SIAM J. DISCRETE MATH.
Vol. 34, No. 1, pp. 742–793

© 2020 Society for Industrial and Applied Mathematics

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

SAMPLING IN UNIQUENESS FROM THE POTTS AND RANDOM-CLUSTER MODELS ON RANDOM REGULAR GRAPHS*

ANTONIO BLANCA†, ANDREAS GALANIS‡, LESLIE ANN GOLDBERG‡, DANIEL ŠTEFANKOVIČ§, ERIC VIGODA†, AND KUAN YANG‡

Abstract. We consider the problem of sampling from the Potts model on random regular graphs. It is conjectured that sampling is possible when the temperature of the model is in the so-called uniqueness regime of the regular tree, but positive algorithmic results have been for the most part elusive. In this paper, for all integers $q \geq 3$ and $\Delta \geq 3$, we develop algorithms that produce samples within error $o(1)$ from the q-state Potts model on random Δ-regular graphs, whenever the temperature is in uniqueness, for both the ferromagnetic and antiferromagnetic cases. The algorithm for the antiferromagnetic Potts model is based on iteratively adding the edges of the graph and resampling a bichromatic class that contains the endpoints of the newly added edge. Key to the algorithm is how to perform the resampling step efficiently since bichromatic classes can potentially induce linear-sized components. To this end, we exploit the tree uniqueness to show that the average growth of bichromatic components is typically small, which allows us to use correlation decay algorithms for the resampling step. While the precise uniqueness threshold on the tree is not known for general values of q and Δ in the antiferromagnetic case, our algorithm works throughout uniqueness regardless of its value. In the case of the ferromagnetic Potts model, we are able to simplify the algorithm significantly by utilizing the random-cluster representation of the model. In particular, we demonstrate that a percolation-type algorithm succeeds in sampling from the random-cluster model with parameters p, q on random Δ-regular graphs for all values of $q \geq 1$ and $p < p_c(q, \Delta)$, where $p_c(q, \Delta)$ corresponds to a uniqueness threshold for the model on the Δ-regular tree. When restricted to integer values of q, this yields a simplified algorithm for the ferromagnetic Potts model on random Δ-regular graphs.

Key words. sampling, Potts model, random regular graphs, phase transitions

AMS subject classifications. 68Q87, 68W20

DOI. 10.1137/18M1219722

1. Introduction.

Random constraint satisfaction problems have been thoroughly studied in computer science in an effort to analyze the limits of satisfiability algorithms and understand the structure of hard instances. Analogously, understanding spin systems on random graphs [27, 28, 34, 4, 26, 6, 13, 8, 9] gives insights about the complexity of counting and the efficiency of approximate sampling algorithms. In this paper, we design approximate sampling algorithms for the Potts model on random regular graphs.

*Received by the editors October 9, 2018; accepted for publication (in revised form) December 19, 2019; published electronically March 23, 2020. A preliminary short version of the manuscript (without the proofs) appeared in the proceedings of RANDOM/APPROX 2018.

https://doi.org/10.1137/18M1219722

Funding: The first and fifth authors' research was supported in part by NSF grants CCF-1617306 and CCF-1563838. The fourth author's research was supported in part by NSF grant CCF-1563757. 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 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.

†School of Computer Science, Georgia Institute of Technology, Atlanta, GA 30332 (ablanca@cc.gatech.edu, vigoda@cc.gatech.edu).

‡University of Oxford, Oxford, OX1 3QD, UK (andreas.galanis@cs.ox.ac.uk, leslie.goldberg@cs.ox.ac.uk, kuan.yang@cs.ox.ac.uk).

§Department of Computer Science, University of Rochester, Rochester, NY 14627 (stefanko@cs.rochester.edu).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The Potts model is a fundamental spin system studied in statistical physics and computer science. The model has two parameters: an integer $q \geq 3$, which represents the number of states/colors of the model, and a real parameter $B > 0$, which corresponds to the so-called temperature. We denote the set of colors by $[q] := \{1, \ldots, q\}$. For a graph $G = (V, E)$, configurations of the model are all possible assignments of colors to the vertices of the graph. Each assignment $\sigma : V \to [q]$ has a weight $w_G(\sigma)$ which is determined by the number $m(\sigma)$ of monochromatic edges under σ, namely, $w_G(\sigma) = B^{m(\sigma)}$. The Gibbs distribution μ_G is defined on the space of all configurations σ and is given by

$$\mu_G(\sigma) = B^{m(\sigma)}/Z_G, \text{ where } Z_G = \sum_{\sigma} B^{m(\sigma)}.$$

We also refer to μ_G as the Potts distribution; the quantity Z_G is known as the partition function. Well-known models closely related to the Potts model are the Ising and colorings models. The Ising model is the special case $q = 2$ of the Potts model, while the q-colorings model is the "zero-temperature" case $B = 0$ of the Potts model, where the distribution is supported on the set of proper q-colorings.

The behavior of the Potts model has significant differences depending on whether B is less or larger than 1. When $B < 1$, configurations where most neighboring vertices have different colors have large weight and the model is called *antiferromagnetic*; in contrast, when $B > 1$, configurations where most neighboring vertices have the same colors have large weight and the model is called *ferromagnetic*. One difference between the two cases that will be relevant later is that the ferromagnetic Potts model admits a random-cluster representation—the details of this representation are given in section 2.1.

Sampling from the Potts model is a problem that is frequently encountered in running simulations in statistical physics or inference tasks in computer science. To determine the efficiency and accuracy of sampling methods, it is relevant to consider the underlying phase transitions, which signify abrupt changes in the properties of the Gibbs distribution when the underlying parameter changes. The so-called uniqueness phase transition captures the sensitivity of the state of a vertex to fixing far-away boundary conditions. As an example, in the case of the ferromagnetic Potts model on the Δ-regular tree, uniqueness holds when root-to-leaves correlations in the Potts distribution vanish as the height of the tree goes to infinity; it is known that this holds iff $B < B_c(q, \Delta)$, where $B_c(q, \Delta)$ is the "uniqueness threshold" (cf. (2.2) for its value). Connecting the uniqueness phase transition with the performance of algorithms is a difficult task that is largely under development. This connection is well-understood on the grid, where it is known that the mixing time of local Markov chains, such as the Glauber dynamics, switches from polynomial to exponential at the corresponding uniqueness threshold; see, for example, [23, 22, 32, 1, 24, 2].

For random Δ-regular graphs or, more generally, graphs with maximum degree Δ, the uniqueness threshold on the Δ-regular tree becomes relevant. For certain two-state models, such as the ferromagnetic Ising model and the hard-core model, it has been proved that Glauber dynamics mixes rapidly when the underlying parameter is in the uniqueness regime of the regular tree, and that the dynamics mixes slowly otherwise (see [27, 28, 9]). The same picture is conjectured to hold for the Potts model as well, but this remains open. For the ferromagnetic case in particular, Bordewich, Greenhill, and Patel [4] prove rapid and slow mixing results for Glauber dynamics on random regular graphs and graphs with maximum degree Δ when the parameter B is within a constant factor from the uniqueness threshold on the regular tree. More generally,

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

there has been significant progress in recent years in understanding the complexity of sampling from the Gibbs distribution in two-state systems, but for multistate systems progress has been slower, especially on the algorithmic side.

In this paper, for all integers $q \geq 3$ and $\Delta \geq 3$, we design approximate sampling algorithms for the q-state Potts model on random Δ-regular graphs (regular graphs with n vertices chosen uniformly at random), when the parameter B lies in the uniqueness regime of the regular tree, for both the ferromagnetic and antiferromagnetic cases. Our algorithms are not based on a Markov chain approach but proceed by iteratively adding the edges of the graph and performing a resampling step at each stage. As such, our algorithms can produce samples that are within error $1/n^{\delta}$ from the Potts distribution for some fixed constant $\delta > 0$ (which depends on B, q, Δ).

Remark 1.1. There are certain "bad" Δ-regular graphs where the algorithms will fail to produce samples with the desired accuracy; saying that the algorithms work on random Δ-regular graphs means that the number of these "bad" graphs with n vertices is a vanishing fraction of all Δ-regular graphs with n vertices for large n. Moreover, we can recognize the "good" graphs (where our algorithms will successfully produce samples with the desired accuracy) in polynomial time.

Our approach is inspired by Efthymiou's algorithm [7, 8] for sampling q-colorings on $G(n, d/n)$; the algorithm there also proceeds by iteratively adding the edges of the graph and exploits the uniqueness on the tree to show that the sampling error is small. However, for the antiferromagnetic Potts model, the resampling step turns out to be significantly more involved and we need a substantial amount of work to ensure that it can be carried out efficiently, as we explain in detail in section 3. Nevertheless, for the ferromagnetic case, we manage to give a far simpler algorithm by utilizing the random-cluster representation of the model (see section 2.1). In particular, we demonstrate that a percolation-type algorithm succeeds in sampling approximately from the random-cluster model with parameters p, q on random Δ-regular graphs for all values of $q \geq 1$ and $p < p_c(q, \Delta)$, where $p_c(q, \Delta)$ corresponds to a uniqueness threshold for the model on the Δ-regular tree. When restricted to integer values of q, this yields a simple algorithm for the ferromagnetic Potts model on random Δ-regular graphs.

To conclude this introductory section, we remark that, for many antiferromagnetic spin systems on random graphs, typical configurations in the Gibbs distribution display absence of long-range correlations even beyond the uniqueness threshold, up to the so-called reconstruction threshold [25, 14]. Note that uniqueness guarantees the absence of long-range correlations under a "worst-case" boundary, while nonreconstruction only asserts the absence of long-range correlations under "typical" boundaries; it is widely open whether this weaker notion is in fact sufficient for sampling on random graphs. On an analogous note, for the ferromagnetic Potts model on random regular graphs, the structure of typical configurations can be fairly well-understood using probabilistic arguments for all temperatures (see, e.g., [6, 13]) and it would be very interesting to exploit this structure for the design of sampling algorithms beyond the uniqueness threshold. In this direction, Jenssen, Keevash, and Perkins [19] very recently designed such an algorithm for all sufficiently large B that works more generally on expander graphs (see also [17] for similar-flavored results on the grid).

2. Definitions and main results. We first review in section 2.1 the definition of the random-cluster model. In section 2.2, we state results from the literature about uniqueness on the regular tree for the Potts and random-cluster models. Then, in

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

section 2.3, we state our algorithmic results for the ferromagnetic Potts and random-cluster models, and, in section 2.4, we state our result for the antiferromagnetic Potts model.

2.1. The random-cluster model. The random-cluster model has two parameters $p \in [0,1]$ and $q > 0$; note that q in this case can take noninteger values. For a graph $G = (V, E)$, we denote the random-cluster distribution on G by φ_G; this distribution is supported on the set of all edge subsets. In particular, for $S \subseteq E$, let $k(S)$ be the number of connected components in the graph $G' = (V, S)$ (isolated vertices do count). Then, the weight of the configuration S is given by $w_G(S) = p^{|S|}(1-p)^{|E\setminus S|}q^{k(S)}$ and

$$\varphi_G(S) = w_G(S)/Z_G^{\mathrm{rc}}, \text{ where } Z_G^{\mathrm{rc}} = \sum_{S\subseteq E} w_G(S) = \sum_{S\subseteq E} p^{|S|}(1-p)^{|E\setminus S|}q^{k(S)}.$$

Following standard terminology, each edge in S will be called open, while each edge in $E\setminus S$ closed.

For integer values of q, there is a well-known connection between the random-cluster and ferromagnetic Potts models, as detailed below.

LEMMA 2.1 (see, e.g., [15]). *Let $q \geq 2$ be an integer, $B > 1$, and $p = 1 - 1/B$. Then, the following hold for any graph $G = (V, E)$:*

- *Let $S \subseteq E$ be distributed according to the random-cluster distribution φ_G with parameters p, q. Consider the configuration σ obtained from S by assigning each component in the graph (V, S) a random color from $[q]$ independently. Then, σ is distributed according to the Potts distribution μ_G with parameter B.*
- *Conversely, suppose that $\sigma : V \to [q]$ is distributed according to the Potts distribution μ_G with parameter B. Consider $S \subseteq E$ obtained by adding to S each monochromatic edge under σ with probability p independently. Then, S is distributed according to the random-cluster distribution φ_G with parameters p, q.*

Lemma 2.1 allows us to translate our sampling algorithm for the random-cluster model to a sampling algorithm for the ferromagnetic Potts model. The benefit of working with the random-cluster model (instead of the Potts model) is that the random-cluster distribution satisfies certain monotonicity properties (cf. Lemma 5.3), which simplifies significantly the analysis of the algorithm.

2.2. Uniqueness for Potts and random-cluster models on the tree. In this section, we review uniqueness on the tree for the Potts and random-cluster models. We start with the Potts model.

For a configuration σ and a set U, we denote by σ_U the restriction of σ to the set U; in the case of a single vertex u, we simply write σ_u to denote the color of u. Denote by $\mathbb{T}_\Delta$ the infinite $(\Delta-1)$-ary tree with root vertex ρ and, for an integer $h \geq 0$, denote by T_h the subtree of $\mathbb{T}_\Delta$ induced by the vertices at distance $\leq h$ from ρ. Let L_h be the set of leaves of T_h.

DEFINITION 2.2. *Let $B > 0$ and $q, \Delta \geq 3$ be integers. The q-state Potts model with parameter $B > 0$ has* uniqueness *on the infinite $(\Delta-1)$-ary tree if, for all colors $c \in [q]$, it holds that*

$$\limsup_{h\to\infty} \max_{\tau : L_h \to [q]} \left| \mu_{T_h}\left(\sigma_\rho = c \mid \sigma_{L_h} = \tau\right) - \frac{1}{q} \right| = 0. \tag{2.1}$$

It has nonuniqueness *otherwise.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

For the ferromagnetic q-state Potts model ($B > 1$), it is known that uniqueness holds on the $(\Delta-1)$-ary tree iff $B < B_c(q,\Delta)$, where

$$B_c(q,\Delta) = 1 + \inf_{y>1} h(y), \text{ where } h(y) := \frac{(y-1)(y^{\Delta-1}+q-1)}{y^{\Delta-1}-y}. \tag{2.2}$$

For the antiferromagnetic Potts model ($B < 1$), the uniqueness threshold on the tree is not yet known in full generality. It is a folklore conjecture that the model has uniqueness iff $q \geq \Delta$ and $B \in (0,1)$, or $q < \Delta$ and $B \geq \frac{\Delta-q}{\Delta}$. It is known that the model has nonuniqueness when $B < \frac{\Delta-q}{\Delta}$ [12]. Establishing the uniqueness side of the conjecture is more difficult; this has been established recently in [11] for small values of q and Δ. In the case $q = 3$, [11] also established the uniqueness threshold for all Δ: for $\Delta \geq 4$, uniqueness holds iff $B \in [(\Delta-3)/\Delta, 1)$, and for $\Delta = 3$, uniqueness holds iff $B \in (0,1)$. For the q-colorings model ($B = 0$), Jonasson [21], building on work of Brightwell and Winkler [5], established that the model has uniqueness iff $q > \Delta$.

Remark 2.3. To summarize the above, a necessary condition for uniqueness on $\mathbb{T}_\Delta$ in the antiferromagnetic q-state Potts model with parameter $B \in (0,1)$ is that $B \geq (\Delta-q)/\Delta$. It is also conjectured that this condition is sufficient but this has only been established for $q = 3$.

Uniqueness for the random-cluster model on the tree is less straightforward to define. Häggström [16] studied uniqueness of random-cluster measures on the infinite $(\Delta-1)$-ary tree where all infinite components are connected "at infinity"—we review his results in more detail in section 7.1. He showed that, for all $q \geq 1$, a sufficient condition for uniqueness is that $p < p_c(q,\Delta)$, where the critical value $p_c(q,\Delta)$ is given by[1]

$$p_c(q,\Delta) = 1 - \frac{1}{1+\inf_{y>1} h(y)}, \text{ where } h(y) := \frac{(y-1)(y^{\Delta-1}+q-1)}{y^{\Delta-1}-y}. \tag{2.3}$$

Note that the critical values in (2.2) and (2.3) are connected for integer values of q via $p_c(q,\Delta) = 1 - \frac{1}{B_c(q,\Delta)}$. Häggström [16] also conjectured that uniqueness for the random-cluster model holds on $\mathbb{T}_\Delta$ when $p > \frac{q}{q+\Delta-2}$ for all $q \geq 1$; this remains open but progress has been made in [20].

Remark 2.4. Using that $h(y) \to \frac{q}{\Delta-2}$ as $y \downarrow 1$, we obtain that $p_c(q,\Delta) \leq \frac{q}{q+\Delta-2}$ for all $q \geq 1$. It can further be shown that $p_c(q,\Delta) = \frac{q}{q+\Delta-2}$ iff $q \leq 2$.

We should note here that the bounds appearing in Remarks 2.3 and 2.4 will be useful to simplify some of our arguments. However, the stronger assumption that the parameters are in the uniqueness region of the Δ-regular tree is crucial for the analysis of our algorithm and, in particular, the proofs of the upcoming Lemmas 5.2 and 6.6 hinge on this assumption.

2.3. Sampling ferro Potts and random-cluster models on random regular graphs. We begin by stating our result for the random-cluster model on random regular graphs.

THEOREM 2.5. *Let $\Delta \geq 3$, $q \geq 1$, and $p < p_c(q,\Delta)$. Then, there exists a constant $\delta > 0$ such that, as $n \to \infty$, the following holds with probability $1 - o(1)$ over the choice of a random Δ-regular graph $G = (V,E)$ with n vertices.*

[1] In [16], $p_c(q,\Delta)$ is defined in a different way, but the two definitions are equivalent for all $q \geq 1$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

There is a polynomial-time algorithm which, on input the graph G, outputs a random set $S \subseteq E$ whose distribution ν_S is within total variation distance $O(1/n^\delta)$ from the random-cluster distribution φ_G with parameters p, q, i.e., $\|\nu_S - \varphi_G\|_{\mathrm{TV}} = O(1/n^\delta)$.

We remark that a simple implementation of the algorithm in Theorem 2.5 runs in time $O^*(n^{6/5})$; see Figure 1 in section 5 for details. The constant $\delta > 0$ that controls the error of the algorithm depends on p, q, Δ and gets smaller as p approaches $p_c(q, \Delta)$.

For integer values of q, Theorem 2.5 combined with the translation between the random-cluster and Potts models (cf. Lemma 2.1) yields a sampling algorithm for the ferromagnetic q-state Potts model on random regular graphs. Since uniqueness for the ferromagnetic Potts model holds iff $B < B_c(q, \Delta)$ and $p_c(q, \Delta) = 1 - \frac{1}{B_c(q,\Delta)}$, we therefore have the following corollary of Theorem 2.5.

COROLLARY 2.6. *Let $\Delta \geq 3$, $q \geq 3$ and $B > 1$ be in the uniqueness regime of the $(\Delta - 1)$-ary tree. Then, there exists a constant $\delta > 0$ such that, as $n \to \infty$, the following holds with probability $1 - o(1)$ over the choice of a random Δ-regular graph $G = (V, E)$ with n vertices.*

There is a polynomial-time algorithm which, on input the graph G, outputs a random assignment $\sigma : V \to [q]$ whose distribution ν_σ is within total variation distance $O(1/n^\delta)$ from the Potts distribution μ_G with parameter B, i.e., $\|\nu_\sigma - \mu_G\|_{\mathrm{TV}} = O(1/n^\delta)$.

2.4. Sampling antiferro Potts on random Δ-regular graphs. The algorithm of Corollary 2.6 for the ferromagnetic Potts model does not extend to the antiferromagnetic case since there is no analogous connection with the random-cluster model in this case. Nevertheless, we are able to design a sampling algorithm on random regular graphs when the parameter B is in uniqueness via a far more elaborate approach which consists of recoloring (large) bichromatic color classes.

THEOREM 2.7. *Let $\Delta \geq 3$, $q \geq 3$, and $B \in (0, 1)$ be in the uniqueness regime of the $(\Delta - 1)$-ary tree with $B \neq (\Delta - q)/\Delta$. Then, there exists a constant $\delta > 0$ such that, as $n \to \infty$, the following holds with probability $1 - o(1)$ over the choice of a random Δ-regular graph $G = (V, E)$ with n vertices.*

There is a polynomial-time algorithm which, on input the graph G, outputs a random assignment $\sigma : V \to [q]$ whose distribution ν_σ is within total variation distance $O(1/n^\delta)$ from the Potts distribution μ_G with parameter B, i.e.,

$$\|\nu_\sigma - \mu_G\|_{\mathrm{TV}} = O(1/n^\delta).$$

Note that the sampling algorithm in the antiferromagnetic case works throughout uniqueness apart from the point $(\Delta - q)/\Delta$, where uniqueness on the tree is expected to hold but the model is conjectured to be at criticality. In particular, for all $B \neq (\Delta - q)/\Delta$ which are in the uniqueness regime of $\mathbb{T}_\Delta$, it can be shown that the decay on the tree is exponentially small in its height. In contrast, even if uniqueness on the tree holds for $B = (\Delta - q)/\Delta$, it can be shown that the decay on the tree is only polynomial in its height.

We remark here that the algorithm for the antiferromagnetic case uses as a black-box a subroutine for sampling from the antiferromagnetic Ising model. The running time of this subroutine, which is based on correlation decay methods, is n^c for some constant $c = c(q, B, \Delta) > 0$; it is an open question whether there is a faster algorithm for the Ising model. Finally, the constant δ that controls the error of the algorithm depends on B, q, Δ and gets smaller as B decreases.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

3. Proof approach. In this section, we outline the main idea behind the algorithms of Theorems 2.5 and 2.7 and the key obstacles that we have to address. We focus on the antiferromagnetic Potts model where the details are much more complex and discuss how we get the simplification for the ferromagnetic case via the random-cluster model later.

DEFINITION 3.1. *For an n-vertex graph $G = (V, E)$ with maximum degree Δ, a cycle is* short *if its length is at most $\frac{1}{5}\log_{\Delta-1} n$ and is* long *otherwise.*

Let G be a random Δ-regular graph with n vertices. Following the approach of Efthymiou [8], our algorithm starts from the subgraph of G consisting of all short cycles, which we denote by G'. It is fairly standard to show that, with probability $1 - o(1)$ over the choice of G, the subgraph G' is a *disjoint* union of short cycles; see Lemma 4.1. It is therefore possible to sample a configuration σ' on G' which is distributed according to the Potts distribution $\mu_{G'}$ (exactly). This can be accomplished in several ways; in fact, since the cycles are disjoint and each cycle has logarithmic length, this initial sampling step can even be done via brute force in polynomial time (though it is not hard to come up with much faster algorithms).

After this initial preprocessing, the algorithm then proceeds by adding sequentially the edges that do not belong to short cycles. At each step, the current configuration is updated with the aim to preserve its distribution close to the Potts distribution of the new graph (with the edge that we just added). Key to this update procedure is a resampling step which is performed only when the endpoints of a newly added edge $\{u, v\}$ happen to have the same colors under the current configuration; intuitively, some action is required in this case because the weight of the current configuration reduces by a factor of $B < 1$ in the new graph (because of the added edge). The resampling step consists of recoloring a bichromatic class, where the latter is defined as follows.

DEFINITION 3.2. *Let $G = (V, E)$ be a graph and $\sigma : V \to [q]$ be a configuration. For colors $c_1, c_2 \in [q]$, let $\sigma^{-1}(c_1, c_2)$ be the set of vertices that have either color c_1 or color c_2 under σ.*

For distinct colors $c_1, c_2 \in [q]$, we say that $U = \sigma^{-1}(c_1, c_2)$ is the (c_1, c_2)-color-class of σ and that U is a bichromatic class under σ. We refer to a connected component of $G[U]$ as a bichromatic component.

In the proper colorings case ($B = 0$), Efthymiou [8] demonstrated that the resampling step when adding an edge $e = \{u, v\}$ can be done by just flipping the colors of a bichromatic component chosen uniformly at random among those containing one of the vertices u and v (say, u). The rough idea there is that when the colorings model is in uniqueness, the bichromatic components on a random graph are typically small in size. At the same time, by the initial preprocessing step, the edge $e = \{u, v\}$ does not belong to a short cycle and therefore u and v are far away in the graph without e. Hence, u and v are unlikely to belong to the same bichromatic component and the flipping step will succeed in giving u and v different colors with good probability.

Unfortunately, this flipping method does not work for the antiferromagnetic Potts model. It turns out that when $q < \Delta$ and even when the Potts model is in uniqueness, bichromatic components can be large and therefore u and v potentially belong to the same bichromatic component. To make matters worse, these bichromatic components can be quite complicated (with many short/long cycles). This necessitates a more elaborate approach in our setting to succeed in giving u and v different colors without introducing significant bias to the sampler.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The key to overcoming these obstacles lies in the observation that the assignment of the two colors in a bichromatic component follows the Ising distribution; see Observation 6.1 for the precise formulation. Hence we can hope to use an approximate sampling algorithm for the Ising model in the resampling step. The natural implementation of this idea, however, fails: known algorithms for the antiferromagnetic Ising model, based on correlation decay, work as long as $B > \frac{\Delta-2}{\Delta}$, where Δ is the maximum degree of the graph [30, 35]. In general, this inequality is not satisfied for us, i.e., there exist B in the uniqueness regime for Potts such that $B < \frac{\Delta-2}{\Delta}$.

Fortunately, we can employ fairly recent technology for two-state models [27, 31, 29] which demonstrates that the graph parameter that matters is not actually the maximum degree of the graph but rather the "average growth" of the graph. While we cannot apply any of the existing results in the literature directly, adapting these ideas to the antiferromagnetic Ising model is fairly straightforward, using results from Mossel and Sly [27]. The more difficult part in our setting is proving that the average growth of the bichromatic components that we consider for resampling is indeed small for "typical" configurations σ (note that in the worst case, the whole graph can be a bichromatic class which has large average growth for our purposes, so a probability estimate over σ is indeed due). Let us first formalize the notion of average growth that we use.

DEFINITION 3.3. *Let M, b be positive constants and $G = (V, E)$ be a graph with n vertices. We say that G has average growth b up to depth $L = \lceil M \log n \rceil$ if for all vertices $v \in V$ the total number of paths with L vertices starting from v is less than b^L.*

The notion of average growth is similar to the notion of connective constant for finite graphs used in [31, 29]; the reason for the slightly different definition is that we will need an explicit handle on the constant M controlling the depth. Note that, since we only consider paths with a fixed logarithmic length, this places a lower bound on the accuracy of the sampling algorithm. Nevertheless, by choosing the constant M sufficiently large, this will still be sufficient to make the error of our sampler polynomially small. In particular, as long as the inequality $b\frac{1-B}{1+B} < 1$ is satisfied, using results from [27], we obtain an approximate sampler for the antiferromagnetic Ising model with parameter B on graphs of average growth b up to depth $L = \lceil M \log n \rceil$; see Theorem 6.3 for details.

We give a few more technical details on how we bound the average growth of bichromatic classes. Here, we utilize the tree uniqueness and the tree-like structure of random Δ-regular graphs (cf. Lemma 4.4) to provide an upper bound on the number of bichromatic paths. For paths of logarithmic length L, we show in Lemma 6.6 that the probability that a path is bichromatic is $\leq K^L$, where K is roughly $(1+B)/(B+q-1)$. Since there are at most $\Delta(\Delta-1)^{L-2}$ paths with L vertices, we therefore obtain that the average growth b of bichromatic components is bounded above by $(\Delta-1)K$. When B is in uniqueness, we have that $B > (\Delta-q)/\Delta$, and therefore the inequality $b\frac{1-B}{1+B} < 1$ that is required for the Ising sampler is satisfied.

The final technical piece is to bound the error that is introduced by the resampling steps. The placement of a new edge $\{u, v\}$ reweights the probability that u and v have different colors and introduces an error in our sampling algorithm that is captured by the correlation between the colors of u and v (see Lemma 6.9). The main idea at this point is that, in the graph without the edge $\{u, v\}$, u and v are far apart (since the edge $\{u, v\}$ does not belong to a short cycle of G) which can be used to show that the correlation between the colors of u and v is relatively small. In Lemma 6.4, we show

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

that the correlation between u and v can in fact be upper bounded as a weighted sum over paths connecting u and v. This allows us to bound the aggregate sampling error of the algorithm as a weighted sum over short cycles of the random graph G, which can in turn be bounded using a simple expectation argument.

The algorithm that we described for the antiferromagnetic Potts model can actually be adapted to the ferromagnetic case as well. However, as mentioned earlier, we follow a different (and surprisingly simpler) route using the random-cluster representation of the model. At a very rough level, the reason behind the simplification is that the components in the random-cluster model provide a much better grip on capturing the properties of the Potts distribution than the bichromatic-component proxy we used earlier. Indeed, just as we described in the antiferromagnetic case, bichromatic components for the ferromagnetic Potts model can also be linear-sized. However, once we translate the Potts configuration to its random-cluster representation (cf. Lemma 2.1), the components in the latter are small in size (when the model is in the uniqueness region $p < p_c(q, \Delta)$) and therefore vertices that are far away do not belong to the same component. This allows us to perform the resampling step in the random-cluster model by a simple percolation procedure. The details can be found in section 5.

3.1. Organization. In section 4, we give the properties of random regular graphs that we are going to use in the analysis of our sampling algorithms. In section 5, we give the algorithm for the random-cluster model and conclude Theorem 2.5 (assuming the upcoming Lemma 5.2). In section 6, we give the algorithm for the antiferromagnetic Potts model and conclude Theorem 2.7 (assuming the upcoming Theorem 6.3 and Lemmas 6.4 and 6.6). In sections 7 and 8, we analyze the random-cluster and antiferromagnetic Potts models on "tree-like" graphs and give the proofs of Lemmas 5.2 and 6.6. Finally, in section 9, we prove Theorem 6.3 and Lemma 6.4, which are about correlation decay and sampling for the antiferromagnetic Ising model on graphs of small average growth.

4. Properties of random regular graphs. In this section, we state and prove structural properties of random Δ-regular graphs which ensure that our algorithms for the random-cluster and Potts models have the desired accuracy (cf. Remark 1.1). While the exact statements of the properties that we need do not seem to appear in the literature, their proofs follow fairly standard techniques in the area.

We will work in the configuration model; see [18, Chapter 9] for more details. Precisely, for Δn even, let $\mathcal{G} := \mathcal{G}_{n,\Delta}$ denote the uniform distribution on Δ-regular graphs which is obtained by taking a perfect matching of the set $[n] \times [\Delta]$ and collapsing for each $u \in [n]$ the elements $(u, 1), \ldots, (u, \Delta)$ into a single vertex u; the elements of the set $[n] \times [\Delta]$ are called *points*. Technically, the distribution $\mathcal{G}$ is supported on multigraphs but it can be shown that the probability that $G \sim \mathcal{G}$ is simple is asymptotically a positive constant as $n \to \infty$; conditioned on that event, G is uniformly distributed over Δ-regular graphs with n vertices, and therefore any event that holds with probability $1 - o(1)$ in $\mathcal{G}_{n,\Delta}$ also holds with probability $1 - o(1)$ over the uniform distribution on Δ-regular graphs with n vertices.

The following lemma guarantees that short cycles are disjoint in a random Δ-regular graph.

LEMMA 4.1. *Let $\Delta \geq 3$ be an integer. Then, with probability $1 - o(1)$ over the choice of a uniformly random Δ-regular graph with n vertices, any two distinct cycles of length $\leq \frac{1}{5} \log_{\Delta-1} n$ are disjoint, i.e., they do not share any common vertices or edges.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Proof. For convenience, let $\ell := \lfloor \frac{1}{5} \log_{\Delta-1} n \rfloor$ and $\mathcal{G} := \mathcal{G}_{n,\Delta}$. Let $G \sim \mathcal{G}$ and $\mathcal{E}$ be the event that G contains two distinct cycles of length $\leq \ell$ which are not disjoint. Let also $\mathcal{F}$ be the event that, for some integer $k \in [1, 2\ell]$, G contains a subgraph with k vertices and $k+1$ edges such that each vertex has at least two incident multiedges (in the subgraph); we call such a subgraph *bad*. When $\mathcal{E}$ occurs, we obtain that $\mathcal{F}$ also occurs, so it suffices to upper bound the probability of the latter.

To bound $\Pr_{\mathcal{G}}(\mathcal{F})$, we will use a union over all possible bad subgraphs with k vertices with $k \in [1, 2\ell]$. Consider an arbitrary integer $k \in [1, 2\ell]$. There are $\binom{n}{k} \leq (\frac{en}{k})^k$ ways to choose the vertices in the subgraph, at most $\binom{\Delta}{2}^k (\Delta k)^2$ ways to choose the points that are going to be paired, and then $\frac{(2(k+1))!}{(k+1)!2^{k+1}} \leq 10(2k/\mathrm{e})^{k+1}$ ways to pair the points. The probability of a particular pairing of $2(k+1)$ points occuring is equal to $\frac{1}{(\Delta n-1)\cdots(\Delta n-(2k+1))} \leq \frac{2}{(\Delta n)^{k+1}}$ (using that $k \leq 2\ell = o(n^{1/2})$). We therefore obtain that

$$\begin{aligned}\Pr_{\mathcal{G}}(\mathcal{F}) &\leq 20 \sum_{k=1}^{2\ell} \binom{n}{k} \frac{(\Delta k)^2 \binom{\Delta}{2}^k (2k/\mathrm{e})^{k+1}}{(\Delta n)^{k+1}} \leq \frac{60\Delta}{n} \sum_{k=1}^{2\ell} k^2 (\Delta-1)^k \\ &\leq \frac{500\Delta\ell^3}{n} (\Delta-1)^{2\ell} = o(1),\end{aligned}$$

as needed. This finishes the proof. ∎

The following lemma guarantees that certain weighted sums over cycles are small; this bound will be used to show that the aggregate error of our samplers is small (cf. section 3).

LEMMA 4.2. *Let $\Delta \geq 3$. Then, for any constant $W > \Delta - 1$ and any constant $\ell_0 > 0$, there exists a constant $\delta > 0$ such that the following holds with probability $1 - O(1/n^\delta)$ over the choice of $G \sim \mathcal{G}_{n,\Delta}$. Let C_ℓ denote the number of cycles of length ℓ. Then,*

$$\sum_{\ell \geq \ell_0 \log n} \frac{\ell C_\ell}{W^\ell} \leq 1/(2n^\delta).$$

Proof. Let $w := (\Delta-1)/W < 1$ and $\delta := \frac{1}{2}\ell_0 \log(1/w) > 0$. We will show that, for all sufficiently large n, it holds that

$$\mathbf{E}_{\mathcal{G}}[X] \leq \frac{1}{(1-w)n^{2\delta}}, \quad \text{where } X := \sum_{\ell \geq \ell_0 \log n} \frac{\ell C_\ell}{W^\ell}. \tag{4.1}$$

Once we show (4.1), the result follows from Markov's inequality.

Fix an arbitrary integer $\ell \in [n]$; we will calculate $\mathbf{E}_{\mathcal{G}}[C_\ell]$. In the configuration model, a cycle with ℓ vertices corresponds to an ordered 2ℓ-tuple of points $(u_1, i_1), (u_1, i_2), \ldots, (u_\ell, i_{2\ell-1}), (u_\ell, i_{2\ell})$ where $u_1, \ldots, u_\ell$ are distinct elements of $[n]$ and $i_1, \ldots, i_{2\ell} \in [\Delta]$ such that (u_j, i_{2j}) is paired to (u_{j+1}, i_{2j+1}) for all $j \in [\ell]$ (with the convention that $u_{\ell+1} = u_1$ and $i_{2l+1} = i_1$).

There are $\ell!\binom{n}{\ell}$ ways to choose and order $u_1, \ldots, u_\ell$ and $(\Delta(\Delta-1))^\ell$ ways to choose $i_1, \ldots, i_{2\ell}$ for a total of $\ell!\binom{n}{\ell}(\Delta(\Delta-1))^\ell$ possible tuples; this overcounts the number of tuples corresponding to distinct cycles by a factor of 2ℓ (the number of ways to root and orient the 2ℓ-tuple). Now, the pairing corresponding to a tuple occurs with probability $\frac{1}{(\Delta n-1)(\Delta n-3)\cdots(\Delta n-(2\ell-1))}$. Since

$$2(\Delta n-1)(\Delta n-3)\cdots(\Delta n-(2\ell-1)) \geq \Delta^\ell n(n-1)\cdots(n-\ell+1)$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

for all $\ell \in [1, n]$ and $\Delta \geq 3$, it follows that $\mathbf{E}_{\mathcal{G}}[C_\ell] \leq (\Delta - 1)^\ell / \ell$ and hence

$$\mathbf{E}_{\mathcal{G}}[X] = \sum_{\ell \geq \ell_0 \log n} w^\ell \leq w^{\ell_0 \log n}/(1-w),$$

which proves (4.1) and therefore concludes the proof. ∎

Our next lemma captures the tree-like structure of random Δ-regular graphs that will be relevant for us. In particular, we give a description of the neighborhood structure around a path. To do this accurately, we will need a few definitions. Let $G = (V, E)$ be a graph. For a vertex $v \in V$ and integer $h \geq 0$, we denote by $\Gamma_h(G, v)$ the set of vertices at distance $\leq h$ from v.

DEFINITION 4.3. *Let G be a graph and P be a path in G with vertices $u_1, \ldots, u_\ell$. Let $G \backslash P$ be the graph obtained from G by removing the edges of the path P. Then, for an integer $h \geq 0$, the* h-graph-neighborhood *of the path P is the subgraph of $G \backslash P$ induced by the vertex set $\bigcup_{i \in [\ell]} \Gamma_h(G \backslash P, u_i)$.*

A connected component of the h-graph-neighborhood will be called isolated *if it contains* exactly one *of the vertices $u_1, \ldots, u_\ell$.*

LEMMA 4.4. *Let $\Delta \geq 3$. Then, for any constant integer $h \geq 0$ and any $\epsilon > 0$, there exists a constant $\ell_1 > 0$ such that the following holds.*

With probability $1 - O(1/n^2)$ over the choice of $G \sim \mathcal{G}_{n,\Delta}$, every path P in G with ℓ vertices with $\ell_1 \leq \ell \leq n^{9/10}$ has an h-graph-neighborhood with at least $(1-\epsilon)\ell$ isolated tree components.

We will use the following version of the well-known Chernoff/Hoeffding inequality.

LEMMA 4.5 (see, e.g., [10, Theorem 21.6 and Corollary 21.9]). *Suppose that $S_n = X_1 + \cdots + X_n$, where $\{X_i\}_{i \in [n]}$ is a collection of independent random variables such that $0 \leq X_i \leq 1$ and $\mathbb{E}[X_i] = \mu_i$ for $i = 1, \ldots, n$. Let $\mu = \mu_1 + \cdots + \mu_n$. Then, for any $c > 1$,*

$$\Pr(S_n \geq c\mu) \leq \exp\big(-c\mu \log(c/\mathrm{e})\big).$$

Proof of Lemma 4.4. Fix an arbitrary integer $h \geq 0$ and constant $\epsilon > 0$. We will show that the lemma holds with $\ell_1 := 200/\epsilon$.

In the configuration model, a path P with ℓ vertices corresponds to an ordered $2(\ell-1)$-tuple of points $(u_1, i_1), (u_2, i_2), (u_2, i_3), \ldots, (u_{\ell-1}, i_{2(\ell-2)}), (u_{\ell-1}, i_{2\ell-3}), (u_\ell, i_{2(\ell-1)})$ where $u_1, \ldots, u_\ell$ are distinct elements of $[n]$ and $i_1, \ldots, i_{2(\ell-1)} \in [\Delta]$ such that (u_j, i_{2j-1}) is paired to (u_{j+1}, i_{2j}) for all $j \in [\ell - 1]$. Fix any such path P with ℓ vertices $u_1, \ldots, u_\ell$ and condition on the event that P appears in G. We will next reveal the h-graph-neighborhood of P in a breadth-first search manner, as follows:

1. Let $U_0 = \{u_1, \ldots, u_\ell\}$. Initialize $U_1, \ldots, U_h, U_{h+1}$ to be empty sets of vertices.
2. For $t = 0, 1, \ldots, h$:
3. Order the vertices in U_t in lexicographic order.
4. For $k = 1, \ldots, \ell\Delta^t$:
5. Pick the kth vertex in U_t, say, u (if $k > |U_t|$, set $u = 0$).
6. For $i = 1, \ldots, \Delta$:
7. If $u \neq 0$ and the point (u, i) is not already paired,
8. Pair (u, i) with a point not already paired chosen uniformly, say, (v, j).
9. If $v \notin U_1 \cup \cdots \cup U_t \cup U_{t+1}$, add v to U_{t+1}.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

By induction, we have that, for all $t \geq 0$, U_t consists of the set of vertices at distance t from a vertex in $\{u_1, \dots, u_\ell\}$. Note also that $|U_t| \leq \ell\Delta^t$. Now, fix arbitrary $t \in \{0, 1, \dots, h\}$, $k \in [\ell\Delta^t]$, and $i \in [\Delta]$. Let $\mathcal{F}_{t,k,i}$ be the pairings that we have revealed about the graph G just before executing lines 7–9. Similarly, let $S_{t,k,i}$ be the set of vertices we have encountered just before executing lines 7–9 (i.e., the union of $U_1, \dots, U_t$ together with the current set U_{t+1}). Let also $Q_{t,k,i}$ be the event that in lines 7–8 all of the following happen: (i) in line 7, $u \neq 0$ and the point (u, i) is not paired, and (ii) in line 8, (u, i) gets paired to a point in $S_{t,k,i} \times [\Delta]$. There are at least $\Delta(n - |S_{t,k,i}|)$ points that have not been paired, so

$$\Pr(Q_{t,k,i} \mid \mathcal{F}_{t,k,i}) \leq \frac{|S_{t,k,i}|}{n - |S_{t,k,i}|} \leq \frac{2\ell\Delta^{h+2}}{n}, \tag{4.2}$$

where in the last inequality we used that

$$|S_{t,k,i}| \leq |U_0| + \cdots + |U_{t+1}| \leq \ell\Delta^{t+2} \leq \ell\Delta^{h+2} \text{ and } n - \ell\Delta^{h+2} \geq n/2$$

for all sufficiently large n (using that $\ell_1 \leq \ell \leq n^{9/10}$). Using (4.2), we obtain that the number of events $\{Q_{t,k,i}\}$ that occur is dominated above by a binomial random variable (r.v.) $X_\ell \sim \mathrm{Bin}(\ell\Delta^{h+2}, \frac{2\ell\Delta^{h+2}}{n})$. By Lemma 4.5, we have that

$$\Pr[X_\ell \geq \epsilon\ell/2] \leq \mathrm{e}^{-\frac{1}{2}\epsilon\ell\log(c_\ell/\mathrm{e})}, \text{ where } c_\ell = \frac{\epsilon n}{4\ell\Delta^{2h+4}}.$$

Since $\ell \leq n^{9/10}$, we have $c_\ell \geq \mathrm{e}n^{1/20}$ for all sufficiently large n and hence

$$\Pr[X_\ell \geq \epsilon\ell] \leq \mathrm{e}^{-\frac{1}{40}\epsilon\ell\log n}.$$

It follows that for any path P with ℓ vertices, with probability $\geq 1 - \mathrm{e}^{-\frac{1}{40}\epsilon\ell\log n}$, at most $\epsilon\ell/2$ of the events $\{Q_{t,k,i}\}$ occur, i.e., the h-graph-neighborhood of P contains at least $(1-\epsilon)\ell$ isolated tree components (every event $Q_{t,k,i}$ that occurs decreases the number of isolated tree components by at most two—on the other hand, if $Q_{t,i,k}$ does not occur, then the number of isolated tree components stays the same). Since there are at most $n\Delta^\ell$ paths with ℓ vertices, we obtain by a union bound that the probability that there exists a path whose h-graph-neighborhood contains less than $(1-\epsilon)\ell$ isolated tree components is upper bounded by

$$\sum_{\ell=\lceil 200/\epsilon\rceil}^{n^{9/10}} \mathrm{e}^{\log n + \ell\log\Delta - \frac{1}{40}\epsilon\ell\log n} = O(1/n^2),$$

where the last bound follows by observing that, for all sufficiently large n, the summands are decreasing functions of ℓ and that for $\ell = \lceil 200/\epsilon\rceil$ we have $\mathrm{e}^{\log n + \ell\log\Delta - \frac{1}{40}\epsilon\ell\log n} = O(1/n^3)$.

This concludes the proof of Lemma 4.4. ∎

To conclude this section, we clarify a small point relevant to Remark 1.1. We will only utilize Lemma 4.4 for paths of logarithmic length (although the lemma is stated for convenience for much longer paths) and therefore the property can be checked in polynomial time. Analogously, the sum in Lemma 4.2 will only be considered for cycles of logarithmic length and therefore the (restricted) inequality can also be checked in polynomial time.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

5. Algorithm for the random-cluster model. In this section, we prove Theorem 2.5. In section 5.1, we first describe the algorithm and analyze how to update a random-cluster configuration when we add a new edge. In section 5.2, we show how to control the aggregate error of our sampling algorithm on random regular graphs. Finally, in section 5.3, we combine these pieces to conclude Theorem 2.5.

5.1. The algorithm. To prove Theorem 2.5, we will consider a simple percolation algorithm for sampling an random-cluster configuration on a random Δ-regular graph G. The algorithm is given in Figure 1 and in section 5.3 we will detail its performance when the input is a random regular graph.

Prior to that, let us first motivate the algorithm SampleRC, by demonstrating how to update an random-cluster configuration when we add a single edge $\{u, v\}$. To control the effect of adding an edge, it will be relevant to consider the event that there is an open path between u and v (for a path P in G and an random-cluster configuration $S \subseteq E$, we say that P is open in S if all of its edges belong to S); we denote this event by $u \leftrightarrow v$.

LEMMA 5.1. *Let $p \in (0,1)$ and $q \geq 1$, and consider arbitrary $\epsilon \in (0, 1/q)$.*

Let $G = (V, E)$ be a graph and u, v be two vertices such that $\{u, v\} \notin E$ and $\varphi_G(u \leftrightarrow v) \leq \epsilon$. Consider the graph $G' = (V, E')$ obtained from G by adding the edge $\{u, v\}$. Sample a random subset of edges $Y \subseteq E'$ as follows: first, sample a subset of edges $X \subseteq E$ according to the random-cluster distribution φ_G and, then, set $Y = X \cup \{e\}$ with probability $p/(p + (1-p)q)$, and $Y = X$ otherwise.

Then, the distribution of Y, denoted by ν_Y, is within total variation distance $2q\epsilon$ from the random-cluster distribution $\varphi_{G'}$ on G' with parameters p, q, i.e.,

$$\|\nu_Y - \varphi_{G'}\|_{\mathrm{TV}} \leq 2q\epsilon.$$

Proof. Fix an arbitrary $\epsilon \in (0, 1/q)$.

Let Ω_{op} be the set of subsets $S \subseteq E$ such that in the graph (V, S), u and v are connected by an open path and let $\Omega_{\mathrm{cl}} = 2^E \backslash \Omega_{\mathrm{op}}$. For $S \subseteq E$, denote for convenience by S_e the set $S \cup \{e\}$. Observe that

Algorithm SampleRC(G)

parameters: reals $p \in (0,1)$ and $q \geq 1$

Input: Graph $G = (V, E)$
Output: Either Fail or a set $S \subseteq E$

$E' := \{e \in E \mid e$ belongs to a short cycle$\}$

if $G' = (V, E')$ has a component which is neither a cycle nor an isolated vertex
  then Fail
else
  Sample an random-cluster configuration $S' \subseteq E'$ on G' (according to $\varphi_{G'}$);
  Add to S' each edge in $E \backslash E'$ independently with probability $p/(p+(1-p)q)$;
  Output the resulting set $S \subseteq E$.

FIG. 1. *Algorithm for sampling a random-cluster configuration. Note that since G' is a disjoint union of short cycles, the initial configuration S' in the algorithm can be obtained quickly in various ways (e.g., even brute force takes time $O^*(n^{6/5})$ since each cycle has length $\leq \frac{1}{5}\log_{\Delta-1} n$ and there are at most n cycles).*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$\forall S \in \Omega_{\rm cl}: \qquad w_{G'}(S_e) = \frac{p}{q} w_G(S), \quad w_{G'}(S) = (1-p) w_G(S), \tag{5.1}$$

$$\forall S \in \Omega_{\rm op}: \qquad w_{G'}(S_e) = p w_G(S), \quad w_{G'}(S) = (1-p) w_G(S). \tag{5.2}$$

Note also that for any $S \subseteq E$, we have

$$\nu_Y(S_e) = \frac{p\,\varphi_G(S)}{p+q(1-p)} = \frac{\frac{p}{q} w_G(S)}{(\frac{p}{q}+1-p) Z_G}, \ \nu_Y(S) = \frac{q(1-p)\varphi_G(S)}{p+q(1-p)} = \frac{(1-p) w_G(S)}{(\frac{p}{q}+1-p) Z_G}.$$

Using (5.1), (5.2), and the assumption $\varphi_G(u \leftrightarrow v) \le \epsilon$, we will also show the following for the partition functions Z_G, Z'_G:

$$|M| \le q\epsilon, \text{ where } M := \frac{(\frac{p}{q}+1-p) Z_G}{Z_{G'}} - 1. \tag{5.3}$$

Let us conclude the proof assuming, for now, (5.3). To do this, we decompose $\|\nu_Y - \varphi_{G'}\|_{\rm TV}$ as

$$\|\nu_Y - \varphi_{G'}\|_{\rm TV} = \frac{1}{2} \sum_{S \subseteq E} |\nu_Y(S) - \varphi_{G'}(S)| + \frac{1}{2} \sum_{S \subseteq E} |\nu_Y(S_e) - \varphi_{G'}(S_e)|. \tag{5.4}$$

For $S \subseteq E$, we have

$$\begin{aligned} |\nu_Y(S) - \varphi_{G'}(S)| &= \left| \frac{(1-p) w_G(S)}{(\frac{p}{q}+1-p) Z_G} - \frac{(1-p) w_G(S)}{Z_{G'}} \right| = \frac{(1-p) w_G(S)}{(\frac{p}{q}+1-p) Z_G} |M| \\ &\le |M|\, \varphi_G(S), \end{aligned}$$

and therefore we can bound the first sum in (5.4) as

$$\sum_{S \subseteq E} |\nu_Y(S) - \varphi_{G'}(S)| \le |M| \le q\epsilon. \tag{5.5}$$

To bound the second sum in (5.4), we consider whether $S \in \Omega_{\rm cl}$ or $S \in \Omega_{\rm op}$. For $S \in \Omega_{\rm cl}$, we have

$$|\nu_Y(S_e) - \varphi_{G'}(S_e)| = \left| \frac{\frac{p}{q} w_G(S)}{(\frac{p}{q}+1-p) Z_G} - \frac{\frac{p}{q} w_G(S)}{Z_{G'}} \right| = \frac{\frac{p}{q} w_G(S)}{(\frac{p}{q}+1-p) Z_G} |M| \le |M|\, \varphi_G(S),$$

while for $S \in \Omega_{\rm op}$, we have

$$\begin{aligned} |\nu_Y(S_e) - \varphi_{G'}(S_e)| &= \left| \frac{\frac{p}{q} w_G(S)}{(\frac{p}{q}+1-p) Z_G} - \frac{p w_G(S)}{Z_{G'}} \right| = \frac{\frac{p}{q} w_G(S)}{(\frac{p}{q}+1-p) Z_G} |qM + (q-1)| \\ &\le 2q\, \varphi_G(S), \end{aligned}$$

where the last inequality follows from $|qM + (q-1)| \le q(|M|+1) \le 2q$ (using that $|M| \le q\epsilon \le 1$ from (5.3)). Hence

$$\sum_{S \subseteq E} |\nu_Y(S_e) - \varphi_{G'}(S_e)| \le |M|\, \varphi_G(\Omega_{\rm cl}) + 2q\, \varphi_G(\Omega_{\rm op}) \le |M| + 2q\epsilon \le 3q\epsilon. \tag{5.6}$$

Plugging (5.5) and (5.6) into (5.4), we obtain that $\|\nu_Y - \varphi_{G'}\|_{\rm TV} \le 2q\epsilon$ as wanted.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

To complete the proof, it only remains to show (5.3). Let

$$Z_{G,\mathrm{op}} = \sum_{S \in \Omega_{\mathrm{op}}} w_G(S), \quad Z_{G,\mathrm{cl}} = \sum_{S \in \Omega_{\mathrm{cl}}} w_G(S),$$

so that $Z_G = Z_{G,\mathrm{op}} + Z_{G,\mathrm{cl}}$. By assumption, we have that

$$\varphi_G(u \leftrightarrow v) = \frac{Z_{G,\mathrm{op}}}{Z_G} \le \epsilon, \text{ so that } \frac{Z_{G,\mathrm{op}}}{Z_{G,\mathrm{cl}}} \le \frac{\epsilon}{1-\epsilon}.$$

Let also

$$Z_{G',\mathrm{op}} = \sum_{S \in \Omega_{\mathrm{op}}} w_{G'}(S) + w_{G'}(S_e), \quad Z_{G',\mathrm{cl}} = \sum_{S \in \Omega_{\mathrm{cl}}} w_{G'}(S) + w_{G'}(S_e),$$

so that $Z_{G'} = Z_{G',\mathrm{op}} + Z_{G',\mathrm{cl}}$. Using (5.1) and (5.2), we obtain that

$$Z_{G',\mathrm{op}} = Z_{G,\mathrm{op}}, \qquad Z_{G',\mathrm{cl}} = \Big(\frac{p}{q} + 1 - p\Big) Z_{G,\mathrm{cl}}$$

and therefore

$$\frac{Z_G}{Z_{G'}} = \frac{Z_{G,\mathrm{op}} + Z_{G,\mathrm{cl}}}{Z_{G',\mathrm{op}} + Z_{G',\mathrm{cl}}} = \frac{\frac{Z_{G,\mathrm{op}}}{Z_{G,\mathrm{cl}}} + 1}{\frac{Z_{G,\mathrm{op}}}{Z_{G,\mathrm{cl}}} + \big(\frac{p}{q} + 1 - p\big)}.$$

Since $\frac{Z_{G,\mathrm{op}}}{Z_{G,\mathrm{cl}}} \in (0, \frac{\epsilon}{1-\epsilon}]$ and the function $f(x) := \frac{x+1}{x+(\frac{p}{q}+1-p)}$ is decreasing in x, we obtain that $f(\frac{\epsilon}{1-\epsilon}) \le \frac{Z_G}{Z_{G'}} \le f(0)$ and hence the bounds

$$\frac{\frac{p}{q} + 1 - p}{\epsilon + (1-\epsilon)(\frac{p}{q} + 1 - p)} \le \frac{(\frac{p}{q} + 1 - p) Z_G}{Z_{G'}} \le 1.$$

The left-hand side (l.h.s.), is $\ge 1 - q\epsilon$ for all $p \in (0,1)$, $q \ge 1$, and $\epsilon \in (0,1)$, and hence (5.3) follows.

This concludes the proof of Lemma 5.1. ∎

5.2. Aggregating the error. To utilize Lemma 5.1, we need to upper bound the probability that two vertices belong to the same component in an random-cluster configuration. In turn, it suffices to bound the probability that there is an open path between the vertices. To this end, we utilize the fact that the parameters p, q are in the uniqueness region of the $(\Delta-1)$-ary tree and the tree-like structure around paths (cf. Definition 4.3) to show the following. The proof of the lemma is given in section 7.

LEMMA 5.2. *Let $\Delta \ge 3$ be an integer, $q \ge 1$, and $p < p_c(q, \Delta)$. There exist constants $K < 1/(\Delta - 1)$ and $\epsilon > 0$ such that the following holds for all sufficiently large integers ℓ and h.*

Let G be a Δ-regular graph and P be a path with ℓ vertices whose h-graph-neighborhood contains $(1-\epsilon)\ell$ isolated tree components. Let φ_G be the random-cluster distribution on G with parameters p, q. Then,

$$\varphi_G(\textit{path } P \textit{ is open}) \le K^{\ell}.$$

Using monotonicity properties of the random-cluster distribution, we can extend Lemma 5.2 to arbitrary subgraphs of a target graph G. In particular, suppose that G, P are as in Lemma 5.2 and that G' is a subgraph of G which contains the path

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

P. Then it also holds that $\varphi_{G'}(P \text{ is open}) \leq K^{\ell}$. We will not define the notion of monotonic distributions in its full generality, but instead we will just state the following property of random-cluster distributions which will be sufficient for our purposes; see [15, sections 2.1 and 2.2] for a detailed exposition.

LEMMA 5.3 (see, e.g., [15, Chapter 2]). *Let $G = (V, E)$ be a graph and consider the random-cluster distribution on G with parameters $p \in (0, 1)$ and $q \geq 1$. Then for any subsets $S, S' \subseteq E$ such that $S \subseteq S'$, it holds that*

$$\varphi_G(\mathcal{F} \mid S \text{ open}) \leq \varphi_G(\mathcal{F} \mid S' \text{ open}) \quad \text{and} \quad \varphi_G(\mathcal{F} \mid S \text{ closed}) \geq \varphi_G(\mathcal{F} \mid S' \text{ closed})$$

for any increasing event $\mathcal{F}$.[2]

Combining Lemmas 4.2, 4.4, and 5.2, we can now conclude the following.

LEMMA 5.4. *Let $\Delta \geq 3$ be an integer, $q \geq 1$, and $p < p_c(q, \Delta)$. Then, there exists a constant $\delta > 0$ such that, as $n \to \infty$, the following holds with probability $1 - o(1)$ over the choice of a uniformly random Δ-regular graph $G = (V, E)$ with n vertices.*

Let $e_1, \ldots, e_t$ be the edges of G that do not belong to short cycles. For $j \in [t]$, let $e_j = \{u_j, v_j\}$ and G_j be the subgraph $G \backslash \{e_1, \ldots, e_j\}$. Then, it holds that

$$\sum_{j=1}^{t} \varphi_{G_j}(u_j \leftrightarrow v_j) \leq 1/n^{\delta}. \tag{5.7}$$

Proof. Let $K < 1/(\Delta - 1)$ and $\epsilon > 0$ be the constants in Lemma 5.2, and let $\ell', h' > 0$ be constants so that Lemma 5.2 applies for all $\ell \geq \ell'$ and $h \geq h'$. Fix h to be any integer greater than h'. Let $\delta > 0$ be the constant in Lemma 4.2 corresponding to $\ell_0 := 1/(5 \log(\Delta - 1))$ and $W = 1/K$ (note that $W > \Delta - 1$). Let $\ell_1 > 0$ be the constant in Lemma 4.4 corresponding to our choice of ϵ. Finally, let $\ell_2 := \max\{4 \log(W/(\Delta - 1)), 2\ell_0\}$.

Taking a union bound over Lemmas 4.2 and 4.4, we have that a uniformly random Δ-regular graph $G = (V, E)$ with n vertices satisfies the following with probability $1 - o(1)$ over the choice of the graph:

1. $\sum_{\ell=L_0}^{L_2} \frac{\ell C_\ell}{W^\ell} \leq 1/(2n^{\delta})$, where $L_0 := \lceil \ell_0 \log n \rceil = \lceil \frac{1}{5} \log_{\Delta-1} n \rceil$, $L_2 := \lfloor \ell_2 \log n \rfloor$ and C_ℓ is the number of cycles of length ℓ in G.
2. Every path P in G with ℓ vertices where $\ell_1 \leq \ell \leq L_2 + 1$ has an h-graph-neighborhood with at least $(1 - \epsilon)\ell$ isolated tree components.

We will show that for any Δ-regular graph G which satisfies items 1 and 2, it holds that

$$\sum_{j=1}^{t} \varphi_{G_j}(u_j \leftrightarrow v_j) \leq 1/n^{\delta}, \tag{5.7}$$

where $e_1 = \{u_1, v_1\}, \ldots, e_t = \{u_t, v_t\}$ are the edges of G that do not belong to short cycles (i.e., cycles of length $\leq \ell_0 \log n$), G_j is the subgraph $G \backslash \{e_1, \ldots, e_j\}$, and φ_{G_j} is the random-cluster distribution on G_j with parameters p, q. Decreasing the value of δ does not affect the validity of item 1, and hence we will assume that $\delta \in (0, 1)$.

[2] An event $\mathcal{F} \subseteq 2^E$ is increasing if, for all $T \subseteq T' \subseteq E$, $T \in \mathcal{F}$ implies that $T' \in \mathcal{F}$ as well. Note also that S' open means that every edge in S' is open, and S' closed means that every edge in S' is closed.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

For $j \in [t]$, consider the edge $e_j = \{u_j, v_j\}$ and let $P_{\ell,j}$ denote the number of paths with ℓ vertices in G whose endpoints are u_j and v_j. Using the fact that G satisfies item 2, we will show shortly that, for all $j \in [t]$, it holds that

$$\varphi_{G_j}(u_j \leftrightarrow v_j) \leq \frac{1}{n^3} + \sum_{\ell=L_0}^{L_2} \frac{P_{\ell,j}}{W^\ell}. \tag{5.8}$$

Let us assume (5.8) for now and conclude the proof of (5.7). Summing (5.8) over $j \in [t]$ (and using the trivial bound $t \leq |E| \leq \Delta n/2$), we obtain that

$$\sum_{j=1}^{t} \varphi_{G_j}(u_j \leftrightarrow v_j) \leq \frac{\Delta}{n^2} + \sum_{\ell=L_0}^{L_2} \frac{\sum_{j=1}^{t} P_{\ell,j}}{W^\ell} \leq \frac{\Delta}{n^2} + \sum_{\ell=L_0}^{L_2} \frac{\ell C_\ell}{W^\ell}, \tag{5.9}$$

where in the last inequality we used that $\sum_{j=1}^{t} P_{\ell,j} \leq \ell C_\ell$, which follows from the observation that every path with ℓ vertices connecting the endpoints of an edge $\{u_j, v_j\}$ maps to a cycle of length ℓ (by adding the edge $\{u_j, v_j\}$) and each cycle of length ℓ can potentially arise at most ℓ times under this mapping. Using (5.9) and the fact that G satisfies item 1, we obtain (5.7), as wanted.

To finish the proof, it only remains to prove (5.8). Since G satisfies item 2 and $W = 1/K$, by Lemma 5.2, we have that for any path P in G with ℓ vertices and $\ell \in [L_0, L_2 + 1]$, it holds that

$$\varphi_G(P \text{ is open}) \leq 1/W^\ell. \tag{5.10}$$

Since G_j is a subgraph of G, any path in G_j is also present in G. Moreover, we have that φ_{G_j} is obtained by conditioning some edges of G to be closed (namely, $e_1, \ldots, e_j$). Therefore, by Lemma 5.3, we conclude from (5.10) that, for any path P in G_j with ℓ vertices and $\ell \in [L_0, L_2 + 1]$, it holds that

$$\varphi_{G_j}(\text{path } P \text{ is open}) \leq 1/W^\ell. \tag{5.11}$$

Since the edge $\{u_j, v_j\}$ does not belong to a short cycle, we have that any path P in G_j connecting u_j and v_j has length at least $\ell_0 \log n$. We can therefore bound the probability of an open path between u_j and v_j by

$$\varphi_{G_j}(u_j \leftrightarrow v_j) \leq \varphi_{G_j}(\mathcal{E}_j) + \varphi_{G_j}(\mathcal{F}_j), \tag{5.12}$$

where $\mathcal{E}_j$ is the event that there exists an open path P with ℓ vertices with $L_0 \leq \ell \leq L_2$ connecting u_j and v_j, whereas $\mathcal{F}_j$ is the event that there exists an open path P with $\ell = L_2 + 1$ vertices starting from u_j (the other endpoint can be v_j or any other vertex of the graph). Using (5.11), we have by a union bound over paths that

$$\varphi_{G_j}(\mathcal{E}_j) \leq \sum_{\ell=L_0}^{L_2} \frac{P_{\ell,j}}{W^\ell}, \quad \varphi_{G_j}(\mathcal{F}_j) \leq \frac{\Delta(\Delta-1)^{\ell_2 \log n}}{W^{\ell_2 \log n}} \leq 1/n^3, \tag{5.13}$$

where in the bound for $\varphi_{G_j}(\mathcal{E}_j)$ we used that there are $P_{\ell,j}$ paths with ℓ vertices connecting u_j and v_j, while in the bound for $\varphi_{G_j}(\mathcal{F}_j)$ we used that there are at most $\Delta(\Delta-1)^{L_2-1}$ paths in G with $L_2 + 1$ vertices starting with u_j (since G has max degree Δ), the trivial inequalities $\ell_2 \log n - 1 \leq L_2 \leq \ell_2 \log n$, and the choice of ℓ_2 which guarantees that $\ell_2 \geq 4 \log(W/(\Delta - 1))$.

Combining (5.12) and (5.13) yields (5.8) (since $\delta \in (0, 1)$), thus completing the proof of Lemma 5.4. ∎

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

5.3. Combining the pieces—Proof of Theorem 2.5. We are now able to prove the following theorem, which details the performance of the algorithm SampleRC on random Δ-regular graphs and yields as an immediate corollary Theorem 2.5.

THEOREM 5.5. *Let $\Delta \geq 3$, $q \geq 1$, and $p < p_c(q, \Delta)$. Then, there exists a constant $\delta > 0$ such that, as $n \to \infty$, the following holds with probability $1 - o(1)$ over the choice of a random Δ-regular graph $G = (V, E)$ with n vertices.*

The output of algorithm SampleRC(G) *(cf. Figure 1) is a set $S \subseteq E$ whose distribution ν_S is within total variation distance $O(1/n^\delta)$ from the random-cluster distribution φ_G with parameters p, q, i.e.,*

$$\|\nu_S - \varphi_G\|_{\mathrm{TV}} = O(1/n^\delta).$$

Proof. By Lemmas 4.1 and 5.4, we have by a union bound that a uniformly random Δ-regular graph G with n vertices satisfies the following with probability $1 - o(1)$:

1. any two distinct cycles of length $\leq \frac{1}{5} \log_{\Delta-1} n$ are disjoint,
2. $\sum_{j=1}^{t} \varphi_{G_j}(u_j \leftrightarrow v_j) \leq 1/n^\delta$, where $e_1 = \{u_1, v_1\}, \ldots, e_t = \{u_t, v_t\}$ are the edges of G that do *not* belong to short cycles, and G_j is the subgraph $G \backslash \{e_1, \ldots, e_j\}$ (for $j \in [t]$).

We will show that for any graph $G = (V, E)$ that satisfies items 1 and 2, the output of the algorithm SampleRC is a random set $S \subseteq E$ whose distribution ν_S is within total variation distance $1/n^\delta$ from the random-cluster distribution φ_G with parameters p, q, therefore proving the result.

Let $G_t = (V, E_t), \ldots, G_1 = (V, E_1)$ be the sequence of subgraphs as in item 2 above and, for convenience, set $E_0 = E$ and let $G_0 = (V, E_0)$ (so that $G_0 = G$). Note that $G_t = G'$ (where G' is the graph considered in the algorithm SampleRC), i.e., G_t is the subgraph of G where only the edges that belong to short cycles appear. By item 1, we have that G_t consists of isolated vertices and disjoint cycles and hence we can conclude that the output of the algorithm SampleRC is not Fail, i.e., on input a graph G satifying items 1 and 2, SampleRC outputs a random set $S \subseteq E$. It therefore remains to show that the distribution ν_S of S satisfies

$$\|\nu_S - \varphi_G\|_{\mathrm{TV}} = O(1/n^\delta). \tag{5.14}$$

For $j = t, t-1, \ldots, 0$, let $S_j = S \cap E_j$ and let ν_{S_j} denote the distribution of S_j. Note that $S_0 = S$ and $S_t = S'$ (where S' is the subset of edges considered in the algorithm SampleRC). We have that

$$\|\nu_S - \varphi_G\|_{\mathrm{TV}} = \|\nu_{S_0} - \varphi_{G_0}\|_{\mathrm{TV}}, \quad \|\nu_{S_t} - \varphi_{G_t}\|_{\mathrm{TV}} = \|\nu_{S'} - \varphi_{G'}\|_{\mathrm{TV}} = 0. \tag{5.15}$$

For $j \in [t]$, we have that S_{j-1} is obtained from S_j by adding the edge e_j with probability $p/(q + (1-p)q)$. Let $\hat{S}_{j-1} \subseteq E_{j-1}$ be a subset of edges obtained by sampling an random-cluster configuration from G_j (according to φ_{G_j}) and adding the edge e_j with probability $p/(q + (1-p)q)$; denote by $\nu_{\hat{S}_{j-1}}$ the distribution of $\hat{S}_{j-1}$. By Lemma 5.1 we have that[3]

$$\left\|\nu_{\hat{S}_{j-1}} - \varphi_{G_{j-1}}\right\|_{\mathrm{TV}} \leq 2q\, \varphi_{G_j}(u_j \leftrightarrow v_j).$$

[3]Note that, to apply Lemma 5.1, we need to ensure that $\varphi_{G_j}(u_j \leftrightarrow v_j) < 1/q$, which is guaranteed by item 2.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Moreover, since in each of S_{j-1} and $\hat{S}_{j-1}$ the edge e_j appears independently with the same probability $p/(q+(1-p)q)$, we have that

$$\left\|\nu_{S_{j-1}} - \nu_{\hat{S}_{j-1}}\right\|_{\mathrm{TV}} = \left\|\nu_{S_j} - \varphi_{G_j}\right\|_{\mathrm{TV}}.$$

Using the triangle inequality and induction, we obtain that for all $j = 0, 1, \ldots, t$ it holds that

$$\left\|\nu_{S_j} - \varphi_{G_j}\right\|_{\mathrm{TV}} \leq \|\nu_{S_t} - \varphi_{G_t}\|_{\mathrm{TV}} + 2q \sum_{j'=j}^{t-1} \varphi_{G_{j'}}(u_{j'} \leftrightarrow v_{j'}).$$

Writing this out for $j = 0$ and using (5.15), we obtain that

$$\|\nu_S - \varphi_G\|_{\mathrm{TV}} \leq 2q \sum_{j=0}^{t-1} \varphi_{G_j}(u_j \leftrightarrow v_j) \leq 2q/n^{\delta},$$

where in the last inequality we used that the graph G satisfies item 2. This finishes the proof of (5.14) and therefore the proof of Theorem 5.5 as well. ∎

6. Algorithm for the antiferromagnetic Potts model. In this section, we give the details of our sampling algorithm for the antiferromagnetic Potts model (outlined in section 3). The section is organized as follows. First, in section 6.1, we formalize the connection between the Potts model on bichromatic classes and the Ising model. Then, in section 6.2, we state the sampling algorithm for the Ising model on graphs with small average growth that we are going to use for resampling bichromatic classes in the Potts model; moreover, we state certain correlation decay properties for the Ising model that will be relevant for analyzing the error of our Potts sampler. In section 6.3, we state the key lemma that allows us to bound the average growth of bichromatic classes in the Potts model on random regular graphs. In section 6.4, we show an "idealized" subroutine that updates a Potts configuration when we add a new edge $\{u, v\}$; the subroutine works by resampling an appropriately chosen bichromatic class and it is "idealized" in the sense that it assumes that certain steps can be carried out efficiently. In section 6.5, we modify the subroutine to make it computationally efficient by considering the average growth of bichromatic classes that get resampled; there, we give the complete description of the actual resampling subroutine used in our Potts sampler. With these pieces in place, we are in position to complete the description and analysis of the Potts sampler in section 6.6.

6.1. Connection between Potts on bichromatic classes and the Ising model. In this section, we describe the connection between the Potts model on bichromatic classes and the Ising model. Recall that the Ising model is the special case $q = 2$ of the Potts model; to distinguish between the models, we will use π_G to denote the Ising distribution on G with parameter B. Sometimes, we will need to replace the binary set of states $\{1, 2\}$ in the Ising model by other binary sets to facilitate the arguments; we use $\pi_G^{c_1,c_2}$ to denote the Ising distribution with binary set of states $\{c_1, c_2\}$ (we will have that $c_1, c_2 \in [q]$).

For a configuration $\sigma : V \to [q]$, we will denote by σ_U the restriction of σ to the set U. Our sampling algorithm is based on the following simple observation.

OBSERVATION 6.1. *Let $q \geq 3$ and $B > 0$. Let $G = (V, E)$ be a graph, U be a subset of V, and c, c' be distinct colors in $[q]$. Then, for any configuration $\eta : U \to \{c, c'\}$, it holds that*

$$\mu_G\big(\sigma_U = \eta \mid \sigma^{-1}(c, c') = U\big) = \pi_{G[U]}^{c,c'}(\eta),$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

i.e., conditioned on U being the (c, c')-color-class in the Potts distribution μ_G, the marginal distribution on U is the Ising distribution $\pi_{G[U]}$ (with set of states $\{c, c'\}$).

The following definition will be notationally convenient.

DEFINITION 6.2. *Let G be a graph, u, v be vertices in G, and c, c' be distinct colors in $[q]$. We write $\pi^{c,c'}_{G,u,v}$ to denote the Ising distribution on G (with set of states $\{c, c'\}$) conditioned on u taking the state c and v the state c'.*

6.2. Sampling the Ising model on graphs with small average growth. Recall from section 3 that our algorithm for sampling the antiferromagnetic Potts model with parameter B will use as a subroutine a sampling algorithm for the Ising model with parameter B to recolor bichromatic classes. In general, these classes may consist of large bichromatic components (with a linear number of vertices), so to carry out this subroutine efficiently, we need to use an approximate sampling algorithm; our leverage point will be that we can bound the average growth of bichromatic components (cf. Definition 3.3). Adapting results of [27], we show the following in section 9.2.

THEOREM 6.3. *Let $B \in (0, 1)$ and $b > 0$ be constants such that $b\frac{1-B}{1+B} < 1$, and let $\Delta \geq 3$ be an integer. Then, there exists $M_0 > 0$ such that the following holds for all $M > M_0$.*

There is a polynomial-time algorithm that, on input an n-vertex graph G with maximum degree at most Δ and average growth b up to depth $L = \lceil M \log n \rceil$, outputs a configuration $\tau : V \to \{1, 2\}$ whose distribution ν_τ is within total variation distance $1/n^{10}$ from the Ising distribution on G with parameter B, i.e.,

$$\|\nu_\tau - \pi_G\|_{\mathrm{TV}} \leq 1/n^{10}.$$

Moreover, the algorithm, when given as additional input two vertices u and v in G, outputs a configuration $\tau : V \to \{1, 2\}$ such that $\tau_u = 1$ and $\tau_v = 2$, and whose distribution ν_τ satisfies

$$\|\nu_\tau - \pi^{1,2}_{G,u,v}(\cdot)\|_{\mathrm{TV}} \leq 1/n^{10},$$

where $\pi^{1,2}_{G,u,v}$ is the Ising distribution on G conditioned on u having state 1 and v having state 2.

In addition, we will use the following spatial mixing result to analyze the accuracy of our algorithm for the antiferromagnetic Potts model. The proof is given in section 9.3.

LEMMA 6.4. *Let $B \in (0, 1)$ and $b > 0$ be constants such that $b\frac{1-B}{1+B} < 1$. Then, there exists $M_0' > 0$ such that the following holds for all $M > M_0'$.*

Let G be an n-vertex graph with average growth b up to depth $L = \lceil M \log n \rceil$, and let u, v be distinct vertices in G. Then

$$\left|\pi_G(\sigma_u = 1 \mid \sigma_v = 1) - \pi_G(\sigma_u = 1 \mid \sigma_v = 2)\right| \leq \frac{1}{n^{10}} + \sum_{\ell=1}^{L} P_\ell(G, u, v) \left(\frac{1-B}{1+B}\right)^\ell,$$

where $P_\ell(G, u, v)$ is the number of paths with ℓ vertices in G that connect u and v.

We will also need the following crude bound.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

LEMMA 6.5. *Let $B \in (0,1)$ and $\Delta \geq 3$ be an integer. Suppose that G is a graph of maximum degree at most Δ and let u be a vertex and Λ be a set of vertices in G such that $u \notin \Lambda$. Then, for every configuration $\tau : \Lambda \to \{1,2\}$ and $s \in \{1,2\}$*

$$\frac{B^{\Delta}}{1+B^{\Delta}} \leq \pi_G(\sigma_u = s \mid \sigma_\Lambda = \tau).$$

Proof. Without loss of generality, we assume that $s = 1$. Let D be the number of neighbors of u and $u_1, \ldots, u_d$ be the neighbors of u in $G \setminus \Lambda$ and note that $d \leq D \leq \Delta$. Let d_0 be the number of $v \in \Lambda$ such that v is u's neighbor and $\tau_v = 1$. Let $s_1, \ldots, s_d \in \{1,2\}$ be arbitrary and let d_1 be the number of the s_i's that are equal to 1. Then we have that

$$\pi_G(\sigma_u = 1 \mid \sigma_{u_1} = s_1, \ldots, \sigma_{u_d} = s_d, \sigma_\Lambda = \tau) = \frac{B^{d_2}}{B^{d_2} + B^{D-d_2}},$$

where $d_2 = d_0 + d_1 \leq d_0 + d \leq D$. Since $u_1, \ldots, u_d \notin \Lambda$ we have that $\pi_G(\sigma_{u_1} = s_1, \ldots, \sigma_{u_d} = s_d \mid \sigma_\Lambda = \tau) \neq 0$ for any choice of $s_1, \ldots s_d$ and therefore by the law of total probability we obtain

$$\min_{d_2 \in \{d_0, \ldots, d_0+d\}} \left\{ \frac{B^{d_2}}{B^{d_2} + B^{D-d_2}} \right\} \leq \pi_G(\sigma_u = 1 \mid \sigma_\Lambda = \tau).$$

The function $\frac{B^x}{B^x + B^{D-x}}$ is decreasing for $x \in [0, D]$ (since $B \in (0,1)$) and hence

$$\min_{d_2 \in \{d_0, \ldots, d_0+d\}} \left\{ \frac{B^{d_2}}{B^{d_2} + B^{D-d_2}} \right\} = \frac{B^D}{1+B^D}.$$

Using that $B \in (0,1)$ and $D \leq \Delta$, we obtain the inequalities in the statement of the lemma. ∎

6.3. Average growth of bichromatic components in the Potts distribution. To utilize Theorem 6.3 and Lemma 6.4 for our sampling algorithms, we will need to bound the average growth of bichromatic components in a typical Potts configuration on a random regular graph. Our key lemma to achieve this will bound the probability that a path is bichromatic[4] in uniqueness, provided that the local neighborhood around the path (in the sense of Definition 4.3) has a tree-like structure. The following lemma quantifies this probability bound and is proved in section 8.2. The proof uses the fact that the parameter B lies in the uniqueness regime of the $(\Delta-1)$-ary tree.

LEMMA 6.6. *Let $\Delta, q \geq 3$ be integers, and let $B \in (0,1)$ be in the uniqueness regime of the $(\Delta-1)$-ary tree with $B \neq (\Delta-q)/\Delta$. Then, for any $\epsilon' > 0$, there exists a positive constant $K < \frac{1+B}{B+q-1} + \epsilon'$ and $\epsilon > 0$ such that the following holds for all sufficiently large integers ℓ and h.*

Let G be a graph of maximum degree Δ and P be a path with ℓ vertices whose h-graph-neighborhood contains $(1-\epsilon)\ell$ isolated tree components. Let μ_G be the Potts measure on G with parameter B. Then,

$$\mu_G(\textit{path } P \textit{ is bichromatic}) \leq K^{\ell}.$$

[4]Let $G = (V, E)$ be a graph and $\sigma : V \to [q]$. We call a path P bichromatic under σ if there exist colors $c_1, c_2 \in [q]$ such that every vertex u of P satisfies $\sigma_u \in \{c_1, c_2\}$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

For a random Δ-regular graph G, paths do have the tree-like structure of Lemma 6.6 (cf. Lemma 4.4), and hence we can aggregate over all paths emanating from an arbitrary vertex (roughly $(\Delta-1)^\ell$ of them) and get a bound of roughly $(\Delta-1)K < \frac{(\Delta-1)(1+B)}{B+q-1}$ for the average growth of bichromatic components in a typical configuration σ. This will allow us to use the upcoming ReSample subroutine for updating a bichromatic class using the Ising sampler of section 6.2.

6.4. Analyzing an ideal ReSample subroutine. In this section, we give a preliminary description and analysis of the ReSample subroutine that updates a Potts configuration when we add a new edge $\{u, v\}$. The ReSample subroutine is inspired by the approach of Efthymiou [8] for colorings.

In fact, for the moment, we will only study an "idealized" version of ReSample which we call IdealReSample; the subroutine is idealized in the sense that it assumes that certain steps can be carried out efficiently. Later, we will modify the subroutine to obtain the actual ReSample subroutine whose running time will be polynomial with respect to the size of the input graph.

The point of analyzing first IdealReSample is to give the key ideas behind the underlying resampling step without bothering for the moment to make the subroutine computationally efficient. Moreover, the detour is going to be smaller than it might appear since the analysis of the actual ReSample subroutine will follow from the analysis of IdealReSample.

The IdealReSample subroutine takes as inputs a graph G, two vertices u and v of G, and a configuration σ on G such that $\sigma_u = \sigma_v$; it outputs a configuration σ' on G by updating the configuration on an appropriately chosen bichromatic class containing the vertices u and v. The details of the subroutine can be found in Figure 2.

Algorithm IdealReSample(G, u, v, σ)

parameters: real $B \in (0, 1)$, integer $q \geq 3$

Input: Graph $G = (V, E)$, vertices $u, v \in V$ with $\{u, v\} \notin E$, configuration $\sigma : V \to [q]$ with $\sigma_u = \sigma_v$.

Output: A configuration $\sigma' : V \to [q]$.

Flip a coin with heads probability $\frac{qB}{B+q-1}$.

if heads **then** $\sigma' = \sigma$

else

Pick u.a.r. a colour c' from $[q]/\{c\}$, where $c = \sigma_u = \sigma_v$.

Let $U = \sigma^{-1}(c, c')$ and set $H = G[U]$.

Sample Ising configuration τ on H conditioned on $\tau_u = c$ and $\tau_v = c'$, i.e.:

Sample $\tau : U \to \{c, c'\}$ with $\tau_u = c$ and $\tau_v = c'$ so that $\tau \sim \pi^{c,c'}_{H,u,v}$

Set: $\sigma'_w = \tau_w$ for $w \in U$; set $\sigma'_w = \sigma_w$ for $w \notin U$.

return $\sigma' : V \to [q]$.

Fig. 2. *The* IdealReSample *subroutine; we will later modify this to obtain the actual* ReSample *subroutine used in Algorithm* SampleAntiPotts *(cf. Figure* 4*).*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The subroutine will be used to update a Potts configuration when we add a new edge $\{u, v\}$; see the upcoming Lemma 6.9.

To control the output distribution of the IdealReSample subroutine, the following definition will be crucial.

Definition 6.7. *Let $B > 0$. Suppose that $G = (V, E)$ is a graph and that u, v are vertices in G. For a set $U \subseteq V$ such that $u, v \in U$, let*

$$\mathsf{Corr}_G(U, u, v) = \left| \frac{\pi_{G[U]}\big(\eta_u = 1, \eta_v = 1\big)}{\pi_{G[U]}\big(\eta_u = 1, \eta_v = 2\big)} - 1 \right|,$$

where η denotes a configuration $U \to \{1, 2\}$ sampled according to $\pi_{G[U]}$. Note that $\mathsf{Corr}_G(U, u, v)$ measures the correlation between u and v in the Ising distribution with parameter B on the subgraph $G[U]$.

To state the main lemma of this section, we will also need the following definition of a "random bichromatic class" containing two specific vertices u and v under a configuration σ.

Definition 6.8. *Let $G = (V, E)$ be a graph, u, v be vertices in G, and $\sigma : V \to [q]$ be a configuration on G. We let $U_\sigma \subseteq V$ be a bichromatic class under σ which contains u and v, chosen uniformly at random among the set of all such classes if there is more than one.*

More precisely, if $\sigma_u \neq \sigma_v$, then U_σ is the (c_1, c_2)-color-class in σ where c_1, c_2 are the colors of u and v under σ. If $\sigma_u = \sigma_v$, then U_σ is the (c, c')-color-class in σ, where c is the common color of u and v under σ and c' is a uniformly random color from $[q] \backslash \{c\}$.

The following lemma will be critical for our Potts sampler. It shows how to update a Potts configuration when we add a new edge $\{u, v\}$, based on the IdealReSample subroutine. It also controls the error introduced based on the "average correlation" between u and v in a random bichromatic class that contains them.

Lemma 6.9. *Let $B \in (0, 1)$ and $q \geq 3$ be an integer, and consider arbitrary $\epsilon \in (0, 1)$.*

Let $G = (V, E)$ be a graph and μ_G be the Potts distribution on G with parameter B. Suppose that u, v are vertices in G such that $\{u, v\} \notin E$ and

$$\mathbf{E}\big[\mathsf{Corr}_G(U_\sigma, u, v)\big] \leq \epsilon,$$

where the expectation is over the choice of a random configuration $\sigma \sim \mu_G$ and the choice of a random bichromatic class $U_\sigma \subseteq V$ containing u and v under σ (cf. Definition 6.8).

Consider the graph $G' = (V, E')$ obtained from G by adding the edge $\{u, v\}$. Sample a configuration $\sigma' : V \to [q]$ as follows. First, sample $\sigma : V \to [q]$ according to μ_G. Then, if $\sigma_u \neq \sigma_v$, set $\sigma' = \sigma$; otherwise, set $\sigma' = \textsc{IdealReSample}(G, u, v, \sigma)$. Then, the distribution of σ', denoted by $\nu_{\sigma'}$, is within total variation distance $2\epsilon/B$ from the Potts distribution $\mu_{G'}$ on G' with parameter B, i.e.,

$$\|\nu_{\sigma'} - \mu_{G'}\|_{\mathrm{TV}} \leq 2\epsilon/B.$$

Proof. We begin with a few definitions that will be used throughout the proof. Fix distinct colors $c_1, c_2 \in [q]$ and a set $U \subseteq V$ such that $u, v \in U$. Let $\Omega(U, c_1, c_2)$ be the set of configurations such that U is the (c_1, c_2)-color-class, i.e.,

$$\Omega(U, c_1, c_2) = \big\{\eta : V \to [q] \mid U = \eta^{-1}(c_1, c_2)\big\}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

We will be interested in two particular types of configurations in $\Omega(U, c_1, c_2)$, those where u, v take the colors c_1, c_2 and those where u, v take the colors c_1, c_1. Namely, let

$$\Omega_{\mathsf{neq}}(U, c_1, c_2) = \big\{\eta \in \Omega(U, c_1, c_2) \mid \eta_u = c_1,\, \eta_v = c_2\big\},$$

$$\Omega_{\mathsf{eq}}(U, c_1, c_2) = \big\{\eta \in \Omega(U, c_1, c_2) \mid \eta_u = c_1,\, \eta_v = c_1\big\}.$$

Note the asymmetry in the above definitions with respect to c_1, c_2, e.g., for $\eta \in \Omega_{\mathsf{eq}}(U, c_1, c_2)$, we have that $\eta_u, \eta_v \neq c_2$. For a configuration $\eta \in \Omega(U, c_1, c_2)$, we also denote by

$$\Omega^{\eta}(U, c_1, c_2) = \big\{\tau \in \Omega(U, c_1, c_2) \mid \tau_{V\setminus U} = \eta_{V\setminus U}\big\}$$

the set of configurations in $\Omega(U, c_1, c_2)$ that agree with η on $V\setminus U$; we define analogously the sets $\Omega^{\eta}_{\mathsf{neq}}(U, c_1, c_2), \Omega^{\eta}_{\mathsf{eq}}(U, c_1, c_2)$. Using Observation 6.1 and the Ising distribution $\pi^{c_1,c_2}_{G[U]}$ (cf. section 6.1), we have that for any $\eta \in \Omega(U, c_1, c_2)$ it holds that

$$\mu_G(\eta) = \mu_G\big(\Omega^{\eta}(U, c_1, c_2)\big)\, \pi^{c_1,c_2}_{G[U]}(\eta_U) \tag{6.1}$$

and

$$\begin{aligned} \frac{\mu_G\big(\Omega^{\eta}_{\mathsf{eq}}(U, c_1, c_2)\big)}{\mu_G\big(\Omega^{\eta}(U, c_1, c_2)\big)} &= \pi^{c_1,c_2}_{G[U]}(\tau_u = \tau_v = c_1), \\ \frac{\mu_G\big(\Omega^{\eta}_{\mathsf{neq}}(U, c_1, c_2)\big)}{\mu_G\big(\Omega^{\eta}(U, c_1, c_2)\big)} &= \pi^{c_1,c_2}_{G[U]}(\tau_u = c_1, \tau_v = c_2). \end{aligned} \tag{6.2}$$

Also, the assumption $\mathbf{E}\big[\mathsf{Corr}_G(U_\sigma, u, v)\big] \leq \epsilon$ translates into

$$\sum_{\substack{c_1,c_2\in[q];\\ c_1\neq c_2}} \sum_{\substack{U\subseteq V;\\ u,v\in U}} \Big(\mu_G\big(\Omega_{\mathsf{neq}}(U, c_1, c_2)\big) + \frac{1}{q-1}\mu_G\big(\Omega_{\mathsf{eq}}(U, c_1, c_2)\big)\Big)\mathsf{Corr}_G(U, u, v) \leq \epsilon. \tag{6.3}$$

To see this, for a set $U \subseteq V$ such that $u, v \in U$, we find how much $\mathsf{Corr}_G(U, u, v)$ contributes to $\mathbf{E}\big[\mathsf{Corr}_G(U_\sigma, u, v)\big]$. Note that for a configuration σ to have $U_\sigma = U$ (cf. Definition 6.8), there must exist distinct colors $c_1, c_2 \in [q]$ such that $U = \sigma^{-1}(c_1, c_2)$ and either (i) $\sigma(u) = c_1$, $\sigma(v) = c_2$, or (ii) $\sigma(u) = c_1$, $\sigma(v) = c_1$. In case (i), we have that $\sigma \in \Omega_{\mathsf{neq}}(U, c_1, c_2)$ and $U_\sigma = U$ with probability 1. In case (ii), we have that $\sigma \in \Omega_{\mathsf{eq}}(U, c_1, c_2)$ and $U_\sigma = U$ with probability $1/(q-1)$. By aggregating over the relevant σ, we therefore obtain (6.3).

We next proceed to the proof. Recalling that $\nu_{\sigma'}$ is the distribution of σ', our goal is to show that

$$\|\nu_{\sigma'} - \mu_{G'}\|_{\mathrm{TV}} \leq 2\epsilon/B. \tag{6.4}$$

Weight of configurations in $\boldsymbol{\nu_{\sigma'}}$. For distinct colors $c_1, c_2 \in [q]$ and a set $U \subseteq V$ such that $u, v \in U$, we first show that

$$\forall \eta \in \Omega_{\mathsf{eq}}(U, c_1, c_2): \quad \nu_{\sigma'}(\eta) = \frac{qB}{B+q-1}\mu_G(\eta), \tag{6.5}$$

$$\forall \eta \in \Omega_{\mathsf{neq}}(U, c_1, c_2): \quad \nu_{\sigma'}(\eta) = \mu_G(\eta) + \frac{(1-B)\pi^{c_1,c_2}_{G[U],u,v}(\eta_U)}{B+q-1}\mu_G\big(\Omega^{\eta}_{\mathsf{eq}}(U, c_1, c_2)\big). \tag{6.6}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

(Recall from Definition 6.2 that $\pi^{c_1,c_2}_{G[U],u,v}$ is the conditional Ising distribution on $G[U]$ where u,v take the colors c_1, c_2, respectively.) To see the expression for $\nu_{\sigma'}(\eta)$ in (6.5), note that we can obtain $\eta \in \Omega_{\mathsf{eq}}(U, c_1, c_2)$ via the subroutine IdealReSample only if $\sigma = \eta$ and the coin flip came up heads (because $\eta_u = \eta_v = c_1$). Analogously, to see the expression for $\nu_{\sigma'}(\eta)$ in (6.6), note that we can obtain $\eta \in \Omega_{\mathsf{neq}}(U, c_1, c_2)$ if one of the following happens (using that $\eta_u = c_1$ and $\eta_v = c_2$):

- We started with the configuration $\sigma = \eta$; this happens with probability $\mu_G(\eta)$. Then, we obtain η with probability 1.
- We started with a configuration σ such that $\sigma_u = \sigma_v = c_1$, $\sigma^{-1}(c_1, c_2) = U$, and $\sigma_{V\setminus U} = \eta_{V\setminus U}$, i.e., $\sigma \in \Omega^{\eta}_{\mathsf{eq}}(U, c_1, c_2)$; this happens with probability $\mu_G\big(\Omega^{\eta}_{\mathsf{eq}}(U, c_1, c_2)\big)$. Then the coin flip came tails, the color c' selected was $c' = c_2$, and the Ising configuration τ sampled was η_U. All that happens with probability $\frac{(1-B)\pi^{c_1,c_2}_{G[U],u,v}(\eta_U)}{B+q-1}$.

Using the assumption $\mathbf{E}\big[\mathsf{Corr}_G(U_\sigma, u, v)\big] \le \epsilon$ (cf. (6.3)), we will later show the following bound for the ratio of the partition functions $Z_G, Z_{G'}$:

$$|M| \le \epsilon/B, \text{ where } M := \frac{(B+q-1)Z_G}{qZ_{G'}} - 1. \tag{6.7}$$

Let us conclude the proof assuming, for now, (6.7). Then, we will prove (6.7) in Step II below.

Step I: Proof of (6.4) (assuming (6.7)). We will decompose $\|\nu_{\sigma'} - \mu_{G'}\|_{\mathrm{TV}}$ as

$$\|\nu_{\sigma'} - \mu_{G'}\|_{\mathrm{TV}} = \frac{1}{2}\sum_{\eta: V\to[q]} |\nu_{\sigma'}(\eta) - \mu_{G'}(\eta)| = \frac{1}{2}\sum_{U;\, u,v\in U}\ \sum_{c_1,c_2\in[q];\, c_1\neq c_2} D(U, c_1, c_2), \tag{6.8}$$

where

$$D(U, c_1, c_2) := \sum_{\eta\in\Omega_{\mathsf{neq}}(U,c_1,c_2)} |\nu_{\sigma'}(\eta) - \mu_{G'}(\eta)| + \frac{1}{q-1}\sum_{\eta\in\Omega_{\mathsf{eq}}(U,c_1,c_2)} |\nu_{\sigma'}(\eta) - \mu_{G'}(\eta)|. \tag{6.9}$$

To see the second equality in (6.8), fix any configuration $\eta : V \to [q]$. If $\eta_u = c_1$ and $\eta_v = c_2$ for distinct colors $c_1, c_2 \in [q]$, then $|\nu_{\sigma'}(\eta) - \mu_{G'}(\eta)|$ appears only in one term $D(U, c_1, c_2)$ in the summation of (6.8), namely for $U = \eta^{-1}(c_1, c_2)$. If $\eta_u = \eta_v = c_1$ for some color $c_1 \in [q]$, then $|\nu_{\sigma'}(\eta) - \mu_{G'}(\eta)|$ appears in exactly $q-1$ terms $D(U, c_1, c_2)$ in the summation of (6.8), once for each color $c_2 \neq c_1$ and $U = \eta^{-1}(c_1, c_2)$.

For $\eta \in \Omega_{\mathsf{eq}}(U, c_1, c_2)$, we have $\mu_{G'}(\eta) = \frac{w_{G'}(\eta)}{Z_{G'}} = \frac{Bw_G(\eta)}{Z_{G'}} = \frac{qB(M+1)\mu_G(\eta)}{B+q-1}$, where the last equality follows from the definition of M in (6.7). Hence, by (6.5),

$$|\nu_{\sigma'}(\eta) - \mu_{G'}(\eta)| = \frac{qB}{B+q-1}\Big|\mu_G(\eta) - (M+1)\mu_G(\eta)\Big| \le |M|\,\mu_G(\eta), \tag{6.10}$$

where the last inequality follows from $\frac{qB}{B+q-1} \le 1$ which holds for all $B \in (0, 1)$.

For $\eta \in \Omega_{\mathsf{neq}}(U, c_1, c_2)$, we have $\mu_{G'}(\eta) = \frac{w_{G'}(\eta)}{Z_{G'}} = \frac{w_G(\eta)}{Z_{G'}} = \frac{q(M+1)}{B+q-1}\mu_G(\eta)$ using again the definition of M in (6.7). Hence, by (6.6),

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$
\begin{aligned}
|\nu_{\sigma'}(\eta) - \mu_{G'}(\eta)| &= \left| \frac{1-B}{B+q-1}\Big(\mu_G(\eta) - \pi^{c_1,c_2}_{G[U],u,v}(\eta_U)\mu_G\big(\Omega^{\eta}_{\mathsf{eq}}(U,c_1,c_2)\big)\Big) \right. \\
&\qquad \left. + \frac{qM}{B+q-1}\mu_G(\eta) \right| \\
&\le \Big|\mu_G(\eta) - \pi^{c_1,c_2}_{G[U],u,v}(\eta_U)\mu_G\big(\Omega^{\eta}_{\mathsf{eq}}(U,c_1,c_2)\big)\Big| + 2|M|\,\mu_G(\eta), \qquad (6.11)
\end{aligned}
$$

where the last inequality follows from the triangle inequality and the inequalities $\frac{1-B}{B+q-1} \le 1$, $\frac{q}{B+q-1} \le 2$. Note that, by the definition of $\pi^{c_1,c_2}_{G[U],u,v}(\cdot)$, for $\eta \in \Omega_{\mathsf{neq}}(U,c_1,c_2)$ it holds that

$$
\pi^{c_1,c_2}_{G[U],u,v}(\eta_U) = \frac{\pi^{c_1,c_2}_{G[U]}(\eta_U)}{\pi^{c_1,c_2}_{G[U]}(\tau_u = c_1, \tau_v = c_2)}
$$

and therefore, using (6.1) and (6.2), we have that

$$
\Big|\mu_G(\eta) - \pi^{c_1,c_2}_{G[U],u,v}(\eta_U)\mu_G\big(\Omega^{\eta}_{\mathsf{eq}}(U,c_1,c_2)\big)\Big| = \mu_G(\eta)\mathsf{Corr}_G(U,u,v), \qquad (6.12)
$$

and hence (6.11) gives that, for all $\eta \in \Omega_{\mathsf{neq}}(U,c_1,c_2)$, it holds that

$$
|\nu_{\sigma'}(\eta) - \mu_{G'}(\eta)| \le \big(2|M| + \mathsf{Corr}_G(U,u,v)\big)\mu_G(\eta). \qquad (6.13)
$$

Summing (6.10) and (6.13) over the relevant configurations η, we obtain that, for all $U \subseteq V$ with $u,v \in U$ and distinct colors $c_1, c_2 \in [q]$, it holds that

$$
D(U,c_1,c_2) \le \big(2|M| + \mathsf{Corr}_G(U,u,v)\big)\left(\mu_G\big(\Omega_{\mathsf{neq}}(U,c_1,c_2)\big) + \frac{\mu_G\big(\Omega_{\mathsf{eq}}(U,c_1,c_2)\big)}{q-1}\right). \qquad (6.14)
$$

To conclude the proof of (6.4), note that analogously to (6.8) we have that

$$
\sum_{U;\, u,v\in U} \sum_{c_1,c_2\in[q];\, c_1\neq c_2} \left(\mu_G\big(\Omega_{\mathsf{neq}}(U,c_1,c_2)\big) + \frac{\mu_G\big(\Omega_{\mathsf{eq}}(U,c_1,c_2)\big)}{q-1}\right) = 1. \qquad (6.15)
$$

Now consider the expression for $\|\nu_{\sigma'} - \mu_{G'}\|_{\mathrm{TV}}$ from (6.8). We bound each term $D(U,c_1,c_2)$ using (6.14) and then apply (6.3) and (6.15) to obtain that

$$
\|\nu_{\sigma'} - \mu_{G'}\|_{\mathrm{TV}} \le \frac{1}{2}(2|M| + \epsilon) \le \tfrac{3}{2B}\epsilon \le 2\epsilon/B,
$$

where the last inequality follows from $|M| \le \epsilon/B$ (cf. (6.7)). This finishes the proof of (6.4), modulo the proof of (6.7), which is given below.

Step II: Proof of (6.7). To prove (6.7), it will be useful to define

$$
\begin{aligned}
X(U,c_1,c_2) &:= \sum_{\eta\in\Omega_{\mathsf{neq}}(U,c_1,c_2)} w_G(\eta), \\
Y(U,c_1,c_2) &:= \sum_{\eta\in\Omega_{\mathsf{eq}}(U,c_1,c_2)} w_G(\eta), \\
Z(U,c_1,c_2) &:= X(U,c_1,c_2) + \frac{1}{q-1}Y(U,c_1,c_2). \qquad (6.16)
\end{aligned}
$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Analogously to (6.8) (and (6.15)), we have that

$$Z_G = \sum_{c_1 \neq c_2} \sum_{U; u,v \in U} Z(U, c_1, c_2). \tag{6.17}$$

Using that $w_{G'}(\eta) = w_G(\eta)$ for $\eta \in \Omega_{\mathsf{neq}}(U, c_1, c_2)$ and $w_{G'}(\eta) = B w_G(\eta)$ for $\eta \in \Omega_{\mathsf{eq}}(U, c_1, c_2)$, we obtain that

$$Z_{G'} = \sum_{c_1 \neq c_2} \sum_{U; u,v \in U} Z'(U, c_1, c_2), \tag{6.18}$$

where $Z'(U, c_1, c_2) := X(U, c_1, c_2) + \frac{B}{q-1} Y(U, c_1, c_2)$.

We first show that, for all $U \subseteq V$ with $u, v \in U$ and distinct colors $c_1, c_2 \in [q]$, it holds that

$$\Big| \frac{(B+q-1) Z(U, c_1, c_2)}{q Z'(U, c_1, c_2)} - 1 \Big| \leq \frac{1}{q} \mathsf{Corr}_G(U, u, v). \tag{6.19}$$

Observe that

$$\frac{(B+q-1) Z(U, c_1, c_2)}{q Z'(U, c_1, c_2)} - 1 = \frac{1-B}{q} \frac{Y(U, c_1, c_2) - X(U, c_1, c_2)}{X(U, c_1, c_2) + \frac{B}{q-1} Y(U, c_1, c_2)}. \tag{6.20}$$

Observe also that (cf. (6.2)) $\big| \frac{Y(U,c_1,c_2)}{X(U,c_1,c_2)} - 1 \big| = \mathsf{Corr}_G(U, u, v)$, which combined with (6.20) (and ignoring the positive term $\frac{B}{q-1} Y(U, c_1, c_2)$ in the denominator of the r.h.s. in the latter) gives (6.19), using also that $B \in (0, 1)$.

Using (6.17) and (6.18), we have that

$$\frac{(B+q-1) Z_G}{q Z_{G'}} - 1 = \sum_{c_1 \neq c_2} \sum_{U;\, u,v \in U} \frac{q Z'(U, c_1, c_2)}{Z_{G'}} \Big(\frac{(B+q-1) Z(U, c_1, c_2)}{q Z'(U, c_1, c_2)} - 1 \Big). \tag{6.21}$$

Since $B \in (0, 1)$, we have for all $U \subseteq V$ with $u, v \in U$ and distinct colors $c_1, c_2 \in [q]$ that $Z'(U, c_1, c_2) \leq Z(U, c_1, c_2)$ and $Z_{G'} \geq B Z_G$; therefore

$$\frac{Z'(U, c_1, c_2)}{Z_{G'}} \leq \frac{Z(U, c_1, c_2)}{B Z_G} = \frac{1}{B} \Big(\mu_G\big(\Omega_{\mathsf{neq}}(U, c_1, c_2)\big) + \frac{\mu_G\big(\Omega_{\mathsf{eq}}(U, c_1, c_2)\big)}{q-1} \Big). \tag{6.22}$$

Combining (6.19), (6.21), and (6.22), we obtain by the triangle inequality that (cf. (6.3))

$$\Big| \frac{(B+q-1) Z_G}{q Z_{G'}} - 1 \Big| \leq \mathbf{E}\big[\mathsf{Corr}_G(U_\sigma, u, v)\big]/B \leq \epsilon/B,$$

thus completing the proof of (6.7) and therefore the proof of Lemma 6.9. ∎

6.5. The ReSample subroutine. In this section, we modify the subroutine IdealReSample of section 6.4 to make it computationally efficient; this will give us the actual ReSample subroutine that we will use in our sampling algorithm for the Potts model. To describe the ReSample subroutine, we will need the following definition.

Definition 6.10. *Let $b, M > 0$ be constants. Let $G = (V, E)$ be an n-vertex graph and let σ be a configuration on G. We say that a bichromatic component in σ is* (b, M)-good *if it has average growth b up to depth $L = \lceil M \log n \rceil$; we say that it is* (b, M)-bad *otherwise. Analogously, we say that σ is* (b, M)-good *if* all *bichromatic components in σ are good; otherwise, we say that σ is (b, M)-bad.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Algorithm ReSample$_{b,M}(G, u, v, \sigma)$

parameters: real $B \in (0,1)$, $b, M > 0$ satisfying Theorem 6.3, integer $q \geq 3$

Input: Graph $G = (V, E)$, vertices $u, v \in V$ with $\{u, v\} \notin E$, configuration $\sigma : V \to [q]$ with $\sigma(u) = \sigma(v)$.

Output: A configuration $\sigma' : V \to [q]$.

Does σ contain a bichromatic component which is (b, M)-bad?
if yes **then return** $\sigma' : V \to [q]$ selected uniformly at random.
else
Flip a coin with heads probability $\frac{qB}{B+q-1}$.
if heads **then return** $\sigma' = \sigma$
else
Pick u.a.r. a colour c' from $[q]/\{c\}$, where $c = \sigma(u) = \sigma(v)$.
Let $U = \sigma^{-1}(c, c')$ and set $H = G[U]$.
Use algorithm of Theorem 6.3 to sample Ising distribution on H, i.e.:
sample $\tau : U \to \{c, c'\}$ with $\tau(u) = c$ and $\tau(v) = c'$ s.t. $\|\nu_\tau - \pi^{c,c'}_{H,u,v}\| \leq n^{-10}$
Set: $\sigma'(w) = \tau(w)$ for $w \in U$; set $\sigma'(w) = \sigma(w)$ for $w \notin U$.
return $\sigma' : V \to [q]$.

FIG. 3. *The* ReSample *subroutine used in Algorithm* SampleAntiPotts *(given in Figure 4).*

We are now able to describe the ReSample subroutine, which takes as inputs a graph G, two vertices u and v of G, and a configuration σ on G such that $\sigma(u) = \sigma(v)$. The subroutine is given in Figure 3.

6.6. Analysis of the Potts algorithm. We now have all the pieces to give our algorithm for sampling from the antiferromagnetic Potts model; see the algorithm SampleAntiPotts in Figure 4. We next prove the following theorem, which details the performance of the algorithm SampleAntiPotts on random Δ-regular graphs and yields as an immediate corollary, Theorem 2.7.

THEOREM 6.11. *Let $\Delta \geq 3$, $q \geq 3$, and $B \in (0,1)$ be in the uniqueness regime of the $(\Delta-1)$-ary tree. Then, there exist constants $b, M, \delta > 0$ such that, as $n \to \infty$, the following holds with probability $1 - o(1)$ over the choice of a random Δ-regular graph $G = (V, E)$ with n vertices.*

The output of the algorithm SampleAntiPotts(G) *(cf. Figure 4) is an assignment $\sigma : V \to [q]$ whose distribution ν_σ is within total variation distance $O(1/n^\delta)$ from the Potts distribution μ_G with parameter B, i.e.,*

$$\|\nu_\sigma - \mu_G\|_{\mathrm{TV}} = O(1/n^\delta).$$

Proof. Since B is in the uniqueness regime of the $(\Delta-1)$-ary tree and $B \neq \frac{\Delta-q}{\Delta}$, we have that $B > \frac{\Delta-q}{\Delta}$ (cf. Remark 2.3). It follows that $\frac{1-B}{B+q-1} < \frac{1}{\Delta-1}$ and therefore there exists $\epsilon' > 0$ such that

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Algorithm SampleAntiPotts(G)

parameters: real $B \in (0,1)$, $b, M > 0$ satisfying Theorem 6.3, integers $q, \Delta \geq 3$

Input: Graph $G = (V, E)$
Output: Either Fail or an assignment $\sigma : V \to [q]$

$E' := \{e \in E \mid e \text{ belongs to a short cycle}\}$

if $G' = (V, E')$ has a component which is neither a cycle nor an isolated vertex
then return Fail
else
Sample a configuration $\sigma' : V \to [q]$ on G' (according to $\mu_{G'}$);
Let $e_1, e_2, \ldots, e_t$ be the edges in $E \backslash E'$; set $G_t = G'$ and $\sigma_t = \sigma'$.
for $j = t$ **downto** 1:
Suppose that $e_j = (u_j, v_j)$;
if $\sigma_j(u_j) = \sigma_j(v_j)$ **then** $\sigma_{j-1} = \text{ReSample}_{b,M}(G_j, u_j, v_j, \sigma_j)$ **else** $\sigma_{j-1} = \sigma_j$;
Obtain the graph G_{j-1} by adding the edge $\{u_j, v_j\}$ in G_j
end
return $\sigma = \sigma_0$.

Fig. 4. *Algorithm for sampling a Potts configuration in the antiferromagnetic case $B \in (0,1)$. The details of the* ReSample *subroutine are given in Figure* 3. *While the algorithm can also be modified to work for the ferromagnetic case $B > 1$, we will instead use a simpler percolation algorithm via the random-cluster representation.*

$$\left(\frac{1+B}{B+q-1} + 3\epsilon'\right)\frac{1-B}{1+B} < \frac{1}{\Delta - 1}. \tag{6.23}$$

Let $K < \frac{1+B}{B+q-1} + \epsilon'$ and $\epsilon > 0$ be the constants in Lemma 6.6 corresponding to ϵ', and let h', ℓ' be positive constants such that Lemma 6.6 applies for all integers $h \geq h'$ and $\ell \geq \ell'$. Fix h to be any integer greater than h'. Let $\ell_1 > 0$ be the constant in Lemma 4.4 corresponding to the values of ϵ and h. Let $\delta > 0$ be the constant in Lemma 4.2 corresponding to $\ell_0 := 1/(5\log(\Delta - 1))$ and $W := 1/(K\frac{1-B}{1+B})$ (note that (6.23) guarantees that $W > \Delta - 1$). Let also

$$b' := (\Delta - 1)\left(\frac{1+B}{B+q-1} + 2\epsilon'\right), \text{ and } b := (\Delta - 1)\left(\frac{1+B}{B+q-1} + 3\epsilon'\right).$$

Let M_0, M_0' be the constants in Theorem 6.3 and Lemma 6.4, respectively. Let M be sufficiently large so that $M > \max\{2\ell_0, 2M_0, 2M_0'\}$ and the following inequalities hold (for all sufficiently large n):

$$\Delta\big((\Delta-1)K\big)^{\lceil M \log n\rceil} \leq (b')^{\lceil M \log n\rceil} \text{ and } (b'/b)^{\lceil M \log n\rceil} \leq 1/n^{11}. \tag{6.24}$$

Note that such an M exists since $(\Delta - 1)K < b' < b$. Finally, set

$$L_0 := \lceil \ell_0 \log n\rceil = \lceil \tfrac{1}{5}\log_{\Delta-1} n\rceil, \quad L := \lceil M \log n\rceil.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Taking a union bound over Lemmas 4.2 and 4.4, we have that, for all sufficiently large n, a uniformly random Δ-regular graph $G = (V, E)$ with n vertices satisfies the following with probability $1 - o(1)$ over the choice of the graph:

1. $\sum_{\ell=L_0}^{L} \ell C_\ell \Big(K\frac{1-B}{1+B}\Big)^\ell \le 1/(2n^\delta)$, where C_ℓ is the number of cycles of length ℓ in G.
2. Every path P in G with ℓ vertices where $\ell_1 \le \ell \le L$ has an h-graph-neighborhood with at least $(1-\epsilon)\ell$ isolated tree components.

Fix any Δ-regular graph G which satisfies items 1 and 2. The theorem will follow by showing that the output of SampleAntiPotts(G) (cf. Figure 4) is an assignment $\sigma : V \to [q]$ whose distribution ν_σ is within total variation distance $O(1/n^\delta)$ from the Potts measure μ_G with parameter B, i.e.,

$$\|\nu_\sigma - \mu_G\|_{\mathrm{TV}} = O(1/n^\delta).$$

To do this, as in the algorithm SampleAntiPotts(G), let $e_1 = \{u_1, v_1\}, \dots, e_t = \{u_t, v_t\}$ be the edges of G that do not belong to short cycles (i.e., cycles of length $\le \ell_0 \log n$). For $j \in \{0, 1, \dots, t\}$, let G_j be the subgraph $G \backslash \{e_1, \dots, e_j\}$ and μ_j be the Potts distribution on G_j with parameter B; note that the graphs G_j are defined exactly as in the algorithm SampleAntiPotts(G). We will use $\hat{\sigma}_j$ to denote a random configuration distributed according to μ_j; note that σ_j is used to denote the configuration considered by the algorithm at the beginning of step j of the algorithm and, as we shall see soon, its distribution is close to that of $\hat{\sigma}_j$ on G_j.

For an integer $\ell \ge 1$, denote by $\mathcal{P}_{\ell,j}$ the set of paths of length ℓ that connect u_j and v_j in G_j and by $P_{\ell,j} = |\mathcal{P}_{\ell,j}|$ the number of all such paths. Note that

$$\sum_{j=1}^{t} P_{\ell,j} \le \ell C_\ell, \tag{6.25}$$

since every path with ℓ vertices connecting the endpoints of an edge $\{u_j, v_j\}$ in G_j maps to a cycle with ℓ vertices in the initial graph G (by adding the edge $\{u_j, v_j\}$), and each cycle with ℓ vertices in G can potentially arise at most ℓ times under this mapping. Let also

$$\epsilon_j := \frac{6}{n^{10}} + \sum_{\ell=L_0}^{L} P_{\ell,j} \left(K\frac{1-B}{1+B}\right)^\ell. \tag{6.26}$$

Let also $\Omega_j(b, M)$ be the set of all (b, M)-bad configurations on G_j. Using the fact that G satisfies item 2, we will show that for all $j = 1, \dots, t$ it holds that

$$\mu_{G_j}\big(\Omega_j(b, M)\big) \le 1/n^{10}, \tag{6.27}$$

$$\mathbf{E}_j\big[\mathsf{Corr}_{G_j}(U_{\hat{\sigma}_j}, u_j, v_j)\big] \le 2\epsilon_j / B^\Delta, \tag{6.28}$$

where the expectation in (6.28) is over the choice of a random configuration $\hat{\sigma}_j$ distributed according to μ_j and over the choice of the random bichromatic class $U_{\hat{\sigma}_j}$ containing u_j and v_j under $\hat{\sigma}_j$.

We will prove (6.27) and (6.28) shortly, but let us assume them for now and conclude the proof of the theorem. We will prove by induction that for all $j \in \{0, 1, \dots, t\}$ it holds that

$$\big\|\nu_{\sigma_j} - \mu_{G_j}\big\|_{\mathrm{TV}} \le \frac{6(t-j)}{n^{10}} + \frac{4}{B^{\Delta+1}} \sum_{j'=j+1}^{t} \epsilon_{j'}. \tag{6.29}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

For $j = t$, the result holds trivially since σ_t is distributed as μ_{G_t} (exactly). Assume that (6.29) holds for j, where $j \in \{1, \dots, t\}$; we will also show that it holds for $j-1$. To do this, let us consider the configuration σ'_{j-1} defined as follows:

$$\text{if } \hat{\sigma}_j(u_j) \neq \hat{\sigma}_j(v_j), \text{ then } \sigma'_{j-1} = \hat{\sigma}_j \text{ else } \sigma'_{j-1} = \textsc{IdealReSample}(G_j, u_j, v_j, \hat{\sigma}_j),$$

i.e., σ'_{j-1} is obtained using at the jth step of the algorithm the random configuration $\hat{\sigma}_j$ distributed according to μ_{G_j} (exactly). In contrast, note that

$$\text{if } \sigma_j(u_j) \neq \sigma_j(v_j), \text{ then } \sigma_{j-1} = \sigma_j \text{ else } \sigma_{j-1} = \textsc{ReSample}_{b,M}(G_j, u_j, v_j, \sigma_j).$$

Let $\nu_{\sigma'_{j-1}}$ denote the distribution of σ'_{j-1}. Then, by Lemma 6.9 applied to the graph G_j and the inequality in (6.28) we have that

$$\left\|\nu_{\sigma'_{j-1}} - \mu_{G_{j-1}}\right\|_{\mathrm{TV}} \leq \frac{2}{B}\mathbf{E}_j[\mathsf{Corr}_{G_j}(U_{\hat{\sigma}_j}, u_j, v_j)] \leq \frac{4}{B^{\Delta+1}}\epsilon_j. \tag{6.30}$$

Since $\hat{\sigma}_j$ is distributed according to μ_{G_j}, by the coupling lemma, there exists a coupling $\Pr(\cdot)$ of σ_j and $\hat{\sigma}_j$ such that

$$\Pr(\sigma_j \neq \hat{\sigma}_j) = \left\|\nu_{\sigma_j} - \mu_{G_j}\right\|_{\mathrm{TV}}. \tag{6.31}$$

Note also that, for a configuration $\eta \notin \Omega_j(b, M)$, conditioned on $\sigma_j = \eta$ and $\hat{\sigma}_j = \eta$, we can couple σ_{j-1} and σ'_{j-1} so that $\sigma_{j-1} \neq \sigma'_{j-1}$ with probability at most $1/n^{10}$. To see this, if $\eta(u_j) = \eta(v_j)$, then we trivially have $\sigma'_{j-1} = \sigma_{j-1} = \eta$. Otherwise, σ_{j-1} and σ'_{j-1} are produced by first choosing a random bichromatic class containing u_j and v_j under η, which we can couple so that it is the same in both $\textsc{ReSample}_{b,M}(G_j, u_j, v_j, \eta)$ and $\textsc{IdealReSample}(G_j, u_j, v_j, \eta)$. Denote this class by U and let $c_1 = \eta(u_j)$, $c_2 = \eta(v_j)$. Then, from the definition of $\textsc{IdealReSample}(G_j, u_j, v_j, \eta)$ we have that the distribution of $\sigma'_{j-1}(U)$ is given by $\pi^{c_1,c_2}_{G_j[U],u_j,v_j}$, while from $\textsc{ReSample}_{b,M}(G_j, u_j, v_j, \eta)$ we have that the distribution of $\sigma_{j-1}(U)$ is $1/n^{10}$-close to $\pi^{c_1,c_2}_{G_j[U],u_j,v_j}$. We therefore have that

$$\Pr\left(\sigma_{j-1} \neq \sigma'_{j-1} \mid \sigma_j = \hat{\sigma}_j \notin \Omega_j(b, M)\right) \leq 1/n^{10}. \tag{6.32}$$

Invoking the coupling lemma again, we therefore obtain that

$$\begin{aligned}\left\|\nu_{\sigma_{j-1}} - \nu_{\sigma'_{j-1}}\right\|_{\mathrm{TV}} &\leq \Pr(\sigma_j \neq \hat{\sigma}_j) + \Pr\left(\sigma_j = \hat{\sigma}_j \in \Omega_j(b, M)\right) \\ &\quad + \Pr\left(\sigma_{j-1} \neq \sigma'_{j-1} \mid \sigma_j = \hat{\sigma}_j \notin \Omega_j(b, M)\right) \\ &\leq \left\|\nu_{\sigma_j} - \mu_{G_j}\right\|_{\mathrm{TV}} + \mu_{G_j}(\Omega_j(b, M)) + 1/n^{10}, \end{aligned} \tag{6.33}$$

where in the last inequality we used (6.31), (6.32), and

$$\Pr\left(\sigma_j = \hat{\sigma}_j \in \Omega_j(b, M)\right) \leq \Pr\left(\hat{\sigma}_j \in \Omega_j(b, M)\right) = \mu_{G_j}(\Omega_j(b, M)).$$

Combining (6.33) with the inductive hypothesis (6.29) and (6.27), we obtain

$$\left\|\nu_{\sigma_{j-1}} - \nu_{\sigma'_{j-1}}\right\|_{\mathrm{TV}} \leq \frac{6(t-j+1)}{n^{10}} + \frac{4}{B^{\Delta+1}} \sum_{j'=j+1}^{t} \epsilon_{j'}.$$

Using this and (6.30) completes (via the triangle inequality) the inductive step, therefore completing the proof of (6.29) for all $j \in \{0, 1, \dots, t\}$ as wanted.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

We have $G_0 = G$ and $\sigma_0 = \sigma$, so (6.29) for $j = 0$ gives

$$\|\nu_\sigma - \mu_G\|_{\mathrm{TV}} \le \frac{6t}{n^{10}} + \frac{4}{B^{\Delta+1}} \sum_{j=1}^{t} \epsilon_j \le \frac{1}{n^8} + \frac{4}{B^{\Delta+1}} \sum_{j=1}^{t} \sum_{\ell=L_0}^{L} P_{\ell,j} \left(K \frac{1-B}{1+B} \right)^{\ell},$$

where the last inequality holds for all sufficiently large n using the values of ϵ_j from (6.26) (and the crude bound $t \le |E| \le \Delta n$). Using (6.25), we therefore have that

$$\|\nu_\sigma - \mu_G\|_{\mathrm{TV}} \le \frac{1}{n^8} + \frac{4}{B^{\Delta+1}} \sum_{\ell=L_0}^{L} \ell C_\ell \left(K \frac{1-B}{1+B} \right)^{\ell} \le \frac{10}{B^{\Delta+1} n^{\delta}},$$

where the last inequality follows from our assumption that G satisfies item 1 (and by assuming that n is sufficiently large). This completes the proof of Theorem 6.11, modulo the proofs of (6.27) and (6.28), which are given below.

To prove (6.27) and (6.28), fix any value $j \in \{1, \dots, t\}$. Since G_j is a subgraph of G and G satisfies item 2 every path P in G_j with ℓ vertices where $\ell \in [\ell_1, L]$ has an h-graph-neighborhood with at least $(1-\epsilon)\ell$ isolated tree components. By Lemma 6.6, which, recall, applies for all $\ell \ge \ell'$, we therefore have that for any path P in G_j with ℓ vertices and $\ell \in [\max\{\ell_1, \ell'\}, L]$ it holds that

$$\mu_{G_j}(P \text{ is bichromatic}) \le K^{\ell}. \tag{6.34}$$

We are now ready to prove (6.27). For a vertex v, let X_v be the r.v. that counts the number of bichromatic paths with L vertices that start from v in a random configuration $\sigma \sim \mu_{G_j}$. Since G_j has maximum degree $\le \Delta$, there are at most $\Delta(\Delta-1)^{L-2}$ paths starting from v with L vertices and each of them is bichromatic with probability at most K^L from (6.34). Hence, the expectation of X_v is at most

$$\Delta(\Delta-1)^{L-2} K^L \le (b')^L,$$

where the last inequality follows from the choice of M (cf. (6.24)). By Markov's inequality we therefore have that for all $v \in V$ it holds that

$$\mu_{G_j}(X_v > b^L) \le (b'/b)^L \le 1/n^{11},$$

where the last inequality also follows from the choice of M (cf. (6.24)). Note that for (b, M)-bad configurations, i.e., configurations in $\Omega_j(b, M)$, at least one of the events $X_v > b^L$ occurs for some $v \in V$, and therefore by a union bound over $v \in V$ we have that

$$\mu_{G_j}\big(\Omega_j(b, M)\big) \le \sum_{v \in V} \mu_{G_j}\big(X_v > b^L\big) \le 1/n^{10}.$$

This finishes the proof of (6.27).

To prove (6.28), for a set $U \subseteq V$ such that $u_j, v_j \in U$, let

$$\widehat{\mathsf{Corr}}_{G_j}(U, u_j, v_j) := \big|\pi_{G_j[U]}(\eta_{u_j} = 1 \mid \eta_{v_j} = 1) - \pi_{G_j[U]}(\eta_{u_j} = 1 \mid \eta_{v_j} = 2)\big|.$$

Note that by the symmetry of the states in the Ising model, we have that $\pi_{G[U]}\big(\tau_{v_j} = 1\big) = \pi_{G[U]}\big(\tau_{v_j} = 2\big) = 1/2$ and therefore

$$\mathsf{Corr}_{G_j}(U, u_j, v_j) = \left| \frac{\pi_{G[U]}\big(\tau_{u_j} = 1, \tau_{v_j} = 1\big)}{\pi_{G[U]}\big(\tau_{u_j} = 1, \tau_{v_j} = 2\big)} - 1 \right| = \left| \frac{\pi_{G[U]}\big(\tau_{u_j} = 1 \mid \tau_{v_j} = 1\big)}{\pi_{G[U]}\big(\tau_{u_j} = 1 \mid \tau_{v_j} = 2\big)} - 1 \right|.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

By Lemma 6.5 and since G_j has maximum degree $\leq \Delta$, we obtain that

$$\mathsf{Corr}_{G_j}(U, u_j, v_j) \leq \frac{2}{B^{\Delta}} \widehat{\mathsf{Corr}}_{G_j}(U, u_j, v_j),$$

and therefore to prove (6.28) we only need to show that

$$\mathbf{E}_j\big[\widehat{\mathsf{Corr}}_{G_j}(U_{\hat{\sigma}_j}, u_j, v_j)\big] \leq \epsilon_j. \tag{6.35}$$

For convenience, let $\mathcal{F}_j$ be the event that the Potts configuration $\hat{\sigma}_j$ is (b, M)-bad, i.e., $\mathcal{F}_j = \{\hat{\sigma}_j \in \Omega_j(b, M)\}$, and denote by $\overline{\mathcal{F}_j}$ the event that $\hat{\sigma}_j \notin \Omega_j(b, M)$.

$$\begin{aligned}\mathbf{E}_j\big[\widehat{\mathsf{Corr}}_{G_j}(U_{\hat{\sigma}_j}, u_j, v_j)\big] &= \mathbf{E}_j\big[\widehat{\mathsf{Corr}}_{G_j}(U_{\hat{\sigma}_j}, u_j, v_j) \mid \mathcal{F}_j\big]\mu_j\big(\mathcal{F}_j\big) \\ &\quad + \mathbf{E}_j\big[\widehat{\mathsf{Corr}}_{G_j}(U_{\hat{\sigma}_j}, u_j, v_j) \mid \overline{\mathcal{F}_j}\,\big]\mu_j\big(\overline{\mathcal{F}_j}\big) \\ &\leq \frac{1}{n^{10}} + \mathbf{E}_j\big[\widehat{\mathsf{Corr}}_{G_j}(U_{\hat{\sigma}_j}, u_j, v_j) \mid \overline{\mathcal{F}_j}\,\big], \end{aligned} \tag{6.36}$$

where the inequality follows from the bounds $\widehat{\mathsf{Corr}}_{G_j}(U_{\hat{\sigma}_j}, u_j, v_j) \leq 1$ and $\mu_j\big(\mathcal{F}_j\big) \leq 1/n^{10}$ from (6.27). Recall that $\mathcal{P}_{\ell,j}$ denotes the set of paths of length ℓ that connect u_j and v_j in G_j. Consider also the r.v.

$$Y_j = \frac{1}{n^{10}} + \sum_{\ell=L_0}^{L} \sum_{P \in \mathcal{P}_{\ell,j}} \mathbf{1}\{P \text{ is bichromatic}\} \times \left(\frac{1-B}{1+B}\right)^{\ell}.$$

Recall that $U_{\hat{\sigma}_j}$ is a bichromatic class under $\hat{\sigma}_j$ which contains u_j and v_j (chosen uniformly at random from among the set of all such classes if there is more than one). Conditioned on the event $\hat{\sigma}_j \notin \Omega_j(b, M)$, we have that every bichromatic component under $\hat{\sigma}_j$ has average growth b up to depth $L = \lceil M \log n \rceil$ and therefore, irrespective of the random choice of $U_{\hat{\sigma}_j}$, we obtain by Lemma 6.4 that $\widehat{\mathsf{Corr}}_{G_j}(U_{\hat{\sigma}_j}, u_j, v_j) \leq Y_j$. Note that in applying Lemma 6.4, we used that $P_{\ell,j} = 0$ for all $\ell \in [1, L_0]$; this holds because the edge $\{u_j, v_j\}$ does not belong to a short cycle in G (and therefore in G_{j-1} as well). We thus have that

$$\mathbf{E}_j\big[\widehat{\mathsf{Corr}}_{G_j}(U_{\hat{\sigma}_j}, u_j, v_j) \mid \overline{\mathcal{F}_j}\big] \leq \mathbf{E}_j\big[Y_j \mid \overline{\mathcal{F}_j}\big] \leq \frac{\mathbf{E}_j[Y_j]}{\mu_{G_j}\big(\overline{\mathcal{F}_j}\big)} \leq \Big(1 + \frac{2}{n^{10}}\Big)\mathbf{E}_j[Y_j], \tag{6.37}$$

where the last inequality follows from $1/\mu_{G_j}\big(\overline{\mathcal{F}_j}\big) \leq 1/(1 - 1/n^{10}) \leq 1 + 2/n^{10}$ (cf. (6.27)). Using (6.34) and the fact that $L_0 \geq \max\{\ell_1, \ell'\}$ for all sufficiently large n, we have that

$$\mathbf{E}_j[Y_j] \leq \frac{1}{n^{10}} + \sum_{\ell=L_0}^{L} P_{\ell,j} \left(K \frac{|B-1|}{B+1}\right)^{\ell}. \tag{6.38}$$

Since G satisfies item 1, we have that the sum in (6.38) is less than $1/n^{\delta}$ and hence (6.37) and (6.38) give that for all sufficiently large n it holds that

$$\mathbf{E}_j\big[\widehat{\mathsf{Corr}}_{G_j}(U_{\hat{\sigma}_j}, u_j, v_j) \mid \overline{\mathcal{F}_j}\,\big] \leq \frac{1}{n^{10}} + \sum_{\ell=L_0}^{L} P_{\ell,j} \left(K \frac{|B-1|}{B+1}\right)^{\ell} + \frac{2}{n^{10}} \left(\frac{1}{n^{10}} + \frac{1}{n^{\delta}}\right).$$

Combining this with (6.36) yields (6.35), therefore concluding the proof of (6.28). This finishes the proof of Theorem 6.11. ∎

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

7. Analyzing random-cluster on graphs with tree-like structure. In this section, we give the proof of Lemma 5.2. We begin by revisiting the uniqueness results of Häggström [16] on the $(\Delta-1)$-ary tree; then, we use these results to obtain Lemma 5.2 in section 7.2.

7.1. Uniqueness on the $(\Delta-1)$-ary tree. Fix an integer $\Delta \geq 3$. In this section, we review the results of Häggström [16] about uniqueness of random-cluster measures on the infinite $(\Delta-1)$-ary tree. In fact, it will be more relevant for us to consider the case of finite trees of large depth (rather than the infinite tree itself); this approach has also been followed in [15, Chapter 10] for the case $\Delta = 3$.

As in section 2.2, let $\mathbb{T}_\Delta$ denote the infinite $(\Delta-1)$-ary tree with root vertex ρ. For integer $h \geq 0$, let $T_h = (V_h, E_h)$ denote the subtree of $\mathbb{T}_\Delta$ induced by the vertices at distance $\leq h$ from ρ and let L_h denote the leaves of T_h. We will consider the random-cluster distribution on T_h with the so-called wired boundary condition where all the leaves are identified into a single vertex or, equivalently, all the leaves are connected to a vertex "at infinity."[5] In particular, for $S \subseteq E_h$, let $k^*(S)$ denote the number of connected components in the graph with vertex set $V_h \cup \{\infty\}$ and edge set $S \cup (L_h \times \{\infty\})$; the purpose of the extra vertex ∞ and the edges $L_h \times \{\infty\}$ connecting the leaves to ∞ is to capture the wired boundary condition that all leaves are in the same cluster. The "wired" random-cluster distribution on T_h is given by

$$\varphi_h^*(S) = \frac{p^{|S|}(1-p)^{|E_h \setminus S|} q^{k^*(S)}}{Z_h^*}, \text{ where } Z_h^* = \sum_{S \subseteq E_h} p^{|S|}(1-p)^{|E_h \setminus S|} q^{k^*(S)}. \tag{7.1}$$

Denote by $\rho \leftrightarrow \infty$ the event that there exists an open path connecting the root ρ to infinity (or, equivalently, that the root is connected via an open path to some leaf). The following lemma is implicitly proved in [16] in the context of the infinite tree; we give an alternative proof following the approach in [15, Chapter 10] which is carried out for the case $\Delta = 3$. (The proof is for the sake of completeness, and the reader might want to skip this.)

LEMMA 7.1 (see [16, Theorems 1.5 and 1.6]). *Let $\Delta \geq 3$, $q \geq 1$ and $p \in [0,1]$. Then, in the wired random-cluster distribution on T_h, the probability of the event $\rho \leftrightarrow \infty$ converges as h grows, i.e.,*

$$\lim_{h \to \infty} \varphi_h^*(\rho \leftrightarrow \infty) = \varphi^*, \text{ where } \varphi^* \in [0,1].$$

For $p_c(q,\Delta)$ as in (2.3), it holds that $\varphi^* = 0$ if $p < p_c(q,\Delta)$ and $\varphi^* > 0$ if $p > p_c(q,\Delta)$.

Remark 7.2. For the sake of completeness, we remark that at criticality, i.e., when $p = p_c(q,\Delta)$, it holds that $\varphi^* = 0$ iff $1 \leq q \leq 2$.

Proof of Lemma 7.1. For convenience, let $d := \Delta - 1$. Let $Z_{h,\infty}^*$ denote the contribution to Z_h^* from $S \subseteq E_h$ such that ρ is connected to infinity and $Z_{h,\neg\infty}^*$ from $S \subseteq E_h$ such that ρ is not connected to infinity. Note that $\varphi_h^*(\rho \leftrightarrow \infty) = \frac{Z_{h,\infty}^*}{Z_{h,\infty}^* + Z_{h,\neg\infty}^*}$ for all $h \geq 0$. Moreover, with $t := \frac{p}{q} + 1 - p$, we have that

[5] To motivate the wired boundary condition, consider the case where we have a connected graph G and we want to upper bound the probability that a vertex v belongs to a large cluster in an random-cluster configuration. Using the monotonicity of random-cluster (cf. Lemma 5.3), we can restrict our attention to the graph induced by the ball of radius h around the vertex v, by conditioning all edges outside the ball to be open. This conditioning has exactly the same effect as the wiring we describe here.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$Z^*_{h+1,\infty} = q\sum_{k=1}^{d}\binom{d}{k}\left(\frac{Z^*_{h,\infty}}{q}\right)^k\left(t\frac{Z^*_{h,\neg\infty}}{q}\right)^{d-k}\left(1-(1-p)^k\right)$$

$$= q\left(\frac{Z^*_{h,\infty}+tZ^*_{h,\neg\infty}}{q}\right)^d - q\left(\frac{(1-p)Z^*_{h,\infty}+tZ^*_{h,\neg\infty}}{q}\right)^d,$$

$$Z^*_{h+1,\neg\infty} = q^2\sum_{k=0}^{d}\binom{d}{k}\left(\frac{(1-p)Z^*_{h,\infty}}{q}\right)^k\left(t\frac{Z^*_{h,\neg\infty}}{q}\right)^{d-k} = q^2\left(\frac{(1-p)Z^*_{h,\infty}+tZ^*_{h,\neg\infty}}{q}\right)^d.$$

It follows that $\varphi^*_{h+1}(\rho\leftrightarrow\infty) = f\big(\varphi^*_h(\rho\leftrightarrow\infty)\big)$, where

$$f(x) := \frac{\left(t+p(1-\frac{1}{q})x\right)^d - \left(t-\frac{p}{q}x\right)^d}{\left(t+p(1-\frac{1}{q})x\right)^d + (q-1)\left(t-\frac{p}{q}x\right)^d}.$$

Note that[6] $\varphi^*_{h+1}(\rho\leftrightarrow\infty) \le \varphi^*_h(\rho\leftrightarrow\infty)$ for all $h\ge 0$ and therefore, as h goes to infinity, $\varphi^*_h(\rho\leftrightarrow\infty)$ converges to a limit φ^* which is the largest root in the interval $[0,1]$ of the equation

$$\varphi^* = f(\varphi^*).$$

Consider the transformation $u = \frac{t+p(1-1/q)\varphi^*}{t-\frac{p}{q}\varphi^*}$, whose inverse transformation is given by $\varphi^* = \frac{(p+q(1-p))(u-1)}{p(u+q-1)}$. It follows that u is the largest root in the interval $[1, \frac{1}{1-p}]$ of the equation

$$\frac{\big(p+q(1-p)\big)(u-1)}{p(u+q-1)} = \frac{u^d-1}{u^d+(q-1)} \text{ or equivalently } p = 1-\frac{1}{1+h(u)}, \tag{7.2}$$

where $h(y) = \frac{(y-1)(y^d+q-1)}{(y^d-y)}$ is the same function as in (2.3).

We next examine for which values of p it holds that the root u of (7.2) is strictly larger than 1 (note that $\varphi^*>0$ iff $u>1$). Recall from (2.3) that the value of $p_c(q,\Delta)$ is given by

$$p_c(q,\Delta) = 1-\frac{1}{1+\inf_{y>1}h(y)}, \text{ where } h(y) := \frac{(y-1)(y^d+q-1)}{(y^d-y)}. \tag{2.3}$$

First, consider the case $p < p_c(q,\Delta)$. For the sake of contradiction, assume that $u>1$. Then, we obtain from (7.2) that

$$p = 1-\frac{1}{1+h(u)} > 1-\frac{1}{1+\inf_{y>1}h(y)} = p_c(q,\Delta),$$

a contradiction. Thus, $\varphi^*=0$ for all $p<p_c(q,\Delta)$.

Next, consider the case $p>p_c(q,\Delta)$. Using the continuity of the function h in the interval $(1,\infty)$, we obtain that there exists $y>1$ such that $p = 1-\frac{1}{1+h(y)}$. Note that

$$1+h(y) = \frac{1}{1-p}$$

and we have that $y \le 1+h(y)$ for all $y>1$, so in fact $y\in(1,\frac{1}{1-p}]$. It follows that u, the largest root in the interval $[1,\frac{1}{1-p}]$ of (7.2), satisfies $u\ge y>1$, and therefore $\varphi^*>0$. ∎

[6] This follows from Lemma 5.3: the event $\rho\leftrightarrow\infty$ is increasing and φ^*_h can be obtained from φ^*_{h+1} by conditioning the edges incident to the leaves of T_{h+1} to be open.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

We will need the following corollary of Lemma 7.1 for a slightly modified tree where the root has degree $\Delta-2$. In particular, for an integer $h \geq 1$, let $\hat{T}_h = \big(\hat{V}_h, \hat{E}_h\big)$ be the tree obtained by taking $\Delta-2$ disjoint copies of T_{h-1} and joining their root vertices into a new vertex ρ (for $h=0$, we let $\hat{T}_h$ to be the single-vertex graph). Analogously to (7.1), we use $\hat{L}_h$ to denote the leaves of the tree and $\hat{\varphi}^*_h$ to denote the wired measure on $\hat{T}_h$ where all the leaves are wired to infinity, i.e.,

$$\hat{\varphi}^*_h(S) = \frac{p^{|S|}(1-p)^{|\hat{E}_h \setminus S|} q^{k^*(S)}}{\hat{Z}^*_h}, \text{ where } \hat{Z}^*_h = \sum_{S \subseteq \hat{E}_h} p^{|S|}(1-p)^{|\hat{E}_h \setminus S|} q^{k^*(S)}, \tag{7.3}$$

and $k^*(S)$ denotes the number of connected components in the graph $\big(\hat{V}_h \cup \{\infty\}, S \cup (\hat{L}_h \times \infty)\big)$.

COROLLARY 7.3. *Let $\Delta \geq 3$, $q \geq 1$, and $p < p_c(q,\Delta)$. Then, in the wired random-cluster distribution on $\hat{T}_h$, the probability of the event $\rho \leftrightarrow \infty$ converges to 0 as h grows, i.e.,*

$$\lim_{h\to\infty} \hat{\varphi}^*_h(\rho \leftrightarrow \infty) = 0.$$

Proof. Let $T_h = (V_h, E_h)$ and $\hat{T}_h = (\hat{V}_h, \hat{E}_h)$. Since $\hat{T}_h$ is a subgraph of T_h, we may assume that $\hat{V}_h \subseteq V_h$ and $\hat{E}_h \subseteq E_h$. Observe that

$$\hat{\varphi}^*_h(\cdot) = \varphi^*_h(\cdot \mid E_h \setminus \hat{E}_h \text{ closed}).$$

Since the event $\rho \leftrightarrow \infty$ is increasing, we have that

$$\varphi^*_h(\rho \leftrightarrow \infty \mid E_h \setminus \hat{E}_h \text{ closed}) \leq \varphi^*_h(\rho \leftrightarrow \infty).$$

It follows that $\hat{\varphi}^*_h(\rho \leftrightarrow \infty) \leq \varphi^*_h(\rho \leftrightarrow \infty)$. Using Lemma 7.1 for $p < p_c(q,\Delta)$, we obtain the corollary. ∎

7.2. Analyzing random-cluster on disjoint trees whose roots are connected via a path. We are now in position to prove Lemma 5.2, which we restate here for convenience.

LEMMA 5.2. *Let $\Delta \geq 3$ be an integer, $q \geq 1$, and $p < p_c(q,\Delta)$. There exist constants $K < 1/(\Delta-1)$ and $\epsilon > 0$ such that the following holds for all sufficiently large integers ℓ and h.*

Let G be a Δ-regular graph and P be a path with ℓ vertices whose h-graph-neighborhood contains $(1-\epsilon)\ell$ isolated tree components. Let φ_G be the random-cluster distribution on G with parameters p, q. Then,

$$\varphi_G(\text{path } P \text{ is open}) \leq K^\ell.$$

Let us first give the rough idea of the proof. For simplicity, we will consider the somewhat special case where $\epsilon = 0$, but the argument can be easily adapted to account for small positive $\epsilon > 0$. In particular, let H be the graph induced by the h-graph-neighborhood of the path P, together with the edges of P. For $\epsilon = 0$, the assumptions of the lemma imply that H is a union of ℓ disjoint trees, each[7] isomorphic

[7] A technical detail, which is not important for this rough outline, is that we have to be a bit careful with the endpoints of the path, since the degree regularity of G implies that the trees "hanging" from the endpoints of the P have to be copies of $(\Delta-1)$-ary tree T_h (where the root has degree $\Delta-1$ instead of $\Delta-2$, which is the case for $\hat{T}_h$).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

to $\hat{T}_h$, whose roots are connected by a path. Using the monotonicity of the random-cluster distribution (cf. Lemma 5.3) to upper bound the probability that P is open in φ_G, we can condition all the edges outside H to be open; let φ_H^* be the conditioned probability distribution (note the analogy with the wired measure we considered in the previous section).

Consider first the graph F which is the disjoint union of the ℓ trees (i.e., F is obtained from H by removing the edges of the path P) and let φ_F^* be the analogue of φ_H^* (i.e., the random-cluster distribution on $G\backslash P$ where all edges outside F are assumed to be open). Crucially, since $p < p_c(q, \Delta)$, we have by Lemma 7.1 that only $\epsilon'\ell$ roots are connected via an open path to infinity, where $\epsilon' > 0$ is a constant that can be made arbitrarily small by taking the depth h of the trees sufficiently large. Denote by R this (random) set of root vertices, so that $|R| \leq \epsilon'\ell$ with high probability.

Now, we add the edges of the path P and consider how this reweights the random-cluster configuration. In particular, we focus on edges of the path which are not incident to a root in R, letting e be such an edge. If e is open, we get a factor of p/q in the weight of the random-cluster configuration (p because e is open and $1/q$ because the total number of connected components decreases by 1); if e is closed, we get a factor of $1-p$ in the weight of the random-cluster configuration (because e is closed—note that the number of connected components stays the same in this case). Since $|R| \leq \epsilon'\ell$, there are at least $\ell - 2\epsilon'\ell$ edges of the path P that are not incident to a vertex in R. Therefore the probability that all of them are open is roughly $\tau^{\ell(1-2\epsilon)}$, where $\tau := \frac{p/q}{p/q+(1-p)}$. For $p < p_c(q, \Delta)$ it holds that $\tau < 1/(\Delta - 1)$ and hence the lemma follows by taking $\epsilon' > 0$ to be a sufficiently small constant (to account also for the roughly $2^{\epsilon'\ell}$ choices of the set R).

Proof of Lemma 5.2. For convenience, let $d := \Delta - 1$. We begin first by specifying the constant K and how large ℓ and h need to be.

Let $\tau := \frac{p}{p+q(1-p)}$ and note that for all $p < p_c(q, \Delta)$ we have that $\tau < 1/d$ (since $p_c(q, \Delta) < \frac{q}{q+d-1}$). Let K be any constant satisfying $\tau < K < 1/d$ and $\epsilon > 0$ be a small constant such that for $\epsilon' := 10q^3\epsilon$ it holds that $\tau^{1-\epsilon'}q^{10\epsilon'} < K$ (note that such an ϵ exists by considering the limit $\epsilon \downarrow 0$). Let ℓ be sufficiently large so that $\epsilon'\ell \geq \epsilon\ell \geq 2$.

As in section 7.1, for an integer $h \geq 0$, let $\hat{T}_h = (\hat{V}_h, \hat{E}_h)$ denote the subtree of the $(\Delta-1)$-ary tree with height h where the root ρ has degree $\Delta - 2$ (and every other nonleaf vertex has degree Δ). Let $\hat{Z}_h^*$ denote the partition function of the random-cluster model where all the leaves are connected to infinity (cf. (7.3)). Let $\hat{Z}_{h,\infty}^*$ be the contribution to $\hat{Z}_h^*$ from $S \subseteq \hat{E}_h$ such that ρ is connected to infinity and $\hat{Z}_{h,\neg\infty}^*$ from $S \subseteq E_h$ such that ρ is not connected to infinity. Since $p < p_c(q, \Delta)$, we have by Corollary 7.3 that for all sufficiently large h it holds that

$$\hat{Z}_{h,\infty}^*/\hat{Z}_h^* \leq \epsilon. \tag{7.4}$$

We are now ready to proceed to the proof of the lemma. Consider a path P with vertices $u_1, \ldots, u_\ell$ whose h-graph neighborhood contains at least $(1-\epsilon)\ell$ isolated tree components. By definition, an isolated tree component contains exactly one of the vertices $u_1, \ldots, u_\ell$ and therefore the set

$$U = \left\{u_i \,\middle|\, u_i \text{ belongs to an isolated tree component and } u_i \neq u_1, u_\ell\right\}$$

satisfies $|U| \geq (1-\epsilon)\ell - 2 \geq (1-2\epsilon)\ell$. Observe also that

$$\text{for } u \in U,\ \Gamma_h(G\backslash P, u) \text{ induces a subgraph in } G \text{ which is isomorphic to } \hat{T}_h.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Indeed, for $u \in U$, denote by $\mathcal{C}_u$ the component of the h-graph-neighborhood of the path P that the vertex u belongs to. Since $u \in U$, $\mathcal{C}_u$ is isolated and therefore the vertex set of $\mathcal{C}_u$ is precisely $\Gamma_h(G\backslash P, u)$ and $\mathcal{C}_u$ is an induced subgraph of G. It remains to observe that $\mathcal{C}_u$ is a tree (since $u \in U$), every nonleaf vertex in $\mathcal{C}_u$ different than u has degree Δ (since the graph G is Δ-regular), and u itself has degree $\Delta - 2$ (since $u \neq u_1, u_\ell$).

Let F be the subgraph of G induced by the vertex set $\bigcup_{u\in U} \Gamma_h(G\backslash P, u)$. Note that F is a disjoint union of copies of $\hat{T}_h$. For convenience, denote by V_F, E_F the vertex and edge sets of F and similarly denote V_P, E_P for the corresponding sets of the path P. Note that E_F is disjoint from E_P (though V_P and V_F intersect at the roots of the trees). By Lemma 5.3, we have that

$$\varphi_G(P \text{ is open}) \leq \varphi_G(P \text{ is open} \mid \text{ all edges in } E\backslash(E_P \cup E_F) \text{ are open}).$$

The conditioning in the r.h.s. has the same effect as "wiring." In particular, let $H = (V_H, E_H)$ be the subgraph of G with vertex set $V_P \cup V_F$ and edge set $E_P \cup E_F$. Note that H is a path with (disjoint) copies of the tree $\hat{T}_h$ "hanging" from most vertices of the path (in particular, the vertices in U). Let L be the set of leaves of all the trees and let $W = (V_P\backslash U) \cup L$. Let φ^*_H be the random-cluster distribution on H, where we wire all vertices in W to infinity, i.e., φ^*_H is given by

$$\varphi^*_H(S) = \frac{p^{|S|}(1-p)^{|E_H\backslash S|}q^{k^*(S)}}{Z^*_H}, \text{ where } Z^*_H := \sum_{S\subseteq E_H} p^{|S|}(1-p)^{|E_H\backslash S|}q^{k^*(S)}, \tag{7.5}$$

where $k^*(S)$ is the number of connected components in the graph with vertex set $V_H \cup \{\infty\}$ and edge set $S \cup (W \times \infty)$. With this definition, we have that

$$\varphi_G\big(P \text{ is open} \mid \text{ all edges in } E\backslash(E_F \cup E_P) \text{ are open}\big) = \varphi^*_H(P \text{ is open}).$$

Let Q be the contribution to Z^*_H from configurations $S \subseteq E_H$ such that the path P is open, i.e.,

$$Q := \sum_{S\subseteq E_H;\, E_P\subseteq S} p^{|S|}(1-p)^{|E_H\backslash S|}q^{k^*(S)}.$$

We will show that for all sufficiently large h (so that (7.4) holds) and all sufficiently large ℓ it holds that

$$Z^*_H \geq \Big(\frac{p}{q} + (1-p)\Big)^{\ell-1}\big(\hat{Z}_h/q\big)^{|U|}, \qquad Q \leq \Big(\frac{p}{q}\Big)^{\ell-1} q^{10\epsilon'\ell}\big(\hat{Z}_h/q\big)^{|U|}. \tag{7.6}$$

Assuming (7.6) for the moment, note that

$$\varphi^*(P \text{ is open}) = \frac{Q}{Z^*_H} \leq \tau^{\ell-1}q^{10\epsilon'\ell} \leq K^\ell,$$

where in the last inequality we used that for all sufficiently large ℓ, we have $\tau^{\ell-1}q^{10\epsilon'\ell} \leq (\tau^{1-\epsilon'}q^{10\epsilon'})^\ell \leq K^\ell$. This proves the lemma. We therefore focus on proving (7.6).

To prove the first inequality in (7.6), note that for any $S \subseteq E_H$, we have

$$\begin{gathered} |S| = |S \cap E_F| + |S \cap E_P|, \quad |E_H\backslash S| = |E_F\backslash S| + (\ell - 1 - |S \cap E_P|), \\ k^*(S) + |S \cap E_P| \geq k^*(S \cap E_F). \end{gathered} \tag{7.7}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The equalities follow from the fact that $\{E_F, E_P\}$ is a partition of E_H, while the inequality follows by noting that deleting an edge in $S \cap E_P$ increases the number of connected components by at most 1. Using (7.7) and $q \geq 1$, we therefore obtain that (with $m = |S \cap E_P|$ and $S' = S \cap E_F$)

$$Z_H^* \geq \sum_{S \subseteq E_H} p^{|S|}(1-p)^{|E_H \setminus S|} q^{k^*(S \cap E_F) - |S \cap E_P|}$$

$$= \sum_{m=0}^{\ell-1} \binom{\ell-1}{m} \left(\frac{p}{q}\right)^m (1-p)^{\ell-1-m} \sum_{S' \subseteq E_F} p^{|S'|}(1-p)^{|E_F \setminus S'|} q^{k^*(S')}. \tag{7.8}$$

In the following, we focus on lower bounding the inner sum in (7.8). Recall that the graph F consists of $|U|$ disjoint copies of the tree $\hat{T}_h$, each rooted at a vertex in U. For an integer $r = 0, \ldots, |U|$, let

$$\Omega_F(r) = \left\{ S' \subseteq E_F \,\middle|\, \begin{array}{c} \text{exactly } r \text{ vertices in } U \text{ are connected to } \infty \text{ in the graph} \\ \text{with vertex set } V_F \cup \{\infty\} \text{ and edge set } S' \cup (L \times \infty) \end{array} \right\}.$$

Then, we have that

$$q^{|U|-r} \sum_{S' \in \Omega_F(r)} p^{|S'|}(1-p)^{|E_F \setminus S'|} q^{k^*(S')} = \binom{|U|}{r} \left(\hat{Z}_{h,\infty}\right)^r \left(\hat{Z}_{h,\neg\infty}\right)^{|U|-r}. \tag{7.9}$$

(Note that the r.h.s. in (7.9) counts $|U| - r + 1$ times the component containing ∞, while $k^*(S')$ counts only once for all $S' \subseteq E_F$, which explains the need for the factor $q^{|U|-r}$ on the l.h.s.). We therefore have that (using $q \geq 1$)

$$\sum_{S' \subseteq E_F} p^{|S'|}(1-p)^{|E_F \setminus S'|} q^{k^*(S')} = \frac{1}{q^{|U|}} \sum_{r=0}^{|U|} q^r \binom{|U|}{r} \left(\hat{Z}_{h,\infty}\right)^r \left(\hat{Z}_{h,\neg\infty}\right)^{|U|-r} \geq \left(\frac{\hat{Z}_h}{q}\right)^{|U|}. \tag{7.10}$$

Plugging (7.10) into (7.8) and using the binomial expansion yields the first inequality in (7.6), as wanted.

To prove the second inequality in (7.6), consider as before the set of configurations $\Omega_F(r)$ where exactly r roots of the trees are connected to infinity and let $\Omega_H(r) = \{S' \cup E_P \mid S' \in \Omega_F(r)\}$. For $S \in \Omega_H(r)$, we will show

$$\begin{aligned} |S| &= |S \cap E_F| + (\ell - 1), \quad |E_H \setminus S| = |E_F \setminus S|, \\ k^*(S) &+ 2(|U| - r) - (\ell - 1) \leq k^*(S \cap E_F), \end{aligned} \tag{7.11}$$

The equalities are an immediate consequence of the equalities in (7.7) and the fact that, by the definition of $\Omega_H(r)$, we have that $E_P \subseteq S$ for all $S \in \Omega_H(r)$. To justify the inequality, note that there are exactly $M = r + \ell - |U|$ vertices of the path P which are connected to infinity. It follows that there are at least $\ell - 1 - 2M = 2(|U| - r) - (\ell + 1)$ edges in $S \cap E_P = E_P$ whose endpoints are not connected to infinity; deleting any of these edges causes the number of components to increase by one. Using (7.11) and the fact that $q \geq 1$, we can bound Q by

$$Q \leq \sum_{r=0}^{|U|} \sum_{S \in \Omega_H(r)} p^{|S|}(1-p)^{|E_H \setminus S|} q^{k^*(S \cap E_F) + (\ell+1) - 2(|U|-r)} = \left(\frac{p}{q}\right)^{\ell-1} \sum_{r=0}^{|U|} A_r, \tag{7.12}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

where

$$A_r := \sum_{S' \in \Omega_F(r)} p^{|S'|}(1-p)^{|E_F \setminus S'|} q^{k^*(S')+2(\ell-|U|-r)}.$$

Using (7.9) again, we obtain that

$$\frac{A_r}{(\hat{Z}_h/q)^{|U|}} = q^{2(l-|U|)+3r}\binom{|U|}{r}\left(\frac{\hat{Z}_{h,\infty}}{\hat{Z}_h}\right)^r\left(\frac{\hat{Z}_{h,\neg\infty}}{\hat{Z}_h}\right)^{|U|-r}.$$

Recall that $\epsilon' = 10q^3\epsilon$; using (7.4), note that for all $r \geq \epsilon'\ell$ we have

$$\binom{|U|}{r}\left(\frac{\hat{Z}_{h,\infty}}{\hat{Z}_h}\right)^r\left(\frac{\hat{Z}_{h,\neg\infty}}{\hat{Z}_h}\right)^{\ell-r} \leq \left(\frac{\mathrm{e}|U|}{r}\right)^r \epsilon^r \leq \left(\frac{\mathrm{e}\ell}{r}\right)^r \epsilon^r \leq \frac{1}{(4q^3)^r}.$$

We also have that $|U| \geq (1-2\epsilon)\ell$, so $\ell - |U| \leq 2\epsilon\ell$. Therefore, for all sufficiently large ℓ we have the bound

$$\begin{aligned}\frac{1}{(\hat{Z}_h/q)^{|U|}}\sum_{r=0}^{|U|} A_r &\leq q^{4\epsilon\ell}\sum_{r=0}^{|U|} q^{3r}\binom{|U|}{r}\left(\frac{\hat{Z}_{h,\infty}}{\hat{Z}_h}\right)^r\left(\frac{\hat{Z}_{h,\neg\infty}}{\hat{Z}_h}\right)^{\ell-r} \\ &\leq q^{4\epsilon\ell}\left(1+\sum_{r=0}^{\lfloor\epsilon'\ell\rfloor} q^{3\epsilon'\ell}\right) \leq q^{10\epsilon'\ell}.\end{aligned}$$

Plugging this into (7.12) yields the second inequality in (7.6), as needed. ∎

8. Analyzing the Potts model on graphs with tree-like structure. The goal of this section is to prove Lemma 6.6.

8.1. Analyzing Potts on trees. Fix an integer $\Delta \geq 3$. As in section 7.1, we use $\mathbb{T}_\Delta$ denote the infinite $(\Delta-1)$-ary tree with root vertex ρ. For integer $h \geq 0$, let $T_h = (V_h, E_h)$ denote the subtree of $\mathbb{T}_\Delta$ induced by the vertices at distance $\leq h$ from ρ and let L_h denote the leaves of T_h.

Recall that uniqueness on $\mathbb{T}_\Delta$ implies that root-to-leaves correlations on T_h tend to 0 as $h \to \infty$; cf. Definition 2.2. The following lemma extends these decay properties to arbitrary subtrees of T_h; this was proved in [8] in the case of the colorings model and the proof for the Potts model is analogous.

LEMMA 8.1. *Let $\Delta, q \geq 3$ be integers, and let $B > 0$ be in the uniqueness regime of the $(\Delta-1)$-ary tree. There exists a function $\vartheta : \mathbb{N} \to \mathbb{R}_{\geq 0}$ with $\vartheta(h) \to 0$ as $h \to \infty$ such that the following holds for any integer $h \geq 0$.*

Let T'_h be an arbitrary subtree of T_h containing the root ρ and let L'_h be the set of vertices in T'_h at distance exactly h from ρ. Then, for any color $c \in [q]$ and any configuration $\tau : L'_h \to [q]$, it holds that

$$\left|\mu_{T'_h}(\sigma_\rho = c \mid \sigma_{L'_h} = \tau) - \frac{1}{q}\right| \leq \vartheta(h).$$

Proof. We will show that the statement of the lemma holds with the function $\vartheta(\cdot)$ given by $\vartheta(0) = 1$ and

$$\vartheta(h) := \max_{c\in[q],\ \tau:L_h\to[q]} \left|\mu_{T_h}[\sigma(\rho) = c \mid \sigma_{L_h} = \tau] - \frac{1}{q}\right|.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Note that since B is assumed to be in the uniqueness regime of the $(\Delta-1)$-ary tree, we have by definition that $\vartheta(h) \to 0$ as $h \to \infty$.

We first show by induction on h that, for all $h \geq 0$, for all subtrees T'_h of T_h containing the root ρ, for all configurations $\tau : L'_h \to [q]$ and any color $c \in [q]$, it holds that

$$\mu_{T'_h}(\sigma_\rho = c \mid \sigma_{L'_h} = \tau) = \mu_{T_h}(\sigma_\rho = c \mid \sigma_{L'_h} = \tau). \tag{8.1}$$

For $h = 0$ the result is trivial. Supposing that $h \geq 1$ and that the result holds for all integers less than h, we prove the result for h as well. Let $U = \{u_1, \ldots, u_{\Delta-1}\}$ be the children of ρ in T_h and let $U' \subseteq U$ be the children of ρ in T'_h. For a vertex $u \in U$, denote by $T_h(u)$ the subtree of T_h rooted at u and by $L_h(u)$ the vertices in L_h that belong to the tree $T_h(u)$. Similarly, for a vertex $u \in U'$, denote by $T'_h(u)$ the subtree of T'_h rooted at u and by $L'_h(u)$ the vertices in L'_h that belong to the tree $T'_h(u)$. Using standard tree recursions (see, for example, [11, Lemma 19]), we have that

$$\mu_{T_h}(\sigma_\rho = c \mid \sigma_{L'_h} = \tau) = \frac{\prod_{u \in U} \left(1 - (1-B)\mu_{T_h(u)}(\sigma_u = c \mid \sigma_{L'_h(u)} = \tau_{L'_h(u)})\right)}{\sum_{c'=1}^{q} \prod_{u \in U} \left(1 - (1-B)\mu_{T_h(u)}(\sigma_u = c' \mid \sigma_{L'_h(u)} = \tau_{L'_h(u)})\right)} \tag{8.2}$$

and

$$\mu_{T'_h}(\sigma_\rho = c \mid \sigma_{L'_h} = \tau) = \frac{\prod_{u \in U'} \left(1 - (1-B)\mu_{T'_h(u)}(\sigma_u = c \mid \sigma_{L'_h(u)} = \tau_{L'_h(u)})\right)}{\sum_{c'=1}^{q} \prod_{u \in U'} \left(1 - (1-B)\mu_{T'_h(u)}(\sigma_u = c' \mid \sigma_{L'_h(u)} = \tau_{L'_h(u)})\right)}. \tag{8.3}$$

For every $u \in U \backslash U'$ we have that $L'_h(u') = \emptyset$ and therefore, using the symmetry among the colors, we have that for every $c' \in [q]$ it holds that

$$\mu_{T_h(u)}(\sigma_u = c' \mid \sigma_{L'_h(u)} = \tau_{L'_h(u)}) = \mu_{T_h(u)}(\sigma_u = c') = 1/q. \tag{8.4}$$

For $u \in U'$, we have that $T_h(u)$ is isomorphic to T_{h-1} and $T'_h(u)$ is a subtree of $T_h(u)$, so by the induction hypothesis we have that for any color $c \in [q]$ it holds that

$$\mu_{T'_h(u)}(\sigma_u = c' \mid \sigma_{L'_h(u)} = \tau_{L'_h(u)}) = \mu_{T_h(u)}(\sigma_u = c' \mid \sigma_{L_h(u)} = \tau_{L_h(u)}). \tag{8.5}$$

Combining (8.2), (8.3), (8.4), and (8.5) yields (8.1), thus completing the induction.

To complete the proof it remains to observe that for any color $c \in [q]$, it holds that

$$\begin{aligned}\mu_{T_h}(\sigma_\rho = c \mid \sigma_{L'_h} = \tau) &\\ = \sum_{\eta : L_h \backslash L'_h \to [q]} &\mu_{T_h}(\sigma_\rho = c \mid \sigma_{L'_h} = \tau, \sigma_{L_h \backslash L'_h} = \eta) \times \mu_{T_h}(\sigma_{L_h \backslash L'_h} = \eta \mid \sigma_{L'_h} = \tau).\end{aligned}$$

Note that by the definition of the function $\vartheta(\cdot)$, for any $\eta : L_h \backslash L'_h \to [q]$ we have that

$$\left|\mu_{T_h}(\sigma_\rho = c \mid \sigma_{L'_h} = \tau, \sigma_{L_h \backslash L'_h} = \eta) - \frac{1}{q}\right| \leq \vartheta(h)$$

and $\sum_{\eta : L_h \backslash L'_h \to [q]} \mu_{T_h}(\sigma_{L_h \backslash L'_h} = \eta \mid \sigma_{L'_h} = \tau) = 1$, which gives that

$$\left|\mu_{T_h}(\sigma_\rho = c \mid \sigma_{L'_h} = \tau) - \frac{1}{q}\right| \leq \vartheta(h),$$

Combining this with (8.1) yields the lemma. ∎

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

8.2. Analyzing Potts on disjoint trees whose roots are connected via a path. To prove Lemma 6.6, we will need the following lemmas.

LEMMA 8.2. *Let $q \geq 3$ and $B \in (0,1)$, and set $\chi := \frac{1+B}{B+q-1}$. Let P be a path with $\ell \geq 2$ vertices and endpoints u, v. Then, for arbitrary colors $c, c' \in [q]$, it holds that*

$$\mu_P(P \text{ is bichromatic} \mid \sigma_u = c, \sigma_v = c') \leq (4q/B)\chi^{\ell-2}.$$

Proof. Let z be the neighbor of v (note that if $\ell = 2$, then $z = u$). We have that

(8.6)
$$\mu_P(\sigma_v = c' \mid \sigma_u = c) = \sum_{c'' \in [q]} \mu_P(\sigma_v = c' \mid \sigma_z = c'', \sigma_u = c) \times \mu_P(\sigma_z = c'' \mid \sigma_u = c).$$

Note also that for arbitrary colors $c, c', c'' \in [q]$ we have that

$$\mu_P(\sigma_v = c' \mid \sigma_z = c'', \sigma_u = c) = \mu_P(\sigma_v = c' \mid \sigma_z = c'') \geq \frac{B^{\mathbf{1}\{c'=c''\}}}{B+q-1} \geq B/(B+q-1).$$

Combining this with (8.6), we obtain that $\mu_P(\sigma_v = c' \mid \sigma_u = c) \geq B/(B+q-1)$ and hence

$$\mu_P(\sigma_u = c, \sigma_v = c') = \frac{1}{q}\mu_P(\sigma_v = c' \mid \sigma_u = c) \geq \frac{B}{q(B+q-1)}. \tag{8.7}$$

Let $c_1, c_2 \in [q]$ be distinct colors in $[q]$ and $\mathcal{E}(c_1, c_2)$ be the event that every vertex in P is colored with c_1, c_2. The lemma will follow by showing that

$$\mu_P(\mathcal{E}(c_1, c_2)) = 2\chi^{\ell-1}/q. \tag{8.8}$$

Let us briefly conclude the lemma assuming (8.8). Indeed, if $c \neq c'$, then applying (8.8) for $c_1 = c$ and $c_2 = c'$ and using the lower bound in (8.7) we obtain that

$$\mu_P(P \text{ is bichromatic} \mid \sigma_u = c, \sigma_v = c') \leq (4/B)\chi^{\ell-2},$$

while, if $c = c'$, we obtain by summing (8.8) for $c_1 = c$ and the $q-1$ possible values of c_2 that

$$\mu_P(P \text{ is bichromatic} \mid \sigma_u = c, \sigma_v = c') \leq (4q/B)\chi^{\ell-2}.$$

It remains to prove (8.8). For convenience, denote by $w_1, \ldots, w_\ell$ the vertices of P in order so that $u = w_1$ and $v = w_\ell$ and let $C = \{c_1, c_2\}$. Note that

$$\mu_P(\mathcal{E}(c_1, c_2)) = \mu_P(\sigma_{w_1} \in C) \prod_{i=2}^{\ell} \mu_P\big(\sigma_{w_i} \in C \mid \sigma_{w_{i-1}} \in C, \ldots, \sigma_{w_1} \in C\big).$$

We have $\mu_P(\sigma_{w_1} \in C) = 2/q$. For $i = 2, \ldots, n$, let P_i be the path induced by the vertices $w_{i-1}, \ldots, w_n$. We have that

$$\mu_P\big(\sigma_{w_i} \in C \mid \sigma_{w_{i-1}} \in C, \ldots, \sigma_{w_1} \in C\big) = \mu_{P_i}\big(\sigma_{w_i} \in C \mid \sigma_{w_{i-1}} \in C\big).$$

Since $\mu_{P_i}(\sigma_{w_i} \in C) = \mu_{P_i}(\sigma_{w_{i-1}} \in C) = 2/q$, we have by the Bayes' rule that

$$\mu_{P_i}(\sigma_{w_i} \in C \mid \sigma_{w_{i-1}} \in C) = \mu_{P_i}(\sigma_{w_{i-1}} \in C \mid \sigma_{w_i} \in C) = \frac{1+B}{B+q-1} = \chi.$$

Combining these, we obtain (8.8), therefore concluding the proof of Lemma 8.2. ∎

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

We will use the following corollary of Lemma 8.2.

COROLLARY 8.3. *Let $q \geq 3$ and $B \in (0,1)$, and set $\chi := \frac{1+B}{B+q-1}$. Let P be a path with $\ell \geq 2$ vertices and let Λ be a subset of the vertices which includes the endpoints of the path. Then, for any configuration $\tau : \Lambda \to [q]$, it holds that*

$$\mu_P(P \text{ is bichromatic} \mid \sigma_\Lambda = \tau) \leq (4q/B)^{|\Lambda|}\chi^{\ell-2|\Lambda|}.$$

Proof. Let $t := |\Lambda|$ and denote the set of vertices in Λ by $u_1, \ldots, u_t$ in the order that they appear in the path. Note that u_1 and u_t are the endpoints of P (since by assumption Λ includes the endpoints of P). For $i = 1, 2, \ldots, t-1$, let P_i be the path induced by the vertices between u_i and u_{i+1} and let ℓ_i be the number of vertices in P_i. Then, we have

$$\begin{aligned}\mu_P(P \text{ is bichromatic} \mid \sigma_\Lambda = \tau) &\leq \mu_P(P_i \text{ is bichromatic for } i = 1, \ldots, t-1 \mid \sigma_\Lambda = \tau) \\ &= \prod_{i=1}^{t-1} \mu_{P_i}(P_i \text{ is bichromatic} \mid \sigma_{u_i} = \tau_{u_i}, \sigma_{u_{i+1}} = \tau_{u_{i+1}}).\end{aligned}$$

By Lemma 8.2, we have that

$$\mu_{P_i}(P_i \text{ is bichromatic} \mid \sigma_{u_i} = \tau_{u_i}, \sigma_{u_{i+1}} = \tau_{u_{i+1}}) \leq (4q/B)\chi^{\ell_i - 2}$$

and therefore, since $\sum_i^t \ell_i = \ell$ and $\chi, B \in (0,1)$, we obtain that

$$\mu_P(P \text{ is bichromatic} \mid \sigma_\Lambda = \tau) \leq (4q/B)^{|\Lambda|-1}\chi^{\ell-2(|\Lambda|-1)} \leq (4q/B)^{|\Lambda|}\chi^{\ell-2|\Lambda|}.$$

This finishes the proof. ∎

We are now ready to prove Lemma 6.6, which we restate here for convenience.

LEMMA 6.6. *Let $\Delta, q \geq 3$ be integers, and let $B \in (0,1)$ be in the uniqueness regime of the $(\Delta-1)$-ary tree with $B \neq (\Delta-q)/\Delta$. Then, for any $\epsilon' > 0$, there exists a positive constant $K < \frac{1+B}{B+q-1} + \epsilon'$ and $\epsilon > 0$ such that the following holds for all sufficiently large integers ℓ and h.*

Let G be a graph of maximum degree Δ and P be a path with ℓ vertices whose h-graph-neighborhood contains $(1-\epsilon)\ell$ isolated tree components. Let μ_G be the Potts measure on G with parameter B. Then,

$$\mu_G(\text{path } P \text{ is bichromatic}) \leq K^\ell.$$

Proof. Let $\chi := \frac{1+B}{B+q-1}$ and consider arbitrary $\epsilon' > 0$. We begin by specifying the constants K, ϵ and how large ℓ and h need to be. In particular, let K be any constant satisfying $\chi < K < \chi + \epsilon'$ and $\epsilon > 0$ be a small constant such that

$$(4q/\chi B)^{2\epsilon}\left(\frac{1/q+\epsilon}{1/q-\epsilon}\chi\right)^{1-2\epsilon} \leq K/(1+\epsilon). \tag{8.9}$$

Note that such an ϵ exists by considering the limit $\epsilon \downarrow 0$. Let ℓ be sufficiently large so that

$$\epsilon\ell \geq 2 \quad \text{and} \quad (1+\epsilon)^\ell \geq q^2. \tag{8.10}$$

Finally, let $\vartheta(\cdot)$ be the function in Lemma 8.1, so that for all sufficiently large h it holds that

$$\vartheta(h) \leq \epsilon. \tag{8.11}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

We are now ready to proceed to the proof of the lemma. Let G be a graph and consider a path P in G, with vertices $u_1, \dots, u_\ell$, whose h-graph neighborhood contains at least $(1-\epsilon)\ell$ isolated tree components. By definition, an isolated tree component contains exactly one of the vertices $u_1, \dots, u_\ell$ and therefore the set

$$U = \big\{ u_i \,\big|\, u_i \text{ belongs to an isolated tree component and } u_i \neq u_1, u_\ell \big\}$$

satisfies $|U| \geq (1-\epsilon)\ell - 2 \geq (1-2\epsilon)\ell$. Note that we exclude the endpoints of the path P from U, even if they belong to isolated tree components. Let also $\overline{U}$ denote the set $\{u_1, \dots, u_\ell\} \backslash U$, so that $|\overline{U}| \leq 2\epsilon\ell$. Recall that T_h is the subtree of the $(\Delta-1)$-ary tree consisting of vertices at distance at most h from the root. By definition of the set U, we have that

(8.12) for $u \in U$, $\Gamma_h(G \backslash P, u)$ induces a subgraph in G which is a subtree of T_h.

For $u \in U$, let for convenience $V_u = \Gamma_h(G \backslash P, u)$, F_u be the subgraph of G induced on V_u and L_u be the set of vertices that are at distance h from u in F_u. Note that F_u is a tree rooted at u all of whose vertices are at distance at most h from u; L_u is thus the set of all leaves in F_u which are at distance h from u (note that there could be other leaves which are closer to u but these will not matter).

Let F be the union of the subgraphs F_u for $u \in U$; note that F is a disjoint union of copies of trees (each of which is a subtree of T_h from (8.12)). Let also $L := \bigcup_{u \in U} L_u$. We will also denote by V_P, E_P the vertex and edge set of the path P. Note that E_F is disjoint from E_P (though V_P and V_F intersect at V_P). Finally, let $H = (V_H, E_H)$ be the subgraph of G with vertex set $V_P \cup V_F$ and edge set $E_P \cup E_F$.

To bound the probability that P is bichromatic, we will condition on a worst case boundary configuration on L and $\overline{U}$. In particular, we have the bound

$$\mu_G(P \text{ is bichromatic}) \leq \max_{\tau : L \to [q], \tau' : \overline{U} \to [q]} \mu_G(P \text{ is bichromatic} \mid \sigma_L = \tau, \sigma_{\overline{U}} = \tau').$$

To bound the r.h.s., fix arbitrary configurations $\tau : L \to [q], \tau' : \overline{U} \to [q]$ and note that

$$\mu_G(P \text{ is bichromatic} \mid \sigma_L = \tau, \sigma_{\overline{U}} = \tau') = \mu_H(P \text{ is bichromatic} \mid \sigma_L = \tau, \sigma_{\overline{U}} = \tau'),$$

so the lemma will follow (since τ, τ' are arbitrary) by showing that, for arbitrary colors $c_1, c_2 \in [q]$, it holds that

$$\mu_H\big(\mathcal{E}(c_1, c_2) \mid \sigma_L = \tau, \sigma_{\overline{U}} = \tau'\big) \leq K^\ell / q^2, \tag{8.13}$$

where $\mathcal{E}(c_1, c_2)$ is the event that each vertex in P is colored with either c_1 or c_2. Since $\overline{U}$ includes the endpoints of the path P, by Corollary 8.3 we have that

$$\mu_P(\mathcal{E}(c_1, c_2) \mid \sigma_{\overline{U}} = \tau') \leq (4q/B)^{|\overline{U}|} \chi^{\ell - 2|\overline{U}|} = (4q/\chi B)^{|\overline{U}|} \chi^{|U|}. \tag{8.14}$$

Let Z_P be the partition function of the path P and let $Z_P(c_1, c_2)$ be the contribution to Z_P from configurations such that P is colored with c_1 or c_2, i.e.,

$$Z_P = \sum_{\sigma : V_P \to [q];\ \sigma_{\overline{U}} = \tau'} w_P(\sigma), \qquad Z_P(c_1, c_2) = \sum_{\sigma : V_P \to \{c_1, c_2\};\ \sigma_{\overline{U}} = \tau'} w_P(\sigma).$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Then, (8.14) translates into

$$Z_P(c_1, c_2)/Z_P \le (4q/\chi B)^{|\overline{U}|}\chi^{|U|}.$$

The trees hanging from vertices in U reweight the probability that the path is bichromatic but do not cause significant distortion since the states of the roots are roughly uniformly distributed (because of uniqueness on the tree). To quantify this, for $u \in U$ and a color $c \in [q]$, let

$$Z_u(c) = \sum_{\substack{\sigma: V_u \to [q];\\ \sigma_{L_u} = \tau_{L_u},\ \sigma_u = c}} w_{F_u}(\sigma).$$

Let also $Z_u = \sum_{c\in[q]} Z_u(c)$. By Lemma 8.1 and the choice of h in (8.11), we have that for all $c \in [q]$ it holds that

$$\Big|\frac{Z_u(c)}{Z_u} - \frac{1}{q}\Big| \le \vartheta(h) \le \epsilon. \tag{8.15}$$

We can now write $\mu_H\big(\mathcal{E}(c_1, c_2) \mid \sigma_L = \tau,\, \sigma_{\overline{U}} = \tau'\big)$ as

$$\mu_H\big(\mathcal{E}(c_1, c_2) \mid \sigma_L = \tau,\, \sigma_{\overline{U}} = \tau'\big) = \frac{\sum_{\sigma: V_P \to \{c_1,c_2\};\ \sigma_{\overline{U}} = \tau'} w_P(\sigma) \prod_{u\in U} Z_u(\sigma_u)}{\sum_{\sigma: V_P \to [q];\ \sigma_{\overline{U}} = \tau'} w_P(\sigma) \prod_{u\in U} Z_u(\sigma_u)}.$$

Dividing both numerator and denominator by $\prod_{u\in U} Z_u$, we obtain using (8.15) that

$$\begin{aligned}\mu_H\big(\mathcal{E}(c_1, c_2) \mid \sigma_L = \tau,\, \sigma_{\overline{U}} = \tau'\big) &\le \left(\frac{1/q + \epsilon}{1/q - \epsilon}\right)^{|U|} (Z_P(c_1, c_2)/Z_P)\\ &\le (4q/\chi B)^{|\overline{U}|} \left(\frac{1/q + \epsilon}{1/q - \epsilon}\chi\right)^{|U|}.\end{aligned}$$

We have $|U| \ge (1 - 2\epsilon)\ell$ and $|\overline{U}| \le 2\epsilon\ell$, so from the choice of ϵ and ℓ in (8.9) and (8.10), we obtain that

$$\mu_H\big(\mathcal{E}(c_1, c_2) \mid \sigma_L = \tau,\, \sigma_{\overline{U}} = \tau'\big) \le K^\ell/(1 + \epsilon)^\ell \le K^\ell/q^2,$$

which is exactly (8.13), as wanted. This concludes the proof of Lemma 6.6. ∎

9. Correlation decay and sampling for antiferromagnetic Ising. In this section, we prove Theorem 6.3 and Lemma 6.4. The proofs follow relatively easily from correlation decay bounds on trees appearing in [27]; the bounds there are stated for the case of the ferromagnetic Ising model, but there is a simple translation of these bounds to the antiferromagnetic case which allows us to conclude the desired results. In section 9.1, we import the results from the literature that we need. We then give the proof of Theorem 6.3 in section 9.2 and the proof of Lemma 6.4 in section 9.3.

9.1. Preliminaries. Following [27], for a graph G and a vertex u in G, we consider the self-avoiding walk tree $T = T_{\mathrm{SAW}}(G, u)$, which consists of all paths starting from u and not intersecting themselves, except possibly at the terminal vertex of the path.[8] The following lemma originates in the work of Weitz for the hard-core model [33]; it is well-known that the lemma holds more generally for any two-state system. The particular version we state here is close to [27, Lemma 13].

[8] More precisely, a self-avoiding walk in our context is a walk $w = (v_1, \ldots, v_t)$ where vertices are pairwise distinct, except that we allow $v_1 = v_t$. The tree $T = T_{\mathrm{SAW}}(G, u)$ is induced by the set of all walks starting from u, where two walks w_1, w_2 are connected by an edge if w_1 can be obtained from w_2 by deleting the last vertex of w_2 (or vice versa). Naturally, we can relabel walks in T according to their last vertex, so that each vertex in T corresponds to a vertex in G (in a many-to-one fashion).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

LEMMA 9.1 (see, e.g., [27, Lemma 13]). *Let $G = (V, E)$ be a graph and let u be a vertex in G. Consider the self-avoiding walk tree $T = T_{\mathrm{SAW}}(G, u)$ starting from u and denote by A the leaves of the tree. Then, there is a configuration $\eta : A \to \{1, 2\}$ (described in [33]) such that the following holds for any set $\Lambda \subseteq V$ and any configuration $\tau : \Lambda \to \{1, 2\}$.*

Let U_Λ be the set of vertices in T which correspond to vertices in Λ and let $\eta' : U_\Lambda \setminus A \to \{1, 2\}$ be the configuration where each vertex in $U_\Lambda \setminus A$ inherits the state of the corresponding vertex in Λ under τ. Then,

$$\pi_G(\sigma_u = 1 \mid \sigma_\Lambda = \tau) = \pi_T(\sigma_u = 1 \mid \sigma_A = \eta, \sigma_{U_\Lambda \setminus A} = \eta').$$

The following lemma follows from a strong spatial mixing result in [27] for the ferromagnetic Ising model, which in turn builds upon a lemma from [3].

LEMMA 9.2. *Let $B \in (0, 1)$. Let $T = (V, E)$ be a tree, $\Lambda \subseteq V$ be a subset of the vertices, and u be an arbitrary vertex. Let $\tau_1, \tau_2 : \Lambda \to \{1, 2\}$ be two configurations on Λ which differ only on a subset $U \subseteq \Lambda$. Then,*

$$\left|\pi_T(\sigma_u = 1 \mid \sigma_\Lambda = \tau_1) - \pi_T(\sigma_u = 1 \mid \sigma_\Lambda = \tau_2)\right| \le \sum_{v \in U} \left(\frac{1 - B}{1 + B}\right)^{\mathsf{dist}(u,v)},$$

where $\mathsf{dist}(u, v)$ denotes the distance between u and v in T.

Proof. It is well-known that, on bipartite graphs $G = (V, E)$, there is a measure-preserving bijection between configurations of the antiferromagnetic Ising model with parameter $B \in (0, 1)$ and configurations of the ferromagnetic Ising model with parameter $1/B$, obtained by flipping the states of each vertex on one side of the bipartition. The desired inequality therefore follows from the strong spatial mixing result for the ferromagnetic Ising model given in [27, Lemma 14]. To translate the parameterization of that result, note that in [27] the weight of an (ferromagnetic) Ising configuration σ is parameterized to be proportional to $\exp(2\beta m(\sigma))$—therefore $1/B = e^{2\beta}$, so $\tanh \beta = \frac{1-B}{1+B}$. Using this translation, we obtain the desired inequality. ∎

9.2. Proof of Theorem 6.3. To prove Theorem 6.3, we will use the following algorithm that allows us to compute conditional marginal probabilities with very small absolute error.

LEMMA 9.3. *Let $B \in (0, 1)$ and $b > 0$ be constants such that $b\frac{1-B}{1+B} < 1$, and let $\Delta \ge 3$ be an integer. Then, there exists $M_0 > 0$ such that the following holds for all $M > M_0$.*

There is a polynomial-time algorithm that, on input (i) *an n-vertex graph $G = (V, E)$ with maximum degree at most Δ and average growth b up to depth $L = \lceil M \log n \rceil$,* (ii) *a subset $\Lambda \subseteq V$ with a configuration $\tau : \Lambda \to \{1, 2\}$, and* (iii) *a vertex $u \in V \setminus \Lambda$, outputs a number $\hat{p} \in [0, 1]$ such that*

$$\left|\hat{p} - p\right| \le \frac{B^\Delta}{2n^{11}(1 + B^\Delta)}, \quad \text{where } p := \pi_G(\sigma_u = 1 \mid \sigma_\Lambda = \tau),$$

i.e., $\hat{p}$ is within absolute error $O(1/n^{11})$ from the marginal probability that $\sigma_u = 1$ conditioned on $\sigma_\Lambda = \tau$, where σ is from the Ising distribution π_G on G with parameter B.

Proof. Letting M_0 be a large constant so that for $n \ge 2$ and $L_0 := \lceil M_0 \log n \rceil$, it holds that

$$b^{L_0} \left(\frac{1 - B}{1 + B}\right)^{L_0} \le \frac{B^\Delta}{2n^{11}(1 + B^\Delta)}. \tag{9.1}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Note that such an M_0 exists since $b\frac{1-B}{1+B} < 1$. Fix M to be an arbitrary constant larger than M_0.

Let G be an arbitrary n-vertex graph with average growth b up to depth $L = \lceil M \log n \rceil$, $\Lambda \subseteq V$ be a subset of the vertices, $\tau : \Lambda \to \{1,2\}$ be a configuration on Λ, and $u \in V \setminus \Lambda$ be a vertex. Consider the self-avoiding walk tree $T = T_{\mathrm{SAW}}(G,u)$ starting from u, and denote by A the leaves of the tree. By Lemma 9.1, we have that there is a configuration $\eta : A \to \{1,2\}$ such that

$$\pi_G(\sigma_u = 1 \mid \sigma_\Lambda = \tau) = \pi_T(\sigma_u = 1 \mid \sigma_A = \eta, \sigma_{U\setminus A} = \eta'), \tag{9.2}$$

where $U := U_\Lambda$ is the set of vertices in T that correspond to some vertex in Λ and η' is the assignment on $U \setminus A$ inherited by τ (see Lemma 9.1 for details). For convenience, let $F := U \cup A$ and $\zeta : F \to \{1,2\}$ be the configuration which agrees with η on A and with η' on $U\setminus A$, so that (9.2) can be rewritten as

$$\pi_G(\sigma_u = 1 \mid \sigma_\Lambda = \tau) = \pi_T(\sigma_u = 1 \mid \sigma_F = \zeta).$$

Let $T' = (V', E')$ be the subtree of T induced by vertices at distance at most $L-1$ from u and let $F' = F \cap V'$. Let

$$p := \pi_T(\sigma_u = 1 \mid \sigma_F = \zeta) \quad \text{and} \quad \hat{p} := \pi_{T'}(\sigma_u = 1 \mid \sigma_{F'} = \zeta_{F'}).$$

Note that $\hat{p}$ can be computed in polynomial time since the tree T' and the configuration $\zeta_{F'}$ can be constructed in polynomial time (since T' is a tree of size $O(\Delta^L)$). So, the lemma will follow by showing that

$$\left|\hat{p} - p\right| \le \frac{B^\Delta}{2n^{11}(1+B^\Delta)}. \tag{9.3}$$

To prove this, let J be the set of vertices in $V'\setminus F'$ whose distance from u is exactly $L-1$ in T and note that $|J| \le b^L$ since G has average growth b up to depth L (cf. Definition 3.3). Note also that, conditioned on the configurations on J and F', the probability that $\sigma_u = 1$ depends only on T' (and not on the configuration on the rest of the tree T). Using the law of total probability, we can therefore expand p as

$$p = \sum_{\iota : J \to \{1,2\}} \pi_{T'}(\sigma_u = 1 \mid \sigma_{F'} = \zeta_{F'}, \sigma_J = \iota) \times \pi_T(\sigma_J = \iota \mid \sigma_F = \zeta_F). \tag{9.4}$$

Therefore, (9.3) will follow by showing that for any configuration $\iota : J \to \{1,2\}$ it holds that

$$\left|\pi_{T'}\big(\sigma_u = 1 \mid \sigma_{F'} = \zeta_{F'},\, \sigma_J = \iota\big) - \hat{p}\right| \le \frac{B^\Delta}{2n^{11}(1+B^\Delta)}. \tag{9.5}$$

Note that we can expand $\hat{p}$ analogously to (9.4) by conditioning on the configuration on J, so to prove (9.4) it suffices to show that for any two configurations $\iota_1, \iota_2 : J \to [q]$ it holds that

$$\kappa \le \frac{B^\Delta}{2n^{11}(1+B^\Delta)},$$

where $\kappa := \left|\pi_{T'}\big(\sigma_u = 1 \mid \sigma_{F'} = \zeta_{F'},\, \sigma_J = \iota_1\big) - \pi_{T'}\big(\sigma_u = 1 \mid \sigma_{F'} = \zeta_{F'},\, \sigma_J = \iota_2\big)\right|$. By the strong spatial mixing result of Lemma 9.2, we have that

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$\kappa \le \sum_{v \in J} \left(\frac{1-B}{B+1}\right)^{\mathrm{dist}(u,v)} = |J| \left(\frac{1-B}{B+1}\right)^{L} \le b^L \left(\frac{1-B}{1+B}\right)^{L}.$$

Combining this with the choice of M_0 (cf. (9.1)), we obtain (9.5), thus concluding the proof of the lemma. ∎

Using Lemma 9.3, the proof of Theorem 6.3 follows by standard techniques. We restate it here for convenience.

THEOREM 6.3. *Let $B \in (0,1)$ and $b > 0$ be constants such that $b\frac{1-B}{1+B} < 1$, and let $\Delta \ge 3$ be an integer. Then, there exists $M_0 > 0$ such that the following holds for all $M > M_0$.*

There is a polynomial-time algorithm that, on input an n-vertex graph G with maximum degree at most Δ and average growth b up to depth $L = \lceil M \log n \rceil$, outputs a configuration $\tau : V \to \{1,2\}$ whose distribution ν_τ is within total variation distance $1/n^{10}$ from the Ising distribution on G with parameter B, i.e.,

$$\big\|\nu_\tau - \pi_G\big\|_{\mathrm{TV}} \le 1/n^{10}.$$

Moreover, the algorithm, when given as additional input two vertices u and v in G, outputs a configuration $\tau : V \to \{1,2\}$ such that $\tau_u = 1$ and $\tau_v = 2$, and whose distribution ν_τ satisfies

$$\big\|\nu_\tau - \pi^{1,2}_{G,u,v}(\cdot)\big\|_{\mathrm{TV}} \le 1/n^{10},$$

where $\pi^{1,2}_{G,u,v}$ is the Ising distribution on G conditioned on u having state 1 and v having state 2.

Proof of Theorem 6.3. Denote by $v_1, v_2, \ldots, v_n$ the vertices of G. The algorithm will sample the state s_i of vertex v_i sequentially for $i = 1, \ldots, n$. We just give the details for the first part of the algorithm; the proof of the second part is completely analogous (namely, it suffices to assume in the following that we first set $v_1 = u, v_2 = v$, then fix the states $s_1 = 1, s_2 = 2$, and finally sample the states s_i for $i = 3, \ldots, n$).

Assume that, at some time $i = 1, \ldots, n$, we have sampled the states $s_1, \ldots, s_{i-1}$ (which can take arbitrary values in $\{1,2\}$). Using the algorithm of Lemma 9.3, we obtain in polynomial time numbers $\hat{p}_i(1), \hat{p}_i(2) \in [0,1]$ such that $\hat{p}_i(1) + \hat{p}_i(2) = 1$ and, for $s \in \{1,2\}$, it holds that

$$\big|a_i(s) - \hat{p}_i(s)\big| \le \frac{B^\Delta}{2n^{11}(1+B^\Delta)}, \tag{9.6}$$

where $a_i(s) := \pi_G(\sigma_{v_i} = s \mid \sigma_{v_1} = s_1, \ldots, \sigma_{v_{i-1}} = s_{i-1})$. We then sample the state s_i by letting $s_i = 1$ with probability $\hat{p}_i(1)$, or else $s_i = 2$ (note that $s_i = 2$ with probability $\hat{p}_i(2)$). Denote by τ the final configuration and by ν_τ its distribution.

We will show that, for any configuration $\eta : V \to \{1,2\}$, it holds that

$$|\nu_\tau(\eta) - \pi_G(\eta)| \le \frac{2}{n^{10}} \pi_G(\eta), \tag{9.7}$$

so by summing over η we obtain

$$\big\|\nu_\tau - \pi_G\big\|_{\mathrm{TV}} = \frac{1}{2} \sum_{\eta\colon V \to \{1,2\}} \big|\nu_\tau(\eta) - \pi_G(\eta)\big| \le \frac{1}{n^{10}} \sum_{\eta\colon V \to \{1,2\}} \pi_G(\eta) = 1/n^{10},$$

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

which proves the first part of the theorem; the second part follows analogously. To prove (9.7), fix an arbitrary configuration $\eta : V \to \{1, 2\}$ and let

$$p_{i,\eta} := \pi_G(\sigma_{v_i} = \eta_{v_i} \mid \sigma_{v_1} = \eta_{v_1}, \dots, \sigma_{v_{i-1}} = \eta_{v_{i-1}}).$$

Note that

$$\pi_G(\eta) = \prod_{i=1}^{n} p_{i,\eta} \quad \text{and} \quad \nu_\tau(\eta) = \prod_{i=1}^{n} \hat{p}_i(\eta_{v_i}). \tag{9.8}$$

Moreover, from (9.6), we have that for all $i = 1, \dots, n$ it holds that

$$\left|p_{i,\eta} - \hat{p}_i(\eta_{v_i})\right| \le \frac{B^\Delta}{2n^{11}(1 + B^\Delta)} \le \frac{p_{i,\eta}}{2n^{11}},$$

where the last inequality follows from the lower bound of Lemma 6.5. Using the inequalities $1 - x \le \mathrm{e}^{-x} \le 1 - x/2$ which hold for all $x \in [0, 1/2]$, we obtain that, for $\epsilon = 1/n^{11}$, it holds that

$$\mathrm{e}^{-\epsilon} p_{i,\eta} \le \hat{p}_i(\eta_{v_i}) \le \mathrm{e}^{\epsilon} p_{i,\eta}.$$

Multiplying over $i = 1, \dots, n$ and combining with (9.8) gives (9.7), thus completing the proof of Theorem 6.3. ∎

9.3. Proof of Lemma 6.4. In this section, we give the proof of Lemma 6.4, which we restate here for convenience.

LEMMA 6.4. *Let $B \in (0, 1)$ and $b > 0$ be constants such that $b\frac{1-B}{1+B} < 1$. Then, there exists $M_0' > 0$ such that the following holds for all $M > M_0'$.*

Let G be an n-vertex graph with average growth b up to depth $L = \lceil M \log n \rceil$, and let u, v be distinct vertices in G. Then

$$\left|\pi_G(\sigma_u = 1 \mid \sigma_v = 1) - \pi_G(\sigma_u = 1 \mid \sigma_v = 2)\right| \le \frac{1}{n^{10}} + \sum_{\ell=1}^{L} P_\ell(G, u, v) \left(\frac{1-B}{1+B}\right)^\ell,$$

where $P_\ell(G, u, v)$ is the number of paths with ℓ vertices in G that connect u and v.

Proof of Lemma 6.4. Let M_0' be sufficiently large so that for all $n \ge 2$ and $L_0 := \lceil M_0' \log n \rceil$, it holds that

$$b^{L_0} \left(\frac{1-B}{1+B}\right)^{L_0} \le 1/(2n^{10}). \tag{9.9}$$

Note that such a constant exists since $b\frac{1-B}{1+B} < 1$. Fix arbitrary $M > M_0'$. Let G be an arbitrary graph with average growth b up to depth $L = \lceil M \log n \rceil$, $\Lambda \subseteq V$ be a subset of the vertices, $\tau : \Lambda \to \{1, 2\}$ be a configuration on Λ, and $u \in V \setminus \Lambda$ be a vertex.

Consider the self-avoiding walk tree $T = T_{\mathrm{SAW}}(G, u)$ starting from u, and denote by A the leaves of the tree and by U the set of vertices in T that correspond to v. For a subset of vertices W of the tree we denote by $\sigma_W = 1$ the event that all vertices in W have state 1 and analogously for $\sigma_W = 2$. By Lemma 9.1, we have that there is a configuration $\eta : A \to \{1, 2\}$ such that for $s \in \{1, 2\}$ it holds that

$$\pi_G(\sigma_u = 1 \mid \sigma_v = s) = p_s, \text{ where } p_s := \pi_T\big(\sigma_u = 1 \mid \sigma_A = \eta, \sigma_{U \setminus A} = s\big). \tag{9.10}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Let $T' = (V', E')$ be the subtree of T induced by vertices at distance at most $L-1$ from u and let $A' = A \cap V'$, $U' = U \cap V'$. For $s \in \{1,2\}$, let

$$p'_s := \pi_{T'}(\sigma_u = 1 \mid \sigma_{A'} = \eta_{A'}, \sigma_{U'\setminus A'} = s).$$

We will show that

$$|p'_1 - p'_2| \le \sum_{\ell=1}^{L} P_\ell(G,u,v) \left(\frac{1-B}{1+B}\right)^\ell, \tag{9.11}$$

$$|p_1 - p'_1| \le 1/(2n^{10}), \quad |p_2 - p'_2| \le 1/(2n^{10}). \tag{9.12}$$

Assuming these for the moment, we obtain by (9.10) and the triangle inequality that

$$\begin{aligned}|\pi_G(\sigma_u = 1 \,|\, \sigma_v = 1) - \pi_G(\sigma_u = 1 \,|\, \sigma_v = 2)| &= |p_1 - p_2| \\ &\le \frac{1}{n^{10}} + \sum_{\ell=1}^{L} P_\ell(G,u,v) \left(\frac{1-B}{1+B}\right)^\ell,\end{aligned}$$

thus proving the lemma. It thus remains to prove (9.11) and (9.12).

To prove (9.11), note that by Lemma 9.2, we have that

$$|p'_1 - p'_2| \le \sum_{w \in U'\setminus A'} \left(\frac{1-B}{1+B}\right)^{\mathsf{dist}(u,w)} \le \sum_{\ell=1}^{L} P_\ell(G,u,v) \left(\frac{1-B}{1+B}\right)^\ell,$$

where the last inequality follows from observing that each vertex $w \in U'\setminus A'$ with $\mathsf{dist}(u,w) = \ell$ corresponds to a distinct path with $\ell+1$ vertices between u and v in G. This proves (9.11).

To prove (9.12), we focus on showing that $|p_1 - p'_1| \le 1/(2n^{10})$, the other inequality being completely analogous. We follow closely a similar argument which was presented in the proof of Lemma 9.3. Let J be the set of vertices in $V'\setminus(U' \cup A')$ whose distance from u is exactly $L-1$ in T and note that $|J| \le b^L$ since G has average growth b up to depth L (cf. Definition 3.3). Note also that, conditioned on the configurations on J and $U' \cup A'$, the probability that $\sigma_u = 1$ depends only on T' (and not on the configuration on the rest of the tree T), i.e., we can write p_1 as

$$\sum_{\iota : J \to \{1,2\}} \pi_{T'}(\sigma_u = 1 \mid \sigma_{A'} = \eta_{A'}, \sigma_{U'\setminus A'} = 1, \sigma_J = \iota) \times \pi_T(\sigma_J = \iota \mid \sigma_A = \eta, \sigma_{U\setminus A} = 1). \tag{9.13}$$

Therefore, $|p_1 - p'_1| \le 1/(2n^{10})$ will follow by showing that, for any configuration $\iota : J \to \{1,2\}$ it holds that

$$\left|\pi_{T'}\big(\sigma_u = 1 \mid \sigma_{A'} = \eta_{A'}, \sigma_{U'\setminus A'} = 1, \sigma_J = \iota\big) - p'_1\right| \le 1/(2n^{10}). \tag{9.14}$$

We can expand p'_1 analogously to (9.13) by conditioning on the configuration on J, so to prove (9.14) it suffices to show that for any two configurations $\iota_1, \iota_2 : J \to [q]$ it holds that $\kappa \le 1/(2n^{10})$, where κ equals

$$\begin{aligned}\Big|&\pi_{T'}\big(\sigma_u = 1 \mid \sigma_{A'} = \eta_{A'}, \sigma_{U'\setminus A'} = 1, \sigma_J = \iota_1\big) \\ &\quad - \pi_{T'}\big(\sigma_u = 1 \mid \sigma_{A'} = \eta_{A'}, \sigma_{U'\setminus A'} = 1, \sigma_J = \iota_2\big)\Big|.\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

By the strong spatial mixing result of Lemma 9.2, we have that

$$\kappa \leq \sum_{v \in J} \left(\frac{1-B}{B+1}\right)^{\operatorname{dist}(u,v)} = |J| \left(\frac{1-B}{B+1}\right)^{L} \leq b^{L} \left(\frac{1-B}{1+B}\right)^{L}.$$

Combining this with the choice of M_0' (cf. (9.9)), we obtain $\kappa \leq 1/(2n^{10})$, thus concluding the proof of $|p_1 - p_1'| \leq 1/(2n^{10})$ and therefore completing the proof of Lemma 6.4. ∎

REFERENCES

[1] K. S. Alexander, *Mixing properties and exponential decay for lattice systems in finite volumes*, Ann. Probab., 32 (2004), pp. 441–487.

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

[3] N. Berger, C. Kenyon, E. Mossel, and Y. Peres, *Glauber dynamics on trees and hyperbolic graphs*, Probab. Theory Related Fields, 131 (2005), pp. 311–340.

[4] M. Bordewich, C. Greenhill, and V. Patel, *Mixing of the Glauber dynamics for the ferromagnetic Potts model*, Random Structures Algorithms, 48 (2016), pp. 21–52.

[5] G. R. Brightwell and P. Winkler, *Random colorings of a Cayley tree*, Contemp. Combin., 10 (2002), pp. 247–276.

[6] A. Dembo, A. Montanari, A. Sly, and N. Sun, *The replica symmetric solution for Potts models on d-regular graphs*, Comm. in Math. Phys., 327 (2014), pp. 551–575.

[7] C. Efthymiou, *A simple algorithm for random colouring $G(n, d/n)$ using $(2+\epsilon)d$ colours*, in Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms, 2012, pp. 272–280.

[8] C. Efthymiou, *A simple algorithm for sampling colorings of $G(n, d/n)$ up to the Gibbs uniqueness threshold*, SIAM J. Comput., 45 (2016), pp. 2087–2116.

[9] C. Efthymiou, T. P. Hayes, D. Štefankovič, and E. Vigoda, *Sampling random colorings of sparse random graphs*, in Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms, 2018, pp. 1759–1771.

[10] A. Frieze and M. Karoński, *Introduction to Random Graphs*, Cambridge University Press, Cambridge, UK, 2015.

[11] A. Galanis, L. A. Goldberg, and K. Yang, *Uniqueness of the 3-State Antiferromagnetic Potts Model on the Tree*, arXiv:1804.03514, 2018.

[12] A. Galanis, D. Štefankovič, and E. Vigoda, *Inapproximability for antiferromagnetic spin systems in the tree nonuniqueness region*, J. ACM, 62 (2015), pp. 50:1–50:60.

[13] A. Galanis, D. Štefankovič, E. Vigoda, and L. Yang, *Ferromagnetic Potts model: Refined #BIS-hardness and related results*, SIAM J. Comput., 45 (2016), pp. 2004–2065.

[14] A. Gerschenfeld and A. Montanari, *Reconstruction for models on random graphs*, in Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science, 2007, pp. 194–204.

[15] G. Grimmett, *The Random-Cluster Model*, Springer, New York, 2006.

[16] O. Häggström, *The random-cluster model on a homogeneous tree*, Probab. Theory Related Fields, 104 (1996), pp. 231–253.

[17] T. Helmuth, W. Perkins, and G. Regts, *Algorithmic Pirogov-Sinai Theory*, arXiv:1806.11548, 2018.

[18] S. Janson, T. Łuczak, and A. Rucinski, *Random Graphs*, Wiley Ser. Discrete Math. Optim. 45, John Wiley & Sons, New York, 2011.

[19] M. Jenssen, P. Keevash, and W. Perkins, *Algorithms for #BIS-Hard Problems on Expander Graphs*, arXiv:1807.04804, 2018.

[20] J. Jonasson, *The random cluster model on a general graph and a phase transition characterization of nonamenability*, Stochastic Process. Appl., 79 (1999), pp. 335–354.

[21] J. Jonasson, *Uniqueness of Uniform Random Colorings of Regular Trees*, Statist. Probab. Lett., 57 (2002), pp. 243–248.

[22] F. Martinelli and E. Olivieri, *Approach to equilibrium of Glauber dynamics in the one phase region.* I. *The attractive case*, Comm. Math. Phys., 161 (1994), pp. 447–486.

[23] F. Martinelli and E. Olivieri, *Approach to equilibrium of Glauber dynamics in the one phase region.* II. *The general case*, Comm. Math. Phys., 161 (1994), pp. 487–514.

[24] F. Martinelli, E. Olivieri, and R. H. Schonmann, *For 2-D lattice spin systems weak mixing implies strong mixing*, Comm. in Math. Phys., 165 (1994), pp. 33–47.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[25] M. Mézard and A. Montanari, *Information, Physics, and Computation*, Oxford University Press, New York, 2009.
[26] E. Mossel and A. Sly, *Rapid mixing of Gibbs sampling on graphs that are sparse on average*, Random Structures Algorithms, 35 (2009), pp. 250–270.
[27] E. Mossel and A. Sly, *Exact thresholds for Ising–Gibbs samplers on general graphs*, Ann. Probab., 41 (2013), pp. 294–328.
[28] E. Mossel, D. Weitz, and N. Wormald, *On the hardness of sampling independent sets beyond the tree threshold*, Probab. Theory Related Fields, 143 (2009), pp. 401–439.
[29] A. Sinclair, P. Srivastava, D. Štefankovič, and Y. Yin, *Spatial mixing and the connective constant: Optimal bounds*, Probab. Theory Related Fields, 168 (2017), pp. 153–197.
[30] A. Sinclair, P. Srivastava, and M. Thurley, *Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs*, J. Statist. Phys., 155 (2014), pp. 666–686.
[31] A. Sinclair, P. Srivastava, and Y. Yin, *Spatial mixing and approximation algorithms for graphs with bounded connective constant*, in 54th Annual IEEE Symposium on Foundations of Computer Science, 2013, pp. 300–309.
[32] L. E. Thomas, *Bound on the mass gap for finite volume stochastic Ising models at low temperature*, Comm. Math. Phys., 126 (1989), pp. 1–11.
[33] D. Weitz, *Counting independent sets up to the tree threshold*, in Proceedings of the 38th Annual ACM Symposium on Theory of Computing, 2006, pp. 140–149.
[34] Y. Yin and C. Zhang, *Sampling in Potts model on sparse random graphs*, in Proceedings of Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 2016, pp. 47:1–47:22.
[35] J. Zhang, H. Liang, and F. Bai, *Approximating partition functions of the two-state spin system*, Inform. Process. Lett., 111 (2011), pp. 702–710.

Downloaded 06/29/20 to 88.104.25.182. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.