Sampling from Potts on Random Graphs of Unbounded Degree via Random-Cluster Dynamics

Antonio Blanca
Department of CSE, Pennsylvania State University, University Park, PA, USA

Reza Gheissari
Department of Statistics and EECS, University of California, Berkeley, CA, USA

Abstract

We consider the problem of sampling from the ferromagnetic Potts and random-cluster models on a general family of random graphs via the Glauber dynamics for the random-cluster model. The random-cluster model is parametrized by an edge probability $p \in (0,1)$ and a cluster weight $q > 0$. We establish that for every $q \geq 1$, the random-cluster Glauber dynamics mixes in optimal $\Theta(n \log n)$ steps on n-vertex random graphs having a prescribed degree sequence with bounded average branching γ throughout the full high-temperature uniqueness regime $p < p_u(q, \gamma)$.

The family of random graph models we consider includes the Erdős–Rényi random graph $G(n, \gamma/n)$, and so we provide the first polynomial-time sampling algorithm for the ferromagnetic Potts model on Erdős–Rényi random graphs for the full tree uniqueness regime. We accompany our results with mixing time lower bounds (exponential in the largest degree) for the Potts Glauber dynamics, in the same settings where our $\Theta(n \log n)$ bounds for the random-cluster Glauber dynamics apply. This reveals a novel and significant computational advantage of random-cluster based algorithms for sampling from the Potts model at high temperatures.

2012 ACM Subject Classification Theory of computation → Random walks and Markov chains; Theory of computation → Generating random combinatorial structures; Theory of computation → Random network models; Mathematics of computing → Probabilistic algorithms; Mathematics of computing → Markov processes

Keywords and phrases Potts model, random-cluster model, random graphs, Markov chains, mixing time, tree uniqueness

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2022.24

Category RANDOM

Related Version *Full Version*: `https://arxiv.org/abs/2107.10246`

Funding *Reza Gheissari*: Thanks the Miller Institute for Basic Research for its support.

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

1 Introduction

The ferromagnetic Potts model is a classical spin system model in statistical physics and computer science. It is defined on a finite graph $G = (V, E)$, by a set of spins (or colors) $[q] = \{1, ..., q\}$ and an edge weight or inverse temperature parameter $\beta > 0$. A configuration $\sigma \in \{1, \dots, q\}^V$ of the model is an assignment of spins to the vertices of V. The probability of σ is given by the Gibbs distribution:

$$\mu_{G,\beta,q}(\sigma) = \frac{1}{Z_{G,\beta,q}} \exp(-\beta D(\sigma)), \tag{1}$$

where $D(\sigma) = |\{\{v, w\} \in E : \sigma(v) \neq \sigma(w)\}|$ is the number of edges whose endpoints have different spins in σ, and $Z_{G,\beta,q}$ is a normalizing factor known as the partition function. The Ising model of ferromagnetism corresponds to the case where $q = 2$.

Sampling from the Potts Gibbs distribution (1) is one of the most frequently encountered problems when running simulations in statistical physics or when solving a variety of inference tasks in computer science; see [33, 32, 54, 52, 25, 24, 49] and the references therein for a sample of these applications. There is a family of powerful sampling algorithms for the Potts model based on its *random-cluster representation*, defined subsequently. Such algorithms, which include the Glauber dynamics of the random-cluster model and the widely-used Swendsen–Wang dynamics [55], are an attractive option computationally since they are often efficient at "low-temperatures" (large β), where standard Markov chains for the Potts model (including the canonical Glauber dynamics) often converge exponentially slowly; see, e.g., [13, 14, 18, 11, 38].

To be more precise, the *random-cluster model* on a finite graph $G = (V, E)$, is defined by an edge probability parameter $p \in (0, 1)$ and a cluster weight $q > 0$. The set of configurations of the model is the set of all subsets of edges $\omega \subseteq E$. The probability of each configuration ω is given by the Gibbs distribution:

$$\pi_{G,p,q}(\omega) = \frac{1}{Z_{G,p,q}} p^{|\omega|}(1-p)^{|E|-|\omega|} q^{c(\omega)}, \tag{2}$$

where $c(\omega)$ is the number of connected components (also called clusters) in the subgraph (V, ω), and $Z_{G,p,q}$ is the corresponding partition function. The random-cluster model was introduced by Fortuin and Kasteleyn [26] as a unifying framework for studying random graphs, spin systems, and electrical networks, and it is also known as the *FK-representation* of the Ising and Potts model.

For integer $q \geq 2$, a sample $\omega \subseteq E$ from the random-cluster Gibbs distribution $\pi_{G,p,q}$ can be easily transformed into one for the ferromagnetic q-state Potts model with inverse temperature $\beta(p) = -\ln(1-p)$, by independently assigning a random spin from $\{1, \ldots, q\}$ to (all vertices in) each connected component of (V, ω) [26, 23, 37]. As such, any sampling algorithm for the random-cluster model yields one for the ferromagnetic Potts model with essentially no computational overhead. This has led to significantly improved sampling algorithms for the Potts model in various low-temperature settings [55, 29, 47, 12, 57, 42] and more generally, to a broad interest in dynamics for the random-cluster model [15, 38, 7, 6, 3, 8, 5].

In this paper, we focus on the Glauber dynamics of the random-cluster model, which for easy distinction we will henceforth call the *FK-dynamics*. From a configuration $\omega_t \subseteq E$, one step of this Markov chain 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 $\omega_{t+1} = \omega_t \setminus \{e_t\}$.

Here, 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). The probabilities in step (2) are exactly the conditional probabilities of e_t being in the configuration ω_t given the remainder of ω_t. As such, this Markov chain is reversible with respect to $\pi_{G,p,q}$ and converges to it. We are interested in its *mixing time* t_{MIX}; i.e., the number of steps until the dynamics is within total variation distance $1/4$ of $\pi_{G,p,q}$, starting from the worst possible initial configuration.

As mentioned, the FK-dynamics is by now well-studied in its own right, though sharp analyses of its mixing time are only available on certain structured graphs like the complete graph [6, 36, 8], boxes in the infinite integer lattice graph $\mathbb{Z}^d$ [7, 5, 34, 35, 40, 31, 14], and

trees [1]. Recently, in [3], the authors studied the FK-dynamics on random regular graphs and established an optimal $\Theta(n \log n)$ mixing time bound throughout the entire high-temperature tree uniqueness regime.

Our aim in this paper is to study the FK-dynamics in settings in which the maximum degree of the underlying graph is much larger than its *average* degree. Such settings introduce hard technical challenges and have been well-studied for other models; see [20] for some early work. For instance, high-degree vertices are an obstruction to the fast convergence of the Ising/Potts Glauber dynamics. We later prove (see Section 1.2) that on a general class of random graphs on n vertices with maximum degree d_{MAX}, the Ising/Potts Glauber dynamics requires $n \cdot \exp(\Omega(d_{\text{MAX}}))$ steps to converge at high temperatures.

We reveal here that, for the same general family of random graphs, random-cluster based algorithms are not affected by the presence of high-degree vertices; both their mixing times and fast mixing parameter regimes are determined instead by the *average degree* of the graph. This reveals a novel and significant computational advantage of random-cluster based algorithms for sampling from the ferromagnetic Potts model *at high temperatures*. Indeed, prior to this work, random-cluster based sampling algorithms were only found to be more efficient than Ising/Potts Glauber dynamics at low temperatures.

More precisely, we study the mixing time of the FK-dynamics on random graphs of average branching $\gamma > 0$ in the full uniqueness (high-temperature) regime $p < p_u(q, \gamma)$. At integer γ, the threshold $p_u(q, \gamma)$, formally defined in (4), was identified in [39] as a uniqueness/non-uniqueness phase transition point of the random-cluster model on the *wired* γ-ary tree, i.e., where the leaves are externally wired to be in the same connected component. For us, $p_u(q, \gamma)$ is the natural extension of that function to non-integer γ, which we prove corresponds to the high-temperature uniqueness threshold of the random-cluster model on general trees of average branching γ for all $q \geq 1$. (This result could be of independent interest.)

Before we describe our general results for random graph models with fixed degree sequence (which we define in the next subsection) we present a special case of our main result of particular interest concerning the FK-dynamics on sparse Erdős–Rényi random graphs.

▶ **Theorem 1.** *Fix $q \geq 1$, $\gamma > 0$ and $p < p_u(q, \gamma)$. If $\mathcal{G}$ is an Erdős–Rényi random graph $\mathcal{G} \sim G(n, \gamma/n)$, then with probability $1 - o(1)$, $\mathcal{G}$ is such that the FK-dynamics on $\mathcal{G}$ satisfies $t_{MIX} = \Theta(n \log n)$.*

This yields a sampler for the Potts distribution on Erdős–Rényi random graphs with near-optimal running time. Let $\beta_u(q, \gamma) = -\ln(1 - p_u(q, \gamma))$ be the corresponding uniqueness point for the Potts model.

▶ **Corollary 2.** *Fix $q \geq 2$, $\gamma > 0$ and $\beta < \beta_u(q, \gamma)$. There is a sampling algorithm that, with probability $1 - o(1)$ over the choice of an Erdős–Rényi random graph $\mathcal{G} \sim G(n, \gamma/n)$, outputs a configuration whose distribution is within total-variation distance $\delta > 0$ of $\mu_{\mathcal{G},\beta,q}$ in time $O(n(\log n)^3 \log(1/\delta))$.*

Corollary 2 is a direct consequence of Theorem 1 and the aforementioned connection between the random-cluster model and the ferromagnetic Potts model. The extra $O((\log n)^2)$ factor in the running time of the algorithm comes from the (amortized) cost of checking whether the chosen edge is a cut-edge in each step of the FK-dynamics (see [43, 56]).

To the best of our knowledge, this is the first polynomial-time sampling algorithm for the Potts model on Erdős–Rényi random graphs for $q \geq 3$ and $\beta = \Omega(1)$. Even for the better understood $q = 2$ case (i.e., the Ising model), Corollary 2 provides the fastest known

sampling algorithm for this parameter regime, improving upon the running time of samplers based on the Glauber dynamics for the Ising model which converges in $n^{1+\Theta(\frac{1}{\log\log n})}$ steps for all $\beta < \beta_u(2,\gamma)$ [51].

We mention that the thresholds $p_u(q,\gamma)$ and $\beta_u(q,\gamma)$ should be sharp, in the sense that the FK-dynamics is conjectured to undergo polynomial or exponential slowdowns (depending on q) at the point $p_u(q,\gamma)$ (and when $q > 2$ in a whole critical window (p_u, p'_u)). This is by analogy with the FK-dynamics on the complete graph [36] and on random regular graphs [17]; see also [30, 42, 19].

1.1 Results on random graphs with general degree sequences

We next provide our main results on random graph models with a fixed degree sequence. Let $\mathbf{d}_n = (d_1, ..., d_n)$ be the degree sequence giving the degree of each vertex $v \in \{1, ..., n\}$. Our results will hold for uniform random graphs with degree sequence $\mathbf{d}_n$ under certain mild conditions on this degree sequence. The first condition we make on $\mathbf{d}_n$ is that the sequence is *graphical*: i.e., that there exists at least one simple graph having degree sequence $\mathbf{d}_n$.

Given a graphical sequence $\mathbf{d}_n$, we define $\mathbb{P}_{\mathrm{RG}(\mathbf{d}_n)}$ as the uniform distribution over all simple graphs on n vertices having degree sequence $\mathbf{d}_n$. The governing quantity in this degree sequence, in terms of the uniqueness thresholds for the Potts and random-cluster models on $\mathcal{G} \sim \mathbb{P}_{\mathrm{RG}(\mathbf{d}_n)}$, will be what we call the *effective offspring distribution* $\mathbb{P}_{\mathbf{d}_n}$. In words, the distribution $\mathbb{P}_{\mathbf{d}_n}$ will correspond to choosing $d-1$ with probability proportional to the total degree of vertices having degree d. This determines the offspring distribution corresponding to the random trees one obtains when looking at balls of small radius around a vertex of a random graph $\mathcal{G} \sim \mathbb{P}_{\mathrm{RG}(\mathbf{d}_n)}$. Specifically, a vertex of degree d is selected to be the next vertex added to the random tree with probability proportional to the total degree of all such vertices, and once it is selected and connected to its parent, it has $d-1$ available edges to connect to other randomly chosen vertices. Formally, $\mathbb{P}_{\mathbf{d}_n}$ is defined as the distribution over the set $\mathcal{M}(\mathbf{d}_n) = \{d_v - 1 : v \in \{1, ..., n\}\}$ where $x \in \mathcal{M}(\mathbf{d}_n)$ is assigned probability:

$$\mathbb{P}_{\mathbf{d}_n}(x) = \frac{\sum_v (x+1)\mathbf{1}_{\{d_v = x+1\}}}{\sum_v d_v}\,. \tag{3}$$

Our results apply to graphical degree sequences where $\mathbb{P}_{\mathbf{d}_n}$ has bounded finite moments, as we detail next.

▶ **Definition 3.** *Let $\mathcal{D}_{\gamma,\kappa}$ be the set of graphical degree sequences $(\mathbf{d}_n)_n$ such that $D \sim \mathbb{P}_{\mathbf{d}_n}$ has mean that is uniformly bounded away from γ and uniformly bounded κ-th moment. Formally,*

1. *There exists $\varepsilon > 0$ such that $\mathbb{E}_{\mathbf{d}_n}[D] < \gamma - \varepsilon$ for all sufficiently large n; and*
2. *There exists finite $K > 0$ such that $\mathbb{E}_{\mathbf{d}_n}[D^\kappa] < K$ for all sufficiently large n.*

This framework is fairly standard in the random graphs literature [10] and is similar to e.g., the setting of [28] for studying sampling from Potts on random graphs with fixed degree sequences at sufficiently low temperatures. While Definition 3 yields a fairly general family of random graphs, we draw attention to some well-studied examples which fall under its umbrella.

▶ **Example 4.** *Δ-regular random graph.* In this case, $\mathbf{d}_n = (\Delta, \ldots, \Delta)$ and the effective offspring distribution simply assigns probability 1 to $\Delta - 1$; thus $(\mathbf{d}_n)_n \in \mathcal{D}_{\gamma,\kappa}$ for every $\gamma > \Delta - 1$ and every κ.

▶ **Example 5.** *Erdős–Rényi random graph $G(n, \lambda/n)$.* It was shown in [45] that if $\mathbf{d}_n$ is drawn as an i.i.d. sequence of Poisson random variables of mean λ, then $\mathbb{P}_{\mathrm{RG}(\mathbf{d}_n)}$ is contiguous with respect to $G(n, \lambda/n)$. (Two random graph models are *contiguous* when any sequence of

events that has a probability of $1-o(1)$ in one has a probability of $1-o(1)$ in the other model as well.) Hence, it suffices to prove the desired results with high probability over such $\mathbf{d}_n$. Standard concentration estimates for Poisson random variables then give that for every $\gamma > \lambda$ and every κ, with high probability, $(\mathbf{d}_n)_n \in \mathcal{D}_{\gamma,\kappa}$.

Our main result is an optimal mixing time bound for the FK-dynamics on $\mathcal{G} \sim \mathbb{P}_{\mathrm{RG}(\mathbf{d}_n)}$, which applies to all the examples above and more generally to random graphs with degree sequences in $\mathcal{D}_{\gamma,\kappa}$.

▶ **Theorem 6.** *Fix $q \geq 1$, $\gamma > 0$, and $p < p_u(q,\gamma)$. There exists κ such that if $(\mathbf{d}_n)_n \in \mathcal{D}_{\gamma,\kappa}$, then with probability $1-o(1)$, $\mathcal{G} \sim \mathbb{P}_{RG(\mathbf{d}_n)}$ is such that the FK-dynamics on $\mathcal{G}$ satisfies $t_{MIX} = \Theta(n \log n)$.*

This parameter regime in Theorem 6 is tight as FK-dynamics have been very recently shown [17] to exponentially slow down as soon as $p > p_u(q,\gamma)$ for random regular graphs (Example 4) at integer $q > 2$.

The proof of the upper bound in Theorem 6 is the main content of this paper. As mentioned, the special case of the Δ-regular random graph (i.e., $\mathbf{d}_n = (\Delta, ..., \Delta)$) was the content of an earlier paper [3]. However, as soon as the degree sequence is not homogeneous, substantial further obstacles arise.

First, the uniqueness threshold for the random-cluster model on wired heterogeneous trees (specifically, with offspring distribution $\mathbb{P}_{\mathbf{d}_n}$) had not been established. In our proof of Theorem 6 we in fact require something much stronger; namely, an exponential decay of connectivities with the correct rate (see Lemma 17). The fact that $p_u(q,\gamma)$ is the uniqueness threshold in the regular case goes back to the work of Häggström [39] (see also [44, 2]). The exponential decay rate was established in [3]. To establish analogous results for the heterogeneous case, we combine the approach of [48] (which considered the special case of the Ising model $q = 2$) with ideas from [2], so as to recurse, not on the marginal of an edge of the tree, but rather on a functional of its probability of downwards connection to infinity.

The second technical obstacle concerns establishing that the FK-dynamics on $\mathcal{G} \sim \mathbb{P}_{\mathrm{RG}(\mathbf{d}_n)}$ *shatters*, i.e., that its components have size at most $O(n^{\varepsilon})$ after $O(n)$ steps of the dynamics. This is proved using a delicate revealing procedure for the random graph with the FK-dynamics configuration on top of it, a technique introduced in [3] for the case of random regular graphs. The heterogeneity of the degrees in the current setting, however, introduces extra correlations between the underlying graph and the FK-dynamics configuration, necessitating substantial modifications to the revealing procedure from [3].

The changes we make to deal with the above-described dependencies include: (i) modifications to the revealing process so that it is based on half-edges rather than vertices and the dynamics is run in continuous time, and (ii) a new criteria to truncate potentially unbounded increments in the revealing procedure. The more robust procedure yields a notable further improvement: we show that the shattering time is $O(n)$ (as opposed to $O(n \log n)$ in [3]). Though this improvement has no impact on the eventual mixing time bound, the more precise understanding of the shattering phase may be useful in other settings. A more detailed proof sketch of this theorem and the new complications that arise is provided in Section 2.

1.2 Slowdown for the Ising/Potts Glauber dynamics

Returning to the advantage of FK-dynamics in the presence of high-degree vertices, the following theorem shows that, in the same setting as Theorem 6, the mixing time of the Ising/Potts Glauber dynamics depends exponentially on the maximum degree.

▶ **Theorem 7.** *Fix $q \geq 2$, $\gamma > 0$ and $\beta < \beta_u(q,\gamma)$. Then there exists κ such that if $(\mathbf{d}_n)_n \in \mathcal{D}_{\gamma,\kappa}$, then with probability $1 - o(1)$, $\mathcal{G} \sim \mathbb{P}_{RG(\mathbf{d}_n)}$ is such that the Glauber dynamics for the ferromagnetic Potts model on $\mathcal{G}$ has $t_{MIX} = n \cdot \exp(\Omega(\|\mathbf{d}_n\|_\infty))$.*

Intuitively, the slowdown comes from the fact that the neighborhood of a vertex of degree $\|\mathbf{d}_n\|_\infty$ is a star graph, where the Ising/Potts Glauber dynamics mixes slowly for $\beta \gg 1/\|\mathbf{d}_n\|_\infty$. In random graphs at high temperatures ($\beta < \beta_u(q,\gamma)$) there is no interference with this effect from the rest of the graph. In contrast, the FK-dynamics in the star graph is fast mixing at all temperatures, so this obstruction is not present.

▶ Remark 8. We remark that under various decay of correlation conditions (see, e.g., [21, 41, 22, 16]) the mixing time of the Ising/Potts Glauber dynamics is known to be poly(n) when (roughly) $\beta \leq 1/\|\mathbf{d}_n\|_\infty$. This does not contradict Theorem 7, which holds when $\beta = \Omega(1)$. In fact, if one tracks the dependence on β in our proof, it gives $t_{\text{MIX}} = n \cdot \exp(\Omega(\beta^2\|\mathbf{d}_n\|_\infty))$.

The known $n^{1+\Omega(\frac{1}{\log\log n})}$ slowdown of the Ising/Potts Glauber dynamics on the Erdős–Rényi random graph [50, 51], where $\|\mathbf{d}_n\|_\infty = \Theta(\frac{\log n}{\log\log n})$, is a special case of Theorem 7. Below are a few examples where this slowdown is more dramatic, indeed stretched exponential in the total number of vertices.

▶ **Example 9.** *Power-law degree distributions.* Consider graphical sequences $(\mathbf{d}_n)_n$ satisfying item (1) in Definition 3, and for which the fraction of degrees of size ℓ is $\Theta(\ell^{-\zeta})$. For every κ, if $\zeta > \kappa + 2$, one would have $(\mathbf{d}_n)_n \in \mathcal{D}_{\gamma,\kappa}$. In such situations, $\|\mathbf{d}_n\|_\infty = \Theta(n^{1/\zeta})$ and $t_{\text{MIX}} = \exp(\Omega(n^{1/\zeta}))$.

▶ **Example 10.** *Planted high-degree vertices.* Consider a random Δ-regular random graph and change the degree of one vertex to $\Theta(n^\varepsilon)$. If $\varepsilon < 1/(\kappa+1)$ and $\gamma > \Delta - 1$, then $(\mathbf{d}_n) \in \mathcal{D}_{\gamma,\kappa}$ and $t_{\text{MIX}} = \exp(\Omega(n^\varepsilon))$.

In the above instances where the maximum degree is polynomial in n but the average degree is constant, there is an exponential vs. polynomial difference in the high-temperature mixing times of the Ising/Potts Glauber dynamics and of the FK-dynamics. At this level, the computational benefits of random-cluster based sampling methods also extend to the often implemented Swendsen–Wang dynamics [55]. In particular, using the comparison inequalities from [57] the upper bounds of Theorems 1 and 6 translate into $O(n^2 \log n)$ upper bounds on the mixing time of the Swendsen–Wang dynamics in those settings.

2 Proof outline

In this section, we present our technical contributions about random graphs (Section 2.1), the exponential decay of correlations and uniqueness on hetereogeneous trees (Section 2.2), and the shattering phenomenon of the FK-dynamics (Section 2.3), all of which we combine in Section 2.4 to yield our main result: the mixing time upper bound of Theorem 6. Throughout the paper, a subset $\omega \subset E$ is naturally identified with an assignment of $\{0,1\}$, or closed and open, to E, via $\omega(e) = 1$ if and only if $e \in \omega$. All our results should also be understood to hold uniformly over all sufficiently large n.

2.1 Random graphs

We start by describing the locally treelike structure and exponential rate of volume growth of random graphs with fixed degree sequence $(\mathbf{d}_n)_n \in \mathcal{D}_{\gamma,\kappa}$. It will be convenient to work with the *configuration model*, a useful and standard tool for studying random graphs with

fixed degree sequence. The configuration model $\mathbb{P}_{\mathrm{CM}(\mathbf{d}_n)}$ is a distribution over multigraphs on n vertices with degree sequence $\mathbf{d}_n$. It is defined by giving d_v half-edges to every vertex v and drawing a uniform random perfect matching on the $\sum_v d_v$ many half-edges to form the $\frac{1}{2}\sum_v d_v$ edges of the graph [9]. It is a standard fact that for any $(\mathbf{d}_n)_n \in \mathcal{D}_{\gamma,\kappa}$, and any sequence of sets A_n of simple graphs on n vertices, we have

$$\mathbb{P}_{\mathrm{RG}(\mathbf{d}_n)}(\mathcal{G} \in A_n) = o(1) \qquad \text{if and only if} \qquad \mathbb{P}_{\mathrm{CM}(\mathbf{d}_n)}(\mathcal{G} \in A_n) = o(1)\,:$$

see [9, 27]. It thus suffices to prove Theorems 6-7 for $\mathcal{G} \sim \mathbb{P}_{\mathrm{CM}(\mathbf{d}_n)}$.

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

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

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

▶ **Definition 11.** *We say that a graph $G = (V, E)$ is L-*Treelike *if there is a set $H \subset E$ with $|H| \le L$ such that the graph $(V, E \setminus H)$ is a tree. We say that G is (L, R)-*Treelike *if for every $v \in V$ the subgraph $(B_R(v), E(B_R(v)))$ is L-*Treelike*.*

The following lemma says that small balls of the random graph $\mathcal{G} \sim \mathbb{P}_{\mathrm{CM}(\mathbf{d}_n)}$ are treelike. Indeed, $B_R(v)$ in $\mathcal{G} \sim \mathbb{P}_{\mathrm{CM}(\mathbf{d}_n)}$ is typically a random tree with offspring distribution approximately $\mathbb{P}_{\mathbf{d}_n}$, defined in (3).

▶ **Lemma 12.** *There exists κ such that if $(\mathbf{d}_n)_n \in \mathcal{D}_{\gamma,\kappa}$ the following holds. For every $\delta > 0$, there exists $L = L(\delta)$ such that if $1 \le R \le (\frac{1}{2} - \delta)\log_\gamma n$, we have $\mathbb{P}_{CM(\mathbf{d}_n)}\big(\mathcal{G}$ is (L, R)-*Treelike*$\big) = 1 - o(n^{-10})$.*

Using standard concentration estimates for the volume of Galton–Watson trees, we also establish that if $(\mathbf{d}_n)_n \in \mathcal{D}_{\gamma,\kappa}$, then $\mathcal{G} \sim \mathbb{P}_{\mathrm{CM}(\mathbf{d}_n)}$ has average exponential rate γ of volume growth.

▶ **Definition 13.** *A graph $G = (V, E)$ on n vertices is said to have (γ, ε)-volume growth if for every $v \in V$ and every integer $r \in [\varepsilon \log_\gamma n, \frac{1}{2}\log_\gamma n]$ the graph has $|B_r(v)| \le \gamma^r$.*

▶ **Lemma 14.** *Fix $\varepsilon \in (0, \frac{1}{2})$. There exists $\kappa(\varepsilon)$ such that if $(\mathbf{d}_n)_n \in \mathcal{D}_{\gamma,\kappa}$, then*

$$\mathbb{P}_{CM(\mathbf{d}_n)}\big(\mathcal{G} \text{ has } (\gamma, \varepsilon)\text{-growth}\big) \ge 1 - o(n^{-10})\,.$$

2.2 Exponential decay and uniqueness on general trees and treelike graphs

Given the local tree structure of the random graphs from $\mathbb{P}_{\mathrm{CM}(\mathbf{d}_n)}$, to control the decay rate of connectivities of the random-cluster model on $\mathcal{G} \sim \mathbb{P}_{\mathrm{CM}(\mathbf{d}_n)}$, we need to first understand how these connectivities decay on heterogeneous (i.e., non-regular) trees. The relevant random-cluster measure on the tree requires the addition of *boundary conditions* mimicking the possible presence of open edges in the random graph outside of the treelike ball. Towards this, let us formally define boundary conditions.

▶ **Definition 15.** *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* (2) *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* (2)*. The boundary condition can alternatively be seen as external "wirings" of the vertices in the same element of ξ.*

▶ Remark 16. The *free* boundary condition, $\xi = 0$, corresponds to the case of no external wirings; i.e., its partition is the one consisting of 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 $\xi \leq \xi'$ if ξ is a finer partition than ξ'. When $q \geq 1$, the random-cluster model has the following monotonicity property: for any two boundary conditions $\xi \geq \xi'$, $\pi^{\xi}_{G,p,q} \succcurlyeq \pi^{\xi'}_{G,p,q}$ where $\succcurlyeq$ denotes stochastic domination [37].

Now define the threshold

$$p_u(q,\gamma) := 1 - \frac{1}{1 + \inf_{y>1} h(y)}\,, \qquad \text{where} \qquad h(y) := \frac{(y-1)(y^\gamma + q - 1)}{y^\gamma - y}\,. \tag{4}$$

The work [39] studied the random-cluster measure on homogeneous, d-ary trees, with wired boundary conditions and identified $p_u(q,d)$ as the *uniqueness threshold* such that whenever $p < p_u(q,d)$, the probability that the root is connected to a distance h in the wired d-ary tree goes to zero as $h \to \infty$; a different proof was given in [2]. In [3], it was shown that this decay is in fact exponential with rate $\hat{p} = p/(p+q(1-p))$. However, the methods of those papers do not easily extend to the non-regular setting, where there may be vertices of unbounded degree, but one would expect the threshold for connectivity decay to only depend on the *average* branching rate. In [48], it was shown that the analogue $\beta_u(2,\gamma)$ of (4) gives the correct uniqueness threshold in the case of the Ising model $q = 2$, for general (non-homogenous) trees of average branching γ. However, the argument there recursed over the single-site spin marginals, and relied on the fact that it was an Ising model whose interactions are nearest-neighbor. In the case of the random-cluster model, interactions between edge-marginals are non-local, and we therefore have to work with a more complicated functional encoding the probability of an edge being downward connected to the wired boundary. Combining ideas from [48] and [2], we are then able to establish uniqueness, and that connectivities decay exponentially with rate $\hat{p}$ on general heterogenous trees of average branching factor γ for all $q \geq 1$ and all $p < p_u(q,\gamma)$. When $p < p_u(q,\gamma)$, we have $\hat{p} < 1/\gamma$ (see e.g., [39, Theorem 1.5]); this indicates by a union bound why there will typically be no connections to the boundary in a tree of average branching γ.

More formally, let $\mathcal{T}_h = (V(\mathcal{T}_h), E(\mathcal{T}_h))$ be an arbitrary finite tree, rooted at ρ, and of height h. Let $\partial\mathcal{T}_h \subset V(\mathcal{T}_h)$ be the set of vertices of $\mathcal{T}_h$ at distance exactly h from ρ. For $v \in V(\mathcal{T}_h)$, let $\mathcal{T}_v$ be the subtree of $\mathcal{T}_h$ rooted at v, let $h(v)$ denotes the height of $\mathcal{T}_v$, and let $\partial\mathcal{T}_v = \partial\mathcal{T}_h \cap \mathcal{T}_v$. For a random-cluster configuration ω on $\mathcal{T}_h$, let $\mathcal{C}_\rho(\omega)$ denote the connected component of ω that contains the root ρ of $\mathcal{T}_h$. Finally, let $(1,\circlearrowleft)$ denote the boundary condition that wires all vertices of $\partial\mathcal{T}_h$ together, and also wires them up to the root, and let $\pi^{(1,\circlearrowleft)}_{\mathcal{T}_h}$ be the random-cluster measure with this boundary condition.

▶ **Lemma 17.** *Fix $q \geq 1$, $\gamma > 1$, $p < p_u(q,\gamma)$, and $\varepsilon \in [0,1)$. Suppose that $|\partial\mathcal{T}_v| \leq \gamma^{h(v)}$ for every $v \in V(\mathcal{T}_h)$ with $h(v) > \varepsilon h$. Then, there exists a constant $C = C(p,q,\gamma)$ such that for any $u \in \partial\mathcal{T}_h$*

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

We note that the condition that $|\partial\mathcal{T}_v| \leq \gamma^{h(v)}$ for every $v \in V(\mathcal{T}_h)$ with $h(v) > \varepsilon h$ in the lemma holds with high probability for random trees with averaging branching γ. In addition, the exponential decay rate in Lemma 17 is essentially optimal, and together with Lemmas 12–14, allows us to derive precise estimates on the exponential decay of connectivities on the

treelike balls around each vertex of the random graph $\mathcal{G} \sim \mathbb{P}_{\mathrm{CM}(\mathbf{d}_n)}$. We will actually need a sharp bound on the rate of influence decay between the boundary and the center of the ball $B_R(v)$; we find that this is the square of the rate of connectivity decay on a corresponding tree of depth R. (Intuitively, this is because *two* disjoint paths are required to reach the center of the ball in order for the boundary to have any effect on it.) To be more precise, let $G = (V, E)$ be a graph and for $v \in V$, let $E_v \subseteq E$ denote the set of edges incident to v.

▶ **Definition 18.** *A random-cluster boundary condition ξ on a graph H is said to be K-*Sparse *if the number of vertices in non-trivial (non-singleton) boundary components of ξ is at most K.*

▶ **Theorem 19.** *Fix $\gamma > 0$, $q \geq 1$, and $p < p_u(q, \gamma)$. Suppose G is (L, R)-*Treelike *for some L and some $R \leq \frac{1}{2}\log_\gamma n$. Also suppose G has (γ, ε)-volume growth for some $\varepsilon > 0$ sufficiently small. There exists a constant $C > 0$ such that for every $v \in G$, and any two K-*Sparse *boundary conditions ξ and τ on $B_R(v)$:*

$$\|\pi^{\xi}_{B_R(v)}(\omega(E_v) \in \cdot) - \pi^{\tau}_{B_R(v)}(\omega(E_v) \in \cdot)\|_{TV} \leq C\hat{p}^{(2-CL\sqrt{\varepsilon})R}\,.$$

2.3 Shattering of the FK-dynamics

With Theorem 19 in hand, the core of our argument becomes establishing that the boundary conditions induced by the FK-dynamics chains on the small balls around each vertex are *shattered*. This is formalized as follows.

▶ **Definition 20.** *A random-cluster configuration ω on $G = (V(G), E(G))$ is (K, R)-*Sparse *if, for every $v \in V(G)$, the boundary conditions induced on $B_R(v)$ by $\omega(E(G) \setminus E(B_r(v)))$ are K-*Sparse*.*

▶ Remark 21. It will be technically convenient to prove our results in continuous time instead of discrete time. In the continuous-time FK-dynamics, each edge of the graph has a rate-1 Poisson clock and every time a clock rings, the corresponding edge is updated as in the discrete-time version of the FK-dynamics; that is, according to the conditional distribution given the configuration off of this edge. It is a standard fact (see e.g., [46, Theorem 20.3]) that the discrete-time mixing time is comparable to $|E(\mathcal{G})|$ times the continuous-time mixing time. It therefore suffices for us to establish the mixing time bounds of Theorems 1 and 6 as $\Theta(\log n)$ bounds for the continuous-time version of the FK-dynamics. From this point on, we let $X_t^{x_0}$ denote the continuous-time FK-dynamics on $\mathcal{G}$ initialized from the configuration x_0, and use the superscripts 1 and 0 to denote the full (all-open) and empty (all-closed) configurations, respectively.

The following theorem says that after $O(1)$ continuous time, the configuration of the FK-dynamics from the all wired, and by monotonicity from any, initialization is (K, R)-Sparse, for $R \leq \frac{1}{2}\log_\gamma n$ and constant K.

▶ **Theorem 22.** *Fix $q \geq 1$, $\gamma > 0$ and $p < p_u(q, \gamma)$. For every $\delta > 0$, there exists κ such that if $(\mathbf{d}_n)_n \in \mathcal{D}_{\gamma,\kappa}$, there exists $T = T(p, q, \gamma)$ and $K = K(p, q, \gamma, \delta)$ such that for any $t \geq T$, and every $R \leq (\frac{1}{2} - \delta)\log_\gamma n$, with probability $1 - o(1)$, $\mathcal{G} \sim \mathbb{P}_{\mathrm{CM}(\mathbf{d}_n)}$ is such that $\mathbb{P}\big(X_t^1$ is (K, R)-Sparse$\big) \geq 1 - o(n^{-5})$.*

Our proof of Theorem 22 relies on a delicate simultaneous revealing procedure for the random graph, along with the connected component of a vertex v in X_t^1, showing that after a burn-in period, the configuration X_t^1 is shattered. This technique was introduced for random regular graphs in [3]. The revealing scheme for the component of a vertex v in the FK-dynamics chain X_t^1 roughly proceeds as follows (see Figure 1). First set the starting vertex v as "exposed", and iteratively, for each exposed vertex u do the following:

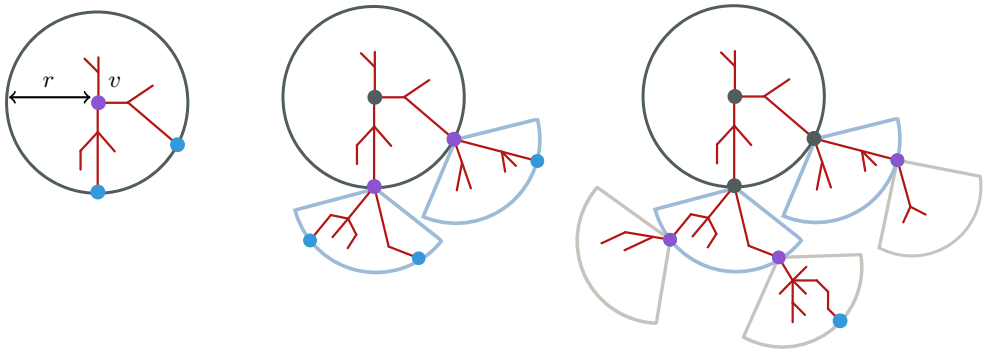

Figure 1 Three "generations" of the revealing procedure. In each figure, the purple vertices are the current generation of exposed vertices; the revealing procedure reveals the ball of radius r around such a vertex v, and a dominating localized FK-dynamics configuration $\tilde{\omega}(B_r(v))$ on that ball. The next generation of exposed vertices (blue) consists of those on the boundary $B_r(v)$ that are in the connected component of v in the configuration $\tilde{\omega}(B_r(v))$. Exposed vertices from previous generations are then colored black.

1. Reveal the ball $B_r(u)$ in the random graph for a large $r = O(1)$;
2. Reveal a configuration $\tilde{\omega}(B_r(u))$ that dominates the configuration of the FK-dynamics at time t on $B_r(u)$. This configuration will come from simulating the FK-dynamics that ignores all updates outside of $B_r(u)$ (effectively inducing the wired boundary condition on $B_r(u)$) and thus can be obtained independently of the dynamics on the rest of the graph;
3. Add to the set of exposed vertices all vertices of $\partial B_r(u)$ that get connected to u in $\tilde{\omega}(B_r(u))$.

The key point of the argument is then to stochastically dominate the exposed vertices by a branching process, which can be shown to be sub-critical (see Lemma 17). In our setting, the heterogeneity of the degrees causes substantial complications to the argument from [3], because in balls where the branching rate is locally larger than γ, the overlayed FK-dynamics configuration will actually be highly connected. The presence of high degrees also destroys the $O(1)$ bounds on the maximum number of new vertices that could possibly get exposed in step (3) above; this complicates relevant concentration arguments, as our branching process martingale will no longer have bounded increments.

2.4 Proof of main result

We end this section by proving the mixing time upper bound of Theorem 6. The proofs of all other results are deferred to the full version of this paper [4]. It suffices to prove the result for $\gamma > 1$ since $\lim_{\gamma \downarrow 1} p_u(q, \gamma) = 1$, and if $\gamma \geq \gamma'$, then $\mathcal{D}_{\gamma',\kappa} \subset \mathcal{D}_{\gamma,\kappa}$.

Proof of Theorem 6: upper bound. Fix $q > 1$, $\gamma > 1$ and $p < p_u(q, \gamma)$. Let $R = (\frac{1}{2} - \delta) \log_\gamma n$, where $\delta > 0$ is a small constant we choose later. For K and L fixed positive constants, $\varepsilon \in (0, 1/2)$ and $t \geq 0$, let $\Gamma_t = \Gamma_t(L, K, \delta, \varepsilon, \gamma)$ be the subset of (multi)graphs on n vertices with degree sequence $\mathbf{d}_n$ given by:

$$\Gamma_t = \{\mathcal{G} : \mathcal{G} \text{ is } (L, R)\text{-Treelike}, \text{ has } (\gamma, \varepsilon)\text{-volume growth}, \\ \text{and } \mathbb{P}(X^1_{\mathcal{G},t} \text{ is } (K, R)\text{-Sparse}) \geq 1 - n^{-5}\}.$$

By Lemmas 12 and 14, as well as Theorem 22, for every $\delta \in (0,1/2)$ and $\varepsilon \in (0,1/2)$, there exist constants $\kappa(p,q,\gamma,\delta)$, $L(\delta)$, $K(p,q,\gamma,\delta)$, and $T(p,q,\gamma)$ such that if $(\mathbf{d}_n)_n \in \mathcal{D}_{\gamma,\kappa}$ then $\mathbb{P}_{\mathrm{CM}(\mathbf{d}_n)}(\Gamma_T^c) = o(1)$. Hence, it suffices for us to prove that the mixing time of the FK-dynamics on any $\mathcal{G} \in \Gamma_T$ is $O(\log n)$.

Fix any $\mathcal{G} \in \Gamma_T$. Let $((X_t^{x_0})_{t\geq 0})_{x_0}$ be the family of FK-dynamics initialized from all possible configurations x_0, coupled via the standard grand coupling for the FK-dynamics; i.e., using the same clock rings and the same uniform random variables to make the edge updates while running the chain from different initializations. Recall that this coupling is monotone when $q \geq 1$ so that for every $t \geq 0$, if $X_t^{x_0} \leq X_t^{y_0}$, then $X_{t'}^{x_0} \leq X_{t'}^{y_0}$ for all $t' \geq t$. Using the standard fact that the coupling time provides a bound on the mixing time (see e.g., [46]), by a union bound over the edges, it suffices to show that under this grand coupling,

$$\mathbb{P}\big(X_{\hat{T}}^1(e) \neq X_{\hat{T}}^0(e)\big) = o(1/|E(\mathcal{G})|) \qquad \text{for every } e \in E(\mathcal{G})\,. \tag{5}$$

Now fix any such $e = \{u,v\}$ and for ease of notation, set $B_v = E(B_R(v))$ and $B_v^c = E(\mathcal{G}) \setminus B_v$. Consider two auxiliary copies of the FK-dynamics Y_t^1 and Y_t^0 that censor (ignore) all updates on edges of B_v^c after time T. The censoring inequality from [53] applied to the FK-dynamics [35, Theorem 2.5] implies that $Y_t^1 \succcurlyeq X_t^1$ and $Y_t^0 \preccurlyeq X_t^0$ for all $t \geq 0$ and thus

$$\mathbb{P}\big(X_t^1(e) \neq X_t^0(e)\big) = \mathbb{P}\big(X_t^1(e)=1\big) - \mathbb{P}\big(X_t^0(e)=1\big) \leq \mathbb{P}\big(Y_t^1(e)=1\big) - \mathbb{P}\big(Y_t^0(e)=1\big)\,.$$

Let $\mathcal{H}_v$ be the set of configurations on B_v^c such that the boundary conditions they induce on B_v are K-Sparse. (Here and throughout the paper, the boundary condition induced by a configuration $\omega(B^c)$ on a set B wires two vertices $w,w' \in V(B)$ if they are in the same connected component of $\omega(B^c)$.) By definition of Γ_T and monotonicity, we have for every $\mathcal{G} \in \Gamma_T$, $\mathbb{P}(Y_T^0(B_v^c) \notin \mathcal{H}_v) \leq \mathbb{P}(Y_T^1(B_v^c) \notin \mathcal{H}_v) \leq n^{-5}$, because up to time T there is no censoring, and so Y_T^1, Y_T^0 have the same distribution as X_T^1, X_T^0, respectively. Therefore, $\mathbb{P}(Y_t^1(e)=1) - \mathbb{P}(Y_t^0(e)=1)$ is bounded by

$$\max_{\phi^1,\phi^0 \in \mathcal{H}_v} \Big[\mathbb{P}(Y_t^1(e)=1 \mid Y_T^1(B_v^c)=\phi^1) - \mathbb{P}(Y_t^0(e)=1 \mid Y_T^0(B_v^c)=\phi^0)\Big] + 2n^{-5}\,.$$

Fix any $\phi^1,\phi^0 \in \mathcal{H}_v$, and let $t = T+s$. By the triangle inequality, the difference in the brackets is at most

$$\big|\mathbb{P}(Y_{T+s}^1(e)=1 \mid Y_T^1(B_v^c)=\phi^1) - \pi_{\mathcal{G}}(\omega(e)=1 \mid \omega(B_v^c)=\phi^1)\big| \tag{6}$$
$$+ \big|\pi_{\mathcal{G}}(\omega(e)=1 \mid \omega(B_v^c)=\phi^1) - \pi_{\mathcal{G}}(\omega(e)=1 \mid \omega(B_v^c)=\phi^0)\big| \tag{7}$$
$$+ \big|\mathbb{P}(Y_{T+s}^0(e)=1 \mid Y_T^0(B_v^c)=\phi^0) - \pi_{\mathcal{G}}(\omega(e)=1 \mid \omega(B_v^c)=\phi^0)\big|\,. \tag{8}$$

Observe that the chain $(Y_{T+s}^1)_{s\geq 0}$ may be viewed as an FK-dynamics on B_v with the boundary condition induced by ϕ^1, initialized from the (random) configuration $Y_T^1(B_v)$ and with stationary distribution $\pi_{\mathcal{G}}(\omega(B_v) \in \cdot \mid \omega(B_v^c)=\phi^1) = \pi_{B_v}^{\phi^1}$; the analogous statement is true for $(Y_{T+s}^0)_{s\geq 0}$ and $\pi_{B_v}^{\phi^0}$.

In order to bound (6) and (8), we use as an input a bound on the rate of convergence of FK-dynamics on treelike graphs with sparse boundary conditions. This bound comes from a straightforward (Dirichlet form) comparison to the FK-dynamics on a tree with free boundary conditions, and its proof is deferred to the full version [4].

▶ **Lemma 23.** *Consider an L-*Treelike *graph $G=(V,E)$ with a K-*Sparse *boundary condition ξ. Let $\pi_{\mathrm{MIN}} = \min_{x\in\Omega} \pi_G^{\xi}(x)$. For every $p \in (0,1)$ and $q>0$, there exists $\alpha_0 = \alpha_0(p,q,L,K) > 0$ such that*

$$\max_{x_0\in\Omega} \|\mathbb{P}(X_t^{x_0} \in \cdot) - \pi_G^{\xi}\|_{TV} \leq e^{-\alpha_0 t}\sqrt{2\log(1/\pi_{\mathrm{MIN}})}\,.$$

Setting $\hat{T} = T + \hat{S}_n$ where $\hat{S}_n = \hat{C} \log n$ for a constant $\hat{C}(p, q, \gamma, L, K)$ sufficiently large, since B_v is L-Treelike and ϕ^1 is K-Sparse, by Lemma 23, we obtain the following, and its analogue for (8):

$$\left|\mathbb{P}(Y^1_{\hat{T}}(e) = 1 \mid Y^1_T(B_v^c) = \phi^1) - \pi_{\mathcal{G}}(\omega(e) = 1 \mid \omega(B_v^c) = \phi^1)\right| \leq n^{-5}\,.$$

Finally, since both ϕ^1 and ϕ^0 induce K-Sparse boundary conditions on B_v and $\mathcal{G}$ is (L, R)-Treelike with (γ, ε)-volume growth, by Theorem 19 there exists $C = C(p, q, L, K, \gamma) > 0$ such that (7) is at most

$$\|\pi^{\phi^1}_{B_v}(\omega(E_v) \in \cdot) - \pi^{\phi^0}_{B_v}(\omega(E_v) \in \cdot)\|_{\mathrm{TV}} \leq C\hat{p}^{2(1-C\sqrt{\varepsilon})R} \leq C\hat{p}^{(1-2\delta)(1-C\sqrt{\varepsilon})\log_\gamma n}\,,$$

where E_v is the set of edges incident to v, and we used $R = (\frac{1}{2} - \delta)\log_\gamma n$. Setting $\theta = (1-2\delta)(1-C\sqrt{\varepsilon})$,

$$\|\pi^{\phi^1}_{B_v}(\omega(E_v) \in \cdot) - \pi^{\phi^0}_{B_v}(\omega(E_v) \in \cdot)\|_{\mathrm{TV}} \leq C\hat{p}^{\theta \log_\gamma n} = Cn^{-\theta(1-\frac{1}{\log_{\hat{p}\gamma}\gamma})}\,. \tag{9}$$

Since $\hat{p} < 1/\gamma$, $\log_{\hat{p}\gamma}\gamma < 0$, there is some $c_{p,\gamma} > 0$ such that the right-hand side is $Cn^{-\theta(1+c_{p,\gamma})}$. By taking ε, δ sufficiently small, θ can be made arbitrarily close to 1, so that (9) is $o(1/n)$.

Now notice that $|E(\mathcal{G})| = O(n)$. To see this, observe that by Jensen's inequality $(\frac{1}{n}\sum_v d_v)^2 \leq \frac{1}{n}\sum_v d_v^2$, and since $(\mathbf{d}_n) \in \mathcal{D}_{\gamma,\kappa}$, we also have $\sum_v d_v^2 \leq (1+\gamma)\sum_v d_v$. Combining these two inequalities we find that $|E(\mathcal{G})| \leq \frac{(1+\gamma)n}{2}$. Therefore, each of (6)–(8) are $o(1/|E(\mathcal{G})|)$, implying (5) as desired. ◀

References

1. Antonio Blanca, Zongchen Chen, Daniel Štefankovič, and Eric Vigoda. The Swendsen-Wang dynamics on trees. In *Proceedings of the 25th International Workshop on Randomization and Computation (RANDOM)*, 2021.
2. Antonio Blanca, Andreas Galanis, Leslie Ann Goldberg, Daniel Štefankovič, Eric Vigoda, and Kuan Yang. Sampling in uniqueness from the Potts and random-cluster models on random regular graphs. In *Proceedings of the 22nd International Workshop on Randomization and Computation (RANDOM)*, 2018.
3. Antonio Blanca and Reza Gheissari. Random-cluster dynamics on random regular graphs in tree uniqueness. *Communications in Mathematical Physics*, 2021. doi:10.1007/s00220-021-04093-z.
4. Antonio Blanca and Reza Gheissari. Sampling from Potts on random graphs of unbounded degree via random-cluster dynamics. *arXiv preprint*, 2021. arXiv:2107.10246.
5. Antonio Blanca, Reza Gheissari, and Eric Vigoda. Random-cluster dynamics in $\mathbb{Z}^2$: Rapid mixing with general boundary conditions. *Ann. Appl. Probab.*, 30(1):418–459, February 2020. doi:10.1214/19-AAP1505.
6. Antonio Blanca and Alistair Sinclair. Dynamics for the mean-field random-cluster model. In *Proceedings of the 19th International Workshop on Randomization and Computation (RANDOM)*, pages 528–543, 2015. doi:10.4230/LIPIcs.APPROX-RANDOM.2015.528.
7. Antonio Blanca and Alistair Sinclair. Random-cluster dynamics in $\mathbb{Z}^2$. *Probab. Theory Related Fields*, 2016. Extended abstract appeared in Proc. of the 27th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2016), pp. 498–513. doi:10.1007/s00440-016-0725-1.
8. Antonio Blanca, Alistair Sinclair, and Xusheng Zhang. The critical mean-field Chayes-Machta dynamics. In *Proceedings of the 25th International Workshop on Randomization and Computation (RANDOM)*, 2021.
9. Béla 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 Béla Bollobás. *Random Graphs*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2 edition, 2001. doi:10.1017/CBO9780511814068.

11 Magnus Bordewich, Catherine Greenhill, and Viresh Patel. Mixing of the glauber dynamics for the ferromagnetic potts model. *Random Structures & Algorithms*, 48(1):21–52, 2016.

12 Christian Borgs, Jennifer Chayes, Tyler Helmuth, Will Perkins, and Prasad Tetali. Efficient sampling and counting algorithms for the Potts model on Z^d at all temperatures. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2020, pages 738–751, New York, NY, USA, 2020. Association for Computing Machinery. doi:10.1145/3357713.3384271.

13 Christian Borgs, Jennifer T. Chayes, Alan Frieze, Jeong H. Kim, Prasad Tetali, Eric Vigoda, and Van H. Vu. Torpid mixing of some Monte Carlo Markov chain algorithms in statistical physics. In *Proc. of the 40th Annual Symposium on Foundations of Computer Science (FOCS 1999)*, pages 218–229, 1999. doi:10.1109/SFFCS.1999.814594.

14 Christian Borgs, Jennifer T. Chayes, and Prasad Tetali. Tight bounds for mixing of the Swendsen-Wang algorithm at the Potts transition point. *Probab. Theory Related Fields*, 152(3-4):509–557, 2012. doi:10.1007/s00440-010-0329-0.

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

16 Xiaoyu Chen, Weiming Feng, Yitong Yin, and Xinyuan Zhang. Rapid mixing of Glauber dynamics via spectral independence for all degrees. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 137–148. IEEE, 2022.

17 Amin Coja-Oghlan, Andreas Galanis, Leslie Ann Goldberg, Jean Bernoulli Ravelomanana, Daniel Štefankovič, and Eric Vigoda. Metastability of the Potts Ferromagnet on Random Regular Graphs. In Mikołaj Bojańczyk, Emanuela Merelli, and David P. Woodruff, editors, *49th International Colloquium on Automata, Languages, and Programming (ICALP 2022)*, volume 229 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 45:1–45:20, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ICALP.2022.45.

18 Paul Cuff, Jian Ding, Oren Louidor, Eyal Lubetzky, Yuval Peres, and Allan Sly. Glauber dynamics for the mean-field Potts model. *Journal of Statistical Physics*, 149(3):432–477, 2012.

19 Amir Dembo, Andrea Montanari, Allan Sly, and Nike Sun. The replica symmetric solution for Potts models on d-regular graphs. *Communications in Mathematical Physics*, 327(2):551–575, 2014. doi:10.1007/s00220-014-1956-6.

20 Martin Dyer, Abraham D Flaxman, Alan M Frieze, and Eric Vigoda. Randomly coloring sparse random graphs with fewer colors than the maximum degree. *Random Structures & Algorithms*, 29(4):450–465, 2006.

21 Martin Dyer, Leslie Ann Goldberg, and Mark Jerrum. Dobrushin conditions and systematic scan. *Combinatorics, Probability and Computing*, 17(6):761–779, 2008.

22 Martin Dyer, Leslie Ann Goldberg, and Mark Jerrum. Matrix norms and rapid mixing for spin systems. *The Annals of Applied Probability*, 19(1):71–107, 2009.

23 Robert G. Edwards and Alan D. Sokal. Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm. *Phys. Rev. D (3)*, 38(6):2009–2012, 1988. doi:10.1103/PhysRevD.38.2009.

24 G. Ellison. Learning, local interaction, and coordination. *Econometrica: Journal of the Econometric Society*, pages 1047–1071, 1993.

25 J. Felsenstein. *Inferring phylogenies*, volume 2. Sinauer Associates, Inc., Sunderland, MA, 2004.

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

27 Alan Frieze and Michał Karoński. *Introduction to random graphs*. Cambridge University Press, 2016.

28 Andreas Galanis, Leslie Ann Goldberg, and James Stewart. Fast mixing via polymers for random graphs with unbounded degree. *Information and Computation*, page 104894, 2022.

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

30 Andreas Galanis, Daniel Štefankovič, Eric Vigoda, and Linji Yang. Ferromagnetic Potts model: Refined #BIS-hardness and related results. *SIAM Journal on Computing*, 45(6):2004–2065, 2016.

31 Shirshendu Ganguly and Insuk Seo. Information percolation and cutoff for the random-cluster model. *Random Structures & Algorithms*, 57(3):770–822, 2020. doi:10.1002/rsa.20931.

32 S. Geman and C. Graffigne. Markov random field image models and their applications to computer vision. In *Proceedings of the International Congress of Mathematicians*, volume 1, pages 1496–1517. Berkeley, CA, 1986.

33 H.-O. Georgii. *Gibbs measures and phase transitions*, volume 9. Walter de Gruyter, 2011.

34 Reza Gheissari and Eyal Lubetzky. Mixing times of critical two-dimensional Potts models. *Comm. Pure Appl. Math*, 71(5):994–1046, 2018.

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

36 Reza Gheissari, Eyal Lubetzky, and Yuval Peres. Exponentially slow mixing in the mean-field Swendsen–Wang dynamics. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 56(1):68–86, 2020. doi:10.1214/18-AIHP955.

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

38 Heng Guo and Mark Jerrum. Random cluster dynamics for the Ising model is rapidly mixing. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1818–1827, 2017.

39 Olle Häggström. The random-cluster model on a homogeneous tree. *Probability Theory and Related Fields*, 104(2):231–253, 1996. doi:10.1007/BF01247839.

40 Matan Harel and Yinon Spinka. Finitary codings for the random-cluster model and other infinite-range monotone models. *Electronic Journal of Probability*, 27:1–32, 2022.

41 Thomas P Hayes. A simple condition implying rapid mixing of single-site dynamics on spin systems. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 39–46. IEEE, 2006.

42 Tyler Helmuth, Matthew Jenssen, and Will Perkins. Finite-size scaling, phase coexistence, and algorithms for the random cluster model on random graphs, 2020. arXiv:2006.11580.

43 Jacob Holm, Kristian De Lichtenberg, and Mikkel Thorup. Poly-logarithmic deterministic fully-dynamic algorithms for connectivity, minimum spanning tree, 2-edge, and biconnectivity. *Journal of the ACM (JACM)*, 48(4):723–760, 2001.

44 Johan 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.

45 Jeong Han Kim. Poisson cloning model for random graphs. In *International Congress of Mathematicians (ICM)*, 2006.

46 David A. Levin and Yuval Peres. Markov chains and mixing times (second edition). *The Mathematical Intelligencer*, 41(1):90–91, 2019. doi:10.1007/s00283-018-9839-x.

47 Yun Long, Asaf Nachmias, Weiyang Ning, and Yuval Peres. A power law of order 1/4 for critical mean field Swendsen-Wang dynamics. *Mem. Amer. Math. Soc.*, 232(1092), 2014. doi:10.1090/memo/1092.

48 Russell Lyons. The Ising model and percolation on trees and tree-like graphs. *Communications in Mathematical Physics*, 125(2):337–353, 1989. doi:cmp/1104179469.

49 A. Montanari and A. Saberi. The spread of innovations in social networks. *Proceedings of the National Academy of Sciences*, 107(47):20196–20201, 2010.

50 Elchanan Mossel and Allan Sly. Rapid mixing of Gibbs sampling on graphs that are sparse on average. *Random Structures & Algorithms*, 35(2):250–270, 2009.

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

52 S. Osindero and G.E. Hinton. Modeling image patches with a directed hierarchy of Markov random fields. In *Advances in neural information processing systems*, pages 1121–1128, 2008.

53 Yuval Peres and Peter Winkler. Can extra updates delay mixing? *Communications in Mathematical Physics*, 323(3):1007–1016, 2013. doi:10.1007/s00220-013-1776-0.

54 S. Roth and M.J. Black. Fields of experts: A framework for learning image priors. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 860–867, 2005.

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

56 Mikkel Thorup. Near-optimal fully-dynamic graph connectivity. In *Proceedings of the 32nd Annual ACM symposium on Theory of computing (STOC)*, pages 343–350, 2000.

57 Mario Ullrich. Swendsen–Wang is faster than single-bond dynamics. *SIAM Journal on Discrete Mathematics*, 28(1):37–48, 2014. doi:10.1137/120864003.