very large, or q large. This is the state of affairs even on the (geometrically trivial) complete graph [8, 30, 38].

Our results are tight in the sense that the FK-dynamics is expected to undergo a slowdown at $p_u(q,\Delta)$, as we describe next. The equilibrium phase transition of the random-cluster model on random Δ-regular graphs should qualitatively resemble those on the Δ-regular tree and the complete graph. Based on this relation, and understandings of those phase diagrams [11, 34, 39, 41], it is expected to involve three critical points $p_u(q,\Delta) \le p_c(q,\Delta) \le p_u^*(q,\Delta)$. The tree uniqueness/non-uniqueness phase transition at $p_u(q,\Delta)$ manifests on the finite Δ-regular tree in the form of existence/non-existence of root-to-leaf paths under wired boundary conditions. The threshold $p_u^*(q,\Delta)$ corresponds to a (conjectured) second non-uniqueness/uniqueness transition; above this point even the Δ-regular tree under free boundary conditions has root-to-leaf connections (see [25, 34, 36, 39] for more details). The threshold $p_c(q,\Delta)$, on the other hand, corresponds to an order-disorder transition captured by the emergence of a "giant component" of linear size on the random graph (which, roughly, imposes "typical" boundary conditions on its treelike balls).

When $q \in (1,2]$ the phase transition is of second-order and these three thresholds coincide; namely $p_u(q,\Delta) = p_c(q,\Delta) = p_u^*(q,\Delta)$. On the other hand when $q > 2$, the phase transition on random Δ-regular graphs is conjectured to be of *first-order* and $p_u(q,\Delta) < p_c(q,\Delta) < p_u^*(q,\Delta)$. Here, the uniqueness threshold $p_u(q,\Delta)$ should mark the onset of the *metastability* phenomenon, and that should persist up to $p_u^*(q,\Delta)$. Metastability has been linked to an exponential slowdown for both random-cluster and Potts Glauber dynamics on the complete graph [7, 13, 24, 30], and the same slowdown is expected to occur on random Δ-regular graphs. Namely, in the window $(p_u(q,\Delta), p_u^*(q,\Delta))$, the ordered and disordered phases should each be "metastable" behaving locally (on treelike balls) like the configurations on wired and free trees, respectively. The coexistence of these metastable phases with exponentially small boundaries, facilitates states from which reversible Markov chains cannot easily escape (i.e., these sets have bad conductance). It is thus expected that on random Δ-regular graphs, for every $q > 2$, the FK-dynamics mixes exponentially slowly throughout $(p_u(q,\Delta), p_u^*(q,\Delta))$. For q sufficiently large, such slowdown was established in [25] at $p = p_c(q,\Delta) \in (p_u(q,\Delta), p_u^*(q,\Delta))$.

From Theorem 1.1 we obtain an efficient MCMC sampling algorithm, for both the random-cluster model and the ferromagnetic Ising/Potts model on random Δ-regular graphs in the uniqueness regime.

Corollary 1.2. *Fix any $q \geq 1, \Delta \geq 3$, $p < p_u(q,\Delta)$ and any accuracy parameter $\delta \in (0,1)$. Then, with probability $1-o(1)$ over the choice of the random Δ-regular n-vertex graph $\mathcal{G}$, there is a sampling algorithm which, given the graph $\mathcal{G}$, outputs a random-cluster configuration ω whose distribution is within total variation distance δ of $\pi_{\mathcal{G},p,q}$. The running time of the algorithm is $O(n(\log n)^3 \log(1/\delta))$.*

The extra $O((\log n)^2)$ factor in the running time of the algorithm comes from the (amortized) computational cost of checking whether the chosen edge is a cut-edge in each step of the FK-dynamics. This is equivalent to the fully dynamic connectivity problem which has been thoroughly studied (see, e.g., [37,51]).

For integer q, the algorithm in Corollary 1.2 combined with the $O(n)$ cost of translating between the random-cluster and Potts configurations mentioned earlier yields a sampling algorithm for the ferromagnetic q-state Potts model on random regular graphs up to the Potts uniqueness threshold (the uniqueness thresholds of both these models coincide). This improves on the best previously known sampling algorithm for both these models in [5], which runs in $\tilde{O}(n^{6/5})$ time, and it is a "weak sampler" in the sense that it outputs samples that are close in total variation distance to the target distribution but with a *fixed* accuracy. (See also the recent work of [36] for a $\mathrm{poly}(n)$ sampler for *all* $p \in (0,1)$ but provided q is sufficiently large.)

As another important corollary of Theorem 1.1, we deduce fast mixing of the standard Swendsen-Wang (SW) algorithm for the ferromagnetic q-state Potts model [50]. This is an extensively-used global-update Markov chain. The dynamics starts from a Potts configuration $\sigma_t \in \{1,\dots,q\}^V$, moves to a "joint" spin/random-cluster configuration (σ_t, ω_t) by including each monochromatic edge independently with probability p and then assigns to each connected component of (V,ω_t) a uniform at random spin from $\{1,...,q\}$ to obtain a new Potts configuration σ_{t+1} (see [18,50]).

Corollary 1.3. *Fix any integer $q \geq 2$ and $\Delta \geq 3$, and let $p < p_u(q,\Delta)$. Consider the Swendsen-Wang dynamics on a uniformly random Δ-regular graph on n vertices. With probability $1-o(1)$ over the choice of the random graph $\mathcal{G}$, the mixing time of the Swendsen–Wang dynamics on $\mathcal{G}$ is $O\big(n^2 \log n\big)$.*

Corollary 1.3 follows immediately from Theorem 1.1 and the comparison results of Ullrich [52,53]. Previously, our understanding of the speed of convergence of the SW dynamics on random Δ-regular graphs was very limited. For the special case of $q=2$, which corresponds to the Ising model, it was established in [4] that the spectral gap of the SW dynamics is $\Omega(1)$ for all $p < p_u(2,\Delta)$; this implies an $O(n)$ mixing time bound. In addition, Guo and Jerrum [33] established an $O(n^{10})$ mixing time bound for the SW dynamics that applies to any graph and any $p \in (0,1)$. The methods in both of these works are specific to the Ising model ($q=2$) and do not generalize to other values of q. Beyond the special case of $q=2$, no sub-exponential bound was previously known for either the FK-dynamics or the SW dynamics throughout the uniqueness regime $p < p_u(q,\Delta)$.

Proof ideas. We comment briefly on the techniques and main innovations in our analysis next: for more details and an extended proof sketch, we refer the reader to Section 3. The main ingredient in our proof is an $O(n \log n)$ bound on the "shattering time" of the FK-dynamics (Theorem 3.2); this is the number of steps the chain requires to break up any configuration into connected components of size at most $O(\log n)$. The bound on the shattering time uses a novel and delicate iterative scheme to simultaneously reveal the underlying random graph and the connected components of the FK-dynamics configuration on it at a given time: see Definition 4.11 and Figures 4.1–4.2. While revealing procedures are a standard tool in the study of both random graphs and of the random-cluster model, their combined analysis is highly non-trivial, as the law of the random-cluster configuration at an edge depends on the global geometry of the graph. To our knowledge, this the first direct upper bound for the shattering time of the FK-dynamics in any setting. In fact, understanding the shattering time is usually the main obstacle for proving rapid mixing of the FK-dynamics on other graphs: e.g., on the complete graph, the shattering time is not known and only loose mixing time bounds (off by $\Theta(n^2)$ factors) can be derived [7].

Once the dynamics has shattered, we use standard methods (i.e., censoring [46]) to reduce the analysis of the FK-dynamics to localized dynamics in balls of radius $o(\sqrt{n})$ centered at each vertex, but with random boundary conditions induced by the current state outside the ball. In random Δ-regular graphs, these balls are "treelike" and, after shattering, their boundary conditions are "almost free", in that only $O(1)$ vertices in their boundaries are connected through the external configuration. This implies that the FK-dynamics mix quickly and satisfy a log-Sobolev inequality akin to a product measure in each of these balls. The last ingredient in our proof is an exponential decay of correlation property (sometimes called spatial mixing) between the root and boundary of such balls. A delicate point is that since these balls have radius $\Theta(\log n)$, we need exact control on the rate of this exponential decay to sustain the union bound over the n balls.

Remark 1.4. We expect our methods for the analysis of the shattering phase to have applications to other locally tree-like graphs, e.g., wired trees and Erdős–Rényi random graphs. In the latter case, however, the possibility of having a small number of vertices of large degree poses technical obstructions to direct extension of our methods. Whereas this should not affect the equilibrium phase diagram of the model, interestingly, in the case of the Glauber dynamics for the Ising model on an Erdős–Rényi random graph, the high maximum degree is known to slow down the high-temperature mixing time to $n^{1+\Omega(\frac{1}{\log\log n})}$ [44].

Organization of paper. The rest of the paper is organized as follows. In Section 2, we provide a number of preliminary definitions and notations we will use. In Section 3, we give a detailed proof overview highlighting some of the key novelties in our arguments. Our revealing procedures to bound the shattering time are the focus of Section 4. In Section 5 we establish the sharp rate of spatial mixing on treelike graphs with sparse boundary conditions. We combine these to conclude the proof of the upper bound of Theorem 1.1 in Section 6. We prove the matching lower bound on the mixing time in Section 7.

2. Preliminaries

In this section, we collect some standard definitions and properties that are necessary to present our proofs, and to which the reader can refer throughout. See the standard texts [10], [31], and [40] for more details on random graphs, the random-cluster model, and Markov chain mixing times, respectively.

2.1. Random Δ-regular graphs.

We begin by considering the underlying geometry we work on. Fix $\Delta \geq 3$ and consider the uniform distribution $\mathbf{P}_{\text{RRG}}$ over Δ-regular graphs on n vertices. (Let us always assume n is such that Δn is even, so that such a graph exists.) We identify the vertices $V(\mathcal{G})$ with the set $\{1,...,n\}$, and the randomness of $\mathbf{P}_{\text{RRG}}$ will be over the edge-subset of $\{ij = ji : 1 \leq i,j \leq n\}$. Throughout this paper, we set $d := \Delta - 1$ for convenience.

Random graphs are treelike. A key ingredient in our proof is the fact that random Δ-regular graphs are locally treelike. While this can be formalized in various ways, we use a notion that is most relevant to this paper, and applies uniformly to all vertices (as opposed to a vertex chosen uniformly at random).

For a graph $\mathcal{G} = (V(\mathcal{G}), E(\mathcal{G}))$ and a vertex $v \in V(\mathcal{G})$, we define the ball of radius R around v as:

$$B_R(v) := \{w \in V(\mathcal{G}) : d(w,v) \leq R\}\,,$$

where $d(w,v)$ is the graph distance. For a vertex set, $B \subset V(\mathcal{G})$, define $E(B) = \{v,w \in B : vw \in E(\mathcal{G})\}$.

Definition 2.1. We say that a graph $G = (V,E)$ is L-Treelike if there is a set $H \subset E$ with $|H| \leq L$ such that the graph $(V, E \setminus H)$ is a tree.

Definition 2.2. We say that a Δ-regular graph $\mathcal{G} = (V(\mathcal{G}), E(\mathcal{G}))$ is (L,R)-Treelike if for every $v \in V(\mathcal{G})$ the subgraph $(B_R(v), E(B_R(v)))$ is L-Treelike.

Fact 2.3. *Fix any $\Delta \geq 3$. For every $\delta > 0$, there exists $L(\delta,\Delta)$ such that if $R = (\frac{1}{2} - \delta)\log_d n$, we have*

$$\mathbf{P}_{\text{RRG}}\big(\mathcal{G} \text{ is } (L,R)\text{-Treelike}\big) = 1 - o(n^{-1})\,.$$

We include a short proof of Fact 2.3 after introducing the configuration model in Section 4.1. It is known that when $R > \frac{1}{2}\log_d n$, the number of cycles in every ball $B_R(v)$ goes to ∞ with n.

2.2. **The random-cluster model.** For a graph $G = (V, E)$, recall the definition of the random-cluster model from (1.1). We say an edge $e \in E$ is *open* or *wired* if $\omega(e) = 1$ and *closed* or *free* if $\omega(e) = 0$. We say two vertices are *connected in* ω if they are in the same connected component of the sub-graph $(V, \{e \in E : \omega(e) = 1\})$. For a vertex set $\mathcal{V} \subset V$, denote by $\mathcal{C}_{\mathcal{V}}(\omega)$ the union of connected components (clusters) containing $v \in \mathcal{V}$ in this sub-graph. For a configuration ω and edge set $A \subset E$, we use $\omega(A)$ for the restriction of ω to A.

Boundary conditions. To help study the random-cluster measure, we introduce boundary conditions.

Definition 2.4. A random-cluster *boundary condition* ξ on $G = (V, E)$ is a partition of V, such that the vertices in each element of the partition are identified with one another. The random-cluster measure with boundary conditions ξ, denoted $\pi^{\xi}_{G,p,q}$, is the same as in (1.1) except the number of connected components $c(\omega) = c(\omega; \xi)$ would be counted with this vertex identification, i.e., if v, w are in the same element of ξ, they are always counted as being in the same connected component of ω in (1.1). In this manner, the boundary condition can alternatively be seen as ghost "wirings" of the vertices in the same element of ξ.

The *free* boundary condition, $\xi = 0$, is the one whose partition consists only of singletons. For a subset $\partial V \subset V$, the *wired* boundary condition on ∂V, denoted $\xi = 1$, is the one whose partition has all vertices of ∂V in the same element and all vertices of $V \setminus \partial V$ as singletons; i.e., $\xi = \{\partial V\} \cup \bigcup\{v : v \in V \setminus \partial V\}$. For boundary conditions ξ, ξ' we say that $\xi \leq \xi'$ if ξ is a finer partition than ξ'. We have the following important monotonicity in boundary conditions: for any two boundary conditions ξ and ξ' with $\xi \geq \xi'$, we have $\pi^{\xi}_{G,p,q} \succcurlyeq \pi^{\xi'}_{G,p,q}$ where $\succcurlyeq$ denotes stochastic domination.

Uniqueness/non-uniqueness transition on the Δ-regular tree. As the geometry of the random graph is locally treelike, its dynamical transition point should be inherited from a transition on the Δ-regular tree. Throughout this paper, we denote by $\mathcal{T}_h := \mathcal{T}_{h,\Delta} = (V(\mathcal{T}_h), E(\mathcal{T}_h))$ the rooted (at ρ) Δ-regular complete tree of depth h (the root has Δ children, and all other vertices have $\Delta - 1$ children and one parent). Since the tree has depth $h < \infty$, evidently it is not actually Δ-regular, and has leaves $\partial\mathcal{T}_h = \{w \in V(\mathcal{T}_h) : d(\rho, w) = h\}$ (where $d(\cdot, \cdot)$ denotes graph distance); observe that

$$|V(\mathcal{T}_h)| = 1 + \Delta \sum_{i=1}^{h} d^{i-1} \leq 2\Delta d^h \,, \qquad \text{and} \qquad |E(\mathcal{T}_h)| = \Delta \sum_{i=1}^{h} d^{i-1} \leq 2\Delta d^h \,, \tag{2.1}$$

and $|\partial\mathcal{T}_h| = \Delta d^{h-1}$. The wired boundary condition "1" is the one that wires all vertices of $\partial\mathcal{T}_h$ together.

For every $\Delta \geq 3$ and $q \geq 1$, the random-cluster measure $\pi^{1}_{\mathcal{T}_h,p,q}$ undergoes a transition at $p_u(q, \Delta)$: when $p < p_u(q, \Delta)$ the probability that ρ is connected to $\partial\mathcal{T}_h$ in ω goes to 0 as $h \to \infty$, whereas when $p > p_u(q, \Delta)$ it stays bounded away from zero [34]. (While in general $p_u(q, \Delta)$ does not have a closed form, it can be expressed as the root of an explicit formula: see [5, 34].) A key fact (see [34, Theorem 1.5]) we will use is that whenever $p < p_u(q, \Delta)$ we have that $\hat{p}$ (the probability of a cut-edge being open) satisfies

$$\hat{p} := \frac{p}{q(1-p) + p} < \frac{1}{d} \,, \qquad \text{where} \qquad d := \Delta - 1 \,. \tag{2.2}$$

2.3. **Markov chain mixing times.** Consider a (discrete-time) Markov chain with transition matrix P on a finite state space Ω, reversible with respect to an invariant distribution π; denote the chain initialized from x_0 by $(X_t^{x_0})_{t \geq 0}$. Its *mixing time* is given by

$$t_{\text{MIX}} = t_{\text{MIX}}(1/4) \,, \qquad \text{where} \qquad t_{\text{MIX}}(\varepsilon) = \min\{t : \max_{x_0 \in \Omega} \|P(X_t^{x_0} \in \cdot) - \pi\|_{\text{TV}} \leq \varepsilon\} \,,$$

where the total-variation distance between μ and ν is given by

$$\|\mu - \nu\|_{\text{TV}} = \frac{1}{2}\|\mu - \nu\|_1 = \inf_{(U,V) \sim \mathbb{P} : U \sim \mu, V \sim \nu} \mathbb{P}(U \neq V) \,.$$

Here the infimum runs over all couplings of μ, ν. By this definition, to bound the mixing time, it suffices to bound the *coupling time* of the dynamics; i.e., if we construct a coupling $\mathbb{P}$ of the steps of the chain

such that for each $x_0, y_0 \in \Omega$, we have $\mathbb{P}(X_T^{x_0} \neq X_T^{y_0}) \leq 1/4$, then $t_{\text{MIX}} \leq T$. It is a standard fact that $t_{\text{MIX}}(\delta) \leq t_{\text{MIX}} \log(2\delta^{-1})$. See chapters 4–5 of [40] for more details.

A coupling for the FK-dynamics. Recall the definition of the FK-dynamics from the introduction. Note that in the presence of boundary conditions ξ, the only change is that in step (2) of the FK-dynamics transitions, the status of e being a cut-edge is dictated by whether its presence changes $c(\omega_t; \xi)$.

For the FK-dynamics, there is a canonical choice of coupling known as the *identity coupling*. This is the coupling that couples the evolution of two copies of the FK-dynamics, $(X_t^{x_0})$ and $(X_t^{y_0})$, by using the same random edge e_t and the same uniform random number $U_{e_t,t}$ to decide whether to add or remove e_t. When $q \geq 1$, the identity coupling is a *monotone coupling*, in the sense that if $X_t^{x_0} \leq X_t^{y_0}$ then $X_{t+1}^{x_0} \leq X_{t+1}^{y_0}$ with probability 1. The identity coupling can also be extended to a simultaneous coupling of all the Markov chains $(X_t^{x_0})$ indexed by their initial configuration $x_0 \in \{0,1\}^E$ (i.e., a *a grand coupling*), so that if $x_0 \leq y_0$ we have $X_t^{x_0} \leq X_t^{y_0}$ for all $t \geq 0$. As a consequence, the coupling time starting from any pair of configurations is bounded by the coupling time starting from the free $x_0 = 0$ and wired $y_0 = 1$ configurations.

3. Extended proof sketch

In this section, we provide a detailed sketch of our proof of Theorem 1.1, outlining the structure of the argument and highlighting some of the key technical difficulties we encountered. Most of the paper is dedicated to upper bounding the mixing time of the FK-dynamics by $O(n \log n)$, so the sequel is dedicating to sketching that proof. The matching lower bound follows from coupling a certain projection of the FK-dynamics to a product chain and is derived in Section 7.

3.1. Proof outline.

Let $\mathcal{G} = (V(\mathcal{G}), E(\mathcal{G}))$ be an n-vertex graph. Let $(X_t^1)_{t\geq 0}$ and $(X_t^0)_{t\geq 0}$ be two realizations of the FK-dynamics started from the all-wired and all-free configurations, respectively, and coupled via the identity coupling as defined in Section 2.

Our goal is to show that there exists $T = O(n \log n)$ such that for every vertex $v \in V(\mathcal{G})$, with probability $1 - o(n^{-1})$, the configurations X_T^1 and X_T^0 agree on the Δ edges incident v, denoted

$$\mathcal{N}_v := \{e \in E(\mathcal{G}) : v \in e\}. \tag{3.1}$$

A union bound over the n vertices would then imply that under the identity coupling $X_T^1 = X_T^0$ with probability $1 - o(1)$. By the monotonicity of the FK-dynamics under the identity coupling, this would show that the mixing time of the FK-dynamics is at most $T = O(n \log n)$.

There are two key stages to establishing this coupling, each of which we describe next.

Stage I. In the first stage of the coupling, we show that after an initial *burn-in period* of $T = O(n \log n)$ steps, the configuration X_T^1 is *shattered*. That is, its connected components have constant size in expectation, and every component is of size $O(\log n)$ with high probability; more precisely, we show that the size of the connected components have exponential tails. Since $X_T^1 \geq X_T^0$, the same holds for X_T^0.

The intuition behind our proof of shattering after T steps goes as follows. Consider the balls $B_r(v)$ for $v \in V(\mathcal{G})$ where $r = O(1)$ is a sufficiently large constant. For each $v \in V(\mathcal{G})$, let (Z_t^v) denote the chain on $B_r(v)$ with a fixed *wired* boundary condition outside of $B_r(v)$. The evolution of (Z_t^v) and (X_t^1) can be coupled with the same identity coupling by ignoring the updates outside of $B_r(v)$ for (Z_t^v); by monotonicity, for every $v \in V(\mathcal{G})$ and $t \geq 0$, we have $Z_t^v \geq X_t^1(B_r(v))$. Consequently, we can even take the *minimum* (intersection) of the chains Z_t^v over $v \in V(\mathcal{G})$, to obtain a configuration ω_t by setting $\omega_t(e) = 1$ if $Z_t^v(e) = 1$ for *all* $v \in V(\mathcal{G})$. Since all these chains are coupled using the same randomness, we maintain the domination $\omega_t \geq X_t^1$ for all $t \geq 0$.

We thus focus on showing the shattering property for ω_T. Notice that we can bound the connected component of a vertex v in ω_T via an iterative exploration process. We initialize a set A as the connected component $\mathcal{C}_v$ of v in $Z_T^v(B_r(v))$ and initialize ∂A to $\mathcal{C}_v \cap \partial B_r(v)$. For each $u \in \partial A$ (while $\partial A \neq \emptyset$), we

(1) Add to A the connected component $\mathcal{C}_u$ of u in $Z_T^u(B_r(u))$
(2) Add to ∂A all vertices in $\mathcal{C}_u \cap \partial B_r(u)$. Remove u from ∂A.

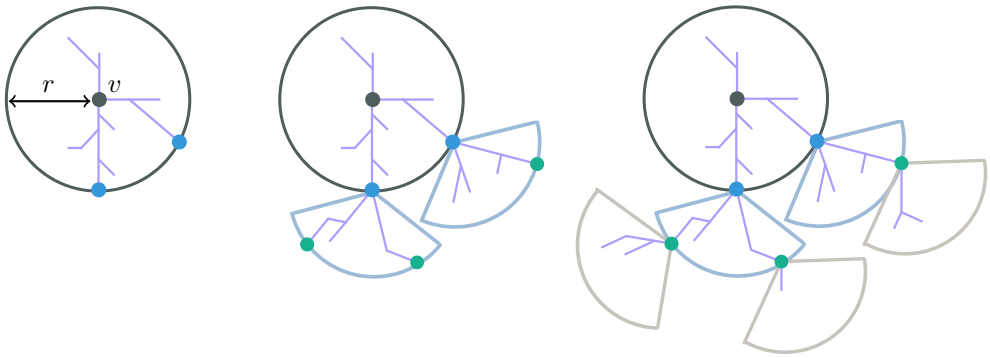

FIGURE 3.1. Left: Upon exposing the localized FK-dynamics Z_T^v on $B_r(v)$, the connected component of v (purple) reaches two boundary points of $B_r(v)$ (blue). Middle: The revealing procedure then exposes their localized configurations, in their balls of radius r. Right: The procedure continues in that manner until these connected components die out.

The procedure ends when ∂A is empty and outputs an edge set A necessarily containing the component of v in ω_T. See the depiction in Figure 3.1. The nature of this exploration process lends itself naturally to comparison with a branching process in which the "children" of u are the vertices connected to u through Z_T^u. (It turns out that the revealing of these configurations can be done in such a way that although they are all coupled, the dependencies between configurations $Z_T^u(B_r(u))$ are negligible: we comment on this later.) We will show that with high probability over $\mathcal{G}$, the resulting branching process is *sub-critical*.

To see this, first note that since the mixing time on $B_r(v)$ is $O(1)$ (r is constant), after $T = \Theta(n \log n)$ steps, enough updates have occurred in each ball $B_r(v)$ so that the chains (Z_t^v) have all mixed with high probability. Hence, up to a small error, we can consider instead the branching process where the number of children of v is given by the number of connections of v to the boundary of $B_r(v)$ in a sample from $\pi^1_{B_r(v)}$.

Now, most $O(1)$-sized balls in a random Δ-regular graph are trees. A key characteristic of the uniqueness regime $p < p_u(q, \Delta)$ is that in the wired Δ-regular tree $\mathcal{T}_r$, the expected number of leaves connected to v under $\pi^1_{\mathcal{T}_r}$ is less than 1 as long as r is large. As long as the role played by non-tree balls in $\mathcal{G}$ is bounded, this would imply the desired sub-criticality of the dominating branching process. We in fact need concentration bounds on the number of explored vertices in this branching process; towards this we show that $\hat{p}$ is the actual exponential decay rate of root-to-leaf connectivities on Δ-regular (wired) trees.

Lemma 3.1. *Let $\mathcal{T}_h$ denote the rooted Δ-regular complete tree of depth h and let $p < p_u(q, \Delta)$. Let $(1, \circlearrowleft)$ be the wired boundary condition on $\partial\mathcal{T}_h$ that additionally wires the root of $\mathcal{T}_h$ to $\partial\mathcal{T}_h$. There exists a constant $C = C(p, q, \Delta)$ such that for every h and every leaf $u \in \partial\mathcal{T}_h$,*

$$\pi^{(1,\circlearrowleft)}_{\mathcal{T}_h}(u \text{ is connected to the root of } \mathcal{T}_h) \le C\hat{p}^h .$$

Since there are $O(d^r)$ leaves in $\mathcal{T}_r$, the lemma implies that the expected number of connections from v to the boundary is $O((\hat{p}d)^r)$, which is less than one for r large (as $\hat{p}d < 1$ when $p < p_u(q, \Delta)$). The reason we establish this decay for the boundary condition $(1, \circlearrowleft)$, instead of simply the wired one, is to eliminate the potential dependencies between the chains (Z_t^u) through their roots.

To conclude our sketch of the ideas in Stage I, we mention two fundamental challenges to implementing the above approach. First, since all the chains (Z_t^v) are coupled via the identity coupling, revealing their configurations while maintaining some independence is delicate (see Lemma 4.17). We perform this revealing by additionally wiring the root to the boundary as hinted by the $(1, \circlearrowleft)$, and for each u, only revealing the new randomness needed to run the resulting chain on $B_R(u)$ up to time T. Roughly speaking, the wired boundary conditions allow us to evolve the un-revealed configuration in $B_r(v)$ in isolation.

Secondly, not every ball $B_r(v)$ in $\mathcal{G}$ will be a tree, and there are strong correlations between the short cycles of the underlying graph and the places where the random-cluster configuration is more wired. A key

contribution of our work is to construct a simultaneous revealing procedure for the random graph $\mathcal{G}$ with the overlayed random-cluster configuration of ω_T in a manner that handles these dependencies and can be approximated by the above sub-critical branching process; see Definition 4.11.

Putting all these ideas together, we establish the following exponential tail bound (shattering estimate) on cluster sizes of X_T^1 after a burn-in period of $O(n \log n)$ time.

Theorem 3.2. *Let $p < p_u(q, \Delta)$ and suppose that $\mathcal{G}$ is sampled from $\mathbf{P}_{\text{RRG}}$, the uniform distribution over Δ-regular graphs on n vertices. Then, for every $v \in V(\mathcal{G})$, $k \geq 1$ and $T \geq Cn \log n$, where $C > 0$ is a sufficiently large constant, with probability $1 - \exp(-\Omega(k)) - O(n^{-5})$, the random graph $\mathcal{G}$ is such that*

$$P(|\mathcal{C}_v(X_T^1)| \geq k) \leq \exp(-\Omega(k)) + O(n^{-5}) .$$

(Recall that $\mathcal{C}_v(X_T^1)$ denotes the component of vertex v in X_T^1.) By a union bound, Theorem 3.2 implies that all components of X_T^1 are of size at most $O(\log n)$ with high probability. Theorem 3.2 is proved in §4.

Using the above arguments, we can further show that for each $v \in V(\mathcal{G})$ the boundary condition induced on the ball $B_R(v)$ of radius $R = (\frac{1}{2} - \delta) \log_d n$ by the configuration of X_T^1 on the edges outside of $B_R(v)$ is typically K-*sparse*, i.e., the boundary condition induces only $K = O(1)$ many connections on $\partial B_R(v)$. Theorem 4.3 establishes that this property holds for all $v \in V(\mathcal{G})$ simultaneously with high probability.

Stage II. After the initial $T = O(n \log n)$ steps of the burn-in phase, the configurations X_T^1 and X_T^0 shatter and induce sparse boundary conditions (with up to $O(1)$ vertices wired through the boundary) on every ball $B_R(v)$ of radius $R = (\frac{1}{2} - \delta) \log_d n$ with high probability. It remains to show that the copies of the FK-dynamics will couple on $\mathcal{N}_v$ except with probability $1 - o(n^{-1})$ in an additional $O(n \log n)$ steps.

Starting at time T, we consider localized copies of the FK-dynamics in each ball $B_R(v)$ with $v \in V(\mathcal{G})$. This is done by ignoring (or *censoring*) the moves of the dynamics outside of $B_R(v)$ which has the effect of "freezing" the two distinct boundary conditions induced by $X_T^1(E(\mathcal{G}) \setminus B_R(v))$ and $X_T^0(E(\mathcal{G}) \setminus B_R(v))$ on the boundary of $B_R(v)$. With the sparse boundaries conditions frozen on $\partial B_R(v)$, the two coupled chains continue to run inside $B_R(v)$, and we can more easily analyze their configurations near v. The censoring technology of [46] implies that if these censored chains are coupled on $\mathcal{N}_v$, then so are the original chains.

In Lemma 6.5, we show that if $X_T^1(E(\mathcal{G}) \setminus B_R(v))$ and $X_T^0(E(\mathcal{G}) \setminus B_R(v))$ induce sparse boundary conditions on $B_R(v)$, and $\mathcal{G}$ is (L, R)-Treelike with $L = O(1)$ (see Definition 2.2), the mixing time of the FK-dynamics on $B_R(v)$ is $O(d^R \log(d^R))$. In fact, we can establish a tight bound on the log-Sobolev constant of the FK-dynamics on $B_r(v)$ under sparse boundary conditions, showing that $t_{\text{MIX}}(\varepsilon) = O(d^R \log(d^R/\varepsilon))$. This slightly stronger fact turns out to be crucial for deducing the tight $O(n \log n)$ bound on the mixing time of the FK-dynamics on $\mathcal{G}$, i.e., without an additional polylog(n) factor.

With this optimal bound on the local mixing on treelike balls, we know that the localized chains have all mixed after $O(n \log n)$ steps of the FK-dynamics. Therefore, the probability that two instances of the FK-dynamics on $B_R(v)$ with distinct sparse boundary conditions ξ and ξ' are not coupled on $\mathcal{N}_v$ is given by the total variation distance between $\pi^{\xi}_{B_R(v)}$ and $\pi^{\xi'}_{B_R(v)}$ on $\mathcal{N}_v$. We show that this distance is $O(\hat{p}^{2R})$.

Proposition 3.3. *Consider a vertex v in a Δ-regular graph $\mathcal{G}$. If $B_R(v)$ is L-*Treelike *and ξ, ξ' are any two K-*sparse *boundary conditions on $\partial B_R(v)$, there exists a constant $C = C(p, q, \Delta, L, K)$ such that*

$$\|\pi^{\xi}_{B_R(v)}(\omega(\mathcal{N}_v) \in \cdot) - \pi^{\xi'}_{B_R(v)}(\omega(\mathcal{N}_v) \in \cdot)\|_{\text{TV}} \leq C\hat{p}^{2R} .$$

(Recall that we say a boundary condition is K-sparse when there are only K boundary wirings.) We stress the importance of obtaining the sharp $\hat{p}^2$ decay rate here for the spatial mixing to support a union bound over n vertices. Since $\hat{p} < 1/d$ and $R = (\frac{1}{2} - \delta) \log_d n$, we have $\hat{p}^{2R} = o(n^{-1})$, but any weaker bound on the decay rate would force us to choose a larger R, which would cross the threshold at which point balls of $\mathcal{G}$ are no longer (L, R)-Treelike for $L = O(1)$, and we would lose control over the mixing time on $B_R(v)$.

The proof of this spatial mixing property is based on the fact that in order for information to travel from the boundary of $B_R(v)$ to $\mathcal{N}_v$ there must be *two* disjoint open paths from $\mathcal{N}_v$ to non-singleton elements of ξ in $\partial B_R(v)$. We contrast this to the more traditional bound on influence by the existence of a *single*

connection from the center of a ball to its boundary, which in our setting would only yield a bound of $\hat{p}^R$. (Such a bound by a single connectivity event is the one traditionally used on amenable graphs like $\mathbb{Z}^2$ to go from spatial mixing with *any* positive rate of exponential decay to fast mixing: see [1, 8, 42].)

4. The FK-dynamics shatters quickly on random graphs

Our first goal in this section is to prove Theorem 3.2 establishing existence of $T_{\text{BURN}} = O(n \log n)$ such that for $t \geq T_{\text{BURN}}$, the configuration $X^1_{\mathcal{G},t}$ is shattered. We will then use this to conclude that the boundary conditions $X^1_{\mathcal{G},t}$ induces on any ball of volume $o(\sqrt{n})$ are $O(1)$-sparse. Let us now be more precise.

Definition 4.1. A random-cluster boundary condition ξ on an edge-subset $H \subset E(\mathcal{G})$ is said to be K-Sparse if the number of vertices in non-trivial (non-singleton) boundary components of ξ is at most K.

Definition 4.2. A random-cluster configuration ω on $\mathcal{G} = (V(\mathcal{G}), E(\mathcal{G}))$ is (K, R)-Sparse if, for every $v \in V(\mathcal{G})$, the boundary conditions induced on $B_R(v)$ by $\omega(E(\mathcal{G}) \setminus E(B_r(v)))$ are K-Sparse.

The following key result asserts that the boundary of every ball about a vertex is $O(1)$-Sparse with high probability after an $O(n \log n)$ burn-in time: this is proven in Section 4.5.

Theorem 4.3. *Fix $p < p_u(q, \Delta)$. There exists $C(p, q, \Delta)$ such that for every $t \geq Cn \log n$, the following holds. For every $\delta > 0$, if $R := (\frac{1}{2} - \delta) \log_d n$, there exists $K(p, q, \Delta, \delta)$ such that with $\mathbf{P}_{\text{RRG}}$-probability $1 - o(1)$, $\mathcal{G}$ is such that*

$$P\big(X^1_{\mathcal{G},t} \text{ is } (K, R)\text{-Sparse}\big) \geq 1 - O(n^{-2}) . \tag{4.1}$$

Remark 4.4. By monotonicity of the FK-dynamics, for every $\mathcal{G}$, we have that $X^1_{\mathcal{G},t} \succeq \pi_{\mathcal{G}}$, from which it follows that both Theorem 4.3, and the exponential tails of Theorem 3.2, hold under $\pi_{\mathcal{G}}$, i.e., if one replaces $X^1_{\mathcal{G},T}$ by an equilibrium configuration $\omega \sim \pi_{\mathcal{G}}$.

In Section 4.1, we construct the relevant revealing procedures for FK-dynamics clusters on random graphs, and define the branching process we dominate it by. In Sections 4.2–4.4, we analyze these processes, and in Section 4.5, we complete the proofs of Theorems 3.2 and 4.3.

4.1. Couplings and revealing schemes for the FK-dynamics on random graphs.

In this section, we summarize the key couplings and revealing schemes for the connected components of $X^1_{\mathcal{G},t}$. These are fundamental to the proof of shattering for $X^1_{\mathcal{G},t}$ in the uniqueness region after an $O(n \log n)$ burn-in time.

4.1.1. *The configuration model.* The configuration model $\mathbf{P}_{\text{CM}}$ is a distribution over multigraphs on n vertices and fixed degree distribution, which we take to be Δ for every vertex, defined as follows [9]. Give every vertex $v \in \{1, ..., n\}$ Δ-half-edges and select a matching on the Δn many half-edges uniformly at random to form the $\Delta n/2$ edges of the graph. Let $\mathfrak{M}_n$ be the set of possible edges (the set of pairs of half-edges).

The configuration model is a useful tool for studying the random Δ-regular graph, as the distribution $\mathbf{P}_{\text{RRG}}$ is equal to the distribution $\mathbf{P}_{\text{CM}}(\cdot \mid \mathcal{G} \in \Gamma_{\text{RRG}})$ where Γ_{RRG} is the event that the graph $\mathcal{G}$ is simple (i.e., has no self-loops or multi-edges). In particular, it is standard (see e.g., [9]) that $\mathbf{P}_{\text{CM}}(\Gamma_{\text{RRG}}) > c$ for some $c(\Delta) > 0$, and therefore for any event Γ,

$$\mathbf{P}_{\text{RRG}}(\Gamma) = \mathbf{P}_{\text{CM}}(\Gamma, \Gamma_{\text{RRG}})\big(\mathbf{P}_{\text{CM}}(\Gamma_{\text{RRG}})\big)^{-1} \leq c^{-1}\mathbf{P}_{\text{CM}}(\Gamma) . \tag{4.2}$$

Refer to the book [22] for more on the configuration model. We will use (4.2), with an iterative revealing scheme of a matching of the Δn half-edges, to analyze the random Δ-regular graph.

The configuration model lends itself to revealing procedures. Towards introducing the joint revealing procedure for the random graph $\mathcal{G} \sim \mathbf{P}_{\text{CM}}$ and the configuration $X^1_{\mathcal{G},t}$, let us first recall a standard revealing procedure for random Δ-regular graphs according to $\mathbf{P}_{\text{CM}}$ on its own. This procedure is useful to proving random graph estimates for the configuration model and Δ-regular random graph. It also serves as a building block for the revealing procedure of the random graph together with the FK-dynamics configuration.

The following iterative algorithm is a way to sample from the configuration model for a given degree sequence. The fact that this gives a valid sample from $\mathbf{P}_{\text{CM}}$ is straightforward after naturally identifying samples from $\mathbf{P}_{\text{CM}}$ with samples from the uniform distribution over matchings on Δn.

Definition 4.5. Assign every $v \in \{1, ..., n\}$, Δ half-edges. Suppose f is a (possibly random) function from edge-sets $\mathcal{A} \subset \mathfrak{M}_n$ (a set of matched pairs of half-edges) to a half-edge not matched in $\mathcal{A}$.

(1) Initialize the set of exposed edges as $\mathcal{A}_0 = \emptyset$.
(2) For every $1 \le m \le \frac{\Delta n}{2}$, match $f(\mathcal{A}_{m-1})$ to a half-edge selected uniformly at random from the remaining un-matched ones to form the edge e_m. Let $\mathcal{A}_m := \mathcal{A}_{m-1} \cup \{e_m\}$.

Observe, importantly, that the choice of next half-edge to match (given by the function f) can be *adaptive*, specifically, adapted to the filtration generated by $(\mathcal{A}_0, ..., \mathcal{A}_{m-1})$.

Definition 4.5 provides an adaptive sampling method from the configuration model distribution $\mathbf{P}_{\text{CM}}$ (see e.g., [43]) and can be used to prove myriad properties of random Δ-regular graphs. In particular, it yields a simple proof of Fact 2.3 that $\mathcal{G} \sim \mathbf{P}_{\text{RRG}}$ is (L, R)-Treelike for $R \le n^{\frac{1}{2}-\delta}$ and $L = O(1)$; see Section 4.2.

Remark 4.6. The definitions of the random-cluster model (1.1), and the FK-dynamics extend naturally to multigraphs, where $G = (V, E)$ is such that V is identified with $\{1, ..., n\}$ and $E \subset \mathfrak{M}_n$ is a multiset. The random-cluster model and FK-dynamics then live over subsets of E, identified with $\omega : E \to \{0, 1\}$, and connectivity components of a configuration ω are understood naturally.

4.1.2. *A coupling of localized FK-dynamics chains.* Our goal is to simultaneously expose edges of $\mathcal{G} \sim \mathbf{P}_{\text{CM}}$ while revealing the FK-dynamics configuration $X^1_{\mathcal{G},t}$ at time t on $\mathcal{G}$. We show that under their joint distribution the size of the connected components of $X^1_{\mathcal{G},t}$ have exponential tails; this in turn implies that the boundary condition on $B_r(v)$ is $O(1)$-Sparse (see Definition 4.1).

Note that a ball of radius $O(\log n)$ about a vertex v may have many cycles—indeed it may encompass the entire graph $\mathcal{G}$—but a typical FK cluster of *size* $O(\log n)$ does not use most of these cycles. Thus, we expose the edges of $\mathbf{P}_{\text{CM}}$ guided by the revealing of the random-cluster component of a vertex v in $X^1_{\mathcal{G},t}$; in this way, to expose the $\mathcal{C}_v(X^1_{\mathcal{G},t})$ we will not have to reveal much of the random graph.

There are two key difficulties to consider when constructing a joint revealing process for $(\mathcal{G}, X^1_{\mathcal{G},t})$:

(1) Under either of $X^1_{\mathcal{G},t}$ or e.g., the random-cluster measure $\pi_{\mathcal{G}}$, the value $\omega(e)$ on an edge e shown to belong to $E(\mathcal{G})$, affects the distribution of the remainder of the underlying random graph.
(2) Unlike the random-cluster measure $\pi_{\mathcal{G}}$, the law of $X^1_{\mathcal{G},t}$ does not satisfy any domain Markov property. Indeed, the distribution of $X^1_{\mathcal{G},t}(e)$ conditionally on some $X^1_{\mathcal{G},t}(A)$ is quite difficult to analyze.

The key to overcoming these obstructions will be to reveal the configurations of a family of FK-dynamics chains that are localized (in the sense that their distribution only depends on a small $O(1)$ sized subset of edges of the graph) and whose concatenation stochastically dominates the distribution of $X^1_{\mathcal{G},t}$. Let us be more precise next and explicitly construct a coupling of a family of localized FK-dynamics chains.

Definition 4.7. For a graph $\mathcal{G}$ and edge subset $A \subset E(\mathcal{G})$, let ∂A be the set of vertices in $V(A)$ that are adjacent to vertices of $V(\mathcal{G}) \setminus V(A)$, and let π^1_A be the random-cluster measure on A with wired boundary conditions on ∂A. Let $(X^1_{A,t})_{t\ge 0}$ be the FK-dynamics chain that starts from all wired on $E(\mathcal{G})$, censors (ignores) all updates in $E(\mathcal{G}) \setminus A$, and makes FK-dynamics updates w.r.t. π^1_A when it updates edges in A.

Importantly, the wiring on ∂A ensures that the law of $X^1_{A,t}$ does not depend on $E(\mathcal{G}) \setminus A$.

Definition 4.8. For any graph $\mathcal{G}$, we can construct $(\mathcal{X}^1_t)_{t\ge 0} = \left\{(X^1_{A,t})_{t\ge 0}\right\}_{A\subset E(\mathcal{G})}$, a grand monotone coupling of the ensemble of FK-dynamics chains $(X^1_{A,t})_{t\ge 0}$ for $A \subset E(\mathcal{G})$ as follows:

(1) Initialize $X^1_{A,0} \equiv 1$ for all $A \subset E(\mathcal{G})$; i.e., the all wired configuration on A.
(2) Let $(e_t)_{t\ge 1} = (e_1, e_2, ...)$ be drawn i.i.d. from $E(\mathcal{G})$.
(3) Let $(U_{e,t})_{e\in E(\mathcal{G}),t\ge 1}$ be a sequence of i.i.d. uniform random variables on $[0, 1]$.

For $A \subset E(\mathcal{G})$, construct $(X^1_{A,t})_{t\geq 1}$ as follows: for each $t \geq 1$, set $X^1_{A,t}(e) = X^1_{A,t-1}(e)$ for $e \neq e_t$ and

$$X^1_{A,t}(e_t) = \begin{cases} 1 & \text{if } e_t \notin A; \\ 1 & \text{if } e_t \in A \text{ and } U_{e_t,t} \leq \varrho; \\ 0 & \text{if } e_t \in A \text{ and } U_{e_t,t} > \varrho; \end{cases}$$

for $\varrho = \pi^1_A\big(\omega(e_t) = 1 \mid \omega(A \setminus \{e_t\}) = X^1_{A,t-1}(A \setminus \{e_t\})\big)$; i.e., if $e_t \in A$, we resample e_t given the remainder of the configuration on A, together with the wired boundary condition on ∂A, using the same uniform random variable $U_{e_t,t}$ for every $X^1_{A,t}$ such that $e_t \in A$.

As in the grand coupling for different initializations, this is a monotone coupling. In particular, we have $X^1_{\mathcal{G},t} \leq X^1_{A,t}$ for all $A \subset E(\mathcal{G})$ and thus $X^1_{\mathcal{G},t} \leq \bigcap_{A\subset E(\mathcal{G})} X^1_{A,t}$.

A key observation for our revealing process is that for every A, the configuration $X^1_{A,t}$ depends only on:

(1) the number of updates amongst $(e_s)_{s\leq t}$ that belong to A, which we denote by $\kappa_{A,t}$;
(2) the choice of edges to be updated on A on those $\kappa_{A,t}$ updates; we denote such set by $\mathcal{O}_{A,\kappa_{A,t}}$; and
(3) the family of uniform random variables on those edges, $(U_{e,s})_{e\in A, s\leq t}$.

With this observation in hand, we can extend this to a coupling of $(X^1_{A,t})$ averaged over $\mathcal{G} \sim \mathbf{P}_{\text{CM}}$.

Definition 4.9. Let $\mathbb{P}^1_t$ be the distribution over pairs $(\mathcal{G}, \omega_t)$ where ω_t is a random-cluster configuration on $\mathcal{G}$ that results by first drawing $\mathcal{G} \sim \mathbf{P}_{\text{CM}}$, then drawing $\omega_t \sim P(X^1_{\mathcal{G},t} \in \cdot)$. Likewise, for every set $A \subset \mathfrak{M}_n$, let $\mathbb{P}^1_{A,t}$ be the distribution over pairs $(\mathcal{G}, \omega_{A\cap E(\mathcal{G}),t})$ where $\omega_{A,t} \sim P(X^1_{A\cap E(\mathcal{G}),t} \in \cdot)$. Couple, under the distribution $\mathbb{P}$, the family of distributions $(\mathbb{P}^1_{A,t})_{A\subset\mathfrak{M}_n, t\geq 1}$ by selecting the same random graph $\mathcal{G} \sim \mathbf{P}_{\text{CM}}$ for all of them, then using the coupling of Definition 4.8 of the family $(X^1_{A,t})_{A,t}$.

In this manner, we have constructed a monotone coupling of the family $(\mathcal{G}, (X^1_{A,t})_{t\geq 1})_{A\subset\mathfrak{M}_n}$. Note that we use this coupling for sets A which we know have $E(\mathcal{G}) \cap A = A$, so that the averaging is only over the edges of $E(\mathcal{G}) \setminus A$, which we earlier noted $X^1_{A,t}$ is independent of; thus the role of this coupling is only to put the random graphs with their random-cluster configurations on the same probability space. We defer detailed discussion of the properties of the coupling to Section 4.2 (after constructing the revealing procedure in the sequel) but emphasize that by construction, if $A \cap B = \emptyset$, the only dependency of $X^1_{A,t}$ and $X^1_{B,t}$ is through the distributions of the binomial random variables $\kappa_{A,t}$ and $\kappa_{B,t}$.

4.1.3. *The joint revealing procedure.* We now construct a revealing procedure for $\mathcal{G}$ and a configuration $\tilde{\omega}_t$ on $\mathcal{G}$ that stochastically dominates $X^1_{\mathcal{G},t}$. Fix r to be chosen as a large constant (depending on p, q, Δ) later.

Definition 4.10. Given an exposed set of edges $\mathcal{A}$ of the random graph $\mathcal{G} \sim \mathbf{P}_{\text{CM}}$, we define $B^{out}_r(v) = B^{out}_r(v; \mathcal{A})$ as the ball of radius r in $E(\mathcal{G}) \setminus \mathcal{N}_v(\mathcal{A})$ where $\mathcal{N}_v(\mathcal{A})$ is the set of edges in $\mathcal{A}$ incident to v. We drop the $\mathcal{A}$ from the notation when understood contextually.

For an edge-set $\mathcal{A}_0 \subset \mathfrak{M}_n$ revealed to be part of $E(\mathcal{G})$ and a vertex set $\mathcal{V}_0 \subset V(\mathcal{G})$, we construct a joint iterative procedure to expose (a set containing) the connected components $\mathcal{C}_{\mathcal{V}_0}(X^1_{\mathcal{G},t}(E(\mathcal{G}) \setminus \mathcal{A}_0))$. The examples to have in mind are (1) $\mathcal{A}_0 = \emptyset$ and $\mathcal{V}_0 = \{v\}$ and (2) $\mathcal{A}_0 = E(B_R(v))$ and $\mathcal{V}_0 = \partial B_R(v)$.

Through this process we will keep track of the following variables at each step:

- $\mathcal{A}_m$: the set of edges of the random graph that have been revealed by step m;
- $\mathcal{V}_k$: the set of vertices in the k-th generation we want to explore out of;
- $\tilde{\omega}_m$: the random-cluster configuration revealed up to step m;
- $\mathcal{F}_m$: elements of the filtration with respect to which the configuration $\tilde{\omega}_m$ on $\mathcal{A}_m$ is measurable.

The process is defined as follows (see Figures 4.1 and 4.2 for a depiction of several steps of this process):

Definition 4.11. Initialize: $k = 0$; $m = 1$; $\mathcal{F}_0 = \emptyset$; $\mathcal{A}_0 \subset E(\mathcal{G})$; $\mathcal{V}_0 \subset V(\mathcal{G})$;

for each $k \geq 0$ **while** $\mathcal{V}_k \neq \emptyset$

for each $v \in \mathcal{V}_k$

1. Set $v_m = v$. Conditionally on $\mathcal{A}_{m-1}$, reveal the edges of the random graph in $B_r^{out}(v_m)$ and set
$$\mathcal{A}_m := \mathcal{A}_{m-1} \cup E(B_r^{out}(v_m));$$
Let $A_m := \mathcal{A}_m \setminus \mathcal{A}_{m-1}$ be the set of new edges revealed in $\mathcal{G}$.
2. Reveal the triplet $\mathcal{H}_{A_m,t} := \{\kappa_{A_m,t}, \mathcal{O}_{A_m,\kappa_{A_m,t}}, (U_{e,s})_{e\in A_m, s\leq t}\}$ conditionally on $\mathcal{F}_{m-1}$. Recall that $\kappa_{A_m,t}$ is the number of updates of the FK-dynamics in A_m, $\mathcal{O}_{A_m,\kappa_{A_m,t}}$ is the sequence of edges to be updated in A_m, and $(U_{e,s})_{e\in A_m,s\leq t}$ is the family of uniform random variables used for the edges updates in A_m.
3. Construct the random-cluster configuration $X^1_{A_m,t}(A_m)$ from the triplet $\mathcal{H}_{A_m,t}$ by simulating the steps of the FK-dynamics on A_m with wired boundary conditions. Concatenate $X^1_{A_m,t}(A_m)$ with $\tilde{\omega}_{m-1}$ to obtain a new configuration $\tilde{\omega}_m$ on $\mathcal{A}_m \setminus \mathcal{A}_0$.
4. Add to $\mathcal{V}_{k+1}$ all vertices of $\partial(\mathcal{A}_m \setminus \mathcal{A}_0)$ that are in the component of $\mathcal{V}_0$ in $\tilde{\omega}_m(\mathcal{A}_m \setminus \mathcal{A}_0)$ and are not in $\bigcup_{\ell\leq k} \mathcal{V}_\ell$.
5. Let $\mathcal{F}_m = (\mathcal{F}_{m-1}, \mathcal{H}_{A_m,t})$ and increase m by 1.

Let $\kappa_\emptyset$ be the first k such that $\mathcal{V}_k = \emptyset$ and let $\mathfrak{m}_{k_\emptyset} = \sum_{k=0}^{k_\emptyset} |\mathcal{V}_k|$. Let $\tilde{\omega}_t = \tilde{\omega}_t(\mathcal{A}_{\mathfrak{m}_{k_\emptyset}} \setminus \mathcal{A}_0)$ be the random-cluster configuration revealed when the process terminates. The key observation about the above process is that we can control the cluster of $\mathcal{V}_0$ in $X^1_t(E(\mathcal{G}) \setminus \mathcal{A}_0)$ by the set $\mathcal{A}_{\mathfrak{m}_{k_\emptyset}}$; the size of this set will then be approximately controlled by comparison to a sub-critical branching process in the following subsection.

Observation 4.12. *The connected components of $\mathcal{V}_0$ in $X^1_t(E(\mathcal{G}) \setminus \mathcal{A}_0)$ are a subset of the connected components of $\mathcal{V}_0$ in $\tilde{\omega}_t(E(\mathcal{G}) \setminus \mathcal{A}_0)$. In particular, the number of vertices in non-trivial (i.e., non-singleton) components of the boundary condition $X^1_t(E(\mathcal{G}) \setminus \mathcal{A}_0)$ induces on $\mathcal{A}_0$ is less than the number of vertices in non-trivial components of the boundary condition $\tilde{\omega}_t(E(\mathcal{G}) \setminus \mathcal{A}_0)$ induces on $\mathcal{A}_0$. The edges in both of these sets of connected components are subsets of the edge-set $\mathcal{A}_{\mathfrak{m}_{k_\emptyset}} \setminus \mathcal{A}_0$.*

With Observation 4.12 in hand, we focus on obtaining the exponential tail bound of Theorem 3.2 for $\mathcal{C}_v(\tilde{\omega}_t)$ (the component of v in $\tilde{\omega}_t$) and likewise, the sparsity bound of Proposition 4.21 for $\mathfrak{V}_{B_R(v)}(\tilde{\omega}_t)$.

4.1.4. *Constructing a dominating branching process.* Towards proving Proposition 4.21, we construct a (non-Markovian, size-dependent) branching process which we will show stochastically dominates the sequence $(\mathcal{V}_k)_{k\geq 0}$ of our joint reveleaing process. This process $(Z_k)_{k\geq 0}$ will then be shown to be sub-critical and satisfy exponential tail bounds on its total population, implying the same for the cluster of $\mathcal{V}_0$ in $\tilde{\omega}_t$.

Definition 4.13. Initialize $Z_0 = |\mathcal{V}_0|$, and let $(Z_k)_{k\geq 1}$ be the (size-dependent) branching process, which for each k, has progeny $(\chi_{i,k})_{i\leq Z_k}$ drawn i.i.d. from the following distribution:

(1) With probability $n^{-1/2}$, let $\chi_{i,k} = |V(\mathcal{T}_r)| \sum_{\ell\leq k} Z_\ell$ and say the progeny number $\chi_{i,k}$ is Bad.
(2) Otherwise, sample $\chi_{i,k}$ from the distribution of the number of leaves in the connected component of the root under $\pi^{(1,\circlearrowleft)}_{\mathcal{T}_r}$ (the random-cluster measure on the Δ-regular tree of depth r with a wired boundary condition and with the root also wired to $\partial\mathcal{T}_r$).

Let $Z_{k+1} = \sum_{i\leq Z_k} \chi_{i,k}$; that is, the i-th member of the k'th generation gets $\chi_{i,k}$ many children.

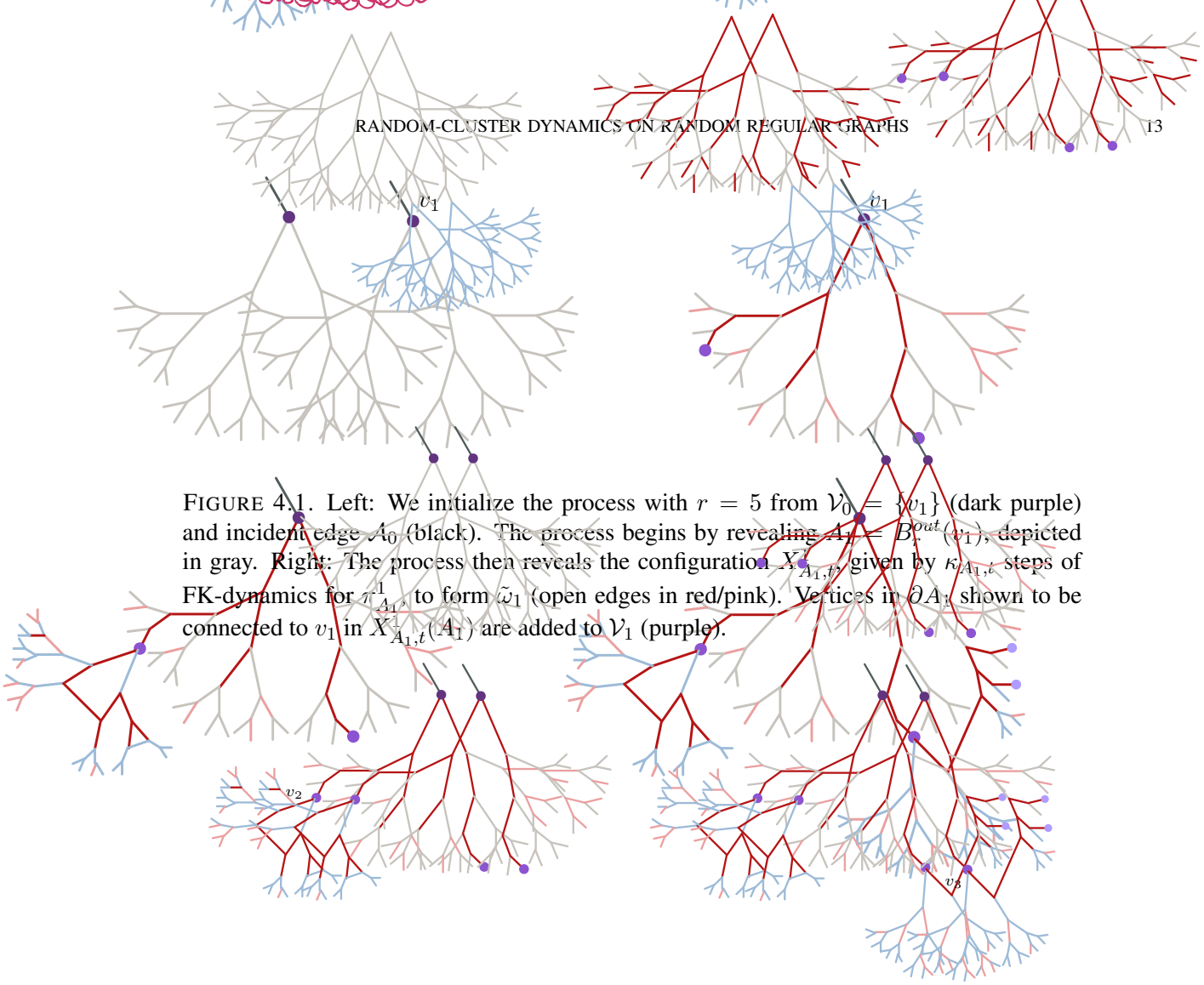

FIGURE 4.1. Left: We initialize the process with $r = 5$ from $\mathcal{V}_0 = \{v_1\}$ (dark purple) and incident edge $\mathcal{A}_0$ (black). The process begins by revealing $A_1 = B_r^{out}(v_1)$, depicted in gray. Right: The process then reveals the configuration $X^1_{A_1,t}$ given by $\kappa_{A_1,t}$ steps of FK-dynamics for $\tilde{\omega}^1_{A_1}$ to form $\tilde{\omega}_1$ (open edges in red/pink). Vertices in ∂A_1 shown to be connected to v_1 in $X^1_{A_1,t}(A_1)$ are added to $\mathcal{V}_1$ (purple).

FIGURE 4.2. Left: Proceeding from above, in the next generation, starting from $v_2 \in \mathcal{V}_1$, reveal the edges of $B_t^{out}(v_2)$ in $\mathcal{G}$; in this case, this is not a tree, but is disjoint from $\mathcal{A}_1$, so that $A_2 = B_r^{out}(v_2)$. The configuration $X^1_{A_2,t}$ is generated and concatenated with $\tilde{\omega}_1$ to form $\tilde{\omega}_2$. Right: For $v_3 \in \mathcal{V}_1$, $B_r^{out}(v_3)$ is a tree, but it intersects $\mathcal{A}_2$. As such, $A_3 = B_r^{out}(v_3)\backslash\mathcal{A}_2$. Running the FK-dynamics on A_3 with all-wired boundary conditions ensures that $X^1_{A_3,t}$ is nonetheless independent of the configuration we had revealed in $\tilde{\omega}_2$. The light purple vertices are connected to $\mathcal{V}_0$ in $\tilde{\omega}_3$ and are added to $\mathcal{V}_2$ to form the next generation.

Note that this is not a branching process in the traditional sense, since the progeny distribution is not i.i.d. and depends on the population up to that generation. Nonetheless, we will show good tail bounds on $(Z_k)_{k\geq 0}$ by dominating it by sub-critical branching processes between the Bad steps.

To justify the above construction, let us formalize the relation between $(Z_k)_k$ and the revealed vertices of the process in Definition 4.11, $(\mathcal{V}_k)$. Intuitively, we want to identify vertices $v_m \in \mathcal{V}_k$ with those of generation k in (Z_k); the *progeny* of v_m will then be those vertices added to $\mathcal{V}_{k+1}$ in step (4) of Definition 4.11. Item (1) from the progeny distribution of Definition 4.13 corresponds to situations where:

(1) $B_r^{out}(v_m)$ intersects $\mathcal{A}_{m-1}$;
(2) $B_r^{out}(v_m)$ is not a tree; or
(3) There are an insufficient number of updates on $B_r^{out}(v_m)$ for $X^1_{A_m,t}$ to mix.

Examples of situations (1)–(2) were depicted in Figure 4.2. The $n^{-1/2}$ probability assigned to these bad situations by the dominating branching process comes from the fact that Theorem 4.3 requires us to consider $|\mathcal{A}_0|$ of size $n^{-1/2+\delta}$, and thus any edge has at least probability $n^{-\frac{1}{2}-\delta}$ of intersecting $\mathcal{A}_0$. On the complement of situations (1)–(3) above, $X^1_{A_m,t}$ is mixed, and is comparable to the $(1,\circlearrowleft)$-tree of depth r.

4.1.5. *Comparing the revealing procedure and branching process.* We conclude the section by stating the main two lemmas comparing the revealing procedure to the branching process defined above. Towards stating these, denote by $t_{\text{MIX}}(\mathcal{T}_r,(1,\circlearrowleft)$ the mixing time of FK-dynamics on $\mathcal{T}_r$ with $(1,\circlearrowleft)$ boundary conditions, and define the burn-in time

$$T_{\text{BURN}} = T_{\text{BURN}}(C_0,r) := C_0 n\log n\cdot\frac{t_{\text{MIX}}(\mathcal{T}_r,(1,\circlearrowleft))}{|E(\mathcal{T}_r)|}\,. \tag{4.3}$$

Recall, the definition of the update numbers $(\kappa_{A_m,t})_m$ and define, for every $t\ge T_{\text{BURN}}$, the event

$$\mathfrak{E}_t = \mathfrak{E}_t^\infty \qquad \text{where} \qquad \mathfrak{E}_t^m := \Big\{(e_s)_{s\le t} : \bigcap_{l=1}^m \Big\{\kappa_{A_l,t}\le \frac{4|E(\mathcal{T}_r)|}{\Delta n}t\Big\}\Big\}\,. \tag{4.4}$$

Standard tail estimates for binomial random variables will imply that $\mathfrak{E}_t$ holds with high probability.

Let $\mathfrak{m}_0=0$ and for each $k\ge 0$, let $\mathfrak{m}_{k+1}=\mathfrak{m}_k+|\mathcal{V}_k|$, i.e., the total number of exposed vertices on the boundaries of explored balls, and in the same connected component as $\mathcal{V}_0$ before the exploration for the $(k+1)$-th generation begins. This will be the quantity which we compare to the population of the branching process of Definition 4.13. More precisely, on the event $\mathfrak{E}_t$, by construction of (Z_k), and the choice of T_{BURN}, we are able to show the following stochastic domination.

Lemma 4.14. *There exists $C_0(p,q,\Delta)$ in the definition of (4.3) such that the following holds for every $t\ge T_{\text{BURN}}$. For every $\mathcal{A}_0,\mathcal{V}_0$ such that $|\mathcal{A}_0|,|\mathcal{V}_0|\le n^{\frac12-\delta}$ for $\delta>0$, every $K>0$ fixed, and every $\ell\ge 1$,*

$$\big(|\mathcal{V}_j|\mathbf{1}\{\mathfrak{E}_t^{\mathfrak{m}_j}\}\mathbf{1}\{\mathfrak{m}_{j-1}\le n^{1/2-\delta/2}\}\big)_{j\le\ell}\preccurlyeq (Z_j)_{j\le\ell}\,.$$

In this manner, we will have reduced the analysis of the set of exposed vertices through the revealing process of $(\mathcal{G},\tilde\omega_t)$, and thus, the clusters of $X^1_{\mathcal{G},t}$, to the analysis of the process (Z_k), which except on some rare Bad increments, is a simple branching process with subcritical progeny distribution dictated by connectivity probabilities in the wired measure $\pi^{(1,\circlearrowleft)}_{\mathcal{T}_r}$. We will establish the following tail estimate for (Z_k).

Lemma 4.15. *Suppose $p<p_u(q,\Delta)$ and fix any $\delta>0$, any $M\ge 1$, and any $1\le Z_0\le n^{\frac12-\delta}$. There exist $r_0(p,q,\Delta),C(p,q,\Delta,M),K_0(p,q,\Delta,M)$ such that for every $r\ge r_0$ fixed and every $0<\lambda\le n^{\frac12-\delta}$,*

$$\mathbb{P}\Big(\sum_{k\ge 0}Z_k\ge K_0Z_0+\lambda\Big)\le C\exp(-\lambda/C)+Ce^Mn^{-\delta M}\,.$$

4.1.6. *Outline of remainder of section.* Having sketched the key revealing procedures and the way they fit together to provide the desired bounds on the clusters of $X^1_{\mathcal{G},t}$, let us prove the various relations and bounds claimed above. In Section 4.2, we prove various key properties of the configuration model revealing process of Definition 4.5 and the coupling of Definition 4.8 that will be central to the analysis of the revealing procedure of Definition 4.11. Then in Section 4.3, we show that the size-dependent branching process (Z_k) of Definition 4.13 stochastically dominates the FK process $(\mathcal{V}_k)$ of Definition 4.11 on a high-probability event, proving Lemma 4.14. In Section 4.4, we analyze the process (Z_k) by comparing its population to the sum of $O(1)$ many sub-critical branching processes to deduce Lemma 4.15. In Section 4.5, we combine these ingredients to conclude Theorem 3.2 and Proposition 4.21, and from that Theorem 4.3.

4.2. **Key properties of the revealing procedure for $(\mathcal{G},\tilde\omega_t)$.** In this section, we describe some of the key properties of the coupling constructed in Definition 4.8, and the revealing procedure constructed for the clusters of $\mathcal{V}_0$ in $\tilde\omega_t$ in Definition 4.11. The following preliminary lemmas describe the law of the random graph edges and overlaying FK configurations through the revealing process.

4.2.1. *Properties of the configuration model revealing procedure.* We begin with the following lemma on the law of the random graph $\mathcal{G}$ conditionally on a set $\mathcal{A}_m$ which we have revealed to be a subset of $E(\mathcal{G})$. Recall the configuration model's revealing procedure from Definition 4.5 and say a vertex is *discovered* if at least one of its half-edges has been matched, and *exhausted* if all of its half-edges have been matched.

Lemma 4.16. *Let $\mathcal{A}$ be any set of edges (pairing of half-edges) revealed to belong to $E(\mathcal{G})$. For every r,*

$$\sup_{v\in\{1,\dots,n\}} \mathbf{P}_{\mathrm{CM}}\big(B_r^{out}(v)\cap\mathcal{A}\neq\emptyset \text{ or } B_r^{out}(v) \text{ is not a tree} \mid \mathcal{A}\big) \le \frac{2\Delta d^r(|V(\mathcal{A})|+d^r)}{n-(|V(\mathcal{A})|+d^r)}\,.$$

Proof. Fix any edge-set $\mathcal{A}$. We can sample from the conditional distribution $\mathbf{P}_{\mathrm{CM}}(\cdot \mid \mathcal{A})$ by defining the adaptive scheme f in Definition 4.5 so that it first matches the half-edges belonging to $\mathcal{A}$, yielding the set $\mathcal{A}_{|\mathcal{A}|/2}=\mathcal{A}$ after $|\mathcal{A}|/2$ steps, then setting f to do a breadth-first search (BFS) of $B_r^{out}(v)$: this latter part is done by choosing f so that it first exhausts v, then exhausts each of the neighbors of v, and so on.

Revealing the entire set $B_r^{out}(v)$ takes at most $|E(\mathcal{T}_r)|$ many steps beyond $|\mathcal{A}|/2$. If for every $m\in\{|\mathcal{A}|/2+1,\dots,|\mathcal{A}|/2+d^r\}$ the half-edge from $f(\mathcal{A}_{m-1})$ is not matched to a half-edge belonging to a vertex in $V(\mathcal{A}_{m-1})$, then evidently $B_r^{out}(v)\cap\mathcal{A}=\emptyset$ and $B_r^{out}(v)$ is a tree.

Since on each of these steps, the half-edge $f(\mathcal{A}_m)$ is being matched to a u.a.r. un-matched half-edge, uniformly over the at most $|E(\mathcal{T}_r)|$ steps it takes to reveal $B_r^{out}(v)$, the probability that the half-edge it is matched to belongs to $\mathcal{A}_{m-1}$ is at most

$$\frac{d(|V(\mathcal{A})|+d^r)}{\Delta n-(d|V(\mathcal{A})|+d^r)} \le \frac{|V(\mathcal{A})|+d^r}{n-(|V(\mathcal{A})|+d^r)}\,.$$

(The first inequality here uses the fact that in the BFS of $B_r^{out}(v)$, there are at most d^r vertices of the ball that have been discovered but not exhausted.) Union bounding over the at most $|E(\mathcal{T}_r)|\le 2\Delta d^r$ such attempts yields the desired bound. □

We can use a similar reasoning as the proof above to deduce a proof of Fact 2.3 as follows.

Proof of Fact 2.3. Fix any v and choose f so that the revealing scheme performs a BFS revealing of $B_R(v)$. In order for $B_R(v)$ to not be L-Treelike, it must be the case that for more than L different m's in the first $|E(\mathcal{T}_R)|$ steps, the half-edge $f(\mathcal{A}_{m-1})$ is being matched to a half-edge belonging to $\mathcal{A}_{m-1}$. (If there were at most L such steps, then the removal of the at-most L edges formed by those at-most L matchings in the revealing scheme, evidently leaves a tree, so that $B_R(v)$ would be L-Treelike.) Uniformly over $\mathcal{A}_{m-1}$, the probability of this in the m'th step is at most $d^{R+1}/(\Delta(n-m))$. Summing over the at most $|E(\mathcal{T}_R)|$ many such attempts while revealing $B_R(v)$, we find that for every $\ell\ge 1$,

$$\mathbf{P}_{\mathrm{CM}}(B_R(v) \text{ is not } \ell\text{-Treelike}) \le \mathbb{P}\Big(\mathrm{Bin}\big(|E(\mathcal{T}_R)|,\frac{d^R}{n-|E(\mathcal{T}_R)|}\big)>\ell\Big)\,. \tag{4.5}$$

Recall that the standard Chernoff bound applied to a Poisson binomial distribution with mean $\mu=Np$ says that for every $s\ge\mu$,

$$\mathbb{P}\big(\mathrm{Bin}(N,p)\ge s\big) \le e^{s-\mu}\Big(\frac{s}{\mu}\Big)^{-s}\,. \tag{4.6}$$

With the choice $R=(\frac12-\delta)\log_d n$, so that $d^R=n^{\frac12-\delta}$ and $|E(\mathcal{T}_R)|\le 2\Delta d^R$, (4.6) implies that the right-hand side of (4.5) is at most $(Cn^{-\delta})^\ell$ for some $C(\Delta)$ and large enough n. As a consequence, choosing $L>2\delta^{-1}$, we would find

$$\sup_{v\in\{1,\dots,n\}} \mathbf{P}_{\mathrm{CM}}(B_R(v) \text{ is not } L\text{-Treelike}) \le o(n^{-2})\,. \tag{4.7}$$

It remains to translate this to a bound under $\mathbf{P}_{\mathrm{RRG}}$. This follows by the following standard comparison argument. Let Γ_{RRG} be the event that the graph $\mathcal{G}\sim\mathbf{P}_{\mathrm{CM}}$ has no self-loops or double edges (i.e., it is a simple graph). Taking $\Gamma=\{\mathcal{G}:\mathcal{G} \text{ is not } (L,R)\text{-Treelike}\}$ in (4.2) and union bounding (4.7) over the n vertices yields the desired bound. □

4.2.2. *Properties of the coupling of localized Markov chains.* The following lemma is the key fact about the construction of the grand coupling of FK dynamics, Definition 4.8, whereby after revealing some $X^1_{A,t}$, we can control the influence that revealing has on $X^1_{B,t}$ for $A \cap B = \emptyset$. In this manner, through the revealing procedure of Definition 4.11, which reveals different localized configurations $X^1_{A_m,t}$ iteratively, as long as $t \geq T_{\text{BURN}}$ these are each close to their respective stationary distributions of $\pi^1_{A_m}$, so that it is approximately a concatenation of localized FK models on treelike graphs, inducing an exponential decay of connectivities.

Lemma 4.17. *Recall the coupling of Definitions 4.8–4.9 of the distributions $(\mathbb{P}^1_{A,t})_{A \subset \mathfrak{M}_n, t \geq 1}$. Suppose we have revealed edges in $E(\mathcal{G})$ showing $E(\mathcal{G}) \cap A = A$.*

(1) The configuration $X^1_{A,t}(A)$ is measurable with respect to $\kappa_{A,t}$ (the number of edge-updates in A), the edges chosen to update $\mathcal{O}_{A,\kappa_{A,t}}$, and the uniform random variables $(U_{e,s})_{e \in A, s \leq t}$ on those edges. The number $\kappa_{A,t}$ is distributed as $\mathrm{Bin}(t, 2|A|/\Delta n)$; the sequence $\mathcal{O}_{A,\kappa_{A,t}}$ is distributed as $(\tilde{e}_j)_{j \leq \kappa_{A,t}}$ drawn i.i.d. from A. The values $(U_{e,s})_{e \in A, s \leq t}$ are distributed as i.i.d. Unif$[0,1]$.

(2) Suppose B is such that $E(\mathcal{G}) \cap B = B$ and $A \cap B = \emptyset$. Conditionally on $\kappa_{A,t}$, $\mathcal{O}_{A,\kappa_{A,t}}$, and $(U_{e,s})_{e \in A, s \leq t}$, the distribution of $X^1_{B,t}(B)$ is given as follows. The number of updates $\kappa_{B,t}$ is drawn from $\mathrm{Bin}(t - k_{A,t}, 2|B|/(\Delta n - 2|A|))$, and the edges chosen are distributed as $(\tilde{e}_j)_{j \leq \kappa_{B,t}}$ drawn i.i.d. amongst B. The random variables $(U_{e,s})_{e \in B, s \leq t}$ are distributed as i.i.d. Unif$[0,1]$.

Proof. Let $\mathcal{G}$ be any graph having $E(\mathcal{G}) \cap A = A$. We claim that uniformly over $\mathcal{G}$, items (1)–(2) above hold. Observe first that $|E(\mathcal{G})| = \Delta n/2$ necessarily, and therefore uniformly over such $\mathcal{G}$, the number of updates on edges in A by time t in the update sequence $(e_s)_{s \leq t}$ is distributed as $\mathrm{Bin}(t, 2|A|/(\Delta n))$. Evidently, the distribution of $\mathcal{O}_{A,\kappa_{A,t}}$ only depends on $\kappa_{A,t}$ and not on the times these updates were; in particular, given that $e_j \in A$ for some j, the law of e_j is clearly uniform at random on A. Finally, notice that for every e, the sequence $(U_{e,s})_{s \leq t}$ is independent of all other sources of randomness, implying the desired item (1).

Turning to item (2), we fix a $\kappa_{A,t}, \mathcal{O}_{A,\kappa_{A,t}}$ and family $(U_{e,s})_{e \in A, s \leq t}$. We can condition further on the exact times of the updates in A, i.e., $(e_s)_{s \leq t} \cap A$. Conditionally on that set of updates, the distribution on the remaining updates is evidently $t - \kappa_{A,t}$ i.i.d. draws from $E(\mathcal{G}) \setminus A$. It is then clear that $\kappa_{B,t}$ counts the number of times, amongst these remaining draws, that the update is in B. As in item (1), the induced distribution on $\mathcal{O}_{B,\kappa_{B,t}}$ is then the same as $\kappa_{B,t}$ i.i.d. draws from the edges of B. Finally, for every $e \in B$, the uniform random variables $(U_{e,s})_{s \leq t}$ are independent of all other sources of randomness. □

4.3. **Domination by the modified branching process (Z_k).** In this section, we establish the stochastic domination of the sequence $(\mathcal{V}_k)_{k \geq 0}$ from Definition 4.11 by the branching process (Z_k) of Definition 4.13.

Proof of Lemma 4.14. We prove the desired stochastic domination by induction over ℓ. The base case, $Z_0 = |\mathcal{V}_0|$, is by construction. Now fix $\ell \geq 1$ and suppose by way of induction that the following stochastic domination holds:

$$(|\mathcal{V}_j| \mathbf{1}_{\mathfrak{E}_t^{\mathfrak{m}_j}} \mathbf{1}_{\{\mathfrak{m}_{j-1} \leq n^{1/2-\delta/2}\}})_{j \leq \ell-1} \preccurlyeq (Z_j)_{j \leq \ell-1} .$$

Thus there exists a monotone coupling of the sequence on the left-hand side, such that it is below the sequence $(Z_j)_{j \leq \ell-1}$ in the natural element-wise ordering on the sequence. Working on that coupling, it suffices for us to then show that on the intersection $\mathfrak{E}_t^{\mathfrak{m}_\ell} \cap \{\mathfrak{m}_{\ell-1} \leq n^{1/2-\delta/2}\}$, for every $m \in \{\mathfrak{m}_{\ell-1}+1, ..., \mathfrak{m}_\ell\}$, the distribution of the children of v_m is stochastically below the progeny distribution of Definition 4.13.

Observe, first of all, that for every $m \in \{\mathfrak{m}_{\ell-1}+1, ..., \mathfrak{m}_\ell\}$, on $\mathfrak{E}_t^{\mathfrak{m}_\ell} \cap \{\mathfrak{m}_{\ell-1} \leq n^{1/2-\delta/2}\}$, deterministically the number of children of v_m is bounded by

$$|V(\mathcal{A}_m \setminus \mathcal{A}_0)| \leq |V(\mathcal{T}_r)| \mathfrak{m}_{\ell-1} \leq |V(\mathcal{T}_r)| \sum_{j \leq \ell-1} |\mathcal{V}_j| \leq |V(\mathcal{T}_r)| \sum_{j \leq \ell-1} Z_j ,$$

where the last inequality is by the inductive hypothesis, and the fact that $\mathfrak{E}_t^{\mathfrak{m}_\ell} \cap \{\mathfrak{m}_{\ell-1} \leq n^{1/2-\delta/2}\}$ implies $\mathfrak{E}_t^{\mathfrak{m}_j} \cap \{\mathfrak{m}_{j-1} \leq n^{1/2-\delta/2}\}$ for all $j < \ell$.

Now, for every set of revealed edges $(A_l)_{l\le m-1}$, define the following events on $\mathcal{F}_{m-1}$ consisting of $(\kappa_{A_l,t})_{l\le m-1}$, edge-values $(\mathcal{O}_{A_l,\kappa_{A_l,t}})_{l\le m-1}$, uniform random variables $((U_{e,s})_{e\in A_l,s\le t})_{l\le m-1}$:

(1) Let $\Gamma_{\text{TREE},m}$ be the event that $B_r^{out}(v_m)\cap\mathcal{A}_{m-1}=\emptyset$ and $A_m=\mathcal{A}_m\setminus\mathcal{A}_{m-1}$ is a tree.
(2) Let $\Gamma_{\text{UPD},m}$ be the event that $\kappa_{A_m,t}\ge d^rT_{\text{BURN}}/(2\Delta n)$

We first claim that these two events each happen with probability $1-\frac{1}{3}n^{-1/2}$, uniformly over $(A_l)_{l\le m-1}$ and elements of $\mathcal{F}_{m-1}$ such that $\mathfrak{E}_t^\ell$ holds and $\mathfrak{m}_{\ell-1}\le n^{1/2-\delta/2}$. Given v_m, the law of A_m is independent of $\mathcal{F}_{m-1}$, and only depends on $\mathcal{A}_{m-1}$. (This can be seen from the explicit construction of the law of $\mathcal{F}_{m-1}$ in Lemma 4.17 as independent of $E(\mathcal{G})\setminus\mathcal{A}_{m-1}$.) Notice that if $\mathfrak{m}_{\ell-1}\le n^{1/2-\delta/2}$ and $|\mathcal{A}_0|\le n^{1/2-\delta}$ then $|V(\mathcal{A}_{m-1})|\le(1+|V(\mathcal{T}_r)|)n^{1/2-\delta/2}$. As such, by Lemma 4.16, for every $\mathcal{A}_0,\mathcal{V}_0$ such that $|\mathcal{A}_0|\le n^{1/2-\delta}$,

$$\begin{aligned}
&\sup_{(\mathcal{A}_{m-1},\mathcal{F}_{m-1})\in\mathfrak{E}_t^{\mathfrak{m}_\ell}\cap\{\mathfrak{m}_{\ell-1}\le n^{\frac12-\frac\delta2}\}}\mathbf{P}_{\text{CM}}(\Gamma^c_{\text{TREE},m}\mid\mathcal{A}_{m-1},\mathcal{F}_{m-1})\\
&\quad\le\sup_{\substack{\mathcal{A}_{m-1}:\\ V(\mathcal{A}_{m-1})\le 2|V(\mathcal{T}_r)|n^{\frac12-\frac\delta2}}}\sup_{v_m}\mathbf{P}_{\text{CM}}(B_r^{out}(v_m)\cap\mathcal{A}_{m-1}\neq\emptyset\text{ or }B_r^{out}(v_m)\text{ is not a tree}\mid\mathcal{A}_{m-1})\\
&\quad\le\sup_{\substack{\mathcal{A}_{m-1}:\\ V(\mathcal{A}_{m-1})\le 4\Delta d^rn^{\frac12-\frac\delta2}}}\frac{2\Delta d^r(|V(\mathcal{A}_{m-1})|+d^r)}{n-(|V(\mathcal{A}_{m-1})|+d^r)}\le\frac{10\Delta^2d^{2r}n^{\frac12-\frac\delta2}}{n-5\Delta d^rn^{\frac12-\frac\delta2}}\,.
\end{aligned}$$

Thus, for n large enough and $r=o(\log n)$, the above is at most $\frac13n^{-1/2}$ as desired.

We next turn to the probability of $\Gamma^c_{\text{UPD},m}\cap\Gamma_{\text{TREE},m}$. Recall from item (2) of Lemma 4.17 that conditionally on $\mathcal{F}_{m-1}$, the distribution of $\kappa_{A_m,t}$ is

$$\kappa_{A_m,t}\sim\text{Bin}\Big(t-\sum_{l\le m-1}\kappa_{A_l,t},\frac{2|A_m|}{\Delta n}\Big)\,.$$

Since we are on the event $\mathfrak{E}_t^{\mathfrak{m}_\ell}$ and thus $\mathfrak{E}_t^{m-1}$, we have that $\sum_{l\le m-1}\kappa_{A_l,t}\le 4m|E(\mathcal{T}_r)|t/(\Delta n)$, from which we deduce, using $m\le\mathfrak{m}_\ell\le|V(\mathcal{T}_r)|\mathfrak{m}_{\ell-1}\le|V(\mathcal{T}_r)|n^{1/2-\delta/2}$, that the number of trials in the binomial is at least

$$t(1-16\Delta d^{2r}n^{-\frac12-\frac\delta2})\ge t/2\,,$$

as long as $r=o(\log n)$. Since we are on the event $\Gamma_{\text{TREE},m}$, we have $d^r\le|A_m|\le|E(\mathcal{T}_r)|\le2\Delta d^r$, and we see from lower tail estimates on binomial random variables that

$$\begin{aligned}
&\sup_{(\mathcal{A}_{m-1},\mathcal{F}_{m-1})\in\mathfrak{E}_t^{m-1}\cap\{\mathfrak{m}_{\ell-1}\le n^{1/2-\delta/2}\}}\mathbb{P}\big(\Gamma^c_{\text{UPD},m}\cap\Gamma_{\text{TREE},m}\mid\mathcal{A}_{m-1},\mathcal{F}_{m-1}\big)\\
&\qquad\le\mathbb{P}\big(\text{Bin}(T_{\text{BURN}}/2,2d^r/(\Delta n))\le d^rT_{\text{BURN}}/(2\Delta n)\big)\le\frac13n^{-1/2}\,,
\end{aligned}$$

as long as C_0 in (4.3) is sufficiently large (depending on r,Δ).

By item (2) of Lemma 4.17, conditionally on *any* $(\mathcal{A}_l)_{l\le m-1}$ and $\mathcal{F}_{m-1}$, and any $A_m\in\Gamma_{\text{TREE},m}$ and $\kappa_{A_m,t}\in\Gamma_{\text{UPD},m}$, the conditional distribution of $X^1_{A_m,t}(A_m)$ is equivalent (up to relabeling of edges) to that of $\kappa_{A_m,t}$ updates of a heat-bath chain $(Y^1_s)_s$ on a subtree $\hat{\mathcal{T}}_r$ of the complete tree $\mathcal{T}_r$ with $(1,\circlearrowleft)$-wired boundary conditions, initialized from $Y^1_0\equiv1$. Notice that the equivalent sub-tree $\hat{\mathcal{T}}_r$ consists of some $k\le d$ of the children of the root, together with their complete sub-trees. In particular, the random-cluster model on A_m with wired boundary conditions is stochastically below the FK model on the corresponding subset of $\mathcal{T}_r$ with its $(1,\circlearrowleft)$ boundary conditions. In particular, the number of leaves in the FK cluster of the root under $\pi^{(1,\circlearrowleft)}_{\hat{\mathcal{T}}_r}$ is stochastically below the same quantity under $\pi^{(1,\circlearrowleft)}_{\mathcal{T}_r}$. It therefore suffices for us to show that as long as A_m is a tree disjoint from $\mathcal{A}_{m-1}$ and $\kappa_{A_m,t}\ge d^rT_{\text{BURN}}/(2\Delta n)$, we have

$$\big\|\mathbb{P}\big(Y^{(1,\circlearrowleft)}_{\kappa_{A_m,t}}\in\cdot\big)-\pi^{(1,\circlearrowleft)}_{\hat{\mathcal{T}}_r}\big\|_{\text{TV}}\le\frac13n^{-1/2}\,.$$

This follows as long as C_0 is sufficiently large (depending on Δ), from the fact that

$$\frac{d^r T_{\text{BURN}}}{2\Delta n} \geq \frac{C_0 d^r}{2\Delta |E(\mathcal{T}_r)|} \log n \cdot t_{\text{MIX}}(\mathcal{T}_r, (1, \circlearrowleft)) \geq \frac{C_0 d^r}{2\Delta |E(\mathcal{T}_r)|} \log n \cdot t_{\text{MIX}}(\hat{\mathcal{T}}_r, (1, \circlearrowleft)) \,,$$

and $|E(\mathcal{T}_r)| \leq 2\Delta d^r$, together with the sub-multiplicativity of total-variation distance. □

4.4. **Sub-criticality and tail bounds for the dominating branching process.** We now analyze the process (Z_k) of Definition 4.13, and show that it indeed is sub-critical, and satisfies good tails on its total population. For ease of notation, let $\mathcal{P}_k = \sum_{\ell \leq k} Z_\ell$ be the total population after k generations.

Proof of Lemma 4.15. Since (Z_k) is a size-dependent, non-Markov process, we cannot directly use results on branching processes to control its growth. Instead, to control the population of the process (Z_k), we compare it to a sum of branching processes in the following manner. Consider the stopping generation κ_λ for exceeding population $K_0 Z_0 + \lambda$, i.e.,

$$\kappa_\lambda = \inf\{k : \mathcal{P}_\kappa > K_0 Z_0 + \lambda\} \,.$$

Our aim is to control the probability that $\kappa_\lambda < \infty$. Let $\Gamma_{M,k}$ be the event that no more than M of the progeny counts $((\chi_{i,\ell})_{i \leq Z_\ell})_{\ell \leq k-1}$ were Bad. By (4.6), we get

$$\mathbb{P}(\Gamma^c_{M,\kappa}) \leq \mathbb{P}\big(\text{Bin}(K_0 Z_0 + \lambda, n^{-\frac{1}{2}}) > M \big) \leq \exp\big(M - \mu - M \log \tfrac{M}{\mu} \big)$$

where μ is the mean of the Binomial, i.e., $\mu = (K_0 Z_0 + \lambda) n^{-\frac{1}{2}}$. As long as n is sufficiently large and $Z_0, \lambda \leq n^{\frac{1}{2} - \delta}$ for $\delta > 0$, so that $\mu \leq 2K_0 n^{-\delta}$, this implies for some $C > 0$,

$$\mathbb{P}(\Gamma^c_{M,\kappa}) \leq C e^M n^{-\delta M} \,.$$

Next consider the event that $\kappa < \infty$ on the event $\Gamma_{M,k}$. On $\Gamma_{M,k}$, we dominate the population $\mathcal{P}_k$ by the following sum of sub-critical branching processes with bounded progeny distributions.

Define $(\tilde{Z}^{(1)}_k)_k$ to be the branching process initialized at $\tilde{Z}^{(1)}_0 = Z_0$ with progeny $(\tilde{\chi}^{(1)}_{i,k})$, distributed i.i.d. from the distribution of the number of leaves connected to the root, in a sample from $\pi^{(1,\circlearrowleft)}_{\mathcal{T}_r}$, i.e., the distribution of $(\chi_{i,k})$ conditionally on the progeny number not being Bad. Let $\tilde{\mathcal{P}}^{(1)}_k = \sum_{\ell \leq k} \tilde{Z}^{(1)}_k$. For each $1 \leq j \leq M$, iteratively let $\tilde{Z}^{(j)}_k$ be an independent branching process with the same progeny distribution, initialized from $\tilde{Z}^{(j)}_0 = |V(\mathcal{T}_r)| \tilde{\mathcal{P}}^{(j-1)}_\infty$, where we recall $|V(\mathcal{T}_r)| \leq 2\Delta d^r$.

The following stochastic domination is clear by construction if we decompose the process (Z_k) revealed in a breadth-first manner, into its excursions between the at most M times (on the event $\Gamma_{M,k}$) when the progeny number $\chi_{i,k}$ was Bad.

Claim 4.18. *Fix any $k \geq 1$. We have the stochastic domination*

$$\mathcal{P}_k \mathbf{1}\{\Gamma_{M,k}\} \preccurlyeq \sum_{j \leq M} \tilde{\mathcal{P}}^{(j)}_\infty \,.$$

With this domination in hand, notice that in order for $\kappa < \infty$ while $\Gamma_{M,\kappa}$ holds, there must exist some $k \leq K_0 Z_0 + \lambda$ such that $\Gamma_{M,k}$ holds and $\mathcal{P}_k \geq K_0 Z_0 + \lambda$. Therefore, by a union bound,

$$\begin{aligned} \mathbb{P}\Big(\mathcal{P}_\infty \geq K_0 Z_0 + \lambda, \Gamma_{M,\kappa} \Big) &\leq \sum_{k \leq K_0 Z_0 + \lambda} \mathbb{P}\Big(\mathcal{P}_k \geq K_0 Z_0 + \lambda, \Gamma_{M,k} \Big) \\ &\leq (K_0 Z_0 + \lambda) \mathbb{P}\Big(\sum_{j \leq M} \tilde{\mathcal{P}}^{(j)}_\infty \geq K_0 Z_0 + \lambda \Big) \,. \end{aligned}$$

We claim that if $\sum_{j \leq M} \tilde{\mathcal{P}}^{(j)}_\infty \geq K_0 Z_0 + \lambda$ holds, there must exist $j \leq M$ for which $\tilde{Z}^{(j)}_0 \leq K_0 Z_0 + \lambda$, and

$$\tilde{\mathcal{P}}^{(j)}_\infty \geq |V(\mathcal{T}_r)|^{-1} \Big(\frac{K_0}{M} \Big)^{1/M} (\tilde{Z}^{(j)}_0 + K_0^{-1} M^{-2} \lambda) =: C_{r,\Delta,M} K_0^{1/M} (\tilde{Z}^{(j)}_0 + K_0^{-1} M^{-2} \lambda) \,.$$

Indeed, if no such j existed, as long as K_0 is sufficiently large, we could bound $\sum_{j\le M}\tilde{\mathcal{P}}_\infty^{(j)}$ by

$$\sum_{j\le M} C_{r,\Delta,M}K_0^{1/M}(\tilde{Z}_0^{(j)}+\lambda) \le M\Big[\frac{K_0}{M}\tilde{Z}_0^{(1)} + K_0^{-1}M^{-2}\lambda(1+\cdots+\frac{K_0}{M})\Big] \le K_0Z_0+\lambda\,.$$

Now fix any $j\le M$, any $\tilde{Z}_0^{(j)}$ and consider the branching process $\tilde{Z}_k^{(j)}$. This is a branching process with progeny distribution having mean $\mathbf{m}=A(\hat{p}d)^r$ for some $A(p,q)$ per Lemma 5.1. Since $\hat{p}<d^{-1}$ when $p<p_u(q,\Delta)$, as long as r is greater than some $r_0(p,q,\Delta)$, for n sufficiently large we have $\mathbf{m}<1$, and $\tilde{Z}_k^{(j)}$ is sub-critical. Additionally, the progeny distribution of $\tilde{Z}_k^{(j)}$ is almost surely bounded by $|\partial\mathcal{T}_r|\le\Delta d^{r-1}$. As such, using the standard breadth-first exploration of the total population of the branching process $\tilde{Z}_k^{(j)}$ (through which $\tilde{\mathcal{P}}_k^{(j)}$ is expressed as the random walk $\tilde{Z}_0^{(j)}+\sum_{\ell\le k}\sum_{i\le\tilde{Z}_k^{(j)}}(\tilde{\chi}_{i,k}^{(j)}-1)$), we can bound

$$\mathbb{P}\Big(\tilde{\mathcal{P}}_\infty^{(j)}\ge N_\lambda^{(j)}\Big)\le\mathbb{P}\Big(\sum_{i\le N_\lambda^{(j)}}\tilde{\chi}_i>N_\lambda^{(j)}-\tilde{Z}_0^{(j)}\Big)\quad\text{for}\quad N_\lambda^{(j)}:=C_{r,\Delta,M}K_0^{1/M}(\tilde{Z}_0^{(j)}+K_0^{-1}M^{-2})\lambda\,,$$

where $\tilde{\chi}_i$ are i.i.d. copies of $\tilde{\chi}_{i,k}^{(j)}$. Now observe that if $K_0(p,q,\Delta,M)$ is sufficiently large, the right-hand in the probability above exceeds the mean $\mathbf{m}N_\lambda^{(j)}$ by some $cN_\lambda^{(j)}$ for $c=c(p,q,\Delta,M,K_0)>0$. As this is a tail probability for a sum of i.i.d. random variables, by Hoeffding's inequality, it is at most

$$\exp\big(-(cN_\lambda^{(j)})^2/(4N_\lambda^{(j)}d^{2r})\big)\le\exp(-c'N_\lambda^{(j)})\,,$$

for some $c'(r,\Delta,M)>0$. Taking a union bound over the M possible values of $j\le M$, we altogether find

$$\sum_{k\le K_0Z_0+\lambda}\mathbb{P}\Big(\mathcal{P}_k\ge K_0Z_0+\lambda,\Gamma_{M,k}\Big)\le(K_0Z_0+\lambda)M\exp\big(-c'N_\lambda^{(j)}\big)\,.$$

It follows from this, and the definition of $N_\lambda^{(j)}$, that for some $C(p,q,\Delta,M,K_0)$ large enough,

$$\mathbb{P}(\kappa<\infty)\le\mathbb{P}(\Gamma_{M,\kappa}^c)+\sum_{k\le K_0Z_0+\lambda}\mathbb{P}\Big(\mathcal{P}_k\ge K_0Z_0+\lambda,\Gamma_{M,k}\Big)\le Cn^{-\delta M}+C\exp\big(-(Z_0+\lambda)/C\big)\,,$$

concluding the proof. □

4.5. **Proof of exponential tail on cluster sizes and shattering.** We are now in position to conclude the proof of the exponential tail bound on clusters of $X_{\mathcal{G},t}^1$, and use that to deduce that $X_{\mathcal{G},t}^1$ is (K,R)-Sparse, except with probability $o(n^{-2})$. We begin by using Lemmas 4.14–4.15 to prove the following tail bound on the sequence $(\mathcal{V}_k)$, which are the roots of the balls revealed through the revealing process of Definition 4.11.

Lemma 4.19. *Fix $\delta>0$ and consider the revealing procedure for any initial subsets $\mathcal{A}_0$ and $\mathcal{V}_0$ having $|\mathcal{A}_0|,|\mathcal{V}_0|\le n^{\frac12-\delta}$. For every $M\ge1$, there exist $C(p,q,\Delta,M)$, $K_0(p,q,\Delta,M)$ and $C_0(p,q,\Delta,\delta,M)$ $r(p,q,\Delta)$ in the definition of T_{BURN} in (4.3) such that for all $t\ge T_{\text{BURN}}$ and all $0\le\lambda\le n^{\frac12-\delta}$,*

$$\mathbb{P}\Big(\sum_{0\le k\le k_\emptyset}|\mathcal{V}_k|\ge K_0|\mathcal{V}_0|+\lambda\Big)\le C\exp(-\lambda/C)+Cn^{-\delta M}\,,$$

Proof. Fix K_0 large to be chosen later, and define the following stopping generation

$$\varsigma=\inf\{\ell:\mathfrak{m}_{\ell-1}>K_0|\mathcal{V}_0|+\lambda\}\,.$$

Recall $\mathfrak{E}_t$ from (4.4). Since for every $\ell\le\varsigma$, we have from Lemma 4.14, that $(|\mathcal{V}_\ell|\mathbf{1}\{\mathfrak{E}_t\})_{j\le\ell}\preccurlyeq(Z_j)_{j\le\ell}$, we have that if C_0 in (4.3) is sufficiently large, the probability of $\{\varsigma<\infty\}$ is bounded by the probability of $\mathcal{P}_\infty=\sum_{k\ge0}Z_k\ge K_0Z_0+\lambda$. By Lemma 4.14, we obtain

$$\mathbb{P}\Big(\sum_{k\le k_\emptyset}|\mathcal{V}_k|\ge K_0|\mathcal{V}_0|+\lambda\Big)\le\mathbb{P}\Big(\sum_{k\ge0}Z_k\ge K_0Z_0+\lambda\Big)+\mathbb{P}(\mathfrak{E}_t^c)\,.$$

Lemma 4.15 implies the existence of $r(p,q,\Delta)$ such that the first-term above is at most $C_1\exp(-\lambda/C_1)+C_1n^{-\delta M}$ for some $C_1(p,q,\Delta,M)$.

Next, consider $\mathbb{P}(\mathfrak{E}_t^c)$. By a union bound and item (1) of Lemma 4.17, with the trivial observations that $\mathfrak{m}_{k_\emptyset}\le n$ and $|A_m|\le|E(\mathcal{T}_r)|\le 2\Delta d^r$ necessarily, we get for every $t\ge T_{\text{BURN}}$,

$$\mathbb{P}(\mathfrak{E}_t^c)\le n\mathbb{P}\Big(\text{Bin}\,\big(t,\frac{2|E(\mathcal{T}_r)|}{\Delta n}\big)>\frac{4|E(\mathcal{T}_r)|}{\Delta n}t\Big)\,.$$

The above entails a deviation of at least $4td^rn^{-1}$ from its mean; as such, by standard tail estimates for binomials, for every $t\ge T_{\text{BURN}}$,

$$\mathbb{P}(\mathfrak{E}_t^c)\le n\exp(-td^rn^{-1})\,, \tag{4.8}$$

which is at most $n^{-\delta M}$ for n large, as long as C_0 in (4.3) is sufficiently large (depending on δM). The desired bound then follows up to a change of the constant C. □

Before proceeding to prove Proposition 4.21, let us translate the tail bound of Lemma 4.19 on $\sum_k|\mathcal{V}_k|$ to a tail bound on the FK cluster of a single vertex under $X^1_{\mathcal{G},t}$ and $\pi_{\mathcal{G}}$. Notice that towards the proofs of Theorem 3.2 and Proposition 4.21, it suffices to show these for $t\ge T_{\text{BURN}}$ for some fixed choices of C_0,r in (4.3) depending on p,q,Δ (as $t_{\text{MIX}}(\mathcal{T}_r,(1,\circlearrowleft))$ is of course independent of n).

Proof of Theorem 3.2. Fix any $v\in\{1,\dots,n\}$, let $\mathcal{A}_0=\emptyset$ and let $\mathcal{V}_0=\{v\}$ in Definition 4.11. By Observation 4.12, for each $\mathcal{G}\sim\mathbf{P}_{\text{CM}}$, the cluster of v in the configuration $X^1_{\mathcal{G},t}$, denoted $\mathcal{C}_v(X^1_{\mathcal{G},t})$ is a subset of $\mathcal{C}_v(\tilde{\omega}_t)$, which in turn is a subset of $V(\mathcal{A}_{\mathfrak{m}_\emptyset})$, so that

$$|\mathcal{C}_v(X^1_{\mathcal{G},t})|\le|\mathcal{C}_v(\tilde{\omega}_t)|\le|V(\mathcal{T}_r)|\sum_{k\le k_\emptyset}|\mathcal{V}_k|\le 2\Delta d^r\sum_{k\le k_\emptyset}|\mathcal{V}_k|\,.$$

By Lemma 4.19 and the above, we find that for each M, there exists $C(p,q,\Delta,M)$ such that

$$\mathbb{P}\big((\mathcal{G},X^1_{\mathcal{G},t}):|\mathcal{C}_v(X^1_{\mathcal{G},t})|\ge 2\Delta d^r(1+\lambda)\big)\le C\exp(-\lambda/C)+Cn^{-\delta M}\,.$$

Observing that $\mathbb{P}(X^1_{\mathcal{G},t}\in\cdot)=\mathbf{E}_{\text{CM}}[P(X^1_{\mathcal{G},t}\in\cdot)]$, we can use Markov's inequality to write

$$\mathbf{P}_{\text{CM}}\Big(\mathcal{G}:P\Big(X^1_{\mathcal{G},t}:|\mathcal{C}_v(X^1_{\mathcal{G},t})|\ge 2\Delta d^r(1+\lambda)\Big)\ge\sqrt{Ce^{-\lambda/C}+Cn^{-\delta M}}\Big)\le\sqrt{Ce^{-\lambda/C}+Cn^{-\delta M}}\,.$$

We can obtain the same bound for $\mathbf{P}_{\text{RRG}}$ by (4.2), up to a multiplicative $c(\Delta)^{-1}$ on the right-hand side. Taking M such that $\delta M>2K$ and using the fact that $\sqrt{a+b}\le\sqrt{a}+\sqrt{b}$ for all $a,b\ge 0$, we deduce the desired tail bound on $|\mathcal{C}_v(X^1_{\mathcal{G},t})|$ up to the change of constant C to $2C$. Using the monotonicity $X^1_{\mathcal{G},t}\succcurlyeq\pi^1_{\mathcal{G}}$ implies the analogous bound for $|\mathcal{C}_v(\omega)|$ where $\omega\sim\pi_{\mathcal{G}}$. □

We now turn to proving that for typical random graphs, the configuration $X^1_{\mathcal{G},t}$ is (K,R)-Sparse with high probability for all $t\ge T_{\text{BURN}}$. This allows us to localize to treelike balls with sparse boundary conditions. Let us define the following subset of the boundary of a set H, which we will apply with the choice $H=B_R(v)$.

Definition 4.20. For a subgraph $H=(V(H),E(H))$ of $\mathcal{G}$ and a configuration ω on $E(\mathcal{G})$, let us define $\mathfrak{V}_H(\omega)$ as the subset of vertices in $V(H)$ in non-trivial components in the boundary condition induced on H by $\omega(E(\mathcal{G})\setminus E(H))$ (a connected component is non-trivial when it has at least two vertices).

We first prove the following proposition, giving a tail bound on $\mathfrak{V}_{B_R(v)}(X^1_{\mathcal{G},t})$; after proving this proposition, we straightforwardly use it to conclude (K,R)-sparsity of $X^1_{\mathcal{G},t}$, i.e., Theorem 4.3.

Proposition 4.21. *Let p,q,Δ be such that $p<p_u(q,\Delta)$. Fix $\delta>0$ and let $R=(\frac{1}{2}-\delta)\log_d n$. There exists $K(p,q,\Delta,\delta)$ such that if $\mathcal{G}\sim\mathbf{P}_{\text{CM}}$, with probability $1-O(n^{-2})$, $\mathcal{G}$ is such that for all $t\ge T_{\text{BURN}}$*

$$\sup_{v\in V(\mathcal{G})}P\big(X^1_{\mathcal{G},t}:|\mathfrak{V}_{B_R(v)}(X^1_{\mathcal{G},t})|>K\big)\le O(n^{-2})\,.$$

Proof of Proposition 4.21. Fix $v \in \{1, ..., n\}$ and $\delta > 0$, and let $R = (\frac{1}{2} - \delta) \log_d n$.

We apply the revealing procedure of Definition 4.11 with the choices $\mathcal{A}_0 = E(B_R(v))$ and $\mathcal{V}_0 = \partial B_R(v)$. Recall from Observation 4.12 that the FK-clusters of $\mathcal{V}_0$ induced by $\tilde{\omega}_t(E(\mathcal{G}) \setminus \mathcal{A}_0)$ ($\tilde{\omega}_t$ was extended to be all wired off of $\mathcal{A}_{\mathfrak{m}_{k_\emptyset}} \setminus \mathcal{A}_0$) are confined to the set $\mathcal{A}_{\mathfrak{m}_{k_\emptyset}} \setminus \mathcal{A}_0$, and the extended configuration $\tilde{\omega}_t$ satisfies $\tilde{\omega}_t \geq X^1_{\mathcal{G},t}$. Thus, the sets $\mathfrak{V}_{B_R(v)}(\tilde{\omega}_t)$ and in turn $\mathfrak{V}_{B_R(v)}(X^1_{\mathcal{G},t})$, are below the number of vertices in $\mathcal{V}_0$ that share a connected component of $\mathcal{A}_{\mathfrak{m}_{k_\emptyset}} \setminus \mathcal{A}_0$ with another vertex of $\mathcal{V}_0$.

Suppose that through the revealing process of Definition 4.11, for each m, the edges of $B^{out}_r(v_m)$ are revealed one at a time per Definition 4.5. Notice then, that $|\mathfrak{V}_{B_R(v)}(\mathcal{A}_{\mathfrak{m}_{k_\emptyset}} \setminus \mathcal{A}_0)|$ is bounded by the number of times through the revealing of $\mathcal{A}_{\mathfrak{m}_{k_\emptyset}}$, that a half-edge is matched up to a half-edge belonging to a vertex that has been discovered at that point. Throughout this process, conditionally on an exposed edge-set $\mathcal{A}$ (and the edge-update sequence, and uniform random variables given by the filtration up to that step of the revealing, but $E(\mathcal{G}) \setminus \mathcal{A}$ is independent of these), the law of the next half-edge to be matched is uniform amongst un-matched half-edges. Thus on any such edge-revealing, uniformly on the history of the revealing, the probability that it matches with a half-edge belonging to a discovered vertex is at most $\frac{|V(\mathcal{A}_{\mathfrak{m}_{k_\emptyset}})|}{n-2|V(\mathcal{A}_{\mathfrak{m}_{k_\emptyset}})|}$.

By a union bound, we obtain for Λ a sufficiently large constant (depending on p, q, Δ, r), for all $k \geq 1$,

$$\begin{aligned}\mathbb{P}\Big((\mathcal{G}, \tilde{\omega}_t) : |\mathfrak{V}_{v,R}(X^1_{\mathcal{G},t})| > k\Big) &\leq \mathbb{P}\Big(|V(\mathcal{A}_{\mathfrak{m}_{k_\emptyset}})| > \Lambda^2|\mathcal{V}_0|\Big) + \mathbb{P}\Big(\mathrm{Bin}\left(\Lambda^2|\mathcal{V}_0|, \frac{2\Lambda^2|\mathcal{V}_0|}{n}\right) > k\Big) \\ &\leq \mathbb{P}\Big(\sum_{k \leq k_\emptyset} |\mathcal{V}_k| \geq \Lambda|\mathcal{V}_0|\Big) + \mathbb{P}\Big(\mathrm{Bin}\left(\Lambda^2|\mathcal{V}_0|, \frac{2\Lambda^2|\mathcal{V}_0|}{n}\right) > k\Big)\,.\end{aligned}$$

By Lemma 4.19 and the fact that $\Lambda|\mathcal{V}_0| \leq n^{\frac{1}{2}-\frac{\delta}{2}}$ for n large, as long as Λ is large enough, the first term is at most n^{-5}. Using the fact that $|\mathcal{V}_0| \leq n^{\frac{1}{2}-\delta}$, we see that the mean of the binomial is at most $n^{-3\delta/2}$, so that by (4.6), for every fixed $k \geq 1$,

$$\mathbb{P}\Big((\mathcal{G}, X^1_{\mathcal{G},t}) : |\mathfrak{V}_{v,R}(X^1_{\mathcal{G},t})| > k\Big) \leq n^{-\delta k \wedge 4}\,, \tag{4.9}$$

for n large enough. Choosing $k = K$ sufficiently large (depending on δ), we can make the right-hand side at most n^{-4}. We deduce the proposition by using Markov's inequality to write

$$\mathbf{P}_{\mathrm{CM}}\Big(\mathcal{G} : P(X^1_{\mathcal{G},t} : |\mathfrak{V}_{B_R(v)}(X^1_{\mathcal{G},t})| \geq K) > n^{-2}\Big) \leq n^2 \mathbf{E}_{\mathrm{CM}}[P(X^1_{\mathcal{G},t} : |\mathfrak{V}_{B_R(v)}(X^1_{\mathcal{G},t})| \geq K)]\,,$$

and noticing that the expectation on the right equals the probability bounded in (4.9). □

Proof of Theorem 4.3. First of all, a union bound of Proposition 4.21 over $v \in \{1, ..., n\}$, with $\mathbf{P}_{\mathrm{CM}}$-probability $1 - O(n^{-1})$, $\mathcal{G}$ is such that

$$P\Big(X^1_{\mathcal{G},T} : \bigcup_{v \in V(\mathcal{G})} \{|\mathfrak{V}_{B_R(v)}(X^1_{\mathcal{G},t})| > K\}\Big) \leq n^{-1}\,.$$

We now translate this to a bound under $\mathbf{P}_{\mathrm{RRG}}$. Taking

$$\Gamma = \Big\{\mathcal{G} : P\Big(X^1_{\mathcal{G},t} : \bigcup_{v \in V(\mathcal{G})} \{|\mathfrak{V}_{B_R(v)}(X^1_{\mathcal{G},t})| > K\}\Big) > n^{-1}\Big\}\,,$$

in (4.2), we deduce that $\mathbf{P}_{\mathrm{RRG}}(\Gamma) \leq c^{-1}\mathbf{P}_{\mathrm{CM}}(\Gamma) \leq c^{-1}n^{-1}$ for some $c(\Delta) > 0$, as needed. □

5. Sharp rates of correlation decay in trees and treelike graphs

In this section we establish the precise exponential decay rate of influence from an $O(1)$-Sparse boundary condition on the root of an $O(1)$-Treelike ball. We recall from Section 3, that getting the right decay rate, (as opposed to e.g., using the decay rate of connectivity from the root to the boundary) is central to pushing our argument through for all $p < p_u$. In particular, the decay rate of influence will be inherited from *twice* the exponential decay rate of the wired tree.

Recall that the uniqueness point $p_u(q,\Delta)$ is defined by a transition on the wired Δ-regular tree, where the measure $\pi^1_{\mathcal{T}_h}$ transitions between exponentially small (in h) probability of a root-to-leaf connection, to giving this event uniformly positive probability. A recursion for this connectivity probability was calculated in [5, Lemma 33]. A careful examination of this recursion will yield the following identification of the rate of the exponential decay with $\hat{p}$ of (2.2).

Lemma 5.1. *Let $p < p_u(q,\Delta)$. There exists $C(p,q,\Delta)$ such that for every h and every leaf $u \in \partial\mathcal{T}_h$,*

$$\pi^{(1,\circlearrowleft)}_{\mathcal{T}_h}(\omega : u \in \mathcal{C}_\rho(\omega)) \le C\hat{p}^h\,.$$

In particular, the probability of the root being connected to $\partial\mathcal{T}_h$ in ω is at most $C(\hat{p}d)^h$.

In Section 5.1, we establish Lemma 5.1. In Section 5.2, we show that influence in the random-cluster model travels through the existence of *two* distinct connections; thus on Treelike graphs, influence has twice the exponential decay rate of root-to-leaf connectivities on the wired tree. This will yield Proposition 3.3.

5.1. **Exponential decay rate in the wired Δ-regular tree.** Because of its recursive structure, connectivity properties of the random-cluster measure on the wired tree can be analyzed sharply. In this section, we pursue this and show that in the uniqueness regime of $p < p_u$, the probability of a connection from the root to a leaf at depth h is $O(\hat{p}^h)$, as one would have for the free tree (corresponding to i.i.d. $\mathrm{Ber}(\hat{p})$ percolation on $\mathcal{T}_h$). We first show that the probability of a root-to-boundary connection decays exponentially in h.

Let $\mathcal{T}_h$ be the complete Δ-regular tree of height h rooted at ρ. The wired "1" boundary conditions on $\mathcal{T}_h$ are those that wire all leaves of $\mathcal{T}_h$ (all vertices in $\partial\mathcal{T}_h$). Define the probability

$$\varphi_h := \pi^1_{\mathcal{T}_h}(\omega : \mathcal{C}_\rho(\omega) \cap \partial\mathcal{T}_h \neq \emptyset)\,,$$

that the root is connected to a leaf of $\mathcal{T}_h$. Using the recursive structure of the tree, it was shown in [5, Lemma 33] that if we define $\mu := \frac{p}{q} + 1 - p$, for every h, we have

$$\varphi_{h+1} = f(\varphi_h)\,, \qquad \text{where} \qquad f(x) = \frac{\big(\mu + p(1-\frac{1}{q})x\big)^d - \big(\mu - \frac{p}{q}x\big)^d}{\big(\mu + p(1-\frac{1}{q})x\big)^d + (q-1)\big(\mu - \frac{p}{q}x\big)^d}\,, \tag{5.1}$$

and for every $p < p_u(q,\Delta)$, this satisfies $\lim_{h\to\infty}\varphi_h = 0$. The following lemma establishes that this convergence is exponentially fast.

Lemma 5.2. *Let $p < p_u(q,\Delta)$. We have $\lim_{h\to\infty}\frac{\varphi_{h+1}}{\varphi_h} = \hat{p}d$. Moreover, $\varphi_h \le (\hat{p}d)^{h+o(h)}$.*

Proof. Consider the recursion of (5.1) for φ_h. Since $\lim_{h\to\infty}\varphi_h = 0$, if $\lim_{x\to 0}\frac{f(x)}{x}$ exists, we would have

$$\lim_{h\to\infty}\frac{\varphi_{h+1}}{\varphi_h} = \lim_{x\to 0}\frac{f(x)}{x} = \lim_{x\to 0}\frac{\big(\mu + p(1-\frac{1}{q})x\big)^d - \big(\mu - \frac{p}{q}x\big)^d}{x\big(\mu + p(1-\frac{1}{q})x\big)^d + x(q-1)\big(\mu - \frac{p}{q}x\big)^d}\,. \tag{5.2}$$

Since both the numerator and denominator of (5.2) are differentiable and have limit 0 as $x \to 0$, using L'Hôpital's rule we get

$$\begin{aligned}\lim_{x\to 0}\frac{f(x)}{x} &= \lim_{x\to 0}\frac{\partial_x\Big[\big(\mu + p(1-\frac{1}{q})x\big)^d - \big(\mu - \frac{p}{q}x\big)^d\Big]}{\partial_x\Big[x\big(\mu + p(1-\frac{1}{q})x\big)^d + x(q-1)\big(\mu - \frac{p}{q}x\big)^d\Big]} \\ &= \frac{dp(1-\frac{1}{q})\mu^{d-1} + d\frac{p}{q}\mu^{d-1}}{\mu^d + (q-1)\mu^d} = \frac{dp}{q\mu} = \frac{dp}{p+q(1-p)} = d\hat{p}\,.\end{aligned}$$

Recall that for every $0 < p < p_u$, we have $0 < d\hat{p} < 1$. Thus, there exists a sequence $\{\varepsilon_h\}$ such that $\lim_{h\to\infty}\varepsilon_h = 0$ and for every h,

$$\varphi_h = \varphi_1 \cdot \frac{\varphi_h}{\varphi_{h-1}} \cdots \frac{\varphi_2}{\varphi_1} = \varphi_1 \cdot \prod_{i=2}^{h}(\hat{p}d + \varepsilon_i)\,.$$

Expanding this out, we deduce the desired

$$\varphi_h = \varphi_1(\hat{p}d)^h \exp\left(\sum_{i=2}^{h} \ln\left(1+\frac{\varepsilon_i}{\hat{p}d}\right)\right) \le \varphi_1(\hat{p}d)^h \exp\left((\hat{p}d)^{-1}\sum_{i=2}^{h}\varepsilon_i\right) = (\hat{p}d)^{h+o(h)}\,. \qquad \square$$

Our aim is to now prove Lemma 5.1, bounding connectivities of the root to a single leaf.

Proof of Lemma 5.1. To prove Lemma 5.1, we write a recursion for the root-to-leaf connection probability. Let ϑ_h be the probability under $\pi^1_{\mathcal{T}_h}$ that the root is connected to the left-most leaf of depth h. Let $\vartheta^{\circlearrowleft}_h$ be the probability of the same event, under $\pi^{(1,\circlearrowleft)}_{\mathcal{T}_h}$ where we recall that the $(1,\circlearrowleft)$ boundary conditions additionally wire the leaves of $\mathcal{T}_h$ to the root. By monotonicity we have $\vartheta_h \le \vartheta^{\circlearrowleft}_h$ and by Lemma 6.3, we have $\vartheta^{\circlearrowleft}_h \le q^2\vartheta_h$.

Let $(I_i)_{i\le\Delta}$ be the indicator function of the event that there is a root-to-boundary path going through the i-th child of the root; set $I = \sum_{i=2}^{\Delta} I_i$. Then, we can write

$$\vartheta_h \le p\cdot\pi^1_{\mathcal{T}_h}(I\ge 1)\cdot\vartheta^{\circlearrowleft}_{h-1} + \hat{p}\vartheta_{h-1} \le \vartheta_{h-1}\left[pq^2\cdot\pi^1_{\mathcal{T}_h}(I\ge 1)+\hat{p}\right]\,,$$

where in the first inequality we used the fact that in order for the root to be connected to the left-most leaf, it is required that the root is connected to its left-most child w_1, and that w_1 is connected to the left-most leaf of its sub-tree. The former event occurs with probability p or $\hat{p}$, depending on whether or not the root is connected to $\partial\mathcal{T}_h$ through any child besides w_1.

By monotonicity, for every $i=2,\dots,\Delta$, the law of I_i under $\pi^1_{\mathcal{T}_h}$ is below its law under $\pi^{(1,\circlearrowleft)}_{\mathcal{T}_h}$ and the same holds for I. Since, by Lemma 6.3 a single external wiring may distort the distribution by at most a q^2 factor, we get $\pi^{(1,\circlearrowleft)}_{\mathcal{T}_h}(I_i=1)\le pq^2\varphi_h$ for all i. Hence, under $\pi^{(1,\circlearrowleft)}_{\mathcal{T}_h}$, I is stochastically below Q, where $Q\sim\mathrm{Bin}(d,pq^2\varphi_h)$. A union bound and Lemma 5.2 then imply

$$\pi^{(1,\circlearrowleft)}_{\mathcal{T}_h}(I\ge 1)\le\mathbb{P}(Q\ge 1)\le dpq^2(\hat{p}d)^{h-o(h)}\le C(\hat{p}d)^{(1-\varepsilon)h}\,,$$

for all h; note that ε can be chosen as small as needed provided the constant $C(p,q,\Delta,\varepsilon)$ is large enough. Thus, setting $a=Cpq^2\hat{p}^{-1}$ we obtain

$$\vartheta_h\le\hat{p}\vartheta_{h-1}\left[1+a(\hat{p}d)^{(1-\varepsilon)h}\right]\le\hat{p}^h\prod_{i=1}^{h}\left[1+a(\hat{p}d)^{(1-\varepsilon)i}\right]\,,$$

by continuing the recursion. Now, observe that since $\hat{p}d<1$ when $p<p_u$,

$$\prod_{i=1}^{h}\left[1+a(\hat{p}d)^{(1-\varepsilon)i}\right]=\exp\left[\sum_{i=1}^{h}\log\left(1+a(\hat{p}d)^{(1-\varepsilon)i}\right)\right]\le\exp\left[a\sum_{i=1}^{h}(\hat{p}d)^{(1-\varepsilon)i}\right]\le\exp\left[\frac{a}{(\hat{p}d)^{1-\varepsilon}}\right].$$

Combining the above two bounds, there exists an absolute constant $A=A(p,q,\Delta)$ such that for every h we have $\vartheta_h\le A\hat{p}^h$ and thus $\vartheta^{\circlearrowleft}_h\le Aq^2\hat{p}^h$. The first inequality in the lemma follows by noticing that all the leaves in $\mathcal{T}_h$ are equivalent, and the second follows from a union bound over the Δd^{h-1}. $\square$

5.2. **Exponential decay rate in (L,R)-Treelike graphs.** Let $G=(V,E)$ be an (L,R)-Treelike graph. For $v\in V$, let $B:=B_R(v)$ denote the ball of radius R around the vertex v. Recall that we use $\mathcal{N}_v\subseteq E$ for the set of edges incident to v. For each $1\le\ell\le R$, let $Q_\ell=\{u\in B: d(u,v)\ge\ell\}$.

For a boundary condition ξ on ∂B, recall the set $\mathfrak{V}_{B,\xi}$ of vertices in non-trivial boundary components of ξ from Definition 4.20. For any $u\in B$ such that $d(u,v)=\ell$, let $u\overset{Q_\ell}{\longleftrightarrow}\mathfrak{V}_{B,\xi}$ denote the event that u is connected to $\mathfrak{V}_{B,\xi}$ by a path of open edges fully contained in Q_ℓ: i.e.,

$$\{u\overset{Q_\ell}{\longleftrightarrow}\mathfrak{V}_{B,\xi}\}:=\{\omega:\mathcal{C}_u(\omega(Q_\ell))\cap\mathfrak{V}_{B,\xi}\ne\emptyset\}\,.$$

Define the event

$$\Upsilon_{B,\xi}:=\{\omega\in\{0,1\}^{E(B)}:|\{u\in B:d(u,v)=\ell\,,\,u\overset{Q_\ell}{\longleftrightarrow}\mathfrak{V}_{B,\xi}\}|\ge 2\text{ for all }1\le\ell\le R\}\,.$$

Notice that $\Upsilon_{B,\xi}$ is an increasing event. We claim that $\Upsilon_{B,\xi}$ controls the propagation of influence from ∂B.

Lemma 5.3. *Fix a graph $G = (V, E)$, a vertex $v \in V$ and consider the ball $B_R(v)$; let $\xi \geq \tau$ denote two boundary conditions on $\partial B_R(v) = \{w \in B_R(v) : d(v, w) = R\}$. Then,*

$$\|\pi^{\xi}_{B_R(v)}(\omega(\mathcal{N}_v) \in \cdot) - \pi^{\tau}_{B_R(v)}(\omega(\mathcal{N}_v) \in \cdot)\|_{\text{TV}} \leq \pi^{\xi}_{B_R(v)}(\Upsilon_{B_R(v),\xi}) \,.$$

Proof of Lemma 5.3. For ease of notation let $B := B_R(v)$, $\Upsilon_\xi = \Upsilon_{B,\xi}$ and $\mathfrak{V}_\xi = \mathfrak{V}_{B,\xi}$. We construct a monotone coupling $\mathbb{P}$ of $\omega^\xi \sim \pi^\xi_B$ and $\omega^\tau \sim \pi^\tau_B$. The coupling $\mathbb{P}$ reveals the configurations $\omega^\xi \sim \pi^\xi_B$ and $\omega^\tau \sim \pi^\tau_B$ on B one edge at a time using i.i.d. uniform random variables $U_e \in [0, 1]$ for each $e \in E(B)$. The same U_e is used to reveal the values $\omega^\xi(e)$ and $\omega^\tau(e)$ from the corresponding conditional measures. The order in which the uniform variables are revealed is irrelevant and can be adaptive; this will allow us to reveal the boundary components. (For more details on the process of revealing random-cluster components under the monotone coupling, see below, as well as e.g., [6, 8].)

We construct an adaptive revealing scheme that ensures that on the event Υ^c_ξ for the top sample ω^ξ, the samples ω^ξ and ω^τ agree on $\mathcal{N}_v$. This implies the desired result as one would then have by the definition of total-variation distance,

$$\|\pi^{\xi}_{B}(\omega(\mathcal{N}_v) \in \cdot) - \pi^{\tau}_{B}(\omega(\mathcal{N}_v) \in \cdot)\|_{\text{TV}} \leq \mathbb{P}(\omega^\xi(\mathcal{N}_v) \neq \omega^\tau(\mathcal{N}_v)) \leq \pi^{\xi}_{B}(\Upsilon_{B,\xi}) \,.$$

We construct $\mathbb{P}$ with the following iterative scheme which proceeds level-by-level from the leaves of B. Recall that for each $\ell \geq 1$, we let $Q_\ell = \{u \in B : d(u, v) \geq \ell\}$ and $E(Q_\ell)$ is the set of edges with both endpoints in Q_ℓ. At any time in the revealment process, we say that a vertex $u \in Q_\ell$ is unsaturated in Q_ℓ if there exists $w \in Q_\ell$ such that the edge-values $(\omega^\xi(uw), \omega^\tau(uw))$ have not been revealed. Let $(U_e)_{e\in E(B)}$ be a family of i.i.d. uniform random variables on $[0, 1]$ and reveal the configuration ω^ξ as follows:

Definition 5.4. Initialize $\mathcal{V}_\xi = \mathfrak{V}_\xi$ and $\mathcal{E}_\xi = \emptyset$;
for $i = 1, 2, ..., R$ **do**
  while $\exists u \in \mathcal{V}_\xi$ such that u is unsaturated in Q_{R-i}
    for each vertex $w \in Q_{R-i} :\ uw \in E(Q_{R-i})$
    1. Reveal $\omega^\xi(uw)$ from $\pi^\xi_B(\cdot \mid \omega(\mathcal{E}_\xi))$ using U_{uw}, i.e., set
$$\omega^\xi(uw) = \begin{cases} 1 & \text{if } \pi^\xi_B(\omega(uw) = 1 \mid \omega(\mathcal{E}_\xi)) \geq U_{uw} \\ 0 & \text{else} \end{cases} ;$$
    2. Add the edge uw to the set $\mathcal{E}_\xi$;
    3. If $\omega^\xi(uw) = 1$, add the vertex w to $\mathcal{V}_\xi$;

Note that we can use the same family $(U_e)_{e\in E(B)}$ in this process to generate coupled samples of ω^ξ and ω^τ. Notice that this coupling is monotone, so that because $\xi \geq \tau$, $\omega^\xi \geq \omega^\tau$ almost surely. Let $\mathcal{C}^i_{\mathfrak{V}}(\omega^\xi)$ denote the set of open edges revealed up to the i-th iteration of the procedure; we observe that $\mathcal{C}^i_{\mathfrak{V}}(\omega^\xi)$ is not necessarily equal to the intersection of $\mathcal{C}_{\mathfrak{V}}(\omega^\xi)$ with $E(Q_{R-i})$, but it is a subset of $\mathcal{C}_{\mathfrak{V}}(\omega^\xi) \cap E(Q_{R-i})$. Refer to Figure 5.1 for a depiction of the above revealing procedure.

Through this revealing process, we see that ω^ξ is open on the edges in the random set $\mathcal{C}^i_{\mathfrak{V}}(\omega^\xi)$ and free on the edges in its outer (edge) boundary in Q_{R-i}. Let $\bar{\mathcal{C}}^i_{\mathfrak{V}}(\omega^\xi)$ be the union $\mathcal{C}^i_{\mathfrak{V}}(\omega^\xi)$ with its outer (edge) boundary in Q_{R-i}, and note that this corresponds to the state of $\mathcal{E}_\xi$ after the i'th iteration. The random set $\mathcal{C}^i_{\mathfrak{V}}(\omega^\xi)$ is measurable with respect to the uniform random variables assigned to edges of $\bar{\mathcal{C}}^i_{\mathfrak{V}}(\omega^\xi)$.

For each $\mathcal{C}^i_{\mathfrak{V}}(\omega^\xi)$, let $\mathcal{V}^i(\omega^\xi)$ be the vertices in $\mathcal{C}^i_{\mathfrak{V}}(\omega^\xi)$ at distance exactly $R - i$ from v. Then,

$$\mathcal{V}^i(\omega^\xi) \subseteq \mathcal{C}_{\mathfrak{V}}(\omega^\xi) \cap \{w : d(w, v) = R - i\} \,.$$

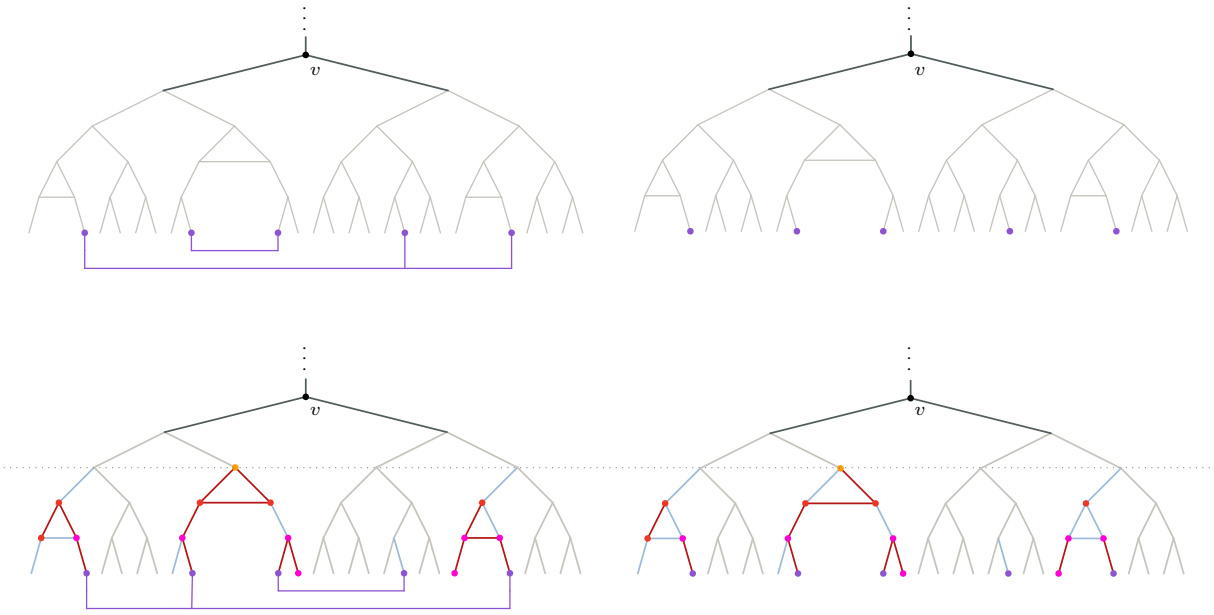

FIGURE 5.1. Top: The ball $B_R(v)$ for $R = 5$, with a K-sparse boundary condition τ for $K = 4$ (left), and the free boundary condition $\xi = 0$ (right). Bottom: The configurations revealed by the procedure of Definition 5.4, showing $\mathcal{C}^{i_0}_{\mathfrak{V}}(\omega^\tau)$ (red, left) along with its outer edge boundary in Q_{i_0} (light blue), revealing the dotted line (depth i_0) to be the largest i for which the set $\mathcal{V}_i$ is a singleton. The vertices that would have been exposed for larger values of i are colored in different colors. The coupled edge configuration $\omega^0(\mathcal{C}^{i_0}_{\mathfrak{V}}(\omega^\tau))$ is depicted on the right (open edges in red, closed edges in blue). The exposed configurations on $\mathcal{C}^{i_0}_{\mathfrak{V}}(\omega^\tau)$ induce free boundary conditions on $E(B) \setminus \bar{\mathcal{C}}^{i_0}_{\mathfrak{V}}(\omega^\tau)$.

On Υ^c_B, there must be some i for which $|\mathcal{C}_{\mathfrak{V}}(\omega^\xi) \cap \{w : d(w,v) = R-i\}| \le 1$, and therefore $|\mathcal{V}^i(\omega^\xi)| \le 1$. Let i_0 be the first i for which $|\mathcal{V}^{i_0}(\omega^\xi)| \le 1$, and for ease of notation set $\mathcal{V}_0 = \mathcal{V}^{i_0}(\omega^\xi)$, $\bar{\mathcal{C}}_{\mathfrak{V},0} = \bar{\mathcal{C}}^{i_0}_{\mathfrak{V}}(\omega^\xi)$ and set $\bar{\mathcal{C}}^c_{\mathfrak{V},0} = E(B) \setminus \bar{\mathcal{C}}_{\mathfrak{V},0}$. Notice the inclusion

$$\bar{\mathcal{C}}^i_{\mathfrak{V}}(\omega^\xi) \subset \bar{\mathcal{C}}^{i+1}_{\mathfrak{V}}(\omega^\xi)\,,$$

and from that deduce that i_0 is measurable with respect to the uniform random variables assigned to edges of $\bar{\mathcal{C}}^{i_0}_{\mathfrak{V}}(\omega^\xi)$. By the domain Markov property (see e.g., [31]), conditionally on $\omega^\xi(\bar{\mathcal{C}}_{\mathfrak{V},0})$, the configuration $\omega^\xi(\bar{\mathcal{C}}^c_{\mathfrak{V},0})$ (respectively, $\omega^\tau(\bar{\mathcal{C}}^c_{\mathfrak{V},0})$) is distributed according to the random-cluster distribution on $\bar{\mathcal{C}}^c_{\mathfrak{V},0}$ with boundary conditions induced by ξ and $\omega^\xi(\bar{\mathcal{C}}_{\mathfrak{V},0})$, respectively τ and $\omega^\tau(\bar{\mathcal{C}}_{\mathfrak{V},0})$.

To conclude the proof, it suffices to see that because $|\mathcal{V}_0| \le 1$, both $\omega^\xi(\bar{\mathcal{C}}_{\mathfrak{V},0})$ and $\omega^\tau(\bar{\mathcal{C}}_{\mathfrak{V},0})$ induce the free boundary conditions on $\bar{\mathcal{C}}^c_{\mathfrak{V},0}$. In that case ω^ξ and ω^τ would agree on $\bar{\mathcal{C}}^c_{\mathfrak{V},0}$ and in particular on $\mathcal{N}_v$. By monotonicity, it suffices for us to show that the boundary conditions induced by ξ and $\omega^\xi(\bar{\mathcal{C}}_{\mathfrak{V},0})$ on $\bar{\mathcal{C}}^c_{\mathfrak{V},0}$ are free. Since the wirings of ξ are only on vertices of $\mathfrak{V}_\xi \subset \bar{\mathcal{C}}_{\mathfrak{V},0}$, the only way for the boundary conditions on $\bar{\mathcal{C}}^c_{\mathfrak{V},0}$ to be not free is if multiple vertices on its boundary are incident to open edges of $\omega^\xi(\bar{\mathcal{C}}_{\mathfrak{V},0})$. By construction, the only vertices in $\bar{\mathcal{C}}_{\mathfrak{V},0}$ which can be incident to an open edge of $\omega^\xi(\bar{\mathcal{C}}_{\mathfrak{V},0})$ must be at distance exactly $R - i_0$ from v. By the assumption that $|\mathcal{V}_0| \le 1$, there can be at most one such vertex, and therefore there are no non-trivial (i.e., non-singleton) boundary components induced on $\bar{\mathcal{C}}^c_{\mathfrak{V},0}$ by the boundary condition $(\xi, \omega^\xi(\bar{\mathcal{C}}_{\mathfrak{V},0}))$, implying the desired conclusion. □

Proof of Proposition 3.3. With Lemma 5.3 in hand, it suffices for us to prove the following: there exists $C(p,q,K,L) > 0$ such that if $G = (V,E)$ is an (L,R)-Treelike graph and ξ is a K-Sparse boundary condition for the L-Treelike ball $B := B_R(v)$ about some $v \in V$, we have

$$\pi_B^{\xi}(\Upsilon_{B,\xi}) \le C\hat{p}^{2R}. \tag{5.3}$$

Let $H \subset E(B)$ be a set of at most L edges such that the subgraph $(V, E(B) \setminus H)$ is a tree; the existence of such a set is guaranteed by the fact that $B_R(v)$ is L-Treelike. Let $\mathcal{Z} = \{d_1, ..., d_k\}$ be the subset of distances (from v) which H intersects, i.e., $\mathcal{Z} = \{1 \le \ell < R : \exists w \in V(H) : d(w,v) = \ell\}$. See Figure 5.2 for a depiction. Observe that each edge of H intersects either one or two consecutive depths in $\mathcal{Z}$. Since B Is L-Treelike, we clearly have $|\mathcal{Z}| \le 2L$. Letting $d_0 = 0$ and $d_{k+1} = R$, for $i = 0, \dots, k$ we define:

$$\mathcal{F}_i := \{u \in B : d_i < d(u,v) < d_{i+1}\}.$$

For each $0 \le i \le k$, the graph $\mathcal{F}_i = (\mathcal{F}_i, E(\mathcal{F}_i))$ is a forest. For each i, let $\mathcal{T}_{ij} = (\mathcal{T}_{ij}, E(\mathcal{T}_{ij}))$ for $j = 0, 1, \dots$ denote the distinct connected components (subtrees) of $\mathcal{F}_i$ so that $\mathcal{F}_i = \bigcup_{j \ge 0} \mathcal{T}_{ij}$. (For some i, this may be empty, and for other i, this may be a single vertex.)

Now, in order for $\Upsilon_{B,\xi}$ to hold, it must be the case that in each $\mathcal{F}_i$, every depth ℓ is intersected by at least two sites in the FK cluster of $\mathfrak{V}_{B,\xi}$ in Q_ℓ. Specifically, for each i, at distance $d_i + 1$ from v there must be at least two distinct vertices connected to $\mathfrak{V}_{B,\xi}$ with paths in Q_{d_i+1}. Thus, for each i there must exist two open monotone paths (each intersecting each height in $\mathcal{F}_i$ at exactly one vertex), $\gamma_i \subset E(\mathcal{T}_{ij})$ and $\gamma_i' \subset E(\mathcal{T}_{ij'})$ with $j \ne j'$ such that γ_i (resp., γ_i') connects the root of $\mathcal{T}_{ij}$ (resp., $\mathcal{T}_{ij'}$) to one of its leaves. If there are multiple such paths, choose according to some predetermined ordering, and call the sequences of paths $\Gamma = \gamma_0, \dots, \gamma_k$ and $\Gamma' = \gamma_0', \dots, \gamma_k'$. See Figure 5.3 for a depiction.

We enumerate over the choices of such sequences of paths and then show that for any two fixed sequences of paths, the probability that they are both open is bounded by $C\hat{p}^{2R}$ for some $C(p,q,\Delta,K,L)$. (We say that a sequence of paths is open if all of its paths are.)

In order to enumerate over the choices of sequences of paths, for each monotone path γ_i, let x_i be its bottom endpoint, and define x_i' for γ_i' similarly. Since ξ is K-Sparse, there are evidently at most K many choices of x_0, and K choices of x_0'. Now observe that since γ_i is a monotone path on a tree, for each i, the bottom endpoint x_i determines the entire path γ_i. Since these paths form parts of the connections to $\mathfrak{V}_{B,\xi}$ the sequence of paths can be required to have endpoints at depths $d_{i+1} - 1$ that are either an ancestor of x_0, or an ancestor of $V(H)$. Here, at each height $h \notin \mathcal{Z}$ an *ancestor* of a vertex u at height h is a vertex along the geodesic from v to u. We make the following observation.

Claim 5.5. *If $B_R(v)$ is L-Treelike, if u is such that $d(u,v) = h$, for every $h' < h$, u has at most 2^L many ancestors at height h'.*

Indeed, except along the edges in H, every vertex has a unique *parent* which is an ancestor of that vertex at one smaller depth. Thus, the geodesics of B are uniquely determined by their endpoints together, possibly, with a subset of edges of H traversed along the geodesic, yielding the at most 2^L available choices.

Returning to the enumeration over Γ, Γ', the heights of the endpoints x_i, x_i' are predetermined by i, and therefore, having chosen x_0, x_0' for each i, there are at most $2L$ many choices of bottom end-point x_i, and likewise of x_i', and therefore at most $2L \cdot 2^L$ many choices of γ_i and γ_i'.

Hence, a union bound implies

$$\pi_B^{\xi}(\Upsilon_{B,\xi}) \le K^2(2L)^{2L}(2^L)^{2L} \sup_{\Gamma,\Gamma'} \pi_B^{\xi}(\omega(\Gamma \cup \Gamma') = 1). \tag{5.4}$$

Now fix any two such sequences of paths Γ, Γ', and consider the probability that $\omega(\Gamma \cup \Gamma') = 1$. Observe that Γ and Γ' are vertex-disjoint by construction. Our aim is to make the events that Γ and Γ' are open in ω independent. For this, let ρ_i be the set of roots of the trees in $\mathcal{F}_i$. We introduce auxiliary wirings (as shown in Figures 5.2–5.3) for all vertices at depths $\{d : \min_{i=0,\dots,k+1} |d - d_i| \le 1\}$. Call the resulting distribution

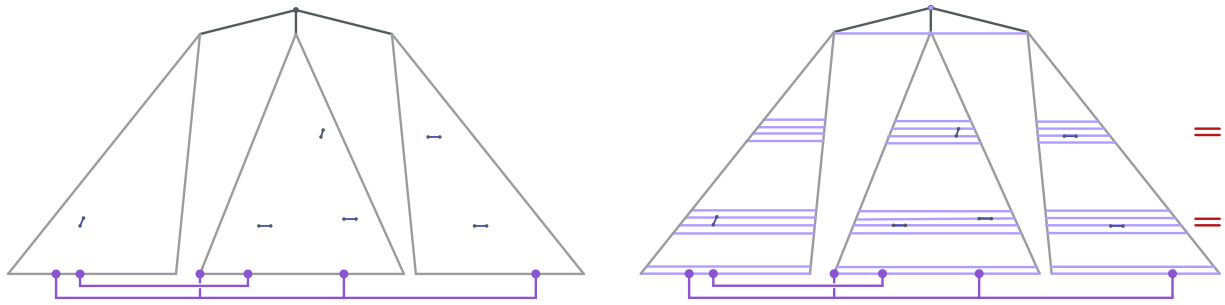

FIGURE 5.2. Left: An L-Treelike ball with K-Sparse boundary conditions is depicted for $L = K = 6$: the L edges that need to be removed to leave a tree are indicated in blue. Right: We modify the boundary conditions to be all wired (the wired component is depicted in purple) at or one away from heights in $\mathcal{Z}$ (marked by red dashes).

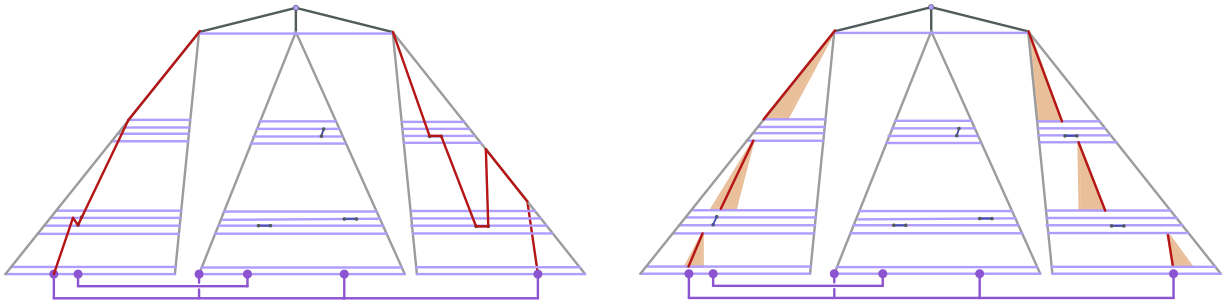

FIGURE 5.3. Left: Two disjoint components (red) of the vertices in $\mathfrak{V}_{B,\xi}$, together intersect every depth in the ball and satisfy the event $\Upsilon_{B,\xi}$. Right: The two components contain corresponding sequences of open leaf-to-root paths (red) in independent wired subtrees (shaded, orange) whose endpoints are amongst the ancestors of vertices of H or $\mathfrak{V}_{B,\xi}$.

$\tilde{\pi}_B$; by monotonicity,

$$\pi_B^{\xi}(\omega(\Gamma \cup \Gamma') = 1) \leq \tilde{\pi}_B(\omega(\Gamma \cup \Gamma') = 1)\,. \tag{5.5}$$

The distribution $\tilde{\pi}_B$ is a product measure over the $\mathcal{T}_{ij}$'s with boundary condition $(1, \circlearrowleft)$ in each $\mathcal{T}_{ij}$ (recall that this boundary condition wires all leaves $\partial\mathcal{T}_{ij}$ together with the root of $\mathcal{T}_{ij}$). Hence, since Γ and Γ' are such that, for each $i \geq 0$, γ_i and γ_i' belong to distinct subtrees $\mathcal{T}_{ij_i}$, $\mathcal{T}_{ij_i'}$ of the forest $\mathcal{F}_i$, and we have

$$\tilde{\pi}_B(\omega(\Gamma \cup \Gamma') = 1) = \prod_{i=0}^{k} \pi_{\mathcal{T}_{ij_i}}^{(1,\circlearrowleft)}(\gamma_i) \prod_{i=0}^{k} \pi_{\mathcal{T}_{ij_i'}}^{(1,\circlearrowleft)}(\gamma_i')\,.$$

Let $h_i = d_{i+1} - d_i$ be the height of the trees in $\mathcal{F}_i$. We deduce from Lemma 5.1 that there exists a constant $A(p, q, \Delta) > 0$ such that uniformly over Γ, Γ',

$$\tilde{\pi}_B(\omega(\Gamma \cup \Gamma') = 1) \leq A^{2L} \prod_{i=0}^{k} \hat{p}^{2h_i} \leq A^{2L} \hat{p}^{2(R-4L)}\,.$$

Plugging this bound into (5.4)–(5.5), we obtain

$$\pi_B^{\xi}(\Upsilon_{B,\xi}) \leq K^2((2L)(2^L)\hat{p}^{-4}A)^{2L}\hat{p}^{2R}\,,$$

from which the required (5.3) follows. □

Remark 5.6. A matching lower bound of $\Omega(\hat{p}^{2R})$ for the decay rate in Proposition 3.3 is easy to construct by e.g., taking the K-Sparse boundary conditions ξ that wires two leaves w_1, w_2 on distinct sub-trees of v, and the free boundary conditions $\xi' = 0$ on $\mathcal{T}_R$. The event that the root is connected to w_1 and its corresponding child is connected to w_2 has probability at least $C\hat{p}^{2R}$ by Lemma 5.1 and the FKG inequality (see e.g., [31]). On this event, the probability that the edge incident v down towards w_2 is open is p under the boundary condition ξ and $\hat{p}$ under $\xi' = 0$.

6. Proof of fast mixing

In this section, we combine the results of Sections 4–5 to conclude the proof of Theorem 1.1. As indicated in Section 3, the analysis of Sections 4–5 reduce the mixing time of the FK-dynamics on a random graph to understanding the convergence to equilibrium on $O(1)$-Treelike balls of volume $O(n^{\frac{1}{2}-\delta})$ with $O(1)$-Sparse boundary conditions. In Section 6.1, we recall the log-Sobolev inequality and comparison bounds for the log-Sobolev constant under different boundary conditions. In Section 6.2, we bound this log-Sobolev constant via straightforward comparison to a product chain. Then in Section 6.3, we proceed to combine all of the above ingredients to deduce the proof of Theorem 1.1 using the censoring inequalities of [46].

6.1. Mixing time preliminaries.

Let us recall some standard tools to help us bound the rate of convergence to equilibrium of the FK-dynamics on treelike balls with sparse boundary conditions.

Log-Sobolev inequalities. Recall, for a Markov chain with transition matrix P, the Dirichlet form

$$\mathcal{E}(f,f) := \frac{1}{2}\sum_{\omega,\omega'\in\{0,1\}^E} \pi(\omega)P(\omega,\omega')(f(\omega)-f(\omega'))^2\,, \tag{6.1}$$

for $f:\Omega\to\mathbb{R}$. Then the *log-Sobolev constant* is given by

$$\alpha(P) := \min_{f:\mathrm{Ent}_\pi[f^2]\neq 0} \frac{\mathcal{E}(f,f)}{\mathrm{Ent}_\pi[f^2]}\,, \qquad \text{where} \qquad \mathrm{Ent}_\pi[f^2] = \mathbb{E}_\pi\Big[f^2\log\frac{f^2}{\mathbb{E}_\pi[f^2]}\Big]\,. \tag{6.2}$$

As such, a *log-Sobolev inequality* takes the form $\mathcal{E}(f,f)\geq\gamma\cdot\mathrm{Ent}_\pi[f^2]$ for all f. A log-Sobolev inequality is stronger than a mixing time bound, in the sense that it implies exponential convergence with rate γ in total-variation distance from the stationary distribution. This is captured by the following standard fact (e.g., a proof in the discrete time setting we consider follows immediately from Lemma 2.8 and Eq. (2.10) of [3]).

Fact 6.1. *Consider an ergodic finite state Markov chain $(X_t)_{t\geq 0}$ with transition matrix P reversible with respect to stationary measure π. If the chain has a log-Sobolev constant $\alpha=\alpha(P)$, then for every $\gamma<\alpha$,*

$$\max_{x_0\in\Omega}\|P(X_t^{x_0}\in\cdot)-\pi\|_{\mathrm{TV}} \leq e^{-\gamma t/2}\Big(\log\frac{1}{\min_{x\in\Omega}\pi(x)}\Big)^{1/2}\,.$$

Boundary condition comparisons for the FK-dynamics. The following formalizes the notion that sparse boundary conditions are "close to free", and allows us to compare the induced mixing time on balls with sparse boundary to those with free boundary.

Definition 6.2 (Definition 2.1 of [6])**.** For two boundary conditions (partitions) $\phi\leq\phi'$, define $D(\phi,\phi') := c(\phi)-c(\phi')$ where $c(\phi)$ is the number of components in ϕ. For two partitions ϕ,ϕ' that are not comparable, let ϕ'' be the smallest partition such that $\phi''\geq\phi$ and $\phi''\geq\phi'$ and set $D(\phi,\phi') = c(\phi)-c(\phi'')+c(\phi')-c(\phi'')$.

Lemma 6.3 (Lemma 2.2 of [6])**.** *Let $G=(V,E)$ be an arbitrary graph, $p\in(0,1)$ and $q>0$. Let ϕ and ϕ' be two partitions of V encoding two distinct external wirings on the vertices of G. Let π_G^{ϕ}, $\pi_G^{\phi'}$ be the resulting random-cluster measures. Then, for all FK configurations $\omega\in\{0,1\}^E$, we have*

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

From Lemma 6.3, and the definition of the Dirichlet form, (6.1), we deduce the following.

Corollary 6.4. *Let $G = (V, E)$ be an arbitrary graph, $p \in (0, 1)$ and $q > 0$. Consider the FK-dynamics on G with boundary conditions ϕ and ϕ', and let $\mathcal{E}_{G^\phi}$, $\mathcal{E}_{G^{\phi'}}$ denote their Dirichlet forms, respectively. Then*

$$q^{-4D(\phi,\phi'))}\mathcal{E}_{G^{\phi'}}(f,f) \le \mathcal{E}_{G^\phi}(f,f) \le q^{4D(\phi,\phi'))}\mathcal{E}_{G^{\phi'}}(f,f)\,, \qquad \text{for all } f : \{0,1\}^E \to \mathbb{R}\,.$$

Together with Corollary 6.4 and Lemma 6.3 again, this controls the change in both *log-Sobolev constant* (6.2), and mixing time, under two boundary conditions with distance $D(\phi, \phi')$.

6.2. **Local mixing: fast mixing on treelike graphs with sparse boundary conditions.** In this section we establish a bound for the speed of convergence of the FK-dynamics on L-Treelike balls with K-Sparse boundary conditions (see Definitions 2.1 and 4.1). Our goal is to prove the following lemma.

Lemma 6.5. *Suppose $B_R(v)$ is L-Treelike. Let ξ be a K-Sparse boundary condition on $\partial B_R(v)$. For every $p \in (0, 1)$ and $q > 0$, the log-Sobolev constant of the FK-dynamics on $B_R(v)$ with boundary condition ξ is $\Omega(|E(B_R(v))|^{-1}) = \Omega(d^{-R})$.*

Lemma 6.5 follows by comparing log-Sobolev on an L-Treelike ball with K-Sparse boundary to a tree with K-Sparse boundary conditions, whose log-Sobolev constant is bounded by comparison to a product chain. We first note the following bound on the log-Sobolev constant on trees with sparse boundaries.

Corollary 6.6. *There exists $c(p, q) > 0$ such that the following holds. For every rooted (not necessarily complete) tree $\hat{\mathcal{T}}_h = (V(\hat{\mathcal{T}}_h), E(\hat{\mathcal{T}}_h))$ of depth h and degree at most Δ, and every K-Sparse boundary condition ϕ on $\hat{\mathcal{T}}_h$, the log-Sobolev constant of the FK-dynamics on $\hat{\mathcal{T}}_h$ with boundary conditions ϕ is at least $cq^{6K}|E(\hat{\mathcal{T}}_h)|^{-1}$.*

Proof. Consider the FK-dynamics on $\hat{\mathcal{T}}_h$ under the free boundary conditions. In this case, the random-cluster measure is a $\mathrm{Ber}(\hat{p})$ product measure and thus the log-Sobolev constant of the FK-dynamics is $c|E(\hat{\mathcal{T}}_h)|^{-1}$ for some $c(p, q) > 0$; see, e.g., [17]. The result then follows from Lemma 6.3 and Corollary 6.4. □

To move from mixing on an L-Treelike ball to mixing on a tree, the following fact will be useful.

Fact 6.7. *Let G be a subgraph of G' such that $V(G) = V(G')$ and $E(G) \subset E(G')$; let $H = E(G') \setminus E(G)$. Suppose ϕ is a boundary condition on G, G' such that for every $e \in H$, the endpoints of e are wired in ϕ. For every $p \in (0, 1)$ and $q > 0$, let P_G and $P_{G'}$ be the transition matrices of the FK-dynamics on G and G', respectively, with boundary conditions ϕ, and let $\alpha(P_G)$ and $\alpha(P_{G'})$ be their log-Sobolev constants. There exists a constant $c(p) > 0$ such that*

$$\alpha(P_{G'}) \ge \min\left\{\frac{|E(G)|}{|E(G)| + |H|} \cdot \alpha(P_G), \frac{c|H|}{|E(G)| + |H|}\right\}.$$

Proof. The FK-dynamics on G' is a product Markov chain on $\{0,1\}^{E(G)} \times \{0,1\}^H$ with stationary distribution $\pi^\phi_{G'} = \pi^\phi_G \otimes \prod_{i=1}^{|H|} \nu_i$, where $(\nu_i)_{1 \le i \le |H|}$ are independent $\mathrm{Ber}(p)$ distributions over edges in H. The result then follows from the tensorization of the log-Sobolev inequality (e.g., [47, Lemma 2.2.11]). □

We can now combine the above ingredients to deduce the bound of Lemma 6.5.

Proof of Lemma 6.5. Let $B = B_R(v)$ and let $H \subset E(B)$ be a set of at most L edges such that $(B, E(B) \setminus H)$ is a tree. Consider the tree $\hat{\mathcal{T}}_R = (V(B), E(B) \setminus H)$ and let ϕ be the boundary condition that includes all the connections from ξ and adds wirings between w and w' for every edge $ww' \in H$.

Corollary 6.6 implies that the log-Sobolev constant for the FK-dynamics on $\hat{\mathcal{T}}_R$ with boundary condition ϕ is at least $cq^{6(K+L)}|E(\hat{\mathcal{T}}_R)|^{-1}$ for some $c(p, q) > 0$. We then get from Fact 6.7 that the log-Sobolev constant for the FK-dynamics on B with boundary condition ϕ is at least $cq^{6(k+L)}|E(B)|^{-1}$. Lemma 6.3 and Corollary 6.4 then imply that the log-Sobolev constant on B with boundary conditions ξ is at least $cq^{6K+12L}|E(B)|^{-1}$. □

6.3. **Proof of Theorem 1.1: upper bound.** Fix $p < p_u(q,\Delta)$, let $\varepsilon = 1 - \hat{p}d$ (positive when $\hat{p} < p_u$) and fix $\delta > 0$ small enough (depending on ε, Δ) such that

$$2\delta + (1-2\delta)\log_d(1-\varepsilon) < 0\,,$$

in which case the following is polynomially decaying in n:

$$n\hat{p}^{(1-2\delta)\log_d n} = nd^{(-1+2\delta)\log_d n}(1-\varepsilon)^{(1-2\delta)\log_d n} = n^{2\delta}(1-\varepsilon)^{(1-2\delta)\log_d n}\,. \tag{6.3}$$

Let $R = (\frac{1}{2}-\delta)\log_d n$ and let K be a constant sufficiently large (depending on p,q,Δ) that both Fact 2.3 and Theorem 4.3 hold for (K,R). For each t, let Γ_t be the set of Δ-regular graphs on n vertices having

$$\Gamma_t = \{\mathcal{G} : \mathcal{G} \text{ is } (K,R)\text{-Treelike}\} \cap \{\mathcal{G} : P(X^1_{\mathcal{G},t} \text{ is } (K,R)\text{-Sparse}) \geq 1 - n^{-2}\}\,.$$

By Fact 2.3 and Theorem 4.3, there exists $C_0(p,q,\Delta)$ such that if $T = C_0 n\log n$, then $\mathbf{P}_{\text{RRG}}(\Gamma_T^c) \leq o(1)$. It suffices for us to prove that the mixing time of the FK-dynamics on any $\mathcal{G} \in \Gamma_T$ is $O(n\log n)$.

Fix any $\mathcal{G} \in \Gamma_T$ and for every configuration ω on $E(\mathcal{G})$, let $X^\omega_t = X^\omega_{\mathcal{G},t}$ be the FK-dynamics chain on $\mathcal{G}$ initialized from $X^\omega_0 = \omega$. Couple the family of chains $((X^\omega_t)_{t\geq 0})_{\omega\in\{0,1\}^{E(\mathcal{G})}}$ using the grand coupling as in Definition 4.8: recall that this is the coupling that in each step picks the same random $e \in E(\mathcal{G})$ to update, and the same uniform random variable $U_{e,t}$ on $[0,1]$ to decide the next state on the edge e. As mentioned earlier, this coupling is monotone when $q > 1$ so that for every $t \geq 0$, if $X^\omega_t \leq X^{\omega'}_t$, then $X^\omega_{t+1} \leq X^{\omega'}_{t+1}$.

It follows from the definition of t_{MIX} and monotonicity of the grand coupling (see e.g., [40]), that it suffices for us to show that there exists $\hat{C}(p,q,\Delta)$ such that if $\hat{T} = T + \hat{C}n\log n$,

$$\mathbb{P}\big(X^1_{\hat{T}} \neq X^0_{\hat{T}}\big) \leq \frac{1}{4}\,.$$

By a union bound over the n edge-neighborhoods $\mathcal{N}_v$ (edges of $\mathcal{G}$ incident v), this reduces to showing

$$\sup_{v\in V(\mathcal{G})} \mathbb{P}\big(X^1_{\hat{T}}(\mathcal{N}_v) \neq X^0_{\hat{T}}(\mathcal{N}_v)\big) \leq \frac{1}{4n}\,. \tag{6.4}$$

Now fix any such v and consider the probability above. For ease of notation, let $B_v = B_R(v)$ and $B_v^c = E(\mathcal{G})\setminus B_v$. Introduce two new Markov chains Y^1_t and Y^0_t that are coupled via the grand coupling to X^1_t, X^0_t except that they censor (ignore) all updates on edges of B_v^c after time $T = C_0 n\log n$. The censoring inequality [46, Lemma 2.3] implies the stochastic relations $Y^1_t \succcurlyeq X^1_t$ and $Y^0_t \preccurlyeq X^0_t$ for all $t \geq 0$ and thus

$$\mathbb{P}\big(X^1_t(\mathcal{N}_v) \neq X^0_t(\mathcal{N}_v)\big) \leq \Delta \sup_{e\in\mathcal{N}_v} \mathbb{P}\big(X^1_t(e) \neq X^0_t(e)\big) \leq \Delta \sup_{e\in\mathcal{N}_v} \big[\mathbb{P}\big(Y^1_t(e)=1\big) - \mathbb{P}\big(Y^0_t(e)=1\big)\big]\,.$$

Fix any $e \in \mathcal{N}_v$ and consider the difference in probabilities on the right-hand side. Let $\mathcal{E}_T$ be the event (measurable with respect to the first T steps of the Markov chain) that the boundary conditions induced by $X^1_T(B_v^c)$ are K-Sparse. Observe that K-sparsity of a boundary condition is a decreasing event, so that on $\mathcal{E}_T$, the boundary conditions induced by $X^0_T(B_v^c)$ are also K-Sparse. As such, for all $t \geq T$,

$$\begin{aligned}
&\mathbb{P}(Y^1_t(e)=1) - \mathbb{P}(Y^0_t(e)=1) \qquad\qquad (6.5)\\
&\quad\leq \mathbb{P}(\mathcal{E}_T^c) + \sup_{\substack{\phi^0,\phi^1\in\{0,1\}^{B_v^c}\\ \phi^1 \text{ is } K-\text{Sparse}\,;\,\phi^0\leq\phi^1}} \mathbb{P}(Y^1_t(e)=1 \mid Y^1_T(B_v^c)=\phi^1) - \mathbb{P}(Y^0_t(e)=1 \mid Y^0_T(B_v^c)=\phi^0)\,.
\end{aligned}$$

Since $\mathcal{G} \in \Gamma_T$, and $Y^1_T = X^1_T$, the first term is at most n^{-2}. Turning to the second term, fix any two configurations ϕ^1, ϕ^0 on B_v^c such that $\phi^0 \leq \phi^1$ and ϕ^1 (and therefore also ϕ^0) induce K-Sparse boundary conditions on B_v, and consider the difference

$$\begin{aligned}
&\mathbb{P}(Y^1_{T+s}(e)=1 \mid Y^1_T(B_v^c)=\phi^1) - \mathbb{P}(Y^0_{T+s}(e)=1 \mid Y^0_T(B_v^c)=\phi^0)\\
&\qquad\leq |\mathbb{P}(Y^1_{T+s}(e)=1 \mid Y^1_T(B_v^c)=\phi^1) - \pi_{\mathcal{G}}(\omega(e)=1 \mid \omega(B_v^c)=\phi^1)| \qquad (6.6)\\
&\qquad\quad + |\pi_{\mathcal{G}}(\omega(e)=1 \mid \omega(B_v^c)=\phi^1) - \pi_{\mathcal{G}}(\omega(e)=1 \mid \omega(B_v^c)=\phi^0)| \qquad (6.7)\\
&\qquad\quad + |\mathbb{P}(Y^0_{T+s}(e)=1 \mid Y^0_T(B_v^c)=\phi^0) - \pi_{\mathcal{G}}(\omega(e)=1 \mid \omega(B_v^c)=\phi^0)|\,. \qquad (6.8)
\end{aligned}$$

Observe that $Y^1_{T+s}(B_v)$ is distributed as a *lazy* FK-dynamics chain Z^1_s on B_v with boundary conditions induced by ϕ^1, initialized from the random configuration $Z^1_0(B_v) = Y^1_T(B_v)$: the laziness is in the choice that at each step, Z^1_s makes an FK-dynamics update on B_v with probability $|E(B_v)|/|E(\mathcal{G})|$ and makes no update otherwise. The analogous statement holds for $Y^0_{T+s}(B_v)$ with respect to some lazy chain Z^0_s. The invariant measure of Z^1_s is easily seen to be

$$\pi_{\mathcal{G}}(\omega(B_v) \in \cdot \mid \omega(B_v^c) = \phi^1) = \pi^{\phi^1}_{B_v}\,,$$

and the analogous statement holds for Z^0_s.

Now let $\hat{T} = T + \hat{S}$ where $\hat{S} = \hat{C} n \log n$ for a constant $\hat{C}$ to be chosen sufficiently large depending on p, q, Δ. The expected number of updates in B_v between time T and $T + \hat{S}$ is

$$\hat{C} n \log n \cdot \frac{|E(B_v)|}{|E(\mathcal{G})|} \geq 2\Delta^{-1}\hat{C}|E(B_v)| \log n\,.$$

Let $C_1(p, q, \Delta, K)$ be a constant such that the $\Omega(d^{-R})$ bound on the log-Sobolev constant guaranteed by Lemma 6.5 (with the choice of $L = K$) is at least $C_1^{-1}d^{-R}$. For any C_2, if $\hat{C}$ is sufficiently large, a Chernoff bound (namely (4.6) applied to $\mathrm{Bin}(\hat{S}, |E(B_v)|/|E(\mathcal{G})|)$) implies that with probability $1 - o(n^{-2})$, at least $C_1C_2|E(B_v)| \log n$ updates are made in $E(B_v)$ between times T and $\hat{T}$. By K-sparsity of ϕ^1, Lemma 6.5, and Fact 6.1, the term in (6.6) is bounded by

$$\begin{aligned}\|\mathbb{P}(Z^1_{\hat{S}}(B_v) \in \cdot) - \pi^{\phi^1}_{B_v}\|_{\mathrm{TV}} &\leq \frac{1}{\sqrt{\log \min_\omega \pi(\omega)}} \exp\Big(-\frac{C_1C_2|E(B_v)| \log n}{2C_3|E(B_v)|}\Big) + o(n^{-2}) \\ &\leq O(\sqrt{n}) \cdot e^{-C_1C_2 \log n/(2C_3)} + o(n^{-2})\,,\end{aligned}$$

for some $C_3(p, q, \Delta)$; we thus have for C_2 sufficiently large (and therefore $\hat{C}$ sufficiently large), that this is at most $o(n^{-2})$. By the same reasoning, by K-sparsity of ϕ^0, the same bound applies to (6.8).

Finally, since both ϕ^1 and ϕ^0 induce K-Sparse boundary conditions on B_v, by Proposition 3.3 there exists a constant $C(p, q, \Delta, K) > 0$ such that (6.7) is at most

$$\|\pi^{\phi^1}_{B_v}(\omega(\mathcal{N}_v) \in \cdot) - \pi^{\phi^0}_{B_v}(\omega(\mathcal{N}_v) \in \cdot)\|_{\mathrm{TV}} \leq C\hat{p}^{2R}\,,$$

which is $o(n^{-1})$ by our choice of δ and (6.3). Putting these three bounds together we see that as long as $\hat{C}$ is sufficiently large (depending on p, q, Δ) the difference in (6.5) is $o(n^{-1})$, from which the bound of (6.4) follows for n sufficiently large, concluding the proof. □

7. Matching lower bound on the mixing time

In this section, we show a matching $\Omega(n \log n)$ lower bound on the mixing time of the FK-dynamics on a random Δ-regular graph and thus complete the proof of Theorem 1.1 from the introduction. A general lower bound for the mixing time of the Glauber dynamics on spin systems was show in [35]. However, the non-locality of the FK-dynamics complicates extending the ideas from [35] to the random-cluster setting. In [8], the argument from [35] was adapted to the random-cluster model on $\mathbb{Z}^2$ when $p \neq p_c(q)$, but the amenability of $\mathbb{Z}^2$ together with the exponential decay of connectivities at $p < p_c$ was key to this extension.

In our setting, the non-amenability of the random Δ-regular graph prevents us from bounding the speed of disagreement percolation under couplings of the FK-dynamics and implementing the argument of [35] directly. Instead, we use the locally treelike structure of the random Δ-regular graph to directly couple a projection of the model on a certain set of n^ε edges to a product measure on n^ε edges, for which the coupon collector problem gives an immediate lower bound.

Claim 7.1. *With $\mathbf{P}_{\mathrm{RRG}}$-probability $1 - o(1)$, $\mathcal{G}$ is such that there exist $n^{1/5}$ vertices whose balls of radius $\frac{1}{5}\log_d n$ are disjoint, and are trees.*

Proof. Per (4.2), it suffices to prove the above under $\mathbf{P}_{\mathrm{CM}}$. We prove the claim by repeated application of Lemma 4.16. Namely, consider the procedure where we repeatedly take an arbitrary vertex v that has not been discovered yet, and reveal its ball of radius R. Let v_i be the i'th vertex to be selected in this procedure, and let $\mathcal{A}_i$ be $\bigcup_{j\le i} E(B_R(v_j))$. Then, for integer $m \le n$ the probability that $(B_R(v_1),...,B_R(v_m))$ are disjoint trees, is at least

$$1 - \sum_{i=1}^{m} \mathbf{P}_{\mathrm{CM}}\big(B_R(v_i)\cap \mathcal{A}_{i-1} = \emptyset \text{ or } B_R(v_i) \text{ is not a tree} \mid \mathcal{A}_{i-1}\big)\,.$$

By Lemma 4.16 (using the fact that each $v_i \notin V(\mathcal{A}_{i-1})$ so that $B_R^{out}(v_i) = B_R(v_i)$) each of the summands is at most $O(md^{2R}/(n - O(md^R)))$. Taking $R = \delta \log_d n$ and $m = n^\delta$, we see that the sum above is at most $O(n^{4\delta}/(n - O(n^{2\delta}))$ which is $o(1)$ as long as $\delta < \frac{1}{4}$. □

Fix $\varepsilon \in (0, 1/5)$ to be taken sufficiently small later. For every $\mathcal{G}$ having $n^{1/5}$ many vertices whose balls of radius $\frac{1}{5}\log_d n$ are disjoint trees, choose arbitrarily some n^ε vertices amongst the $n^{1/5}$ of Claim 7.1, and for each vertex collect a representative edge incident to it to form the set $\mathcal{C} = \mathcal{C}_\varepsilon(\mathcal{G})$. Our proof will rely on a coupling of the restrictions of $X_{t,\mathcal{G}}$ and $\pi_{\mathcal{G}}$ to $\mathcal{C}$ to $\mathrm{Ber}(\hat{p})$ product chains. For this, let:

(1) $X_t = X_{t,\mathcal{G}}$ be a realization of the FK-dynamics;
(2) $Y_t = Y_{t,\mathcal{G}}$ be a realization of the FK-dynamics that censors all updates in $E(\mathcal{G}) \setminus \mathcal{C}$;
(3) ν as the product measure over $|\mathcal{C}|$ many $\mathrm{Ber}(\hat{p})$ random variables.

As before, let Y_t^0 be the chain Y_t initialized from the all-0 configuration.

Lemma 7.2. *Let $\mathcal{G}$ be a graph with at least $n^{1/5}$ vertices whose balls of radius $\frac{1}{5}\log_d n$ are disjoint trees. For every $q > 1$, integer $\Delta \ge 3$, and $p < p_u(q,\Delta)$, there exists $\varepsilon > 0$ sufficiently small such that we have the following for $\mathcal{C} = \mathcal{C}_\varepsilon(\mathcal{G})$:*

(1) For all $t \le T = O(n\log n)$,

$$\|P(X_t^0(\mathcal{C}) \in \cdot) - P(Y_t^0(\mathcal{C}) \in \cdot)\|_{\mathrm{TV}} \le o(1)\,.$$

(2) $\|\pi_{\mathcal{G}}(\omega(\mathcal{C}) \in \cdot) - \nu\|_{\mathrm{TV}} \le o(1)$.

Proof. We start with part (1). Our aim is to show that under the grand coupling of X_t^0 and Y_t^0, for every $t \le T = O(n\log n)$, we have $\mathbb{P}(X_t^0 \ne Y_t^0) \le o(1)$. Under the grand coupling, let $\mathscr{T}_T = (t_1, t_2, ..., t_{s(T)})$ denote the sequence of times on which the updated edge is in $\mathcal{C}$, so that $s(T)$ counts the number of updates in $\mathcal{C}$ by time T. We can then bound

$$\mathbb{P}(X_t^0 \ne Y_t^0) \le \mathbb{P}(s(T) > n^{2\varepsilon}) + \mathbb{P}(X_t^0 \ne Y_t^0, s(T) \le n^{2\varepsilon})\,.$$

The first term on the right-hand side is at most the probability that $\mathrm{Binom}(T, |\mathcal{C}|/|E(\mathcal{G})|) \ge n^{2\varepsilon}$ which is $o(1)$ by the Chernoff bound (4.6). It thus suffices to work on the event $s(T) \le n^{2\varepsilon}$.

Let $R := \frac{1}{6}\log_d n$ and let Z_t be the FK-dynamics chain (coupled to X_t, Y_t through the grand coupling) that freezes the configuration on $\mathcal{C} \cup (E(\mathcal{G}) \setminus \bigcup_{e\in\mathcal{C}} E(B_R(e)))$ to be all-1. Let Z_t^0 be the chain Z_t initialized from the configuration that is all-0 on $\bigcup_{e\in\mathcal{C}} E(B_R(e))$ (but all-1 on the frozen edges). Observe, trivially, that $X_t^0 \le Z_t^0$ for all $t \ge 0$. Also, observe that the updates of Z_t^0 are stochastically dominated by Glauber updates on the union of $2|\mathcal{C}|$ many d-ary trees $(\mathcal{T}_{e,1}, \mathcal{T}_{e,2})_{e\in\mathcal{C}}$ of depth R, rooted at the endpoints of the edges of $\mathcal{C}$, and each having $(1, \circlearrowleft)$ boundary conditions. By the monotonicity of the FK-dynamics, for all $t \ge 0$, we have that

$$\mathbb{P}\Big(Z_t^0\Big(\bigcup_{e\in\mathcal{C}} \{E(B_R(e)) \setminus \{e\}\}\Big) \in \cdot\Big) \preceq \bigotimes_{e\in\mathcal{C}} \bigotimes_{i\in\{1,2\}} \pi_{\mathcal{T}_{e,i}}^{(1,\circlearrowleft)}\,. \tag{7.1}$$

For each time $t_i \in \mathscr{T}_T$, when an edge $e_{t_i} \in \mathcal{C}$ is updated, $Y_{t_i}^0(e_{t_i})$ is drawn from $\mathrm{Ber}(\hat{p})$. At the same time, $X_{t_i}^0(e_{t_i})$ is drawn from $\mathrm{Ber}(\hat{p})$ if the endpoints of e_{t_i} are not connected in $X_{t_i}^0$, which in turn must occur if none of $(\mathcal{T}_{e,1}, \mathcal{T}_{e,2})_{e\in\mathcal{C}}$ have an open root-to-leaf path in Z_t^0, as $X_t^0 \le Z_t^0$.

By the stochastic domination of (7.1) on Z_t^0, and Lemma 5.1, the probability that the endpoints of e_{t_i} are connected in $Z_{t_i}^0$ is at most $2C(\hat{p}d)^R$; for ε sufficiently small (depending on p, q, d), the above is $O(n^{-3\varepsilon})$. On the event that $\{s(T) \le n^{2\varepsilon}\}$, we can union bound the above probability over the $s(T)$ times in $\mathscr{T}_T$, to find that $\mathbb{P}(X_t^0 \neq Y_t^0, s(T) \le n^{2\varepsilon})$ is at most $O(n^{-\varepsilon}) = o(1)$ as desired.

For part (2), consider the $2|\mathcal{C}|$ many d-ary trees $(\mathcal{T}_{e,1}, \mathcal{T}_{e,2})_{e\in\mathcal{C}}$ emanating from the endpoints of the edges of $\mathcal{C}$. Notice that if none of $(\mathcal{T}_{e,1}, \mathcal{T}_{e,2})_{e\in\mathcal{C}}$ have an open root-to-leaf path, then the values $\omega(\mathcal{C})$ are conditionally distributed as a product of $\mathrm{Ber}(\hat{p})$ random variables, i.e., $\omega(\mathcal{C})$ would conditionally be distributed as $\nu(A)$.

As such, the total-variation distance $\|\pi_{\mathcal{G}}(\omega(\mathcal{C}) \in \cdot) - \nu\|_{\mathrm{TV}}$ is bounded by the $\pi_{\mathcal{G}}$-probability that one of $(\mathcal{T}_{e,1}, \mathcal{T}_{e,2})_{e\in\mathcal{C}}$ has an open root-to-leaf path. By the stochastic domination

$$\pi_{\mathcal{G}}(\omega(\bigcup_{e\in\mathcal{C}} \mathcal{T}_{e,1} \cup \mathcal{T}_{e,2}) \in \cdot) \preceq \bigotimes_{e\in\mathcal{C}} \bigotimes_{i\in\{1,2\}} \pi_{\mathcal{T}_{e,i}}^{(1,\circlearrowleft)} .$$

By a union bound, the $\pi_{\mathcal{G}}$-probability that one of $(\mathcal{T}_{e,1}, \mathcal{T}_{e,2})_{e\in\mathcal{C}}$ has an open root-to-leaf path is at most

$$\sum_{e\in\mathcal{C}} \sum_{i\in\{1,2\}} \pi_{\mathcal{T}_{e,i}}^{(1,\circlearrowleft)}(e \leftrightarrow \partial\mathcal{T}_{e,i}) ,$$

which by Lemma 5.1 is at most $2n^{\varepsilon} \cdot C(\hat{p}d)^R$. For ε sufficiently small (depending on p, q, d) this is $o(1)$. □

Proof of Theorem 1.1: lower bound. Take any n-vertex graph $\mathcal{G}$ with $n^{1/5}$ many vertices whose balls of radius $\frac{1}{5}\log_d n$ are disjoint trees. Note that by Claim 7.1, such graphs have $\mathbf{P}_{\mathrm{RRG}}$-probability $1 - o(1)$. Take ε sufficiently small per Lemma 7.2. Consider the event $A^+ \subset \{0,1\}^{\mathcal{C}}$ that at least $\hat{p}n^{\varepsilon} - n^{2\varepsilon/3}$ of the edges in $\mathcal{C}$ are open. Let $(\overline{Y}_s)$ be the standard product chain over $|\mathcal{C}| = n^{\varepsilon}$ many i.i.d. $\mathrm{Ber}(\hat{p})$ random variables, coupled to $Y_t(\mathcal{C})$ via $\overline{Y}_{s(t)} = Y_t(\mathcal{C})$ for all t, where $s(t)$ counts the number of updates in $\mathcal{C}$ by time t. By item (1) of Lemma 7.2, for every $T = O(n\log n)$,

$$\begin{aligned}\mathbb{P}(X_T^0(\mathcal{C}) \in A^+) &\le \mathbb{P}(s(T) > cn^{\varepsilon}\log n^{\varepsilon}) + \mathbb{P}(Y_T^0 \in A^+, s(T) \le cn^{\varepsilon}\log n^{\varepsilon}) + o(1)\\ &\le \mathbb{P}(s(T) > cn^{\varepsilon}\log n) + \sup_{s\le cn^{\varepsilon}\log n} \mathbb{P}(\overline{Y}_s^0 \in A^+) + o(1) .\end{aligned}$$

Taking $T := c^2 n\log n^{\varepsilon} = \Theta(n\log n)$ for $c > 0$ sufficiently small, the probability that $s(T)$ is more than $cn^{\varepsilon}\log n^{\varepsilon}$ is $o(1)$ by the Chernoff bound (4.6). Turning to the middle term above, by the standard coupon collector bound, for every $c > 0$ sufficiently small,

$$\sup_{s\le cn^{\varepsilon}\log n^{\varepsilon}} \mathbb{P}(\overline{Y}_s^0 \in A^+) \le o(1) .$$

Combining the above, we obtain

$$\mathbb{P}(X_T^0(\mathcal{C}) \in A^+) = o(1) .$$

At the same time, by a Chernoff bound,

$$\nu(A^+) = \mathbb{P}(\mathrm{Bin}(n^{\varepsilon}, \hat{p}) \ge \hat{p}n^{\varepsilon} - n^{2\varepsilon/3}) = 1 - o(1) ,$$

so that by item (2) of Lemma 7.2, we have $\pi_{\mathcal{G}}(A^+) = 1 - o(1)$. This implies the mixing time is at least $T = \Omega(n\log n)$ as claimed. □

Acknowledgements. The authors thank the anonymous referee for detailed comments on the manuscript. The research of A.B. was supported in part by NSF grant CCF-1850443. R.G. thanks the Miller Institute for Basic Research in Science for its support.

REFERENCES

[1] K. S. Alexander. On weak mixing in lattice models. *Probab. Theory Related Fields*, 110(4):441–471, 1998.

[2] I. Bezáková, A. Blanca, Z. Chen, D. Štefankovič, and E. Vigoda. Lower bounds for testing graphical models: Colorings and antiferromagnetic Ising models. *Journal of Machine Learning Research*, 21(25):1–62, 2020.

[3] A. Blanca, P. Caputo, D. Parisi, A. Sinclair, and E. Vigoda. Entropy decay in the Swendsen-Wang dynamics. 2020. Preprint available at arXiv:2007.06931.

[4] A. Blanca, Z. Chen, and E. Vigoda. Swendsen-Wang dynamics for general graphs in the tree uniqueness region. *Random Structures & Algorithms*, 56(2):373–400, 2020.

[5] A. Blanca, A. Galanis, L. Goldberg, D. Štefankovic, E. Vigoda, and K. Yang. Sampling in uniqueness from the Potts and random-cluster models on random regular graphs. In *Proceedings of APPROX RANDOM*, 2018.

[6] A. Blanca, R. Gheissari, and E. Vigoda. Random-cluster dynamics in $\mathbb{Z}^2$: Rapid mixing with general boundary conditions. *Ann. Appl. Probab.*, 30(1):418–459, 02 2020.

[7] A. Blanca and A. Sinclair. Dynamics for the mean-field random-cluster model. In *Proc. of the 19th International Workshop on Randomization and Computation (RANDOM 2015)*, pages 528–543, 2015.

[8] A. Blanca and A. Sinclair. Random-cluster dynamics in $\mathbb{Z}^2$. *Probab. Theory Related Fields*, 2016. Extended abstract appeared in Proc. of the 27th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2016), pp. 498–513.

[9] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.

[10] B. Bollobás. *Random Graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2001.

[11] B. Bollobás, G. Grimmett, and S. Janson. The random-cluster model on the complete graph. *Probability Theory and Related Fields*, 104(3):283–317, 1996.

[12] M. Bordewich, C. Greenhill, and V. Patel. Mixing of the glauber dynamics for the ferromagnetic potts model. *Random Structures & Algorithms*, 48(1):21–52, 2016.

[13] P. Cuff, J. Ding, O. Louidor, E. Lubetzky, Y. Peres, and A. Sly. Glauber dynamics for the mean-field Potts model. *Journal of Statistical Physics*, 149(3):432–477, 2012.

[14] A. Dembo and A. Montanari. Ising models on locally tree-like graphs. *Ann. Appl. Probab.*, 20(2):565–592, 04 2010.

[15] A. Dembo, A. Montanari, A. Sly, and N. Sun. The replica symmetric solution for Potts models on d-regular graphs. *Communications in Mathematical Physics*, 327(2):551–575, 2014.

[16] A. Dembo, A. Montanari, and N. Sun. Factor models on locally tree-like graphs. *Ann. Probab.*, 41(6):4162–4213, 11 2013.

[17] P. Diaconis and L. Saloff-Coste. Logarithmic sobolev inequalities for finite markov chains. *Ann. Appl. Probab.*, 6(3):695–750, 08 1996.

[18] R. G. Edwards and A. D. Sokal. Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm. *Phys. Rev. D (3)*, 38(6):2009–2012, 1988.

[19] C. Efthymiou. A simple algorithm for random colouring $\mathrm{G}(n, d/n)$ using $(2+\varepsilon)d$ colours. In *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete algorithms*, pages 272–280. SIAM, 2012.

[20] C. Efthymiou. A simple algorithm for sampling colorings of $\mathrm{G}(n, d/n)$ up to the Gibbs uniqueness threshold. *SIAM Journal on Computing*, 45(6):2087–2116, 2016.

[21] C. M. Fortuin and P. W. Kasteleyn. On the random-cluster model. I. Introduction and relation to other models. *Physica*, 57:536–564, 1972.

[22] A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, 2016.

[23] A. Galanis, D. Štefankovič, and E. Vigoda. Inapproximability for antiferromagnetic spin systems in the tree nonuniqueness region. *Journal of the ACM (JACM)*, 62(6):1–60, 2015.

[24] A. Galanis, D. Štefankovic, and E. Vigoda. Swendsen-Wang Algorithm on the Mean-Field Potts Model. In *Proc. of the 19th International Workshop on Randomization and Computation (RANDOM 2015)*, pages 815–828, 2015.

[25] A. Galanis, D. Štefankovic, E. Vigoda, and L. Yang. Ferromagnetic Potts model: Refined #BIS-hardness and related results. *SIAM Journal on Computing*, 45(6):2004–2065, 2016.

[26] S. Ganguly and I. Seo. Information percolation and cutoff for the random-cluster model. *Random Structures & Algorithms*, 2018.

[27] R. Gheissari and E. Lubetzky. The effect of boundary conditions on mixing of 2D Potts models at discontinuous phase transitions. *Electron. J. Probab.*, 23:30 pp., 2018.

[28] R. Gheissari and E. Lubetzky. Mixing times of critical two-dimensional Potts models. *Comm. Pure Appl. Math*, 71(5):994–1046, 2018.

[29] R. Gheissari and E. Lubetzky. Quasi-polynomial mixing of critical two-dimensional random cluster models. *Random Structures and Algorithms*, 2019.

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

[31] G. Grimmett. The random-cluster model. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 73–123. Springer, Berlin, 2004.
[32] G. R. Grimmett and C. J. McDiarmid. On colouring random graphs. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 77, pages 313–324. Cambridge University Press, 1975.
[33] H. Guo and M. Jerrum. Random cluster dynamics for the Ising model is rapidly mixing. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017*, pages 1818–1827, 2017.
[34] O. Häggström. The random-cluster model on a homogeneous tree. *Probability Theory and Related Fields*, 104(2):231–253, 1996.
[35] T. P. Hayes and A. Sinclair. A general lower bound for mixing of single-site dynamics on graphs. In *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, pages 511–520. IEEE, 2005.
[36] T. Helmuth, M. Jenssen, and W. Perkins. Finite-size scaling, phase coexistence, and algorithms for the random cluster model on random graphs, 2020.
[37] S.-E. Huang, D. Huang, T. Kopelowitz, and S. Pettie. Fully dynamic connectivity in $o(\log n(\log\log n)^2)$ amortized expected time. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 510–520. SIAM, 2017.
[38] M. Huber. Perfect sampling using bounding chains. *The Annals of Applied Probability*, 14(2):734–753, 2004.
[39] J. Jonasson. The random cluster model on a general graph and a phase transition characterization of nonamenability. *Stochastic Processes and their Applications*, 79(2):335–354, 1999.
[40] D. A. Levin and Y. Peres. Markov chains and mixing times (second edition). *The Mathematical Intelligencer*, 41(1):90–91, 2019.
[41] M. Luczak and T. Luczak. The phase transition in the cluster-scaled model of a random graph. *Random Structures & Algorithms*, 28(2):215–246, 2006.
[42] F. Martinelli, E. Olivieri, and R. H. Schonmann. For 2-D lattice spin systems weak mixing implies strong mixing. *Comm. Math. Phys.*, 165(1):33–47, 1994.
[43] B. D. McKay and N. C. Wormald. Asymptotic enumeration by degree sequence of graphs with degrees $o(n^{1/2})$. *Combinatorica*, 11(4):369–382, 1991.
[44] E. Mossel and A. Sly. Rapid mixing of Gibbs sampling on graphs that are sparse on average. *Random Structures & Algorithms*, 35(2):250–270, 2009.
[45] E. Mossel and A. Sly. Exact thresholds for Ising–Gibbs samplers on general graphs. *Ann. Probab.*, 41(1):294–328, 01 2013.
[46] Y. Peres and P. Winkler. Can extra updates delay mixing? *Communications in Mathematical Physics*, 323(3):1007–1016, 2013.
[47] L. Saloff-Coste. *Lectures on finite Markov chains*, pages 301–413. Springer Berlin Heidelberg, Berlin, Heidelberg, 1997.
[48] A. Sly. Computational transition at the uniqueness threshold. In *Proceedings of the 51st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 287–296. IEEE, 2010.
[49] A. Sly and N. Sun. The computational hardness of counting in two-spin models on d-regular graphs. In *Proceedings of 53rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 361–369. IEEE, 2012.
[50] R. H. Swendsen and J.-S. Wang. Nonuniversal critical dynamics in Monte Carlo simulations. *Phys. Rev. Lett.*, 58:86–88, Jan 1987.
[51] M. Thorup. Near-optimal fully-dynamic graph connectivity. In *Proceedings of the 32nd annual ACM symposium on Theory of computing (STOC)*, pages 343–350, 2000.
[52] M. Ullrich. Comparison of Swendsen-Wang and heat-bath dynamics. *Random Structures and Algorithms*, 42(4):520–535, 2013.
[53] M. Ullrich. Rapid mixing of Swendsen-Wang dynamics in two dimensions. *Dissertationes Math. (Rozprawy Mat.)*, 502:64, 2014.
[54] L. Zdeborová and F. Krzkakala. Phase transitions in the coloring of random graphs. *Physical Review E*, 76(3):031131, 2007.

A. BLANCA
DEPARTMENT OF CSE, PENNSYLVANIA STATE UNIVERSITY
Email address: ablanca@cse.psu.edu

R. GHEISSARI
DEPARTMENTS OF STATISTICS AND EECS, UNIVERSITY OF CALIFORNIA, BERKELEY
Email address: gheissari@berkeley.edu