The Critical Mean-Field Chayes-Machta Dynamics

Antonio Blanca
Pennsylvania State University, University Park, PA, USA

Alistair Sinclair
University of California at Berkeley, CA, USA

Xusheng Zhang
Pennsylvania State University, University Park, PA, USA

Abstract

The random-cluster model is a unifying framework for studying random graphs, spin systems and electrical networks that plays a fundamental role in designing efficient Markov Chain Monte Carlo (MCMC) sampling algorithms for the classical ferromagnetic Ising and Potts models. In this paper, we study a natural non-local Markov chain known as the *Chayes-Machta dynamics* for the mean-field case of the random-cluster model, where the underlying graph is the complete graph on n vertices. The random-cluster model is parametrized by an *edge probability* p and a *cluster weight* q. Our focus is on the critical regime: $p = p_c(q)$ and $q \in (1, 2)$, where $p_c(q)$ is the threshold corresponding to the order-disorder phase transition of the model. We show that the mixing time of the Chayes-Machta dynamics is $O(\log n \cdot \log\log n)$ in this parameter regime, which reveals that the dynamics does not undergo an exponential slowdown at criticality, a surprising fact that had been predicted (but not proved) by statistical physicists. This also provides a nearly optimal bound (up to the $\log\log n$ factor) for the mixing time of the mean-field Chayes-Machta dynamics in the only regime of parameters where no non-trivial bound was previously known. Our proof consists of a multi-phased coupling argument that combines several key ingredients, including a new local limit theorem, a precise bound on the maximum of symmetric random walks with varying step sizes, and tailored estimates for critical random graphs. In addition, we derive an improved comparison inequality between the mixing time of the Chayes-Machta dynamics and that of the local Glauber dynamics on general graphs; this results in better mixing time bounds for the local dynamics in the mean-field setting.

2012 ACM Subject Classification Mathematics of computing → Markov processes; Theory of computation → Design and analysis of algorithms; Theory of computation → Random walks and Markov chains; Theory of computation → Random network models; Mathematics of computing → Random graphs

Keywords and phrases Markov Chains, Mixing Times, Random-cluster Model, Ising and Potts Models, Mean-field, Chayes-Machta Dynamics, Random Graphs

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2021.47

Category RANDOM

Related Version *Full Version*: `https://arxiv.org/abs/2102.03004v2`

Funding *Antonio Blanca*: Research supported in part by NSF grant CCF-1850443.
Alistair Sinclair: Research supported in part by NSF grant CCF-1815328.
Xusheng Zhang: Research supported in part by NSF grant CCF-1850443.

1 Introduction

The *random-cluster model* generalizes classical random graph and spin system models, providing a unifying framework for their study [12]. It plays an indispensable role in the design of efficient Markov Chain Monte Carlo (MCMC) sampling algorithms for the ferromagnetic Ising/Potts model [23, 6, 17] and has become a fundamental tool in the study of phase transitions [1, 11, 10].

The random-cluster model is defined on a finite graph $G = (V, E)$ with 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 $A \subseteq E$. The probability of each configuration A is given by the Gibbs distribution:

$$\mu_{G,p,q}(A) = \frac{1}{Z} \cdot p^{|A|}(1-p)^{|E|-|A|}q^{c(A)}; \tag{1}$$

where $c(A)$ is the number of connected components in (V, A) and $Z := Z(G, p, q)$ is the normalizing factor called the *partition function.*

The special case when $q = 1$ corresponds to the independent bond percolation model, where each edge of the graph G appears independently with probability p. Independent bond percolation is also known as the Erdős-Rényi random graph model when G is the complete graph.

For integer $q \geq 2$, the random-cluster model is closely related to the ferromagnetic q-state Potts model. Configurations in the q-state Potts model are the assignments of spin values $\{1, \ldots, q\}$ to the vertices of G; the $q = 2$ case corresponds to the Ising model. A sample $A \subseteq E$ from the random-cluster distribution can be easily transformed into one for the Ising/Potts model by independently assigning a random spin from $\{1, \ldots, q\}$ to each connected component of (V, A). Random-cluster based sampling algorithms, which include the widely-studied Swendsen-Wang dynamics [22], are an attractive alternative to Ising/Potts Markov chains since they are often efficient at "low-temperatures" (large p). In this parameter regime, several standard Ising/Potts Markov chains are known to converge slowly.

In this paper we investigate the *Chayes-Machta (CM) dynamics* [9], a natural Markov chain on random-cluster configurations that converges to the random-cluster measure. The CM dynamics is a generalization to non-integer values of q of the widely studied Swendsen-Wang dynamics [22]. As with all applications of the MCMC method, the primary object of study is the *mixing time*, i.e., the number of steps until the dynamics is close to its stationary distribution, starting from the worst possible initial configuration. We are interested in understanding how the mixing time of the CM dynamics grows as the size of the graph G increases, and in particular how it relates to the phase transition of the model.

Given a random-cluster configuration (V, A), one step of the CM dynamics is defined as follows:

- **(i)** activate each connected component of (V, A) independently with probability $1/q$;
- **(ii)** remove all edges connecting active vertices;
- **(iii)** add each edge between active vertices independently with probability p, leaving the rest of the configuration unchanged.

We call (i) the *activation* sub-step, and (ii) and (iii) combined the *percolation* sub-step. It is easy to check that this dynamics is reversible with respect to the Gibbs distribution (1) and thus converges to it [9]. For integer q, the CM dynamics may be viewed as a variant of the Swendsen-Wang dynamics. In the Swendsen-Wang dynamics, each connected component of (V, A) receives a random color from $\{1, \ldots, q\}$, and the edges are updated within each color class as in (ii) and (iii) above; in contrast, the CM dynamics updates the edges of exactly *one* color class. However, note that the Swendsen-Wang dynamics is only well-defined for integer q, while the CM dynamics is feasible for any real $q > 1$. Indeed, the CM dynamics was introduced precisely to allow this generalization.

The study of the interplay between phase transitions and the mixing time of Markov chains goes back to pioneering work in mathematical physics in the late 1980s. This connection for the specific case of the CM dynamics on the complete n-vertex graph, known as the *mean-field model*, has received some attention in recent years (see [5, 13, 16]) and is the focus of this

paper. As we shall see, the mean-field case is already quite non-trivial, and has historically proven to be a useful starting point in understanding various types of dynamics on more general graphs. We note that, so far, the mean-field is the only setting in which there are tight mixing time bounds for the CM dynamics; all other known bounds are deduced indirectly via comparison with other Markov chains, thus incurring significant overhead [6, 4, 15, 3, 23, 5].

The phase transition for the mean-field random-cluster model is fairly well-understood [8, 20]. In this setting, it is natural to re-parameterize by setting $p = \zeta/n$; the phase transition then occurs at the *critical* value $\zeta = \zeta_{\text{CR}}(q)$, where $\zeta_{\text{CR}}(q) = q$ when $q \in (0, 2]$ and $\zeta_{\text{CR}}(q) = 2(\frac{q-1}{q-2})\log(q-1)$ for $q > 2$. For $\zeta < \zeta_{\text{CR}}(q)$ all components are of size $O(\log n)$ with high probability (w.h.p.); that is, with probability tending to 1 as $n \to \infty$. On the other hand, for $\zeta > \zeta_{\text{CR}}(q)$ there is a unique giant component of size $\approx \theta n$, where $\theta = \theta(\zeta, q)$. The phase transition is thus analogous to that in $G(n, p)$ corresponding to the emergence of a giant component.

The phase structure of the mean-field random-cluster model, however, is more subtle and depends crucially on the second parameter q. In particular, when $q > 2$ the model exhibits *phase coexistence* at the critical threshold $\zeta = \zeta_{\text{CR}}(q)$. Roughly speaking, this means that when $\zeta = \zeta_{\text{CR}}(q)$, the set of configurations with all connected components of size $O(\log n)$, and set of configurations with a unique giant component, contribute each a constant fraction of the probability mass. For $q \leq 2$, on the other hand, there is no phase coexistence.

Phase coexistence at $\zeta = \zeta_{\text{CR}}(q)$ when $q > 2$ has significant implications for the speed of convergence of Markov chains, including the CM dynamics. The following detailed connection between the phase structure of the model and the mixing time $\tau_{\text{mix}}^{\text{CM}}$ of the CM dynamics was recently established in [5, 2, 16]. When $q > 2$, we have:

$$\tau_{\text{mix}}^{\text{CM}} = \begin{cases} \Theta(\log n) & \text{if } \zeta \notin [\zeta_{\text{L}}, \zeta_{\text{R}}); \\ \Theta(n^{1/3}) & \text{if } \zeta = \zeta_{\text{L}}; \\ e^{\Omega(n)} & \text{if } \zeta \in (\zeta_{\text{L}}, \zeta_{\text{R}}), \end{cases} \tag{2}$$

where $(\zeta_{\text{L}}, \zeta_{\text{R}})$ is the so-called *metastability window*. It is known that $\zeta_{\text{R}} = q$, but ζ_{L} does not have a closed form; see [5, 20]; we note that $\zeta_{\text{CR}}(q) \in (\zeta_{\text{L}}, \zeta_{\text{R}})$ for $q > 2$.

When $q \in (1, 2]$, there is no metastability window, and the mixing time of the mean-field CM dynamics is $\Theta(\log n)$ for all $\zeta \neq \zeta_{\text{CR}}(q)$. In view of these results, the only case remaining open is when $q \in (1, 2]$ and $\zeta = \zeta_{\text{CR}}(q)$. Our main result shown below concerns precisely this regime, which is particularly delicate and had resisted analysis until now for reasons we explain in our proof overview.

▶ **Theorem 1.** *The mixing time of the CM dynamics on the complete n-vertex graph when $\zeta = \zeta_{\text{CR}}(q) = q$ and $q \in (1, 2)$ is $O(\log n \cdot \log\log n)$.*

A $\Omega(\log n)$ lower bound is known for the mixing time of the mean-field CM dynamics that holds for all $p \in (0, 1)$ and $q > 1$ [5]. Therefore, our result is tight up to the lower order $O(\log\log n)$ factor, and in fact even better as we explain in Remark 16. The conjectured tight bound when $\zeta = \zeta_{\text{CR}}(q)$ and $q \in (1, 2)$ is $\Theta(\log n)$. We mention that the $\zeta = \zeta_{\text{CR}}(q)$ and $q = 2$ case, which is quite different and not covered by Theorem 1, was considered earlier in [19] for the closely related Swendsen-Wang dynamics, and a tight $\Theta(n^{1/4})$ bound was established for its mixing time. The same mixing time bound is expected for the CM dynamics in this regime; see Remark 7 for further comments about the $\zeta = \zeta_{\text{CR}}(q)$, $q = 2$ case.

Our result establishes a striking behavior for random-cluster dynamics when $q \in (1, 2)$. Namely, there is no slowdown (exponential or power law) in this regime at the critical threshold $\zeta = \zeta_{\text{CR}}(q)$. Note that for $q > 2$, as described in (2) above, the mixing time of the

dynamics undergoes an exponential slowdown, transitioning from $\Theta(\log n)$ when $\zeta < \zeta_{\text{L}}$, to a power law at $\zeta = \zeta_{\text{L}}$, and to exponential in n when $\zeta \in (\zeta_{\text{L}}, \zeta_{\text{R}})$. The absence of a critical slowdown for $q \in (1, 2)$ was in fact predicted by the statistical physics community [14], and our result provides the first rigorous proof of this phenomenon.

Our second result concerns the local *Glauber dynamics* for the random-cluster model. In each step, the Glauber dynamics updates a single edge of the current configuration chosen uniformly at random; a precise definition of this Markov chain is given in the full version of the present paper [7]. In [5], it was established that any upper bound on the mixing time $\tau_{\text{mix}}^{\text{CM}}$ of the CM dynamics can be translated to one for the mixing time $\tau_{\text{mix}}^{\text{GD}}$ of the Glauber dynamics, at the expense of a $\tilde{O}(n^4)$ factor; the $\tilde{O}$ notation hides polylogarithmic factors. In particular, it was proved in [5] that $\tau_{\text{mix}}^{\text{GD}} \leq \tau_{\text{mix}}^{\text{CM}} \cdot \tilde{O}(n^4)$. We provide here an improvement of this comparison inequality.

▶ **Theorem 2.** *For all $q > 1$ and all $\zeta = O(1)$, $\tau_{\text{mix}}^{\text{GD}} \leq \tau_{\text{mix}}^{\text{CM}} \cdot O(n^3(\log n)^2)$.*

To prove this theorem, we establish a general comparison inequality that holds for any graph, any $q \geq 1$ and any $p \in (0, 1)$; see the full version of this paper [7] for a precise statement. When combined with the known mixing time bounds for the CM dynamics on the complete graph, Theorem 2 yields that the random-cluster Glauber dynamics mixes in $\tilde{O}(n^3)$ steps when $q > 2$ and $\zeta \notin (\zeta_{\text{L}}, \zeta_{\text{R}})$, or when $q \in (1, 2)$ and $\zeta = O(1)$. In these regimes, the mixing time of the Glauber dynamics was previously known to be $\tilde{O}(n^4)$ and is conjectured to be $\tilde{O}(n^2)$; the improved comparison inequality in Theorem 2 gets us closer to this conjectured tight bound. We note, however, that even if one showed the conjectured optimal bound for the mixing time of the Glauber dynamics, the CM is faster, even if we take into account the computational cost associated to implementing its steps.

We conclude this introduction with some brief remarks about our analysis techniques, which combine several key ingredients in a non-trivial way. Our bound on the mixing time uses the well-known technique of *coupling*: in order to show that the mixing time is $O(\log n \cdot \log\log n)$, it suffices to couple the evolutions of two copies of the dynamics, starting from two arbitrary configurations, in such a way that they arrive at the *same* configuration after $O(\log n)$ steps with probability $\Omega(1/\log\log n)$. (The moves of the two copies can be correlated any way we choose, provided that each copy, viewed in isolation, is a valid realization of the dynamics.) Because of the delicate nature of the phase transition in the random-cluster model, combined with the fact that the percolation sub-step of the CM dynamics is critical when $\zeta = q$, our coupling is somewhat elaborate and proceeds in multiple phases. The first phase consists of a *burn-in period*, where the two copies of the chain are run independently and the evolution of their largest components is observed until they have shrunk to their "typical" sizes. This part of the analysis is inspired by similar arguments in earlier work [5, 19, 13].

In the second phase, we design a coupling of the activation of the connected components of the two copies which uses: (i) a local limit theorem, which can be thought of as a stronger version of a central limit theorem; (ii) a precise understanding of the distribution of the maximum of symmetric random walks on $\mathbb{Z}$ with varying step sizes; and (iii) precise estimates for the component structure of random graphs. We develop tailored versions of these probabilistic tools for our setting and combine them to guarantee that the same number of vertices from each copy are activated in each step w.h.p. for sufficiently many steps. This phase of the coupling is the main novelty in our analysis, and allows us to quickly converge to the same configuration.

The rest of the paper is organized as follows. In Section 2, we give a detailed overview of our proof. The proof of the key step in our coupling construction is provided in Section 3; all others proofs are deferred to the full version of this paper [7].

2 Proof sketch and techniques

We now give a detailed sketch of the multi-phased coupling argument for proving Theorem 1. We start by formally defining the notions of mixing and coupling times. Let Ω_{RC} be the set of random-cluster configurations of a graph G; let $\mathcal{M}$ be the transition matrix of a random-cluster Markov chain with stationary distribution $\mu = \mu_{G,p,q}$, and let $\mathcal{M}^t(X_0, \cdot)$ be the distribution of the chain after t steps starting from $X_0 \in \Omega_{\text{RC}}$. The *$\varepsilon$-mixing time* of $\mathcal{M}$ is given by

$$\tau_{\text{mix}}^{\mathcal{M}}(\varepsilon) := \max_{X_0 \in \Omega_{\text{RC}}} \min_{t \geq 0} \left\{ ||\mathcal{M}^t(X_0, \cdot) - \mu(\cdot)||_{\text{TV}} \leq \varepsilon \right\},$$

where $|| \cdot ||_{\text{TV}}$ denotes total variation distance. In particular, the *mixing time* of $\mathcal{M}$ is $\tau_{\text{mix}}^{\mathcal{M}} := \tau_{\text{mix}}^{\mathcal{M}}(1/4)$.

A *(one step) coupling* of the Markov chain $\mathcal{M}$ specifies, for every pair of states $(X_t, Y_t) \in \Omega_{\text{RC}} \times \Omega_{\text{RC}}$, a probability distribution over (X_{t+1}, Y_{t+1}) such that the processes $\{X_t\}$ and $\{Y_t\}$ are valid realizations of $\mathcal{M}$, and if $X_t = Y_t$ then $X_{t+1} = Y_{t+1}$. The *coupling time*, denoted T_{coup}, is the minimum T such that $\Pr[X_T \neq Y_T] \leq 1/4$, starting from the worst possible pair of configurations in Ω_{RC}. It is a standard fact that $\tau_{\text{mix}}^{\mathcal{M}} \leq T_{\text{coup}}$; moreover, when $\Pr[X_T = Y_T] \geq \delta$ for some coupling, then $\tau_{\text{mix}}^{\mathcal{M}} = O(T\delta^{-1})$ (see, e.g., [18]).

We provide first a high level description of our coupling for the CM dynamics. For this, we require the following notation. For a random cluster configuration X, let $L_i(X)$ denote the size of the i-th largest connected component in (V, X), and let $\mathcal{R}_i(X) := \sum_{j \geq i} L_j(X)^2$; in particular, $\mathcal{R}_1(X)$ is the sum of the squares of the sizes of all the components of (V, X). Our coupling has three main phases:

1. *Burn-in period:* run two copies $\{X_t\}$, $\{Y_t\}$ independently, starting from a pair of arbitrary initial configurations, until $\mathcal{R}_1(X_T) = O(n^{4/3})$ and $\mathcal{R}_1(Y_T) = O(n^{4/3})$.
2. *Coupling to the same component structure:* starting from X_T and Y_T such that $\mathcal{R}_1(X_T) = O(n^{4/3})$ and $\mathcal{R}_1(Y_T) = O(n^{4/3})$, we design a two-phased coupling that reaches two configurations with the same component structure as follows:
   - 2a. A two-step coupling after which the two configurations agree on all "large components";
   - 2b. A coupling that after $O(\log n)$ additional steps reaches two configurations that will also have the same "small component" structure.
3. *Coupling to the same configuration:* starting from two configurations with the same component structure, there is a straightforward coupling that couples the two configurations in $O(\log n)$ steps w.h.p.

We proceed to describe each of these phases in detail.

2.1 The burn-in period

During the initial phase, two copies of the dynamics evolve independently. This is called a *burn-in period* and in our case consists of three sub-phases.

In the first sub-phase of the burn-in period the goal is to reach a configuration X such that $\mathcal{R}_2(X) = O(n^{4/3})$. For this, we use a lemma from [2], which shows that after $T = O(\log n)$ steps of the CM dynamics $\mathcal{R}_2(X_T) = O(n^{4/3})$ with at least constant probability; this holds when $\zeta = q$ for any initial configuration X_0 and any $q > 1$.

▶ **Lemma 3** ([2], Lemma 3.42). *Let $q > 1$ and $\zeta = q$, and let X_0 be an arbitrary random-cluster configuration. Then, for any constant $C \geq 0$, after $T = O(\log n)$ steps $\mathcal{R}_2(X_T) = O(n^{4/3})$ and $L_1(X_T) > Cn^{2/3}$ with probability $\Omega(1)$.*

In the second and third sub-phases of the burn-in period, we use the fact that when $\mathcal{R}_2(X_t) = O(n^{4/3})$, the number of activated vertices is well concentrated around n/q (its expectation). This is used to show that the size of the largest component contracts at a constant rate for $T = O(\log n)$ steps until a configuration X_T is reached such that $\mathcal{R}_1(X_T) = O(n^{4/3})$. This part of the analysis is split into two sub-phases because the contraction for $L_1(X_t)$ requires a more delicate analysis when $L_1(X_t) = o(n)$; this is captured in the following two lemmas.

▶ **Lemma 4.** *Let $\zeta = q$ and $q \in (1,2)$. Suppose $\mathcal{R}_2(X_0) = O(n^{4/3})$. Then, for any constant $\delta > 0$, there exists $T = T(\delta) = O(1)$ such that $\mathcal{R}_2(X_T) = O(n^{4/3})$ and $L_1(X_T) \le \delta n$ with probability $\Omega(1)$.*

▶ **Lemma 5.** *Let $\zeta = q$ and $q \in (1,2)$. Suppose $\mathcal{R}_2(X_0) = O(n^{4/3})$ and that $L_1(X_0) \le \delta n$ for a sufficiently small constant δ. Then, with probability $\Omega(1)$, after $T = O(\log n)$ steps $\mathcal{R}_1(X_T) = O(n^{4/3})$.*

Lemmas 4 and 5 are proved in the full paper [7]. Combining them with Lemma 3 immediately yields the following theorem.

▶ **Theorem 6.** *Let $\zeta = q$, $q \in (1,2)$ and let X_0 be an arbitrary random-cluster configuration of the complete n-vertex graph. Then, with probability $\Omega(1)$, after $T = O(\log n)$ steps $\mathcal{R}_1(X_T) = O(n^{4/3})$.*

▶ Remark 7. The contraction of $L_1(X_t)$ established by Lemmas 4 and 5 only occurs when $q \in (1,2)$; when $q > 2$ the quantity $L_1(X_t)$ may increase in expectation, whereas for $q = 2$ we have $\mathrm{E}[L_1(X_{t+1}) \mid X_t] \approx L_1(X_t)$, and the contraction of the size of the largest component is due instead to fluctuations caused by a large second moment. (This is what causes the power law slowdown when $\zeta = q = 2$.)

▶ Remark 8. Sub-steps (ii) and (iii) of the CM dynamics are equivalent to replacing the active portion of the configuration by a $G(m, q/n)$ random graph, where m is the number of active vertices. Since $\mathrm{E}[m] = n/q$, one key challenge in the proofs of Lemmas 4 and 5, and in fact in the entirety of our analysis, is that the random graph $G(m, q/n)$ is critical or almost critical w.h.p. since $m \cdot q/n \approx 1$; consequently its structural properties are not well concentrated and cannot be maintained for the required $O(\log n)$ steps of the coupling. This is one of the key reasons why the $\zeta = \zeta_{\mathrm{CR}}(q) = q$ regime is quite delicate.

2.2 Coupling to the same component structure

For the second phase of the coupling, we assume that we start from a pair of configurations X_0, Y_0 such that $\mathcal{R}_1(X_0) = O(n^{4/3})$, $\mathcal{R}_1(Y_0) = O(n^{4/3})$. The goal is to show that after $T = O(\log n)$ steps, with probability $\Omega(1/\log\log n)$, we reach two configurations X_T and Y_T with the same component structure; i.e., $L_j(X_T) = L_j(Y_T)$ for all $j \ge 1$. In particular, we prove the following.

▶ **Theorem 9.** *Let $\zeta = q$, $q \in (1,2)$ and suppose X_0, Y_0 are random-cluster configurations such that $\mathcal{R}_1(X_0) = O(n^{4/3})$ and $\mathcal{R}_1(Y_0) = O(n^{4/3})$. Then, there exists a coupling of the CM steps such that after $T = O(\log n)$ steps X_T and Y_T have the same component structure with probability $\Omega\left((\log\log n)^{-1}\right)$.*

Our coupling construction for proving Theorem 9 has two main sub-phases. The first is a two-step coupling after which the two configurations agree on all the components of size above a certain threshold $B_\omega = n^{2/3}/\omega(n)$, where $\omega(n)$ is a slowly increasing function. For convenience and definiteness we set $\omega(n) = \log\log\log\log n$. In the second sub-phase we take care of matching the small component structures.

We note that when the same number of vertices are activated from each copy of the chain, we can easily couple the percolation sub-step (with an arbitrary bijection between the activated vertices) and replace the configuration on the active vertices in both chains with the same random sub-graph; consequently, the component structure in the updated sub-graph would be identical. Our goal is thus to design a coupling of the activation of the components that activates the same number of vertices in both copies in every step.

In order for the initial two-step coupling to succeed, certain (additional) properties of the configurations are required. These properties are achieved with a continuation of the initial burn-in phase for a small number of $O(\log \omega(n))$ steps. For a random-cluster configuration X, let $\widetilde{\mathcal{R}}_\omega(X) = \sum_{j:L_j(X)\le B_\omega} L_j(X)^2$ and let $I(X)$ denote the number of isolated vertices of X. Our extension of the burn-in period is captured by the following lemma.

▶ **Lemma 10.** *Let $\zeta = q$, $q \in (1,2)$ and suppose X_0 is such that $\mathcal{R}_1(X_0) = O(n^{4/3})$. Then, there exists $T = O(\log \omega(n))$ and a constant $\beta > 0$ such that $\widetilde{\mathcal{R}}_\omega(X_T) = O(n^{4/3}\omega(n)^{-1/2})$, $\mathcal{R}_1(X_T) = O(n^{4/3})$ and $I(X_T) = \Omega(n)$ with probability $\Omega(\omega(n)^{-\beta})$.*

With these bounds on $\widetilde{\mathcal{R}}_\omega(X_T)$, $\widetilde{\mathcal{R}}_\omega(Y_T)$, $I(X_T)$ and $I(Y_T)$, we construct the two-step coupling for matching the *large* component structure. The construction crucially relies on a new local limit theorem (Theorem 17). In particular, under our assumptions, when $\omega(n)$ is small enough, there are few components with sizes above B_ω. Hence, we can condition on the event that all of them are activated simultaneously. The difference in the number of active vertices generated by the activation of these large components can then be "corrected" by a coupling of the activation of the smaller components; for this we use our new local limit theorem.

Specifically, our local limit theorem applies to the random variables corresponding to the number of activated vertices from the small components of each copy. We prove it using a result of Mukhin [21] and the fact that, among the small components, there are (roughly speaking) many components of many different sizes. To establish the latter we require a refinement of known random graph estimates (see Lemma 23).

To formally state our result we introduce some additional notation. Let $\mathcal{S}_\omega(X)$ be the set of connected components of X with sizes greater than B_ω. At step t, the activation of the components of two random-cluster configurations X_t and Y_t is done using a maximal matching W_t between the components of X_t and Y_t, with the restriction that only components of equal size are matched to each other. For an increasing positive function g and each integer $k \ge 0$, define $\hat{N}_k(t,g) := \hat{N}_k(X_t, Y_t, g)$ as the number of matched pairs in W_t whose component sizes are in the interval

$$\mathcal{I}_k(g) = \Big[\frac{\vartheta n^{2/3}}{2g(n)^{2^k}}, \frac{\vartheta n^{2/3}}{g(n)^{2^k}}\Big],$$

where $\vartheta > 0$ is a fixed large constant (independent of n).

▶ **Lemma 11.** *Let $\zeta = q$, $q \in (1,2)$ and suppose X_0, Y_0 are random-cluster configurations such that $\mathcal{R}_1(X_0) = O(n^{4/3})$, $\widetilde{\mathcal{R}}_\omega(X_0) = O(n^{4/3}\omega(n)^{-1/2})$, $I(X_0) = \Omega(n)$ and similarly for Y_0. Then, there exists a two-step coupling of the CM dynamics such that $\mathcal{S}_\omega(X_2) = \mathcal{S}_\omega(Y_2)$ with probability $\exp\big(-O(\omega(n)^9)\big)$.*

Moreover, $L_1(X_2) = O(n^{2/3}\omega(n))$, $\mathcal{R}_2(X_2) = O(n^{4/3})$, $\widetilde{\mathcal{R}}_\omega(X_2) = O(n^{4/3}\omega(n)^{-1/2})$, $I(X_2) = \Omega(n)$, $\hat{N}_k(2, \omega(n)) = \Omega(\omega(n)^{3\cdot 2^{k-1}})$ for all $k \ge 1$ such that $n^{2/3}\omega(n)^{-2^{k-1}} \to \infty$, and similarly for Y_2.

From the first part of the lemma we obtain two configurations that agree on all of their large components, as desired, while the second part guarantees additional structural properties for the resulting configurations so that the next sub-phase of the coupling can also succeed with the required probability.

In the second sub-phase, after the large component are matched, we can design a coupling that activates exactly the same number of vertices from each copy of the chain. To analyze this coupling we use a precise estimate on the distribution of the maximum of symmetric random walks over integers (with steps of different sizes). We are first required to run the chains coupled for $T = O(\log \omega(n))$ steps, so that certain additional structural properties appear. Let $M(X_t)$ and $M(Y_t)$ be the components in the matching W_t that belong to X_t and Y_t, respectively, and let $D(X_t)$ and $D(Y_t)$ be the complements of $M(X_t)$ and $M(Y_t)$. Let $Z_t = \sum_{\mathcal{C} \in D(X_t) \cup D(Y_t)} |\mathcal{C}|^2$.

▶ **Lemma 12.** *Let $\zeta = q$, $q \in (1, 2)$. Suppose X_0 and Y_0 are random-cluster configurations such that $\mathcal{S}_\omega(X_0) = \mathcal{S}_\omega(Y_0)$, and $\hat{N}_k(0, \omega(n)) = \Omega(\omega(n)^{3 \cdot 2^{k-1}})$ for all $k \geq 1$ such that $n^{2/3}\omega(n)^{-2^{k-1}} \to \infty$. Suppose also that $L_1(X_0) = O(n^{2/3}\omega(n))$, $\mathcal{R}_2(X_0) = O(n^{4/3})$, $\widetilde{\mathcal{R}}_\omega(X_0) = O(n^{4/3}\omega(n)^{-1/2})$, $I(X_0) = \Omega(n)$, and similarly for Y_0.*

Then, there exists a coupling of the CM steps such that with probability $e^{-O((\log \omega(n))^2)}$ after $T = O(\log \omega(n))$ steps: $\mathcal{S}_\omega(X_T) = \mathcal{S}_\omega(Y_T)$, $Z_T = O(n^{4/3}\omega(n)^{-1/2})$, $\hat{N}_k(T, \omega(n)^{1/2}) = \Omega(\omega(n)^{3 \cdot 2^{k-2}})$ for all $k \geq 1$ such that $n^{2/3}\omega(n)^{-2^{k-1}} \to \infty$, $\mathcal{R}_1(X_T) = O(n^{4/3})$, $I(X_T) = \Omega(n)$, and similarly for Y_T.

The proof of Lemma 12 also uses our local limit theorem (Theorem 17).

The final step of our construction is a coupling of the activation of the components of size less than B_ω, so that exactly the same number of vertices are activated from each copy in each step w.h.p.

▶ **Lemma 13.** *Let $\zeta = q$, $q \in (1, 2)$ and suppose X_0 and Y_0 are random-cluster configurations such that $\mathcal{S}_\omega(X_0) = \mathcal{S}_\omega(Y_0)$, $Z_0 = O(n^{4/3}\omega(n)^{-1/2})$, and $\hat{N}_k\left(0, \omega(n)^{1/2}\right) = \Omega(\omega(n)^{3 \cdot 2^{k-2}})$ for all $k \geq 1$ such that $n^{2/3}\omega(n)^{-2^{k-1}} \to \infty$. Suppose also that $\mathcal{R}_1(X_0) = O(n^{4/3})$, $I(X_0) = \Omega(n)$ and similarly for Y_0. Then, there exist a coupling of the CM steps and a constant $\beta > 0$ such that after $T = O(\log n)$ steps, X_T and Y_T have the same component structure with probability $\Omega\left((\log \log \log n)^{-\beta}\right)$.*

We comment briefly on how we prove this lemma. Our starting point is two configurations with the same "large" component structure; i.e., $\mathcal{S}_\omega(X_0) = \mathcal{S}_\omega(Y_0)$. We use the maximal matching W_0 to couple the activation of the large components in X_0 and Y_0. The small components *not matched* by W_0, i.e., those counted in Z_0, are then activated independently. This creates a discrepancy $\mathcal{D}_0$ between the number of active vertices from each copy. Since $\mathrm{E}[\mathcal{D}_0] = 0$ and $\mathrm{Var}(\mathcal{D}_0) = \Theta(Z_0) = \Theta(n^{4/3}\omega(n)^{-1/2})$, it follows from Hoeffding's inequality that $\mathcal{D}_0 \leq n^{2/3}\omega(n)^{-1/4}$ w.h.p. To fix this discrepancy, we use the small components *matched* by W_0. Specifically, under the assumptions in Lemma 13, we can construct a coupling of the activation of the small components so that the difference in the number of activated vertices from the small components from each copy is exactly $\mathcal{D}_0$ with probability $\Omega(1)$. This part of the construction utilizes random walks over the integers; in particular, we use a lower bound for the maximum of such a random walk.

We need to repeat this process until $Z_t = 0$; this takes $O(\log n)$ steps since $Z_t \approx (1 - 1/q)^t Z_0$. However, there are a few complications. First, the initial assumptions on the component structure of the configurations are not preserved for this many steps w.h.p., so we need to relax the requirements as the process evolves. This is in turn possible because the discrepancy $\mathcal{D}_t$ decreases with each step, which implies that the probability of success of the coupling increases at each step.

Proof of Lemma 10, 12 and 13 is provided in the full version of the present paper [7]. We now indicate how these lemmas lead to a proof of Theorem 9 stated earlier.

Proof of Theorem 9. Suppose $\mathcal{R}_1(X_0) = O(n^{4/3})$ and $\mathcal{R}_1(Y_0) = O(n^{4/3})$. It follows from Lemma 10, 11, 12 and 13 that there exists a coupling of the CM steps such that after $T = O(\log n)$ steps, X_T and Y_T could have the same component structure. This coupling succeeds with probability at least

$$\rho = \Omega(\omega(n)^{-\beta_1}) \cdot \exp\big(-O(\omega(n)^9)\big) \cdot \exp\big(-O\big((\log \omega(n))^2\big)\big) \cdot \Omega\big((\log\log\log n)^{-\beta_2}\big),$$

where $\beta_1, \beta_2 > 0$ are constants. Thus, $\rho = \Omega\big((\log\log n)^{-1}\big)$, since $\omega(n) = \log\log\log\log n$. ◀

▶ Remark 14. We pause to mention that this delicate coupling for the activation of the components is not required when $\zeta = q$ and $q > 2$. In that regime, the random-cluster model is super-critical, so after the first $O(\log n)$ steps, the component structure is much simpler, with exactly one large component. On the other hand, when $\zeta = q$ and $q \in (1, 2]$ the model is critical, which, combined with the fact mentioned earlier that the percolation sub-step of the dynamics is also critical when $\zeta = q$, makes the analysis of the CM dynamics in this regime quite subtle.

2.3 Coupling to the same configuration

In the last phase of the coupling, suppose we start with two configurations X_0, Y_0 with the same component structure. We are still required to bound the number of steps until the same configuration is reached. The following lemma from [5] supplies the desired bound.

▶ **Lemma 15** ([5], Lemma 24). *Let $q > 1$, $\zeta > 0$ and let X_0, Y_0 be two random-cluster configurations with the same component structure. Then, there exists a coupling of the CM steps such that after $T = O(\log n)$ steps, $X_T = Y_T$ w.h.p.*

Combining the results for each of the phases of the coupling, we now prove Theorem 1.

Proof of Theorem 1. By Theorem 6, after $t_0 = O(\log n)$ steps, with probability $\Omega(1)$, we have $\mathcal{R}_1(X_{t_0}) = O(n^{4/3})$ and $\mathcal{R}_1(Y_{t_0}) = O(n^{4/3})$. If this is the case, Theorem 9 and Lemma 15 imply that there exists a coupling of the CM steps such that with probability $\Omega\left((\log\log n)^{-1}\right)$ after an additional $t_1 = O(\log n)$ steps, $X_{t_0+t_1} = Y_{t_0+t_1}$. Consequently, we obtain that $\tau_{\text{mix}}^{\text{CM}} = O(\log n \cdot \log\log n)$ as claimed. ◀

▶ Remark 16. The probability of success in Theorem 9, which governs the lower order term $O(\log\log n)$ in our mixing time bound, is controlled by our choice of the function $\omega(n)$ for the definition of "large components". By choosing $\omega(n)$ that goes to ∞ more slowly, we could improve our mixing time bound to $O(\log n \cdot g(n))$ where $g(n)$ is any function that tends to infinity arbitrarily slowly. However, it seems that new ideas are required to obtain a bound of $O(\log n)$ (matching the known lower bound). In particular, the fact that $\omega(n) \to \infty$ is crucially used in some of our proofs. Our specific choice of $\omega(n)$ yields the $O(\log n \cdot \log\log n)$ bound and makes our analysis cleaner.

3 Coupling to the same component structure: proof of Lemma 11

To prove Lemma 11, we use a local limit theorem to construct a two-step coupling of the CM dynamics that reaches two configurations with the same large component structure. The construction of Markov chain couplings using local limit theorems is not common (see [19] for another example), but it appears to be a powerful technique that may have other interesting applications. We provide next a brief introduction to local limit theorems.

3.1 Local limit theorem

Let $c_1 \leq \cdots \leq c_m$ be integers and for $i = 1, \ldots, m$ let X_i be the random variable that is equal to c_i with probability $r \in (0,1)$, and it is zero otherwise. Let us assume that $X_1, \ldots, X_m$ are independent random variables. Let $S_m = \sum_{i=1}^m X_i$, $\mu_m = \mathrm{E}[S_m]$ and $\sigma_m^2 = \mathrm{Var}(S_m)$. We say that a *local limit theorem* holds for S_m if for every integer $a \in \mathbb{Z}$:

$$\Pr[S_m = a] = \frac{1}{\sqrt{2\pi}\sigma_m} \exp\Big(- \frac{(a-\mu_m)^2}{2\sigma_m^2}\Big) + o(\sigma_m^{-1}). \tag{3}$$

We prove, under some conditions, a local limit theorem that applies to the random variables corresponding to the number of active vertices from small components. Recall that for an increasing positive function g and each integer $k \geq 0$, we defined the intervals

$$\mathcal{I}_k(g) = \left[\frac{\vartheta n^{2/3}}{2g(n)^{2^k}}, \frac{\vartheta n^{2/3}}{g(n)^{2^k}} \right],$$

where $\vartheta > 0$ is a fixed large constant.

▶ **Theorem 17.** *Let $c_1 \leq \cdots \leq c_m$ be integers, and suppose $X_1, ..., X_m$ are independent random variables such that X_i is equal to c_i with probability $r \in (0,1)$, and X_i is zero otherwise. Let $g : \mathbb{N} \to \mathbb{R}$ be an increasing positive function such that $g(m) \to \infty$ and $g(m) = o(\log m)$. Suppose $c_m = O\left(m^{2/3} g(m)^{-1}\right)$, $\sum_{i=1}^m c_i^2 = O\left(m^{4/3} g(m)^{-1/2}\right)$ and $c_i = 1$ for all $i \leq \rho m$, where $\rho \in (0,1)$ is independent of m. Let $\ell > 0$ be the smallest integer such that $m^{2/3} g(m)^{-2^\ell} = o(m^{1/4})$. If for all $1 \leq k \leq \ell$, we have $|\{i : c_i \in \mathcal{I}_k(g)\}| = \Omega(g(m)^{3 \cdot 2^{k-1}})$, then a local limit theorem holds for $S_m = \sum_{i=1}^m X_i$.*

Theorem 17 follows from a general local limit theorem proved in [21]; a proof is given in the full paper [7]. We next compile a number of (mostly standard) facts about the $G(n,p)$ random graph model which will be used in our proof of Lemma 11.

3.2 Random graphs estimates

We use $G \sim G(n,p)$ to denote a random graph G sampled from the standard $G(n,p)$ model, in which every edge appears independently with probability p. For a graph G, with a slight abuse of notation, let $L_i(G)$ denote the size of the i-th largest connected component in G, and let $\mathcal{R}_i(G) := \sum_{j \geq i} L_j(G)^2$; note that the same notation is used for the components of a random-cluster configuration, but it will always be clear from context which case is meant.

▶ **Lemma 18** ([19], Lemma 5.7)**.** *Let $I(G)$ denote the number of isolated vertices in G. If $np = O(1)$, then there exists a constant $C > 0$ such that $\Pr[I(G) > Cn] = 1 - O(n^{-1})$.*

▶ **Lemma 19** ([2], Lemma 2.16)**.** *If $np > 0$, we have $\mathrm{E}[\mathcal{R}_2(G)] = O\left(n^{4/3}\right)$.*

▶ **Lemma 20.** *Let $G \sim G\left(n, \frac{1+\varepsilon}{n}\right)$ with $\varepsilon = o(1)$. For any positive constant $\rho \leq 1/10$, there exist constants $C \geq 1$ and $c > 0$ such that if $\varepsilon^3 n \geq C$, then*

$$\Pr\left[|L_1(G) - 2\varepsilon n| > \rho\varepsilon n\right] = O(\exp(-c\varepsilon^3 n)).$$

For the next results, suppose that $G \sim G(n, \frac{1+\lambda n^{-1/3}}{n})$, where $\lambda = \lambda(n)$ may depend on n.

▶ **Lemma 21.** *If $|\lambda| = O(1)$, then $\mathrm{E}[\mathcal{R}_1(G)] = O\left(n^{4/3}\right)$.*

All the random graph facts stated so far can be either found in the literature, or follow directly from well-known results. The following lemmas are slightly more refined versions of similar results in the literature.

▶ **Lemma 22.** *Suppose* $|\lambda| = O(h(n))$ *and let* $B_h = n^{2/3}h(n)^{-1}$*, where* $h : \mathbb{N} \to \mathbb{R}$ *is a positive increasing function such that* $h(n) = o(\log n)$*. Then, for any* $\alpha \in (0,1)$ *there exists a constant* $C = C(\alpha) > 0$ *such that, with probability at least* α*,*

$$\sum_{j:L_j(G)\le B_h} L_j(G)^2 \le Cn^{4/3}h(n)^{-1/2}.$$

▶ **Lemma 23.** *Let* $S_B = \{j : B \le L_j(G) \le 2B\}$ *and suppose there exists a positive increasing function* g *such that* $g(n) \to \infty$*,* $g(n) = o(n^{1/3})$*,* $|\lambda| \le g(n)$ *and* $B \le n^{2/3}g(n)^{-2}$*. If* $B \to \infty$*, then there exists constants* $\delta_1, \delta_2 > 0$ *independent of* n *such that*

$$\Pr\left[|S_B| \le \frac{\delta_1 n}{B^{3/2}}\right] \le \frac{\delta_2 B^{3/2}}{n}.$$

Finally, the following corollary of Lemma 23 will also be useful. For a graph H, let $N_k(H, g)$ be the number of components of H whose sizes are in the interval $\mathcal{I}_k(g)$. We note that with a slight abuse of notation, for a random-cluster configuration X, we also use $N_k(X, g)$ for the number of connected components of X in $\mathcal{I}_k(g)$.

▶ **Lemma 24.** *Let* $m \in (n/2q, n]$ *and let* g *be an increasing positive function that such that* $g(n) = o(m^{1/3})$*,* $g(n) \to \infty$ *and* $|\lambda| \le g(m)$*. If* $H \sim G\big(m, \frac{1+\lambda m^{-1/3}}{m}\big)$*, there exists a constant* $b > 0$ *such that, with probability at least* $1 - O\left(g(n)^{-3}\right)$*,* $N_k(H, g) \ge bg(n)^{3\cdot 2^{k-1}}$ *for all* $k \ge 1$ *such that* $n^{2/3}g(n)^{-2^k} \to \infty$*.*

The proofs of Lemmas 20-24 are given in the full version of the paper [7].

3.3 Proof of Lemma 11

For a random-cluster configuration X, let $A(X)$ denote the random variable corresponding to the number of vertices activated by step (i) of the CM dynamics from X. We provide next the proof of Lemma 11.

Proof of Lemma 11. First, both $\{X_t\}$, $\{Y_t\}$ perform one independent CM step from the initial configurations X_0, Y_0. We start by establishing that X_1 and Y_1 preserve the structural properties assumed for X_0 and Y_0.

By assumption $\mathcal{R}_1(X_0) = O(n^{4/3})$, so Hoeffding's inequality implies that the number of activated vertices from X_0 is such that

$$A(X_0) \in I := \left[n/q - O(n^{2/3}), n/q + O(n^{2/3})\right]$$

with probability $\Omega(1)$. Then, the percolation step is distributed as a

$$G\left(A(X_0), \frac{1 + \lambda A(X_0)^{-1/3}}{A(X_0)}\right)$$

random graph, with $|\lambda| = O(1)$ with probability $\Omega(1)$. Conditioning on this event, from Lemma 18 we obtain that $I(X_1) = \Omega(n)$ w.h.p. Moreover, from Lemma 21 and Markov's inequality we obtain that $\mathcal{R}_1(X_1) = O(n^{4/3})$ with probability at least 99/100 and from Lemma 22 that $\widetilde{\mathcal{R}}_\omega(X_1) = O(n^{4/3}\omega(n)^{-1/2})$ also with probability at least 99/100.

We show next that X_1 and Y_1, in addition to preserving the structural properties of X_0 and Y_0, also have many connected components with sizes in certain carefully chosen intervals. This fact will be crucial in the design of our coupling. When $A(X_0) \in I$, by Lemmas 23 and 24 and a union bound, for all integer $k \ge 0$ such that $n^{2/3}\omega(n)^{-2^k} \to \infty$,

$N_k(X_1,\omega) = \Omega(\omega(n)^{3\cdot 2^{k-1}})$ w.h.p. (Recall, that $N_k(X_1,\omega)$ denotes the number of connected components of X_1 with sizes in the interval $\mathcal{I}_k(\omega)$.) We will also require a bound for the number of components with sizes in the interval

$$J = \left[\frac{cn^{2/3}}{\omega(n)^6}, \frac{2cn^{2/3}}{\omega(n)^6}\right],$$

where $c > 0$ is a constant such that J does not intersect any of the $\mathcal{I}_k(\omega)$'s intervals. Let W_X (resp., W_Y) be the set of components of X_1 (resp., Y_1) with sizes in the interval J. Lemma 23 then implies that for some positive constants δ_1, δ_2 independent of n,

$$\Pr\left[|W_X| \geq \delta_1 n\left(\frac{\omega(n)^6}{cn^{2/3}}\right)^{3/2}\right] \geq 1 - \frac{\delta_2}{n}\left(\frac{cn^{2/3}}{\omega(n)^6}\right)^{3/2} = 1 - O(\omega(n)^{-9}).$$

All the bounds above apply also to the analogous quantities for Y_1 with the same respective probabilities. Therefore, by a union bound, all these properties hold simultaneously for both X_1 and Y_1 with probability $\Omega(1)$. We assume that this is indeed the case and proceed to describe the second step of the coupling, in which we shall use each of the established properties for X_1 and Y_1.

Let C_X and C_Y be the set of components in X_1 and Y_1, respectively, with sizes larger than B_ω. (Recall that $B_\omega = n^{2/3}\omega(n)^{-1}$, where $\omega(n) = \log\log\log\log n$.) Since $\mathcal{R}_1(X_1) = O(n^{4/3})$, the total number of components in C_X is $O(\omega(n)^2)$; moreover, it follows from the Cauchy–Schwarz inequality that the total number of vertices in the components in C_X, denoted $\|C_X\|$, is $O(n^{2/3}\omega(n))$; the same holds for C_Y.

Without loss of generality, let us assume that $\|C_X\| \geq \|C_Y\|$. Let

$$\mathit{\Gamma} = \{C \subset W_Y : \|C_Y \cup C\| \geq \|C_X\|\},$$

and let $C_{\min} = \arg\min_{C\in\mathit{\Gamma}} \|C_Y \cup C\|$. In words, $C_{\min}$ is the smallest subset C of components of W_Y so that $\|C_Y \cup C\| \geq \|C_X\|$. Since every component in W_Y has size at least $cn^{2/3}\omega(n)^{-6}$ and $|W_Y| = \Omega(\omega(n)^9)$, the number of vertices in W_Y is $\Omega(n^{2/3}\omega(n)^3)$ and so $\mathit{\Gamma} \neq \emptyset$. In addition, the number components in $C_{\min}$ is $O(\omega(n)^9)$. Let $C'_Y = C_Y \cup C_{\min}$ and observe that the number of components in C'_Y is also $O(\omega(n)^9)$ and that

$$0 \leq \|C'_Y\| - \|C_X\| \leq 2cn^{2/3}\omega(n)^{-6}.$$

Note that $\|C_X\| - \|C_Y\|$ may be $\Omega(n^{2/3}\omega(n))$ (i.e., much larger than $\|C'_Y\| - \|C_X\|$). Hence, if all the components from C_Y and C_X were activated, the difference in the number of active vertices could be $\Omega(n^{2/3}\omega(n))$. This difference cannot be corrected by our coupling for the activation of the small components. We shall require instead that all the components from C'_Y and C_X are activated so that the difference is $O(n^{2/3}\omega(n)^{-6})$ instead.

We now describe a coupling of the activation sub-step for the second step of the CM dynamics. As mentioned, our goal is to design a coupling in which the same number of vertices are activated from each copy. If indeed $A(X_1) = A(Y_1)$, then we can choose an arbitrary bijective map φ between the activated vertices of X_1 and the activated vertices of Y_1 and use φ to couple the percolation sub-step. Specifically, if u and v were activated in X_1, the state of the edges $\{u, v\}$ in X_2 and $\{\varphi(u), \varphi(v)\}$ in Y_2 would be the same. This yields a coupling of the percolation sub-step such that X_2 and Y_2 agree on the subgraph update at time 1.

Suppose then that in the second CM step all the components in C_X and C'_Y are activated simultaneously. If this is the case, then the difference in the number of activated vertices is $d \leq 2cn^{2/3}\omega(n)^{-6}$. We will use a local limit theorem (i.e., Theorem 17) to argue that

there is a coupling of the activation of the remaining components in X_1 and Y_1 such that the total number of active vertices in both copies is the same with probability $\Omega(1)$. Since all the components in C_X and C'_Y are activated with probability $\exp(-O(\omega(n)^9))$, the overall success probability of the coupling will be $\exp(-O(\omega(n)^9))$.

Now, let $x_1, x_2, \ldots, x_m$ be the sizes of the components of X_1 that are not in C_X (in increasing order). Let $\hat{A}(X_1)$ be the random variable corresponding to the number of active vertices from these components. Observe that $\hat{A}(X_1)$ is the sum of m independent random variables, where the j-th variable in the sum is equal to x_j with probability $1/q$, and it is 0 otherwise. We claim that sequence $x_1, x_2, \ldots, x_m$ satisfies all the conditions in Theorem 17.

First, note that since the number of isolated vertices in X_1 is $\Omega(n)$, $m = \Theta(n)$ and so $x_m = O(m^{2/3}\omega(m)^{-1})$, $\sum_{i=1}^m x_i^2 = \tilde{R}_\omega(X_1) = O(m^{4/3}\omega(m)^{-1/2})$ and $x_i = 1$ for all $i \leq \rho m$, where $\rho \in (0,1)$ is independent of m. Moreover, since $N_k(X_1, \omega) \geq \Omega(\omega(n)^{3\cdot 2^{k-1}})$ for all $k \geq 1$ such that $n^{2/3}\omega(n)^{-2^k} \to \infty$,

$$|\{i : x_i \in \mathcal{I}_k(\omega)\}| = \Omega(\omega(m)^{3\cdot 2^{k-1}}).$$

Since $N_0(X_1, \omega) = \Omega(\omega(n)^{3/2})$, we also have

$$\sum_{i=1}^m x_i^2 \geq N_0(X_1, \omega) \cdot \frac{\vartheta^2 n^{4/3}}{4\omega(n)^2} = \Omega(m^{4/3}\omega(m)^{-1/2}).$$

Let $\mu_X = \mathrm{E}[\hat{A}(X_1)] = q^{-1}\sum_{i=1}^m x_i$ and let

$$\sigma_X^2 = \mathrm{Var}(\hat{A}(X_1)) = q^{-1}(1-q^{-1})\sum_{i=1}^m x_i^2 = \Theta(m^{4/3}\omega(m)^{-1/2}).$$

Hence, Theorem 17 implies that $\Pr[\hat{A}(X_1) = a] = \Omega\left(\sigma_X^{-1}\right)$ for any $a \in [\mu_X - \sigma_X, \mu_X + \sigma_X]$. Similarly, we get $\Pr[\hat{A}(Y_1) = a] = \Omega(\sigma_Y^{-1})$ for any $a \in [\mu_Y - \sigma_Y, \mu_Y + \sigma_Y]$, with $\hat{A}(Y_1)$, μ_Y and σ_Y defined analogously for $Y_1 \setminus C'_Y$. Note that $\mu_X - \mu_Y = O(n^{2/3}\omega(n)^{-6})$ and $\sigma_X, \sigma_Y = \Theta(n^{2/3}\omega(n)^{-1/4})$. Without loss of generality, suppose $\sigma_X < \sigma_Y$. Then for any $a \in [\mu_X - \sigma_X/2, \mu_Y + \sigma_X/2]$ and $d = O(n^{2/3}\omega(n)^{-6})$, we have

$$\min\left\{\Pr[\hat{A}(X_1) = a], \Pr[\hat{A}(Y_1) = a - d]\right\} = \min\left\{\Omega(\sigma_X^{-1}), \Omega(\sigma_Y^{-1})\right\} = \Omega(\sigma_Y^{-1}).$$

Hence, there exists a coupling $\mathbb{P}$ of $\hat{A}(X_1)$ and $\hat{A}(Y_1)$ so that $\mathbb{P}[\hat{A}(X_1) = a, \hat{A}(Y_1) = a - d] = \Omega(\sigma_Y^{-1})$ for all $a \in [\mu_X - \sigma_X/2, \mu_Y + \sigma_X/2]$. Therefore, there is a coupling of $\hat{A}(X_1)$ and $\hat{A}(Y_1)$ such that

$$\Pr[\hat{A}(X_1) - \hat{A}(Y_1) = d] = \Omega\left(\sigma_X/\sigma_Y\right) = \Omega(1).$$

Putting all these together, we deduce that $A(X_1) = A(Y_1)$ with probability $e^{-O(\omega(n)^9)}$. If this is the case, the edge re-sampling step is coupled bijectively (as described above) so that $\mathcal{S}_\omega(X_2) = \mathcal{S}_\omega(Y_2)$.

It remains for us to guarantee the additional desired structural properties of X_2 and Y_2, which follow straightforwardly from the random graph estimates we stated at the beginning of the section. First note that by Hoeffding's inequality, with probability $\Omega(1)$,

$$\left|A(X_1) - \frac{n}{q} - \frac{(q-1)|C_X|}{q}\right| = O(n^{2/3}).$$

Hence, in the percolation sub-step the active subgraph is replaced by

$$F \sim G\left(A(X_1), \frac{1 + \lambda A(X_1)^{-1/3}}{A(X_1)}\right),$$

where $|\lambda| = O(\omega(n))$ with probability $\Omega(1)$ since $|C_X| = O(n^{2/3}\omega(n))$. Conditioning on this event, since the components of F contribute to both X_2 and Y_2, Lemma 24 implies that w.h.p. $\hat{N}_k(2, \omega(n)) = \Omega(\omega(n)^{3\cdot 2^{k-1}})$ for all $k \geq 1$ such that $n^{2/3}\omega(n)^{-2^k} \to \infty$. Moreover, from Lemma 18 we obtain that $I(X_2) = \Omega(n)$ w.h.p. From Lemma 19 and Markov's inequality, we obtain that $\mathcal{R}_2(X_2) = O(n^{4/3})$ with probability at least 99/100 and from Lemma 22 that $\widetilde{\mathcal{R}}_\omega(X_2) = O(n^{4/3}\omega(n)^{-1/2})$ also with probability at least 99/100. All these bounds apply also to the analogous quantities for Y_2 with the same respective probabilities.

Finally, we derive the bound for $L_1(X_2)$ and $L_1(Y_2)$. First, notice $L_1(F)$ is stochastically dominated by $L_1(F')$, where

$$F' \sim G\Big(A(X_1), \frac{1 + |\lambda|A(X_1)^{-1/3}}{A(X_1)}\Big).$$

Under the assumption that $|\lambda| = O(\omega(n))$, if $|\lambda| \to \infty$, then Lemma 20 implies that $L_1(F') = O(|\lambda|A(X_1)^{2/3}) = O(n^{2/3}\omega(n))$ w.h.p.; otherwise, $|\lambda| = O(1)$ and by Lemma 21 and Markov's inequality, $L_1(F') = O(n^{2/3})$ with probability at least 99/100. Thus, $L_1(F) = O(n^{2/3}\omega(n))$ with probability at least 99/100. We also know that the largest inactivated component in X_1 has size less than $n^{2/3}\omega(n)^{-1}$, so $L_1(X_2) = O(n^{2/3}\omega(n))$ with probability at least 99/100. The same holds for Y_2. Therefore, by a union bound, all these properties hold simultaneously for both X_2 and Y_2 with probability $\Omega(1)$, as claimed. ◀

References

1 V. Beffara and H. Duminil-Copin. The self-dual point of the two-dimensional random-cluster model is critical for $q \geq 1$. *Probability Theory and Related Fields*, 153:511–542, 2012.

2 A. Blanca. *Random-cluster dynamics*. PhD thesis, UC Berkeley, 2016.

3 A. Blanca and R. Gheissari. Random-cluster dynamics on random regular graphs in tree uniqueness. *Communications in Mathematical Physics*, 2021.

4 A. Blanca, R. Gheissari, and E. Vigoda. Random-cluster dynamics in $\mathbb{Z}^2$: rapid mixing with general boundary conditions. *Annals of Applied Probability*, 30(1):418–459, 2020.

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

6 A. Blanca and A. Sinclair. Random-Cluster Dynamics in $\mathbb{Z}^2$. *Probability Theory and Related Fields*, 168:821–847, 2017.

7 Antonio Blanca, Alistair Sinclair, and Xusheng Zhang. The critical mean-field chayes-machta dynamics, 2021. `arXiv:2102.03004`.

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

9 L. Chayes and J. Machta. Graphical representations and cluster algorithms II. *Physica A*, 254:477–516, 1998.

10 H. Duminil-Copin, M. Gagnebin, M. Harel, I. Manolescu, and V. Tassion. Discontinuity of the phase transition for the planar random-cluster and Potts models with $q > 4$. *Annales de l'ENS*, 2016. To Appear.

11 H. Duminil-Copin, V. Sidoravicius, and V. Tassion. Continuity of the Phase Transition for Planar Random-Cluster and Potts Models with $1 \leq q \leq 4$. *Communications in Mathematical Physics*, 349(1):47–107, 2017.

12 C.M. Fortuin and P.W. Kasteleyn. On the random-cluster model I. Introduction and relation to other models. *Physica*, 57(4):536–564, 1972.

13 A. Galanis, D. Štefankovič, and E. Vigoda. Swendsen-Wang algorithm on the mean-field Potts model. *Proceedings of the 19th International Workshop on Randomization and Computation (RANDOM)*, pages 815–828, 2015.

14 T. Garoni. Personal communication, 2015.

15 R. Gheissari and E. Lubetzky. Quasi-polynomial mixing of critical two-dimensional random cluster models. *Random Structures & Algorithms*, 56(2):517–556, 2020.

16 R. Gheissari, E. Lubetzky, and Y. Peres. Exponentially slow mixing in the mean-field Swendsen-Wang dynamics. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1981–1988. SIAM, 2018.

17 H. Guo and M. Jerrum. Random cluster dynamics for the Ising model is rapidly mixing. In *Proceedings of the 28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1818–1827. SIAM, 2017.

18 D.A. Levin and Y. Peres. *Markov Chains and Mixing Times*. MBK. American Mathematical Society, 2017. URL: `https://books.google.com/books?id=f208DwAAQBAJ`.

19 Y. Long, A. Nachmias, W. Ning, and Y. Peres. A power law of order 1/4 for critical mean-field Swendsen-Wang dynamics. *Memoirs of the American Mathematical Society*, 232(1092), 2011.

20 Malwina Luczak and Tomasz Luczak. The phase transition in the cluster-scaled model of a random graph. *Random Structures & Algorithms*, 28(2):215–246, 2006.

21 A.B. Mukhin. Local limit theorems for lattice random variables. *Theory of Probability & Its Applications*, 36(4):698–713, 1992.

22 R.H. Swendsen and J.S. Wang. Nonuniversal critical dynamics in Monte Carlo simulations. *Physical Review Letters*, 58:86–88, 1987.

23 M. Ullrich. Swendsen-Wang is faster than single-bond dynamics. *SIAM Journal on Discrete Mathematics*, 28(1):37–48, 2014.