

Swendsen-Wang Algorithm on the Mean-Field Potts Model*

Andreas Galanis[†]

Daniel Štefankovič[‡]

Eric Vigoda[§]

November 9, 2017

Abstract

We study the q -state ferromagnetic Potts model on the n -vertex complete graph known as the mean-field (Curie-Weiss) model. We analyze the Swendsen-Wang algorithm which is a Markov chain that utilizes the random cluster representation for the ferromagnetic Potts model to recolor large sets of vertices in one step and potentially overcomes obstacles that inhibit single-site Glauber dynamics. Long et al. studied the case $q = 2$, the Swendsen-Wang algorithm for the mean-field ferromagnetic Ising model, and showed that the mixing time satisfies: (i) $\Theta(1)$ for $\beta < \beta_c$, (ii) $\Theta(n^{1/4})$ for $\beta = \beta_c$, (iii) $\Theta(\log n)$ for $\beta > \beta_c$, where β_c is the critical temperature for the ordered/disordered phase transition. In contrast, for $q \geq 3$ there are two critical temperatures $0 < \beta_u < \beta_{rc}$ that are relevant. We prove that the mixing time of the Swendsen-Wang algorithm for the ferromagnetic Potts model on the n -vertex complete graph satisfies: (i) $\Theta(1)$ for $\beta < \beta_u$, (ii) $\Theta(n^{1/3})$ for $\beta = \beta_u$, (iii) $\exp(n^{\Omega(1)})$ for $\beta_u < \beta < \beta_{rc}$, and (iv) $\Theta(\log n)$ for $\beta \geq \beta_{rc}$. These results complement refined results of Cuff et al. on the mixing time of the Glauber dynamics for the ferromagnetic Potts model.

Keywords: mean-field Ferromagnetic Potts model, Curie-Weiss model, Swendsen-Wang algorithm, phase transitions.

*An extended abstract of this paper appeared in the proceedings of RANDOM/APPROX 2015.

[†]University of Oxford, Wolfson Building, Parks Road, Oxford, OX1 3QD, UK. andreas.galanis@cs.ox.ac.uk. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) ERC grant agreement no. 334828. The paper reflects only the authors' views and not the views of the ERC or the European Commission. The European Union is not liable for any use that may be made of the information contained therein.

[‡]Department of Computer Science, University of Rochester, Rochester, NY 14627. stefanko@cs.rochester.edu. Research supported in part by NSF grant CCF-1318374.

[§]School of Computer Science, Georgia Institute of Technology, Atlanta GA 30332. vigoda@cc.gatech.edu. Research supported in part by NSF grant CCF-1217458.

1 Introduction

The mixing time of Markov chains is of critical importance for simulations of statistical physics models. It is especially interesting to understand how phase transitions in these models manifest in the behavior of the mixing time; these connections are the topic of this paper.

We study the q -state ferromagnetic Potts model. In the following definition the case $q = 2$ corresponds to the Ising model and $q \geq 3$ is the Potts model. For a graph $G = (V, E)$ the configurations of the model are assignments $\sigma : V \rightarrow [q]$ of spins to vertices; denote by Ω the set of all configurations. The model is parameterized by $\beta > 0$, known as the (inverse) temperature. For a configuration $\sigma \in \Omega$ let $m(\sigma)$ be the number of edges in E that are monochromatic under σ and let its weight be $w(\sigma) = \exp(\beta m(\sigma))$. Then the Gibbs distribution μ is defined as follows: for $\sigma \in \Omega$, $\mu(\sigma) = w(\sigma)/Z(\beta)$, where $Z(\beta) = \sum_{\sigma \in \Omega} w(\sigma)$ is the normalizing constant, known as the partition function.

A useful feature for studying the ferromagnetic Potts model is its alternative formulation known as the random-cluster model. Here configurations are subsets of edges and the weight of such a configuration $S \subseteq E$ is

$$w(S) = p^{|S|} (1-p)^{|E \setminus S|} q^{k(S)},$$

where $p = 1 - \exp(-\beta)$ and $k(S)$ is the number of connected components in the graph $G' = (V, S)$ (isolated vertices do count). The corresponding partition function $Z_{rc} = \sum_{S \subseteq E} w(S)$ satisfies $Z_{rc} = (1-p)^{|E|} Z$.

The focus of this paper is the Curie-Weiss model which in computer science terminology is the n -vertex complete graph $G = (V, E)$. The interest in this model is that it allows more detailed results and these results are believed to extend to other graphs of particular interest such as random regular graphs. For convenience we parameterize the model in terms of a constant $B > 0$ such that the Gibbs distribution is as follows:

$$\mu(\sigma) = \frac{1}{Z(\beta)} (1 - B/n)^{-m(\sigma)}. \quad (1)$$

(Note that $\beta = -\ln(1 - B/n) \sim B/n$ for large n .) The following critical points $\mathfrak{B}_u < \mathfrak{B}_o < \mathfrak{B}_{rc}$ for the parameter B are well-studied¹ and relevant to our study of the Potts model on the complete graph:

$$\mathfrak{B}_u = \sup \left\{ B \geq 0 \mid \frac{B-z}{B+(q-1)z} \neq e^{-z} \text{ for all } z > 0 \right\} = \min_{z \geq 0} \left\{ z + \frac{qz}{e^z - 1} \right\}, \quad (2)$$

$$\mathfrak{B}_o = \frac{2(q-1) \ln(q-1)}{q-2}, \quad \mathfrak{B}_{rc} = q. \quad (3)$$

These thresholds correspond to the critical points for the infinite Δ -regular tree \mathbb{T}_Δ and random Δ -regular graphs by taking appropriate limits as $\Delta \rightarrow \infty$. More specifically, if $B(\Delta)$ is a threshold on \mathbb{T}_Δ or the random Δ -regular graph then $\lim_{\Delta \rightarrow \infty} \Delta(B(\Delta) - 1)$ is the corresponding threshold in the Curie-Weiss model. In this perspective, \mathfrak{B}_u corresponds to the uniqueness/non-uniqueness threshold on \mathbb{T}_Δ ; \mathfrak{B}_o corresponds to the ordered/disordered phase transition; and \mathfrak{B}_{rc} was conjectured by Häggström to correspond to a second uniqueness/non-uniqueness threshold for the random-cluster model on \mathbb{T}_Δ with periodic boundaries (in particular, he conjectured that non-uniqueness holds iff $B \in (\mathfrak{B}_u, \mathfrak{B}_{rc})$). For a detailed exposition of these critical points we refer the reader to [11] (see

¹ \mathfrak{B}_o is β_c in [10, Equation (3.1)] and \mathfrak{B}_u is equivalent to β_s in [11, Equation (1.1)] under the parametrization $z = B(qx - 1)/(q - 1)$. We follow the convention of counting monochromatic edges [10] as opposed to counting monochromatic pairs of vertices [11]; hence our thresholds are larger than those in [11] by a factor of 2.

also [12] for their relevance for random regular graphs). We should finally remark that in the case of the Ising model ($q = 2$), the three points $\mathfrak{B}_u, \mathfrak{B}_o, \mathfrak{B}_{rc}$ coincide.

The Glauber dynamics is a classical tool for studying the Gibbs distribution. This is the class of Markov chains with “local” transitions that update the configuration at a randomly chosen vertex and which are designed so that the stationary distribution is the Gibbs distribution. The limitation of local Markov chains, such as the Glauber dynamics, is that they are typically slow to converge at low temperatures (large B). The Swendsen-Wang algorithm is a more sophisticated Markov chain that utilizes the random cluster representation of the Potts model to potentially overcome bottlenecks that obstruct the simpler Glauber dynamics.

Specifically, the Swendsen-Wang algorithm is a Markov chain (X_t) whose transitions $X_t \rightarrow X_{t+1}$ are as follows. From a configuration $X_t \in \Omega$:

- Let M be the set of monochromatic edges in X_t .
- *Percolation step*: for each edge $e \in M$, keep it independently with probability B/n . Let M' denote the set of the remaining monochromatic edges.
- *Coloring step*: in the graph (V, M') , independently for each connected component, choose a color uniformly at random from $[q]$ and assign to all vertices in that component the chosen color. Let X_{t+1} denote the resulting spin configuration.

It is a standard fact that the chain is reversible with respect to the Gibbs distribution μ (and thus converges to it). We will be interested in the mixing time T_{mix} of the chain, which is defined as the number of steps from the worst initial state to get within total variation distance $1/4$ of the distribution μ .

For the Swendsen-Wang algorithm for the ferromagnetic Ising model ($q = 2$), Cooper et al. [8] showed for the complete graph with n vertices that the mixing time satisfies $T_{\text{mix}} = n^{1/2+o(1)}$ for all temperatures except for $\beta = \beta_c$, where β_c is the uniqueness/non-uniqueness threshold. Long et al. [19] showed more refined results for the complete graph establishing that the mixing time is $\Theta(1)$ for $\beta < \beta_c$, $\Theta(n^{1/4})$ for $\beta = \beta_c$, and $\Theta(\log n)$ for $\beta > \beta_c$. For square boxes of \mathbb{Z}^2 , Ullrich [26, 27] proved that the mixing time of Swendsen-Wang is polynomial for all temperatures (building upon results for the Glauber dynamics by Martinelli and Olivieri [21, 22] and Lubetzky and Sly [20]). Very recently, Guo and Jerrum [16] showed that the mixing time of Swendsen-Wang is polynomial for any graph G for all temperatures.

For the Swendsen-Wang algorithm for the ferromagnetic Potts model ($q \geq 3$), it has been demonstrated that the mixing time can be of order $\exp(n^{\Omega(1)})$ at the ordered/disordered phase transition point (*phase coexistence*). In particular, Gore and Jerrum [15] showed for the complete graph that the mixing time is $\exp(\Omega(\sqrt{n}))$ at the critical point $B = \mathfrak{B}_o$. Similar slow mixing results have been established for other classes of graphs at the analogous critical point: Cooper and Frieze [9] showed this for $G(n, p)$ when $p = \Omega(n^{-1/3})$, Galanis et al. [12] for random Δ -regular graphs when $q \geq 2\Delta/\log \Delta$, and Borgs et al. [5, 6] for the d -dimensional integer lattice for $q \geq 25$. For square boxes of \mathbb{Z}^2 , Ullrich [26, 27] proves polynomial mixing time at all temperatures except criticality building upon the results of Beffara and Duminil-Copin [2]. On the torus $(\mathbb{Z}/n\mathbb{Z})^2$, Gheissari and Lubetzky [13] recently showed the following bounds on the mixing time at criticality: polynomial upper bound for $q = 3$, quasi-polynomial upper bound for $q = 4$ and exponential lower bound for $q > 4$.

In this paper, we study the mixing time of the Swendsen-Wang dynamics for the ferromagnetic Potts model on the complete graph. Previously, Cuff et al. [11] had detailed the mixing time of the Glauber dynamics for the ferromagnetic Potts model on the complete graph (their results

are significantly more precise than what we state here for convenience): $\Theta(n \log n)$ for $B < \mathfrak{B}_u$, $\exp(\Omega(n))$ for $B > \mathfrak{B}_u$, and $\Theta(n^{4/3})$ mixing time for $B = \mathfrak{B}_u$ (and a scaling window of $O(n^{-2/3})$ around \mathfrak{B}_u).

Our main result is a complete classification of the mixing time of the Swendsen-Wang dynamics on the complete graph when the parameter B is a constant independent of n .

Theorem 1. *For all integer $q \geq 3$, the mixing time T_{mix} of the Swendsen-Wang algorithm on the n -vertex complete graph satisfies:*

1. For all $B < \mathfrak{B}_u$, $T_{\text{mix}} = \Theta(1)$.
2. For $B = \mathfrak{B}_u$, $T_{\text{mix}} = \Theta(n^{1/3})$.
3. For all $\mathfrak{B}_u < B < \mathfrak{B}_{rc}$, $T_{\text{mix}} = \exp(n^{\Omega(1)})$.
4. For all $B \geq \mathfrak{B}_{rc}$, $T_{\text{mix}} = \Theta(\log n)$.

In an independent work, Blanca and Sinclair [3] analyze a closely related chain to the Swendsen-Wang dynamics, known as the Chayes-Machta dynamics, which is also suitable for sampling random cluster configurations (works more generally for $q \geq 1$ with $q \in \mathbb{R}$). They provide an analogue of Theorem 1, though their analysis excludes the critical points $B = \mathfrak{B}_u$ and $B = \mathfrak{B}_{rc}$. Very recently, Gheissari, Lubetzky, and Peres [14] improved the lower bound on the mixing time in the window $\mathfrak{B}_u < B < \mathfrak{B}_{rc}$ to $\exp(\Omega(n))$, both for the Swendsen-Wang and the Chayes-Machta dynamics.

In the following section, we give an overview of our proof approach. First, we discuss the critical points $\mathfrak{B}_u, \mathfrak{B}_o, \mathfrak{B}_{rc}$ in more detail. Then, we present a function F which captures a simplified view of the Swendsen-Wang dynamics, and then we connect the behavior of F with the critical points. We also present in Section 2 a high-level sketch of the proof of Theorem 1. In Section 4, we collect facts for the $G(n, c/n)$ random graph which will be relevant for analyzing one step of the Swendsen-Wang algorithm. In Section 5, we prove the slow mixing result (Part 3 of Theorem 1). We then prove the rapid mixing results for $B > \mathfrak{B}_{rc}$ in Section 7, for $B = \mathfrak{B}_{rc}$ in Section 8, for $B < \mathfrak{B}_u$ in Section 10, and for $B = \mathfrak{B}_u$ in Section 11.

2 Proof Approach

2.1 Critical Points for Phase Transitions

In this section, we review the thresholds $\mathfrak{B}_u, \mathfrak{B}_o, \mathfrak{B}_{rc}$ for the mean-field Potts model and their connections to the critical points of the partition function which will be relevant later. The reader is referred to [4, 10] for further details ([4] also applies to the random-cluster model).

We first need to introduce some notation for the complete graph $G = (V, E)$ with n vertices. For a configuration $\sigma : V \rightarrow [q]$ and a color $i \in [q]$, let $\alpha_i(\sigma)$ be the fraction of vertices with color i in σ , i.e., $\alpha_i(\sigma) = |\{v \in V : \sigma(v) = i\}|/n$. We denote by $\alpha(\sigma)$ the vector $(\alpha_1(\sigma), \dots, \alpha_q(\sigma))$, and refer to it as the *phase* of σ . There are $q + 1$ phases that are most relevant, the *uniform* phase $\mathbf{u} := (1/q, \dots, 1/q)$ and the q permutations of the *majority* phase $\mathbf{m} := (a, b, \dots, b)$, for some appropriate $a > 1/q$ and b given by $a + (q - 1)b = 1$. Roughly, these phases correspond to the configurations that have dominant contribution to the partition function.

More precisely, for a q -dimensional probability vector α , let Ω^α be the set of configurations σ

whose phase is α .² Let

$$Z^\alpha = \sum_{\sigma \in \Omega^\alpha} w(\sigma) \text{ and } \Psi(\alpha) := \lim_{n \rightarrow \infty} \frac{1}{n} \ln Z^\alpha.$$

To simplify the formulas, it turns out that it is enough to consider the following one-dimensional version of Ψ corresponding to configurations where one color has density α and the remaining colors have density β where $\alpha + (q-1)\beta = 1$. Namely, let

$$\Psi_1(\alpha) := \Psi(\alpha, \beta, \dots, \beta) = \Psi\left(\alpha, \frac{1-\alpha}{q-1}, \dots, \frac{1-\alpha}{q-1}\right).$$

It is not hard to see that Z^α is given by $\binom{n}{\alpha_1 n, \dots, \alpha_q n} (1 - B/n)^{-\sum_{i \in [q]} \binom{\alpha_i n}{2}}$, so using Stirling's approximation we obtain the explicit expression

$$\Psi_1(\alpha) = -\alpha \ln \alpha - (1-\alpha) \ln \frac{1-\alpha}{q-1} + \frac{B}{2} \left(\alpha^2 + \frac{(1-\alpha)^2}{q-1} \right). \quad (4)$$

With these definitions, we next relate the thresholds $\mathfrak{B}_u, \mathfrak{B}_o, \mathfrak{B}_{rc}$ to the critical points/local maxima of Ψ_1 . Depending on the value of B , there are two points that are relevant, $u = 1/q$ and $a > 1/q$, where a is a critical point of Ψ_1 and hence satisfies³

$$\ln \frac{(q-1)a}{1-a} = B \frac{qa-1}{q-1}. \quad (5)$$

The following folklore lemma illustrates the relevant thresholds, see also Figure 1. For completeness, we give the proof in Section 3.1.

Lemma 2. *Let $q \geq 3$. For the function Ψ_1 ,*

1. *For $B < \mathfrak{B}_u$, Ψ_1 has a unique local maximum, at $u = 1/q$, and there are no other critical points of Ψ_1 .*
2. *For $B = \mathfrak{B}_u$, Ψ_1 has two critical points, at $u = 1/q$ and $a > 1/q$ (satisfying (5)). Of these, $u = 1/q$ is the only local maximum of Ψ_1 .*
3. *For $\mathfrak{B}_u < B < \mathfrak{B}_{rc}$, Ψ_1 has two local maxima, at $u = 1/q$ and $a > 1/q$ (satisfying (5)). Further,*
 - *For $B \in (\mathfrak{B}_u, \mathfrak{B}_o)$, u is the only global maximum of Ψ_1 .*
 - *For $B = \mathfrak{B}_o$, u and a are the global maxima of Ψ_1 .*
 - *For $B \in (\mathfrak{B}_o, \mathfrak{B}_{rc})$, a is the only global maximum of Ψ_1 .*
4. *For $B \geq \mathfrak{B}_{rc}$, Ψ_1 has one local maximum in the interval $[1/q, 1]$, at a point $a > 1/q$ (satisfying (5)).*

²Technically, for integrality reasons, Ω^α are the configurations σ whose phase is within $O(1/n)$ from α . This does not have any effect in the subsequent asymptotic considerations.

³Such a critical point $a > 1/q$ exists when $B \geq \mathfrak{B}_u$ (see Lemma 7). In the regime $\mathfrak{B}_u < B < \mathfrak{B}_{rc}$ there are two critical points of $\Psi_1(a)$ with value $a > 1/q$; the relevant value of a is then given by the point where $\Psi_1(\alpha)$ has a local maximum, see Lemma 2 for details and Figure 1 for a depiction.

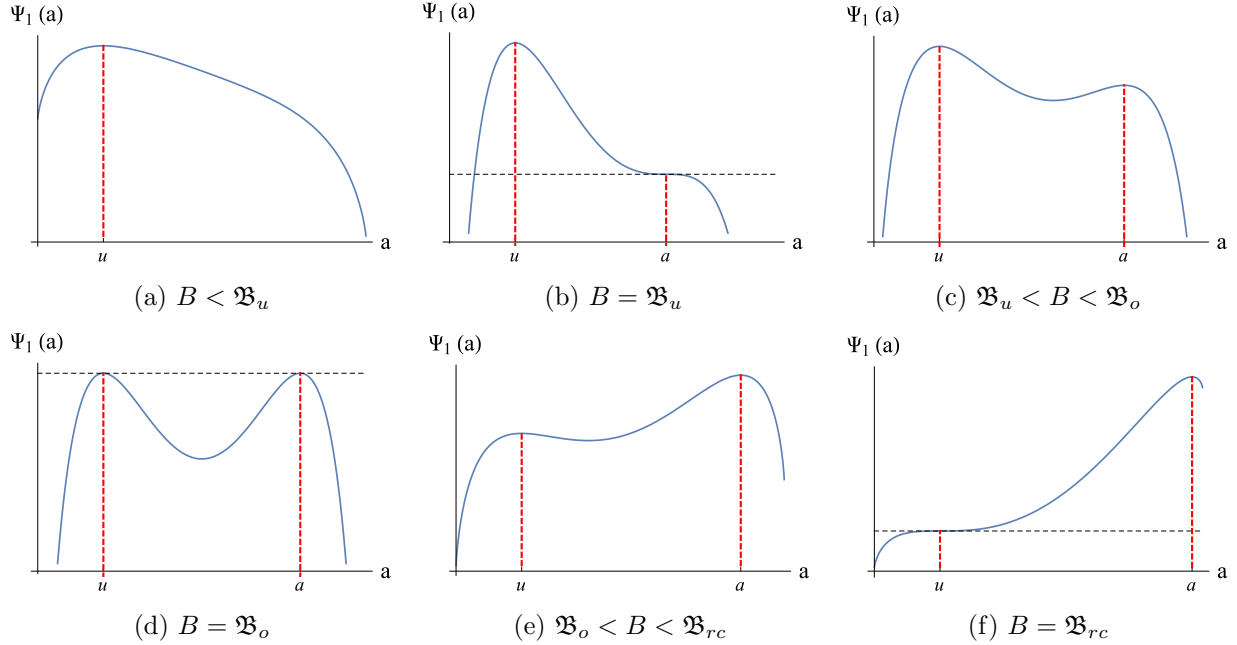


Figure 1: The function Ψ_1 (free energy) plotted in different regimes of B (defined in (4)). The critical points $\mathfrak{B}_u, \mathfrak{B}_o, \mathfrak{B}_{rc}$ are given by (2) and (3). In the regime $B < \mathfrak{B}_u$ (figure 1a), the function Ψ_1 has a unique local maximum at the disordered phase. At $B = \mathfrak{B}_u$ (figure 1b), the function Ψ_1 has a saddle point at the ordered phase. In the regime $\mathfrak{B}_u < B < \mathfrak{B}_{rc}$ (figures 1c, 1d and 1e) the function Ψ_1 has two local maxima; these are both global maxima iff $B = \mathfrak{B}_o$. In the regime $B \geq \mathfrak{B}_{rc}$ (figure 1f), the function Ψ_1 has a unique local maximum in the interval $[1/q, 1]$.

While we will not need the following fact explicitly in our arguments, we remark for the sake of completeness that the local maxima of the multivariable function Ψ correspond to the local maxima of the function Ψ_1 as follows. The phases where Ψ can have a local maximum is the uniform phase $\mathbf{u} = (1/q, \dots, 1/q)$ and the q permutations of the majority phase $\mathbf{m} = (a, b, \dots, b)$, where $a > 1/q$ is as in Lemma 2 and b is given by $a + (q-1)b = 1$. More precisely, \mathbf{u} is a local maximum of Ψ iff $u = 1/q$ is a local maximum of Ψ_1 , the majority phase \mathbf{m} is a local maximum of Ψ iff a is a local maximum of Ψ_1 , and there are no other local maxima of Ψ . Both \mathbf{u} and \mathbf{m} are global maxima of Ψ only at the point $B = \mathfrak{B}_o$.

2.2 Connections to Simplified Swendsen-Wang

The following function⁴ from $[1/q, 1]$ to $[0, 1]$ will capture the behavior of the Swendsen-Wang algorithm. Namely, let

$$F(z) := \frac{1}{q} + \left(1 - \frac{1}{q}\right)zx, \quad (6)$$

where $x = 0$ for $z \leq 1/B$ and for $z > 1/B$, $x \in (0, 1]$ is the (unique) solution of

$$x + \exp(-zBx) = 1. \quad (7)$$

⁴The argument of F will typically be the density of the largest color class — we could have extended the domain of the function F to be the interval $[0, 1]$ by further defining the value of F in the interval $[0, 1/B)$ to be $1/q$.

The function F captures the size of the largest color class when there is a single heavy color where heavy means that the color class is supercritical in the percolation step of the Swendsen-Wang process. Hence after the percolation step this heavy color will have a giant component and the other color classes will all be broken into small components. So say initially the one heavy color has size zn for $1/B < z < 1$ and let's consider its size after one step of the Swendsen-Wang dynamics. After the percolation step, this heavy color will have a giant component of size roughly xzn (where x is as in (7)) and all other components will be of size $O(\log n)$. Then, a $1/q$ fraction of the small components will be recolored the same as the giant component, and hence the size of the largest color class will be (roughly) $nF(z)$ after this one step of the Swendsen-Wang dynamics.

Our next goal is to tie together the functions F and Ψ_1 so that we can relate the behavior of the Swendsen-Wang dynamics with the underlying phase transitions of the model. We first need some terminology. A critical point a of a function $f : \mathbb{R} \rightarrow \mathbb{R}$ is a *hessian maximum* if the second derivative of f at a is negative (this is a sufficient condition for a to be a local maximum). For an integer $n \geq 1$, we will denote by $f^{(n)}$ the n -th iterate of the function f . A fixpoint a of f is *attractive* if there exists $\delta > 0$ such that for all $x \in (a - \delta, a + \delta)$ it holds that $f^{(n)}(x) \rightarrow a$; it is *repulsive* otherwise. The fixpoint a is *jacobian attractive* if $|F'(a)| < 1$; this is a sufficient condition for a to be attractive. The fixpoint a is *jacobian repulsive* if $|F'(a)| > 1$; this is a sufficient condition for a to be repulsive.

Our first lemma connects the local maxima of Ψ_1 with the attractive fixpoints of F (we restrict our attention to the interval $[1/q, 1]$ since the function F will be considered only in this interval).

Lemma 3. *In the interval $[1/q, 1]$, the hessian maxima of Ψ_1 correspond to jacobian attractive fixpoints of F .*

Lemma 3 is proved in Section 3.2. A relevant fact we should remark here and we will prove later is that, in the half-open interval $(1/q, 1]$, the critical points of Ψ_1 correspond to fixpoints of F (see Lemma 9); this actually holds for the left endpoint $1/q$ as well but only when $B \leq \mathfrak{B}_{rc}$ (for $B > \mathfrak{B}_{rc}$, $1/q$ is a critical point of Ψ_1 but not a fixpoint of F , see Lemma 10).

The second lemma studies the behavior of F around the fixpoints and it is the main tool for proving Theorem 1. Recall the earlier discussion of the uniform vector $\mathbf{u} := (1/q, \dots, 1/q)$ and the q permutations of the majority phase $\mathbf{m} := (a, b, \dots, b)$, where $a > 1/q$ is as in Lemma 2. The following lemma (proved in Section 3.3) provides some basic intuition about the proof of Theorem 1, as we shall explain shortly. A depiction of the various regimes is given in Figure 2.

Lemma 4. *Let $q \geq 3$. For the function F ,*

1. *For $B < \mathfrak{B}_u$, $u = 1/q$ is the unique fixpoint and it is jacobian attractive.*
2. *For $B = \mathfrak{B}_u$, there are 2 fixpoints: u and a . Of these, only u is (jacobian) attractive. The fixpoint a is repulsive but not jacobian repulsive.*
3. *For $\mathfrak{B}_u < B < \mathfrak{B}_{rc}$ there are 2 jacobian attractive fixpoints: u and a .*
4. *For $B = \mathfrak{B}_{rc}$, there are 2 fixpoints: u and a . The fixpoint u is jacobian repulsive, while the fixpoint a is jacobian attractive.*
5. *For $B > \mathfrak{B}_{rc}$, a is the only fixpoint and it is jacobian attractive.*

The reason that u abruptly changes from a jacobian attractive fixpoint ($B < \mathfrak{B}_{rc}$) to a jacobian repulsive fixpoint ($B = \mathfrak{B}_{rc}$) stems from the fact that in the regime $B < \mathfrak{B}_{rc}$, F is constant in a small neighborhood around $1/q$ (precisely, in the interval $[1/q, 1/B]$), which is no longer the case for $B = \mathfrak{B}_{rc}$.

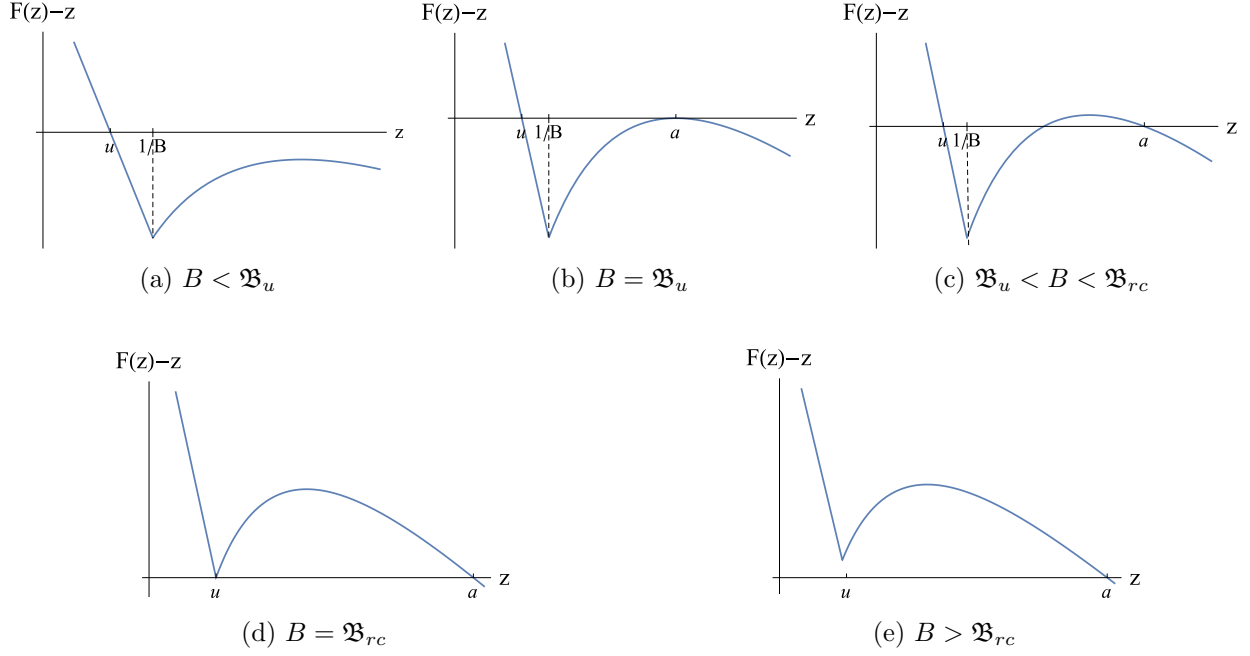


Figure 2: The drift function $F(z) - z$, where F is defined by (6), (7). The critical points $\mathfrak{B}_u, \mathfrak{B}_o, \mathfrak{B}_{rc}$ are given by (2) and (3). In the regime $B < \mathfrak{B}_u$ (figure 2a), the function F has a unique attractive fixpoint at the disordered phase. At $B = \mathfrak{B}_u$ (figure 2b), F also has a (non-jacobian) *repulsive* fixpoint at the ordered phase. In the regime $\mathfrak{B}_u < B < \mathfrak{B}_{rc}$ (figures 2c), F has attractive fixpoints at the ordered and disordered phases. At $B = \mathfrak{B}_{rc}$ (figure 2d), the disordered phase is no longer attractive; it is jacobian repulsive. Finally, in the regime $B > \mathfrak{B}_{rc}$ (figure 2e), the function F has a unique attractive fixpoint at the ordered phase.

2.3 Proof Sketches

We explain the high-level proof approach for the various parts of Theorem 1 before presenting the detailed proofs in subsequent sections.

Slow mixing for $B \in (\mathfrak{B}_u, \mathfrak{B}_{rc})$: The main idea is that the function F has 2 attractive fixpoints (see Lemma 4). At least one of the corresponding phases, \mathbf{u} or \mathbf{m} , is a global maximum for Ψ . Consider the other phase, say it is \mathbf{u} for concreteness. Consider the local ball around \mathbf{u} , these are configurations that are close in ℓ_∞ distance from \mathbf{u} . The key is that since $u = 1/q$ is an attractive fixpoint for F , if the initial state is in this local ball around \mathbf{u} then with very high probability after one step of the Swendsen-Wang dynamics it will still be in the local ball (see Lemma 22, and Lemma 23 for the analogous lemma for \mathbf{m}). The result then follows since one needs to sample from the local ball around the phase which corresponds to the global maximum of Ψ to get close to the stationary distribution. The full argument is given in Section 5.

Fast mixing for $B > \mathfrak{B}_{rc}$: For a configuration σ and spin i , say the color class is heavy if the number of vertices with spin i is $> n/B$ and light if it is $< n/B$. If a color class is heavy then it is supercritical for the percolation step of Swendsen-Wang and hence there will be a giant component. The key is that for any initial state X_0 , with constant probability, the largest components from all of the colors will choose the same new color and consequently there will be only one heavy color class and the other $q - 1$ colors will be light. Hence we can assume there is one heavy color class and $q - 1$ light color classes, and then the function F suitably describes the size of the largest color class

during the evolution of the Swendsen-Wang dynamics. Since the only fixpoint of F corresponds to the majority phase \mathbf{m} , after $O(\log n)$ steps we'll be close to \mathbf{m} – the difference will be due to the stochastic nature of the process. The remaining bit of the proof is then to define a coupling for two chains (X_t, Y_t) whose initial states X_0, Y_0 are close to \mathbf{m} so that after $T = O(\log n)$ steps we have that $X_T = Y_T$ (this latter part is fairly standard). The proof of the upper bound on the mixing time is given in Section 7; the lower bound on the mixing time is proved in Section 9.

Fast mixing for $B = \mathfrak{B}_{rc}$: The basic outline is similar to the $B > \mathfrak{B}_{rc}$ case except here the argument is more intricate when the heaviest color lies in the scaling window (for the onset of a giant component). We need a more involved argument that we get away from initial configurations that are close to the uniform phase; informally, the uniform fixpoint of F is jacobian repulsive, so an initial displacement increases geometrically by a constant factor. The proof of the upper bound on the mixing time is given in Section 8; the lower bound on the mixing time is proved in Section 9.

Fast mixing for $B < \mathfrak{B}_u$: Here the argument is similar to the $B > \mathfrak{B}_{rc}$ case; namely, the evolution of the density of the largest color class is captured by the iterates of the function F . Now, the only fixpoint of F corresponds to the uniform phase \mathbf{u} , so after $O(\log n)$ steps the chain will get close to \mathbf{u} . In fact, this bound can now be improved to $O(1)$ steps: once the largest color class reaches density $< 1/B$ (which happens in $O(1)$ steps), then in the next step the chain jumps close to $1/q$, i.e., we get close to the uniform phase *abruptly*; this is the reason that the mixing time for $B < \mathfrak{B}_u$ is $O(1)$ rather than $O(\log n)$. Once we are close to the uniform phase, we can then adapt a symmetry argument of [19] to show that we can couple two copies of the SW chain in one more step. The details can be found in Section 10.

Fast mixing for $B = \mathfrak{B}_u$: This is the most difficult part. As in the $B > \mathfrak{B}_{rc}$ case with constant probability there will be at most one heavy color class after one step. We then track the evolution of the size of the heavy color class. The difficulty arises because the size of the component does not decrease in expectation at the majority fixpoint. However variance moves the size of the component into a region where the size of the component decreases in expectation. The formal argument uses a carefully engineered potential function that decreases because of the variance (the function is concave around the fixpoint) and expectation (the function is increasing) of the size of the largest color class, see Section 11.

3 Phases of the Gibbs distribution and stability analysis of fixpoints of F

3.1 Analysis of the local maxima of Ψ_1 : Proof of Lemma 2

In this section, we analyze the critical points/local maxima of Ψ_1 and prove Lemma 2.

The following formulas for the derivatives of Ψ_1 will be useful:

$$\Psi_1'(\alpha) = -\ln \frac{(q-1)\alpha}{1-\alpha} + B \frac{q\alpha-1}{q-1}, \quad \Psi_1''(\alpha) = B \frac{q}{q-1} - \frac{1}{\alpha(1-\alpha)}. \quad (8)$$

Recall that at a critical a of Ψ_1 it holds that $\Psi_1'(a) = 0$ and hence a satisfies

$$\ln \frac{(q-1)a}{1-a} = B \frac{qa-1}{q-1}. \quad (5)$$

We will need the following bound on the critical points of Ψ_1 in the interval $(1/q, 1]$.

Lemma 5. *Let $a > 1/q$ be a critical point of Ψ_1 and $b = (1-a)/(q-1)$. Then $aB > 1$ and $bB < 1$.*

Proof. Since $a > 1/q$, there is $z > 0$ such that $a = (z + 1)/(z + q)$. Equation (5) becomes

$$\ln(1 + z) = \frac{zB}{z + q}. \quad (9)$$

Then, using (9), we have that

$$\begin{aligned} aB &= \frac{B(z + 1)}{z + q} = (1 + 1/z) \ln(1 + z) > 1, \\ bB &= (1 - a)B/(q - 1) = \frac{B}{z + q} = \frac{1}{z} \ln(1 + z) < 1. \end{aligned}$$

where the inequalities hold for any $z > 0$. This finishes the proof. \square

Lemma 6. *Let $B > \mathfrak{B}_u$. A critical point $a > 1/q$ of Ψ_1 has non-zero second derivative.*

Proof. For the sake of contradiction, let $a > 1/q$ be a critical point of Ψ_1 such that $\Psi_1''(a) = 0$. Using (8), $\Psi_1''(a) = 0$ yields that $1/q = 1 - Ba(1 - a)$. Plugging the value of q into (5) we obtain

$$\ln \frac{Ba^2}{1 - Ba(1 - a)} = \frac{Ba - 1}{a}. \quad (10)$$

Let $w = B - 1/a$. Since $a > 1/q$, by Lemma 5 we have $w > 0$. Equation (10) becomes

$$\ln(1 - w(1 - w/B)) = -w. \quad (11)$$

We thus obtain the following parametrization of B, a, q in terms of w :

$$B = \frac{w^2}{e^{-w} + w - 1}, \quad a = \frac{1}{1 - e^{-w}} - \frac{1}{w}, \quad q = \frac{e^w + e^{-w} - 2}{e^{-w} + w - 1}. \quad (12)$$

Since $B > \mathfrak{B}_u$, by the definition (2) of the threshold \mathfrak{B}_u , there exists $B' < B$ and $z > 0$ such that

$$\frac{B' - z}{B' + (q - 1)z} = \exp(-z)$$

and hence

$$B' = z + \frac{qz}{e^z - 1}. \quad (13)$$

We will now prove that, for B and q as in (12), for any $z > 0$ we have

$$B \leq z + \frac{qz}{e^z - 1}, \quad (14)$$

contradicting (13) and $B' < B$.

To prove (14), our goal is to show that for any $w > 0$ and any $z > 0$

$$\frac{w^2}{e^{-w} + w - 1} \leq z + \frac{e^w + e^{-w} - 2}{e^{-w} + w - 1} \frac{z}{e^z - 1}.$$

Since $e^{-w} + w - 1 > 0$ and $e^z - 1 > 0$, multiplying out this inequality yields the equivalent

$$0 \leq z(e^z - e^w)(e^{-w} - 1) - w(w - z)(e^z - 1) =: G_1(w, z). \quad (15)$$

We have

$$G_1(s+y, 2s) = (s^2 - y^2)(e^{2s} - 1) - 2s(e^s - e^y)(e^s - e^{-y}) =: G_2(s, y).$$

We will show $G_2(s, y) \geq 0$ for all $s > 0$ and $y \geq -s$. We have $G_2(s, -s) = 0$ and $\lim_{y \rightarrow \infty} G_2(s, y) = \infty$. Thus it is enough to explore the critical points of $G_3(y) := G_2(s, y)$ for each s . We have

$$\frac{\partial}{\partial y} G_3(y) = 2e^s \left(s(e^y - e^{-y}) - y(e^s - e^{-s}) \right).$$

The function $y \mapsto (e^y - e^{-y})/y$ is monotone for $y \geq 0$ (this follows from the series expansion) and hence the only critical points of $G_3(y)$ are $y = 0$ and $y = \pm s$. For $y = \pm s$ we have $G_3(y) = 0$. For $y = 0$ we have

$$G_3(0) = s^2(e^{2s} - 1) - 2s(e^s - 1)^2 = \sum_{i=5}^{\infty} \frac{2^i(i-5) + 16}{4(i-1)!} s^i > 0.$$

This establishes non-negativity of $G_3(y)$ for $y \geq -s$ for all $s > 0$. This completes the proof of (14) and hence the proof of the lemma. \square

The following lemma details the number of critical points of Ψ_1 in the interval $(1/q, 1]$.

Lemma 7. *Let N be the number of critical points of Ψ_1 in the interval $(1/q, 1]$. Then,*

1. *for $B < \mathfrak{B}_u$, N equals 0,*
2. *for $B = \mathfrak{B}_u$, N equals 1,*
3. *for $\mathfrak{B}_u < B < \mathfrak{B}_{rc}$, N equals 2,*
4. *for $B \geq \mathfrak{B}_{rc}$, N equals 1.*

Proof. Consider the function $g(z) = z + \frac{qz}{e^z - 1}$ for $z \geq 0$. By the definition (3) of \mathfrak{B}_u , we have that

$$\mathfrak{B}_u = \min_{z \geq 0} g(z).$$

Since $g''(z) = \frac{qe^z(e^z(z-2)+z+2)}{(e^z-1)^3}$ and $\lim_{z \downarrow 0} g'(z) = 1 - q/2$, we have that $g(z)$ is a convex function of z and that, for $q \geq 3$, its minimum is attained (uniquely) at a point $z_0 > 0$.

Note that if $B \geq \mathfrak{B}_u$ and $z > 0$ satisfy $B = g(z)$, then $\alpha = \frac{e^z}{e^z + q - 1} > 1/q$ is a critical point of Ψ_1 (cf. (5)); similarly a critical point of Ψ_1 in the interval $(1/q, 1]$ yields $z > 0$ such that $B = g(z)$. It follows that for $B < \mathfrak{B}_u$, Ψ_1 has no critical point in the interval $(1/q, 1]$. Since $\lim_{z \uparrow +\infty} g(z) = +\infty$ and g is continuous, we have that for $B \geq \mathfrak{B}_u$, Ψ_1 has at least a critical point in the interval $(1/q, 1]$. Since the function $g(z)$ is convex and $\lim_{z \downarrow 0} g(z) = \mathfrak{B}_{rc}$, we obtain that Ψ_1 , in the interval $(1/q, 1]$, has exactly two critical points for $B \in (\mathfrak{B}_u, \mathfrak{B}_{rc})$ and exactly one critical point for $B \geq \mathfrak{B}_{rc}$. \square

We are now ready to prove Lemma 2.

Proof of Lemma 2. Note that $\Psi'_1(\alpha) \uparrow \infty$ for $\alpha \downarrow 0$ and $\Psi'_1(\alpha) \downarrow -\infty$ for $\alpha \uparrow 1$, so all the local maxima correspond to critical points of Ψ_1 .

By (8), we have that $\Psi'_1(1/q) = 0$ and hence $u = 1/q$ is a critical point of Ψ_1 for all $B > 0$. In fact, we have that $\Psi''_1(1/q) < 0$ for $B < \mathfrak{B}_{rc}$ and $\Psi''_1(1/q) > 0$ for $B > \mathfrak{B}_{rc}$. At $B = \mathfrak{B}_{rc}$, we have $\Psi''_1(1/q) = 0$ and $\Psi'''_1(1/q) \neq 0$ (using $q \geq 3$). It follows that

$$u = 1/q \text{ is a local maximum of } \Psi_1 \text{ iff } B < \mathfrak{B}_{rc}. \quad (16)$$

We also have that Ψ_1'' is monotone in the interval $[0, 1/2]$ (since $1/(\alpha(1-\alpha))$ is monotone). For $B < \mathfrak{B}_{rc}$, we have that $\Psi_1''(1/q) < 0$ and $\lim_{\alpha \downarrow 0} \Psi_1''(\alpha) < 0$, so we obtain that Ψ_1' is decreasing in the interval $[0, 1/q]$ and hence

$$\text{for } B < \mathfrak{B}_{rc}, \text{ there are no critical points/local maxima of } \Psi_1 \text{ in the interval } [0, 1/q]. \quad (17)$$

We next search for the existence of critical points/local maxima in the interval $(1/q, 1]$. We have the following case analysis.

Case I. For $B < \mathfrak{B}_u$, by (16), (17) and Item 1 of Lemma 7, we have that $u = 1/q$ is the unique critical point of Ψ_1 and it is a local maximum of Ψ_1 .

Case II. For $B = \mathfrak{B}_u$, by (17) and Item 2 of Lemma 7 we have that Ψ_1 has exactly two critical points, at $u = 1/q$ and $a > 1/q$. By (16), we have that Ψ_1 has a local maximum at $u = 1/q$. Ψ_1 cannot have a local maximum at a , otherwise Ψ_1 must have at least one critical point in the interval (u, a) which contradicts the fact that Ψ_1 has exactly one critical point in $(1/q, 1]$ (Item 2 of Lemma 7).

Case III. For $B \in (\mathfrak{B}_u, \mathfrak{B}_{rc})$, by (17) and Item 3 of Lemma 7, we have that Ψ_1 has exactly three critical points, at $u = 1/q$ and $a_1, a_2 > 1/q$ with $a_1 < a_2$. By (16), we have that Ψ_1 has a local maximum at $u = 1/q$. From this, it follows that Ψ_1 does not have a local maximum at a_1 , otherwise Ψ_1 would have a critical point in the interval (u, a_1) ; so, $\Psi_1''(a_1) \geq 0$. In fact, by Lemma 6, we can conclude that Ψ_1 has a local minimum at a_1 . It follows that Ψ_1 cannot have a local minimum at a_2 (otherwise there would be a critical point of Ψ_1 between a_1 and a_2). Again by Lemma 6, we conclude that Ψ_1 has a local maximum at a_2 .

The analysis of the values of B where the two local maxima of Ψ_1 correspond to global maxima is well-known and can be found in, e.g., [10]. Roughly, denoting by a the point where Ψ_1 has a local maximum in the interval $(1/q, 1]$, it can be shown that $\Psi_1(a) - \Psi_1(u)$ is increasing with respect to B ; then, one only needs to observe that, at $B = \mathfrak{B}_o$, it holds that $a = (q-1)/q$ and $\Psi_1(a) = \Psi_1(u)$.

Case IV. For $B \geq \mathfrak{B}_{rc}$, by (17) and Item 4 of Lemma 7, we have that Ψ_1 has exactly two critical points in the interval $[1/q, 1]$, at $u = 1/q$ and $a > 1/q$. By (16), we have that Ψ_1 does not have a local maximum at $u = 1/q$. Since $\Psi_1'(\alpha) \downarrow -\infty$ for $\alpha \uparrow 1$, we obtain that Ψ_1 cannot have a local minimum at a (otherwise there would be a critical point in the interval $(a, 1)$). By Lemma 6, we conclude that Ψ_1 has a local maximum at a . \square

3.2 Connection: Proof of Lemma 3

In this section, we prove Lemma 3 presented in Section 2 connecting the critical points of the function Ψ_1 with the fixpoints of the function F . Recall that the function F captures the density of the largest color class after one step of the SW algorithm (see (6) and (7) for the definition of F).

We first prove the following lemma. The lemma corresponds to the intuitive fact that $F(z)$ is an increasing function of the initial density z and that the rate of increase, i.e., $F'(z)$, is a decreasing function of z .

Lemma 8. *For every $B > 0$, the function F satisfies $F'(z) > 0$ and $F''(z) < 0$ for all $z \in (1/B, 1]$, i.e., F is strictly increasing and concave in the interval $[1/B, 1]$.*

Proof. We may assume that $B > 1$ (otherwise there is nothing to prove). Let $z \in (1/B, 1]$ and recall that $x \in (0, 1)$ is the (unique) solution of

$$x + \exp(-zBx) = 1. \quad (7)$$

We view (7) as an equation that defines x as an implicit function of z . Differentiating (7) two times we obtain

$$\begin{aligned}\frac{\partial x}{\partial z} &= \frac{Bxe^{-zBx}}{1 - zBe^{-zBx}}, \\ \frac{\partial^2 x}{\partial z^2} &= -\frac{B^2ze^{-zBx}(2e^{-zBx}(1 - zBe^{-zBx}) + x)}{(1 - zBe^{-zBx})^3}.\end{aligned}$$

Since $F(z) = \frac{1}{q} + \left(1 - \frac{1}{q}\right)zx$, we obtain

$$\begin{aligned}F'(z) &= \frac{q-1}{q} \left(x + z \frac{\partial x}{\partial z}\right) = \frac{(q-1)x}{q(1 - zBe^{-zBx})}, \\ F''(z) &= \frac{q-1}{q} \left(2 \frac{\partial x}{\partial z} + z \frac{\partial^2 x}{\partial z^2}\right) = -\frac{(q-1)Bxe^{-zBx}(zB(x + 2e^{-zBx}) - 2)}{q(1 - zBe^{-zBx})^3}.\end{aligned}\tag{18}$$

We first show that $F'(z) > 0$ for all $z \in (1/B, 1]$. Since x is positive for all $z > 1/B$, it suffices to show that $1 - zBe^{-zBx} > 0$. Since x satisfies (7), we have $zB = -\ln(1-x)/x$, so we have

$$1 - zBe^{-zBx} = \frac{x + (1-x)\ln(1-x)}{x} > 0,\tag{19}$$

for all $0 < x < 1$ (the inequality holds since the derivative of the numerator is $-\ln(1-x)$ and its value at $x = 0$ is 0). Thus $F'(z) > 0$ for $z > 1/B$.

We next show that $F''(z) < 0$ for all $z \in (1/B, 1]$. We have already shown that the denominator in the expression for $F''(z)$ is positive, so we only need to show that $zB(x + 2e^{-zBx}) - 2 > 0$. Using again that $zB = -\ln(1-x)/x$, we have

$$zB(x + 2e^{-zBx}) - 2 = -\frac{2x + (2-x)\ln(1-x)}{x} > 0,$$

for all $0 < x < 1$ (the inequality holds since the numerator at $x = 0$ is 0 and the first derivative of the numerator is $-\frac{x+(1-x)\ln(1-x)}{1-x} < 0$ from (19)). This concludes the proof. \square

We next prove the following correspondence.

Lemma 9. *Let $B > 0$. For any $a > 1/q$, Ψ_1 has a critical point at a iff F has a fixpoint at a .*

Proof. Consider first a critical point a of Ψ_1 (with $a > 1/q$). We use the same parametrization as in the proof of Lemma 5, i.e., we set $a = (z+1)/(z+q)$ where $z > 0$, so that z satisfies

$$\ln(1+z) = \frac{zB}{z+q}.\tag{9}$$

Now, consider a fixpoint a of F (with $a > 1/q$). Note that $a > 1/B$ (this is immediate for $B \geq q$ since then $1/q \geq 1/B$; for $B < q$, we have that $F(z) = 1/q$ for all $z \in [1/q, 1/B]$, so there is no fixpoint of F with $a \in (1/q, 1/B]$). Therefore, from (6), we have $F(a) = \frac{1}{q} + \left(1 - \frac{1}{q}\right)ax$, where $x \in (0, 1]$ is the unique solution of

$$x + \exp(-aBx) = 1.\tag{7}$$

Under the parametrization $a = (z+1)/(z+q)$, equation $F(a) = a$ becomes

$$x = \frac{z}{z+1},\tag{20}$$

and (7) becomes

$$x + \exp\left(-\frac{z+1}{z+q}Bx\right) = 1. \quad (21)$$

Plugging (20) into (21) and taking logarithm of both sides yields (9). This proves the lemma. \square

Lemma 10. *The function F has a fixpoint at $u = 1/q$ iff $B \leq \mathfrak{B}_{rc}$. For $B < \mathfrak{B}_{rc}$, the fixpoint $u = 1/q$ of F is jacobian attractive. For $B = \mathfrak{B}_{rc}$, the fixpoint $u = 1/q$ is jacobian repulsive.*

Proof. Recall from (6) that $F(z) = \frac{1}{q} + \left(1 - \frac{1}{q}\right)zx$, where $x = 0$ for $z \leq 1/B$ and for $z > 1/B$, $x \in (0, 1]$ is the (unique) solution of

$$x + \exp(-zBx) = 1. \quad (7)$$

Note that when $z = u = 1/q$, we obtain that $x = 0$ for $B \leq \mathfrak{B}_{rc}$ and $x > 0$ for $B > \mathfrak{B}_{rc}$. Hence $F(u) = u$ iff $B \leq \mathfrak{B}_{rc}$, proving the first part of the lemma.

For $B < \mathfrak{B}_{rc}$, we have that F is constant throughout $[1/q, 1/B]$, so trivially $F'(1/q) = 0$ and hence u is jacobian attractive.

For $B = \mathfrak{B}_{rc}$, rewrite (7) as

$$zq = f(x), \text{ where } f(x) := -\frac{\ln(1-x)}{x}. \quad (22)$$

Note that as $x \downarrow 0$, we have $z \downarrow 1/q$. Then, for all sufficiently small $x > 0$, an expansion of f around $x = 0$ yields

$$z = \frac{1}{q} \left(1 + \frac{x}{2} + \frac{x^2}{3}\right) + O(x^3).$$

It is not hard from here to conclude

$$x = 2q(z - 1/q) + O((z - 1/q)^2),$$

for all z in a small neighborhood of $1/q$. It follows that

$$F'(1/q) = 2(q-1)/q > 1,$$

for all $q \geq 3$, and hence u is jacobian repulsive. \square

We are now ready to give the proof of Lemma 3.

Proof of Lemma 3. Our goal is to show that, in the interval $[1/q, 1]$, the hessian local maxima of Ψ_1 and the jacobian attractive fixpoints of F are in one-to-one correspondence.

We first prove the correspondence in the half-open interval $(1/q, 1]$. By Lemma 9, we have that every critical point $a > 1/q$ of Ψ_1 is also a fixpoint of F (and vice versa). Therefore, it suffices to show that a critical point a of Ψ_1 is a hessian local maximum of Ψ_1 iff a is also a jacobian attractive fixpoint of F .

From (8), we have

$$\Psi_1''(a) = B \frac{q}{q-1} - \frac{1}{a(1-a)}. \quad (8)$$

By Lemma 5, we have that $a > 1/B$, so from (18), we have that

$$F'(a) = \left(1 - \frac{1}{q}\right) \frac{x}{1 - aB \exp(-aBx)}, \quad (23)$$

where $x \in (0, 1]$ satisfies

$$x + \exp(-aBx) = 1. \quad (7)$$

Since a is also a fixpoint of F , we have $F(a) = a$, which yields $x = (qa - 1)/((q - 1)a)$. From (7), we also have $\exp(-aBx) = 1 - x$. Plugging these values in (23), we obtain

$$F'(a) = 1 + \frac{B \frac{q}{q-1} - \frac{1}{a(1-a)}}{\frac{q}{1-a} - B \frac{q}{q-1}} = 1 + \frac{\Psi_1''(a)}{\frac{q}{1-a} - B \frac{q}{q-1}}. \quad (24)$$

The denominator of (24) is positive (since $a > 1/B$) and hence we have

$$F'(a) < 1 \iff \Psi_1''(a) < 0. \quad (25)$$

We also have by Lemma 8 that $F'(a) > 0$, so we can rewrite (25) as

$$|F'(a)| < 1 \iff \Psi_1''(a) < 0. \quad (26)$$

This establishes the lemma in the interval $(1/q, 1]$.

We next consider the left-endpoint of the interval $[1/q, 1]$, i.e., the point $u = 1/q$. From (8), we have that u is a critical point of Ψ_1 for all $B > 0$ and it is a hessian local maximum of Ψ_1 (i.e., it holds that $\Psi_1''(u) < 0$) iff $B < \mathfrak{B}_{rc}$. By Lemma 10, we have that $u = 1/q$ is a jacobian attractive fixpoint of F precisely when $B < \mathfrak{B}_{rc}$.

This concludes the proof of the lemma. \square

3.3 Analysis of the fixpoints of F : Proof of Lemma 4

In this section, we prove Lemma 4, i.e., we analyze the fixpoints of the function F in the interval $[1/q, 1]$. Lemma 10 details when $u = 1/q$ is a (jacobian attractive) fixpoint of F , therefore we will focus on the interval $(1/q, 1]$.

Recall, by Lemma 9, a point $a \in (1/q, 1]$ is a fixpoint of F iff a is a critical point of Ψ_1 . Therefore, the number of fixpoints of F in the interval $(1/q, 1]$ is the same as the number of critical points of Ψ_1 in the interval $(1/q, 1]$. Lemma 7 therefore yields the following corollary.

Corollary 11. *For $B < \mathfrak{B}_u$, there is no fixpoint of F in the interval $(1/q, 1]$. For $B = \mathfrak{B}_u$, there is a unique fixpoint of F in the interval $(1/q, 1]$. For $B \in (\mathfrak{B}_u, \mathfrak{B}_{rc})$, there are two fixpoints of F in the interval $(1/q, 1]$. For $B \geq \mathfrak{B}_{rc}$, there is a unique fixpoint of F in the interval $(1/q, 1]$.*

By Lemma 6, every local maximum of Ψ_1 in the interval $(1/q, 1]$ is in fact a hessian maximum of Ψ_1 . By Lemma 3, a hessian maximum of Ψ_1 is also a jacobian attractive fixpoint of F . Therefore, Lemma 2, which details the local maxima of Ψ_1 , yields the following.

Corollary 12. *For $B > \mathfrak{B}_u$, the function F has a unique jacobian attractive fixpoint in the interval $(1/q, 1]$, namely the point $a > 1/q$ where Ψ_1 has a local maximum.*

Corollaries 11 and 12 classify the number of fixpoints of F and when these are jacobian attractive for all $B \neq \mathfrak{B}_u$. The following lemma addresses the case $B = \mathfrak{B}_u$.

Lemma 13. *For $B = \mathfrak{B}_u$, consider the fixpoint a of F in the interval $(1/q, 1]$. Then, $F'(a) = 1$.*

Remark 1. *Note, the non-attractiveness of the fixpoint a for $B = \mathfrak{B}_u$ follows from $F'(a) = 1$ and $F''(a) \neq 0$ (Lemma 8).*

Proof. From (24), it suffices to show that $\Psi_1''(a) = 0$.

Recall also that $\Psi_1'(a) = 0$, i.e., a is a critical point of Ψ_1 . Using Lemma 7, we therefore have that the critical points of Ψ_1 in the interval $[1/q, 1]$ are precisely $u = 1/q$ and a .

By Lemma 2, $u = 1/q$ is the unique local maximum of Ψ_1 and hence it must be the case that $\Psi_1''(a) \geq 0$ (otherwise a would also be a local maximum). We also have that $\Psi_1''(a) \leq 0$: otherwise, Ψ_1 has a local minimum at a . Since $\Psi_1'(\alpha) \downarrow -\infty$ as $\alpha \uparrow 1$, we would then obtain that Ψ_1 has a critical point in the interval $(a, 1)$, contradicting that, for $B = \mathfrak{B}_u$, Ψ_1 has a unique critical point in the interval $(1/q, 1]$ (Lemma 7).

Thus, $\Psi_1''(a) = 0$, as wanted. \square

We are now ready to prove Lemma 4 from Section 2.

Proof of Lemma 4. Lemma 10 details when $u = 1/q$ is a fixpoint of F and when it is jacobian attractive. It therefore remains to classify the fixpoints in the interval $(1/q, 1]$.

For $B < \mathfrak{B}_u$, there are no fixpoints of F in the interval $(1/q, 1]$ by Corollary 11.

For $B > \mathfrak{B}_u$, by Corollary 12, there is precisely one jacobian attractive fixpoint of F in the interval $(1/q, 1]$ and it coincides with the point where Ψ_1 has a local maximum.

For $B = \mathfrak{B}_u$, by Corollary 12, there is precisely one fixpoint a of F in the interval $(1/q, 1]$. By Lemma 13 and Lemma 8, we have that $F'(a) = 1$ and $F''(a) \neq 0$, so a is repulsive but not jacobian repulsive.

This completes the proof of the lemma. \square

4 Random Graph Lemmas

In this section, we collect relevant results from the literature for the sizes of the components in $G(n, p)$ where $p \sim 1/n$. We will use these to analyze one step of the SW algorithm.

For a graph G , we denote by C_1, C_2, \dots the connected components of G in decreasing order of size; throughout the paper we refer to the size of a component C as the number of vertices in it and use $|C|$ to denote its size. Roughly, in one iteration of the SW algorithm, the size of the largest component after the percolation step controls the size of the largest color class, and the fluctuations are determined by the sum of squares of the sizes of the components.

4.1 The supercritical regime

We will need several known results on the $G(n, p)$ model in the supercritical regime ($p = c/n$, where $c > 1$). The size of the giant component is asymptotically normal [25] and satisfies moderate deviation inequalities around its mean value [1]. We will use the following moderate deviation inequalities for the sizes of the largest and second largest components of G . These are used to track the evolution of the SW dynamics for an exponential number of steps in the slow mixing regime $\mathfrak{B}_u < B < \mathfrak{B}_{rc}$.

Lemma 14. *Let $G \sim G(n, c/n)$ where $c > 1$. Let $\beta \in (0, 1)$ be the solution of $x + \exp(-cx) = 1$. Let C_1, C_2 be the largest and second largest components of G respectively. Then, for every constant $\varepsilon \in (0, 1/3]$ it holds that*

$$P(|C_1| - \beta n \geq n^{1/2+\varepsilon}) \leq \exp(-\Theta(n^{2\varepsilon})), \quad (27)$$

$$P(|C_2| \geq n^\varepsilon) \leq \exp(-\Theta(n^\varepsilon)). \quad (28)$$

Proof. Equation (27) is proved in [19, Lemma 5.4]. We next prove equation (28). All the elements are contained in the proof of [17, Theorem 5.4]. The probability that there exists a component of size from the interval $\{n^\varepsilon, \dots, n^{2/3}\}$ is bounded by (see [17, p. 110, line 11]):

$$n^2 \exp(-((c-1)^2/(9c))n^\varepsilon). \quad (29)$$

The probability that there exist two or more components of size at least $n^{2/3}$ is bounded by (see [17, p. 110, line 24]):

$$n^2 \exp(-((c-1)^2 c/4)n^{1/3}). \quad (30)$$

Using the union bound (combining (29) and (30)) we obtain (28), that is, with high probability we have only one component of size $\geq n^\varepsilon$. \square

The following lemma will be used to analyze the evolution of the SW chain when $B = \mathfrak{B}_u$.

Lemma 15. *Let $G \sim G(n, c/n)$ where $c_0 < c < c_1$ for absolute constants $c_0, c_1 > 1$ (c may otherwise depend on n). Let $\beta \in (0, 1)$ be the unique solution of $\beta + \exp(-\beta c) = 1$. Denote by C_1 the largest component in G .*

Then, for every constant $\varepsilon > 0$, for all sufficiently large n it holds that

$$n\beta - n^\varepsilon \leq E[|C_1|] \leq n\beta + n^\varepsilon. \quad (31)$$

Moreover, there exist constants $K_1, K_2, K_3 > 0$ (depending only on c_0, c_1) such that for all sufficiently large n it holds that

$$K_1 n \leq \text{Var}[|C_1|] \leq K_2 n, \quad E\left[\sum_{j \geq 2} |C_j|^2\right] \leq K_3 n. \quad (32)$$

Finally, there exists a constant $U > 0$ (depending only on c_0, c_1) such that for all sufficiently large n , for all $u \geq U$, it holds that

$$P(|C_1| - n\beta \geq u\sqrt{n}) \leq U/u^2. \quad (33)$$

Proof. The bounds on $E[|C_1|]$ and $\text{Var}[|C_1|]$ can be found in [7, Theorem 5]. The bound on $E\left[\sum_{j \geq 2} |C_j|^2\right]$ is an immediate corollary of [19, Corollary 5.6]. The probability bound in (33) can be derived by Chebyshev's inequality using the bounds on $E[|C_1|]$ and $\text{Var}[|C_1|]$. \square

4.2 The scaling window & subcritical regimes

We use the following well-known result about the size of the giant component in the subcritical regime.

Lemma 16 (see, e.g., [17], p.109). *Let $t \in (0, 1]$ be a constant. Let $G \sim G(n, c/n)$ where $c < 1$ is a constant, and C_1 be the largest component of G . Then,*

$$P(|C_1| \geq n^t) \leq \exp(-\Theta(n^t)).$$

The following lemma considers the size of the components in the scaling window.

Lemma 17. *There exist constants $K, c, c' > 0$ such that for any n and*

1. any $\varepsilon \in (0, 1)$ for random G from $G(n, (1 - \varepsilon)/n)$ we have

$$E\left[\sum_{i \geq 1} |C_i|^2\right] \leq \frac{Kn}{\varepsilon},$$

2. for any $\varepsilon \in [1/n^{1/3}, c]$ for random G from $G(n, (1 + \varepsilon)/n)$ we have

$$E\left[\sum_{i \geq 2} |C_i|^2\right] \leq \frac{Kn}{\varepsilon},$$

3. for any $\varepsilon \in [c'/n^{1/3}, c']$ for random G from $G(n, (1 + \varepsilon)/n)$ we have

$$P\left([|C_1| < (7/4)\varepsilon n] \cup [|C_1| > 3\varepsilon n]\right) \leq K \exp(-c\varepsilon^3 n).$$

Proof. Part 1 follows from [19, Lemma 5.3 & Theorem 5.12]. Part 2 is [19, Theorem 5.13, Part (ii)]. Part 3 follows from [19, Lemma 5.4 & Theorem 5.9]. \square

Lemma 18. *Let $G \sim G(n, p)$, $p \geq (1 - An^{-1/3})/n$ where A is a large constant. Let C_1, C_2, \dots be the connected components of G in decreasing order of size. Then, for all sufficiently large constant $L > 0$, there exists a positive constant p' such that for all n sufficiently large it holds that $P(|C_1| \geq Ln^{2/3}, \sum_{j \geq 2} |C_j|^2 \leq n^{4/3}) \geq p'$.*

The proof of Lemma 18 is based on [19, Proof of Lemma 8.26]. We will use the following special case of [17, Theorem 5.20].

Corollary 19 ([17, Theorem 5.20]). *Let t be a positive integer and $d, a_1, \dots, a_t, b_1, \dots, b_t$ be such that $\infty \geq a_1 > b_1 > a_2 > b_2 > \dots > a_t > b_t > d > 0$. Let c be a constant (not necessarily positive) and let $p = (1 + cn^{-1/3})/n$.*

For $G \sim G(n, p)$ denote by C_1, C_2, \dots the connected components of G in decreasing order of their sizes. There exists $\ell := \ell(c, t, d, a_1, \dots, a_t, b_1, \dots, b_t) > 0$ such that for all sufficiently large n , it holds that

$$P\left(a_1 \geq \frac{|C_1|}{n^{2/3}} \geq b_1, \dots, a_t \geq \frac{|C_t|}{n^{2/3}} \geq b_t, d \geq \frac{|C_{t+1}|}{n^{2/3}}\right) \geq \ell.$$

Proof. The statement of [17, Theorem 5.20] is for the Erdős-Rényi random graph model $G(n, M)$ with $M = (n/2) + cn^{2/3}$. Since for $G \sim G(n, p)$ with $p = (1 + 2cn^{-1/3})/n$ the number of edges is $(n/2) + cn^{2/3} + O(\sqrt{n})$ with probability $\Omega(1)$, the corollary follows. \square

Proof of Lemma 18. Let $A > 0$ be a large constant. We consider two cases. If $p \geq (1 + An^{-1/3})/n$, we have from Part 2 of Lemma 17 that $E[\sum_{j \geq 2} |C_j|^2] \leq Kn^{4/3}/A$, so by Markov's inequality

$$P\left(\sum_{j \geq 2} |C_j|^2 \leq n^{4/3}\right) \geq 1 - \frac{K}{A}.$$

From Corollary 19 (with $t = 1$, $b_1 = L$) we obtain that for $p = 1/n$, $|C_1|$ is greater than $Ln^{2/3}$ with asymptotically positive probability p_1 for any constant $L > 0$. Note that for $p > 1/n$ we can couple $G \sim G(n, 1/n)$ and $G' \sim G(n, p)$ so that G is a subgraph of G' . Since $|C_1|$ is monotone, it

follows that for $p > 1/n$, $|C_1|$ is greater than $Ln^{2/3}$ with positive probability p_1 . Provided that A is sufficiently large (depending on K, p_1), by a union bound we have that

$$P\left(|C_1| \geq Ln^{2/3}, \sum_{j \geq 2} |C_j|^2 \leq n^{4/3}\right) \geq p_1 - \frac{K}{A} > 0.$$

If $(1 - An^{-1/3})/n \leq p \leq (1 + An^{-1/3})/n$, we have from Corollary 19 (with $t = 1$, $d = 1$, $b_1 = L$) and the argument in [19, Proof of Lemma 8.26] that for all sufficiently large L , it holds that

$$P\left(|C_1| \geq Ln^{2/3}, \sum_{j \geq 2} |C_j|^2 \leq n^{4/3}\right) \geq p_2 > 0,$$

where p_2 is a constant. The lemma follows. \square

We will also use the following upper bound on the size of the giant component in the critical window.

Lemma 20 ([24, Corollary 5.6], see also [23, Theorems 1 & 7]). *Let $G \sim G(n, p)$ with $p = (1 \pm cn^{-1/3})/n$ where c is a sufficiently large constant. Let C_1 be the largest component in G . Then there exists a constant $r > 0$ such that for positive A larger than an absolute constant, it holds that*

$$P(|C_1| > An^{2/3}) \leq \exp(-rA^3).$$

4.3 Concentration Inequalities

We conclude this section by recording the following version of Azuma's inequality that we will use.

Lemma 21 (Azuma's inequality, see, e.g., [17, p.37]). *Let X_1, \dots, X_n be independent random variables such that, for $i = 1, \dots, n$ it holds that $a_i \leq X_i \leq b_i$. Let $X = X_1 + \dots + X_n$. Then, for all $t \geq 0$, it holds that*

$$\Pr(|X - E(X)| \geq t) \leq 2 \exp\left(-\frac{t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).$$

5 Slow Mixing for $\mathfrak{B}_u < B < \mathfrak{B}_{rc}$

In this section, we show that the SW chain mixes slowly when $B \in (\mathfrak{B}_u, \mathfrak{B}_{rc})$.

Let $\mathcal{B}(\mathbf{v}, \delta)$ be the ℓ_∞ -ball of configuration vectors of the q -state Potts model in K_n around \mathbf{v} of radius δ , that is,

$$\mathcal{B}(\mathbf{v}, \delta) = \{\mathbf{w} \in \mathbb{Z}^q \mid \|\mathbf{w}/n - \mathbf{v}\|_\infty \leq \delta\}. \quad (34)$$

We will show that for $B < \mathfrak{B}_{rc}$ the Swendsen-Wang algorithm is extremely unlikely to leave the vicinity of the uniform configuration. More precisely, we show the following.

Lemma 22. *Assume $B < \mathfrak{B}_{rc}$. There exists $\varepsilon_0 > 0$ such that, for all constant $\varepsilon \in (0, \varepsilon_0)$, for $S = \mathcal{B}(\mathbf{u}, \varepsilon)$, it holds that*

$$\Pr(X_1 \in S \mid X_0 \in S) \geq 1 - \exp(-\Theta(n^{1/2})).$$

The reason for Lemma 22 failing for $B > \mathfrak{B}_{rc}$ is that the percolation step of the Swendsen-Wang algorithm on a cluster of size n/q yields linear sized connected components, and these allow the algorithm to escape the neighborhood of \mathbf{u} (a somewhat similar argument applies for $B = \mathfrak{B}_{rc}$ as well, though in this case one has to account more carefully for the fluctuations of the largest components since the percolation step of the SW dynamics is in the critical window for such configurations).

Proof of Lemma 22. Let $X_0 \in S$. The first step of the Swendsen-Wang algorithm chooses, for each color class, a random graph from $G(m, p)$, where $p = B/n$ and m is the number of vertices of that color. For all sufficiently small ε we have

$$p = \frac{B}{m} \frac{m}{n} \leq \frac{d}{m},$$

where $d < 1$ (we used $B < q$ and $m \leq n/q + \varepsilon n$). Now Lemma 16 (with $t = 1/2$) implies that with probability at least

$$1 - n \exp(-\Theta(n^{1/2})) \quad (35)$$

all components after the first step have size $\leq n^{1/2}$. The second step of the Swendsen-Wang algorithm colors each component by a uniformly random color; call the resulting state X_1 . Let Z_i be the number of vertices of color i in X_1 . By symmetry, $E[Z_i] = n/q$.

Now assume that all components have size $\leq n^{1/2}$. By Azuma's inequality (see Lemma 21),

$$\Pr(|Z_i - n/q| \geq \varepsilon n) \leq \exp(-\Theta(n^{1/2})), \quad (36)$$

and hence $\Pr(X_1 \in S) \geq 1 - n \exp(-\Theta(n^{1/2}))$, which combined with (35) yields the lemma. \square

We also analyze the behavior of the algorithm around the majority configuration \mathbf{m} (recall, for the configuration to exist we need $B \geq \mathfrak{B}_u$).

Lemma 23. *Assume $B > \mathfrak{B}_u$ and let $\mathbf{m} = (a, b, \dots, b)$ where $a > 1/q$ is the jacobian attractive fixpoint of F of Lemma 4. There exists constant $\varepsilon_0 > 0$ such that, for all sufficiently large n , for all $\varepsilon \in (n^{-1/7}, \varepsilon_0)$, for $S = \mathcal{B}(\mathbf{m}, \varepsilon)$, we have*

$$\Pr(X_1 \in S \mid X_0 \in S) \geq 1 - \exp(-\Theta(n^{1/3})). \quad (37)$$

The reason that Lemma 23 does not hold for $B = \mathfrak{B}_u$ is that the fixpoint $a > 1/q$ of F is no longer attractive; indeed, in Section 11 we show that the Swendsen-Wang algorithm escapes the vicinity of this fixpoint in $O(n^{1/3})$ steps.

Proof of Lemma 23. Let $X_0 \in S$ and let $\gamma := F'(a)$ (recall that $|\gamma| < 1$, since a is Jacobian attractive fixpoint by Lemma 4). The first step of the Swendsen-Wang algorithm chooses, for each color class, a random graph from $G(m, p)$, where $p = B/n$ and m is the number of vertices of that color. Let m_1 be the number of vertices of the dominant color. Since $X_0 \in S$ we have $m_1/n = a + \tau =: a'$ where $|\tau| \leq \varepsilon$. We can write

$$p = (m_1 B/n)/m_1 = (a' B)/m_1,$$

where $a' B > 1$ for sufficiently small $\varepsilon_0 > 0$ (using $a B > 1$ from Lemmas 5 and 9). This means that the $G(m, p)$ process for the dominant color class is supercritical. Let $\beta \in (0, 1]$ be the root of $x + \exp(-a' B x) = 1$. By Lemma 14 the random graph will have, with probability $\geq 1 - \exp(-\Theta(n^{1/3}))$, one component of size $a' \beta n \pm n^{2/3}$ and all the other components will have size at most $n^{1/3}$.

Let m_2 be the number of vertices in one of the non-dominant colors. Since $X_0 \in S$ we have $m_2/n =: b'$ where

$$b - \varepsilon_0 \leq b - \varepsilon \leq b' \leq b + \varepsilon \leq b + \varepsilon_0. \quad (38)$$

We can write

$$p = (m_2 B/n)/m_2 = (b' B)/m_2,$$

where $b' B < 1$ for sufficiently small $\varepsilon_0 > 0$ (using $bB < 1$ from Lemmas 5 and 9). This means that the $G(m, p)$ process in this component is subcritical. By Lemma 16 (with $t = 1/3$), with probability $\geq 1 - \exp(-\Theta(n^{1/3}))$ the random graph will have all components of size at most $n^{1/3}$.

To summarize: starting from a configuration in S after the first step of the Swendsen-Wang algorithm we have, with probability $\geq 1 - q \exp(-\Theta(n^{1/3}))$ one large component of size $a' \beta n \pm n^{2/3}$ and the remaining components are of size $\leq n^{1/3}$ (small components). In the second step of the algorithm the components get colored by a random color. By symmetry, in expectation each color obtains $(n - a' \beta n \mp n^{2/3})/q$ vertices from the small components and by Azuma's inequality this number is $(n - a' \beta n \mp n^{2/3})/q \pm n^{5/6}$ with probability $\geq 1 - \exp(-\Theta(n^{1/3}))$. Combining the analysis of the first and the second step we obtain that at the end with probability $\geq 1 - 2q \exp(-\Theta(n^{1/3}))$ we have

$$\left\| \alpha(X_{t+1}) - \left(F(a'), \frac{1 - F(a')}{q - 1}, \dots, \frac{1 - F(a')}{q - 1} \right) \right\|_\infty \leq 2n^{-1/6}. \quad (39)$$

For sufficiently small $\varepsilon_0 > 0$ there exists $\gamma' \in (\gamma, 1)$ such that for all $|\tau| < \varepsilon_0$ we have $|F(a + \tau) - a| < \gamma' \tau \leq \gamma' \varepsilon$. Therefore, for all sufficiently large n and $\varepsilon \in (n^{-1/7}, \varepsilon_0)$, we have

$$|F(a') \pm 2n^{-1/6} - a| \leq \varepsilon \text{ and } \left| \frac{1 - F(a')}{q - 1} \pm 2n^{-1/6} - b \right| \leq \varepsilon. \quad (40)$$

Combining (39) and (40) gives that $X_1 \in S$ with probability at least $1 - \exp(-\Omega(n^{1/3}))$, which finishes the proof of the lemma. \square

Combining Lemmas 22 and 23 we obtain Part 3 of Theorem 1.

Corollary 24. *For $B \in (\mathfrak{B}_u, \mathfrak{B}_{rc})$ the Swendsen-Wang algorithm has mixing time $\exp(\Omega(n^{1/3}))$.*

Proof. For some small constant $\varepsilon > 0$, let $S_{\mathbf{u}} = \mathcal{B}(\mathbf{u}, \varepsilon)$ and $S_{\mathbf{m}} = \mathcal{B}(\mathbf{m}, \varepsilon)$. We can choose ε so that $S_{\mathbf{u}} \cap S_{\mathbf{m}} = \emptyset$ (since $\mathbf{u} \neq \mathbf{m}$) and further, by Lemmas 22 and 23,

$$\Pr(X_1 \in S_{\mathbf{u}} \mid X_0 \in S_{\mathbf{u}}) \geq 1 - \exp(-Cn^{1/3}), \quad \Pr(X_1 \in S_{\mathbf{m}} \mid X_0 \in S_{\mathbf{m}}) \geq 1 - \exp(-Cn^{1/3}), \quad (41)$$

where $C > 0$ is a constant independent of n .

Let μ be the stationary distribution of the SW chain, i.e., μ is the Potts distribution given in (1). Let $S = S_{\mathbf{u}}$ if $\mu(S_{\mathbf{u}}) \leq \mu(S_{\mathbf{m}})$ and $S = S_{\mathbf{m}}$ otherwise, so that $\mu(S) \leq 1/2$. We will use \bar{S} to denote the set of configurations which are not in S .

Let $X_0 \in S$ and $T = \frac{1}{10} \exp(Cn^{1/3})$. Then, using (41), we have that

$$\Pr(X_T \in S) \geq (1 - \exp(-Cn^{1/3}))^T \geq 1 - T \exp(-Cn^{1/3}) \geq 9/10.$$

Observe now that

$$d_{TV}(X_T, \mu) = \max_{A \subseteq \Omega} |\mu(A) - \Pr(X_T \in A)| \geq |\mu(\bar{S}) - \Pr(X_T \in \bar{S})| \geq \frac{1}{2} - \frac{1}{10} > \frac{1}{4}.$$

It follows from the definition of mixing time that $T_{\text{mix}} \geq T$, as claimed. \square

We remark that for $B \in (\mathfrak{B}_u, \mathfrak{B}_{rc})$ and $B \neq \mathfrak{B}_o$, the subset of initial configurations where the mixing of Swendsen-Wang is slow has exponentially small mass in the Gibbs distribution (known as *essential mixing*, see [11]). More precisely, for $B \neq \mathfrak{B}_o$, the Swendsen-Wang algorithm started from a typical configuration of the Gibbs distribution gets within total variation distance $1/\text{poly}(n)$ from the stationary distribution in $O(\log n)$ steps. For $B \in (\mathfrak{B}_u, \mathfrak{B}_o)$, this follows by considering starting configurations which are close to uniform and then using the upcoming Lemmas 25, 26 and 42; for $B \in (\mathfrak{B}_o, \mathfrak{B}_{rc})$, this follows by considering starting configurations which are close to a majority phase and then using Lemmas 25, 27 and 31.

6 Basic rapid mixing results

Recall from Section 2.1 the definition of a phase of a configuration. In this section, we will consider two copies of the SW chain and, utilizing the symmetry of the complete graph, we give sufficient conditions on their phases that ensure the existence of a coupling.

The first lemma asserts that once the phases of the two chains align, we can couple the chains (so that the configurations agree).

Lemma 25 ([8, Lemma 4]). *For any constant $B > 0$, for all $q \geq 2$, all constant $\varepsilon > 0$, for $T = O(\log n)$ there is a coupling where $\Pr(X_T \neq Y_T \mid \alpha(X_0) = \alpha(Y_0)) \leq \varepsilon$.*

Lemma 25 is essentially identical to [8, Lemma 4], which is also used in [19, Lemma 4.1]. For completeness, we include the proof of the lemma.

Proof of Lemma 25. Let $A_t = \{v : X_t(v) = Y_t(v)\}$ and $D_t = V \setminus A_t$. We will define a one-step coupling which maintains $\alpha(X_t) = \alpha(Y_t)$ and where

$$E[|D_{t+1}| \mid X_t, Y_t] = (1 - 1/q)|D_t|. \quad (42)$$

We'll define a matching $\tau : V \rightarrow V$. For $v \in A_t$ let $\tau(v) = v$. For $V \setminus A_t$ define τ so that for all $v \in V$, $X_t(v) = Y_t(\tau(v))$. In words, τ matches vertices with the same color (this is always possible since $\alpha(X_t) = \alpha(Y_t)$) and it uses the identity matching on those vertices whose colors agree in the two chains. In the percolation step of the Swendsen-Wang process, first perform the step for chain X_t , then for Y_t for a pair v, w where $Y_t(v) = Y_t(w)$ we delete the edge iff the edge $(\tau(v), \tau(w))$ is deleted. Therefore, the component sizes are identical for the two chains and we can couple the recoloring in the same manner so that if $v \in A_t$ then $v \in A_{t+1}$ and (42) holds. Then, by applying Markov's inequality,

$$\Pr(X_t \neq Y_t \mid X_0, Y_0) \leq n(1 - 1/q)^t \leq \varepsilon$$

for $t = O(\log n)$. □

It is enough to get the phases within $O(\sqrt{n})$ distance and then there is a coupling so that with constant probability the phases will be identical after one additional step. More precisely, we have the following lemmas which are analogous to [19, Theorem 6.5] for the $q = 2$ case.

Lemma 26. *Let $B < \mathfrak{B}_{rc}$ and $\mathbf{u} = (1/q, \dots, 1/q)$. Let X_0, Y_0 be a pair of configurations where $\|\alpha(X_0) - \mathbf{u}\|_\infty \leq Ln^{-1/2}$, $\|\alpha(Y_0) - \mathbf{u}\|_\infty \leq Ln^{-1/2}$, for an arbitrarily large constant $L > 0$. For all sufficiently large n , there exists a coupling such that with prob. $\Theta(1)$, $\alpha(X_1) = \alpha(Y_1)$.*

Lemma 27. *Let $B \geq \mathfrak{B}_u$ and $\mathbf{m} = (a, b, \dots, b)$ where $a > 1/q$ is the attractive fixpoint of F of Lemma 4. Let X_0, Y_0 be a pair of configurations where $\|\alpha(X_0) - \mathbf{m}\|_\infty \leq Ln^{-1/2}$, $\|\alpha(Y_0) - \mathbf{m}\|_\infty \leq Ln^{-1/2}$, for an arbitrarily large constant $L > 0$. For all sufficiently large n , there exists a coupling such that with prob. $\Theta(1)$, $\alpha(X_1) = \alpha(Y_1)$.*

For completeness, we include the proof of the lemmas.

Proof of Lemmas 26 and 27. We focus on proving Lemma 27 which is (slightly) more involved than Lemma 26, and then explain the small modification needed to obtain Lemma 26. Our proof closely follows the approach in [19, Theorem 6.5] (which is for the case $q = 2$) with small differences in some of the technical details.

Perform the percolation step of the Swendsen-Wang algorithm independently for the chains X_0 and Y_0 . By Lemma 5.7 in [19], there is a constant $C > 0$ such that with probability $1 - O(1/n)$, there are $\geq Cn$ isolated vertices in each chain (i.e., components of size 1). Our goal will be to couple the colorings of the components using the Cn isolated vertices to guarantee that $\alpha(X_1) = \alpha(Y_1)$.

In each chain, order the components by decreasing size. Next, couple the coloring step so that the largest component in each chain gets the same color. For the remaining components, color them independently in each chain in order of decreasing size, but leave the last Cn components uncolored. As noted earlier, the remaining Cn uncolored components in each chain are isolated vertices (with probability $1 - O(1/n)$). Let \hat{X}_1, \hat{Y}_1 denote the configuration except on these Cn uncolored components and denote by x_i, y_i the number of vertices which are assigned color i under \hat{X}_1, \hat{Y}_1 respectively.

We will show that under this coupling, for a (large) constant $L' > 0$, with probability $\Theta(1)$, it holds that

$$|x_i - y_i| \leq L' \sqrt{n} \text{ for all } i = 1, \dots, q. \quad (43)$$

We will do this shortly, let us assume (43) for the moment and conclude the coupling argument. For $i \in [q]$, let $\ell_i := x_i - y_i$, so that $|\ell_i| \leq L' \sqrt{n}$ and denote by ℓ the vector with coordinates ℓ_1, \dots, ℓ_q . Further, denote the remaining Cn uncolored vertices as v_1, \dots, v_{Cn} . Let Z_i be the r.v. which denotes the number of vertices from v_1, \dots, v_{Cn} that get color i in X_1 , and let Z'_i denote the respective r.v. for Y_1 . We will couple $\mathbf{Z} := (Z_1, \dots, Z_q)$ with $\mathbf{Z}' := (Z'_1, \dots, Z'_q)$ so that

$$\Pr(\mathbf{Z}' = \mathbf{Z} + \ell) = \Omega(1). \quad (44)$$

From this, we clearly obtain a coupling such that with probability $\Theta(1)$ we have $\alpha_i(X_1) = \alpha_i(Y_1)$ for $i \in [q]$. The coupling in (44) is nearly identical to the one used in [19, Lemma 6.7], we give the details for completeness.

Consider $\mathbf{W} := (W_1, \dots, W_q)$, where \mathbf{W} follows the multinomial distribution $\text{Mult}(Cn, (\frac{1}{q}, \dots, \frac{1}{q}))$ and note that \mathbf{Z}, \mathbf{Z}' have the same distribution as \mathbf{W} . For $t > 0$, let

$$I(t) := \left\{ \mathbf{w} = (w_1, \dots, w_q) \in \mathbb{Z}^q \mid w_1, \dots, w_q \in \left[\frac{Cn}{q} - t\sqrt{n}, \frac{Cn}{q} + t\sqrt{n} \right], w_1 + \dots + w_q = Cn \right\}.$$

Standard deviation bounds (or, alternatively, using Stirling's approximation) yield that, for every constant $t > 0$, for $\mathbf{w} = (w_1, \dots, w_q) \in I(t)$, it holds that

$$\Pr(\mathbf{W} = \mathbf{w}) \geq \frac{C_0}{(\sqrt{n})^{q-1}}, \quad (45)$$

for some absolute constant $C_0 > 0$ (depending only on q, C, t). Note that the variance of any coordinate W_i is $\Theta(n)$, but since the sum of W_i 's is equal to Cn , the random vector \mathbf{W} lies in a $(q-1)$ -dimensional space, yielding the denominator $(\sqrt{n})^{q-1}$ in (45).

The coupling μ of \mathbf{Z}, \mathbf{Z}' will be defined to be optimal on pairs of the form $(\mathbf{w}, \mathbf{w} + \ell)$ with $\mathbf{w} \in I(L')$. More precisely, for $\mathbf{w} = (w_1, \dots, w_q) \in I(L')$, we set

$$\mu(\mathbf{Z} = \mathbf{w}, \mathbf{Z}' = \mathbf{w} + \ell) := \min \{ \Pr(\mathbf{W} = \mathbf{w}), \Pr(\mathbf{W} = \mathbf{w} + \ell) \} \geq \Omega(n^{-(q-1)/2}), \quad (46)$$

where in the last inequality we used (45) for $t = 2L'$ (recall that the coordinates of ℓ are bounded in absolute value by $L'\sqrt{n}$). For pairs $(\mathbf{w}, \mathbf{w}') \notin \{(\mathbf{w}, \mathbf{w} + \ell) \mid \mathbf{w} \in I(L')\}$, the coupling is independent (the construction is analogous to the one used in the proof of the Coupling lemma, see [18, Section 4.2]). Now note that

$$\mu(\mathbf{Z} = \mathbf{Z}' + \ell) \geq \sum_{\mathbf{w} \in I(L')} \mu(\mathbf{Z} = \mathbf{w}, \mathbf{Z}' = \mathbf{w} + \ell) = \Omega(1),$$

where in the last inequality we used (46) and the fact that the number of \mathbf{w} in $I(L')$ is $\Omega((\sqrt{n})^{q-1})$. This proves (44) with the coupling μ , and hence, modulo the proof of (43) which is given below, the proof of Lemma 27 is complete.

To prove (43), we may assume w.l.o.g. that the largest component received color 1 (in each of the chains, by the coupling). Let $n' = n - Cn$ and denote by $C_{1,X}, C_{1,Y}$ the largest components after the percolation step of the SW dynamics on X_0, Y_0 respectively. Since the configurations X_0 and Y_0 are close to $\mathbf{m} = (a, b, \dots, b)$, in each of these configurations, exactly one color class is supercritical and the remaining color classes are subcritical in the percolation step (using that $aB > 1$ and $bB < 1$ from Lemmas 5 and 9). Therefore, by Lemma 15, we have with probability $\Theta(1)$ that

$$||C_{1,X}| - |C_{1,Y}|| \leq K_0\sqrt{n}$$

for some (large) constant $K_0 > 0$. We will further show that for a (large) constant $K_1 > 0$, with probability $\Theta(1)$, it holds that

$$\left| x_1 - \left(\frac{n'}{q} + \left(1 - \frac{1}{q}\right)|C_{1,X}| \right) \right| \leq K_1\sqrt{n} \text{ and } \left| x_i - \frac{n' - |C_{1,X}|}{q} \right| \leq K_1\sqrt{n} \text{ for } i \neq 1, \quad (47)$$

and, by an identical argument, the analogous inequalities for the y_i 's. Combining these, we obtain (43) with $L' = 2(K_0 + K_1)$.

It remains to show (47). Consider the configuration X_0 . W.l.o.g. we may assume that color 1 induces the largest color class in X_0 , so that the assumption $\|\alpha(X_0) - \mathbf{m}\|_\infty \leq Ln^{-1/2}$ translates into

$$|\alpha_1(X_0) - a| \leq Ln^{-1/2}, \quad |\alpha_i(X_0) - b| \leq Ln^{-1/2} \text{ for } i \neq 1.$$

From this, we have that color 1 is supercritical in the coloring step of the SW dynamics, while the colors $2, \dots, q$ subcritical (since it holds that $aB > 1$ and $bB < 1$ by Lemmas 5 and 9). Let C_1, C_2, \dots be the components in decreasing order of size after performing the percolation step in X_0 and note that $C_1 = C_{1,X}$. We have

$$E\left[\sum_{j \geq 2} |C_j|^2\right] \leq Kn$$

for some absolute constant $K > 0$. To see this, use the bound in Lemma 15 and equation (32) for the (supercritical) color class 1 and Item 1 of Lemma 17 for each of the (subcritical) color classes $2, \dots, q$. By Markov's inequality (and restricting our attention to components other than the isolated vertices $\{v_1, \dots, v_{C_n}\}$) we obtain that with probability $\Theta(1)$ it holds that

$$\sum_{j \geq 2; C_j \neq \{v_1\}, \dots, \{v_{C_n}\}} |C_j|^2 \leq K'n \quad (48)$$

for some absolute constant $K' > 0$. Now, for $i = 1, \dots, q$ let J_i be the number of vertices colored with i among the vertices other than v_1, \dots, v_{C_n} and those that belonged to the component $C_{1,X}$. Note that

$$x_1 = |C_{1,X}| + J_1, \quad x_i = J_i \text{ for } i \neq 1. \quad (49)$$

Observe that $E[J_i] = (n' - |C_{1,X}|)/q$. Further, using (48) and Azuma's inequality, we obtain that with probability $\Theta(1)$ it holds that

$$\left| J_i - \frac{n' - |C_{1,X}|}{q} \right| \leq K''\sqrt{n} \text{ for } i = 1, \dots, q. \quad (50)$$

for some absolute constant $K'' > 0$. Combining (49) and (50) yields (47) (with $K_1 = K''$), as wanted. In turn, this completes the proof of (43) and hence the proof of Lemma 27.

As mentioned earlier, the proof of Lemma 26 is completely analogous. The only difference is that now, where the configurations X_0, Y_0 are close to $\mathbf{u} = (1/q, \dots, 1/q)$, all color classes are subcritical in the percolation step of the SW algorithm (using that $B < \mathfrak{B}_{rc}$). Hence, there is no need to consider the size of the biggest components $C_{1,X}$ and $C_{1,Y}$. In particular, adapting the above arguments yields the following analogue of (47):

$$\left| x_i - \frac{n'}{q} \right| \leq K_1\sqrt{n} \text{ for } i \in [q]. \quad (51)$$

Using (51) (and the analogous inequalities for y_i 's), we obtain (43); the remaining bit of the proof of Lemma 26 is in all other respects identical to the proof of Lemma 27 (i.e., using the isolated vertices to couple X_1 and Y_1).

This concludes the proofs. \square

7 Fast mixing for $B > \mathfrak{B}_{rc}$

In this section, we prove that the SW algorithm mixes in $O(\log n)$ steps for all $B > \mathfrak{B}_{rc}$.

Let $\varepsilon > 0$ and consider a state X_t of the SW algorithm. We say that a color i is ε -heavy if $\alpha_i(X_t) \geq (1 + \varepsilon)/B$; it is ε -light if $\alpha_i(X_t) \leq (1 - \varepsilon)/B$. We will show that the SW algorithm has a reasonable chance of moving into a state where one color is ε -heavy and the remaining $q - 1$ colors are ε -light.

Lemma 28. *Assume $B > \mathfrak{B}_{rc}$ is a constant. There exists a constant $\varepsilon > 0$ such that the following hold for all sufficiently large n . For any initial state X_0 , with probability $\Theta(1)$ the next state X_1 has one ε -heavy color and the remaining $q - 1$ colors are ε -light. Moreover, if X_0 has one ε -heavy color and the remaining $q - 1$ colors are ε -light, the same is true for X_1 with probability $1 - o(1)$.*

Before proving Lemma 28 we will need the following function $g : [0, 1] \rightarrow [0, 1]$ which roughly captures the size of the largest component in $G(zn, B/n)$. Specifically, for $z \leq 1/B$ we set $g(z) = 0$; for $z > 1/B$ we set $g(z) = zx$, where x is the unique solution of $x + \exp(-zBx) = 1$ in $(0, 1]$. Note that the functions F and g are connected by the relation

$$F(z) = \frac{1}{q} + \left(1 - \frac{1}{q}\right)g(z) \text{ for all } z \in (1/B, 1].$$

The following inequality will be used to conclude the existence of heavy colors.

Lemma 29. *Assume $B > \mathfrak{B}_{rc}$. Then, for all $\alpha_1, \dots, \alpha_q \geq 0$ with $\alpha_1 + \dots + \alpha_q = 1$, it holds that*

$$\sum_{i \in [q]} g(\alpha_i) \geq g\left(1 - \frac{q-1}{B}\right) > 1 - \frac{q}{B}.$$

Proof of Lemma 29. For convenience, let $W := \sum_{i \in [q]} g(\alpha_i)$. Note that $g(z)$ is increasing and concave for $z > 1/B$ (this follows by Lemma 8 since $F(z) = \frac{1}{q} + (1 - \frac{1}{q})g(z)$ for $z \in (1/B, 1]$).

To minimize W , observe that

1. If $\alpha_i > 1/B$ and $\alpha_j < 1/B$ then we can decrease the value of W by decreasing α_i and increasing α_j (since $g(z) = 0$ for all $z \leq 1/B$ and $g(z)$ is increasing for $z > 1/B$).
2. If $1/B < \alpha_i < \alpha_j$ then we can decrease the value of W by decreasing α_i and increasing α_j (since $g(z)$ is concave for $z > 1/B$).

Since $B > \mathfrak{B}_{rc} = q$ and $\alpha_1 + \dots + \alpha_q = 1$, we have that at least one of the α_i 's is strictly larger than $1/B$. Thus, from Items 1 and 2, it follows that W is minimized when all but one of the α_i 's are equal to $1/B$ (the value of the remaining α_i is given by the condition $\alpha_1 + \dots + \alpha_q = 1$). Since $g(1/B) = 0$, it follows that

$$W \geq g\left(1 - \frac{q-1}{B}\right).$$

It remains to show that $g(z) > 1 - \frac{q}{B}$, where $z := 1 - (q-1)/B$. Note that $z > 1/B$ from $B > q$. Let $x \in (0, 1)$ be the solution of $x + \exp(-zBx) = 1$. The inequality $g(z) > 1 - \frac{q}{B}$ is equivalent to

$$x > \frac{B-q}{B-q+1}. \quad (52)$$

For the sake of contradiction, suppose that (52) is false, that is, $x \leq (B-q)/(B-q+1)$. Then,

$$1 - \frac{q-1}{B} = -\frac{\ln(1-x)}{Bx} \leq \frac{(B-q+1)\ln(B-q+1)}{B(B-q)}, \quad (53)$$

where the equality follows from $x + \exp(-zBx) = 1$ and the inequality follows from the fact that $x \mapsto -\frac{\ln(1-x)}{x}$ is an increasing function on $(0, 1)$. Inequality (53) yields that $B-q \leq \ln(1+B-q)$, which is false (since $B-q > 0$), and hence we have a contradiction. This shows that (52) is true. \square

We are now ready to prove Lemma 28.

Proof of Lemma 28. Let $W := g(1 - \frac{q-1}{B})$. By Lemma 29, there exists a small constant $\varepsilon > 0$ such that

$$W - \varepsilon \geq 1 - \frac{q}{B}(1 - 2\varepsilon).$$

Since the function $g(z)$ is continuous and $g(z) = 0$ for all $z \leq 1/B$, there exists a small constant $\eta > 0$ such that for all $z \leq (1 + \eta)/B$ it holds that $g(z) \leq \varepsilon/q$.

For $i \in [q]$, let m_i be the number of vertices of color i in X_0 and let $\alpha_i = m_i/n$. By Lemma 29,

$$\sum_{i \in [q]} g(\alpha_i) \geq W.$$

Perform the percolation step of the SW algorithm on the color class i and denote by G_i be the resulting graph. Moreover, let $C_1^{(i)}, C_2^{(i)}, \dots$ be the components of G_i in decreasing order of size. Note that G_i is distributed as $G(n\alpha_i, B/n)$.

To prove the first part of the lemma, note that for each color $i \in [q]$ the following hold with probability $1 - o(1)$:

- If $B\alpha_i \geq 1 + \eta$, the size of the largest component in G_i is $ng(\alpha_i) + o(n)$ (by Lemma 15).

- If $B\alpha_i \leq 1 + \eta$, by the choice of η we have $g(\alpha_i) \leq \varepsilon/q$ and therefore the largest component in G_i is trivially at least $g(\alpha_i)n - \frac{\varepsilon}{q}n$.

Moreover, with A being the constant in Lemma 18, we have that for each color $i \in [q]$ the following hold with positive probability (not depending on n):

1. If $B\alpha_i \geq (1 - Am_i^{-1/3})/m_i$, then $\sum_{j>1} |C_j^{(i)}|^2 \leq m_i^{4/3} \leq n^{4/3}$ (by Lemma 18).
2. If $(1 - Am_i^{-1/3})/m_i > B\alpha_i$, then $\sum_{j \geq 1} |C_j^{(i)}|^2 = O(n^{4/3})$ (by Item 1 of Lemma 17).

It follows that for all sufficiently large n , with probability $\Theta(1)$, after the percolation step of the SW algorithm, it holds that

$$\sum_{i \in [q]} |C_1^{(i)}| \geq (W - \varepsilon)n \geq \left(1 - \frac{q}{B}(1 - 2\varepsilon)\right)n \quad \text{and} \quad \sum_{i \in [q]} \sum_{j \geq 2} |C_j^{(i)}|^2 = o(n^2).$$

Now, in the coloring step of the SW algorithm, with probability $q^{-q} = \Theta(1)$, all of the components $C_1^{(i)}$ with $i \in [q]$ receive color 1. Conditioned on that, the expected number of vertices which get the color $k \neq 1$ after the coloring step of the SW algorithm is

$$\frac{n - \sum_{i \in [q]} |C_1^{(i)}|}{q} \leq n(1 - 2\varepsilon)/B.$$

Thus, using Azuma's inequality, we obtain that with probability $\Theta(1)$, for all colors $k \neq 1$, the number of vertices which get the color k after the coloring step of the SW algorithm is at most $n(1 - \varepsilon)/B$, which implies that the number of vertices which get the color 1 is at least $n(1 - (q - 1)(1 - \varepsilon)/B) \geq n(1 + \varepsilon)/B$. Thus, combining all the above, we obtain that, after one step of the SW algorithm, with probability $\Theta(1)$, color 1 is ε -heavy and all other colors are ε -light.

The second part of the lemma is analogous, the only difference is that now there is a unique ε -heavy color class i , which is therefore supercritical in the percolation step; all other color classes are ε -light and therefore subcritical. This allows us to improve the probability bounds in the previous analysis. In particular, by Lemma 15 (applied to the supercritical color) and Lemma 16 (applied to the subcritical colors), we obtain that with probability $1 - o(1)$, after the percolation step of the SW algorithm, there is just one linear-sized component of size $g(\alpha_i)n + o(n)$ and the remaining components have size $o(n)$. Since $g(\alpha_j) = 0$ for all $j \neq i$, Lemma 29 yields that $g(\alpha_i) \geq W$. W.l.o.g., we may assume that this unique linear-sized component receives the color 1. Then, using Azuma's inequality just as above, we obtain that, after one step of the SW algorithm, with probability $1 - o(1)$, all colors other than color 1 are ε -light and color 1 is ε -heavy.

This completes the proof of Lemma 28. \square

After applying Lemma 28 the behavior of the SW algorithm in one step will be controlled by the function F (cf. Section 2.2). We use this to show that, after $O(1)$ steps, with constant probability, the state of SW will be close to the majority phase \mathbf{m} ; recall that $\mathbf{m} = (a, b, \dots, b)$ where $a > 1/q$ is the unique fixpoint of F and $b = (1 - a)/(q - 1)$.

Lemma 30. *Assume $B > \mathfrak{B}_{rc}$ is a constant. For any constant $\delta > 0$, for all sufficiently large n and any starting state X_0 , after $T = O(1)$ steps, with probability $\Theta(1)$ the SW algorithm moves to a state X_T such that $\|\alpha(X_T) - \mathbf{m}\|_\infty \leq \delta$.*

Proof. Let $\varepsilon > 0$ be the constant in Lemma 28. Then, with probability $\Theta(1)$, the state X_1 has one ε -heavy color and the remaining $q - 1$ colors are ε -light.

Assume that at time $t \geq 1$ we are at a state X_t with one ε -heavy color and $q - 1$ colors which are ε -light. Then by the second part of Lemma 28, the same is true for the state X_{t+1} with probability $1 - o(1)$. In fact, let zn be the number of vertices of the heavy color class in X_t ; we claim that with probability $1 - o(1)$, in X_{t+1} the heavy color class has $F(z)n + o(n)$ vertices, while all the other color classes have $\frac{1-F(z)}{q-1}n + o(n)$ vertices. Indeed, in the percolation step of the SW dynamics, exactly one color class is supercritical and the remaining $q - 1$ color classes are subcritical. By Lemma 15 (applied to the supercritical color) and Lemma 16 (applied to the subcritical colors), we obtain that with probability $1 - o(1)$, after the percolation step of the SW algorithm, there is just one linear-sized component C of size $g(z)n + o(n)$ and the remaining components have size $o(n)$. W.l.o.g., we may assume that the component C receives the color 1. Then, using Azuma's inequality just as in the proof of Lemma 28, we obtain that, with probability $1 - o(1)$, for each color $k \neq 1$, $\frac{1-g(z)}{q}n + o(n)$ vertices receive the color k and the remaining $\frac{n}{q} + \frac{q-1}{q}g(z)n + o(n) = F(z)n + o(n)$ vertices receive the color 1, as claimed.

We thus obtain that for any constant integer $T \geq 2$, with probability $\Theta(1)$ the SW algorithm moves to a state X_T where one color class has $\alpha n + o(n)$ vertices and each of the remaining color classes has $\frac{1-\alpha}{q-1}n + o(n)$ vertices, where α belongs to the interval $F^{(T)}([1/q, 1])$ (recall that $F^{(T)}$ is the T -th iterate of the function F). Since F is increasing (Lemma 8), we have $F^{(T)}([1/q, 1]) = [F^{(T)}(1/q), F^{(T)}(1)]$. Since $B > \mathfrak{B}_{rc}$, by Lemma 4 we have that F has a unique fixpoint $a > 1/q$. Hence, using also again that F is increasing, for any constant $\delta > 0$, there is a constant T such that $[F^{(T)}(1/q), F^{(T)}(1)] \subseteq [a - \delta/2, a + \delta/2]$. Thus in T steps, with probability $\Theta(1)$, we are within ℓ_∞ -distance δ of \mathbf{m} (with room to spare to absorb the $o(n)$ fluctuations of the color classes). \square

Then we show that once we are at constant distance from \mathbf{m} then in $O(\log n)$ steps the distance to \mathbf{m} further decreases to $O(n^{-1/2})$.

Lemma 31. *For $B > \mathfrak{B}_u$, there exist $\delta, L > 0$ such that the following is true. Suppose that we start at a state X_0 such that $\|\alpha(X_0) - \mathbf{m}\|_\infty \leq \delta$. Then in $T = O(\log n)$ steps with probability $\Theta(1)$ the SW algorithm ends up in a state X_t such that*

$$\|\alpha(X_T) - \mathbf{m}\|_\infty \leq Ln^{-1/2}. \quad (54)$$

Proof. Recall that $\mathbf{m} = (a, b, \dots, b)$ where $a > 1/q$ is a jacobian attractive fixpoint of F and $b = (1 - a)/(q - 1)$. Moreover, by Lemmas 5 and 9, it holds that $aB > 1$ and $bB < 1$.

Let $\delta > 0, c \in (0, 1)$ be constants such that for all $z \in [a - \delta, a + \delta]$ it holds that $|F(z) - a| \leq c|z - a|$ and $zB > 1, (1 - z)B/(q - 1) < 1$. Note that the existence of such constants δ, c is guaranteed by the jacobian attractiveness of the fixpoint a throughout the regime $B > \mathfrak{B}_u$ (Lemma 4) and the facts $aB > 1, bB < 1$.

Define the geometrically decreasing sequence $\{w_t\}_{t \geq 0}$ by setting $w_0 = \delta n^{1/2}$ and $w_t = \frac{1+c}{2}w_{t-1}$. Further, let $T := \left\lceil \frac{\frac{1}{2} \log n}{\log \frac{2}{1+c}} \right\rceil - K$ where $K > 0$ is a large constant to be chosen later. Note that for any constant K , it holds that

$$\frac{1+c}{2}L \leq w_T \leq L, \text{ where } L := \delta(2/(1+c))^K.$$

Thus, to prove the lemma, it suffices to show the following (slightly stronger) statement: there exists a constant $K > 0$ such that with probability $\Theta(1)$,

$$\text{for all } t = 0, 1, \dots, T, \text{ it holds that } \|\alpha(X_t) - \mathbf{m}\|_\infty \leq w_t n^{-1/2}. \quad (55)$$

The main step in the proof is to track one step of the SW dynamics. Specifically, we will show that there exist constants $L', C > 0$ such that for all $w_t \in [L', \delta n^{1/2}]$, for any state X_t such that $\|\alpha(X_t) - \mathbf{m}\|_\infty \leq w_t n^{-1/2}$, with probability at least $\exp(-C/w_t)$ it holds that

$$\|\alpha(X_{t+1}) - \mathbf{m}\|_\infty \leq w_{t+1} n^{-1/2}. \quad (56)$$

To conclude (55) from (56), note that by choosing K large, we can ensure that $w_0 \geq \dots \geq w_T \geq L'$ and hence the probability of the event in (55) is at least $\prod_{t=0}^T \exp(-C/w_t)$. The latter product is bounded by a positive constant, since w_t is a geometrically decreasing sequence.

It remains to show (56). In particular, assume that at time t it holds that $\|\alpha(X_t) - \mathbf{m}\|_\infty \leq w_t n^{-1/2}$ where $w_t \in [L', \delta n^{1/2}]$ for some large constant L' to be specified later. By the choice of the constant δ , in the percolation step of the SW dynamics, exactly one color class is supercritical and the remaining $q-1$ color classes are subcritical. Denote by C_1, C_2, \dots all the connected components after the percolation step, sorted in decreasing order of size. By the second inequality in (32) of Lemma 15 (applied to the supercritical color) and part 1 of Lemma 17 (applied to the subcritical colors), we obtain that there exists a constant $K' > 0$ such that

$$E\left[\sum_{i \geq 2} |C_i|^2\right] \leq K'n.$$

Let $w'_t := \frac{(1-c)}{2(1+\sqrt{K'})} w_t$; the choice of w'_t will become apparent shortly. Note that by choosing L' to be a large constant, we can ensure that w'_t is larger than any desired constant (whenever $w_t \in [L', \delta n^{1/2}]$).

By Markov's inequality, it holds that

$$P\left(\sum_{i \geq 2} |C_i|^2 \leq w'_t K'n\right) \geq 1 - 1/w'_t. \quad (57)$$

Assuming that the event in (57) occurred, by Azuma's inequality, in the coloring step of the SW algorithm the number Z_i of vertices in $C_2 \cup C_3 \dots$ that receive color i is concentrated around the expectation, i.e.,

$$P\left(|Z_i - E[Z_i]| \geq w'_t \sqrt{K'n}\right) \leq 2 \exp(-w'_t/2). \quad (58)$$

Let zn be the number of vertices in the largest color class of X_t ; by the choice of δ in the beginning, we have that $zB > 1$ and hence the largest color class is supercritical in the percolation step of SW. Therefore, by Lemma 15 (equation (33)),

$$P(|C_1| - g(z)n \geq w'_t \sqrt{n}) \leq U/(w'_t)^2. \quad (59)$$

Combining (57), (58), and (59) (and choosing L' to be a large constant relative to $K', U, 1/(1-c), q$), we obtain that with probability at least

$$(1 - 1/w'_t)(1 - 2q \exp(-w'_t/2) - U/(w'_t)^2) \geq \exp(-10/w'_t) = \exp(-C/w_t), \quad C := \frac{20(1 + \sqrt{K'})}{1 - c}, \quad (60)$$

we have that

$$\left\| \alpha(X_{t+1}) - \left(F(z), \frac{1 - F(z)}{q-1}, \dots, \frac{1 - F(z)}{q-1} \right) \right\|_\infty \leq w'_t (1 + \sqrt{K'}) n^{-1/2}. \quad (61)$$

By the choice of the constants δ, c , we have

$$\left\| \left(F(z), \frac{1-F(z)}{q-1}, \dots, \frac{1-F(z)}{q-1} \right) - \mathbf{m} \right\|_{\infty} \leq c \|\alpha(X_t) - \mathbf{m}\|_{\infty} \leq c w_t n^{-1/2}. \quad (62)$$

Equations (61) and (62) combined yield that with probability $\geq \exp(-C/w_t)$, it holds that

$$\|\alpha(X_{t+1}) - \mathbf{m}\|_{\infty} \leq w'_t(1 + \sqrt{K'})n^{-1/2} + c w_t n^{-1/2} \leq \frac{c+1}{2} w_t n^{-1/2} = w_{t+1} n^{-1/2}, \quad (63)$$

where the last inequality follows from $w'_t = \frac{(1-c)}{2(1+\sqrt{K'})} w_t$. This proves (56) and therefore completes the proof of Lemma 31. \square

From Lemmas 25, 27, 30 and 31 we conclude the following.

Corollary 32. *Let $B > \mathfrak{B}_{rc}$ be a constant. The mixing time of the Swendsen-Wang algorithm on the complete graph on n vertices is $O(\log n)$.*

Proof. Let $\varepsilon > 0$ be a small constant and consider two copies X_t, Y_t of the SW chain. We will show that for some $T = O(\log n)$, there exists a coupling such that $\Pr(X_T \neq Y_T) \leq \varepsilon$.

Let δ, L be as in Lemma 31. By Lemma 30, for some $T_1 = O(1)$ with probability $\Theta(1)$ we have that

$$\|\alpha(X_{T_1}) - \mathbf{m}\|_{\infty} \leq \delta \text{ and } \|\alpha(Y_{T_1}) - \mathbf{m}\|_{\infty} \leq \delta. \quad (64)$$

By Lemma 31, for some $T_2 = O(\log n)$ with probability $\Theta(1)$, we have that

$$\|\alpha(X_{T_1+T_2}) - \mathbf{m}\|_{\infty} \leq L n^{-1/2} \text{ and } \|\alpha(Y_{T_1+T_2}) - \mathbf{m}\|_{\infty} \leq L n^{-1/2}. \quad (65)$$

Let $T' := T_1 + T_2$. Conditioning on (65), by Lemma 27 there exists a coupling that with probability $\Theta(1)$, for $T_3 = T' + 1$, it holds that $\alpha(X_{T_3}) = \alpha(Y_{T_3})$. Conditioned on $\alpha(X_{T_3}) = \alpha(Y_{T_3})$, by Lemma 25, for every constant $\varepsilon' > 0$ there exists $T_4 = O(\log n)$ and a second coupling such that $\Pr(X_{T_3+T_4} \neq Y_{T_3+T_4}) \leq \varepsilon'$. By letting ε' to be a sufficiently small constant, we obtain a coupling and some $T = O(\log n)$ such that $\Pr(X_T \neq Y_T) \leq \varepsilon$, as wanted. \square

8 Fast mixing for $B = \mathfrak{B}_{rc}$

The proof resembles the case $B > \mathfrak{B}_{rc}$, though we have to account more carefully for the mixing time of the chain for configurations which are close to uniform. In particular, for starting configurations which are ε -far from being uniform, a straightforward modification of the proof for $B > \mathfrak{B}_{rc}$ gives that the SW chain mixes rapidly. The main difficulty in the case $B = \mathfrak{B}_{rc}$ is to show that the chain escapes from starting configurations which are close to uniform. We will show that this happens after roughly $\log n$ steps. More precisely, we have the following lemma.

Lemma 33. *Assume $B = \mathfrak{B}_{rc}$. There exists constant $\varepsilon > 0$ such that for any n and any initial state X_0 with probability $\Theta(1)$ after $T_1 = O(\log n)$ steps, X_{T_1} has an ε -heavy color and the remaining $q-1$ colors are ε -light.*

Lemma 33 yields the following analogue of Lemma 30 (note here the logarithmic bound on T).

Lemma 34. *Assume $B = \mathfrak{B}_{rc}$. For any constant $\delta > 0$ and any starting state X_0 , after $T = O(\log n)$ steps, with probability $\Theta(1)$ the SW algorithm moves to state X_T with $\|\alpha(X_T) - \mathbf{m}\|_{\infty} \leq \delta$.*

Proof of Lemma 34. From Lemma 33, for some (small) constant $\varepsilon > 0$, we have that for $T_1 = O(\log n)$, with probability $\Theta(1)$, X_{T_1} has an ε -heavy color and the remaining $q - 1$ colors are ε -light. Using Lemma 8 (F is increasing), the second part of Lemma 10 (the uniform fixpoint is jacobian repulsive) and Corollary 11 (there exists a unique fixpoint of F in the interval $(1/q, 1]$), we obtain that for constant T_2 (depending on δ) we have $F^{(T_2)}([(1+\varepsilon)/q, 1]) \subseteq [a - \delta/2, a + \delta/2]$, so the same arguments as in the proof of Lemma 30 yield that $\|\alpha(X_{T_1+T_2}) - \mathbf{m}\|_\infty \leq \delta$ with probability $\Theta(1)$, as wanted. \square

Using Lemma 25 (note that it applies to all $B > 0$) and Lemmas 27 and 31 (note that these apply to all $B > \mathfrak{B}_u$), we may conclude the following from Lemma 34.

Corollary 35. *Let $B = \mathfrak{B}_{rc}$. The mixing time of the Swendsen-Wang algorithm on the complete graph on n vertices is $O(\log n)$.*

Proof. The proof is completely analogous to the proof of Corollary 32, the only difference is that now we use Lemma 34 to argue that (64) holds with probability $\Theta(1)$ for $T_1 = O(\log n)$. \square

We next turn to the proof of Lemma 33. We will use the following definition. For $W > 0$, a state X will be called W -good if X has a W -heavy color and the remaining $q - 1$ colors are $(W/2q)$ -light.

Lemma 36. *Let $B = \mathfrak{B}_{rc}$. For any starting state X_0 and an arbitrary constant $w > 0$, with probability at least $p(w) > 0$ (not depending on n) the next state X_1 of the SW dynamics is $wn^{-1/3}$ -good.*

Lemma 37. *Let $B = \mathfrak{B}_{rc}$. There exist absolute constants $c_1, c_2, C > 0$ such that for all n sufficiently large the following holds. For all w such that $c_1 \leq w \leq c_2 n^{1/3}$, for every $wn^{-1/3}$ -good starting state X_0 , the next state of the SW dynamics X_1 is $(13/12)wn^{-1/3}$ -good with probability at least $\exp(-C/w)$.*

Before proceeding, let us briefly motivate Lemmas 36 and 37. First, we explain the origin of the constant $13/12$ in Lemma 37, whose value is somewhat arbitrary, any constant strictly smaller than $4/3$ (and greater than 1) would work for all $q \geq 3$. To understand where the constant $4/3$ comes from, recall from Lemma 10 that the uniform phase $u = 1/q$ is a jacobian repulsive fixpoint of F (for $B = \mathfrak{B}_{rc}$) and, more precisely, $F'(1/q) = 2(q-1)/q$ (note that $F'(1/q) > 1$ for all $q > 2$). Then, just observe that $\min_{q \geq 3} \{2(q-1)/q\} = 4/3$.

Thus, for any $4/3 > c > 1$ (or, slightly less loosely, when $F'(1/q) > c > 1$), whenever $\|\alpha(X_t) - \mathbf{u}\|_\infty$ is sufficiently small, for all sufficiently large n , one would expect that

$$\|\alpha(X_{t+1}) - \mathbf{u}\|_\infty \geq c \|\alpha(X_t) - \mathbf{u}\|_\infty.$$

We show that this indeed holds by accounting carefully for color classes which are in the critical window for the percolation step of the SW dynamics (technically, to establish the probability bound in Lemma 37, we need that $\|\alpha(X_t) - \mathbf{u}\|_\infty = \Omega(n^{-1/3})$). Lemma 37 thus proves that an initial displacement of $\Omega(n^{-1/3})$, which is guaranteed with constant probability from Lemma 36, increases geometrically.

Lemma 33 follows immediately from Lemmas 36 and 37.

Proof of Lemma 33. Let c_1, c_2, C be the constants from Lemma 37. Define w_t by $w_1 = c_1$ and $w_t = (13/12)w_{t-1}$. Moreover, let $0 < \varepsilon_0 < c_2$ and set $t_0 = \lfloor \log(\varepsilon_0 n) / \log(13/12) \rfloor$. By Lemmas 36 and 37, for any starting state X_0 , the state X_t is w_t -good for all $t = 1, \dots, t_0$ with probability at

least $p(w_1) \prod_{t=2}^{t_0} \exp(-C/w_t) =: L$. Note that the product in the expression for L is bounded by an absolute positive constant, since the series $\sum_{t \geq 1} 1/w_t$ converges.

It follows that for any positive $\varepsilon < \varepsilon_0/(10q)$, with positive probability (not depending on n), X_{t_0} has an ε -heavy color and the remaining $q-1$ colors are ε -light, as wanted. \square

We next prove Lemmas 36 and 37.

Proof of Lemma 36. We will write α_i as a shorthand for $\alpha_i(X_0)$, and denote $m_i = n\alpha_i$. In each step of the Swendsen-Wang algorithm, the percolation step for color i picks a graph G_i from $G(m_i, q\alpha_i/m_i)$. Let $C_1^{(i)}, C_2^{(i)}, \dots$ be the components of G_i in decreasing order of size.

Let A, L be the constants in Lemma 18 and let $w \geq L$. For each color i the following hold with positive probability (not depending on n):

1. If $q\alpha_i \geq (1 - Am_i^{-1/3})/m_i$, then $|C_1^{(i)}| \geq 100wq^2n^{2/3}$, $\sum_{j>1} |C_j^{(i)}|^2 \leq m_i^{4/3} \leq n^{4/3}$ (by Lemma 18).⁵
2. If $(1 - Am_i^{-1/3})/m_i > q\alpha_i$, then $\sum_{j \geq 1} |C_j^{(i)}|^2 \leq n^{4/3}$ (by Item 1 of Lemma 17).

Note that for at least 1 color we have $q\alpha_i \geq 1$ (since the α_i 's sum to 1). Let $S = \{i \in [q] : q\alpha_i \geq 1\}$ and consider all the components *different* from $C_1^{(i)}$, $i \in S$. Color these components independently by a uniformly random color from $[q]$. Let A_i be the number of vertices of color i . By Azuma's inequality and a union bound we have that with probability at least $1 - 2q \exp(-50w^2q)$, for each $i \in [q]$ it holds that

$$\left| A_i - \frac{n - \sum_{i \in S} |C_1^{(i)}|}{q} \right| \leq (10wq)n^{2/3}.$$

With probability at least q^{-q} each of $C_1^{(i)}$ with $i \in S$ receives color 1. Let A'_i be the number of vertices of color i after the coloring step of the SW algorithm. Note, we have $A'_1 = A_1 + \sum_{i \in S} |C_1^{(i)}|$ and $A'_i = A_i$ for $i \geq 2$. We obtain that with probability at least $q^{-q}(1 - 2q \exp(-50w^2q)) > 0$

$$|A'_1| \geq \frac{n}{q} + \left(\sum_{i \in S} |C_1^{(i)}| \right) \left(1 - \frac{1}{q} \right) - (10wq)n^{2/3} \geq \frac{n}{q} + (80wq^2)n^{2/3},$$

and for all $i \in \{2, \dots, q\}$

$$|A'_i| \leq \frac{n}{q} - \frac{1}{q} \left(\sum_{i \in S} |C_1^{(i)}| \right) + (10wq)n^{2/3} \leq \frac{n}{q} - (90wq)n^{2/3}.$$

This concludes the proof. \square

Proof of Lemma 37. W.l.o.g., we may assume that the color classes S_1, S_2, \dots, S_q of X_0 satisfy

$$|S_1| \geq \frac{n}{q} + wn^{2/3} \quad \text{and} \quad |S_i| \leq \frac{n}{q} - \frac{w}{2q}n^{2/3} \quad \text{for } i \in \{2, \dots, q\}. \quad (66)$$

Now we make a step of the Swendsen-Wang algorithm. Let C_1, C_2, \dots , be all the connected components after the percolation step of the Swendsen-Wang algorithm, listed in decreasing size. By

⁵We remark that the choice of the constant 100 in the bound for $|C_1^{(i)}|$ is somewhat arbitrary, any sufficient large constant would work; similar remarks apply for the explicit constants 80 and 50 appearing in the proof of Lemma 36.

Lemma 17 (first part for the color classes $i = 2, \dots, q$ and second part for the color class $i = 1$) we have

$$E\left[\sum_{j \geq 2} |C_j|^2\right] \leq \frac{2Kn^{4/3}}{w}.$$

By Markov's inequality

$$P\left(\sum_{j \geq 2} |C_j|^2 \geq n^{4/3}\right) \leq \frac{2K}{w}. \quad (67)$$

By Lemma 17 (part 3), there exists a constant $c > 0$ such that

$$P\left(|C_1| \leq (7/4)wn^{2/3}\right) \leq K \exp(-cq^2w^3). \quad (68)$$

For all sufficiently large w , we may assume that the events in (67) and (68) occurred, that is, $|C_1| \geq (7/4)wn^{2/3}$ and $\sum_{i \geq 2} |C_i|^2 \leq n^{4/3}$. Now we color the components C_2, C_3, \dots independently by a uniformly random color from $[q]$ (for now we leave the component C_1 uncolored). Let A_i be the number of vertices of color i . We have by Azuma's inequality that

$$P\left(\left|A_i - \frac{n - |C_1|}{q}\right| \geq \frac{wn^{2/3}}{4q}\right) \leq 2 \exp(-w^2/(32q^2)). \quad (69)$$

Now we color C_1 , and assume w.l.o.g. that it receives color 1. Let A'_i be the number of vertices of color i now (we have $A'_1 = A_1 + |C_1|$ and $A'_i = A_i$ for $i \geq 2$). Applying union bound to (69) we obtain that with probability at least $1 - 2q \exp(-w^2/(32q^2))$ we have

$$|A'_1| \geq \frac{n}{q} + wn^{2/3} \left(\frac{7}{4}(1 - 1/q) - 1/(4q)\right) \geq \frac{n}{q} + \frac{13}{12}wn^{2/3}, \quad (70)$$

and for all $i \in \{2, \dots, q\}$

$$|A'_i| \leq \frac{n}{q} - wn^{2/3} \left(\frac{7}{4}(1/q) - 1/(4q)\right) \leq \frac{n}{q} - \frac{13}{12}w/(2q)n^{2/3}. \quad (71)$$

Note that, in the second inequality in (70), we used the fact that $q \geq 3$.

Let $w' = (13/12)w$. Summarizing all the steps we obtain that from a state satisfying (66) we get to a state satisfying

$$|S_1| \geq \frac{n}{q} + w'n^{2/3} \quad \text{and} \quad |S_i| \leq \frac{n}{q} - \frac{w'}{2q}n^{2/3} \quad \text{for } i \in \{2, \dots, q\}, \quad (72)$$

with probability at least

$$\left(1 - \frac{2K}{w} - K \exp(-cq^2w^3)\right) \left(1 - 2q \exp(-w^2/(32q^2))\right). \quad (73)$$

For all sufficiently large w , the last expression is greater than $\exp(-C/w)$, where C is a positive constant (depending on K, c, q), as wanted. \square

9 Lower bound on the mixing time for $B \geq \mathfrak{B}_{rc}$

In this section, we prove that the SW algorithm mixes in $\Omega(\log n)$ steps for all $B > \mathfrak{B}_{rc}$.

Recall from Section 5 that $\mathcal{B}(\mathbf{v}, \delta)$ is the ℓ_∞ -ball of configuration vectors of the q -state Potts model in K_n around \mathbf{v} of radius δ , cf. equation (34). Let

$$S := \mathcal{B}(\mathbf{m}, n^{-1/7}),$$

and denote the set of configuration vectors which are not in S by \bar{S} .

We first establish the following (crude) bound on the probability mass of configurations in \bar{S} in the Potts distribution. (Far more precise bounds are known and can be found in, e.g., [10]; the following estimate follows easily from our upper bound on the mixing time.)

Lemma 38. *Let $B \geq \mathfrak{B}_{rc}$. For the Potts distribution μ in (1), for all sufficiently large n , it holds that $\mu(\bar{S}) \leq 1/8$.*

Proof. For any starting state X_0 , we have that for $T = O(\log n)$, it holds that

$$\Pr(X_T \in S) \geq \varepsilon,$$

where $\varepsilon > 0$ is a constant independent of n . (For $B > \mathfrak{B}_{rc}$ this follows by Lemmas 30 and 31, and for $B = \mathfrak{B}_{rc}$ this follows by Lemmas 34 and 31.) It follows that for all non-negative integers j it holds that

$$\Pr(X_{(j+1)T} \in S \mid X_{jT} \notin S) \geq \varepsilon.$$

Further, by Lemma 23, for integer $t \geq 0$, it holds that

$$\Pr(X_{t+1} \in S \mid X_t \in S) \geq 1 - \exp(-\Omega(n^{1/3})).$$

We thus obtain that for some positive integer $j = j(\varepsilon)$, for all sufficiently large n , for all integer $t \geq jT$, it holds that

$$\Pr(X_t \in S) \geq 15/16. \tag{74}$$

Let $T^* = \max\{jT, 2T_{\text{mix}}\}$. Recall that $T_{\text{mix}} = O(\log n)$ (cf. Corollaries 32 and 35), so $T^* = O(\log n)$ as well. Since T_{mix} is the time needed to get within total variation distance $\leq 1/4$ from μ , we have that for any $\varepsilon' > 0$, for $t \geq T_{\text{mix}} \log_2(1/\varepsilon')$, it holds that $d_{TV}(X_t, \mu) \leq \varepsilon'$ (see [18, Section 4.5]). Thus, we have that

$$\mu(\bar{S}) - \Pr(X_{T^*} \in \bar{S}) \leq \max_{A \subseteq \Omega} |\mu(A) - \Pr(X_{T^*} \in A)| = d_{TV}(X_{T^*}, \mu) \leq 1/16. \tag{75}$$

Combining (74) and (75) yields $\mu(\bar{S}) \leq 1/8$, as wanted. \square

Lemma 39. *For $B \geq \mathfrak{B}_{rc}$, there exist constants $\delta_1, \delta_2 > 0$ such that the following is true. Suppose that we start at a state X_0 such that $X_0 \notin S$ and $\delta_2 \leq \|\alpha(X_0) - \mathbf{m}\|_\infty \leq \delta_1$. Then for some $T = \Omega(\log n)$, with probability $\geq 1/2$, it holds that $X_T \notin S$.*

Proof of Lemma 39. Recall that $\mathbf{m} = (a, b, \dots, b)$ where $a > 1/q$ is a fixpoint of F . Let $\delta > 0$ be such that for some $0 < c_l < c_u < 1$ for all $z \in [a - \delta, a + \delta]$ we have

$$c_l |z - a| \leq |F(z) - a| \leq c_u |z - a|. \tag{76}$$

Note that the existence of such δ is guaranteed throughout the regime $B \geq \mathfrak{B}_{rc}$, since $|F'(a)| < 1$ by Lemma 4, $F'(a) > 0$ by Lemma 8 and F' is continuous in a neighbourhood around a . Let δ_1, δ_2 be arbitrary constants satisfying $0 < \delta_2 < \delta_1 < \delta$.

Suppose that we are at X_t such that $n^{-1/7} < \|\alpha(X_t) - \mathbf{m}\|_\infty \leq \delta$ (note that for such X_t , we have $X_t \notin S$). Let m_1 be the number of vertices in the largest color class and note that $m_1/n = a + \tau =: a'$ where $|\tau| < \delta$. Exactly as in the proof of Lemma 23 (cf. equation (39)), we obtain that with probability $\geq 1 - 2q \exp(-\Theta(n^{1/3}))$ it holds that

$$\left\| \alpha(X_{t+1}) - \left(F(a'), \frac{1 - F(a')}{q-1}, \dots, \frac{1 - F(a')}{q-1} \right) \right\|_\infty \leq n^{-1/6}. \quad (77)$$

Using (76), we have

$$c_l \|\alpha(X_t) - \mathbf{m}\|_\infty \leq \left\| \left(F(a'), \frac{1 - F(a')}{q-1}, \dots, \frac{1 - F(a')}{q-1} \right) - \mathbf{m} \right\|_\infty \leq c_u \|\alpha(X_t) - \mathbf{m}\|_\infty. \quad (78)$$

Equations (77) and (78) combined yield that for all sufficiently large n we have the following two bounds:

$$\|\alpha(X_{t+1}) - \mathbf{m}\|_\infty \geq c_l \|\alpha(X_t) - \mathbf{m}\|_\infty - n^{-1/6} \geq \frac{c_l}{2} \|\alpha(X_t) - \mathbf{m}\|_\infty, \quad (79)$$

$$\|\alpha(X_{t+1}) - \mathbf{m}\|_\infty \leq n^{-1/6} + c_u \|\alpha(X_t) - \mathbf{m}\|_\infty \leq \delta. \quad (80)$$

Let $c' = -\frac{1}{8}/\log(\frac{c_l}{2})$. Applying (79) for $t = 0, \dots, \lfloor c' \log n \rfloor$ (note that (80) guarantees that we remain sufficiently close to \mathbf{m} so that (79) indeed applies), we obtain that with probability $1 - o(1)$ it holds that

$$\|\alpha(X_{c' \log n}) - \mathbf{m}\|_\infty \geq n^{-1/8} \|\alpha(X_0) - \mathbf{m}\|_\infty \geq \delta_2 n^{-1/8} > n^{-1/7}.$$

This completes the proof. \square

Using Lemma 39, we obtain the following corollary.

Corollary 40. *Let $B \geq \mathfrak{B}_{rc}$. Then the mixing time T_{mix} of the SW dynamics on the n -vertex complete graph satisfies $T_{mix} = \Omega(\log n)$.*

Proof. Let δ_1, δ_2 be as in Lemma 39. Consider X_0 such that $X_0 \notin S$ and $\delta_2 \leq \|\alpha(X_0) - \mathbf{m}\|_\infty \leq \delta_1$. Then, by Lemma 39, for some $T = \Omega(\log n)$ we have that

$$\Pr(X_T \notin S) \geq 1/2.$$

On the other hand, by Lemma 38 we have that $\mu(\bar{S}) \leq 1/8$. It follows that

$$d_{TV}(X_T, \mu) = \max_{A \subseteq \Omega} |\mu(A) - \Pr(X_T \in A)| \geq \Pr(X_T \in \bar{S}) - \mu(\bar{S}) \geq 1/2 - 1/8 > 1/4.$$

It follows from the definition of mixing time that $T_{mix} \geq T$, as claimed. \square

10 Fast mixing for $B < \mathfrak{B}_u$

In this section, we prove that the SW algorithm mixes in $O(1)$ steps for all $B > \mathfrak{B}_{rc}$. The proof for establishing mixing in the uniqueness regime will be similar to the $B > \mathfrak{B}_{rc}$ case. We begin with the following analogue of Lemma 28.

Lemma 41. *Assume $B < \mathfrak{B}_{rc}$ is a constant. There exists a constant $\varepsilon > 0$ such that for, any initial state X_0 , with probability $\Theta(1)$ the next state X_1 has at least $q - 1$ colors that are ε -light.*

Proof. The proof is analogous to that of Lemma 28, the only difference is that now we do not need to argue that there is an ε -heavy color (and hence the proof is simpler).

Let $\varepsilon \in (0, 1/10)$ be a small enough constant such that $B(1 + 2\varepsilon) < q$. As in the proof of the first part of Lemma 28 with probability $q^{-q} = \Theta(1)$ all the biggest components of each color class receive the color 1 and the sum of squares of (the sizes of) the remaining components is $o(n^2)$ with probability $\Theta(1)$. Condition on these events happening. Then, the expected number of vertices that receive a color $i = 2, \dots, q$ is at most n/q . Therefore, using Azuma's inequality, with probability $\Theta(1)$, there are at most $n(1 + \varepsilon/2)/q$ vertices which have color $i = 2, \dots, q$ in X_1 . By the choice of ε , we have that $(1 + \varepsilon/2)/q \leq (1 + \varepsilon/2)/(B(1 + 2\varepsilon)) \leq (1 - \varepsilon)/B$, and therefore there are $q - 1$ colors which are ε -light in X_1 . \square

We then have the following lemma, which is an analogue of Lemmas 30 and 31 in the $B > \mathfrak{B}_{rc}$ case, showing that we get within distance $O(n^{-1/2})$ from the uniform phase.

Lemma 42. *Assume $B < \mathfrak{B}_u$ is a constant. There exists a constant L such that for any starting state X_0 after $T = O(1)$ steps with probability $\Theta(1)$ the SW algorithm moves to state X_T such that $\|\alpha(X_T) - \mathbf{u}\|_\infty \leq Ln^{-1/2}$.*

Proof of Lemma 42. Let $\varepsilon > 0$ be as in Lemma 41. By Lemma 41 starting from any X_0 with constant probability we move to X_1 where $q - 1$ colors are ε -light. As in Lemma 30, the evolution of the largest color class is then captured by the iterates of the function F . Since $1/q$ is the only fixpoint of F (by Lemma 4), we have that for any constant $\delta > 0$ there exists constant T such that $F^{(T)}([0, 1]) \subseteq [1/q - \delta/2, 1/q + \delta/2]$. Therefore, with probability $1 - o(1)$, after at most T steps the size of the largest color class becomes less than $1/q + \delta$ (see the proof of Lemma 30 for details). In the next step even the largest color class is subcritical (by taking δ to be a small constant) and we end up, with probability $1 - o(1)$, in a state where each color occurs $(1 + o(1))n/q$ times. In the next step the components sizes after the percolation step satisfy, by Lemma 17

$$E\left[\sum_i |C_i|^2\right] = O(n).$$

Hence, after the coloring step, with constant probability (using the same argument as in (57) and (58)) we have color classes of size $(n + O(n^{1/2}))/q$. \square

To show that the mixing time of SW is $O(1)$ when $B < \mathfrak{B}_u$, we extend the strategy of [19] for $q = 2$ to $q \geq 3$. In [19], a certain projection of the SW chain is defined, called the magnetization chain. For us, the magnetization chain can be defined as follows. Let $\{V_1, \dots, V_q\}$ be a fixed partition of the vertex set of the complete graph into q parts. The magnetization chain is a Markov chain $\mathcal{A}_t = (A_{ij,t})_{i,j \in [q]}$ with $A_{ij,t}$ being the number of vertices in V_i with color j at time t (the fact that the magnetization chain is a Markov chain is due to the symmetry). Note that for every $t = 0, 1, \dots$, for every $i \in [q]$ it holds that $\sum_j A_{ij,t} = |V_i|$.

The following lemma is the analogue of [19, Proposition 7.3] and can be proved analogously to Lemma 27.

Lemma 43. *Assume $B < \mathfrak{B}_{rc}$ is a constant. Let $\{V_1, \dots, V_q\}$ be a partition of the vertex set of the complete graph on n vertices into q parts. Let \mathcal{A}_t and \mathcal{A}'_t be two copies of the magnetization chain. Further, denote by $a_{j,t}, a'_{j,t}$ the total number of vertices with color j in \mathcal{A}_t and \mathcal{A}'_t , respectively, i.e.,*

$$a_{j,t} = \sum_{i \in [q]} A_{ij,t}, \quad a'_{j,t} = \sum_{i \in [q]} A'_{ij,t}.$$

Let $L > 0$ be an arbitrarily large constant and suppose that at time t it holds that

$$|a_{j,t} - n/q| \leq L\sqrt{n}, \quad |a'_{j,t} - n/q| \leq L\sqrt{n} \text{ for all } j \in [q].$$

Then, there exists a coupling of $\mathcal{A}_{t+1}, \mathcal{A}'_{t+1}$ such that with probability $\Theta(1)$, it holds that $\mathcal{A}_{t+1} = \mathcal{A}'_{t+1}$.

Proof. The proof is completely analogous to [19, Proof of Proposition 7.3] and resembles the proof of Lemmas 26 and 27 given earlier. We therefore highlight the key differences.

Perform the percolation step of the Swendsen-Wang algorithm independently for the chains \mathcal{A}_t and \mathcal{A}'_t . By Lemma 5.7 in [19], there is a constant $c > 0$ such that with probability $\Theta(1)$, in each chain, in each part V_i , there are $\geq c|V_i|$ isolated vertices (i.e., components of size 1). Next, perform the coloring step in each of the two chains independently but leaving, in each chain and for each part V_i , these $c|V_i|$ isolated vertices uncolored. For $i, j \in [q]$, let $\hat{a}_{ij}, \hat{a}'_{ij}$ be the number of vertices which are assigned color j in part V_i (excluding the $c|V_i|$ isolated vertices which are not yet colored). We claim that there exists a (large) constant $L > 0$ such that with probability $\Theta(1)$, for all $i, j \in [q]$, it holds that

$$|\hat{a}_{ij} - \hat{a}'_{ij}| \leq L\sqrt{|V_i|}. \quad (81)$$

Assuming this, then, just as in the proof of Lemmas 27 and 26 (cf. (43) and the coupling thereafter), we can couple the coloring of the $c|V_i|$ isolated vertices in each part V_i to equalize the counts with probability $\Theta(1)$, i.e., with probability $\Theta(1)$, the coupling of the two chains satisfies $\mathcal{A}_{t+1} = \mathcal{A}'_{t+1}$.

We focus therefore on proving (81). Let $\{C_k\}_{k \geq 1}$ denote the components in the first chain after the percolation step. Then, for each $i \in [q]$, we will show that with probability $\Theta(1)$ it holds that

$$\sum_{k \geq 1} |C_k \cap V_i|^2 = O(|V_i|). \quad (82)$$

To see this, for a vertex v , let $C(v)$ be the component that v belongs to after the percolation step. Then, note that

$$\sum_{k \geq 1} |C_k \cap V_i|^2 \leq \sum_{v \in V_i} |C(v)|.$$

Since by the assumption of the lemma all colors are subcritical in the percolation step, we have that $E[|C(v)|] = O(1)$ for all $v \in V$. Using Markov's inequality, we therefore obtain (82).

Let \hat{n}_i be the number of vertices in V_i excluding the isolated vertices. We obtain (using Azuma's inequality) that, with probability $\Theta(1)$, for all $i, j \in [q]$ it holds that

$$|\hat{a}_{ij} - \hat{n}_i/q| = O(\sqrt{|V_i|})$$

Identically, we obtain an analogous bound for \hat{a}'_{ij} which yields (81), as needed. \square

Using Lemmas 42 and 43, we conclude the following corollary.

Corollary 44. *Let $B < \mathfrak{B}_u$ be a constant. The mixing time of the Swendsen-Wang algorithm on the complete graph on n vertices is $\Theta(1)$.*

Proof. Let μ be the stationary distribution of the Swendsen-Wang algorithm (cf. (1)). Consider two copies of the SW algorithm X_t and Y_t , where X_0 is an arbitrary starting configuration and Y_0 is distributed according to μ . It suffices to show that there is $T = O(1)$ and a coupling of X_T, Y_T such that $X_T = Y_T$ with probability $\Omega(1)$.

We will use the magnetization chain for an appropriate partition $\{V_1, \dots, V_q\}$ of the vertices of the complete graph. Namely, for a color $i \in [q]$, let V_i be the set of vertices with color i in X_0 . Let $\mathcal{A}_t = \{A_{ij,t}\}_{i,j \in [q]}$, $\mathcal{A}'_t := \{A'_{ij,t}\}_{i,j \in [q]}$ be such that $A_{ij,t}$, $A'_{ij,t}$ is the number of vertices with color j in V_i in X_t and Y_t , respectively. The key idea is that, due to symmetry, the probability that the SW chain at time t is at a particular configuration σ depends only on the counts $|V_i \cap \sigma^{-1}(j)|$ for $i \in [q]$ and $j \in [q]$. It follows that for every t , it holds that

$$d_{TV}(X_t, Y_t) = d_{TV}(\mathcal{A}_t, \mathcal{A}'_t). \quad (83)$$

It thus suffices to show that for $T = O(1)$, there is a coupling of \mathcal{A}_T and \mathcal{A}'_T such that $\mathcal{A}_T = \mathcal{A}'_T$ with probability $\Theta(1)$.

Let L be the constant in Lemma 42. By Lemma 42, we have that for $T_1 = O(1)$, with probability $\Theta(1)$ it holds that

$$\|\alpha(X_{T_1}) - \mathbf{u}\|_\infty \leq Ln^{-1/2}, \quad \|\alpha(Y_{T_1}) - \mathbf{u}\|_\infty \leq Ln^{-1/2}. \quad (84)$$

Conditioned on (84), Lemma 43 shows that there exists a coupling of \mathcal{A}_{T_1+1} and \mathcal{A}'_{T_1+1} such that with probability $\Theta(1)$ it holds that $\mathcal{A}_{T_1+1} = \mathcal{A}'_{T_1+1}$. Using (83), we thus conclude that the mixing time of the Swendsen-Wang algorithm is $O(1)$, as wanted. \square

11 Mixing Time at $B = \mathfrak{B}_u$

For $B = \mathfrak{B}_u$, our goal is to show that the SW chain reaches the uniform phase in $O(n^{1/3})$ steps. To do this, let S_t be the size of the largest color class in state X_t of the SW chain; throughout this section, we will focus on tracking S_t .

In Section 11.1, we first give some relevant statistics of S_t after one iteration of the SW chain; the main lemma we will use later is Lemma 47. In Sections 11.2 and 11.3, we use these statistics to outline our potential function argument for deriving the upper and lower bounds on the mixing time. Finally, in Section 11.4, we give in detail the construction of the potential function which is the most technical part of the proof.

11.1 Tracking one iteration of the SW dynamics

As a starting point, we have the following analogue of Lemma 28.

Lemma 45. *For sufficiently small (constant) $\varepsilon > 0$, for any state X_t of the SW chain, with probability $\Theta(1)$, there are at least $q - 1$ colors in state X_{t+1} which are ε -light. Further, if state X_t has $q - 1$ colors which are ε -light, then with probability $1 - \exp(-n^{\Omega(1)})$, the same is true for X_{t+1} .*

Proof. For a color $i \in [q]$, we will write α_i as a shorthand for $\alpha_i(X_t)$, and denote $m_i = n\alpha_i$. In each step of the Swendsen-Wang algorithm, the percolation step for color i picks a graph G_i from $G(m_i, B\alpha_i/m_i)$. Let $C_1^{(i)}, C_2^{(i)}, \dots$ be the components of G_i in decreasing order of size.

The beginning of the proof is analogous to the beginning of the proof of Lemma 36. Let A be the constant in Lemma 18. For each color $i \in [q]$ the following hold with positive probability (not depending on n):

1. If $B\alpha_i \geq (1 - Am_i^{-1/3})/m_i$, then $\sum_{j \geq 1} |C_j^{(i)}|^2 \leq m_i^{4/3} \leq n^{4/3}$ (by Lemma 18).
2. If $(1 - Am_i^{-1/3})/m_i > B\alpha_i$, then $\sum_{j \geq 1} |C_j^{(i)}|^2 \leq n^{4/3}$ (by Item 1 of Lemma 17).

Let $S = \{i \in [q] : B\alpha_i \geq 1\}$ (note that the set S may be empty). Consider all the components *different* from $C_1^{(i)}$, $i \in S$. Color these components independently by a uniformly random color from $[q]$. For $i \in [q]$, let A'_i be the number of vertices of color i . Let $w > 0$ be a constant such that $1 > 2q \exp(-w^2/2)$. By Azuma's inequality and a union bound we have that with probability at least $1 - 2q \exp(-w^2/2) > 0$, for each $i \in [q]$ it holds that

$$\left| A'_i - \frac{n - \sum_{i \in S} |C_1^{(i)}|}{q} \right| \leq wn^{2/3}.$$

For $i \in [q]$, let A_i be the number of vertices of color i after the coloring step of the SW algorithm. With probability at least q^{-q} each of $C_1^{(i)}$ with $i \in S$ receives color 1. Note, we have $A_1 = A'_1 + \sum_{i \in S} |C_1^{(i)}|$ and $A_i = A'_i$ for $i \geq 2$. We obtain that with probability at least $q^{-q}(1 - 2q \exp(-w^2/2)) > 0$, for all $i \geq 2$,

$$|A_i| \leq \frac{n}{q} - \frac{1}{q} \left(\sum_{i \in S} |C_1^{(i)}| \right) + wn^{2/3} \leq \frac{n}{q} + wn^{2/3}. \quad (85)$$

Since $\mathfrak{B}_u < q$, we have that for sufficiently small constant $\varepsilon > 0$, for all n sufficiently large, it holds that $|A_i| \leq (1 - \varepsilon)n/B$ for all $i \neq 1$, and thus the colors $2, \dots, q$ are ε -light with probability $\Theta(1)$ as wanted.

For the second part of the lemma where we know that in X_t there are $q - 1$ colors which are ε -light, the proof is analogous. The difference is that now we need upper bounds for the sum of squares of the components (other than the largest component — there can be at most one of those by the assumption) which hold with probability $1 - \exp(-n^{\Omega(1)})$. Note, for a color class i , we have the (crude) bounds

$$\sum_{j \geq 1} |C_j^{(i)}|^2 \leq n |C_1^{(i)}| \text{ and } \sum_{j \geq 2} |C_j^{(i)}|^2 \leq n |C_2^{(i)}|. \quad (86)$$

For each of the $(q - 1)$ ε -light colors, the first inequality in (86) together with Lemma 16 bounds the sum of squares of the components by $n^{7/4}$ with probability $1 - \exp(-\Theta(n^{3/4}))$. For the remaining color class (i.e., the one that we do not have an upper bound on its density by the assumption), to bound the sum of squares of the components we obtain the same bound $n^{7/4}$ with probability $1 - \exp(-n^{\Omega(1)})$ by considering cases. If the color class is supercritical we use Lemma 14 and the second inequality in (86). If the color class is in the critical window we use Lemma 20 and the first inequality in (86). If the color class is subcritical we use Lemma 16 and the first inequality in (86). The only modification needed in the argument is to replace $wn^{2/3}$ in (85) by $n^{9/10}$ and the remaining part holds verbatim. \square

The key part of our arguments is to track the evolution of the size S_t of the largest colors when there are $q - 1$ colors which are ε -light.

We first do this in the easier case when S_t has density close to $1/B$ (in the complementary regime, we will need more statistics of S_t). In this regime, the following lemma roughly says that a step of the SW dynamics makes the density of the largest color class roughly $1/q$. (Intuitively, this follows by a “continuity” argument since $F(1/B) = 1/q$.)

Lemma 46. *Let $\varepsilon > 0$ be a sufficiently small constant. Suppose that X_t is such that $q - 1$ colors are ε -light and that $S_t < (1 + \varepsilon)n/B$. Then with probability $1 - \exp(-n^{\Omega(1)})$ it holds that $S_{t+1} < (1 + 3q\varepsilon)n/q$.*

Proof of Lemma 46. The proof is analogous to the proof of Lemma 45 and as such we follow the notation in there. The only difference is that now we have to account slightly more accurately for the size of the largest color class in X_{t+1} .

Assume that the $q - 1$ ε -light colors in X_t are $2, \dots, q$ and assume w.l.o.g. that (the perhaps linear sized) $C_1^{(1)}$ gets colored with color 1 (in state X_{t+1}). The color classes of $2, \dots, q$ in X_t are subcritical and thus fall into Item 2 of the analysis in the proof of Lemma 45. For the remaining color class 1 in X_t , it may fall either into Item 1 or 2.

It follows that the bounds for A'_i in (85) still hold and in particular the colors $2, \dots, q$ have size at most $(1/q)n + o(n)$ (since they did not receive a giant component).

For the color class 1 in X_{t+1} , note that $A_1 = |C_1^{(1)}| + A'_1$. For all sufficiently small (constant) $\varepsilon > 0$, the largest component $C_1^{(1)}$, with probability $1 - \exp(-n^{\Omega(1)})$, has size at most $3\varepsilon(n/B)$ (by Item 3 of Lemma 17). Note that $\mathfrak{B}_u \geq 1$ for all $q \geq 3$ (follows, e.g., by definition (3)) and hence $3\varepsilon(n/B) \leq 3\varepsilon n$. It follows that for all sufficiently large n , A_1 is at most $(1 + 3q\varepsilon)n/q$, as wanted. \square

The following lemma gives some statistics of S_t/n throughout the range $(1/B, 1]$, i.e., when the largest color class is supercritical in the percolation step of the SW dynamics. Recall the function F defined in (6),(7).

Lemma 47. *Let $\varepsilon > 0$ be an arbitrarily small constant and condition on the event that X_t has $q - 1$ colors which are ε -light. Assume that ζ satisfies $(1 + \varepsilon)/B \leq \zeta/n \leq 1$. Let $W := E[S_{t+1} | S_t = \zeta]$.*

For all constant $\varepsilon' > 0$, for all sufficiently large n , it holds that

$$nF(\zeta/n) - n^{\varepsilon'} \leq W \leq nF(\zeta/n) + n^{\varepsilon'}. \quad (87)$$

Also, there exist absolute constants Q_1, Q_2 (depending only on ε) such that

$$nQ_2 \leq \text{Var}[S_{t+1} | S_t = \zeta] \leq nQ_1, \quad (88)$$

Finally, for every integer $k \geq 3$ and constant $\varepsilon' > 0$, there exists a constant $c > 0$ such that

$$E[|S_{t+1} - W|^k | S_t = \zeta] \leq cn^{k/2 + \varepsilon'}. \quad (89)$$

Proof. To avoid overloading notation, we assume throughout that we condition on $S_t = \zeta$.

We will write α_i as a shorthand for $\alpha_i(X_t)$, and denote $m_i = n\alpha_i$. W.l.o.g. we will assume that the color class with largest size is the one corresponding to color 1, so that $\alpha_1 = \zeta/n \geq (1 + \varepsilon)/B$. Since the remaining $(q - 1)$ colors are ε -light, for each $i \in \{2, \dots, q\}$ we have $\alpha_i \leq (1 - \varepsilon)/B$.

In each step of the Swendsen-Wang algorithm, the percolation step for color i picks a graph G_i from $G(m_i, B\alpha_i/m_i)$. Let $C_1^{(i)}, C_2^{(i)}, \dots$ be the components of G_i in decreasing order of size. Note that G_1 is in the supercritical regime, while G_2, \dots, G_q are in the subcritical regime. By Lemma 15, for every constant $\varepsilon' > 0$ we have that

$$E[|C_1^{(1)}|] = \beta\zeta \pm \zeta^{\varepsilon'} = \beta\zeta \pm n^{\varepsilon'}, \quad (90)$$

where $\beta \in (0, 1)$ satisfies $\beta + \exp(-\beta \frac{B\zeta}{n}) = 1$. Note that

$$nF(\zeta/n) = \frac{n}{q} + \left(1 - \frac{1}{q}\right)\beta\zeta.$$

Let A_i be the number of vertices with color i in X_{t+1} and w.l.o.g. assume that $C_1^{(1)}$ receives the color 1 in the coloring step of the SW dynamics. We will show that with probability $1 - \exp(-n^{\Omega(1)})$ it holds that $S_{t+1} = A_1$, so the estimates on the moments of S_{t+1} will follow from those of A_1 .

More precisely, with a scope to also prove (89), we will show that for every sufficiently small constant $\varepsilon' > 0$ it holds with probability $1 - \exp(-\Theta(n^{-\varepsilon'}))$ that

$$|A_1 - nF(\zeta/n)| \leq 2n^{1/2+\varepsilon'} \text{ and } A_i \leq (n - \beta\zeta)/q + n^{1/2+\varepsilon'} \text{ for } i \in \{2, \dots, q\}. \quad (91)$$

Since $\beta\zeta = \Omega(n)$, we will then obtain that $A_1 > A_i$ for all $i \neq 1$.

From Lemma 14 equation (27) (applied to color $i = 1$) and Lemma 16 (applied to colors $i = 2, \dots, q$), with probability $1 - q \exp(-\Theta(n^{\varepsilon'}))$, we have

$$|C_j^{(1)}| \leq n^{\varepsilon'} \text{ for } j \geq 2, \quad |C_j^{(i)}| \leq n^{\varepsilon'} \text{ for } i \in \{2, \dots, q\}, j \geq 1. \quad (92)$$

From Lemma 14 equation (28), with probability $1 - \exp(-\Theta(n^{\varepsilon'}))$, we also have

$$|C_1^{(1)} - \beta\zeta| \leq n^{1/2+\varepsilon'}. \quad (93)$$

(Note that $\beta\zeta = \Omega(n)$.) Condition on the event that the bounds in (92) and (93) hold. From (92), we have the crude bound

$$\sum_{j \geq 2} (|C_j^{(1)}|)^2 + \sum_{q \geq i \geq 2} \sum_{j \geq 1} (|C_j^{(i)}|)^2 \leq n^{1+\varepsilon'}. \quad (94)$$

Consider now the coloring step of the SW algorithm and, in particular, color independently all the components *different* from $C_1^{(1)}$ by a uniformly random color from $[q]$. Let A'_i be the number of vertices of color i in this process. Note that $A_1 = |C_1^{(1)}| + A'_1$ and $A_i = A'_i$ for $i = 2, \dots, q$. Using (94), by Azuma's inequality we have that with probability $1 - 2q \exp(-n^{\varepsilon'})$ for all $i \in [q]$ it holds that

$$\left| A'_i - \frac{n - |C_1^{(1)}|}{q} \right| \leq n^{1/2+\varepsilon'}. \quad (95)$$

From (93) and (95) we obtain that for all sufficiently large n , with probability $1 - \exp(-\Theta(n^{\varepsilon'}))$ it holds that $S_{t+1} = A_1$. It follows that $E[S_{t+1}] = E[A_1] + o(1)$ and $E[|S_{t+1} - E[S_{t+1}]|^k] = E[|A_1 - E[A_1]|^k] + o(1)$ for all integer $k \geq 2$. Thus, the bounds in (87), (88), (89) will follow from

$$E[A_1] = nF(\zeta/n) \pm n^{\varepsilon'}, \quad (96)$$

$$Q_1 n \leq \text{Var}[A_1] \leq Q_2 n, \quad (97)$$

$$E[|A_1 - E[A_1]|^k] \leq K n^{k/2+\varepsilon'}, \quad (98)$$

where $k \geq 3$ is an integer, $\varepsilon' > 0$ is an arbitrarily small constant, $Q_1, Q_2 > 0$ are absolute constants and K is a constant depending on k .

We start by proving (96) and (97) where we need more precise bounds. By the second inequality in (32) of Lemma 15 (applied to color 1) and part 1 of Lemma 17 (applied to colors $i = 2, \dots, q$), we have for some constants $K_1, K_2, K_3 > 0$ that

$$K_1 n \leq \text{Var}[|C_1^{(1)}|] \leq K_2 n, \quad E\left[\sum_{j \geq 2} (|C_j^{(1)}|)^2 + \sum_{q \geq i \geq 2} \sum_{j \geq 1} (|C_j^{(i)}|)^2\right] \leq K_3 n. \quad (99)$$

Denote by \mathcal{C} the random vector $\{|C_j^{(i)}|\}_{i \in [q], j \geq 1}$. We first estimate the moments of A_1 conditioned on \mathcal{C} . We have

$$E[A_1 | \mathcal{C}] = |C_1^{(1)}| + \frac{n - |C_1^{(1)}|}{q} = \frac{n}{q} + \left(1 - \frac{1}{q}\right)|C_1^{(1)}|, \quad (100)$$

$$\text{Var}[A_1 | \mathcal{C}] = \frac{1}{q} \left(1 - \frac{1}{q}\right) \left[\sum_{j \geq 2} (|C_j^{(1)}|)^2 + \sum_{q \geq i \geq 2} \sum_{j \geq 2} (|C_j^{(i)}|)^2 \right]. \quad (101)$$

It follows from (100) that $E[A_1] = \frac{n}{q} + (1 - \frac{1}{q})E[|C_1^{(1)}|]$, so (96) follows from (90). Also, by the law of total variance we have $\text{Var}[A_1] = \text{Var}[E[A_1 | \mathcal{C}]] + E[\text{Var}[A_1 | \mathcal{C}]]$, so from (90), (99), (100), we obtain (97).

Finally, it remains to prove (98). Let $\varepsilon'' := \varepsilon'/k > 0$. By the triangle inequality and (96) (applied for the constant ε''), we have that

$$|A_1 - E[A_1]| \leq |A_1 - nF(\zeta/n)| + |E[A_1] - nF(\zeta/n)| \leq |A_1 - nF(\zeta/n)| + n\varepsilon'',$$

and hence by the AM-GM inequality we have

$$|A_1 - E[A_1]|^k \leq 2^{k-1} (|A_1 - nF(\zeta/n)|^k + n^{k\varepsilon'}).$$

By integrating the first inequality in (91) (applied for the constant ε''), we obtain that $E[|A_1 - nF(\zeta/n)|^k] \leq 2^k n^{k/2+\varepsilon'} + o(1)$. Combining these bounds yields (98) with $K = 2^{3k}$ (to absorb the lower order terms $n^{\varepsilon'}$ and $o(1)$).

This concludes the proof of Lemma 47. \square

11.2 Upper bound on the mixing time at $B = \mathfrak{B}_u$

In this section, we prove that the mixing time of the SW chain satisfies $T_{\text{mix}} = O(n^{1/3})$ at the critical point $B = \mathfrak{B}_u$.

The most difficult part of our arguments is to argue that the SW chain escapes the vicinity of the majority phase in $O(n^{1/3})$ steps, i.e., when the size S_t of the largest color class is in the window $|S_t - na| \leq \delta n^{2/3}$ for some small constant $\delta > 0$ (recall that a is the marginal of the majority phase and satisfies $F(a) = a$, see also Lemma 4). Note that from (87) we have that $E[S_{t+1} | S_t] \approx nF(S_t/n)$ and hence the drift of the process inside the window is very weak; for example, when $S_t/n = a$, the expected value of S_{t+1}/n remains very close to a . More generally, an expansion of F around the point a yields that $F(z) \approx z - c(z - a)^2$ for all $z \in (a - \varepsilon, a + \varepsilon)$ for some constants $c, \varepsilon > 0$. Therefore, we obtain that $E[S_{t+1} | S_t] \approx S_t - c(S_t - an)^2/n$ for some constant $c > 0$, so the change (in expectation) of S_{t+1} relative to S_t is bounded above by roughly $\delta^2 n^{1/3}$. In particular, how does the process escape the window $|S_t - na| \leq \delta n^{2/3}$ in $O(n^{1/3})$ steps?

The rough intuition is that inside the window the variance of the process aggregates the right way and the process gets displaced (with constant probability) by the square root of the “aggregate variance”. That is, after $\Omega(n^{1/3})$ steps, S_t is displaced by roughly $\Omega(\sqrt{n^{1/3}n}) = \Omega(n^{2/3})$ from na . In the meantime, it holds that $F(z) \leq z$ for all $z \in [1/B, 1]$ so S_t is bound to escape from the lower end of the window. Once S_t escapes the window, the drift $F(z) - z$ coming from the expectation of S_t/n takes over and the trajectory of S_t/n is close to a deterministic process $z(t)$ which satisfies the differential equation $dz = (F(z) - z)dt$. Since $F(z) \approx z - c(z - a)^2$ for all $z \in (a - \varepsilon, a + \varepsilon)$, we obtain that the number of steps needed so that S_t/n goes from $a - n^{-1/3}$ to $a - \varepsilon$ is roughly $\int_{a-n^{-1/3}}^{a-\varepsilon} \frac{1}{F(z)-z} dz \approx n^{1/3}$; from that point on, the SW chain will get within constant distance from

the uniform phase in $O(1)$ steps.⁶ Rather than formalizing explicitly this intuition, we will capture the progress of the chain towards the uniform phase by a potential function argument.

The potential function is designed so that its maximum value is at most $O(n^{1/3})$ and, at each step of the SW chain, the expected decrease of the potential function is at least a constant. More precisely, we show the following lemma in Section 11.4.

Lemma 48. *Let $B = \mathfrak{B}_u$. There exist constants $M_1, M_2, \tau > 0$ such that for all sufficiently small $\varepsilon > 0$, for all sufficiently large n the following holds. There exists an increasing three-times differentiable potential function $G : [1/q, 1] \rightarrow [0, M_1 n^{1/3}]$ with $G(1/q) = 0$ and $\max_{z \in [1/q, 1/B]} G'(z) \leq M_2$ such that for any $\zeta \geq (1 + \varepsilon)n/B$, if X_t has $(q - 1)$ colors which are ε -light, then it holds that*

$$E[G(S_{t+1}/n) | S_t = \zeta] \leq G(\zeta/n) - \tau. \quad (102)$$

The proof of Lemma 48 is quite technical, so let us briefly discuss the main ideas underlying the proof. The crucial ingredient is to specify the potential function G so that (102) is satisfied. To motivate the choice of G , by taking expectations in the second order Taylor expansion of $G(S_{t+1}/n)$ around $E[S_{t+1}/n | S_t = \zeta] \approx F(\zeta/n)$ we obtain

$$E[G(S_{t+1}/n) | S_t = \zeta] \approx G(F(\zeta/n)) + \frac{1}{2} \text{Var}[S_{t+1}/n | S_t = \zeta] G''(F(\zeta/n)). \quad (103)$$

(The precise conditions on the derivatives of G such that the approximation in (103) is sufficiently accurate are given in Lemma 55.) From (103), in order to satisfy (102), the function G has to be carefully chosen to control the interplay between $G(F(x)) - G(x)$ and $G''(F(x))$. The first derivative of G should correspond to the drift $F(x) - x$ of the process coming from its expectation while the second derivative of G to the variance of the process. More precisely, when x is outside the critical window, the choice of the potential function is such that $G(F(x)) - G(x)$ is bounded above by a negative constant (i.e., its derivative is $1/(x - F(x))$); by our earlier remarks this should be sufficient to establish progress outside the critical window. Indeed, with this choice it turns out that $|G''(x)|/n$ is bounded above by a small constant outside the critical window, so that (102) is satisfied. Inside the critical window, where $x \approx F(x)$ and hence $G(F(x)) - G(x) \approx 0$, we choose G so that $G''(x)$ is negative. More precisely, to satisfy (102), since $\text{Var}[S_{t+1}/n | S_t = \zeta] = \Theta(1/n)$ from Lemma 47, we set $G''(x) = -Cn$ for some constant $C > 0$. The remaining part is then to interpolate between these two regimes keeping $G'(x)/G''(x)$ sufficiently large (so that (102) is satisfied) and $G(x)$ small (i.e., $O(n^{1/3})$); this is possible due to the quadratic behaviour of $F(z) - z$ around $z = a$. (See Lemma 56 and its proof for the explicit specification of G .)

We next combine Lemmas 45, 46 and 48 to show the following.

Lemma 49. *For $B = \mathfrak{B}_u$, there exists $L > 0$ such that the following is true. In $T = O(n^{1/3})$ steps, for any starting state X_0 , with probability $\Theta(1)$ the SW algorithm ends up in a state X_T such that $\|\alpha(X_T) - \mathbf{u}\|_\infty \leq Ln^{-1/2}$.*

Proof. Let $T := \lceil 3M_1 n^{1/3} / \tau \rceil$, where M_1, τ are the constants in Lemma 48.

Let $\varepsilon > 0$ be a sufficiently small constant, to be picked later. We will assume that the state X_1 has $q - 1$ colors which are ε -light since (by the first part of Lemma 45) this event happens with probability $\Theta(1)$. Henceforth, we will condition on this event.

⁶Heuristically, the exponent $1/3$ in our target mixing time bound $O(n^{1/3})$ is the value $\rho \geq 0$ obtained by balancing (i) the number of steps that the process needs to get out from the interval $(na - n^{1-\rho}, na + n^{1-\rho})$ using its variance which we expect to happen in roughly $n^{1-2\rho}$ steps (since $\sqrt{n^{1-2\rho}n} = n^{1-\rho}$), and (ii) the number of steps that the process needs to cross the intervals $(n(a - \varepsilon), na - n^{1-\rho})$ and $(na + n^{1-\rho}, n(a + \varepsilon))$ using the drift $z - F(z) \approx c(z - a)^2$ (which requires roughly n^ρ steps).

Recall that S_t is the size of the largest color class at time t . We will show that with probability $\Theta(1)$ it holds that $S_T < (1 + \varepsilon)n/B$. Assuming this for the moment, then in the next step, i.e., at time $T+1$, by Lemma 46 all color classes have size at most $(1 + 3q\varepsilon)n/q$ and (for all sufficiently small ε) are thus subcritical in the percolation step of the SW dynamics. It follows that the components' sizes after the percolation step satisfy, by Item 1 in Lemma 17, $E\left[\sum_i |C_i|^2\right] = O(n)$. Hence, after the coloring step, using Azuma's inequality with constant probability we have color classes of size $(n + O(n^{1/2}))/q$ (see for example the derivation of (57) and (58) for details).

It remains to argue that with probability $\Theta(1)$ it holds that $S_T < (1 + \varepsilon)n/B$. Let P_t be the probability that at time t it holds that $S_t < (1 + \varepsilon)n/B$. We will show that $P_T \geq 1/10$. We will use Lemma 48 and the potential function G therein to bound P_T . In particular, we will show that for all n sufficiently large, for all $t = 1, \dots, T$, it holds that

$$E[G(S_{t+1}/n)] \leq E[G(S_t/n)] - \tau(1 - P_t) + \tau/2, \quad (104)$$

where τ is the constant in Lemma 48. Prior to that, let us conclude that $P_T \geq 1/10$ assuming (104). Note that if $S_t < (1 + \varepsilon)n/B$ then $S_{t+1} < (1 + \varepsilon)n/B$ with probability at least $1 - \exp(-n^{\Omega(1)})$ (by Lemma 46), so $P_t \leq P_{t+1} + O(1/n)$. Since $T = O(n^{1/3})$, we have $P_t \leq P_T + O(n^{-2/3})$ for all $t = 1, \dots, T$ and hence $\sum_{t=1}^T P_t \leq TP_T + o(1)$. By applying (104) recursively, it hence follows that

$$E[G(S_{T+1}/n)] \leq E[G(S_1/n)] - \tau T(1/2 - P_T) + o(1).$$

Using that $0 \leq G(z) \leq M_1 n^{1/3}$ for all $z \in [1/q, 1]$, we obtain that $P_T \geq 1/2 - M_1 n^{1/3}/(\tau T) + o(1)$. For $T = \lceil 3M_1 n^{1/3}/\tau \rceil$ we thus have $P_T \geq 1/10$ as wanted.

Finally, we prove (104) for $t = 1, \dots, T$. Note that Lemmas 48 and 46 apply whenever X_t has $q - 1$ ε -light colors, so we will need to account for the (small-probability) event that this fails. Namely, let \mathcal{E}_t denote the event that X_t has $q - 1$ ε -light colors. Since we condition on the event that \mathcal{E}_1 holds, we have that $\bigcap_{t=2}^T \mathcal{E}_t$ holds with probability at least $1 - \exp(-n^{\Omega(1)})$ (by the second part of Lemma 45).

Let \mathcal{F}_t be the event that $S_t < (1 + \varepsilon)n/B$ and note that $P_t = \Pr(\mathcal{F}_t)$. By taking expectations in inequality (102) of Lemma 48, we have

$$E[G(S_{t+1}/n) \mid \mathcal{E}_t, \overline{\mathcal{F}_t}] \leq E[G(S_t/n) \mid \mathcal{E}_t, \overline{\mathcal{F}_t}] - \tau. \quad (105)$$

Note that if $S_t < (1 + \varepsilon)n/B$, then by Lemma 46, with probability $1 - \exp(-n^{\Omega(1)})$ we have $S_{t+1} < (1 + 3q\varepsilon)n/q$. From Lemma 48, we have $G(1/q) = 0$ and $\max_{z \in [1/q, 1/B]} G'(z) \leq M_2$ where M_2 is an absolute constant independent of n . It follows that for all sufficiently small constant $\varepsilon > 0$, when $S_{t+1} < (1 + 3q\varepsilon)n/q$, it holds that $G(S_{t+1}/n) \leq \tau/3$. It follows that

$$E[G(S_{t+1}/n) \mid \mathcal{E}_t, \mathcal{F}_t] \leq \tau/3. \quad (106)$$

Note that G is positive throughout the interval $[1/q, 1]$ since $G(1/q) = 0$ and G is increasing. By the positivity of G , we thus obtain the crude inequality

$$\Pr(\overline{\mathcal{F}_t} \mid \mathcal{E}_t) E[G(S_t/n) \mid \mathcal{E}_t, \overline{\mathcal{F}_t}] \leq E[G(S_t/n) \mid \mathcal{E}_t]. \quad (107)$$

Let P'_t be the probability that at time t it holds that $S_t < (1 + \varepsilon)n/B$ conditioned on the event \mathcal{E}_t , i.e., $P'_t := \Pr(\mathcal{F}_t \mid \mathcal{E}_t)$. Note that $P_t \geq P'_t(1 - \exp(-n^{\Omega(1)})) \geq P'_t - \exp(-n^{\Omega(1)})$. Combining (105), (106) and (107), we obtain

$$E[G(S_{t+1}/n) \mid \mathcal{E}_t] \leq E[G(S_t/n) \mid \mathcal{E}_t] - \tau(1 - P'_t) + \tau/3. \quad (108)$$

Since G is bounded by a polynomial in n and the probability of the event $\overline{\mathcal{E}}_t$ is $\exp(-n^{\Omega(1)})$, removing the conditioning in (108) only affects the inequality by an additive $o(1)$. Similarly, replacing P'_t with P_t in (108) only affects the inequality by an additive $o(1)$. This proves that (104) holds for all sufficiently large n , thus concluding the proof of Lemma 49. \square

Using Lemma 49, it is not hard to obtain the following corollary.

Corollary 50. *Let $B = \mathfrak{B}_u$. The mixing time of the Swendsen-Wang algorithm on the complete graph on n vertices is $O(n^{1/3})$.*

Proof. Consider two copies $(X_t), (Y_t)$ of the SW chain. As in the proof of Corollary 32, it suffices to show that for $T = O(n^{1/3})$, there exists a coupling of (X_t) and (Y_t) such that $\Pr(X_T = Y_T) = \Omega(1)$.

By Lemma 49, for $T_1 = O(n^{1/3})$, it holds that with probability $\Theta(1)$

$$\|\alpha(X_{T_1}) - \mathbf{u}\|_\infty \leq Ln^{-1/2} \text{ and } \|\alpha(Y_{T_1}) - \mathbf{u}\|_\infty \leq Ln^{-1/2}. \quad (109)$$

Conditioning on (109), by Lemma 26, there exists a coupling such that with probability $\Theta(1)$ for $T_2 = T_1 + 1$, it holds that $\alpha(X_{T_2}) = \alpha(Y_{T_2})$. Conditioning on $\alpha(X_{T_2}) = \alpha(Y_{T_2})$, by Lemma 25 there exists $T_3 = O(\log n)$ and a coupling such that $\Pr(X_{T_2+T_3} = Y_{T_2+T_3} \mid \alpha(X_{T_2}) = \alpha(Y_{T_2})) = \Omega(1)$. It is now immediate to combine the couplings to obtain a coupling such that $\Pr(X_T = Y_T) = \Omega(1)$ with $T = T_2 + T_3 = O(n^{1/3})$, as desired. \square

11.3 Lower bound on the mixing time at $B = \mathfrak{B}_u$

In this section, we prove that the mixing time of the SW algorithm at $B = \mathfrak{B}_u$ satisfies $T_{\text{mix}} = \Omega(n^{1/3})$.

As in the proof of the upper bound, the lower bound on the mixing time follows by carefully accounting for the number of steps that the SW algorithm needs to escape the window around the majority phase. In this section, our goal is to show that it takes $\Omega(n^{1/3})$ steps to escape the window. The following lemma provides the “reverse” direction of Lemma 48. Recall that for a state X_t of the SW algorithm, the size of the largest color class is denoted by S_t .

Lemma 51. *Let $B = \mathfrak{B}_u$. There exist constants $M_1, M_2, \rho > 0$ such that for all sufficiently small $\varepsilon > 0$, for all sufficiently large n the following holds. There exists a three-times differentiable increasing function $G : [1/q, 1] \rightarrow [0, M_1 n^{1/3}]$ which satisfies $G(1/B) = O(1)$, $G(1) \geq M_2 n^{1/3}$ such that for any $\zeta \geq n/q$, if X_t has $(q-1)$ colors which are ε -light, then it holds that*

$$E[G(S_{t+1}/n) \mid S_t = \zeta] \geq G(\zeta/n) - \rho. \quad (110)$$

We remark here that the potential function in Lemmas 48 and 51 will be chosen to be identical. We thus refer the reader to the discussion after Lemma 48 for an overview of the construction of G and to Section 11.4 for the actual construction and the proof of Lemma 51.

Analogously to Section 9, we will also need a (crude) bound on the probability mass of configurations which are far from the uniform phase in the Potts distribution. Recall from Section 5 that $\mathcal{B}(\mathbf{v}, \delta)$ is the ℓ_∞ -ball of configuration vectors of the q -state Potts model in K_n around \mathbf{v} of radius δ , cf. equation (34). For a constant $\eta > 0$, let

$$U(\eta) := \mathcal{B}(\mathbf{u}, \eta).$$

The following lemma is analogous to Lemma 38 and its proof hinges on the arguments used to derive the upper bound for the mixing time at $B = \mathfrak{B}_u$. (Similarly to Lemma 38, more precise bounds can be found in, e.g., [10]; the following estimate follows easily from our upper bound on the mixing time.)

Lemma 52. *Let $B = \mathfrak{B}_u$ and $\eta > 0$ be a constant. For all sufficiently large n , the Potts distribution μ (given in (1)) satisfies $\mu(\overline{U(\eta)}) \leq 1/8$.*

Proof. For convenience, denote $U := U(\eta)$. By Lemma 49, for all sufficiently large n and any starting state X_0 , we have that for $T = O(n^{1/3})$, it holds that

$$\Pr(X_T \in U) \geq \varepsilon,$$

where $\varepsilon > 0$ is a constant independent of n . It follows that for all non-negative integers j it also holds that

$$\Pr(X_{(j+1)T} \in U \mid X_{jT} \notin U) \geq \varepsilon.$$

Further, by Lemma 22, for integer $t \geq 0$, it holds that

$$\Pr(X_{t+1} \in U \mid X_t \in U) \geq 1 - \exp(-\Omega(n^{1/3})).$$

We thus obtain that for any starting state X_0 , for some positive integer $j = j(\varepsilon)$, for all sufficiently large n , for all integer $t \geq jT$, it holds that

$$\Pr(X_t \in U) \geq 15/16. \quad (111)$$

Let $T^* = \max\{jT, 2T_{\text{mix}}\}$. By Corollary 50, we have $T_{\text{mix}} = O(n^{1/3})$, so $T^* = O(n^{1/3})$ as well. The same arguments as in the proof of Lemma 38 (cf. equation (75)) yield

$$\mu(\overline{U}) - \Pr(X_{T^*} \in \overline{U}) \leq 1/16. \quad (112)$$

Combining (111) and (112) yields $\mu(\overline{U}) \leq 1/8$, as wanted. \square

The following lemma can be derived from Lemma 51 by suitably adapting the proof of Lemma 49.

Lemma 53. *For $B = \mathfrak{B}_u$, there exists a constant $\eta > 0$ such that the following is true for all n . Suppose that we start at a state X_0 where all the vertices are assigned the color 1. Then, for some $T = \Omega(n^{1/3})$, with probability $\geq 1/2$, it holds that $X_T \notin U(\eta)$.*

Proof. Let M_1, M_2, ρ be the constants in Lemma 51 and let $T := \lceil M_2 n^{1/3} / (6\rho) \rceil$.

Recall that S_t is the size of the largest color class at time t . Let $\varepsilon > 0$ be a sufficiently small constant, to be picked later. We will prove that with probability $\geq 1/2$ it holds that

$$\Pr(S_T > (1 + \varepsilon)n/q) \geq 1/2. \quad (113)$$

Let $\eta := \varepsilon/q$ and note that η is a constant. The lemma then follows by just observing that $\Pr(X_T \notin U(\eta)) \geq \Pr(S_T > (1 + \varepsilon)n/q)$.

We next argue that (113) holds. To do this, we will show that for all n sufficiently large, for all $t = 0, \dots, n$, it holds that

$$E[G(S_{t+1}/n)] \geq E[G(S_t/n)] - 2\rho, \quad (114)$$

where G, ρ are the potential function and the constant from Lemma 51, respectively. Prior to proving (114), let us conclude the argument assuming (114). Lemma 51 asserts that the constants M_1, M_2 are such that

$$0 \leq G(z) \leq G(1) \text{ for all } z \in [1/q, 1], \text{ with } G(1) = Cn^{1/3} \text{ and } C \text{ satisfying } M_2 \leq C \leq M_1. \quad (115)$$

Applying (114) for $t = 0, \dots, T-1$, we obtain that

$$E[G(S_T/n)] \geq G(S_0/n) - 2\rho T.$$

Since $S_0 = n$ and $G(1) = Cn^{1/3}$, it thus follows that $E[G(S_T/n)] \geq (2/3)Cn^{1/3}$. Let $\varepsilon > 0$ be such that $(1 + \varepsilon)/q < 1/B$; such an ε exists since at $B = \mathfrak{B}_u$ it holds that $1/q < 1/B$. From $G(1/B) = O(1)$ and the fact that G is increasing, we obtain that there exists a constant $\xi > 0$ such that $G((1 + \varepsilon)/q) \leq \xi$. It is immediate now to conclude that with probability $\geq 1/2$ it holds that $S_T > (1 + \varepsilon)n/q$; otherwise, using (115), we would have that for sufficiently large n , it holds that $E[G(S_T/n)] \leq (3/5)Cn^{1/3}$, contradicting our lower bound for $E[G(S_T/n)]$.

Finally, we prove (114) for $t = 0, \dots, n$. We will use Lemmas 45 and 51. Let $\varepsilon > 0$ be a small constant as in the statement of Lemma 45. Note that Lemma 51 applies whenever X_t has $q - 1$ ε -light colors, so we will need to account for the (small probability) event that this fails. Namely, let \mathcal{E}_t denote the event that X_t has $q - 1$ ε -light colors. Since the event \mathcal{E}_0 holds (by the choice of the starting state X_0), we have that $\bigcap_{t=0}^n \mathcal{E}_t$ holds with probability at least $1 - \exp(-n^{\Omega(1)})$ (by the second part of Lemma 45).

Let t be an integer between 0 and n . By taking expectations in inequality (110) of Lemma 51, we have

$$E[G(S_{t+1}/n) \mid \mathcal{E}_t] \geq E[G(S_t/n) \mid \mathcal{E}_t] - \rho. \quad (116)$$

Since G is bounded by a polynomial (cf. (115)) and the probability of the event $\overline{\mathcal{E}_t}$ is exponentially small, removing the conditioning in (116) only affects the inequality by an additive $o(1)$. This proves that (114) holds for all sufficiently large n , thus concluding the proof of Lemma 49. \square

Using Lemmas 52 and 53, we obtain the following corollary.

Corollary 54. *Let $B = \mathfrak{B}_u$. The mixing time T_{mix} of the Swendsen-Wang algorithm on the complete graph on n vertices satisfies $T_{\text{mix}} = \Omega(n^{1/3})$.*

Proof. Let η be as in Lemma 53 and let $U := U(\eta)$. Consider the starting state X_0 where all the vertices are assigned the color 1. Then, by Lemma 53, for some $T = \Omega(n^{1/3})$ we have that

$$\Pr(X_T \notin U) \geq 1/2.$$

On the other hand, by Lemma 52 we have that $\mu(\overline{U}) \leq 1/8$. It follows that

$$d_{TV}(X_T, \mu) = \max_{A \subseteq \Omega} |\mu(A) - \Pr(X_T \in A)| \geq \Pr(X_T \in \overline{U}) - \mu(\overline{U}) \geq 1/2 - 1/8 > 1/4.$$

It follows from the definition of mixing time that $T_{\text{mix}} \geq T$, as claimed. \square

11.4 Constructing the potential function - Proof of Lemmas 48 and 51

In this section, we prove Lemmas 48 and 51, i.e., construct the potential function G . We split the argument in several lemmas.

The first lemma achieves two goals: first, it quantifies the bounds that the function G must satisfy so that the approximation

$$E[G(S_{t+1}/n) \mid S_t = \zeta] \approx G(F(\zeta/n)) + \frac{1}{2} \text{Var}[S_{t+1}/n \mid S_t = \zeta] G''(F(\zeta/n)), \quad (103)$$

which we described in Section 11.2 is valid; the bounds are given in (117). Second, it gives an inequality that the function G must satisfy (cf. equation (118)) which allows to deduce, using the approximation (103), the bounds on $E[G(S_{t+1}/n) \mid S_t = \zeta] - G(\zeta/n)$ claimed in Lemmas 48 and 51 (see (119) below).

Lemma 55. Let $\varepsilon > 0$. Suppose that, for all n sufficiently large, S_t and S_{t+1} are random variables that satisfy (87),(88),(89) when $\zeta \geq (1 + \varepsilon)n/B$.

Let G be a three-times differentiable potential function defined on the interval $[1/q, 1]$ such that

$$\min_x G'(x) > 0, \max_x |G'(x)| = O(n^{2/3}), \max_x |G''(x)| = O(n), \sup_x |G'''(x)| = O(n^{4/3}). \quad (117)$$

Further, assume that for each $x > 1/B$, it holds that

$$\begin{aligned} -\tau_2 &< G(F(x)) - G(x) + G''(F(x))Q_1/(2n) < -\tau_1, \\ -\tau_2 &< G(F(x)) - G(x) + G''(F(x))Q_2/(2n) < -\tau_1, \end{aligned} \quad (118)$$

where $\tau_1, \tau_2 > 0$ are constants (independent of n) and Q_1, Q_2 are as in (88).

Then, for any $\zeta \geq (1 + \varepsilon)n/B$, it holds that

$$G(\zeta/n) - 2\tau_2 \leq E[G(S_{t+1}/n) | S_t = \zeta] \leq G(\zeta/n) - \tau_1/2. \quad (119)$$

Recall that for $B = \mathfrak{B}_u$, the function $F(z)$ has exactly one fixpoint in the interval $(1/B, 1]$ at $z = a$. The following lemma specifies a potential function G which will be used to verify the conditions (117) and (118) in Lemma 55. We have already described in Section 11.2, the high-level approach for the construction of G . The actual definition of G is quite technical due to the requirement that G should be three times differentiable. We pulled out the important bits in the construction of G that will also be relevant in verifying (118).

For positive real numbers A, B we will use the notation $A \gg B$ to denote that for some (large) constant $C > 1$, it holds that $A > BC$.

Lemma 56. Let L, L' be positive constants which satisfy $L \gg L'$. There exist positive constants M, C_0, C_1, C_2 such that the following holds.

For all sufficiently large n , there exists a strictly increasing three-times differentiable function $G : [1/q, 1] \rightarrow [0, Mn^{1/3}]$ with $G(1/q) = 0$ which satisfies (117) and

$$\begin{aligned} |G'(z)|, |G''(z)| &\leq C_0 \text{ for } z \in [1/q, 1/B], \\ G'(z) &= \frac{1}{z - F(z)} \text{ for } z \in [1/B, a - Ln^{-1/3}] \cup [a + Ln^{-1/3}, 1], \\ G'(z) &\geq C_1 n^{2/3}, \quad |G''(z)| \leq (10^2 C_1/L)n \text{ for } z \in [a - Ln^{-1/3}, a - L'n^{-1/3}], \\ G''(z) &\leq -C_2 n \text{ for } z \in [a - L'n^{-1/3}, a + Ln^{-1/3}]. \end{aligned} \quad (120)$$

Lemma 57. Let L, L' be positive constants which satisfy $L \gg L' \gg 1$. Then, there exist constants $\tau_1, \tau_2 > 0$, such that, for any function G satisfying (117) and (120), inequality (118) holds for every $x > 1/B$.

The following lemma will be useful throughout the rest of this section.

Lemma 58. Let $B = \mathfrak{B}_u$. Then it holds that

1. $F'(z) = 1$ iff $z = a$.
2. $F''(z) < 0$ for all $z \in (1/B, 1]$.
3. $F(z) \leq z$ for all $z \in [1/B, 1]$ with equality iff $z = a$.

Proof. The proofs for the first two parts are given in Lemmas 13 and 8, respectively. For the third part, note that the function $z - F(z)$ is convex in $[1/B, 1]$ and has a unique critical point at $z = a$. Thus, $z - F(z) \geq a - F(a) = 0$ with equality if $z = a$. \square

We are now ready to prove Lemmas 48 and 51 (assuming Lemmas 55, 56 and 57).

Proof of Lemma 48. Let L, L' be positive constants satisfying $L \gg L' \gg 1$. By Lemmas 56 and 57, there exist constants $M, \tau_1, \tau_2 > 0$ such that for all sufficiently large n there exists a three-times differentiable function $G : [1/q, 1/B] \rightarrow [0, Mn^{1/3}]$ which satisfies both (117) and (118). Note that (117) guarantees that G is increasing. Further, by Lemma 56, it holds that $G(1/q) = 0$ and $\max_{z \in [1/q, 1/B]} G'(z) \leq C_0$ where C_0 is a constant. We will use this function G to prove Lemma 48 with $M_1 = M$, $M_2 = C_0$ and $\tau = \tau_1/2$.

Let $\varepsilon > 0$ be a sufficiently small constant and suppose that X_t has $(q - 1)$ colors which are ε -light. Recall that S_t is the size of the largest color class in X_t . By Lemma 47, we have that for all sufficiently large n , for all $\zeta \geq (1 + \varepsilon)n/B$, the random variables S_t, S_{t+1} satisfy (87), (88), (89). It follows by Lemma 55 that

$$E[G(S_{t+1}/n) | S_t = \zeta] \leq G(\zeta/n) - \tau_1/2.$$

This completes the verification of all the conditions that G must satisfy, concluding the proof of the lemma. \square

Proof of Lemma 51. We begin by specifying some constants. Let $\varepsilon_0 > 0$ be a constant such that $(1 + \varepsilon_0)/B < a$ and let

$$W_0 := \min_{z \in [1/B, (1 + \varepsilon_0)/B]} \{z - F(z)\}. \quad (121)$$

By Lemma 58 and the choice of ε_0 , we have that $W_0 > 0$.

Consider positive constants L, L' satisfying $L \gg L' \gg 1$. By Lemmas 56 and 57, there exist constants $M, \tau_1, \tau_2 > 0$ such that for all sufficiently large n there exists a function $G : [1/q, 1/B] \rightarrow [0, Mn^{1/3}]$ which satisfies all of (117), (118) and (120). Further, by Lemma 56, there exist positive constants C_0, C_1 such that

$$G(1/q) = 0, \quad \max_{z \in [1/q, 1/B]} G'(z) \leq C_0, \quad \min_{z \in [a - Ln^{-1/3}, a - L'n^{-1/3}]} G'(z) \geq C_1 n^{2/3}.$$

Therefore, $G(1/B) = O(1)$ and $G(1) \geq C_3(L - L')n^{1/3}$ (for the latter we also need that G is increasing which is guaranteed from (117)). We will also need a bound on the variation of G on the interval $[1/q, (1 + \varepsilon)/B]$. Using (120) and (121), we have that $\max_{z \in [1/B, (1 + \varepsilon_0)/B]} G'(z) \leq 1/W_0$. It follows that for $\eta_0 := \max\{C_0, 1/W_0\}$, it holds that $G'(z) \leq \eta_0$ for all $z \in [1/q, (1 + \varepsilon_0)/B]$ and thus there exists a constant $\eta > 0$ such that

$$|G(z_1) - G(z_2)| \leq \eta \text{ for all } z_1, z_2 \in [1/q, (1 + \varepsilon_0)/B]. \quad (122)$$

We will use G to prove Lemma 51 with $M_1 = M$, $M_2 = C_3(L - L')$ and $\rho = 2\tau_2 + 2\eta$.

Let $\varepsilon > 0$ be a sufficiently small constant and n be sufficiently large. Suppose that X_t has $(q - 1)$ colors which are ε -light. Recall that S_t is the size of the largest color class in X_t and suppose that $S_t = \zeta$ where $\zeta \geq n/q$. We will split the proof into cases depending on whether $\zeta \geq (1 + \varepsilon)n/B$.

Consider first the case where $\zeta \geq (1 + \varepsilon)n/B$. By Lemma 47, we have that the random variables S_t, S_{t+1} satisfy (87), (88), (89). It follows by Lemma 55 that

$$E[G(S_{t+1}/n) | S_t = \zeta] \geq G(\zeta/n) - 2\tau_2.$$

Consider now the case where $\zeta \leq (1 + \varepsilon)n/B$ so that $S_t \leq (1 + \varepsilon)n/B$. Let \mathcal{E}_t be the event that $S_{t+1} \leq (1 + \varepsilon)n/B$. By Lemma 46, we have that $\Pr(\mathcal{E}_t) = 1 - \exp(-n^{\Omega(1)})$. Also, using (122), we have that

$$E[G(S_{t+1}/n) | S_t = \zeta, \mathcal{E}_t] \geq G(\zeta/n) - \eta. \quad (123)$$

Recall that G is non-negative with values that are polynomially bounded. Since \mathcal{E}_t holds with exponentially large probability, it follows that removing the conditioning on the event \mathcal{E}_t in (123) only affects the inequality by $o(1)$. Hence, for all sufficiently large n , it holds that

$$E[G(S_{t+1}/n) | S_t = \zeta] \geq G(\zeta/n) - 2\eta.$$

This completes the verification of all the conditions that G must satisfy, concluding the proof of the lemma. \square

Proof of Lemma 55. Let $x = \zeta/n$ and $y = E[S_{t+1}/n | S_t = \zeta]$. Let $Z = S_{t+1}/n - y$. Note that Z is a random variable, $E[Z | S_t = \zeta] = 0$, and by Lemma 47,

$$Q_1/n \leq \text{Var}[Z | S_t = \zeta] = E[Z^2 | S_t = \zeta] \leq Q_2/n.$$

By Taylor's expansion, we have

$$G(y + Z) = G(y) + G'(y)Z + G''(y)\frac{Z^2}{2} + G'''(\rho)\frac{Z^3}{6}, \quad (124)$$

for some ρ which lies between y and $y + Z$ (note that ρ is also a random variable).

From inequality (89) of Lemma 47 we have for all sufficiently small $\varepsilon' > 0$

$$E[|Z|^3 | S_t = \zeta] \leq Kn^{-3/2+\varepsilon'}.$$

Taking expectations of (124) we obtain

$$E[G(S_{t+1}/n) | S_t = \zeta] = E[G(y + Z) | S_t = \zeta] = G(y) + G''(y)\frac{E[Z^2 | S_t = \zeta]}{2} + C, \quad (125)$$

where $|C| \leq Kn^{-3/2+\varepsilon'} \sup_x |G'''(x)| = o(1)$ since $\sup_x |G'''(x)| = O(n^{4/3})$. Using (87) of Lemma 47 (for $\varepsilon' = 1/10$), we have

$$|G(y) - G(F(x))| \leq \frac{n^{1/10}}{n} \sup_x |G'(x)| \quad \text{and} \quad |G''(y) - G''(F(x))| \leq \frac{n^{1/10}}{n} \sup_x |G'''(x)|.$$

Plugging these estimates in (125) we obtain

$$\left| E[G(S_{t+1}/n) | S_t = \zeta] - \left(G(F(x)) + G''(F(x))\frac{E[Z^2 | S_t = \zeta]}{2} \right) \right| \leq R, \quad (126)$$

where R is an error term satisfying

$$|R| \leq \frac{n^{1/10}}{n} \sup_x |G'(x)| + \frac{n^{1/10}}{n} \sup_x |G'''(x)|\frac{E[Z^2 | S_t = \zeta]}{2} + C.$$

From $\sup_x |G'(x)| = O(n^{2/3})$, $\sup_x |G'''(x)| = O(n^{4/3})$ and $E[Z^2 | S_t = \zeta] \leq Q_2/n$, we obtain that $|R| = o(1)$.

It thus follows from (126) that

$$E[G(S_{t+1}/n) | S_t = \zeta] - G(\zeta/n) = G(F(x)) - G(x) + G''(F(x)) \frac{E[Z^2 | S_t = \zeta]}{2} + o(1). \quad (127)$$

We also have that

$$\begin{aligned} G(F(x)) - G(x) + G''(F(x)) \frac{E[Z^2 | S_t = \zeta]}{2} &\leq G(F(x)) - G(x) + \frac{\max\{Q_1 G''(F(x)), Q_2 G''(F(x))\}}{2n} \\ &\leq -\tau_1 \end{aligned} \quad (128)$$

where in the first inequality we used that $Q_1/n \leq E[Z^2 | S_t = \zeta] \leq Q_2/n$ (note that both estimates are needed since we do not know the sign of G'') and in the second inequality we used (118). Analogously, one has

$$\begin{aligned} G(F(x)) - G(x) + G''(F(x)) \frac{E[Z^2 | S_t = \zeta]}{2} &\geq G(F(x)) - G(x) + \frac{\min\{Q_1 G''(F(x)), Q_2 G''(F(x))\}}{2n} \\ &\geq -\tau_2 \end{aligned} \quad (129)$$

Combining (127), (128) and (129), it follows that for all sufficiently large n it holds that

$$-2\tau_2 \leq -\tau_2 + o(1) \leq E[G(S_{t+1}/n) | S_t = \zeta] - G(\zeta/n) \leq -\tau_1 + o(1) \leq -\tau_1/2$$

This proves that (119) holds, as wanted. \square

We next prove Lemmas 56 and 57. It is more instructive to use Lemma 56 as a black box for now and prove Lemma 57 first.

Proof of Lemma 57. Let $L \gg L' \gg 1$ be constants and n be large. Let

$$z_- := a - Ln^{-1/3}, \quad z'_- := a - L'n^{-1/3}, \quad z_+ := a + Ln^{-1/3},$$

and consider the intervals

$$I_0 = [1/q, 1/B], \quad I_1 = [1/B, z_-], \quad I_2 = [z_-, z'_-], \quad I_3 = [z'_-, z_+], \quad I_4 = [z_+, 1].$$

Let G be a function defined on the interval $[1/q, 1]$ that satisfies (117) and (120), i.e.,

$$\min_x G'(x) > 0, \quad \max_x |G'(x)| = O(n^{2/3}), \quad \max_x |G''(x)| = O(n), \quad \sup_x |G'''(x)| = O(n^{4/3}), \quad (117)$$

and

$$\begin{aligned} |G'(z)|, |G''(z)| &\leq C_0 \text{ for } z \in I_0, \\ G'(z) &= \frac{1}{z - F(z)} \text{ for } z \in I_1 \cup I_4, \\ G'(z) &\geq C_1 n^{2/3}, \quad |G''(z)| \leq (10^2 C_1 / L)n \text{ for } z \in I_2, \\ G''(z) &\leq -C_2 n \text{ for } z \in I_3, \end{aligned} \quad (120)$$

where recall that C_0, C_1, C_2 are positive constants.

Our goal is to show that there exist constants $\tau_1, \tau_2 > 0$, so that for all $x \in (1/B, 1]$ it holds that

$$\begin{aligned} -\tau_2 &< G(F(x)) - G(x) + G''(F(x))Q_1/(2n) < -\tau_1, \\ -\tau_2 &< G(F(x)) - G(x) + G''(F(x))Q_2/(2n) < -\tau_1, \end{aligned} \quad (118)$$

where Q_1, Q_2 are positive constants satisfying $Q_2 \leq Q_1$.

We first show the inequality (118) in the easier regime where $x \in (1/B, 1]$ and $x \notin (a - \varepsilon, a + \varepsilon)$, for any arbitrarily small constant $\varepsilon > 0$ when n is sufficiently large. By Lemma 58, for all $x \neq a$ such that $x \in I_1 \cup I_4$ it holds that $F(x) < x$. Let $\rho := F(\alpha + \varepsilon) - \alpha$, so that $\rho \in (0, \varepsilon)$. Set

$$W_1 := \min_{x \in (1/B, 1], x \notin (a - \rho, a + \rho)} \{x - F(x)\}, \quad W_2 := \max_{x \in (1/B, 1], x \notin (a - \rho, a + \rho)} \{x - F(x)\}.$$

Since $x - F(x)$ is continuous, we obtain that $W_1, W_2 > 0$. By taking n sufficiently large, we obtain that any $x \in (1/B, 1]$ such that $x \notin (a - \rho, a + \rho)$ belongs to $I_1 \cup I_4$ and therefore, from (120), $G'(x)$ is upper and lower bounded by the absolute constants $1/W_1$ and $1/W_2$ for all $x \notin (a - \rho, a + \rho)$. Hence, there exist constants $W'_1, W'_2 > 0$ such that for all $x \notin (a - \varepsilon, a + \varepsilon)$, it holds that

$$-W'_2 \leq G(F(x)) - G(x) \leq -W'_1.$$

A similar argument shows that $|G''(x)|$ is bounded by a constant for all $x \in [1/q, 1]$ with $x \notin (a - \rho, a + \rho)$. It follows that for all $x \in (1/B, 1]$ and $x \notin (a - \varepsilon, a + \varepsilon)$ it holds that

$$\begin{aligned} \max_{i \in \{1, 2\}} G(F(x)) - G(x) + G''(F(x))Q_i/(2n) &\leq -W'_1 + o(1), \\ \min_{i \in \{1, 2\}} G(F(x)) - G(x) + G''(F(x))Q_i/(2n) &\geq -W'_2 + o(1). \end{aligned} \tag{130}$$

This proves (118) when $x \notin (a - \varepsilon, a + \varepsilon)$.

We next prove (118) when $x \in (a - \varepsilon, a + \varepsilon)$ for some appropriate constant $\varepsilon > 0$ to be specified next. Let $c = -F''(a)/2$ and note that $c > 0$ by Lemma 58. Using again Lemma 58 and Taylor's Theorem, there exists $\varepsilon'' > 0$ such that for all $z \in (a - \varepsilon'', a + \varepsilon'')$, it holds that

$$F(z) = z - c(z - a)^2 + O((z - a)^3).$$

Hence, there exists $\varepsilon' > 0$ so that for all $z \in (a - \varepsilon', a + \varepsilon')$ it holds that

$$\begin{aligned} \frac{1}{2}c(z - a)^2 &\leq z - F(z) \leq 2c(z - a)^2, \\ c|z - a| &\leq |F'(z) - 1| \leq 4c|z - a|. \end{aligned} \tag{131}$$

Let $\varepsilon > 0$ be a small constant such that $\varepsilon + 2c\varepsilon^2 < \varepsilon'$ and $4c\varepsilon, 4c\varepsilon^2 < 1/8$. With this choice of ε , we will be able to use the expansion of $F(z)$ around $z = a$. Before we proceed, we give a few intermediate inequalities that will be later used to establish the desired inequalities in (118).

For $x \in (a - \varepsilon, a + \varepsilon)$, we will use the parametrization $x = a + Kn^{-1/3}$ so that $|K| \leq \varepsilon n^{1/3}$. From (131), we have that

$$\frac{1}{2}cK^2n^{-2/3} \leq x - F(x) \leq 2cK^2n^{-2/3}. \tag{132}$$

By the Mean Value Theorem, we also have that there exists $\xi \in (F(x), x)$ such that

$$G(F(x)) - G(x) = G'(\xi)(F(x) - x). \tag{133}$$

Since $\xi \in (F(x), x)$, we have by (132) that $\xi = x - \kappa cK^2n^{-2/3}$ for some $1/2 \leq \kappa \leq 2$. By the choice of ε , it follows that $\xi \in (a - \varepsilon', a + \varepsilon')$ and hence, using (131), we obtain $\frac{1}{2}c(\xi - a)^2 \leq \xi - F(\xi) \leq 2c(\xi - a)^2$. Note that $\xi = a + Kn^{-1/3} - \kappa cK^2n^{-2/3}$, so using that $4c\varepsilon, 4c\varepsilon^2 < 1/8$, we obtain

$$\frac{1}{4}cK^2n^{-2/3} \leq \xi - F(\xi) \leq 4cK^2n^{-2/3}. \tag{134}$$

Finally, for the lower bounds we will use sometimes the following immediate consequences of (117): there exist constants $C'_1, C'_2 > 0$ such that for all $x \in [1/q, 1]$ it holds that

$$0 \leq G'(x) \leq C'_1 n^{2/3}, \quad |G''(x)| \leq C'_2 n. \quad (135)$$

We are now ready to give the proof of (118) for $x \in (a - \varepsilon, a + \varepsilon)$. The proof splits into cases depending on the value of K in the parametrization $x = a + Kn^{-1/3}$.

Case I. $K \leq -L$ or $K \geq L$. We will do the case $K \leq -L$, the proof for $K \geq L$ is analogous. For $K \leq -L$, we have that $x \in I_1$. From (132), we also have $F(x) \in I_1$. In fact, our choice of ε guarantees that $F(x) \in (a - \varepsilon', a + \varepsilon')$, where recall that ε' is as in (131).

For $z \in (a - \varepsilon', a + \varepsilon')$, we have $G'(z) = 1/(z - F(z))$ and thus $G''(z) = \frac{F'(z)-1}{(z-F(z))^2}$. From (131), we thus obtain that $|G''(z)| \leq \frac{16}{c|z-a|^3}$. Applying this for $z = F(x)$ and observing that $F(x) - a \leq -Ln^{-1/3}$, we obtain

$$|G''(F(x))| \leq \frac{16n}{cL^3},$$

so that

$$\max_{i \in \{1,2\}} |G''(F(x))Q_i/(2n)| \leq \frac{8Q_1}{cL^3}.$$

Let ξ be as in (133). Since $\xi \in (F(x), x)$, we have from (120) that $G'(\xi) = 1/(\xi - F(\xi))$. From (134), we have $1/(4cK^2n^{-2/3}) \leq G'(\xi) \leq 4/(cK^2n^{-2/3})$. It follows from (132) and (133) that

$$-8 \leq G(F(x)) - G(x) \leq -1/8.$$

Combining the above estimates, we can conclude that

$$\begin{aligned} \max_{i \in \{1,2\}} G(F(x)) - G(x) + G''(F(x))Q_i/(2n) &\leq -\frac{1}{8} + \frac{8Q_1}{cL^3}, \\ \min_{i \in \{1,2\}} G(F(x)) - G(x) + G''(F(x))Q_i/(2n) &\geq -8 - \frac{8Q_1}{cL^3}. \end{aligned} \quad (136)$$

Since we can choose L to be an arbitrarily large constant, we can make the right-side quantities in (136) to be negative constants, as needed. This completes the proof for Case I.

Case II. $-L \leq K \leq -L'$. In this case, we have $x \in I_2$. It follows from (132) that

$$-\frac{1}{2}c(L')^2n^{-2/3} \geq F(x) - x \geq -2cL^2n^{-2/3}. \quad (137)$$

Since $x \in I_2$, from (137), for all sufficiently large n , we clearly have that either $F(x) \in I_2$ or $F(x) \in I_1$.

Suppose first that $F(x) \in I_2$. From (120) we have that $|G''(F(x))| \leq (10^2C_1/L)n$, so that

$$\max_{i \in \{1,2\}} |G''(F(x))Q_i/n| \leq 10^2C_1Q_1/L.$$

Since $F(x) \in I_2$ and $\xi \in (F(x), x)$, we have $\xi \in I_2$ as well, so from (120) we have $G'(\xi) \geq C_1n^{2/3}$. We also have from (135) that $G'(\xi) \leq C'_1n^{2/3}$. Thus, together with (133) and (137), we obtain

$$-2cL^2C'_1 \leq G(F(x)) - G(x) \leq -\frac{1}{2}c(L')^2C_1. \quad (138)$$

It follows that

$$\begin{aligned} \max_{i \in \{1,2\}} G(F(x)) - G(x) + G''(F(x))Q_i/(2n) &\leq -C_1 \left(\frac{1}{2}c(L')^2 - 10^2 Q_1/L \right), \\ \min_{i \in \{1,2\}} G(F(x)) - G(x) + G''(F(x))Q_i/(2n) &\geq - \left(2cC'_1 L^2 + 10^2 C_1 Q_1/L \right). \end{aligned} \quad (139)$$

Suppose next that $F(x) \in I_1$ so that $a - Ln^{-1/3} \geq F(x)$. From the lower bound in (137), we obtain $F(x) \geq a - Ln^{-1/3} - 2cL^2n^{-2/3}$. Using that $\sup_x |G'''(x)| = O(n^{4/3})$ from (117), it thus follows that $||G''(F(x))| - |G''(a - Ln^{-1/3})|| = O(n^{2/3})$. Moreover, using that $\max_x |G''(x)| = O(n)$ from (117) and $\xi \geq F(x)$, we see that $||G'(\xi)| - |G'(a - Ln^{-1/3})|| = O(n^{1/3})$. Combining these estimates yields again (139) (up to a $o(1)$ term which can be ignored for large n).

Since we can choose L to be an arbitrarily large constant, we can make the right-side quantities in (139) to be negative constants, as needed. This completes the proof for Case II.

Case III. $-L' \leq K \leq L$. In this case, we have $x \in I_3$. Observe that (137) holds in this case as well, so for all sufficiently large n , we have that either $F(x) \in I_2$ or $F(x) \in I_3$.

If $F(x) \in I_3$ then from (120) and (135), we have $-C'_2 n \leq G''(F(x)) \leq -C_2 n$. Since G is increasing and $F(x) \leq x$, we trivially have $G(F(x)) - G(x) \leq 0$. The lower bound on $G(F(x)) - G(x)$ from (138) is valid in this case as well (since both (135) and (137) hold), so we obtain

$$-2cL^2C'_1 \leq G(F(x)) - G(x) \leq 0.$$

It follows that

$$\begin{aligned} \max_{i \in \{1,2\}} G(F(x)) - G(x) + G''(F(x))Q_i/(2n) &\leq -C_2 Q_2, \\ \min_{i \in \{1,2\}} G(F(x)) - G(x) + G''(F(x))Q_i/(2n) &\geq -2cL^2C'_1 - C'_2 Q_1. \end{aligned} \quad (140)$$

If $F(x) \in I_2$ then $a - L'n^{-1/3} \geq F(x)$. From (137) we obtain $F(x) \geq a - L'n^{-1/3} - 2cL^2n^{-2/3}$. It follows that $|G(F(x)) - G(a - L'n^{-1/3})| = o(1)$ and $||G''(F(x))| - |G''(a - L'n^{-1/3})|| = O(n^{2/3})$, yielding again (140) (up to a $o(1)$ term which can be ignored for large n).

The right-side quantities in (140) are negative constants, as needed. This completes the proof for Case III.

We have shown that (118) holds for all $x \in (1/B, 1]$, thus finishing the proof of Lemma 57. \square

We conclude by giving the proof of Lemma 56.

Proof of Lemma 56. Let L, L' be positive constants satisfying $L \gg L'$. To keep better track of the various subintervals involved in the construction of the potential function G , define L_-, L_+, L_m by setting $L_- = L_+ = L$ and $L_m = L'$ and note that $L_+, L_- \gg L_m$. Further, set

$$z_- := a - L_- n^{-1/3}, \quad z_m := a - L_m n^{-1/3}, \quad z_+ := a + L_+ n^{-1/3}.$$

The function G will be more complicated to construct in an interval around a . To help the reader keep track of the notation, we note that z_- refers to the left-most point of the interval around a that will be interesting, while z_+, z_m to the right-most and “middle” points of the interval, respectively.

We will define piecewise the function $G(z)$ in the intervals

$$I_0 = [1/q, 1/B], \quad I_1 = [1/B, z_-], \quad I_2 = [z_-, z_m], \quad I_3 = [z_m, z_+], \quad I_4 = [z_+, 1].$$

Specifically, for $j \in \{0, 1, 2, 3, 4\}$, let $G_j(z)$ be a strictly increasing three-times differentiable function defined on the interval I_j which satisfies

$$\sup_{z \in I_j} |G_j(z)| = O(n^{1/3}), \quad \sup_{z \in I_j} |G'_j(z)| = O(n^{2/3}), \quad \sup_{z \in I_j} |G''_j(z)| = O(n), \quad \sup_{z \in I_j} |G'''_j(z)| = O(n^{4/3}). \quad (141)$$

For $z \in I_j$, we will set $G(z) = G_j(z) + w_j$, where the w_j 's are such that $G(1/q) = 0$ and G is well-defined on the interval (for example, $w_0 = -G_0(1/q)$, $w_1 = -G_1(1/B) + G_0(1/B) - G_0(1/q)$ and so on). Note, from (141), the w_j 's satisfy $|w_j| = O(n^{1/3})$.

The construction of G so far ensures that $G(1/q) = 0$, G is continuous and strictly increasing in the interval $[1/q, 1]$. From (141) and the fact that the w_j 's satisfy $|w_j| = O(n^{1/3})$, we also obtain that there exists a constant $M > 0$ such that $G(z) \leq Mn^{1/3}$ for all $z \in [1/q, 1]$.

The main part of the argument is to specify strictly increasing functions G_j so that:

- i. The properties (120) and (141) hold.
- ii. G is three-times differentiable.

Provided that these conditions are met, we obtain that the function G also satisfies (117) (which completes the proof of the lemma). The roadmap of the construction is as follows:

1. We first specify the functions G_1, G_4 . In particular, we will have

$$G'_1(z) = 1/(z - F(z)) \text{ for } z \in I_1, \quad G'_4(z) = 1/(z - F(z)) \text{ for } z \in I_4. \quad (142)$$

G_1, G_4 are strictly increasing three-differentiable functions which also satisfy (141).

2. The derivatives of G_0 at $z = 1/B$ need to match the derivatives of G_1 at $z = 1/B$, i.e.,

$$G'_0(1/B) = G'_1(1/B), \quad G''_0(1/B) = G''_1(1/B), \quad G'''_0(1/B) = G'''_1(1/B). \quad (143)$$

We will see that $G'_1(1/B), G''_1(1/B), G'''_1(1/B)$ are constants that do not depend n . Thus, G_0 can be chosen to be a function that does not depend on n whatsoever; any strictly increasing three-times differentiable function which satisfies (143) will do. This yields that G_0 in fact satisfies the following bounds (which are stronger than those given in (141)):

$$\max_{z \in I_0} |G_0(z)| = O(1), \quad \max_{z \in I_0} |G'_0(z)| = O(1), \quad \max_{z \in I_0} |G''_0(z)| = O(1), \quad \max_{z \in I_0} |G'''_0(z)| = O(1). \quad (144)$$

3. The function $G_3(z)$ will be chosen to be quadratic. The requirement (146) will thus completely specify G_3 (up to an additive constant). We will see that $G''_4(z_+)$ is negative, so the function G_3 will be concave. Our goal here is to ensure that for constants $C_2, C_3 > 0$ it holds that

$$G''_3(z) \leq -C_2 n \text{ for } z \in I_3, \quad (145)$$

$$G'_3(z_+) = G'_4(z_+), \quad G''_3(z_+) = G''_4(z_+). \quad (146)$$

Note that G_3 is strictly increasing (since $G'_3(z_+) = G'_4(z_+) > 0$ and G'_3 is decreasing from (145)) and three-times differentiable (since G_3 is quadratic). G_3 will also satisfy (141).

4. The function G_2 will satisfy the following constraints (in addition to (148)):

$$G'_2(z) \geq C_1 n^{2/3}, \quad |G''_2(z)| \leq (10^2 C_1 / L) n \text{ for } z \in I_2, \quad (147)$$

$$G'_2(z_-) = G'_1(z_-), \quad G''_2(z_-) = G''_1(z_-), \quad (148)$$

$$G'_2(z_m) = G'_3(z_m), \quad G''_2(z_m) = G''_3(z_m), \quad (149)$$

where C_1 is a positive constant. Note that G_2 is clearly strictly increasing (from (147)). G_2 will also be three-times differentiable and it will satisfy (141).

We will see that $G'_1(z_-) < G'_3(z_m)$ and $G''_1(z_-) > 0 > G''_3(z_m)$. Recall that we also need that the first derivative of G_2 is positive. Thus, the first derivative G'_2 will increase overall in the interval I_2 , yet at the same time G'_2 should change monotonicity at some point inside the interval.

Let us assume for now that the functions G_j satisfy all of the Items 1—4 and conclude that the function G satisfies Conditions i and ii. For Condition i, first observe that (141) is satisfied for all $j \in \{0, 1, 2, 3, 4\}$ by Items 1—4. Also, equations (142), (144), (145) and (147) show that G satisfies (120). This proves that G satisfies Condition i. Relative to Condition ii, using (143), (146), (148), (149) and the three-times differentiability of the G_j 's, we have that G is two-times continuously differentiable with a third derivative which exists everywhere apart (possibly) from the points $z = z_-, z_m, z_+$. For each of these points, we interpolate G''' in an (infinitesimally) small neighborhood of the point using a steep linear function; the use of the linear function guarantees that the order of G''' is still $O(n^{4/3})$. The infinitesimally small length of the interpolation interval guarantees that the effect on G, G', G'' by this modification of G''' can safely be ignored. It follows that G satisfies Conditions i and ii, as wanted.

It remains to obtain Items 1—4. We start with Item 1.

To specify the functions G_1 and G_4 , first consider a function h on the interval $I_1 \cup I_4$ which satisfies $h(1/B) = h(z_+) = 0$ and $h'(z) = 1/(z - F(z))$ for $z \in I_1 \cup I_4$. This well-defines h on $I_1 \cup I_4$. We then set $G_1(z) = h(z)$ for $z \in I_1$ and $G_4(z) = h(z)$ for $z \in I_4$. For $z \in I_1 \cup I_4$, note that $z > F(z)$ (using that $z \neq a$ and Lemma 58) and thus $h'(z) > 0$, so G_1 and G_4 are strictly increasing. It remains to show (141) for $j = 1, 4$.

Note that

$$h''(z) = \frac{F'(z) - 1}{(z - F(z))^2}, \quad h'''(z) = \frac{2(F'(z) - 1)^2 + F''(z)(z - F(z))}{(z - F(z))^3}. \quad (150)$$

Let $c := -F''(a)/2$. By Lemma 58, we have that $c > 0$. By Taylor's theorem, we have that for all sufficiently small $\varepsilon > 0$, for all z in the interval $I := (a - \varepsilon, a + \varepsilon)$, it holds that

$$F(z) = z - c(z - a)^2 + R_3(z) \quad (151)$$

for a remainder function $R_3(z)$ which satisfies $\max_{z \in I} |R_3(z)| = O(|z - a|^3)$. From (151), it also follows that

$$F'(z) = 1 - 2c(z - a) + R_2(z), \quad F''(z) = -2c + R_1(z),$$

for remainder functions $R_1(z), R_2(z)$ which satisfy $\max_{z \in I} |R_1(z)| = O(|z - a|)$ and $\max_{z \in I} |R_2(z)| = O(|z - a|^2)$. We thus obtain that there exist constants $U_1, U_2, U_3 > 0$ such that for $z \in I \setminus \{a\}$, it holds that

$$\begin{aligned} \left| \frac{1}{z - F(z)} - \frac{1}{c}(z - a)^{-2} \right| &\leq U_1 |z - a|^{-1}, \quad \left| \frac{F'(z) - 1}{(z - F(z))^2} + \frac{2}{c}(z - a)^{-3} \right| \leq U_2 |z - a|^{-2}, \\ \left| \frac{2(F'(z) - 1)^2 + F''(z)(z - F(z))}{(z - F(z))^3} - \frac{6}{c}(z - a)^{-4} \right| &\leq U_3 |z - a|^{-3}. \end{aligned} \quad (152)$$

Using (150) and (152), it is immediate to show that $\max_{z \in I_1 \cup I_4} |h'(z)| = O(n^{2/3})$, $\max_{z \in I_1 \cup I_4} |h''(z)| = O(n)$, $\max_{z \in I_1 \cup I_4} |h'''(z)| = O(n^{4/3})$ and thus these bounds carry over to G_1, G_4 as well. We next show that $\max_{z \in I_1} h(z) = O(n^{1/3})$, the proof for $\max_{z \in I_4} h(z) = O(n^{1/3})$ being completely analogous.

In the interval $z \in [1/B, a - \varepsilon]$, we have that h' is bounded above by an absolute constant throughout the interval, so we clearly have that $h(a - \varepsilon) - h(1/B) = O(1)$. Consider next $z \in (a - \varepsilon, z_-)$, and parameterize z as $z = a - Kn^{-1/3}$ for some K which satisfies $L_- < K < \varepsilon n^{1/3}$. Using (152), we have the bound

$$h'(z) \leq \frac{1 + \varepsilon U_1}{cK^2} n^{2/3}.$$

Thus

$$\begin{aligned} h(z_-) - h(a - \varepsilon) &= \int_{a-\varepsilon}^{z_-} h'(z) dz = n^{-1/3} \int_{L_-}^{\varepsilon n^{1/3}} h'(a - Kn^{-1/3}) dK \\ &\leq (1 + \varepsilon U_1) n^{1/3} \int_{L_-}^{\varepsilon n^{1/3}} \frac{1}{cK^2} dK \leq Mn^{1/3}, \end{aligned}$$

for some absolute constant M . This concludes the construction for Item 1.

For Item 2, we only need to show that $G'_1(1/B), G''_1(1/B), G'''_1(1/B)$ are constants. This is clear for $G'_1(1/B)$ which is equal to $h'(1/B) = 1/(1/B - 1/q)$; for $G''_1(1/B)$ and $G'''_1(1/B)$, it follows from the expressions in (150) (note, using the method in Lemma 10, one can show that the right derivative of F at $1/B$ is equal to $2(q-1)/q$, while the right second derivative of F at $1/B$ is equal to $-\frac{4B(q-1)}{3q}$). This yields Item 2.

For Items 3 and 4, we will need the values of the derivatives of G_1 and G_4 at the points z_- and z_+ , respectively. Set

$$d'_- := G'_1(z_-), \quad d''_- := G''_1(z_-), \quad d'_+ := G'_4(z_+), \quad d''_+ := G''_4(z_+).$$

From the first two inequalities in (152), we obtain

$$\lim_{n \rightarrow \infty} \frac{d'_\pm}{n^{2/3}} = \frac{1}{cL_\pm^2}, \quad \lim_{n \rightarrow \infty} \frac{d''_-}{n} = \frac{2}{cL_-^3}, \quad \lim_{n \rightarrow \infty} \frac{d''_+}{n} = -\frac{2}{cL_+^3}. \quad (153)$$

From (153), we obtain that for all sufficiently large n , there exist $D'_\pm, D''_\pm > 0$ such that

$$d'_\pm = D'_\pm n^{2/3}, \quad d''_- = D''_- n, \quad d''_+ = -D''_+ n,$$

and

$$\frac{1}{cL_\pm^2} (1 - 10^{-5}) \leq D'_\pm \leq (1 + 10^{-5}) \frac{1}{cL_\pm^2}, \quad \frac{2}{cL_\pm^3} (1 - 10^{-5}) \leq D''_\pm \leq (1 + 10^{-5}) \frac{2}{cL_\pm^3}. \quad (154)$$

Note that D'_\pm, D''_\pm depend on n , but as (154) shows they satisfy $D'_\pm, D''_\pm = \Theta(1)$.

We are now ready to show Item 3. For $z \in I_3$, we will set $G_3(z) = u_1 n^{2/3}(z - a) + u_2 n(z - a)^2$ for u_1, u_2 which we next specify. To satisfy (146), we will choose

$$2u_2 = -D''_+, \quad u_1 + 2u_2 L_+ = D'_+. \quad (155)$$

Observe that $u_2 < 0$, so G_3 is not only a quadratic function but also concave. Note that u_1, u_2 satisfy $|u_1|, |u_2| = \Theta(1)$ from where it easily follows that (141) is satisfied (for $j = 3$). For (145), just observe that $G'''_3(z) = 2u_2 = -D''_+$ and hence the bound on G'''_3 follows from (154). This completes the construction for Item 3.

For the construction in Item 4, we will need a handle of the derivatives of G_3 at the endpoint z_m of the interval I_3 (we will also use these later in the construction for Item 4). Let

$$D'_m := G'_3(z_m)/n^{2/3} \text{ and } D''_m := -G''_3(z_m)/n.$$

We will show that

$$D'_m = (1 \pm 10^{-4}) \frac{3}{cL_{\pm}^2}, \quad D''_m = (1 \pm 10^{-4}) \frac{2}{cL_{\pm}^3}. \quad (156)$$

By the definition of D'_m , we have that $D'_m = u_1 - 2u_2L_m$ and hence, by the choice (155) of u_1, u_2 , we have

$$D'_m = D'_+ + D''_+(L_+ + L_m).$$

Also, we have $D''_m = D''_+$ since the function G_3 is quadratic. It is immediate thus to conclude (156) using the bounds in (154) and $L_{\pm} \gg L_m$.

We are now ready to give the construction for Item 4. To define the function $G_2(z)$ on the interval I_2 , we will set

$$G_2(z) = n^{1/3} g(n^{1/3}(z - a)),$$

where g is a three times differentiable function on the interval $I := [-L_-, -L_m]$ such that

$$g'(-L_-) = D'_-, \quad g''(-L_-) = D''_-, \quad (157)$$

$$g'(-L_m) = D'_m, \quad g''(-L_m) = -D''_m, \quad (158)$$

$$\min_{x \in I} g'(x) \geq \frac{1}{2cL_-^2}, \quad \max_{x \in I} |g''(x)| \leq \frac{25}{cL_+^3}. \quad (159)$$

Equations (157), (158) and (159) ensure that the function G_2 satisfies (147), (148) and (149). Also it will be clear from the specification of g that all of g, g', g'', g''' are bounded by absolute constants, which thus implies that G_2 satisfies (141) (for $j = 2$).

It remains to specify such a function g , we do this by specifying its second derivative. More precisely, for $z \in I$, we will set

$$g'(z) := D'_- + \int_{-L_-}^z h(x) dx, \text{ so that } g''(z) = h(z), \quad (160)$$

where $h(z)$ is a differentiable function on I satisfying

$$h(-L_-) = D''_-, \quad h(-L_m) = -D''_m, \quad \int_{-L_-}^{-L_m} h(x) dx = D'_m - D'_-, \quad (161)$$

$$\max_{x \in I} |h(x)| \leq \frac{25}{cL_+^3}, \quad \int_{-L_-}^z h(x) dx \geq 0 \text{ for all } z \in I. \quad (162)$$

Using (161) and (162), it is immediate to verify that the function g , as specified in (160), satisfies (157), (158) and (159) (for the first inequality in (159), note that $g'(z) \geq D'_-$ for all $z \in I$ and then use the bound for D'_- from (154)).

To specify the function h , we will need two parameters $K_1, K_2 > 0$ such that

$$-L_- < -K_1 < -K_2 < -L_m. \quad (163)$$

We will specify the parameters K_1, K_2 shortly but for now it will be more instructive to assume that K_1, K_2 just satisfy (163); the freedom to specify K_1, K_2 will be helpful at a slightly later point.

So, consider the function h defined on $[-L_-, -L_m]$ by

$$h(z) = \begin{cases} \frac{D''_-}{(L_- - K_1)^2} (z + K_1)^2, & \text{if } -L_- \leq z \leq -K_1 \\ \frac{100(D'_m - D'_-)}{3(K_1 - K_2)^5} (z + K_1)^2 (z + K_2)^2, & \text{if } -K_1 < z < -K_2 \\ \frac{-D''_m}{(K_2 - L_m)^2} (z + K_2)^2, & \text{if } -K_2 \leq z \leq -L_m \end{cases}$$

Note that

$$h(-L_m) = -D''_m, \quad h(-L_-) = D''_-, \quad (164)$$

and that h is differentiable throughout the interval $[-L_-, -L_m]$ since at the points $z = -K_1, -K_2$ it holds that $h(-K_1) = h(-K_2) = h'(-K_1) = h'(-K_2) = 0$. Further, by a direct calculation, the function h satisfies the following:

$$\int_{-L_-}^{-K_1} h(z) dz = \frac{D''_-}{3}(L_- - K_1), \quad \int_{-K_1}^{-K_2} h(z) dz = \frac{10}{9}(D'_m - D'_-), \quad \int_{-K_2}^{-L_m} h(z) dz = -\frac{D''_m}{3}(K_2 - L_m). \quad (165)$$

Since $D'_m > D'_- > 0$ and $D''_-, D''_m > 0$ (cf. (154) and (156)), we also have that

$$\begin{aligned} 0 &\leq h(z) \leq D''_- \text{ for } z \in [-L_-, -K_1], \\ 0 &\leq h(z) \leq \frac{100(D'_m - D'_-)}{48(K_1 - K_2)} \text{ for } z \in (-K_1, -K_2), \\ -D''_m &\leq h(z) \leq 0 \text{ for } z \in [-K_2, -L_m]. \end{aligned} \quad (166)$$

It follows that

$$\max_{z \in I} |h(z)| \leq M, \text{ where } M := \max \left\{ D''_-, D''_m, \frac{3(D'_m - D'_-)}{K_1 - K_2} \right\}. \quad (167)$$

It remains to choose K_1, K_2 satisfying (163) so that the specifications for h in (161) and (162) are satisfied. We set

$$K_1 = L_- - \frac{D'_-}{3D''_-}, \quad K_2 = L_m + \frac{D'_m}{3D''_m}. \quad (168)$$

Using (154) and (156) and $L_{\pm} \gg L_m$, we have

$$K_1 = \left(\frac{5}{6} \pm 10^{-3} \right) L_-, \quad K_2 = \left(\frac{1}{2} \pm 10^{-3} \right) L_+. \quad (169)$$

Since $L_+ = L_- \gg L_m$, we obtain that K_1, K_2 satisfy (163) as desired.

We next check that the specifications for h in (161) and (162) are satisfied. First, combining (165) and (168), we obtain that

$$\int_{-L_-}^{-L_m} h(z) dz = D'_m - D'_-. \quad (170)$$

Equations (164) and (170) show that h does indeed satisfy (161).

We next show that h satisfies the inequalities in (162). To show that $\max_{z \in I} |h(z)| \leq 25/(cL_+^3)$ it suffices to show that $M \leq 25/(cL_+^3)$, where M is as in (167). This is immediate to verify using (154), (156) and (169). For the second inequality in (162), note from (166) that the function $h(z)$ is non-negative when $z < -K_2$ and negative when $z > -K_2$. Thus, it suffices to check the inequality in (162) when $z = -L_-$ and $z = -L_m$. For $z = -L_-$, the inequality holds (trivially) at equality while for $z = -L_m$ the inequality follows from (170) and $D'_m > D'_-$. This completes the construction for Item 4.

We have thus shown how to do the construction of the functions G_0, G_1, G_2, G_3, G_4 so that Items 1—4 hold, completing the proof of Lemma 56. \square

References

- [1] J. Ameskamp and M. Löwe. Moderate deviations for the size of the largest component in a super-critical Erdős-Rényi graph. *Markov Process. Related Fields*, 17(3):369–390, 2011.
- [2] 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(3):511–542, 2012.
- [3] A. Blanca and A. Sinclair. Dynamics for the Mean-field Random-cluster Model. In Naveen Garg, Klaus Jansen, Anup Rao, and José D. P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2015)*, volume 40, pages 528–543, 2015.
- [4] B. Bollobás, G. Grimmett, and S. Janson. The random-cluster model on the complete graph. *Probability Theory and Related Fields*, 104(3):283–317, 1996.
- [5] C. Borgs, J. T. Chayes, J. H. Kim, A. Frieze, P. Tetali, E. Vigoda, and V. H. Vu. Torpid mixing of some Monte Carlo markov chain algorithms in statistical physics. In *Proceedings of the 40th Annual Symposium on Foundations of Computer Science (FOCS '99)*, pages 218–229, 1999.
- [6] C. Borgs, J. T. Chayes, and P. Tetali. Tight bounds for mixing of the Swendsen–Wang algorithm at the Potts transition point. *Probability Theory and Related Fields*, 152(3):509–557, 2012.
- [7] A. Coja-Oghlan, C. Moore, and V. Sanwalani. Counting connected graphs and hypergraphs via the probabilistic method. *Random Structures & Algorithms*, 31(3):288–329, 2007.
- [8] C. Cooper, M. E. Dyer, A. M. Frieze, and R. Rue. Mixing properties of the Swendsen–Wang process on the complete graph and narrow grids. *Journal of Mathematical Physics*, 41(3):1499–1527, 2000.
- [9] C. Cooper and A. M. Frieze. Mixing properties of the Swendsen–Wang process on classes of graphs. *Random Structures & Algorithms*, 15(3-4):242–261, 1999.
- [10] M. Costeniuc, R. S. Ellis, and H. Touchette. Complete analysis of phase transitions and ensemble equivalence for the Curie–Weiss–Potts model. *Journal of Mathematical Physics*, 46(6):063301, 2005.
- [11] P. Cuff, J. Ding, O. Louidor, E. Lubetzky, Y. Peres, and A. Sly. Glauber dynamics for the mean-field Potts model. *Journal of Statistical Physics*, 149(3):432–477, 2012.
- [12] A. Galanis, D. Štefankovič, E. Vigoda, and L. Yang. Ferromagnetic potts model: Refined #bis-hardness and related results. *SIAM Journal on Computing*, 45(6):2004–2065, 2016.
- [13] R. Gheissari and E. Lubetzky. Mixing times of critical 2d potts models. *CoRR*, abs/1607.02182, 2016.
- [14] R. Gheissari, E. Lubetzky, and Y. Peres. Exponentially slow mixing in the mean-field swendsen–wang dynamics. *CoRR*, abs/1702.05797, 2017.
- [15] V. K. Gore and M. R. Jerrum. The Swendsen–Wang process does not always mix rapidly. *Journal of Statistical Physics*, 97(1):67–86, 1999.

- [16] H. Guo and M. Jerrum. Random cluster dynamics for the Ising model is rapidly mixing. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '17)*, pages 1818–1827, 2017.
- [17] S. Janson, T. Luczak, and A. Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, 2000.
- [18] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2008.
- [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 AMS. American Mathematical Society, 2014.
- [20] E. Lubetzky and A. Sly. Critical ising on the square lattice mixes in polynomial time. *Communications in Mathematical Physics*, 313(3):815–836, 2012.
- [21] F. Martinelli and E. Olivieri. Approach to equilibrium of glauher dynamics in the one phase region. i. the attractive case. *Communications in Mathematical Physics*, 161(3):447–486, 1994.
- [22] F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region. ii. the general case. *Communications in Mathematical Physics*, 161(3):487–514, 1994.
- [23] A. Nachmias and Y. Peres. The critical random graph, with martingales. *Israel Journal of Mathematics*, 176(1):29–41, 2010.
- [24] A. D. Scott and G. B. Sorkin. Solving sparse random instances of max cut and max 2-csp in linear expected time. *Combinatorics, Probability and Computing*, 15(1-2):281–315, 2006.
- [25] V. E. Stepanov. On the probability of connectedness of a random graph $\mathcal{G}_m(t)$. *Theory of Probability & Its Applications*, 15(1):55–67, 1970.
- [26] M. Ullrich. Rapid mixing of Swendsen–Wang dynamics in two dimensions. *Dissertationes Mathematicae*, 502:1–64, 2014.
- [27] M. Ullrich. Swendsen–Wang is faster than single-bond dynamics. *SIAM Journal on Discrete Mathematics*, 28(1):37–48, 2014.