

Flip Dynamics for Sampling Colorings: Improving $(11/6 - \varepsilon)$ Using A Simple Metric

Charlie Carlson* Eric Vigoda*

Abstract

We present improved bounds for randomly sampling k -colorings of graphs with maximum degree Δ ; our results hold without any further assumptions on the graph. The Glauber dynamics is a simple single-site update Markov chain. Jerrum (1995) proved an optimal $O(n \log n)$ mixing time bound for Glauber dynamics whenever $k > 2\Delta$ where Δ is the maximum degree of the input graph. This bound was improved by Vigoda (1999) to $k > (11/6)\Delta$ using a “flip” dynamics which recolors (small) maximal 2-colored components in each step. Vigoda’s result was the best known for general graphs for 20 years until Chen et al. (2019) established optimal mixing of the flip dynamics for $k > (11/6 - \varepsilon)\Delta$ where $\varepsilon \approx 10^{-5}$. We present the first substantial improvement over these results. We prove an optimal mixing time bound of $O(n \log n)$ for the flip dynamics when $k \geq 1.809\Delta$. This yields, through recent spectral independence results, an optimal $O(n \log n)$ mixing time for the Glauber dynamics for the same range of k/Δ when $\Delta = O(1)$. Our proof utilizes path coupling with a simple weighted Hamming distance for “unblocked” neighbors.

1 Introduction

A problem of great interest and considerable study at the intersection of theoretical computer science, discrete mathematics, and statistical physics is the random sampling of k -colorings of a given input graph G . Given a graph $G = (V, E)$ of maximum degree Δ and an integer $k \geq 2$, let Ω denote the collection of proper vertex k -colorings of G , that is Ω is the collection of assignments $X_t : V \rightarrow \{1, \dots, k\}$ where for all $(v, w) \in E$, $X_t(v) \neq X_t(w)$. Let π denote the uniform distribution over Ω . The colorings problem is a natural example of a non-binary graphical model [Mur12, KF09] and, in statistical physics, it is the zero-temperature limit of the antiferromagnetic Potts model [SS97].

We study algorithms for the approximate counting problem of estimating $|\Omega|$, the number of k -colorings, and the approximate sampling problem of generating random k -colorings from a nearly uniform distribution. In particular, given a graph $G = (V, E)$ and a $\delta > 0$, for the sampling problem, our goal is to sample from a distribution μ which is within total variation distance $\leq \varepsilon$ of the uniform distribution π in time polynomial in $n = |V|$ and $\log(1/\varepsilon)$. In the approximate counting problem, a graph $G = (V, E)$, an ε , and a $\delta > 0$ are given, and the goal is to obtain a FPRAS to estimate $|\Omega|$, which is an algorithm to estimate $|\Omega|$ within a $(1 \pm \varepsilon)$ multiplicative factor with probability $\geq 1 - \delta$ in time $\text{poly}(n, 1/\varepsilon, \log(1/\delta))$. These approximate sampling/counting problems are polynomial-time inter-reducible to each other. Relevant for our work, an $O(n \log(n/\varepsilon))$ sampling algorithm yields an $O^*(n^2)$ time approximate counting algorithm [ŠVV09, Hub15, Kol18].

The Markov chain Monte Carlo (MCMC) method is a natural algorithmic approach to approximate sampling. The Glauber dynamics (also known as the Gibbs sampler) is the simplest example

*Department of Computer Science, University of California, Santa Barbara. Email: {charlieanneCarlson,vigoda}@ucsb.edu. Research supported in part by NSF grant CCF-2147094.

of the MCMC method. The Glauber dynamics is a Markov chain on the collection of k -colorings and the transitions update the coloring at a randomly chosen vertex in each step as follows. From a coloring $X_t \in \Omega$, choose a vertex $v \in V$ and a color $c \in \{1, \dots, k\}$ uniformly at random. If no neighbor of v has color c in the current coloring X_t then we recolor v to color c in X_{t+1} and all other vertices maintain the same color $X_{t+1}(w) = X_t(w)$ for all $w \neq v$; and if color c is not available for v then we set $X_{t+1} = X_t$. The Glauber dynamics is ergodic whenever $k \geq \Delta + 2$; hence, the unique stationary distribution is the uniform distribution π .

The mixing time T_{mix} is the number of steps from the worst initial state X_0 to guarantee that the total variation distance from the stationary distribution π is $\leq 1/4$. A mixing time of $O(n \log n)$ is referred to as an optimal mixing time as this matches the lower bound established by Hayes and Sinclair [HS07] for any graph of constant maximum degree Δ .

Jerrum [Jer95] (see also Salas and Sokal [SS97]) established an optimal mixing time of $O(n \log n)$ for the Glauber dynamics whenever $k > 2\Delta$. This was a seminal result in the development of coupling techniques, including the path coupling method of Bubley and Dyer [BD97]. Vigoda [Vig99] improved Jerrum's result to $k > (11/6)\Delta$ by proving $O(n \log n)$ mixing time of the following *flip dynamics*, which implied $O(n^2)$ mixing time of the Glauber dynamics.

The flip dynamics is a generalization of the Glauber dynamics, which recolors maximal two-colored components in each step. For a coloring X_t , vertex $v \in V$, and color $c \in \{1, \dots, k\}$, let $S_{X_t}(v, c)$ denote the set of vertices w which have an alternating path on colors $(X_t(v), c)$ between v and w ; we refer to the set $S_{X_t}(v, c)$ as a cluster. The flip dynamics is defined by a set of parameters $(P_i)_{i \geq 1}$ for the flip probabilities. The dynamics operates by choosing a random vertex v and color c , and then flipping the cluster $S_{X_t}(v, c)$ by interchanging the colors $X_t(v)$ and c on the chosen cluster with probability P_ℓ/ℓ where $\ell = |S_{X_t}(v, c)|$.

More formally, for $X_t \in \Omega$, the transitions $X_t \rightarrow X_{t+1}$ for the flip dynamics, with flip probabilities $(P_i)_{i \geq 1}$, are defined as follows:

- Choose $v \in V$ uniformly at random.
- Choose $c \in [k] = \{1, \dots, k\}$ uniformly at random.
- Let $S = S_{X_t}(v, c)$ be the cluster in X_t defined by colors $\{X_t(v), c\}$ which contains v . Let $\ell = |S|$ denote the size of the cluster.
- With probability P_ℓ/ℓ , let X_{t+1} denote the coloring obtained by interchanging colors $X_t(v)$ and c in S , and for all $w \notin S$ set $X_{t+1}(w) = X_t(w)$. With the remaining probability, set $X_{t+1} = X_t$.

Observe that if $P_1 = 1$ and $P_i = 0$ for all $i > 1$ then the flip dynamics only recolors the selected vertex if no neighbor has the chosen color, and is equivalent to the Glauber dynamics (see Jerrum [Jer95]). In Vigoda's original work, the parameters satisfy the basic properties: $P_1 = 1$, $P_i \geq P_{i+1}$ for all $i \geq 1$, and $P_j = 0$ for $j \geq 7$, see Section 3.1 for more details. All subsequent works (including this paper) follow these broad settings but differ in the detailed setting.

Since Vigoda's result, there was a myriad of improved results for various restricted classes of graphs, including optimal mixing on triangle-free graphs when $k > \alpha^* \Delta$ for $\Delta = O(1)$ where $\alpha^* \approx 1.763$ [CGSV21, FGYZ21, CLV21] (see also [JPV22, LSS19, DFHV13] for related results), optimal mixing on large girth graphs when $k \geq \Delta + 3$ [CLMM23], and further improvements for trees [MSW04], planar graphs [HVV15], and sparse random graphs [EHSV18].

The first improvement on Vigoda's result for general graphs was 20 years later by Chen, Delcourt, Moitra, Perarnau, and Postle [CDM⁺19] who proved the mixing of $O(n \log n)$ of the flip dynamics when $k > (11/6 - \varepsilon_0)\Delta$ where $\varepsilon_0 \approx 10^{-5}$ is a fixed positive constant. Their result (as well

as Vigoda's result [Vig99]) was obtained for a specific setting of the flip parameters. We present the first substantial improvement over Vigoda's result in obtaining optimal mixing of the flip dynamics when $k \geq 1.809\Delta$ for any Δ .

Theorem 1.1. *For all $\Delta \geq 125$, for all $k \geq 1.809\Delta$, there exists a setting of the parameters for the flip dynamics with $P_j = 0$ for all $j \geq 7$, so that for any graph G on n vertices with maximum degree Δ , the flip dynamics has mixing time $O(n \log(n))$.*

The flip probabilities are presented in Section 3.1. Note we use a universal setting of the flip probabilities for all $1.809 \leq k/\Delta$ that differ from those used in [CDM⁺19, Vig99].

As in [Vig99, CDM⁺19] this implies polynomial mixing of the Glauber dynamics for the same range of k/Δ . In particular, by comparison of the spectral gaps of the transition matrices, it implies $O(n^2)$ mixing time of the Glauber dynamics where the hidden constant is polynomial in Δ and k . Moreover, recent work of [BCC⁺22, Liu21] utilizing spectral independence [CLV21, ALO20], implies $O(n \log n)$ mixing time of the Glauber dynamics when Δ is constant; the same result held for the previous work of [CDM⁺19] for the corresponding range of parameters k and Δ .

Corollary 1.2. *For all $\Delta \geq 125$, all $k \geq 1.809\Delta$, there exists a constant $C = C(\Delta, k)$ such that for any and for any graph G on n vertices with maximum degree Δ , the mixing time of the Glauber dynamics is $\leq Cn \log n$.*

A recent algorithm of Chen, Feng, Guo, Zhang, and Zou [CFG⁺24] uses our coupling proof to obtain a deterministic approximate counting algorithm (FPTAS) when k and Δ are constant, see Remark 4.13 for more details.

1.1 Proof Overview

Our proof of Theorem 1.1 utilizes a novel distance metric described here at a high level.

Jerrum's bound of $k > 2\Delta$ can be proved using path coupling in which we consider a pair of configurations X_t, Y_t that differ at exactly one vertex, say v^* . We then analyze the expected Hamming distance of X_{t+1}, Y_{t+1} after one step of a coupled transition $(X_t, Y_t) \rightarrow (X_{t+1}, Y_{t+1})$. The coupling in this setting is fairly simple as it is the identity coupling (i.e., both chains attempt the same (vertex, color) pair (v, c)) for the Glauber update except when the updated vertex v is a neighbor w of v^* ; in which case we couple trying to recolor w to color $X_t(v^*)$ in one chain with color $Y_t(v^*)$ in the other chain. Under this coupling, there is at most one coupled transition per neighbor which can increase the Hamming distance by at most one, and this yields the $k > 2\Delta$ bound.

Vigoda's result for $k > (11/6)\Delta$ uses the same path coupling framework to analyze the expected Hamming distance for a pair X_t, Y_t that differ at a single vertex v^* . In flip dynamics, the coupling is more complicated than in Jerrum's analysis because of additional moves.

Recall that the transitions of the flip dynamics correspond to flipping maximal 2-colored components, where flipping refers to interchanging the 2 colors. An alternative and equivalent view of the transitions of the flip dynamics is as follows. For $X_t \in \Omega$, consider the collection of all clusters (where a cluster is a maximal 2-colored component); note that there are at most nk clusters. Choose a cluster S with probability $P_{|S|}/(nk)$ and then flip S to obtain X_{t+1} (with the remaining probability set $X_{t+1} = X_t$).

We give a brief high-level overview of Vigoda's coupling $(X_t, Y_t) \rightarrow (X_{t+1}, Y_{t+1})$ which is the one-step coupling that for a Hamming distance one pair of cooringss minimizes the expected Hamming distance at time $t+1$ and hence is called the greedy coupling in [CDM⁺19]. Consider a pair X_t, Y_t that differ at a single vertex v^* , thus $X_t(w) = Y_t(w)$ for all $w \neq v^*$. The coupling uses the identity

coupling for all clusters S that are the same in chains X_t and Y_t . This means that the cluster S is flipped in both chains or in neither chain. The only nontrivial couplings are for those clusters that appear in only one chain (and not in the other chain).

What are these clusters that appear in exactly one chain? They are clusters that include a neighbor w of v^* and a color c' which is $X_t(v^*)$ or $Y_t(v^*)$. In other words, if we try to recolor some $w \in N(v^*)$ to a color $c' \in \{X_t(v^*), Y_t(v^*)\}$ then this yields a different cluster in the two chains, as the cluster includes v^* in one chain but not in the other chain. These are the clusters that use a nontrivial coupling. Moreover, this coupling depends on the current color $c = X_t(w) = Y_t(w)$ of the neighbor w of v^* ; we partition the clusters involving color c into a set $\mathcal{D}_{t,c}$ and clusters within $\mathcal{D}_{t,c}$ are coupled with other clusters in the same set $\mathcal{D}_{t,c}$.

Vigoda demonstrated a choice of the flip probabilities and a coupling so that the expected increase in the Hamming distance is $\leq 5/6$ in an amortized cost per neighbor w of v^* , which yields the bound $k > (11/6)\Delta$. Chen et al. [CDM⁺19] identified 6 extremal configurations in Vigoda's analysis (these are the configurations that maximize the expected increase in the Hamming distance) and presented a slightly different setting of the flip probabilities with only two extremal configurations. They considered a weighted Hamming distance where for every neighbor $w \in N(v^*)$, if the local configuration around w is different from the two extremal configurations, then the definition of the distance (between X_t and Y_t) is decreased by η/Δ for a fixed constant η where $1/2 > \eta > 0$. Using Vigoda's greedy coupling with this new metric, they established that the expected distance decreases when $k > (11/6 - \varepsilon_0)\Delta$ where $\varepsilon_0 \approx 10^{-5}$.

We take a complementary approach. Whereas Chen et al. [CDM⁺19] reweight the worst configurations (or equivalently all non-worst case configurations), we instead consider a particularly “good” configuration. Namely, we consider the local configuration where the neighbor w is *unblocked*, which means that the colors $X_t(v^*)$ and $Y_t(v^*)$ do not appear in $N(w) \setminus \{v^*\}$, see Fig. 1. Unblocked neighborhoods are the best local configurations for the greedy coupling (with respect to the Hamming distance). For unblocked neighbors w , the potentially problematic recolorings of $w \in N(v^*)$ with colors $X_t(v^*)$ or $Y_t(v^*)$ can be coupled with flips of clusters of size 2 (containing v^* and w). Consequently, the expected increase in the Hamming distance for w is $1 - P_2$, significantly less than $5/6$ for any setting of the flip probabilities considered.

We choose the flip probabilities to optimize for this new metric, which results in a setting for the flip probabilities that are suboptimal (with respect to Hamming distance) for the extremal configurations considered by Chen et al. [CDM⁺19]. However, for our choice of distance metric, these extremal configurations improve as they have a reasonable probability of moving to the unblocked configurations, yielding a decrease in the distance.

An essential aspect of our proof is that if a neighbor w is unblocked, then it may have a considerable probability of becoming blocked (which increases the distance with respect to our new metric); however, in this case, if w itself becomes a disagreement, then its neighbors will be unblocked (and hence this new disagreement at w has a smaller weight in our new metric). This is the crucial trade-off in our argument: For an unblocked neighbor w , either w is unlikely to become blocked, or if it becomes a disagreement, it has many unblocked neighbors. Our overall proof is of a similar technical level of difficulty as in [CDM⁺19] but yields a substantially improved bound of $k \geq 1.809\Delta$.

We present some basic definitions, including a more formal definition of the flip dynamics, and the path coupling lemma in Section 2. In Section 3 we define our new metric. We then present Vigoda's greedy coupling, analyze our new metric, and prove Theorem 1.1 in Section 4.

2 Preliminaries

In this section we detail the basic definitions and concepts that are required background for our proofs.

For a graph $G = (V, E)$ with vertex set V and edge set E , for $v \in V$, let $N(v)$ denote the neighbors of v , and let $d(v) = |N(v)|$ denote the degree of v in G . Let $\Delta = \max_{v \in V} d(v)$ be the maximum degree.

2.1 List Colorings

We prove our results in the more general context of list colorings. Fix a graph $G = (V, E)$ and a list $L(v)$ of colors for each $v \in V$. A *list labeling* of G is a function σ that maps each vertex $v \in V$ to a color $\sigma(v) \in L(v)$. A list labeling is called a *list coloring* if for all $(u, v) \in E$, $\sigma(u) \neq \sigma(v)$. For any positive integer k , let $[k] = \{1, \dots, k\}$. Observe that if $L(v) = [k]$ for all $v \in V$, then a list labeling is a k -labeling and a list coloring is a k -coloring.

In the remainder of this paper we will work with the general concept of list colorings and list labelings. Let $\mathcal{L} = \bigcup_{v \in V} L(v)$, let Ω be the collection of all list labels of G , and let $\Omega^* \subset \Omega$ be the collection of list colorings.

For a pair $X_t, Y_t \in \Omega$, let $H(X_t, Y_t) = |\{v \in V : X_t(v) \neq Y_t(v)\}|$ denote the Hamming distance between X_t and Y_t . For $v \in V$, let $\Omega_v^2 \subset \Omega^2$ denote the pairs $(X_t, Y_t) \in \Omega^2$ where $X_t(v) \neq Y_t(v)$ and $X_t(u) = Y_t(u)$ for all $u \neq v$; thus, Ω_v^2 is the set of pairs of labelings that only differ at vertex v . Note, $\bigcup_{v \in V} \Omega_v^2$ is the set of all pairs that differ at exactly one vertex. We will refer to pairs (X_t, Y_t) that differ at exactly one vertex as neighboring colorings. For simplicity, we use the term coloring throughout since the distinction between colorings and labelings is clear from the notation Ω vs. Ω^* .

2.2 Clusters, Flip Dynamics, and Mixing Time

For a coloring $X_t \in \Omega$, vertex $v \in V$, and color $c \in \mathcal{L}$, let $\widehat{S}_{X_t}(v, c)$ denote the set of vertices reachable from v by a $(X_t(v), c)$ -alternating path in X_t . Note that X_t may not be a proper coloring, but we still require that the colors alternate along the path. If for all $w \in \widehat{S}_{X_t}(v, c)$ it holds that $\{c, X_t(v)\} \subset L(w)$ then we set $S_{X_t}(v, c) = \widehat{S}_{X_t}(v, c)$, and otherwise we set $S_{X_t}(v, c) = \emptyset$. Note that if the flip of the cluster $\widehat{S}_{X_t}(v, c)$ is invalid (namely, for some $w \in \widehat{S}_{X_t}(v, c)$ the new color is not in its list $L(w)$) then we have set the cluster to be the empty set. Hence, all clusters $S_{X_t}(v, c) \neq \emptyset$ can be flipped in X_t (i.e., if $X_t \in \Omega$ then $X_{t+1} \in \Omega$ where X_{t+1} is obtained from X_t by interchanging the colors $X_t(v)$ and c on the set $S_{X_t}(v, c)$).

For all $v \in V$, note that $S_{X_t}(v, X_t(v)) = S_{Y_t}(v, Y_t(v)) = \{v\}$ and if $c \notin L(v)$ then $S_{X_t}(v, c) = S_{Y_t}(v, c) = \emptyset$. For a coloring $X_t \in \Omega$ and a cluster $S = S_{X_t}(v, c)$, we refer to flipping cluster S with the operation of interchanging colors $X_t(v)$ and c on the set S ; let X_{t+1} denote the resulting coloring. If X_t is a proper coloring, then X_{t+1} is a proper coloring. Moreover, if X_t, X_{t+1} are proper colorings then by flipping $S' = S_{X_{t+1}}(v, X_t(v))$ in X_{t+1} we obtain X_t , and hence the operation is symmetric on Ω^* .

We can now define the flip dynamics for the more general setting of list colorings. Consider probabilities $(P_i)_{i \geq 1}$. For $X_t \in \Omega$ the transitions $X_t \rightarrow X_{t+1}$ of the flip dynamics are defined as follows:

- Choose $v \in V$ uniformly at random.
- For each $c \in L(v)$, let $S = S_{X_t}(v, c)$ with probability $1/k$. Let $\ell = |S|$ denote the size of the cluster.

- If $\ell \geq 1$ then with probability P_ℓ/ℓ , let X_{t+1} denote the coloring obtained by interchanging colors $X_t(v)$ and c in S (and for $w \notin S$ set $X_{t+1}(w) = X_t(w)$).
- Otherwise, let $X_{t+1} = X_t$.

Note, that in the second step, if the flip of the cluster $\hat{S}_{X_t}(v, c)$ is invalid then the corresponding set $S_{X_t}(v, c) = \emptyset$. Moreover, when $P_1 > 0$ and $k \geq \Delta + 2$, then the unique stationary distribution of the flip dynamics is the uniform distribution π over Ω^* , which is the set of proper list colorings; note, the states in $\Omega \setminus \Omega^*$ have zero probability in the stationary distribution.

Our interest is the mixing time T_{mix} , which measures the speed of convergence to the unique stationary distribution π from the worst initial state $X_0 \in \Omega$. For a pair of distributions μ, π on Ω , the total variation distance is defined as $d_{\text{TV}}(\mu, \pi) = \frac{1}{2} \sum_{X_t \in \Omega} |\mu(X_t) - \pi(X_t)|$. For $\varepsilon > 0$, define the mixing time as:

$$T_{\text{mix}}(\varepsilon) = \max_{X_0 \in \Omega} \min\{t : d_{\text{TV}}(P^t(X_0, \cdot), \pi) \leq \varepsilon\},$$

where π is the stationary distribution. We will often refer to $T_{\text{mix}} = T_{\text{mix}}(1/4)$ as the mixing time since $T_{\text{mix}}(\varepsilon) \leq T_{\text{mix}} \times \lceil \log_2(1/\varepsilon) \rceil$ for any $\varepsilon > 0$.

2.3 Path Coupling

We will utilize the coupling method to upper-bound the mixing time. For a pair of states $X_t, Y_t \in \Omega$, a coupling for the flip dynamics is a joint evolution $(X_t, Y_t) \rightarrow (X_{t+1}, Y_{t+1})$ such that when the individual transitions $(X_t \rightarrow X_{t+1})$ and $(Y_t \rightarrow Y_{t+1})$ are viewed in isolation of each other then they are identical to the flip dynamics, see [Jer03] for a more detailed introduction.

We will bound the mixing time using the path coupling framework of Bubley and Dyer [BD97].

Theorem 2.1. [BD97, DG99] *Consider a Markov chain with transition matrix P , state space Ω , and unique stationary distribution π . Let $\Sigma \subset \Omega^2$ denote a subset of pairs of states such that the graph (Ω, Σ) is connected. Consider weights $w(X_t, Y_t)$ defined for all pairs $(X_t, Y_t) \in \Sigma$. Assume there exists a constant $1/2 \leq C \leq 1$ where $w(X_t, Y_t) \in [C, 1]$ for all $(X_t, Y_t) \in \Sigma$. For arbitrary pairs $(X_{t+1}, Y_{t+1}) \in \Omega^2$, define $w(X_{t+1}, Y_{t+1})$ by the length of the shortest path in the graph (Ω, Σ) where edges $(X_t, Y_t) \in \Sigma$ have weight $w(X_t, Y_t)$.*

If there exists a $\delta > 0$ and for all $(X_t, Y_t) \in \Sigma$ there exists a coupling $(X_t, Y_t) \rightarrow (X_{t+1}, Y_{t+1})$ where:

$$\mathbf{E}[w(X_{t+1}, Y_{t+1})] \leq (1 - \delta)w(X_t, Y_t),$$

then the mixing time is bounded as $T_{\text{mix}}(\varepsilon) \leq O(\log(n/\varepsilon)/\delta)$.

3 New Metric Definition

This section introduces our new metric, which is the heart of the proof of [Theorem 1.1](#). First, we describe our flip probabilities and identify some of their key properties. Then, we introduce some new notation. Finally, we define the new metric.

3.1 Cluster and Flip Dynamics

To prove [Theorem 1.1](#), we use the following setting of flip probabilities, which we will refer to as [Parameter Setting 1](#):

$$P_1 := 1, P_2 := 0.324, P_3 := 0.154, P_4 := 0.088, P_5 := 0.044, P_6 := 0.011, \text{ and } P_i := 0 \text{ for } i \geq 7. \quad (1)$$

We also include the following variable which we will use when defining our new metric:

$$\eta := (P_2 - P_3) \frac{\Delta}{2k}. \quad (2)$$

We will assume the following properties, which we will refer to as [Parameter Properties 3.1](#):

$$\begin{aligned} P_1 &= 1, \quad P_2 \leq 1/3, \text{ and } P_7 = 0 & (\text{FP0}) \\ P_j &\leq (2/3)P_{j-1} \text{ for all } j \geq 3, & (\text{FP1}) \\ (1 - P_2) &\geq (P_2 - P_3) \geq 2(P_3 - P_4) & (\text{FP2}) \\ 2(P_3 - P_4) &\geq (j-1)(P_j - P_{j+1}) \text{ for } j \geq 4 & (\text{FP3}) \\ (P_4 - P_5) &\geq (P_5 - P_6) \geq (P_6 - P_7) & (\text{FP4}) \\ 2P_2 &\leq 1 - 4P_4 & (\text{FP5}) \\ 2P_3 &\leq 4P_4 - P_5 & (\text{FP6}) \\ \eta &\leq (6/19)P_2 & (\text{FP7}) \end{aligned}$$

Note that these properties hold for [Parameter Setting 1](#) when $k \geq 9/5$ and the settings originally considered by Vigoda. As we will see later in our analysis, this assumption allows us to quickly identify those initial configurations with the worst expected change for our new metric (introduced in the following sections). The relevant lemma statements include any further assumptions on the flip probabilities.

Our setting for the flip probabilities also differs from those in previous works [Vig99, CDM⁺19]. One of the key differences is that the previous work sets $P_3 \approx 1/6$. In particular, the setting $P_3 = 1/6$ was critical in [Vig99] and any setting of $P_3 \neq 1/6$ yields a worse bound of $k > C\Delta$ for a constant $C > 11/6$ using Vigoda's analysis. In [CDM⁺19] they also fix $P_3 = 1/6$ for their analysis with a modified metric; the argument using a variable length coupling sets slightly above 1/6, namely $P_3 = 0.166762 > 1/6$. In contrast, our setting of $P_3 < 1/6$ is somewhat counterintuitive at first glance as it increases the expected change in Hamming distance for the extremal configurations in Vigoda's analysis, but the introduction of our new metric offsets this effect.

3.2 Our New Metric

Before we introduce our new metric, we need several new definitions. For an integer $s \geq 0$, consider a pair of colorings $X_s, Y_s \in \Omega$; note this pair X_s, Y_s may differ at an arbitrary number of vertices.

For a vertex $z \in V$ such that $X_s(z) \neq Y_s(z)$, we partition the neighbors of z based on how many occurrences of the colors $X_s(z), Y_s(z)$ occur in their neighborhood (besides at z). For a vertex $y \in N(z)$ and $(\sigma, \tau) \in \Omega_z^2$, let

$$B^z(y, \sigma, \tau) := \{w \in N(y) \setminus \{z\} : \{\sigma(w), \tau(w)\} \cap \{\sigma(z), \tau(z)\} \neq \emptyset\}$$

as the blocking neighbors of y with respect to z in σ and τ . For ease of notation, since we are often considering a pair of chains (X_t) and (Y_t) , we simplify the notation as follows. For integer $s \geq 0$,

$$B_s^z(y) := B^z(y, X_s, Y_s).$$

We say that a neighbor $y \in N(z)$ is *unblocked* with respect to z if it has no blocking neighbors (i.e. $|B_s^z(y)| = 0$), *singly blocked* with respect to z if there is exactly one blocking neighbor (i.e. $|B_s^z(y)| = 1$), and *multiblocked* otherwise (i.e. $|B_s^z(y)| \geq 2$).

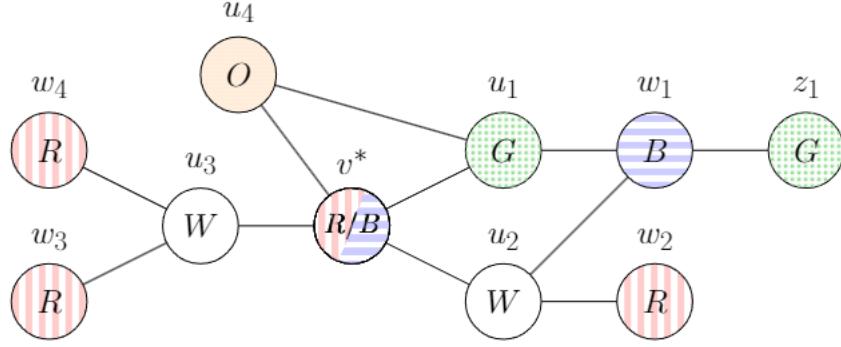


Figure 1: A small graph showing how vertices can be unblocked, singly blocked, and multiblocked. The vertex u_1 is singly blocked with respect to v^* since it has a neighbor w_1 that is colored B . The vertex u_2 is multiblocked with respect to v^* since it has a B and R neighbor, w_1 and w_2 respectively. Likewise, u_3 is multiblocked since it has two R neighbors, w_3 and w_4 . Finally, u_4 is unblocked since it has no neighbors that are R or B .

For integer $i \geq 0$ and $(\sigma, \tau) \in \Omega_z^2$, let

$$F^i(z, \sigma, \tau) := \{y \in N(z) : |B^z(y, \sigma, \tau)| = i\} \text{ and } F^{\geq i}(z, \sigma, \tau) := \bigcup_{j \geq i} F^j(z, \sigma, \tau)$$

be the neighbors of z with exactly i (or at least i) blocking neighbors with respect to σ and τ . Moreover, if $\sigma(z) = \tau(z)$ then let $F^i(z, \sigma, \tau) = \emptyset$. Let

$$d^i(z, \sigma, \tau) := |F^i(z)| \text{ and } d^{\geq i}(z, \sigma, \tau) := \sum_{j \geq i} d^j(z, \sigma, \tau)$$

be the number of neighbors of z with exactly i (or at least i) blocking neighbors with respect to σ and τ . For ease of notation, since we are often considering a pair of chains (X_t) and (Y_t) we simplify the notation as follows. For integers $i \geq 0$ and $s \geq 0$, if $X_s(z) \neq Y_s(z)$ then let

$$F_s^i(z) := F^i(z, X_s, Y_s) \text{ and } F_s^{\geq i}(z) := F^{\geq i}(z, X_s, Y_s).$$

Moreover, if $X_s(z) = Y_s(z)$ then let $F_s^i(z) = \emptyset$. Finally, let

$$d_s^i(z) := d^i(z, X_s, Y_s) \text{ and } d_s^{\geq i}(z) := d_s^{\geq i}(z, X_s, Y_s).$$

Notice that $F_s^0(z)$ is the set of unblocked neighbors of z in (X_s, Y_s) , $F_s^1(z)$ is the set of singly blocked neighbors of z , and $F_s^{\geq 2}(z)$ is the set of multiblocked neighbors of z .

We can now formally define our new metric. Recall the definition of the constant η from Eq. (2). Consider an arbitrary pair $X_t, Y_t \in \Omega$. We first define for a vertex $z \in V$,

$$H_z(X_t, Y_t) := \begin{cases} 1 & \text{if } X_t(z) \neq Y_t(z) \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

and

$$\tilde{H}_z(X_t, Y_t) := \begin{cases} 1 - \frac{\eta}{\Delta} d_t^0(z) & \text{if } X_t(z) \neq Y_t(z) \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

If $(X_t, Y_t) \in \Omega_{v^*}^2$ then we define the distance between these neighboring colorings as follows:

$$\tilde{H}(X_t, Y_t) := \sum_{z \in V} \tilde{H}_z(X_t, Y_t) = 1 - \frac{\eta}{\Delta} d_t^0(v^*). \quad (5)$$

Extend \tilde{H} to define a metric \tilde{H} over all pairs in Ω^2 by considering the path metric defined by the shortest path distance in the graph (Ω, Σ) where neighboring colorings have weight defined by Eq. (5).

The following lemma upper bounds \tilde{H} by the Hamming metric H and the number of unblocked neighbors. The bound is tight if the two states differ at exactly one vertex, v . Recall that if $X_t(u) = Y_t(u)$ then $F_t^0(u) = \emptyset$ and hence $d_t^0(u) = 0$. This lemma has nothing to do with the coupling; it means that the path metric \tilde{H} , which we implicitly defined for pairs (X_t, Y_t) that differ at more than one vertex, can be bounded naturally by the Hamming distance and the number of unblocked neighbors of disagreements in X_t, Y_t .

Lemma 3.1. *For any $X_t, Y_t \in \Omega$,*

$$\tilde{H}(X_t, Y_t) \leq \sum_{z \in V} \tilde{H}_z(X_t, Y_t).$$

Proof. Let $U = \{u \in V : Y_t(u) \neq X_t(u)\}$ be the set of vertices that X_t and Y_t disagree on, and let $U = \{u_1, \dots, u_{|U|}\}$ be an arbitrary ordering of U . Let $\rho_0 = X_t$, $\rho_{|U|} = Y_t$, and for $1 \leq i \leq |U| - 1$, let $\rho_i(v) = X_t(v)$ for $v \notin U$, $\rho_i(u_j) = \rho_{i-1}(u_j)$ for $j \neq i$, and $\rho_i(u_i) = Y_t(u_i)$. It follows that

$$\tilde{H}(X_t, Y_t) \leq \sum_i \tilde{H}(\rho_{i-1}, \rho_i) \quad (6)$$

since $\tilde{H}(X_t, Y_t)$ is defined to be the length of the shortest path between X_t and Y_t and the path $\rho_0, \dots, \rho_{|U|}$ is a particular path. We will prove that for all $u_i \in U$,

$$d^0(u_i, \rho_{i-1}, \rho_i) \geq d^0(u_i, X_t, Y_t). \quad (7)$$

Assuming Eq. (7) and making use of Eq. (6) we can conclude the lemma as follows:

$$\begin{aligned} \tilde{H}(X_t, Y_t) &\leq \sum_i \tilde{H}(\rho_{i-1}, \rho_i) && \text{(by Eq. (6))} \\ &= \sum_{u_i \in U} \left(1 - \frac{\eta}{\Delta} d^0(u_i, \rho_{i-1}, \rho_i)\right) && \text{(by Eq. (5))} \\ &= H(X_t, Y_t) - \frac{\eta}{\Delta} \sum_{u_i \in U} d^0(u_i, \rho_{i-1}, \rho_i) && (|U| = H(X_t, Y_t)) \\ &\leq H(X_t, Y_t) - \frac{\eta}{\Delta} \sum_{u_i \in U} d^0(u_i, X_t, Y_t). && \text{(by Eq. (7))} \end{aligned}$$

It remains to prove Eq. (7), for which it suffices to show that, for all $u_i \in U$,

$$F^0(u_i, \rho_{i-1}, \rho_i) \supseteq F^0(u_i, X_t, Y_t).$$

Fix i and suppose $w \in F^0(u_i, X_t, Y_t)$. We will show that $w \in F^0(u_i, \rho_{i-1}, \rho_i)$. By the definition of $F^0(u_i, X_t, Y_t)$, if $w \in F^0(u_i, X_t, Y_t)$ then $B^{u_i}(w, X_t, Y_t) = \emptyset$. Consider $z \in N(w) \setminus \{u_i\}$. By the

definition of $B^{u_i}(w, X_t, Y_t)$, since we know that $B^{u_i}(w, X_t, Y_t) = \emptyset$ and $z \in N(w) \setminus \{u_i\}$ then we have that:

$$\{X_t(z), Y_t(z)\} \cap \{X_t(u_i), Y_t(u_i)\} = \emptyset.$$

Recall that by construction, we have the following:

$$\rho_{i-1}(z) \in \{X_t(z), Y_t(z)\}, \rho_i(z) \in \{X_t(z), Y_t(z)\}, \rho_{i-1}(u_i) = X_t(u_i), \text{ and } \rho_i(u_i) = Y_t(u_i).$$

Thus, for all $z \in N(w) \setminus \{u_i\}$,

$$\{\rho_{i-1}(z), \rho_i(z)\} \cap \{\rho_{i-1}(u_i), \rho_i(u_i)\} \subset \{X_t(z), Y_t(z)\} \cap \{X_t(u_i), Y_t(u_i)\} = \emptyset.$$

Since this holds for all $z \in N(w) \setminus \{u_i\}$, we have shown that $w \in F^0(u_i, \rho_{i-1}, \rho_i)$ as desired, which completes the proof of the lemma. \square

4 Coupling Analysis

We start our analysis by giving a brief description of the greedy coupling. The coupling is referred to as the greedy coupling as it minimizes the expected (unweighted) Hamming distance after the coupled move for initial pairs of configurations that differ at a single vertex.

To define the greedy coupling for the flip dynamics, let us first observe an alternative formulation of the dynamics. For a state $X_t \in \Omega$, every cluster S in X_t has an associated flip probability $P_{|S|}$. To simulate the flip dynamics we choose a cluster S with probability $P_{|S|}/(nk)$ and then flip it to obtain the new state X_{t+1} , and with the remaining probability, we stay in the same state $X_{t+1} = X_t$.

If a cluster S exists in both chains, the greedy coupling will flip S in both chains or in neither chain; this is referred to as the identity coupling. The only non-identity coupling is for clusters that potentially differ in the two chains. These potential disagreeing clusters involve v^* or neighbors of v^* and can be partitioned according to the current colors of the neighbors of v^* .

We now partition the neighbors of a vertex $z \in V$ on a per color basis. For any coloring $X_s \in \Omega$, color $c \in \mathcal{L}$, and vertex $z \in V$, let

$$N_c(z, X_s) := \{w \in N(z) : X_s(w) = c\}$$

be the neighbors of z that are color c in X_s . We extend this notation for an arbitrary pair $X_s, Y_s \in \Omega$ by letting

$$N_{s,c}(z) := N_c(z, X_s) \cup N_c(z, Y_s).$$

Note that if $(X_t, Y_t) \in \Omega_{v^*}^2$ then $N_{t,c}(v^*) = N_c(v^*, X_t) = N_c(v^*, Y_t)$. Let

$$d_{t,c}(v^*) := |N_{t,c}(v^*)|$$

be the number of neighbors of v^* that are colored c in X_t or Y_t .

Now we will give a brief overview of the greedy coupling. Fix a pair $(X_t, Y_t) \in \Omega_{v^*}^2$. The clusters involving color $c \in \mathcal{L}$ that we need to couple (using the greedy coupling) are the following:

$$\mathcal{D}_{t,c} := \{S_{X_t}(v^*, c)\} \cup \{S_{Y_t}(v^*, c)\} \cup \bigcup_{u \in N_{t,c}(v^*)} \{S_{X_t}(u, Y_t(v^*)), S_{Y_t}(u, X_t(v^*))\}. \quad (8)$$

Let us digest this collection of clusters $\mathcal{D}_{t,c}$. Consider the case when $d_{t,c}(v^*) = 0$, then $\mathcal{D}_{t,c}$ consists of two clusters $S_{X_t}(v^*, c)$ and $S_{Y_t}(v^*, c)$, and both of these clusters are the same:

$S_{X_t}(v^*, c) = S_{Y_t}(v^*, c) = \{v^*\}$ since color c does not appear in the neighborhood of v^* . These clusters are coupled with the identity coupling, which means that we flip the cluster in both chains or neither chain. Note that if we flip these clusters $S_{X_t}(v^*, c) = S_{Y_t}(v^*, c)$ when $d_{t,c}(v^*) = 0$ then the resulting colorings are the same $X_{t+1} = Y_{t+1}$ (as $X_{t+1}(v^*) = Y_{t+1}(v^*) = c$); these are the “good” moves which decrease the Hamming distance by one.

The non-trivial case is where $d_{t,c}(v^*) \geq 1$. The clusters of $\mathcal{D}_{t,c}$ which occur in X_t , are the $S_{X_t}(w, Y_t(v^*))$ cluster for every $w \in N_{t,c}(v^*)$ and the $S_{X_t}(v^*, c)$ cluster; and in Y_t we have the $S_{Y_t}(w, X_t(v^*))$ for $w \in N_{t,c}(v^*)$ and the $S_{Y_t}(v^*, c)$ cluster. Notice that these two (v^*, c) clusters are large clusters that consist of the union of the other small clusters plus v^* , namely,

$$S_{X_t}(v^*, c) = \{v^*\} \cup \bigcup_{w \in N_{t,c}(v^*)} S_{Y_t}(w, X_t(v^*)) \quad \text{and} \quad S_{Y_t}(v^*, c) = \{v^*\} \cup \bigcup_{w \in N_{t,c}(v^*)} S_{X_t}(w, Y_t(v^*)).$$

The only non-identity coupling involves clusters in $\mathcal{D}_{t,c}$ for some c where $d_{t,c}(v^*) \geq 1$. The clusters in $\mathcal{D}_{t,c}$ are coupled with each other (or with nothing corresponding to a self-loop in the other chain); the greedy coupling in these cases is detailed in [Section 5.1](#).

We define the set of vertices besides v^* that are contained in a cluster of $\mathcal{D}_{t,c}$ as

$$\Lambda_{t,c} := \{u \in V \setminus \{v^*\} : \text{there exists } S \in \mathcal{D}_c(X_t, Y_t) \text{ such that } u \in S\}.$$

That is, $\Lambda_{t,c}$ is every vertex besides v^* that can be reached from v^* with a $(c, X_t(v^*))$ or $(c, Y_t(v^*))$ alternating path. These are precisely the vertices where new disagreements can form after a single step of the greedy coupling. Observe that $N_{t,c}(v^*) \subseteq \Lambda_{t,c}$.

4.1 Relating \tilde{H} to H

Fix a pair $(X_t, Y_t) \in \Omega_{v^*}^2$. Let

$$W_t := \mathbf{E}[H(X_{t+1}, Y_{t+1}) - H(X_t, Y_t)]$$

and

$$\tilde{W}_t := \mathbf{E}[\tilde{H}(X_{t+1}, Y_{t+1}) - \tilde{H}(X_t, Y_t)] \tag{9}$$

denote the expected change over one step of the greedy coupling, $(X_t, Y_t) \rightarrow (X_{t+1}, Y_{t+1})$, for H and \tilde{H} respectively. Note, the quantities W_t and \tilde{W}_t are the expected change for the Hamming and weighted Hamming distance, respectively, between the chains (X_t) and (Y_t) for the update at time t ; both of these quantities W_t and \tilde{W}_t are with respect to the same pair of chains (X_t) and (Y_t) . Moreover, the chains (X_t) and (Y_t) are coupled using the greedy coupling, which minimizes the expectation of the unweighted Hamming distance $H(X_{t+1}, Y_{t+1})$; the greedy coupling is not necessarily optimal for the weighted Hamming distance $\tilde{H}(X_{t+1}, Y_{t+1})$. The following subsections aim to bound \tilde{W}_t by decomposing it with respect to each color.

4.2 Analysis by Color

We want to decompose the expected change in the Hamming distance W_t and our metric \tilde{W}_t with respect to each color c .

Recall H_z is the Hamming distance at a vertex z , see [Eq. \(3\)](#). Fix colorings $X_s, Y_s \in \Omega$. We define the Hamming distance with respect to an arbitrary subset $R \subseteq V$ as

$$H_R(X_s, Y_s) := \sum_{z \in R} H_z(X_s, Y_s).$$

Similarly, we define the new metric distance with respect to an arbitrary subset $R \subseteq V$ and vertex $z \notin R$ as

$$\tilde{H}_{z,R}(X_s, Y_s) := H_R(X_s, Y_s) - \frac{\eta}{\Delta} \left(|F_s^0(z) \cap R| + \sum_{y \in R} |F_s^0(y)| \right). \quad (10)$$

We will give some intuition for the definition of $\tilde{H}_{z,R}(X_s, Y_s)$ in Eq. (10) after the statement of Lemma 4.1 below.

We can now define for all pairs $(X_t, Y_t) \in \Omega_{v^*}^2$ and for all c such that $d_{t,c}(v^*) \geq 1$,

$$W_t^c := \mathbf{E} [H_{\Lambda_{t,c}}(X_{t+1}, Y_{t+1}) - H_{\Lambda_{t,c}}(X_t, Y_t)]$$

and

$$\tilde{W}_t^c := \mathbf{E} [\tilde{H}_{v^*, \Lambda_{t,c}}(X_{t+1}, Y_{t+1}) - \tilde{H}_{v^*, \Lambda_{t,c}}(X_t, Y_t)] \quad (11)$$

as the expected change concerning $\Lambda_{t,c}$, which are the vertices related to color c . Notice that both terms for W_t^c (and also \tilde{W}_t^c) the set $R = \Lambda_{t,c}$ which is defined for time t .

We will show that if we bound these new functions W_t^c and \tilde{W}_t^c for every color c such that $d_{t,c}(v^*) \geq 1$, then we obtain an upper bound on the total change as follows:

Lemma 4.1. *For $(X_t, Y_t) \in \Omega_{v^*}^2$,*

$$\tilde{W}_t \leq \mathbf{E} [H_{v^*}(X_{t+1}, Y_{t+1}) - H_{v^*}(X_t, Y_t)] + \sum_{\substack{c \in \mathcal{L}: \\ d_{t,c}(v^*) \geq 1}} \tilde{W}_t^c.$$

In Lemma 4.1, the $\mathbf{E} [H_{v^*}(X_{t+1}, Y_{t+1}) - H_{v^*}(X_t, Y_t)]$ term is capturing the change in Hamming distance at v^* . For the change in the new metric at v^* we also need to capture the change in the number of unblocked neighbors of v^* ; these terms are considered based on the colors of the neighbors of v^* and are captured in the second term in Eq. (10), namely $|F_t^0(v^*) \cap \Lambda_{t,c}|$. Finally, the change in the new metric for all other vertices (besides v^*) are also considered on a per color basis and captured by the last summation in Eq. (10).

Proof of Lemma 4.1. Since $(X_t, Y_t) \in \Omega_{v^*}^2$ and $\eta \leq 1/2$ we have that:

$$\begin{aligned} \tilde{H}(X_t, Y_t) &= \tilde{H}_{v^*}(X_t, Y_t) \\ &= H_{v^*}(X_t, Y_t) + \sum_{\substack{c \in \mathcal{L}: \\ d_{t,c}(v^*) \geq 1}} \tilde{H}_{v^*, \Lambda_{t,c}}(X_t, Y_t). \end{aligned} \quad (12)$$

For a vertex $u \in V$, observe that if $u \neq v^*$ then $X_t(u) = Y_t(u)$, and if for all colors $c \in L(\cup u)$ we have $u \notin \Lambda_{t,c}$ then every cluster that contains u is in both X_t and Y_t and hence $X_{t+1}(u) = Y_{t+1}(u)$, since the greedy coupling either flips a cluster containing u in both chains or neither (see Section 5.1). Therefore, $H_u(X_{t+1}, Y_{t+1}) = 0$ and $\tilde{H}_u(X_{t+1}, Y_{t+1}) = 0$ by definition. In summary, we have the following:

$$\sum_{\substack{u \notin \cup_c \Lambda_{t,c}: \\ u \neq v^*}} \tilde{H}_u(X_{t+1}, Y_{t+1}) = 0. \quad (13)$$

This yields the following decomposition of the new metric at time $t + 1$:

$$\begin{aligned}
\tilde{H}(X_{t+1}, Y_{t+1}) &\leq \sum_{u \in V} \tilde{H}_u(X_{t+1}, Y_{t+1}) && \text{(by Lemma 3.1)} \\
&\leq \tilde{H}_{v^*}(X_{t+1}, Y_{t+1}) + \sum_{\substack{c \in \mathcal{L}: \\ d_{t,c}(v^*) \geq 1}} \sum_{u \in \Lambda_{t,c}} \tilde{H}_u(X_{t+1}, Y_{t+1}) && \text{(by Eq. (13))} \\
&= H_{v^*}(X_{t+1}, Y_{t+1}) + \sum_{\substack{c \in \mathcal{L}: \\ d_{t,c}(v^*) \geq 1}} \tilde{H}_{v^*, \Lambda_{t,c}}(X_{t+1}, Y_{t+1}) && \text{(by Definition 10)} \quad (14)
\end{aligned}$$

where the second inequality is in fact an equality for a graph with sufficiently high girth.

We can now use Eqs. (12) and (14) to finish the proof:

$$\begin{aligned}
\widetilde{W}_t &= \mathbf{E} \left[\tilde{H}(X_{t+1}, Y_{t+1}) - \tilde{H}(X_t, Y_t) \right] && \text{(by Definition 9)} \\
&\leq \mathbf{E} [H_{v^*}(X_{t+1}, Y_{t+1}) - H_{v^*}(X_t, Y_t)] + \sum_{\substack{c \in \mathcal{L}: \\ d_{t,c}(v^*) \geq 1}} \mathbf{E} \left[\tilde{H}_{v^*, t, c}(X_{t+1}, Y_{t+1}) - \tilde{H}_{v^*, t, c}(X_t, Y_t) \right] \\
&= \mathbf{E} [H_{v^*}(X_{t+1}, Y_{t+1}) - H_{v^*}(X_t, Y_t)] + \sum_{\substack{c \in \mathcal{L}: \\ d_{t,c}(v^*) \geq 1}} \widetilde{W}_t^c && \text{(by Definition 11),}
\end{aligned}$$

where the second line follows from Eqs. (12) and (14). This completes the proof of the lemma. \square

4.3 Vertices Changing Between Blocked and Unblocked

Fix $(X_t, Y_t) \in \Omega_{v^*}^2$. Our goal now is to bound \widetilde{W}_t^c for all c . Recall $d_{t,c}(v^*) = |N_{t,c}(v^*)|$. Now let

$$d_{t,c}^i(v^*) := |N_{t,c}(v^*) \cap F_t^i(v^*)|$$

be the number of neighbors of v^* that are colored c in X_t or Y_t and have exactly i blocking neighbors with respect to v^* . For $i \geq 1$, let

$$\Delta_i = \Delta_{i,t}(v^*) := |\{u \in N(v^*) : d_{t,X_t(u)}(v^*) = i \text{ and } X_t(u) \in L(v^*)\}|$$

denote the number of neighbors whose color is in $L(v^*)$ and appears exactly i times in the neighborhood of v^* . Finally, let $\mathcal{A}_t(v^*) = \{c \in L(v^*) : d_{t,c}(v^*) = 0\}$ be the set of “available” colors for v^* in X_t and Y_t .

The following function will serve as an upper bound on \widetilde{W}_t^c and is a function of W_t^c , $d_{t,c}^0(v^*)$, and $d_{t,c}^1(v^*)$. The quantities α and β will appear in later lemmas.

Definition 4.2. Let $\alpha = \alpha(v^*, t) := (k - \Delta - 2)$, $\beta = \beta(v^*, t) := |L(v^*)| - P_2(\Delta_1 + \Delta_2) + (1 + P_2)d(v^*)$, and $c \in L(v^*)$. Then for $(X_t, Y_t) \in \Omega_{v^*}^2$, let

$$nk\widetilde{Z}_t^c := nkW_t^c + d_{t,c}^0(v^*)\frac{\eta\beta}{\Delta} - d_{t,c}^1(v^*)\frac{\eta\alpha}{\Delta}.$$

Observe that if $L(u) = [k]$ for all $u \in V$ (k -coloring case) then $\beta \leq k - P_2(\Delta_1 + \Delta_2) + (1 + P_2)\Delta$.

The following lemma shows that it suffices to bound \widetilde{Z}_t^c instead of \widetilde{W}_t^c directly. We will later be able to show that \widetilde{Z}_t^c is maximized in just a few cases, which we can analyze individually.

Lemma 4.3. For $(X_t, Y_t) \in \Omega_{v^*}^2$ and $c \in L(v^*)$,

$$\widetilde{W}_t^c \leq \widetilde{Z}_t^c.$$

To prove [Lemma 4.3](#), we use the following two lemmas that bound the expected number of vertices colored c in X_t and Y_t that become blocked or unblocked after a single step of the greedy coupling.

We first give a lower bound on the number of newly unblocked neighbors of v^* . In particular, for a color c , we lower bound the probability that for a neighbor w which is colored c in X_t and Y_t , that w becomes unblocked after a single step of the greedy coupling.

Lemma 4.4. For $(X_t, Y_t) \in \Omega_{v^*}^2$ and color $c \in L(v^*)$ such that $d_{t,c}(v^*) \geq 1$,

$$\mathbf{E} [| (F_{t+1}^0(v^*) \setminus F_t^0(v^*)) \cap N_{t,c}(v^*) |] \geq d_{t,c}^1(v^*) \frac{\alpha}{nk}.$$

Proof. Let $u \in F_t^1(v^*) \cap N_{t,c}(v^*)$. Since u is a singly blocked neighbor of v^* there must exist a unique neighbor $w \in N(u) \setminus (N(v^*) \cup \{v^*\})$ where $X_t(w) \in \{X_t(v^*), Y_t(v^*)\}$. There are at least $|\mathcal{A}_t(w) \setminus \{X_t(v^*), Y_t(v^*)\}| \geq (|L(w)| - d(w) - 2) \geq (k - \Delta - 2) = \alpha$ colors (other than $X_t(v^*)$ and $Y_t(v^*)$) that do not appear in the neighborhood of w in X_t and Y_t . Thus, for each color $c' \in \mathcal{A}_t(w) \setminus \{X_t(v^*), Y_t(v^*)\}$, there is a cluster of size 1 that contains just w and flipping such a cluster results in $X_{t+1}(w) = Y_{t+1}(w) = c'$. Therefore, for each $u \in N_{t,c}^1(v^*)$ there are at least $\alpha = (k - \Delta - 2)$ clusters of size 1 (and thus flip with probability $P_1/(nk) = 1/(nk)$) for which flipping one of these clusters results in $u \in F_{t+1}^0(v^*)$.

Summarizing the above calculations we have the following:

$$\mathbf{E} [| (F_{t+1}^0(v^*) \setminus F_t^0(v^*)) \cap N_{t,c}(v^*) |] \geq \mathbf{E} [| (F_{t+1}^0(v^*) \cap F_t^1(v^*)) \cap N_{t,c}(v^*) |] \geq d_{t,c}^1(v^*) \frac{\alpha}{nk}.$$

□

We now want to bound the probability that an unblocked neighbor u of v^* is no longer unblocked after a single step of the greedy coupling. Note that it does not suffice to only consider the probability u becomes blocked because there may no longer be a disagreement at v^* after a step of the greedy coupling, which results in u being neither blocked or unblocked (by definition).

There is a subtle and important trade-off in the following lemma. There are some unblocked neighbors u for which the probability u becomes blocked is relatively high; in these scenarios we will argue that there is a reasonable probability of having new unblocked vertices, namely, when u becomes a disagreement it will have many unblocked neighbors.

To capture the above trade-off, we need to consider two terms together, namely:

$$\mathbf{E} [| (F_t^0(v^*) \setminus F_{t+1}^0(v^*)) \cap N_{t,c}(v^*) |] \text{ and } \sum_{u \in \mathcal{L}_{t,c}} \mathbf{E} [d_{t+1}^0(u)]. \quad (15)$$

Note that $(F_t^0(v^*) \setminus F_{t+1}^0(v^*)) \cap N_{t,c}(v^*)$ includes those neighbors $w \in F_t^0(v^*) \cap N_{t,c}(v^*)$ such that: (i) $w \in F_{t+1}^1(v^*)$ if $X_{t+1}(v^*) \neq Y_{t+1}(v^*)$, or (ii) $w \in F_{t+1}^{\geq 2}(v^*)$ if $X_{t+1}(v^*) = Y_{t+1}(v^*)$. Hence, the first term of [Eq. \(15\)](#) is the expected number of c colored vertices in the neighborhood of v^* that go from unblocked in (X_t, Y_t) to one of the two scenarios (i) or (ii). The summand $\mathbf{E} [d_{t+1}^0(u)]$ in the second term of [Eq. \(15\)](#) is the expected number of unblocked neighbors of a vertex u in (X_{t+1}, Y_{t+1}) assuming u is a disagreement at time $t + 1$ since it equals 0 if $X_{t+1}(u) = Y_{t+1}(u)$ (by definition). This trade-off is a key idea in our improved bound.

Lemma 4.5. *If Property FP5 and Property FP6 hold then for $(X_t, Y_t) \in \Omega_{v^*}^2$ and color $c \in L(v^*)$ such that $d_{t,c}(v^*) \geq 1$,*

$$\mathbf{E} [| (F_t^0(v^*) \setminus F_{t+1}^0(v^*)) \cap N_{t,c}(v^*) |] - \sum_{u \in \Lambda_{t,c}} \mathbf{E} [d_{t+1}^0(u)] \leq d_{t,c}^0(v^*) \frac{\beta}{nk}.$$

Proof. Recall that $N_{t,c}(v^*) \subseteq \Lambda_{t,c}$. Thus, it will suffice to show that for all $u \in F_t^0(v^*) \cap N_{t,c}(v^*)$ the following holds:

$$\mathbf{Pr} (u \notin F_{t+1}^0(v^*) \mid u \in F_t^0(v^*) \cap N_{t,c}(v^*)) - \mathbf{E} [d_{t+1}^0(u)] \leq \frac{\beta}{nk}. \quad (16)$$

Consider $u \in F_t^0(v^*) \cap N_{t,c}(v^*)$. Hence, u is an unblocked neighbor of v^* , and $X_t(u) = Y_t(u) = c$.

We begin by focusing on $\mathbf{Pr} (u \notin F_{t+1}^0(v^*) \mid u \in F_t^0(v^*) \cap N_{t,c}(v^*))$. If u is not an unblocked neighbor of v^* after a single step of the greedy coupling (i.e., $u \notin F_{t+1}^0(v^*)$) then from the definition of $F_{t+1}^0(v^*)$ it follows that v^* was recolored in at least one of the chains or a neighbor $w \in N(u) \setminus \{v^*\}$ was recolored to $X_t(v^*)$ or $Y_t(v^*)$ in one of the chains. Let \mathcal{E}_{v^*} be the event that v^* is recolored in at least one chain (i.e., $X_{t+1}(v^*) \neq X_t(v^*)$ or $Y_{t+1}(v^*) \neq Y_t(v^*)$) and let $\overline{\mathcal{E}_{v^*}}$ be the event that v^* is not recolored in either chain (i.e., $X_{t+1}(v^*) = X_t(v^*)$ and $Y_{t+1}(v^*) = Y_t(v^*)$). Then, we can write

$$\begin{aligned} \mathbf{Pr} (u \notin F_{t+1}^0(v^*) \mid u \in F_t^0(v^*) \cap N_{t,c}(v^*)) \\ \leq \mathbf{Pr} (\mathcal{E}_{v^*}) + \sum_{w \in N(u) \setminus \{v^*\}} \mathbf{Pr} (X_{t+1}(w) \in \{X_t(v^*), Y_t(v^*)\} \mid \overline{\mathcal{E}_{v^*}}). \end{aligned}$$

Therefore, to prove Eq. (16) it suffices to show the following:

$$\mathbf{Pr} (\mathcal{E}_{v^*}) + \sum_{w \in N(u) \setminus \{v^*\}} \mathbf{Pr} (X_{t+1}(w) \in \{X_t(v^*), Y_t(v^*)\} \text{ and } \overline{\mathcal{E}_{v^*}}) - \mathbf{E} [d_{t+1}^0(u)] \leq \frac{\beta}{nk}. \quad (17)$$

We now bound $\mathbf{Pr} (\mathcal{E}_{v^*})$ which is the probability that v^* is recolored in at least one of the chains. If $X_{t+1}(v^*) \neq X_t(v^*)$ or $Y_{t+1}(v^*) \neq Y_t(v^*)$ then a cluster containing v^* must have flipped in X_t or Y_t . There are k clusters that contain v^* in each chain, one for each color in $L(v^*)$. For every color c that does not appear in the neighborhood of v^* , there is a cluster in both chains that contains only v^* ; the number of such colors is $|\mathcal{A}_t(v^*)| = |L(v^*)| - |\{c \in \mathcal{L} : d_{t,c}(v^*) \geq 1\}|$. Each of these clusters is of size 1 and thus flips with probability $1/(nk)$. For every color c that appears in the neighborhood of v^* (i.e., $d_{t,c}(v^*) \geq 1$), there are at most two clusters containing v^* (and all neighbors of v^* that are colored c), $S_{X_t}(v^*, c)$ and $S_{Y_t}(v^*, c)$. Note that $S_{X_t}(v^*, c)$ and $S_{Y_t}(v^*, c)$ flip with probability $P_{|S_{X_t}(v^*, c)|}/(nk)$ and $P_{|S_{Y_t}(v^*, c)|}/(nk)$ respectively. Also note that $|S_{X_t}(v^*, c)| \geq d_{t,c}(v^*) + 1$ since it

contains v^* and $N_{t,c}(v^*)$ and similarly $|S_{Y_t}(v^*, c)| \geq d_{t,c}(v^*) + 1$. Thus,

$$\begin{aligned}
\mathbf{Pr}(\mathcal{E}_{v^*}) &= \frac{1}{nk} \left(|\mathcal{A}_t(v^*)| + \sum_{\substack{c \in \mathcal{L}: \\ d_{t,c}(v^*) \geq 1}} \left(P_{|S_{X_t}(v^*, c)|} + P_{|S_{Y_t}(v^*, c)|} \right) \right) \\
&= \frac{1}{nk} \left(|L(v^*)| - \sum_{\substack{c \in \mathcal{L}: \\ d_{t,c}(v^*) \geq 1}} \left(1 - P_{|S_{X_t}(v^*, c)|} - P_{|S_{Y_t}(v^*, c)|} \right) \right) \\
&\leq \frac{1}{nk} \left(|L(v^*)| - \sum_{\substack{c \in \mathcal{L}: \\ d_{t,c}(v^*) \geq 1}} \left(1 - 2P_{d_{t,c}(v^*)+1} \right) \right) \\
&\leq \frac{1}{nk} (|L(v^*)| - (1 - 2P_2) \Delta_1 - (1 - 2P_3) \Delta_2/2) \\
&\leq \frac{1}{nk} (|L(v^*)| - P_2(\Delta_1 + \Delta_2)). \tag{18}
\end{aligned}$$

where the last inequality holds because $P_2 \leq 1/3$ by [Property FP0](#) and $1/2 \geq P_2 + P_3$ which is obtained by summing [Property FP5](#) and [Property FP6](#) and dividing by 2.

We now want to show for all $w \in N(u) \setminus \{v^*\}$:

$$\mathbf{Pr}(X_{t+1}(w) \in \{X_t(v^*), Y_t(v^*)\} \text{ and } \overline{\mathcal{E}_{v^*}}) - \mathbf{Pr}(w \in F_{t+1}^0(u)) \leq \frac{1 + P_2}{nk}. \tag{19}$$

If [Eq. \(19\)](#) holds then summing it over all $w \in N(u) \setminus \{v^*\}$ and combining it with [Eq. \(18\)](#) proves [Eq. \(17\)](#), which completes the proof of the lemma.

It remains to prove that [Eq. \(19\)](#) holds for all $w \in N(u) \setminus \{v^*\}$. Let $w \in N(u) \setminus \{v^*\}$ and recall that $w \in F_t^0(v^*) \cap N_{t,c}(v^*)$. We consider two cases: case (i) is that w has at least one neighbor besides u with color $X_t(v^*)$ or $Y_t(v^*)$ in X_t or in Y_t , and case (ii) is that w has no neighbors in X_t with colors $X_t(v^*)$ or $Y_t(v^*)$ and w has no neighbors in Y_t with colors $X_t(v^*)$ or $Y_t(v^*)$.

Suppose case (ii) occurs, hence no neighbors of w are colored $X_t(v^*)$ or $Y_t(v^*)$ in X_t or Y_t . Then $S_{X_t}(w, X_t(v^*))$ and $S_{Y_t}(w, Y_t(v^*))$ contain only the vertex w and flip in each chain with probability $1/(nk)$. Thus, the probability of w recoloring to $X_t(v^*)$ or $Y_t(v^*)$ is $2/(nk)$. Notice that if $S_{X_t}(w, X_t(v^*))$ and $S_{Y_t}(w, Y_t(v^*))$ flip then v^* does not flip since $w \neq v^*$. With probability at least $(1 - P_2)/(nk)$ the greedy coupling flips $S_{X_t}(u, Y_t(v^*))$ in X_t and flips $S_{Y_t}(u, X_t(v^*))$ in Y_t (see [Section 5.1](#)) and no other clusters flip at that time (hence no other vertices change colors). Hence, with probability at least $(1 - P_2)/(nk)$, $X_{t+1}(u) = Y_t(v^*)$, $Y_{t+1}(u) = X_t(v^*)$ and for all $z \neq u$, $X_{t+1}(z) = X_t(z)$ and $Y_{t+1}(z) = Y_t(z)$. Thus, with probability at least $(1 - P_2)/(nk)$ we have $w \in F_{t+1}^0(u)$ since w has no neighbor besides u that are colored $X_{t+1}(v^*) = X_t(v^*)$ or $Y_{t+1}(v^*) = Y_t(v^*)$. Therefore, in this case,

$$\mathbf{Pr}(X_{t+1}(w) \in \{X_t(v^*), Y_t(v^*)\} \text{ and } \overline{\mathcal{E}_{v^*}}) - \mathbf{Pr}(w \in F_{t+1}^0(u)) \leq \frac{2}{nk} - \frac{1 - P_2}{nk} = \frac{1 + P_2}{nk}$$

and [Eq. \(19\)](#) holds in this case.

Now suppose that case (i) holds so there is at least one neighbor $z \in N(w) \setminus \{u\}$ such that $X_t(z) = X_t(v^*)$ or $Y_t(z) = Y_t(v^*)$. Suppose without loss of generality that $X_t(z) = X_t(v^*)$. If $S_{X_t}(w, X_t(z)) = S_{Y_t}(w, X_t(z))$ then it contains w and z and thus flips with probability at

most $P_2/(nk)$. If $S_{X_t}(w, X_t(z)) \neq S_{Y_t}(w, X_t(z))$, then it must be the case that $S_{X_t}(w, X_t(z)) = S_{Y_t}(w, X_t(z)) \cup \{v^*\}$ since $(X_t, Y_t) \in \Omega_{v^*}^2$. Thus, flipping $S_{X_t}(w, X_t(z))$ in X_t will recolor v^* . Moreover, $S_{Y_t}(w, X_t(z))$ must contain w and z and thus flip with probability at most $P_2/(nk)$. Hence, the probability of recoloring w to $X_t(v^*)$ conditioned on v^* not being recolored is at most $P_2/(nk)$. Likewise, if there exists $z' \in N(w)$ such that $Y_t(z) = Y_t(v^*)$ then the probability of recoloring w to $X_t(v^*)$ conditioned on v^* not being recolored is at most $P_2/(nk) \leq 1/(nk)$. Finally, if there exists no $z' \in N(w)$ such that $Y_t(z') = Y_t(v^*)$ then the probability of recoloring w to $X_t(v^*)$ is at most $1/(nk)$ since, similar to the previous case, $S_{Y_t}(w, Y_t(v^*)) = S_{X_t}(w, Y_t(v^*))$ is a size 1 cluster and flips with probability $1/(nk)$. Thus,

$$\mathbf{Pr}(X_{t+1}(w) \in \{X_t(v^*), Y_t(v^*)\} \text{ and } \overline{\mathcal{E}_{v^*}}) \leq \frac{1+P_2}{nk}.$$

Therefore, in this case, [Eq. \(19\)](#) holds since $\mathbf{Pr}(w \in F_{t+1}^0(u)) \geq 0$. \square

We now have the tools to prove [Lemma 4.3](#), which states that $\widetilde{W}_t^c \leq \widetilde{Z}_t^c$.

Proof of Lemma 4.3. We start by observing:

$$\begin{aligned} |F_{t+1}^0(v^*) \cap N_{t,c}(v^*)| - d_{t,c}^0(v^*) &= |F_{t+1}^0(v^*) \cap N_{t,c}(v^*)| - |F_t^0(v^*) \cap N_{t,c}(v^*)| \\ &= |(F_{t+1}^0(v^*) \setminus F_t^0(v^*)) \cap N_{t,c}(v^*)| - |(F_t^0(v^*) \setminus F_{t+1}^0(v^*)) \cap N_{t,c}(v^*)|, \end{aligned} \quad (20)$$

where the second line uses the basic fact that for any sets A and B then $|A| - |B| = |A \setminus B| - |B \setminus A|$. We can now bound \widetilde{W}_t^c as follows:

$$\begin{aligned} \widetilde{W}_t^c &= \mathbf{E} \left[\widetilde{H}_{v^*, \Lambda_{t,c}}(X_{t+1}, Y_{t+1}) - \widetilde{H}_{v^*, \Lambda_{t,c}}(X_t, Y_t) \right] \\ &= \mathbf{E} [H_{v^*, c}(X_{t+1}, Y_{t+1})] \quad (\text{by definition}) \\ &\quad - \frac{\eta}{\Delta} \mathbf{E} \left[\sum_{z \in \Lambda_{t,c}} |F_{t+1}^0(z)| + |F_{t+1}^0(v^*) \cap N_{t,c}(v^*)| \right] - H_{v^*, c}(X_t, Y_t) + \frac{\eta}{\Delta} d_{t,c}^0(v^*) \\ &\leq W_t - \frac{\eta}{\Delta} \sum_{z \in N_{c,t}(v^*)} \mathbf{E} [|F_{t+1}^0(z)|] - \frac{\eta}{\Delta} (\mathbf{E} [|F_{t+1}^0(v^*) \cap N_{t,c}(v^*)|] + d_{t,c}^0(v^*)) \quad (N_{t,c}(v^*) \subseteq \Lambda_{t,c}) \\ &= W_t - \frac{\eta}{\Delta} \sum_{z \in N_{c,t}(v^*)} \mathbf{E} [|F_{t+1}^0(z)|] \\ &\quad - \frac{\eta}{\Delta} \mathbf{E} [| (F_{t+1}^0(v^*) \setminus F_t^0(v^*)) \cap N_{t,c}(v^*) |] \quad (\text{by Eq. (20)}) \\ &\quad + \frac{\eta}{\Delta} \mathbf{E} [| (F_t^0(v^*) \setminus F_{t+1}^0(v^*)) \cap N_{t,c}(v^*) |] \\ &\leq W_t - d_t^1(v^*) \frac{\alpha}{nk} \quad (\text{by Lemma 4.4}) \\ &\quad - \frac{\eta}{\Delta} \sum_{z \in N_{c,t}(v^*)} \mathbf{E} [|F_{t+1}^0(z)|] + \mathbf{E} \left[\frac{\eta}{\Delta} |(F_t^0(v^*) \setminus F_{t+1}^0(v^*)) \cap N_{t,c}(v^*)| \right] \\ &\leq W_t + d_{t,c}^0(v^*) \frac{\beta\eta}{\Delta nk} - d_t^1(v^*) \frac{\alpha\eta}{\Delta nk} \quad (\text{by Lemma 4.5}) \\ &= \widetilde{Z}_t^c. \end{aligned}$$

\square

4.4 Bounding \widetilde{W}_t^c and \widetilde{Z}_t^c

Fix a graph $G = (V, E)$, a pair of states $(X_t, Y_t) \in \Omega_{v^*}^2$, and a color $c \in \mathcal{L}$. The following two lemmas significantly reduce the initial configurations we have to consider by showing it suffices to consider the case where v^* has at most 2 neighbors that are color c , and all those neighbors are of the same type: unblocked (which is then considered in [Lemma 4.9](#)), singly blocked ([Lemma 4.10](#)), or multiblocked ([Lemma 4.11](#)). The proofs of these lemmas are deferred to [Section 5](#) but follow from [Lemma 4.3](#) and [Parameter Properties 3.1](#).

This first lemma handles the case where $d_{t,c} \geq 3$. In the proof of this lemma, we observe that as $d_{t,c}(v^*)$ grows, the expected change divided by $d_{t,c}(v^*)$ shrinks. This means that the gain from having additional colors in $\mathcal{A}_c(v^*)$ (from having more neighbors colored c) quickly exceeds the cost of having additional clusters that could flip and cause new disagreements.

Recall,

$$\beta = |L(v^*)| - P_2(\Delta_1 + \Delta_2) + (1 + P_2)d(v^*) \quad \text{and} \quad \alpha = (k - \Delta - 2)$$

from [Lemmas 4.4](#) and [4.5](#) respectively. In the following lemma we will assume that $\eta\beta/\Delta \leq P_2$, we will show that this holds in our parameter region in the upcoming proof of [Theorem 1.1](#).

Lemma 4.6. *If [Parameter Properties 3.1](#) hold, $c \in L(v^*)$, $\eta\beta/\Delta \leq P_2$, and $d_{t,c}(v^*) \geq 3$ then*

$$nk\widetilde{Z}_t^c \leq -1 + d_{t,c}(v^*) \left(\frac{4}{3} + P_2 + \frac{4}{3}P_4 \right).$$

The following lemma handles the case where $c \notin L(v^*)$. In this case, we prove a weaker bound that will suffice in our proof of [Theorem 1.1](#) because if $c \notin L(v^*)$ and $d_{t,c}(v^*) \geq 1$, then there exists an extra $c' \in L(v^*)$ such that $d_{t,c}(v^*) = 0$.

Lemma 4.7. *If [Parameter Properties 3.1](#) hold and $c \notin L(v^*)$ then*

$$nk\widetilde{W}_t^c \leq d_{t,c}(v^*)(1 + P_2).$$

This next lemma handles asymmetric cases: those cases where $d_{t,c}^i(v^*) > 0$ and $d_{t,c}^j(v^*) > 0$ for $i \neq j$. The proof follows from observing that the definition of \widetilde{Z}_t^c is linear in $d_{t,c}^0(v^*)$, $d_{t,c}^1(v^*)$, and $d_{t,c}^{\geq 2}(v^*)$.

Lemma 4.8. *If [Parameter Properties 3.1](#) hold, $c \in L(v^*)$, and $d_{t,c}(v^*) \leq 2$ then the function \widetilde{Z}_t^c is maximized when $d_{t,c}(v^*) = d_{t,c}^0(v^*)$, $d_{t,c}(v^*) = d_{t,c}^1(v^*)$, or $d_{t,c}(v^*) = d_{t,c}^{\geq 2}(v^*)$.*

We now consider the case when $c \in L(v^*)$. Based on the above lemmas we can assume that all neighbors with a specific color $c \in L(v^*)$ are all unblocked (i.e., $d_{t,c}(v^*) = d_{t,c}^0(v^*)$), all singly blocked (i.e., $d_{t,c}(v^*) = d_{t,c}^1(v^*)$), or all multiblocked (i.e., $d_{t,c}(v^*) = d_{t,c}^{\geq 2}(v^*)$). The following three lemmas will handle each of these three cases.

This first lemma handles the case where all neighbors of color c are unblocked. The critical observation is that while these colors will be “charged” a lot since unblocked neighbors can become blocked, their expected change in Hamming distance is relatively small because P_2 is so big.

Lemma 4.9. *If [Parameter Properties 3.1](#) hold, $c \in L(v^*)$, and $d_{t,c}^0(v^*) = d_{t,c}(v^*)$ then*

$$nk\widetilde{Z}_t^c \leq -1 + d_{t,c}(v^*) \left(2 - P_2 + \frac{\eta\beta}{\Delta} \right).$$

The following lemma is similar to the previous one, but handles the case when all neighbors colored c are singly blocked. Without the new metric (i.e., setting $\eta = 0$) these bounds would not satisfy the desired bound of $nk\tilde{Z}_t^c \leq -1 + d_{t,c}(v^*) \cdot 1.809$. However, since $\eta \neq 0$ and each of these cases has at least one singly blocked vertex, we get some benefit from the probability that a blocked neighbor becomes unblocked.

Lemma 4.10. *If [Parameter Properties 3.1](#) hold, $c \in L(v^*)$, and $d_{t,c}^1(v^*) = d_{t,c}(v^*)$ then*

$$nk\tilde{Z}_t^c \leq -1 + d_{t,c}(v^*) \left(2 - P_3 - \frac{\eta\alpha}{\Delta} \right).$$

Finally, the following lemma handles the case when all neighbors of color $c \in L(v^*)$ are multi-blocked. Since none of these cases will be affected by the new metric, it will be enough to bound the expected change in Hamming distance.

Lemma 4.11. *If [Parameter Properties 3.1](#) hold, $c \in L(v^*)$, and $d_{t,c}^{>2}(v^*) = d_{t,c}(v^*)$ then*

$$nk\tilde{Z}_t^c \leq -1 + d_{t,c}(v^*) (2 - P_2 + 2P_3 - 2P_4).$$

[Lemmas 4.8 to 4.11](#) are proved in [Section 5](#). We now prove the main result [Theorem 1.1](#).

4.5 Proof of [Theorem 1.1](#)

To prove [Theorem 1.1](#), we will prove the following stronger statement that holds for list colorings.

Theorem 4.12. *For all $\Delta \geq 125$, for all $k \geq 1.809\Delta$, for all lists such that $k \geq |L(v)| \geq d(v) + (k/\Delta - 1)\Delta$ for all $v \in V$, there exists a setting of the parameters for the flip dynamics with $P_j = 0$ for all $j \geq 7$, so that for any graph G on n vertices with maximum degree Δ ,*

$$nk\tilde{W}_t < -10^{-5}\Delta.$$

We can now prove [Theorem 1.1](#) as a direct corollary of [Theorems 2.1](#) and [4.12](#).

Proof of [Theorem 1.1](#). We apply [Theorem 2.1](#) with $\delta = 10^{-5}\Delta/(nk)$ to get that $T_{mix}(\varepsilon) \leq O(n \log(n/\varepsilon))$, which completes the proof of the theorem (See [Eq. \(9\)](#)). \square

We now prove [Theorem 4.12](#).

Proof of [Theorem 4.12](#). We will, in fact, prove that the mixing time is $O(n \log n)$ when $k > 1.809\Delta$, which is slightly stronger than the $k \geq 1.809\Delta$ statement in [Theorem 4.12](#).

We apply the Path Coupling Theorem ([Theorem 2.1](#)) with the metric \tilde{H} . We use the flip probabilities defined in [Parameter Setting 1](#). Recall $\eta := \Delta(P_2 - P_3)/(2k)$. Observe that when $k \geq 9/5$, $P_2 = 0.324$, and $P_3 = 0.154$,

$$\eta = \Delta(P_2 - P_3)/(2k) \leq (6/19)P_2.$$

Let

$$\lambda := \max\{2 - P_2 + \eta\beta/\Delta, 2 - P_3 - \eta\alpha/\Delta, 2 - P_2 + 2P_3 - 2P_4\}.$$

Consider $v^* \in V$ and $(X_t, Y_t) \in \Omega_{v^*}^2$. Then for all $c \in L(v^*)$ where $1 \leq d_{t,c}(v^*) \leq 2$, it follows from [Lemmas 4.9 to 4.11](#) that

$$nk\tilde{Z}_t^c \leq -1 + \lambda d_{t,c}(v^*). \tag{21}$$

And for all $c \in L(v^*)$ where $d_{t,c}(v^*) \geq 3$, it follows from [Lemma 4.6](#) that

$$nk\tilde{Z}_t^c \leq -1 + d_{t,c}(v^*) \left(\frac{4}{3} + P_2 + \frac{4}{3}P_4 \right) \leq -1 + 1.775d_{t,c}(v^*) \quad (22)$$

where we used that $P_2 = 0.324$ and $P_4 = 0.088$.

Recall that for $i \geq 1$, $\Delta_i = |\{u \in N(v^*) : d_{X_t(u)} = i \text{ and } X_t(u) \in L(v^*)\}|$ denotes the number of neighbors of v^* that are colored with a color in $L(v^*)$ and appears exactly i times in the neighborhood of v^* . Also, recall that $\Delta_{\geq 3} = \sum_{i \geq 3} \Delta_i$. Let $\Delta' = |\{u \in N(v^*) : X_t(u) \notin L(v^*)\}|$.

When $P_2 = 0.324$, we can combine the above bounds as follows:

$$\begin{aligned} nk\tilde{W}_t &\leq -|\mathcal{A}_t(v^*)| + nk \sum_{\substack{c \in \mathcal{L}: \\ d_{t,c}(v^*) \geq 1}} \tilde{W}_t^c && \text{(by Lemma 4.1)} \\ &\leq -|\mathcal{A}_t(v^*)| + nk \sum_{\substack{c \in L(v^*): \\ d_{t,c}(v^*) \geq 1}} \tilde{Z}_t^c + nk \sum_{\substack{c \in \mathcal{L} \setminus L(v^*): \\ d_{t,c}(v^*) \geq 1}} \tilde{W}_t^c && \text{(by Lemma 4.3)} \\ &\leq -|\mathcal{A}_t(v^*)| + nk \sum_{\substack{c \in L(v^*): \\ d_{t,c}(v^*) \geq 1}} \tilde{Z}_t^c + 1.324\Delta' && \text{(by Lemma 4.7 and since } P_2 = 0.324) \\ &\leq -|\mathcal{A}_t(v^*)| - |\{c \in L(v^*) : d_{t,c}(v^*) \geq 1\}| + \lambda(\Delta_1 + \Delta_2) \\ &\quad + 1.775\Delta_{\geq 3} + 1.324\Delta' && \text{(by Eqs. (21) and (22))} \\ &\leq -k + \lambda(\Delta_1 + \Delta_2) + 1.775\Delta_{\geq 3} + 1.324\Delta'. \end{aligned}$$

We will show that

$$\lambda(\Delta_1 + \Delta_2) + 1.775\Delta_{\geq 3} + 1.324\Delta' \leq (1 - 10^{-5})k \quad (23)$$

when $k \geq 1.8089\Delta$ and hence, this implies

$$nk\tilde{W}_t < -10^{-5}\Delta$$

when $k \geq 1.8089\Delta$ as desired.

All that remains is to establish [Eq. \(23\)](#). Recall the definition of λ , and we will consider the three corresponding cases.

Suppose $\lambda = 2 - P_2 + 2P_3 - 2P_4$, then plugging in the settings of $P_2 = 0.324$, $P_3 = 0.154$ and $P_4 = 0.088$ from [Parameter Setting 1](#) we see that $\lambda = 1.808$ and hence [Eq. \(23\)](#) holds in this case.

Now suppose $\lambda = 2 - P_3 - \eta\alpha/\Delta$. Plugging in $\alpha = (k - \Delta - 2)$, $P_3 = 0.154$ and $\eta = \Delta(P_2 - P_3)/(2k)$ we have that for $|L(v^*)| = k \geq 1.8089\Delta$:

$$\lambda = 2 - P_3 - \eta\alpha/\Delta \quad (24)$$

$$\begin{aligned} &= 1.846 - (P_2 - P_3)\alpha/(2k) \\ &\leq 1.846 - 0.17(0.5 - \Delta/(2k) - 1/k) \\ &\leq 1.846 - 0.17(0.223588 - 1/k) \\ &\leq 1.808 + 0.17/k \\ &\leq 1.8089 \end{aligned} \quad (25)$$

where the last inequality requires $\Delta \geq 104$. This establishes [Eq. \(23\)](#) for this case of λ when $\Delta \geq 104$.

Finally, suppose $\lambda = 2 - P_2 + \eta\beta/\Delta$. Plugging in $\beta = |L(v^*)| - P_2(\Delta_1 + \Delta_2) + (1 + P_2)d(v^*)$, $P_2 = 0.324$, $\eta = \Delta(P_2 - P_3)/(2k)$, and $k \geq 1.8089\Delta$,

$$\begin{aligned}
\lambda &= 2 - P_2 + (P_2 - P_3)\beta/(2k) \\
&\leq 1.676 + 0.17(L(v^*) - P_2(\Delta_1 + \Delta_2) + 1.324d(v^*))/(2k) \\
&\leq 1.676 + 0.17(k - P_2(\Delta_1 + \Delta_2) + 1.324\Delta)/(2k) \\
&\leq 1.676 + 0.17(0.5 - 0.324(\Delta_1 + \Delta_2)/(2k) + 1.324\Delta/(2k)) \\
&\leq 1.823212 - 0.015223(\Delta_1 + \Delta_2)/\Delta.
\end{aligned} \tag{26}$$

Note that $\Delta - \Delta' \geq d(v^*) - \Delta' \geq \Delta_1 + \Delta_2 + \Delta_{\geq 3}$. Let $x = (\Delta_1 + \Delta_2)/\Delta$ and $y = \Delta'/\Delta$. Recall that our goal is to establish [Eq. \(23\)](#). Plugging in [Eq. \(26\)](#) into [Eq. \(23\)](#) we have the following:

$$\begin{aligned}
\lambda(\Delta_1 + \Delta_2) + 1.775\Delta_{\geq 3} + 1.324\Delta' &\leq \lambda\Delta x + 1.775\Delta(1 - y - x) + 1.324\Delta y \\
&\leq 1.823212\Delta x - 0.015223\Delta x^2 + 1.775\Delta(1 - x - y) + 1.324\Delta y.
\end{aligned} \tag{27}$$

The maximum of [Eq. \(27\)](#) is 1.80799, which occurs when $x = 1$ and $y = 0$. Thus, [Eq. \(23\)](#) holds in all three cases. \square

4.6 Proof of [Corollary 1.2](#)

[Theorem 1.1](#) established a mixing time bound of $O(n \log(n))$ for the flip dynamics. By comparison of the associated Dirichlet forms, Vigoda [[Vig99](#)] showed that $O(n \log n)$ mixing time for the flip dynamics (with constant sized flips) implies $O(n^2 \log n)$ mixing time for the Glauber dynamics. In fact, since the dependence on ε in the mixing time is of the form $O(n \log(n/\varepsilon))$ then one obtains $O(n^2)$ mixing time of the Glauber dynamics, see [[DJV02](#), Corollary 2] or [[LP17](#), Chapter 14 notes].

To obtain $O(n \log n)$ mixing time of the Glauber dynamics for constant Δ , we utilize the work of Chen, Liu, and Vigoda [[CLV21](#)], who showed optimal mixing via spectral independence. For an introduction to spectral independence, see [[SV22](#)]. The following proof outlines how [Corollary 1.2](#) follows from [Theorem 4.12](#) using the method outlined in [[BCC⁺22](#)].

Proof of [Corollary 1.2](#). To prove [Corollary 1.2](#), we first observe that the flip dynamics is contractive under any pinning when $k \geq 1.809\Delta$ where a pinning is a fixed assignment for an arbitrary subset of vertices. To this end, we start with the assumption that $L(v) = [k]$ for all v . Then for any $U \subset V$ and pinning $\tau : U \rightarrow [k]$, observe that

$$k \geq |L(w) \setminus \bigcup_{z \in N(w) \cap U} \{\tau(z)\}| \geq d(w) + (k/\Delta - 1)\Delta.$$

Hence, it follows from [Theorem 4.12](#) that flip dynamics are contractive for any pair $X_0, Y_0 \in \Omega$, for any pinning and with respect to a distance metric which is 2-equivalent to the Hamming distance (in the terminology of [[BCC⁺22](#)]), it then follows from [[BCC⁺22](#), Theorem 1.11] that for every pinning, spectral independence holds. Thus, [[BCC⁺22](#), Theorem 1.7] yields $O(n \log n)$ mixing time for the Glauber dynamics for constant k and Δ as desired. \square

Remark 4.13. As noted in the proof of [Corollary 1.2](#), we show for list colorings under the conditions in [Theorem 4.12](#) that there is a contractive coupling with respect to a metric which is 2-equivalent to the Hamming distance. Recent work of Chen, Feng, Guo, Zhang, and Zou [[CFG⁺24](#)] shows that this implies a deterministic approximate counting algorithm (FPTAS) when k and Δ are constant.

5 Greedy Coupling and Remaining Proofs

To complete the proof of [Theorem 1.1](#) it remains to prove [Lemma 4.6](#) through [Lemma 4.11](#). Most of these proofs are straightforward but require additional information about the greedy coupling or the consideration of several specific cases. We first give additional details on the greedy coupling in [Section 5.1](#).

5.1 Greedy Coupling: Detailed Definition

In this section, we formally define the details of the greedy coupling and analyze its expected change in Hamming distance. Recall the definition of the clusters in the set $\mathcal{D}_{t,c}$ from [Eq. \(8\)](#). As stated earlier, for each color $c \in \mathcal{L}$, flips in $\mathcal{D}_{t,c}$ are coupled with other flips within the set (or coupled with no flip corresponding to a self-loop). Consequently, we can consider each set $\mathcal{D}_{t,c}$ separately.

Fix a pair $(X_t, Y_t) \in \Omega_{v^*}^2$ for some $v^* \in V$. Fix a color $c \in \mathcal{L}$ and we will define the coupling for the flips of clusters in the set $\mathcal{D}_{t,c}$.

If $c \notin X_t(N(v^*))$ and $c \in L(v^*)$ (and thus c is an available color for v^*) then $\mathcal{D}_{t,c}$ consists of two clusters $S_{X_t}(v^*, c)$ and $S_{Y_t}(v^*, c)$, and both of these clusters are simply the vertex v^* : $S_{X_t}(v^*, c) = S_{Y_t}(v^*, c) = \{v^*\}$. The coupling is the identity coupling for these clusters, and hence, we choose the same (v^*, c) in both chains and flip the cluster in both chains or in neither chain.

Now suppose $c \in X_t(N(v^*))$ and $c \in L(v^*)$. To define the coupling within $\mathcal{D}_{t,c}$ let us begin with the simpler case $d_{t,c}(v^*) = 1$; let $N_c = \{u\}$. In this case, within $\mathcal{D}_{t,c}$ we have 2 clusters for each chain; in X_t we have $S_{X_t}(v^*, c)$ and $S_{X_t}(u, Y_t(v^*))$, and in Y_t we have $S_{Y_t}(v^*, c)$ and $S_{Y_t}(u, X_t(v^*))$. Let $a = |S_{X_t}(u, Y_t(v^*))|$ and $A = |S_{Y_t}(v^*, c)|$, and similarly $b = |S_{Y_t}(u, X_t(v^*))|$ and $B = |S_{X_t}(v^*, c)|$. Note that $S_{Y_t}(v^*, c) = \{v^*\} \cup S_{X_t}(u, Y_t(v^*))$, and hence, $A = 1 + a_1$; similarly, $B = 1 + b_1$.

With probability P_B , we couple the flip of $S_{X_t}(v^*, c)$ (of size B) in X_t with $S_{Y_t}(u, X_t(v^*))$ (of size P_b) in Y_t ; this does not change the Hamming distance as the new chains X_{t+1}, Y_{t+1} only differ at v^* . Similarly, we couple the flip of $S_{Y_t}(v^*, c)$ (of size A) in Y_t with $S_{X_t}(u, Y_t(v^*))$ (of size P_a) in X_t ; again, this does not change the Hamming distance. There remains probability $P_a - P_A$ for flipping $S_{X_t}(u, Y_t(v^*))$ (of size a) in X_t and probability $P_b - P_B$ for flipping $S_{Y_t}(u, X_t(v^*))$ (of size b) in Y_t ; we use the maximal coupling for these flips. Hence, with probability $\min\{P_a - P_A, P_b - P_B\}$ we flip both $S_{X_t}(u, Y_t(v^*))$ in X_t and $S_{Y_t}(u, X_t(v^*))$ in Y_t ; this increases the Hamming distance by $|S_{X_t}(u, Y_t(v^*)) \cup S_{Y_t}(u, X_t(v^*))| \leq a + b - 1$. With the remaining probability, the remaining cluster flips by itself (self-loop in the other chain).

Summarizing, in the case $d_{t,c}(v^*) = 1$, the expected change in Hamming distance is at most

$$a(P_a - P_A) + b(P_b - P_B) - \min\{P_a - P_A, P_b - P_B\}.$$

Now consider the general case $d_{t,c}(v^*) \geq 1$. Let $N_c = \{u_1, \dots, u_{d_{t,c}(v^*)}\}$. Recall, $\mathcal{D}_{t,c}$ consists of the following clusters in X_t :

$$\{S_{X_t}(v^*, c)\} \cup \bigcup_i \{S_{X_t}(u_i, Y_t(v^*))\}$$

and in Y_t we have:

$$\{S_{Y_t}(v^*, c)\} \cup \bigcup_i \{S_{Y_t}(u_i, X_t(v^*))\}$$

For $1 \leq i \leq d_{t,c}(v^*)$, let $a_i = |S_{X_t}(u_i, Y_t(v^*))|$ and $b_i = |S_{Y_t}(u_i, X_t(v^*))|$ denote the sizes of the $(c, Y_t(v^*))$ and $(c, X_t(v^*))$, respectively, clusters containing u_i . Moreover, if $u_i \in S_{X_t}(u_j, Y_t(v^*))$ for some $j < i$ then redefine $a_i = 0$ (this will avoid double-counting); similarly, if $u_i \in S_{Y_t}(u_j, X_t(v^*))$ for some $j < i$ then $b_i = 0$.

Let $a_{\max} = \max_i a_i$ and $b_{\max} = \max_i b_i$. Let i_{\max} be an index i where $a_i = a_{\max}$ and similarly let i'_{\max} be an index i' where $b_{i'} = b_{\max}$. Moreover, if possible it sets $i_{\max} = i'_{\max}$. Finally, let $A = |S_{Y_t}(v^*, c)|$ and $B = |S_{X_t}(v^*, c)|$. If $c \neq Y_t(v^*)$ then $A = 1 + \sum_i a_i$ and if $c \neq X_t(v^*)$ then $B = 1 + \sum_i b_i$. Note that if $c \notin L(v^*)$ then $A = 0$ and $B = 0$. For notational convenience, we will define $P_0 = 0$.

We try to couple the flips of the large clusters of size A and B with the largest of the small clusters of sizes a_{\max} and b_{\max} , respectively. If $c \neq X_t(v^*)$ then with probability P_A we couple the flip of $S_{X_t}(v^*, c)$ in X_t with $S_{Y_t}(u_{i_{\max}}, X_t(v^*))$ in Y_t ; this changes the Hamming distance by $A - a_{\max} - 1$ if $c \neq Y_t(v^*)$ and $A - a_{\max} - 2$ if $c = Y_t(v^*)$ since $X_{t+1}(v^*) = Y_{t+1}(v^*)$ in this case. Similarly, with probability P_B we couple the flip of $S_{Y_t}(v^*, c)$ in Y_t with $S_{X_t}(u_{i_{\max}}, Y_t(v^*))$ in X_t ; this changes the Hamming distance by $B - b_{\max} - 1$ if $c \neq X_t(v^*)$ and changes it by $B - b_{\max} - 2$ if $c = X_t(v^*)$ since $X_{t+1}(v^*) = Y_{t+1}(v^*)$ in this case. For $1 \leq i \leq d_{t,c}(v^*)$, denote the remaining flip probabilities as $Q_i = P_{a_i} - P_A \mathbf{1}(i = i_{\max})$, and $Q'_i = P_{b_i} - P_B \mathbf{1}(i = i'_{\max})$. Then, for each $1 \leq i \leq d_{t,c}(v^*)$, with probability $\min\{Q_i, Q'_i\}$ we flip both $S_{X_t}(u_i, Y_t(v^*))$ in X_t and $S_{Y_t}(u_i, X_t(v^*))$ in Y_t ; this increases the Hamming distance by $a_i + b_i - m_i$ where $m_i = |S_{X_t}(u_i, Y_t(v^*)) \cap S_{Y_t}(u_i, X_t(v^*))|$. With the remaining probability, the remaining cluster flips by itself (self-loop in the other chain).

Hence, for any c such that $d_{t,c}(v^*) \geq 1$, observe that

$$nkW_t^c = (A - a_{\max} - 1)P_A + (B - b_{\max} - 1)P_B + \sum_i (a_i Q_i + b_i Q'_i - m_i \min\{Q_i, Q'_i\}).$$

Note that for all i , $m_i \geq 1$ since $u_i \in S_{X_t}(u_i, Y_t(v^*)) \cap S_{Y_t}(u_i, X_t(v^*))$. Moreover, larger values of m_i only decrease the expected change in Hamming distance. One can view the tree as the worst case since if the graph considered is a tree (or has high girth), then $m_i = 1$ since $S_{X_t}(u_i, Y_t(v^*)) \cap S_{Y_t}(u_i, X_t(v^*)) = \{u_i\}$ for all i . We define the following function Z_t^c , which is an upper bound on W_t^c for all graphs.

Definition 5.1. For any c such that $d_{t,c}(v^*) \geq 1$, let

$$nkZ_t^c := (A - a_{\max} - 1)P_A + (B - b_{\max} - 1)P_B + \sum_i (a_i Q_i + b_i Q'_i - \min\{Q_i, Q'_i\}).$$

When G has sufficiently large girth (namely, girth ≥ 15 suffices) then $W_t^c = Z_t^c$. Moreover, for any G we have $W_t^c \leq Z_t^c$. Recall, [Definition 4.2](#) and thus

$$\tilde{Z}_t^c \leq Z_t^c + d_{t,c}^0(v^*) \frac{\eta\beta}{\Delta} - d_{t,c}^1(v^*) \frac{\eta\alpha}{\Delta}. \quad (28)$$

The proofs of [Lemma 4.6](#) through [Lemma 4.11](#) are provided in the full version of the paper.

5.2 Proof of [Lemma 4.6](#): Color Appearing At Least 3 Times

In this section, we prove [Lemma 4.6](#).

Proof of [Lemma 4.6](#). Let $N_{t,c}(v^*) = \{u_1, \dots, u_{d_{t,c}(v^*)}\}$ be the neighbors of v^* that are colored c in X_t and Y_t . Since $d_{t,c}(v^*) \geq 3$, it follows that $A \geq 4$ and $B \geq 4$. Hence, $(A - a_{\max} - 1)P_A \leq (A - 2)P_A$ and then it follows from [Property FP1](#) that $(A - 2)P_A \leq \max_j \{(j - 2)P_j\} \leq 2P_4$, and similarly, $(B - b_{\max} - 1)P_B \leq 2P_4$.

Fix $u_i \in N_{t,c}(v^*)$ and suppose without loss of generality that $b_i \geq a_i$. If u_i is unblocked then $a_i = 1$, $b_i = 1$, and $a_i Q_i + b_i Q'_i - \min\{Q_i, Q'_i\} = Q_i + Q'_i - \min\{Q_i, Q'_i\} = \max\{Q_i, Q'_i\} \leq 1$. If u_i is not unblocked, then it must be the case that $a_i \geq 2$ or $b_i \geq 2$; thus $a_i Q_i + b_i Q'_i - \min\{Q_i, Q'_i\} =$

$a_i Q_i + (b_i - 1) Q'_i \leq 1 + P_2$ since $a_i Q_i \leq 1$ by [Property FP0](#) and [Property FP1](#), and $(b_i - 1) Q'_i \leq P_2$ by [Property FP2](#) and [Property FP3](#). Putting together these bounds and using [Definition 5.1](#) we have:

$$nkZ_t^c \leq 4P_4 + d_{t,c}(v^*) + (d_{t,c}(v^*) - d_{t,c}^0(v^*))P_2.$$

Thus, using [Definition 5.1](#) and [Eq. \(28\)](#) we get

$$\begin{aligned} nk\tilde{Z}_t^c &\leq nkZ_t^c + d_{t,c}^0(v^*)\frac{\eta\beta}{\Delta} \\ &\leq d_{t,c}(v^*) + (d_{t,c}(v^*) - d_{t,c}^0(v^*))P_2 + 4P_4 + d_{t,c}^0(v^*)\frac{\eta\beta}{\Delta} \\ &\leq -1 + d_{t,c}(v^*)\left(\frac{4}{3} + P_2 + \frac{4}{3}P_4\right). \end{aligned}$$

where the last inequality holds because $\eta\beta/\Delta \leq P_2$ and $d_{t,c}(v^*) \geq 3$. \square

5.3 Proof of [Lemma 4.7](#): Color Not Available

Now we prove [Lemma 4.7](#). The proof of this lemma follows almost immediately from our coupling and [Parameter Properties 3.1](#).

Proof of Lemma 4.7. Since $c \notin L(v^*)$ it follows that $S_{X_t}(v^*, c) = S_{Y_t}(v^*, c) = \emptyset$. Thus,

$$\begin{aligned} nk\tilde{Z}_t^c &\leq \sum_i (a_i P_{a_i} + b_i P_{b_i} - \min\{P_{a_i}, P_{b_i}\}) \\ &\leq d_{t,c}(v^*)(1 + P_2) \end{aligned}$$

where the last inequality follows since $\max\{iP_i\} = 1$ and $\max\{(i-1)P_i\} = P_2$ by [Parameter Properties 3.1](#). \square

5.4 Proof of [Lemma 4.9](#): All Unblocked

Now we prove [Lemma 4.9](#). The key to this proof is the observation that since $d_{t,c}^0(v^*) = d_{t,c}(v^*)$ (by [Lemma 4.8](#)), there is a unique configuration to consider for each value of $d_{t,c}(v^*)$, and hence we can directly compute W_t^c for any value of $d_{t,c}(v^*)$.

Proof of Lemma 4.9. We first consider the case when $d_{t,c}(v^*) = 1$. Since $d_{t,c}^0(v^*) = d_{t,c}(v^*)$, it follows that $a_1 = 1$ and $b_1 = 1$. Hence, $A = 2$, $B = 2$, and by [Eq. \(28\)](#),

$$nk\tilde{Z}_t^c = 1 - P_2 + \frac{\eta\beta}{\Delta} = -1 + d_{t,c}(v^*)\left(2 - P_2 + \frac{\eta\beta}{\Delta}\right).$$

We now consider the case when $d_{t,c}(v^*) = 2$. Since $d_{t,c}^0(v^*) = d_{t,c}(v^*)$, it follows that $a_1 = a_2 = b_1 = b_2 = 1$. Hence, $A = 3$, $B = 3$, by [Definition 5.1](#),

$$Z_t^c = -1 + d_{t,c}(v^*)\left(\frac{3}{2} + P_3\right).$$

Thus, by [Eq. \(28\)](#),

$$nk\tilde{Z}_t^c \leq nkZ_t^c + d_{t,c}(v^*)\frac{\eta\beta}{\Delta} \leq -1 + d_{t,c}(v^*)\left(\frac{3}{2} + P_3 + \frac{\eta\beta}{\Delta}\right) \leq -1 + d_{t,c}(v^*)\left(2 - P_2 + \frac{\eta\beta}{\Delta}\right)$$

where the last inequality follows since $P_2 + P_3 \leq 1/2 - P_5/2 \leq 1/2$ follows from summing [Property FP5](#) and [Property FP6](#) and dividing by 2. \square

5.5 Proof of Lemma 4.10: All Singly Blocked

Next, we prove [Lemma 4.10](#). With the help of [Lemma 4.8](#), the proof of this lemma is very similar to that of [Lemma 4.9](#). However, unlike unblocked vertices, singly blocked vertices can be in clusters of arbitrary size. We have to show that the worst case is when the clusters are relatively small.

Proof of Lemma 4.10. Using [Lemma 5.3](#) we can assume $a_i = 1$ and $b_i = 2$ for $1 \leq i \leq d_{t,c}(v^*)$.

We first consider the case when $d_{t,c}(v^*) = 1$. In this case, $A = 2$ and $B = 3$. Hence, by [Eq. \(28\)](#),

$$nk\tilde{Z}_t^c \leq 1 - P_3 - \frac{\eta\alpha}{\Delta} = -1 + d_{t,c}(v^*) \left(2 - P_3 - \frac{\eta\alpha}{\Delta} \right).$$

We now consider the $d_{t,c}(v^*) = 2$ case. In this case, $A = 3$, $B = 5$, and by [Definition 5.1](#),

$$nkZ_t^c = -1 + d_{t,c}(v^*) \left(\frac{3}{2} + P_2 + \frac{P_5}{2} \right) \leq -1 + d_{t,c}(v^*) (2 - P_3)$$

since $P_2 + P_3 \leq 1/2 - P_5/2 \leq 1/2$ follows from summing [Property FP5](#) and [Property FP6](#) and dividing by 2. Hence, by [Eq. \(28\)](#),

$$nk\tilde{Z}_t^c \leq nkZ_t^c - d_{t,c}(v^*) \frac{\eta\alpha}{\Delta} \leq -1 + d_{t,c}(v^*) \left(2 - P_3 - \frac{\eta\alpha}{\Delta} \right).$$

□

5.6 Proof of Lemma 4.11: All Multiblocked

Finally, we prove [Lemma 4.11](#). Again, this proof is similar to the proofs of [Lemmas 4.9](#) and [4.10](#).

Proof of Lemma 4.11. Since $d_{t,c}^0(v^*) = d_t^1(v^*) = 0$ it follows from [Eq. \(28\)](#) that $nk\tilde{Z}_t^c \leq nkZ_t^c$ and it suffices to show $nkZ_t^c \leq -1 + d_{t,c}(v^*)(2 - P_2 + 2(P_3 - P_4))$.

We first consider the case when $d_{t,c}(v^*) = 1$. Using [Lemma 5.2](#) we can assume $a_1 = 1$ and $b_1 = 3$. In this case, $A = 2$, $B = 4$, and by [Definition 5.1](#):

$$nk\tilde{Z}_t^c \leq 1 - P_2 + 2(P_3 - P_4) = -1 + d_{t,c}(v^*)(2 - P_2 + 2(P_3 - P_4)).$$

We now consider the case where $d_{t,c}(v^*) = 2$. It follows from [Lemma 5.3](#) that we can assume $a_i = 1$ and $b_i = 3$ for all i . In this case, $A = 3$, $B = 7$, and by [Definition 5.1](#),

$$nkZ_t^c = 2 + 4P_3 = -1 + d_{t,c}(v^*) \left(\frac{3}{2} + 2P_3 \right) \leq -1 + d_{t,c}(v^*)(2 - P_2 + 2(P_3 - P_4))$$

where the last inequality follows from the fact that $1/2 \geq P_2 + 2P_4$ by [Property FP5](#). □

5.7 Proof of Lemma 4.8: Extremal Cases

In this section, we prove [Lemma 4.8](#) using the following two lemmas; these two lemmas are analogous to similar claims in previous works, namely [[Vig99](#), Claim 6] and [[CDM⁺19](#), Observation B.1].

Lemma 5.2. *Assume [Parameter Properties 3.1](#) hold. For $c \in \mathcal{L}$ where $d_{t,c}(v^*) = 1$, then Z_t^c is maximized when $a_1 = 1$ and $b_1 = 2$. Moreover, if $d_{t,c}(v^*) = 1$ and $b_1 \geq 3$, then Z_t^c is maximized when $a_1 = 1$ and $b_1 = 3$.*

Proof. Assume without loss of generality that $b_1 \geq a_1$. Then

$$nkZ_t^c = a_1(P_{a_1} - P_{a_1+1}) + (b_1 - 1)(P_{b_1} - P_{b_1+1}).$$

If $a_1 = 1$ then we get

$$nkZ_t^c = 1 - P_2 + (b_1 - 1)(P_{b_1} - P_{b_1+1}). \quad (29)$$

Observe that [Eq. \(29\)](#) is maximized when $b_1 = 2$ by [Property FP2](#) and [Property FP3](#). Moreover, if $b_1 \geq 3$ then [Eq. \(29\)](#) is maximized when $b_1 = 3$ by [Property FP3](#).

Now let us consider the case where $a_1 \geq 2$. First, suppose $b_1 = 2$. If $b_1 = 2$ and $a_1 = 1$, then $nkZ_t^c = 1 - P_3$. If $b_1 = 2$ and $a_1 = 2$, then $nkZ_t^c = 2(P_2 - P_3) \leq 2P_2 \leq 2/3 \leq 1 - P_2 \leq 1 - P_3$ by [Property FP0](#) and [Property FP1](#). Thus, if $b_1 = 2$ then nkZ_t^c is maximized when $a_1 = 1$.

Now let us suppose $b_1 \geq 3$. If $b_1 = 3$ and $a_1 = 1$ then $nkZ_t^c = 1 - P_2 + 2(P_3 - P_4)$. If $b_1 \geq 3$ and $a_1 \geq 2$ then $(b_1 - 1)(P_{b_1} - P_{b_1+1})$ is maximized when $b_1 = 3$ by [Property FP3](#). Observe that $2(P_2 - P_3) \geq 2(j-1)(P_j - P_{j+1}) \geq j(P_j - P_{j+1})$ for all $j \geq 2$ by [Property FP2](#) and [Property FP3](#). Hence, $a_1(P_{a_1} - P_{a_1+1})$ is maximized when $a_1 = 2$. Therefore, if $b_1 \geq 3$ and $a_1 \geq 2$, then nkZ_t^c is maximized with $a_1 = 2$ and $b_1 = 3$ which yields $nkZ_t^c \leq 2(P_2 - P_3) + 2(P_3 - P_4) \leq 1 - P_2 + 2(P_3 - P_4)$ since $2(P_2 - P_3) \leq 2/3 \leq 1 - P_2$ by [Property FP0](#). Therefore, if $b_1 \geq 3$ then nkZ_t^c is maximized when $b_1 = 3$ and $a_1 = 1$. \square

Lemma 5.3. *Assume [Parameter Properties 3.1](#) hold. For $c \in \mathcal{L}$ where $d_{t,c}(v^*) = 2$, then Z_t^c is maximized when $a_1 = a_2 = 1$ and $b_{\max} \leq 3$.*

Proof. Assume without loss of generality that $b_1 \geq \max\{b_2, a_1, a_2\}$. Recall from [Definition 5.1](#)

$$nkZ_t^c = (A - a_{\max} - 1)P_A + (B - b_{\max} - 1)P_B + \sum_i (a_i Q_i + b_i Q'_i - \min\{Q_i, Q'_i\}).$$

We first show that we can assume $a_1 \leq a_2$. Suppose $a_2 < a_1$. Observe that since $b_1 \geq \max\{a_1, a_2\}$, then we have that $Q'_1 = P_{b_1} - P_B \leq \min\{Q_1 = P_{a_1} - P_A, Q_2 = P_{a_2}\} = P_{a_1} - P_A$ since $P_{a_1} - P_A = \min\{Q_1 = P_{a_1} - P_A, Q_2 = P_{a_2}\}$ by [Property FP1](#) and $P_{b_1} - P_B \leq P_{a_1} - P_A$ since $(P_j - P_{j+1}) \leq (P_{j-1} - P_j)$ which holds by [Properties FP2](#) to [FP4](#). Thus, the switching of a_1 and a_2 can only change the $\min\{Q_2, Q'_2\}$ term in nkZ_t^c . Moreover, setting $a_2 = \min\{a_1, a_2\}$ can only decrease $\min\{Q_2, Q'_2\}$. Therefore, Z_t^c is maximized when $a_2 \geq a_1$.

We can assume $b_1 \geq b_2$, $b_1 \geq a_1$, and $a_2 \geq a_1$. Then we can write

$$nkZ_t^c = (A - 2a_2 - 1)P_A + (B - 2b_1)P_B + a_1 P_{a_1} + a_2 P_{a_2} + (b_1 - 1)P_{b_1} + b_2 P_{b_2} - \min\{P_{a_2} - P_A, P_{b_2}\}.$$

Observe that when $a_1 = a_2 = 1$ and $b_1 = b_2 = 3$ then $nkZ_t^c = 2 + 4P_3$. We now consider the case where $P_{b_2} \leq P_{a_2} - P_A$ and $P_{b_2} > P_{a_2} - P_A$ and show in each that the maximum is at most $2 + 4P_3$.

Suppose $P_{b_2} \leq P_{a_2} - P_A$. Then we can write:

$$nkZ_t^c = (A - 2a_2 - 1)P_A + (B - 2b_1)P_B + a_1 P_{a_1} + a_2 P_{a_2} + (b_1 - 1)P_{b_1} + (b_2 - 1)P_{b_2}.$$

Observe that $(b_i - 1)P_{b_i}$ is maximized when $b_i = 2$ by [Property FP1](#). Moreover, $(B - 2b_1)P_B$ is maximized when $b_1 + b_2 \leq 6$ since $P_7 = 0$ by [Property FP0](#). Thus, Z_t^c is maximized when $b_i = 3$ for all i . If $a_1 \geq 2$ then $a_2 \geq 2$, $(A - 2a_2 - 1)P_A \leq P_5$ by [Property FP1](#) and $a_1 P_{a_1} + a_2 P_{a_2} \leq 2P_2$ by [Property FP1](#). The claim holds in this case since $2P_2 + P_5 \leq 2 + 4P_3$ by [Property FP0](#) and [Property FP1](#).

Now suppose $P_{b_2} > P_{a_2} - P_A$. Then we can write:

$$nkZ_t^c = (A - 2a_2)P_A + (B - 2b_1)P_B + a_1 P_{a_1} + (a_2 - 1)P_{a_2} + (b_1 - 1)P_{b_1} + b_2 P_{b_2}.$$

Suppose $a_2 \geq 2$. Notice that $a_1 P_{a_1} \leq 1$ by [Property FP1](#). If $a_1 < a_2$ then $(A - 2a_2)P_A + a_1 P_{a_1} + (a_2 - 1)P_{a_2} \leq 1 - P_{2a_2+1} + (a_2 - 1)P_{a_2} \leq 1 + (a_2 - 1)P_{a_2} \leq 1 + P_2$ by [Property FP1](#). If $a_1 = a_2$ then $(A - 2a_2)P_A + a_1 P_{a_1} + (a_2 - 1)P_{a_2} \leq a_1 P_{a_2} + (a_2 - 1)P_{a_2} \leq 3P_2 \leq 1$ by [Property FP1](#) and [Property FP0](#). Similarly, we get that $(B - 2b_1)P_B + a_1 P_{a_1} + (b_1 - 1)P_{b_1} + b_2 P_{b_2} \leq 1$. The claim then holds since $2 \leq 2 + 4P_3$. \square

We can now prove [Lemma 4.8](#). Using [Lemma 5.3](#) we will write \tilde{Z}_t^c as a linear function in $d_{t,c}^0(v^*)$, $d_{t,c}^1(v^*)$, and $d_{t,c}^{\geq 2}(v^*)$.

Proof of Lemma 4.8. Note that the claim is trivially true if $d_{t,c}(v^*) = 1$. Suppose $d_{t,c}(v^*) = 2$. Let $\{u_1, u_2\} = N_{t,c}(v^*)$. Without loss of generality, assume $A \leq B$, and $b_1 \geq b_2$. Note that if $u_i \in F_0^0(t)$ then $a_i = b_i = 1$ by definition. Likewise, if $u_i \in F_1^1(t)$ then $\max\{a_i, b_i\} = 2$ and $\min\{a_i, b_i\} = 2$. Finally, if $u_i \in F_2^{\geq 2}(t)$ then $a_i + b_i \geq 4$. Recall from [Eq. \(28\)](#) that

$$nk\tilde{Z}_t^c \leq nk\tilde{Z}_t^c + d_{t,c}^0(v^*)\frac{\eta\beta}{\Delta} - d_{t,c}^1(v^*)\frac{\eta\alpha}{\Delta}$$

and this is an equality when the underlying graph has sufficiently high girth. Notice that the only term affected by the value a_1, a_2, b_1 and b_2 is $nk\tilde{Z}_t^c$. Thus, using [Lemma 5.3](#) it suffices to assume $a_1 = a_2 = 1$ and

$$b_i = \begin{cases} 1 & \text{if } u_i \in F_t^0(v^*) \\ 2 & \text{if } u_i \in F_t^1(v^*) \\ 3 & \text{if } u_i \in F_t^{\geq 2}(v^*) \end{cases} \quad (30)$$

for $1 \leq i \leq 2$. Then by [Definition 5.1](#),

$$\begin{aligned} Z_t^c &= (b_2 - 1)P_{b_1+b_2+1} + P_3 + (1 - P_2) + (b_1 - 1)(P_{b_1} - P_{b_1+b_2+1}) + 1 + (b_2 - 1)P_{b_2} \\ &= (b_2 - b_1)P_{b_1+b_2+1} + 2 - P_2 + P_3 + (b_1 - 1)P_{b_1} + (b_2 - 1)P_{b_2}. \end{aligned}$$

Thus,

$$\begin{aligned} \tilde{Z}_t^c &= (b_2 - b_1)P_{b_1+b_2+1} + 2 - P_2 + P_3 + (b_1 - 1)P_{b_1} + (b_2 - 1)P_{b_2} + d_{t,c}^0(v^*)\frac{\eta\beta}{\Delta} - d_{t,c}^1(v^*)\frac{\eta\alpha}{\Delta} \\ &= 2 - P_2 + P_3 + (b_1 - 1)P_{b_1} + (b_2 - 1)P_{b_2} + d_{t,c}^0(v^*)\frac{\eta\beta}{\Delta} - d_{t,c}^1(v^*)\frac{\eta\alpha}{\Delta} \quad (\text{since } b_1 \geq b_2) \\ &= 2 - P_2 + P_3 + d_{t,c}^0(v^*)\frac{\eta\beta}{\Delta} + d_{t,c}^1(v^*)\left(P_2 - \frac{\eta\alpha}{\Delta}\right) + 2d_{t,c}^1(v^*)P_3. \quad (\text{by Eq. (30)}) \quad (31) \end{aligned}$$

Note that [Eq. \(31\)](#) is linear in $d_{t,c}^0(v^*)$, $d_{t,c}^1(v^*)$, and $d_{t,c}^{\geq 2}(v^*)$. It follows that [Eq. \(31\)](#) is maximized when $d_{t,c}^i(v^*) = d_{t,c}(v^*)$ for some $i \in \{1, 2, 3\}$. \square

6 Conclusions

The major open question is to obtain a substantial improvement over [Theorem 1.1](#) by establishing rapid mixing of the flip dynamics (or any other dynamics) for general graphs when $k > \alpha_1\Delta$ for a constant $\alpha_1 < 1.8$. If we restrict attention to triangle-free graphs the best known rapid mixing result holds for $k > \alpha_2\Delta$ where $\alpha_2 \approx 1.763\dots$ is the solution of $\alpha_2 = \exp(1/\alpha_2)$ using spectral independence [[FGYZ21](#), [CGŠV21](#)] (or assuming girth ≥ 5 using burn-in and local uniformity properties [[DFHV13](#)]). Can we utilize triangle-freeness to achieve similar bounds using a modified metric as in this paper? It would be interesting to see if the threshold α_2 is only an obstacle for certain proof techniques (which utilize properties of the stationary distribution) or if it corresponds to the onset of worst-case mixing obstacles for local Markov chains on locally dense graphs.

References

- [ALO20] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. In *Proceedings of the 61st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 1319–1330, 2020.
- [BCC⁺22] Antonio Blanca, Pietro Caputo, Zongchen Chen, Daniel Parisi, Daniel Štefankovič, and Eric Vigoda. On Mixing of Markov Chains: Coupling, Spectral Independence, and Entropy Factorization. In *Proceedings of the 33rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 3670–3692, 2022.
- [BD97] Russ Bubley and Martin E. Dyer. Path coupling: a technique for proving rapid mixing in Markov chains. In *Proceedings of the 38th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 223–231, 1997.
- [CDM⁺19] Sitan Chen, Michelle Delcourt, Ankur Moitra, Guillem Perarnau, and Luke Postle. Improved bounds for randomly sampling colorings via linear programming. In *Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2216–2234, 2019.
- [CFG⁺24] Xiaoyu Chen, Weiming Feng, Heng Guo, Xinyuan Zhang, and Zongrui Zou. Deterministic counting from coupling independence. Available from arXiv at: <https://arxiv.org/abs/2410.23225>, 2024.
- [CGŠV21] Zongchen Chen, Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Rapid mixing for colorings via spectral independence. In *Proceedings of the 32nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1548–1557, 2021.
- [CLMM23] Zongchen Chen, Kuikui Liu, Nitya Mani, and Ankur Moitra. Strong spatial mixing for colorings on trees and its algorithmic applications. In *Proceedings of the 64th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 810–845, 2023.
- [CLV21] Zongchen Chen, Kuikui Liu, and Eric Vigoda. Optimal mixing of Glauber dynamics: Entropy factorization via high-dimensional expansion. In *Proceedings of the 53rd Annual ACM Symposium on Theory of Computing (STOC)*, pages 1537–1550, 2021.
- [DFHV13] Martin E. Dyer, Alan M. Frieze, Thomas P. Hayes, and Eric Vigoda. Randomly coloring constant degree graphs. *Random Structures & Algorithms*, 43:181–200, 2013.
- [DG99] Martin Dyer and Catherine Greenhill. Random walks on combinatorial objects. *Surveys in Combinatorics*, page 101–136, 1999.
- [DJV02] Martin Dyer, Mark Jerrum, and Eric Vigoda. Rapidly mixing Markov chains for dismantlable constraint graphs. In *Randomization and Approximation Techniques in Computer Science (RANDOM)*, pages 68–77, 2002.
- [EHSV18] Charilaos Efthymiou, Thomas P. Hayes, Daniel Stefankovic, and Eric Vigoda. Sampling random colorings of sparse random graphs. In *Proceedings of the 29th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1759–1771, 2018.

[FGYZ21] Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. Rapid mixing from spectral independence beyond the Boolean domain. In *Proceedings of the 32nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1558–1577, 2021.

[HS07] Thomas P. Hayes and Alistair Sinclair. A general lower bound for mixing of single-site dynamics on graphs. *The Annals of Applied Probability*, 17(3):931–952, 2007.

[Hub15] Mark Huber. Approximation algorithms for the normalizing constant of Gibbs distributions. *The Annals of Applied Probability*, 25(2):974–985, 2015.

[HVV15] Thomas P. Hayes, Juan Carlos Vera, and Eric Vigoda. Randomly coloring planar graphs with fewer colors than the maximum degree. *Random Structures & Algorithms*, 47(4):731–759, 2015.

[Jer95] Mark Jerrum. A very simple algorithm for estimating the number of k -colorings of a low-degree graph. *Random Structures & Algorithms*, 7(2):157–165, 1995.

[Jer03] Mark Jerrum. *Counting, Sampling and Integrating: Algorithms and Complexity*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, second edition, 2003.

[JPV22] Vishesh Jain, Huy Tuan Pham, and Thuy-Duong Vuong. Spectral independence, coupling, and the spectral gap of the Glauber dynamics. *Information Processing Letters*, 177:106268, 2022.

[KF09] Daphne Koller and Nir Friedman. *Probabilistic Graphical Models: Principles and Techniques*. MIT Press, 2009.

[Kol18] Vladimir Kolmogorov. A faster approximation algorithm for the Gibbs partition function. In *Proceedings of the 31st Annual Conference on Learning Theory (COLT)*, volume 75, pages 1–22, 2018.

[Liu21] Kuikui Liu. From coupling to spectral independence and blackbox comparison with the down-up walk. In *Randomization and Approximation Techniques in Computer Science (RANDOM)*, pages 32:1–32:21, 2021.

[LP17] David A. Levin and Yuval Peres. *Markov chains and mixing times*. American Mathematical Society, 2017.

[LSS19] Jingcheng Liu, Alistair Sinclair, and Piyush Srivastava. Correlation decay and partition function zeros: Algorithms and phase transitions. In *Proceedings of the 60th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 1380–1404, 2019.

[MSW04] Fabio Martinelli, Alistair Sinclair, and Dror Weitz. Fast mixing for independent sets, colorings and other models on trees. In *Proceedings of the 15th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 456–465, 2004.

[Mur12] Kevin P. Murphy. *Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning)*. MIT Press, 2012.

[SS97] Jesús Salas and Alan D Sokal. Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem. *Journal of Statistical Physics*, 86(3–4):551–579, 1997.

[ŠV22] Daniel Štefankovič and Eric Vigoda. Lecture notes on spectral independence and bases of a matroid: Local-to-global and trickle-down from a Markov chain perspective. Available from arXiv at: <https://arxiv.org/abs/2307.13826>, 2022.

[ŠVV09] Daniel Štefankovič, Santosh Vempala, and Eric Vigoda. Adaptive simulated annealing: A near-optimal connection between sampling and counting. *Journal of the ACM*, 56(3):1–36, 2009.

[Vig99] Eric Vigoda. Improved bounds for sampling colorings. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 51–59, 1999.