Randomly coloring graphs of logarithmically bounded pathwidth

Shai Vardi*

Abstract

We consider the problem of sampling a proper k-coloring of a graph of maximal degree Δ uniformly at random. We describe a new Markov chain for sampling colorings, and show that it mixes rapidly on graphs of logarithmically bounded pathwidth if $k \geq (1+\epsilon)\Delta$, for any $\epsilon > 0$, using a new *hybrid paths* argument.

*California Institute of Technology, Pasadena, CA, 91125, USA. E-mail: `svardi@caltech.edu`.

1 Introduction

A (proper) k-coloring of a graph $G = (V, E)$ is an assignment $\sigma : V \to \{1, \ldots, k\}$ such that neighboring vertices have different colors. We consider the problem of sampling (almost) uniformly at random from the space of all k-colorings of a graph.[1] The problem has received considerable attention from the computer science community in recent years, e.g., [12, 17, 24, 26, 33, 44, 45]. It also has applications in Combinatorics (e.g., [5]) and Statistical Physics (e.g., [42]).

Sampling colorings (as well as other combinatorial objects, e.g., [6, 27, 37]) is commonly done using Markov Chain Monte Carlo (MCMC) methods. A large body of work on sampling colorings is devoted to analyzing a particular Markov chain, known as *Glauber dynamics*: Choose a vertex v uniformly at random; choose a color c uniformly at random from the set of available colors (the complement of the set of colors of the neighbors of v); recolor v with c. Jerrum [26] showed that the Glauber dynamics mix in time $O(n \log n)$ when $k > 2\Delta$, where Δ is the maximal degree of the graph. Vigoda [45] improved the bound on the number of colors to $k > 11\Delta/6$ using a different Markov chain, and showed that it mixes in time $O(nk \log n)$. This remains the best known bound on k for general graphs. A major open question is for what values of k can we sample colors efficiently (i.e., in polynomial time)? It is conjectured (e.g., [16]) that $k = \Delta + 2$ colors suffice, and furthermore, that the Glauber dynamics mix rapidly for any $k \geq \Delta + 2$.

A lot of work has focused on improving the bounds of Vigoda on restricted families of graphs. Dyer and Frieze [11] showed that if the maximal degree and girth are $\Omega(\log n)$, the Glauber dynamics mix in $O(n \log n)$ for $k > \alpha\Delta$, where $\alpha \approx 1.763$. The degree and girth requirements and the value of α were improved in a line of works [13, 20, 24, 25, 36]; see Table 1 for a comparison and summary of some milestones. The current state of the art results exhibit a tradeoff between the value of α and the degree and girth requirements. Hayes and Vigoda [24] showed that on graphs with $\Delta = \Omega(\log n)$ and girth at least 9, $(1 + \epsilon)\Delta$ colors suffice to ensure fast mixing. On the other hand, the Glauber dynamics have been shown to mix rapidly on graphs with girth at least 5 (resp. 6) and $\Delta > \Delta_0$ (where Δ_0 is some absolute constant) using roughly 1.763Δ (resp. 1.489Δ) colors [13]. Stronger bounds have only been shown on highly specialized families of graphs, such as trees [35], planar graphs [22], Erdös-Rényi graphs [12] and cubic graphs [33].

The two results that are closest to ours are the following: Hayes [21] showed that if for graphs with maximum eigenvalue ρ, the mixing time of the Glauber dyanmics is $O(n \log n)$, when $k > \Delta + c\rho$, where c is slightly greater than 1. Planar graphs, trees and graphs of bounded treewidth have a maximum eigenvalue of $O(\sqrt{\Delta})$, implying fast mixing for these graphs. Berger et al. [2] showed that for graph with bounded degree and logarithmically bounded cutwidth, the Glauber dynamics mix in polynomial time. We formally define treewidth, pathwidth and cutwidth later on, but we note that the pathwidth of any graph G is at least as large as its cutwidth, which is at least as large as its treewidth. In addition, the pathwidth is at most $O(\log n)$ times the treewidth.

1.1 Results and Techniques

Our main result is an algorithm that efficiently samples a $((1 + \epsilon)\Delta)$-coloring (almost) uniformly at random if the input graph has logarithmically bounded pathwidth, for any $\epsilon > 0$.[2]

[1] We define precisely what we mean by "almost" in Section 2.2.

[2] Assuming $(1 + \epsilon)\Delta \geq \Delta + 1$.

Table 1: Comparison of results on sampling k-colorings.

Degree	Girth	Graph family	$k >$	Dynamics	Mixing time	Reference
any	any	any	2Δ	Glauber	$O(n \log n)$	[26]
any	any	any	$(1.833\ldots)\Delta$	Flip	$O(n \log n)$	[45]
$\Omega(\log n)$	$\Omega(\log n)$	any	$(1.763\ldots)\Delta$	Glauber	$O(n \log n)$	[11]
$\Omega(\log n)$	≥ 9	any	$(1+\epsilon)\Delta$	Glauber	$O(n \log n)$	[24]
$\geq \Delta_0$ †	≥ 5	any	$(1.763\ldots)\Delta$	Glauber	$O(n \log n)$	[13]
$\geq \Delta_0$ †	≥ 6	any	$(1.489\ldots)\Delta$	Glauber	$O(n \log n)$	[13]
$O(1)$	∞	Trees	4	Glauber	poly(n)	[34]
any	any	treewidth= $O(1)$	$\Delta + \Omega\left(\sqrt{\Delta}\right)$	Glauber	poly(n)	[21]
$O(1)$	any	cutwidth= $O(\log n)$	$\Delta + 2$	Glauber	poly(n)	[2]
any	any	pathwidth $= O(\log n)$	$(1+\epsilon)\Delta$	Single-Flaw	poly(n)	Here

† Δ_0 is some absolute constant.

Theorem 1.1. *(Informal) Let $\epsilon > 0$ and G be a graph with n vertices and maximal degree Δ. There exists an algorithm for sampling a $((1+\epsilon)\Delta)$-proper coloring of G (almost) uniformly at random, whose running time is polynomial in n, Δ, and $\epsilon^{-O(\mathrm{pw}(G))}$, where $\mathrm{pw}(G)$ is the pathwidth of G. In particular, when $\mathrm{pw}(G) = O(\log n)$, the algorithm runs in polynomial time.*

The two main methods of bounding the mixing time of random walks are *coupling* and bounding the *spectral gap* of the transition matrix [19]; methods of bounding the conductance or congestion (which are used to bound the spectral gap) are generally considered to be stronger than coupling methods [19,30]. Despite this, most of the work on sampling colorings uses various coupling methods [10–13,20,22–26,36,45]; the Glauber dynamics do not lend themselves easily to the techniques that are usually used for bounding the spectral gap. In particular, bounding the *congestion* of the underlying graph of the transition matrix[3] typically involves defining flows between states in the underlying graph. If one can describe a flow in the underlying graph such that the congestion of each edge is not too large, it implies that the spectral gap is large and the Markov chain mixes rapidly [9,43]. It is not clear how to construct such flows for the Glauber dynamics, as a transition involves changing a vertex's color to one of its available colors: if all of the neighbors of a vertex v that is colored `blue` are colored `red`, how do we go about changing v's color to `red`?

We introduce a new Markov chain, which we call *Single-Flaw dynamics*. The difference between the Single-Flaw and Glauber dynamics is that the Single-Flaw dynamics also allow colorings that have a "single flaw" – there is at least one monochromatic edge, and all monochromatic edges share a vertex. In other words, the coloring is not proper, but there is a single vertex v such that we can reach a proper coloring by changing v's color only. We call such colorings *singly-flawed*. Concretely, the Single-Flaw dynamics Markov chain is the following: choose a vertex v and a color c at random. If changing v's color to c results in a coloring that is either proper or singly-flawed, change v's color to c. Otherwise do not.

The main advantage afforded by this Markov chain is that it allows us to define paths between two states (colorings) α and β: select some order on the vertices, $v_1, \ldots, v_n$. Starting from v_1, for each vertex v_i, change its color to $\beta(v_i)$. If the transition leads to a proper coloring, continue to v_{i+1}. Otherwise, "fix" the monochromatic edges by recoloring the neighbors of v_i that are

[3] In the underlying graph of the transition matrix A of a Markov chain, the states are represented by vertices, there is an edge (u, v) between two states $u, v \in \Omega$ with weight $w_e = A_{u,v}$ if $A_{u,v} \neq 0$.

also colored $\beta(v_i)$ (as $k \geq \Delta + 1$ there is always at least one available color). When there are no monochromatic edges remaining, continue to vertex v_{i+1}. We note that it is possible to first "fix" the neighbors of v and only then change the color of v, and this corresponds to the Glauber dynamics; we only lose a single color choice per neighbor (in the Single-Flaw dynamics we can recolor with the original color of v, whereas in the Glauber dynamics we cannot). Therefore, our analysis can be easily modified to apply to the Glauber dynamics. We chose to analyze the Single-Flaw dynamics for two reasons (in addition to using one fewer color): 1) The analysis is cleaner this way, and 2) It is straightforward to extend the analysis techniques to "Multi-Flaw" dynamics, which we believe could prove useful in other sampling and counting settings as well.

We first describe a simple attempt to adapt the canonical paths argument of Jerrum and Sinclair [27, 28] to our setting, using the canonical paths described above. Although it fails in all non-trivial cases, it is instructive as it exemplifies an important part of our method. For every edge in the underlying graph, (try to) describe an injective function from paths going through the edge to the state space of the chain. If we can describe such a function, it would mean that at most $|\Omega|$ paths use each edge. If there were no "fixing phase" (i.e., the graph was disconnected), this would be easy: Let t be the transition from the state σ such that the color of the j^{th} vertex is changed to c. The injective function would map the path from α to β to the following coloring σ^*: for vertices $v_i : i = 1, 2, \ldots j$, $\sigma^* = \alpha(v_i)$, for all other vertices $v_i : i = j + 1, \ldots, n$, $\sigma^*(v_i) = \beta(v_i)$. The mapping is injective because (i) σ^* a proper coloring and (ii) knowing t and σ^* allows us to recover α and β, as

$$\sigma(v_i) = \begin{cases} \beta(v_i) & i = 1, 2, \ldots, j \\ \alpha(v_i) & \text{otherwise.} \end{cases}$$

This would imply that at most $|\Omega|$ paths would use each transition. We note that it is not necessary to show that *at most* $|\Omega|$ paths use each transition to show polynomial time mixing; it suffices to show that $|\Omega| \cdot \mathrm{poly}(n)$ paths use each transition [27]. If we are guaranteed that at every round, at most r vertices will be colored differently from α and β, this means that at most

$$(k-1)^r|\Omega| \tag{1}$$

paths use each transition. This observation was used by Berger et al. [2], who used the cutwidth to bound r. (The cutwidth of a graph G is the smallest integer r such that there exists a labeling $v_1, \ldots, v_n$ of the vertices such that for all $1 \leq k \leq n$, the number of edges from $\{v_1, \ldots, v_k\}$ to $\{v_{k+1}, \ldots, v_n\}$ is at most r.) We use a similar argument, but are able to leverage the pathwidth of the graph, which gives a stronger result, as the pathwidth of any graph is at least as large as its cutwidth. In essence, a small pathwidth implies that there is some order on the vertices, such that if our canonical paths obey this order, not too many vertices will need to be fixed at any time. We note that finding such an order is NP-hard [4], however we only need that such an order exists for the canonical paths argument.

The main contribution in this paper is removing the requirement that the maximal degree is bounded by a constant; in other words, removing the number of colors from the base in Expression (1). In order to do this, we use a multicommodity flow argument [9, 43]. Instead of specifying a single path for each pair of states, we describe a flow between them in the underlying graph. Whenever we fix a vertex's color, we split the flow evenly among all available options. This is similar to the argument of Morris and Sinclair [37], in which flow is also split up among different paths; it helps route the flow more evenly, thereby avoiding the case where one edge is heavily congested while other "available" edges are not. In contrast to [37], we only split the flow in the fixing stages; in fact, whenever a vertex v_i is colored $\beta(v_i)$, this consolidates the flow! Our technique can be thought of as a hybrid argument between the canonical paths proof technique of [27] and the multicommodity flow argument of [37]. We call these paths *hybrid paths* and believe they could be useful in other scenarios as well.

Finally, recall that the state space of the Single-Flaw dynamics includes flawed colorings. We show that there are not too many singly-flawed colorings relative to proper colorings, hence executing the chain polynomially many times will guarantee that we output a proper coloring w.h.p.

1.2 Related work

Most of the work on sampling colorings has focused on the Glauber dynamics (e.g., [12, 13, 17, 23, 26, 33, 34, 36, 38, 44]). Other Markov chains have been analyzed, notably the Flip dynamics of Vigoda [45], which is closely related to the chain proposed by Wang, Swendsen, and Kotecký [46]: when a vertex v is required to change color from c to c', the colors of the entire neighborhood that is colored with c and c' are flipped (the chain of Vigoda only performs some flips with some probability). It is also possible to sample colorings using approaches that do not use MCMC methods; for example, Efthymiou [14] proposed a combinatorial method for sampling colors that does not use a Markov chain, and used it to show that it is possible to sample colorings on $G(n, d/n)$ using $k > (1+\epsilon)d$ colors. A caveat is that the run time is only polynomial w.p. $1 - 2n^{-2/3}$. The main difference between Single-Flaw dynamics and other work on sampling colors is that we allow improper states. There are other Markov chains that also consist of "flawed" states that are not part of the space we wish to sample from. An example is the Markov chain for sampling perfect matchings in bipartite graphs, proposed by Broder [7] and analyzed by Jerrum and Sinclair [27] and Jerrum, Sinclair and Vigoda [28]: the chain consists of perfect matchings (of size n), and imperfect matchings of size $n-1$. An interesting distinction is that Broder's chain needs the imperfect matchings to transition between perfect matchings (otherwise, it is unclear how to transition). We do not need the imperfect states to show convergence: the Glauber dynamics are known to converge to the uniform distribution for $k \geq \Delta + 2$; we only use the imperfect colorings to bound the mixing time. We do get the additional benefit that the Single-Flaw dynamics are guaranteed to converge for $k \geq \Delta + 1$ (in contrast to the Glauber dynamics, which are not—a simple counterexample is a triangle graph); we do not leverage this in this work.

The method of bounding the conductance or congestion of the transition matrix of a Markov chain has been used to great success in sampling and counting of various problems e.g., [18, 28, 37]. To our knowledge, the only places that these types of arguments have been successfully applied to sampling colorings is in bounding the mixing time of the Glauber dynamics on trees with bounded degree [34] and hyperbolic graphs [2]. The arguments of [34] and [2] both rely heavily on the upper bound on the degree and it does not appear possible to extend their techniques to more general settings.

Sampling colorings corresponds to sampling configurations of the zero temperature k-state antiferromagnetic Potts model [39]. One can draw an analogy between our technique and temperature-tuned walks that also include higher energy levels (see e.g., [32]), though instead of walking at a fixed temperature, which would allow some Poisson-like distribution of the number of flaws, we allow exactly one flaw, and correct it before allowing the next flaw.

The terms *treewidth* and *pathwidth* were introduced by Robertson and Seymour [40, 41]; the concept of treewidth was discovered independently several times, and was originally introduced under a different name by Bertelè and Brioschi [3]. Of particular interest is a work by Chekuri, Khanna and Shepherd [8] that consider multicommodity flows on graphs of bounded treewidth.

Their techniques and results are incomparable to ours; they study the flow on graphs, while we use a flow to bound the congestion on the underlying graph of the Single-Flaw dynamics.

2 Preliminaries

We denote the set $\{1, 2, \ldots, m\}$ by $[m]$. Let $G = (V, E)$ be a graph, and denote $|V| = n$. We assume that the vertices of G are uniquely identified by $\{1, 2, \ldots, n\}$. For any (not necessarily simple) path p in G, let $|p|$ denote the length of p (i.e., the number of edges in p, where if an edge appears k times in p, it is counted k times).

2.1 Colorings

For any k-coloring of G, $\sigma : V \to [k]$, let $MCE(G) = \{(u, v) : \sigma(u) = \sigma(v)\}$ denote the set of monochromatic edges. σ is a proper coloring if $MCE(G) = \emptyset$. σ is a *singly-flawed coloring* if $MCE(G) \neq \emptyset$ and there is a vertex that is common to all edges in $MCE(G)$, i.e., $\exists v : \forall e \in MCE(G), v \in e$. We say that such a vertex v is a *flawed vertex* of σ. Note that a singly-flawed coloring has exactly two flawed vertices if $|MCE(G)| = 1$ and one flawed vertex otherwise. We denote the set of proper k-colorings of G colors by $\mathcal{C}_p(G, k)$ and the set of all singly-flawed colorings by $\mathcal{C}_{sf}(G, k)$. We drop G and k when they are clear from context. Let σ be a coloring. If, after recoloring some $v \in V$ with a color c, there is no monochromatic edge (u, v), we say that c is *available* to v in σ. Note that a color's availability does not depend on whether σ or the coloring obtained by recoloring v with c is proper, singly-flawed or otherwise.

We first show two results that will be useful later on, regarding proper and singly-flawed colorings: (1) the ratio of singly-flawed colorings to proper colorings is "not too large" (Corollary 2.2), and (2) there is a mapping from singly-flawed colorings to proper colorings, such that "not too many" singly-flawed colorings are mapped to any proper coloring (Corollary 2.3). Both results are corollaries of the following simple lemma.

Lemma 2.1. *For any $G = (V, E)$ such that $|V| = n$ and $k \geq \Delta + 2$, there exists a surjective function*

$$g : \mathcal{C}_p(G, k) \times [k] \times [n] \to \mathcal{C}_{sf}(G, k).$$

Proof. For every coloring $\sigma \in \mathcal{C}_p(G, k)$, every vertex $v \in V$ and every color $c \in [k]$, let

$$\sigma'_{c,v} = \begin{cases} \sigma(u) & \text{if } u \neq v \\ c & \text{if } u = v \end{cases}$$

If $\sigma'_{c,v} \in \mathcal{C}_{sf}(G, k)$, let $g(\sigma, c, v) = \sigma'_{c,v}$, otherwise let $g(\sigma, c, v)$ be some arbitrary coloring in $\mathcal{C}_{sf}(G, k)$. It is easy to see that every $\sigma' \in \mathcal{C}_{sf}(G, k)$ is in the range of g: the reverse operation of changing the color of a flawed vertex v in σ' to some available color gives a proper coloring. $\square$

The two corollaries that we require are the following.

Corollary 2.2. *For any $G = (V, E)$ such that $|V| = n$ and $k \geq \Delta + 2$,*

$$|\mathcal{C}_{sf}(G, k)| \leq kn|\mathcal{C}_p(G, k)|.$$

Proof. Immediate from the surjectivity of the function g in Lemma 2.1. $\square$

Corollary 2.3. *For any $G = (V, E)$ such that $|V| = n$ and $k \geq \Delta + 2$, there exists a function*

$$g' : \mathcal{C}_{sf}(G, k) \to \mathcal{C}_p(G, k),$$

for which each element in the co-domain has at most kn pre-images in the domain.

Proof. For every $\sigma' \in \mathcal{C}_{sf}$, arbitrarily select one pre-image (σ, v, c) w.r.t. g, and set σ as the image for σ' under g'. □

2.2 Markov chains and rapid mixing

In this section we review some of the results on the mixing time of Markov chains that we will require. The reader is referred to [31] for an excellent introduction to Markov chains and modern techniques on bounding their mixing time.

Consider a discrete-time Markov chain $\mathcal{MC}$ with finite state space Ω and symmetric transition probability matrix P (i.e., $P(\sigma, \sigma') = P(\sigma', \sigma)$ for all $\sigma, \sigma' \in \Omega$). The chain is said to be *irreducible* if for every pair of states $\sigma, \sigma' \in \Omega$, there exists some t such that $P^t(\sigma, \sigma') > 0$; in other words, it is possible to get from any state to any state using a finite number of transitions. It is *aperiodic* if for any $\sigma, \in \Omega$, $\gcd\{t : P^t(\sigma, \sigma) > 0\} = 1$. It is *lazy* if for all $\sigma \in \Omega, P(\sigma, \sigma) > 1/2$. A fundamental theorem of stochastic processes states that an irreducible and aperiodic Markov chain converges to a unique *stationary distribution* π over Ω, i.e., $\lim_{t\to\infty} P^t(\sigma, \sigma') = \pi(\sigma')$ for all $\sigma, \sigma' \in \Omega$. If in addition P is symmetric, then π is uniform over Ω (e.g., [1]).

Our goal is to describe a *fully-polynomial almost uniform sampler for proper colorings*; namely, a randomized algorithm that, given as inputs a graph $G = (V, E)$ and a bias parameter δ, outputs a random proper coloring of G from a distribution D that satisfies $d_{TV}(D, U) \leq \delta$, where U is the uniform distribution on the proper colorings of G and d_{TV} is the total variation distance, defined as follows:[4] For any two distributions μ, ν on Ω,

$$d_{TV}(\mu, \nu) = \max_{S \subseteq \Omega} |\mu(S) - \nu(S)|. \tag{2}$$

We are interested in the rate at which a Markov chain converges to its stationary distribution π. We define the *mixing time* from a state σ to be

$$\tau_\sigma(\delta) = \min\{\bar{t} : d_{TV}(P^t(\sigma, \cdot), \pi) \leq \delta \;\; \text{for all } t \geq \bar{t}\}, \tag{3}$$

We further define the mixing time of the Markov chain to be $\tau(\delta) = \max_\sigma \tau_\sigma(\delta)$. We say that a Markov chain is *rapidly mixing* if $\tau(1/2e)$ is polynomial in n. The constant $1/2e$ is arbitrary, as a bound on $\tau(1/2e)$ implies a bound on $\tau(\delta)$ for any $\delta > 0$ (e.g., [1]):

$$\tau(\delta) \leq (1 - \log \delta) \cdot \tau(1/2e).$$

In order to bound the mixing time, we describe a multicommodity flow on the underlying graph $H = (\Omega, F)$ of the Markov chain, where $F = \{(\sigma, \sigma'), P(\sigma, \sigma') > 0\}$ is the set of all transitions that have positive probability.

We denote by

$$q(\sigma, \sigma') = \pi(\sigma) P(\sigma, \sigma'), \tag{4}$$

[4] Alternatively, we can define it as $d_{TV}(\mu, \nu) = \frac{1}{2} \sum_{\sigma \in \Omega} |\mu(\sigma) - \nu(\sigma)|$. It is easy to verify that the two definitions are equivalent.

the *ergodic flow* through the edge (σ, σ') of H (an intuitive way to think about $q(\sigma, \sigma')$ is the probability of traversing edge (σ, σ') at stationarity.)

For all ordered pairs $(\alpha, \beta) \in \Omega^2$, let $\mathcal{P}_{\alpha,\beta}$ denote a set of (not necessarily simple) directed paths from α to β in H. A *flow* is a function $f : \mathcal{P} \to \mathcal{R}^+ \cup \{0\}$ where $\mathcal{P} = \bigcup_{\alpha,\beta} \mathcal{P}_{\alpha,\beta}$ that satisfies

$$f_{\alpha,\beta} \equiv \sum_{p \in \mathcal{P}_{\alpha,\beta}} f(p) = \pi(\alpha)\pi(\beta), \tag{5}$$

for every $\alpha, \beta \in \Omega$. We define the congestion on an edge (σ, σ') with respect to a flow f by

$$\rho_f(\sigma, \sigma') = \frac{1}{q(\sigma, \sigma')} \sum_{\alpha,\beta \in \Omega} \sum_{p : (\sigma,\sigma') \in p \in \mathcal{P}_{\alpha,\beta}} f(p)|p|, \tag{6}$$

and the congestion of f by

$$\rho_f = \max_{(\sigma,\sigma') \in F} \rho_f(\sigma, \sigma').$$

We use the following theorem, due to Sinclair [43] and Diaconis and Stroock [9], that relates the mixing time to the congestion of a flow. Note that it holds for any flow; in order to bound the mixing time, we need to find *some* flow that has low congestion.

Theorem 2.4. *[43] For any irreducible, aperiodic, lazy and symmetric Markov chain $\mathcal{MC}$ with transition matrix P on state space Ω, any flow f on the underlying graph of $\mathcal{MC}$, and any state $\sigma_0 \in \Omega$,*

$$\tau_{\sigma_0}(\delta) \le \rho_f \left(\ln \pi(\sigma_0)^{-1} + \ln \delta^{-1} \right).$$

2.3 Pathwidth and vertex separation

A *path decomposition* of a graph $G = (V, E)$ is a path P with m nodes, where each node of P represents a subset of V: $X_i \subseteq V, i \in [m]$ such that the following hold:

1. $\bigcup_{i=1}^m X_i = V$,
2. For every $(u, v) \in E$, $u, v \in X_i$ for some $i \in [m]$.
3. For all $i, j, k \in [m]$, if X_j is on the (unique) path between X_i and X_k, then $X_i \cap X_k \subseteq X_j$.

The first requirement guarantees that every vertex of G is in at least one node of P, the second that every two neighboring vertices in G share at least one node in P, and the third that if some vertex $v \in V$ is in both X_i and X_k, it is in every node of the path between X_i and X_k in P. The width of P is defined as $\max_{i \in [m]}\{|X_i| - 1\}$. The *pathwidth* of G, denoted $\mathrm{pw}(G)$, is the minimal ω such that there exists some path-decomposition of G with width ω. The *treewidth* is similarly defined, when P is a tree.

A linear ordering of a graph $G = (V, E)$ is a bijective mapping of vertices to integers; $L : V \to \{1, 2, \dots, n\}$. Given a graph G, a linear ordering L, and an integer $j \in \{1, \dots, n\}$, let A_j be the set of vertices mapped to the integers $1, \dots, j$ by L; i.e., $A_j = \{v : L(v) \le j\}$. Let B_j denote the set of vertices that are mapped to integers greater than j by L: $B_j = V \setminus A_j$. A *minimal vertex separator* for an index $j \in \{1, \dots n\}$ (denoted $\mathrm{MVS}(G, L, j)$) is a minimal set of vertices $S_j \subset V$ such that the following hold:[5]

[5]Traditionally, a vertex separator for j is defined such that the separating subset appears before j in the ordering [15]. This is identical to our definition with the order inverted.

1. For all $u \in S_j$, $L(u) > j$.

2. A_j and $B_j \setminus S_j$ are disconnected; that is, there is no edge $(u, v) \in E$ such that $u \in A_j, v \in B_j \setminus S_j$.

W.l.o.g., we henceforth assume for simplicity that L is the identity ordering; i.e., $\forall j \in \{1, \ldots, n\}$, $L(j) = j$.

The *vertex separator number* for a graph G, denoted $\mathrm{VSN}(G)$ is the minimal size of the largest minimal vertex separator over all possible linear orderings L of G.

$$\mathrm{VSN}(G) = \min_L \max_{j \in \{1, \ldots, n\}} \{\mathrm{MVS}(G, L, j)\}.$$

We call an order L for which $\mathrm{VSN}(G)$ is minimized a *minimal order* for G.

We require the following claim and theorem.

Claim 2.5. *Let G and L be a graph and a linear order thereof. For $j = 1, \ldots, n$, set $S_j = \mathrm{MVS}(G, L, j)$. For all $j \in \{2, \ldots, n\}$,*

$$S_{j-1} \subseteq S_j \cup \{j\}.$$

Proof. For any $j \in \{1, \ldots, n\}$, S_j is uniquely determined: it is exactly the neighbors of the vertices of A_j in B_j. This is because for every $v \in A_j$, the set $N(v) \cap B_j$ must all be in S_j, otherwise there is an edge between v and B_j. From the minimality of S_j, therefore, $S_j = \bigcup_{v \in A_j} N(v) \cap B_j$. Hence,

- If $j \in S_{j-1}$ then $S_j = S_{j-1} \setminus \{j\} \cup (N(j) \cap B_j)$.
- If $j \notin S_{j-1}$ then $S_j = S_{j-1} \cup (N(j) \cap B_j)$.

This completes the proof of the claim. $\square$

Theorem 2.6. *[29] For any graph G,* $\mathrm{VSN}(G) = \mathrm{pw}(G)$.

3 Single-Flaw dynamics

Let $G = (V, E)$ be a graph with maximal degree Δ and $\epsilon > 0$ such that $(1 + \epsilon)\Delta \geq \Delta + 2$. The state space Ω of Markov chain $\mathcal{MC}(G, \epsilon)$ (or simply $\mathcal{MC}$) is the set of all proper and singly-flawed k-colorings of G, for $k = \lceil (1 + \epsilon)\Delta \rceil$: $\Omega = \mathcal{C}_p(G, k) \cup \mathcal{C}_{sf}(G, k)$. For simplicity, we henceforth assume that $\epsilon\Delta$, $(1 + \epsilon)\Delta$ and $(1 + \epsilon)\epsilon^{-1}$ are integers. It is easy to generalize the results to real values thereof. For $\sigma \in \Omega$, the transitions $\sigma \to \sigma'$ of $\mathcal{MC}$ are the following

- Let $\sigma' = \sigma$.
- With probability 1/2, do nothing (laziness).
- Otherwise, choose a vertex v and color c uniformly at random from V and $[k]$ respectively. Tentatively, set $\sigma'(v) = c$.
- If $\sigma' \notin \Omega$, set $\sigma'(v) = \sigma(v)$.

It is easy to verify that the chain is irreducible, aperiodic, lazy and symmetric; hence the conditions of Theorem 2.4 hold, and it remains to describe a flow with low congestion of the underlying graph of $\mathcal{MC}$. Our main result describes such a flow.

Lemma 3.1. *Let $\epsilon > 0$ and G be a graph with maximal degree Δ. Then there exists a flow f on the underlying graph of the Markov chain $\mathcal{MC}(G,\epsilon)$ such that*

$$\rho_f \leq 8\,\mathrm{pw}(G)(1+\epsilon)^3\Delta^3 n^5 \left((1+\epsilon)\epsilon^{-1}\right)^{2\,\mathrm{pw}(G)}.$$

There are at most k^n possible colorings of G. Because the stationary distribution is uniform, for all $\sigma \in \Omega$, $\pi(\sigma) \geq \frac{1}{k^n}$, hence $\ln \pi(\sigma)^{-1} = O(n \log n)$.

Theorem 2.4 and Lemma 3.1 imply the following theorem, which is our main result.

Theorem 3.2. *Let $\epsilon > 0$ and G be a graph with maximal degree Δ. The mixing time of $\mathcal{MC}(G,\epsilon)$ satisfies*

$$\tau(\delta) = O\left(\mathrm{pw}(G)(1+\epsilon)^3\Delta^3 n^5 \left((1+\epsilon)\epsilon^{-1}\right)^{2\,\mathrm{pw}(G)} \left(n\log n + \ln \delta^{-1}\right)\right).$$

In particular, if $\mathrm{pw}(G) = O(\log n)$, the chain mixes in polynomial time.

In order to prove Lemma 3.1, we design a flow for $\mathcal{MC}$. To do so, we first describe the set of paths that we will route the flow through.

3.1 Hybrid paths

Let L be some minimal order of G. For simplicity and w.l.o.g. we assume that L is the identity order, i.e., $L(i) = i$ for all $i \in [n]$. We remark that we do not need to explicitly find L; we only require its existence for the hybrid paths argument. Let $\lambda = \frac{\mathrm{VSN}(G)}{\log n}$. If $\mathrm{pw}(G) = O(\log n)$, as is assumed here, λ is a constant.

For each $j \in [n]$, let $S_j = \mathrm{MVS}(G, L, j)$ be a minimal vertex separator. We denote the set of hybrid paths from α to β by $\gamma_{\alpha,\beta}$. *For the rest of this subsection and the next, we assume that α and β are both proper* colorings; we will extend the sets of paths to include ones that start and/or end at singly-flawed colorings in Section 3.3. We divide each path into n phases, where phase j consists of $|S_j| + 1$ steps, for a total of $\ell = n + \sum_{j=1}^{n} |S_j| \leq (\lambda + 1)(n \log n)$ steps. Each step is a recoloring of some vertex; it is possible that a vertex is "recolored" with the same color. In that case, the state (coloring) does not change, but we still count this redundant recoloring as a step, as it guarantees that all paths are of the same length; this will help to make the analysis more concise. A state that appears at the start of the ℓ^{th} step of the j^{th} phase of a hybrid path in $\gamma_{\alpha,\beta}$ is said to be at *distance* (j,ℓ) from α; alternatively, we say that it happens at *time* (j,ℓ). We denote the states of $\gamma_{\alpha,\beta}$ that are at distance (j,ℓ) from α by $\Phi_{\alpha,\beta}(j,\ell)$. The set of hybrid paths $\gamma_{\alpha,\beta}$ can be thought of as a layered graph, where all states of $\Phi_{\alpha,\beta}(j,\ell)$ are placed in the same layer.

3.1.1 A phase of the hybrid paths

We describe a single phase of $\gamma_{\alpha,\beta}$. For any $j \in [n]$, all states in $\Phi_{\alpha,\beta}(j,1)$ are proper colorings. Note that $\Phi_{\alpha,\beta}(1,1) = \{\alpha\}$. For **the first step** of the j^{th} phase, for every $\sigma_i \in \Phi_{\alpha,\beta}(j,1)$, set

$$\sigma_i'(v) = \begin{cases} \sigma_i(v) & \text{if } v \neq j \\ \beta(j) & \text{if } v = j \end{cases}.$$

We therefore have that $\Phi_{\alpha,\beta}(j,2) = \bigcup_i \sigma_i'$. It is clear that there is only one way to route the flow entering σ_i: it is all routed to σ_i' on (σ_i, σ_i'). It is possible that flow becomes consolidated in this step: the flow from all states $\sigma_i \in \Phi_{\alpha,\beta}(j,1)$ that differ only in the j^{th} coordinate is routed to the same σ_i'. Note that every $\sigma' \in \Phi_{\alpha,\beta}(j,2)$ is either a proper or a singly-flawed coloring.

The ℓ^{th} step in the j^{th} phase, $\ell \in \{2, 3, \ldots |S_j| + 1\}$ is a *splitting step*, and is the following: Let u_ℓ be the (lexicographically) $(\ell - 1)^{th}$ vertex of S_j. For every $\sigma_i \in \Phi_{\alpha,\beta}(j, \ell)$, let $C_i(u_\ell)$ be the set of colors available to u_ℓ under σ_i. For each $\sigma_i \in \Phi_{\alpha,\beta}(j, \ell)$ and color $c \in C_i(u_\ell)$, let

$$\sigma'_{i,c}(v) = \begin{cases} \sigma_i(v) & \text{if } v \neq u_\ell \\ c & \text{if } v = u_\ell \end{cases}.$$

We have that $\Phi_{\alpha,\beta}(j, \ell + 1) = \bigcup_i \sigma'_{i,c}$. From each state $\sigma_i \in \Phi_{\alpha,\beta}(j, \ell)$, the flow is split evenly among the transitions (i.e., a $1/|C_i(u_\ell)|$ fraction of the flow entering σ_i is routed on each $(\sigma_i, \sigma'_{i,c})$). Note that all states in $\Phi_{\alpha,\beta}(j, |S_j| + 2)$ are proper colorings, as any edge that may have been monochromatic in any $\sigma \in \Phi_{\alpha,\beta}(j, 2)$ will have been recolored. For all $1 \leq j < n$, set $\Phi_{\alpha,\beta}(j + 1, 1) = \Phi_{\alpha,\beta}(j, |S_j| + 2)$. Note that $S_n = \emptyset$ and $\Phi_{\alpha,\beta}(n, 2) = \{\beta\}$.

Because for all j, $|S_j| = O(\log n)$, given α, β, j and ℓ, the color of most vertices in $\Phi_{\alpha,\beta}(j, \ell)$ is uniquely determined. In particular, for $\ell > 1$, denote $A_j = \{v : v \leq j\}$ and $B_j = V \setminus (A_j \cup S_j)$.[6] It must hold that for any $\sigma \in \Phi_{\alpha,\beta}(j, \ell)$, $v_a \in A_j$, and $v_b \in B_j$, $\sigma(v_a) = \beta(v_a)$ and $\sigma(v_b) = \alpha(v_b)$. This is because all the vertices in A_j have been recolored to β and will not be recolored again, while the vertices in B_j have no neighbors in A_j, hence they have not been recolored yet.

We denote by $QS(\alpha, \beta, j, \ell)$ (QS stands for "quantum set") the number of vertices whose color is not uniquely defined by α, β, j, ℓ. By Claim 2.5, it holds that $QS(\alpha, \beta, j, \ell) \subseteq S_j$, but equality does not necessarily hold: assume that $S_{j-1} \subset S_j$, and let $u \neq j$ be some vertex in $S_j \setminus S_{j-1}$. Then u's color is still $\alpha(u)$ at time $(j, 2)$, as it has not yet been recolored, even though $u \in S_j$.

We note that $QS(\alpha, \beta, j, \ell)$ does not in fact depend on α or β. In fact,

Observation 3.3. *$QS(\alpha, \beta, j, \ell)$ is uniquely determined by either*

1. *j and ℓ, or*
2. *j and (σ, σ').*

Proof. For any α, β, the same vertex is recolored at (j, ℓ): at $\ell = 1$, vertex j is recolored; in all other instances, u_ℓ is recolored by at least $\epsilon\Delta$ different colors, regardless of α, β. Further, note that once a vertex u is in such a "quantum state", it will remain in quantum state until it is colored $\beta(u)$ at distance $(u, 1)$. Therefore, although we cannot recover the exact transitions used without knowledge of α, β, the set of vertices whose color is unknown at any given time is fixed. For the second observation, notice that (σ, σ') recolors some specific vertex u_ℓ (even if it is an idle recoloring), hence ℓ can be inferred. □

Due to Observation 3.3, we sometimes refer to $QS(\alpha, \beta, j, \ell)$ by $QS(j, (\sigma, \sigma'))$ or $QS(j, \ell)$.

3.2 Bounding the flow

Let α and β be proper colorings, $t = (\sigma, \sigma') \in F$ be some transition, and $j \in [n]$ be an integer. Denote the flow routed through t from α to β in phase j by $f_{j,t,\alpha,\beta}$. Note that we are only considering the flow routed through t **during phase** j; it is possible that the hybrid path passes through t in several phases, possibly carrying a different flow each time. Intuitively, it seems natural that after the flow was split evenly several times, "not too much" flow is routed through any state,

[6] For $\ell = 1$, we consider $(j - 1, |S_{j-1}| + 2)$ instead of $(j, 1)$, unless $j = 1$, in which case the vertices are all colored by α.

as it has a specific combination of colors of the vertices of $QS(j,t)$, and many such combinations are possible. It is not straightforward to show this, however, as the colors of different vertices in each state are not independent. The difficulty is compounded by the fact that flow is consolidated at the first step of every phase. Nevertheless, we can prove the following claim by rearranging the vertices of $QS(j,t)$ and using an inductive reasoning on this new order.

Claim 3.4. *The flow routed from α to β through any $t = (\sigma, \sigma') \in F$ in any phase $j \in [n]$ is at most*

$$f_{j,t,\alpha,\beta} \leq \frac{\pi(\alpha)\pi(\beta)}{(\epsilon\Delta)^{|QS(j,t)|}}.$$

Proof. From Observation 3.3, (j,t) uniquely defines (j,ℓ). If no flow is routed from α to β through t in phase j, the claim is trivially satisfied. Otherwise, we show that

$$f_{j,\ell,\alpha,\beta} \leq \frac{\pi(\alpha)\pi(\beta)}{(\epsilon\Delta)^{|QS(j,\ell)|}},$$

where $f_{j,\ell,\alpha,\beta}$ is the maximal flow from α to β through any t at distance (j,ℓ) from α. Order the vertices of $QS(j,\ell)$ in reverse order of the time since their last color change (possibly a null color change). That is, the vertex whose color changed most recently is last in the order. Let $M = |QS(j,\ell)|$ and relabel the vertices of $QS(j,\ell)$ by $1,\ldots,M$ according to their place in this order. Similarly, relabel $\Phi_{\alpha,\beta}(j',\ell')$, by $\Phi_1,\ldots,\Phi_M$, where Φ_i is the set of states at the time just *after* vertex i last changed its color. In other words, $\Phi_M = \Phi_{\alpha,\beta}(j,\ell+1)$,$\Phi_{M-1} = \Phi_{\alpha,\beta}(j,\ell)$, and so on. It is possible that for some m, Φ_m corresponds to states in the previous phase, i.e., $\Phi_m = \Phi_{\alpha,\beta}(j-1,\ell')$. Note that we drop the α,β from the notation for clarity, but we are still only considering the flow from α to β.

We now show by that for any set of $m \leq M$ colors $c_1,\ldots,c_m$, at most $\frac{\pi(\alpha)\pi(\beta)}{(\epsilon\Delta)^m}$ flow is routed into $\{\sigma \in \Phi_m : \sigma(i) = c_i, i \in [m]\}$. In other words, fix the colors $c_1,\ldots,c_m$. We want to bound the flow that passes through (into) the states of Φ_m, where vertices $1,\ldots,m$ are colored with $c_1,\ldots,c_m$ respectively. We do this by induction on m.

The base case:

The total flow from α to β through Φ_i, for any i, is exactly $\pi(\alpha)\pi(\beta)$. For any $c_1 \in [k]$, at most $\frac{\pi(\alpha)\pi(\beta)}{\epsilon\Delta}$ flow is routed through $\{\sigma \in \Phi_1 : \sigma(1) = c_1\}$. This is because the last time vertex 1 changed color, at most $1/\epsilon\Delta$ of all the flow was routed to states where v's color is c_1.

The inductive step:

From the inductive hypothesis, at most $\frac{\pi(\alpha)\pi(\beta)}{(\epsilon\Delta)^{m-1}}$ flow is routed through $\{\sigma \in \Phi_{m-1} : \sigma(i) = c_i, i \in [m-1]\}$. From the construction of the hybrid paths, for each of these states, at most $1/\epsilon\Delta$ of the flow entering it flows to a state where vertex m is colored c_m. □

We want to bound the total flow through a transition in any single phase. The following set of recoloring functions χ is useful. Let C be a set of (available) colors. χ_C is a function, parameterized by C, that takes as an input a color $c \in [k]$. Its output is a color from C, such that each color in C has the same number of pre-images, up to one. We do not explicitly define χ, only note that such a set of functions exists. For example, if $k = 13$, $C = \{1,2,3,4,5\}$, χ_C could allocate (c mod 5) + 1 to every $c \in [k]$, giving each color in C either two or three pre-images. We make the following observation (recall we assume $\epsilon\Delta$, $(1+\epsilon)\Delta$ and $(1+\epsilon)\epsilon^{-1}$ are integers).

Observation 3.5. *For χ_C as defined above, if $k = (1+\epsilon)\Delta$ and $|C| \geq \epsilon\Delta$, each color in C has at most $(1+\epsilon)\epsilon^{-1}$ pre-images in $[k]$.*

Armed with Claim 3.4 and the functions χ, we are now ready to bound the total flow through a transition in any single phase.

Lemma 3.6. *The flow $f_{j,t}$ routed through any $t = (\sigma, \sigma') \in F$ in any phase $j \in [n]$ satisfies*

$$f_{j,t} \le \pi(\cdot)^2 |\mathcal{C}_p| \cdot \left((1+\epsilon)\epsilon^{-1}\right)^{2|S_j|},$$

where $\pi(\cdot)$ is the probability of any state at stationarity.

Proof. For each (j,t), where $j \in [n]$ and $t = (\sigma, \sigma') \in F$, denote by $\texttt{pairs}_{j,t}$ the set of pairs of states $\alpha, \beta \in \mathcal{C}_p^2$ whose paths pass through t in phase j. We describe a function $\mu_{j,t}$ whose domain is $\texttt{pairs}_{j,t}$. We view the co-domain of $\mu_{j,t}$ as the Cartesian product of 3 sets X, Y and Z: $\mu_{j,t} : \texttt{pairs}_{j,t} \to X \times Y \times Z$; the output of $\mu_{j,t}(\cdot)$ is a triple (x, y, z). The function will be injective, therefore the size of the co-domain of $\mu_{j,t}$ will serve as an upper bound to $|\texttt{pairs}_{j,t}|$.

The sets X, Y, Z are the following.

- X is the set of all proper colorings. Assume that the input to $\mu_{j,t}$ is some pair (α, β). In the coloring specified by x, all vertices in A_j[7] are colored by α. All vertices in B_j are colored by β. Before specifying the coloring of S_j under x, note that already, together with j and t, this allows us to deduce α completely on all vertices in $V \setminus QS(j,t)$ and β on all vertices $V \setminus S_j$. To determine the colors of S_j in x, we color them one at a time, using χ. This information, while not characterizing $\beta(v)$ completely for $v \in S_j$, allows allows us to restrict the possible value of $\beta(v)$ to a set of size at most $(1+\epsilon)\epsilon^{-1}$ possible values.

- Y is $\left[(1+\epsilon)\epsilon^{-1}\right]^{|S_j|}$, allowing us to pinpoint $\beta(v)$ for every $v \in S_j$.

- Finally, Z is simply all possible colorings of the vertices of $QS(j,t)$ under α.

Clearly x, y, z, j and t allow us to recover α and β. The size of the co-domain is at most

$$|\mathcal{C}_p| \cdot \left((1+\epsilon)\epsilon^{-1}\right)^{|S_j|} \cdot k^{|QS(j,t)|}.$$

Combining with Claim 3.4 we get that the total flow through any transition t at phase j is at most

$$\begin{aligned}
f_{j,t} &\le |\mathcal{C}_p| \cdot \left((1+\epsilon)\epsilon^{-1}\right)^{|S_j|} \cdot k^{|QS(j,t)|} \cdot \frac{\pi(\cdot)\pi(\cdot)}{(\epsilon\Delta)^{|QS(j,t)|}} \\
&= \pi(\cdot)^2 |\mathcal{C}_p| \left((1+\epsilon)\epsilon^{-1}\right)^{|S_j|} \cdot \frac{((1+\epsilon)\Delta)^{|QS(j,t)|}}{(\epsilon\Delta)^{|QS(j,t)|}} \\
&\le \pi(\cdot)^2 |\mathcal{C}_p| \cdot \left((1+\epsilon)\epsilon^{-1}\right)^{2|S_j|},
\end{aligned}$$

where the last inequality is because $QS(j,t) \subseteq S_j$ for any t. □

[7] As before, $A_j = \{v : v \le j\}$, except for $\ell = 0$, for which $A_j = \{v : v < j\}$.

3.2.1 The congestion of an edge

We are ready to prove our main result of the section, that the congestion of any edge $(\sigma, \sigma') \in F$ under the flow defined by the hybrid paths from proper coloring to proper colorings, is polynomial in the number of vertices.

Lemma 3.7. *The congestion of any transition t under f, when f is restricted to flows from proper colorings to proper colorings, satisfies*

$$\rho_f(t) \leq 2k(\lambda+1)n^3 \log n \left((1+\epsilon)\epsilon^{-1}\right)^{2\,\mathrm{pw}(G)}.$$

Proof. From the definition of the congestion on an edge (Equation (6)), we have

$$\begin{aligned}
\rho_f(t) &= \frac{1}{q(t)} \sum_{\alpha,\beta\in\Omega} \sum_{p:t\in p\in\mathcal{P}(\alpha,\beta)} f(p)|p| \\
&\leq \frac{(\lambda+1)n\log n}{q(t)} \sum_{\alpha,\beta\in\Omega} \sum_{p:t\in p\in\mathcal{P}(\alpha,\beta)} f(p) && \text{(7a)} \\
&= 2|\Omega|k(\lambda+1)n^2 \log n \sum_{\alpha,\beta\in\Omega} \sum_{p:t\in p\in\mathcal{P}(\alpha,\beta)} f(p) && \text{(7b)} \\
&= 2|\Omega|k(\lambda+1)n^2 \log n \sum_{j=1}^{n} \sum_{\alpha,\beta\in\Omega} f_{j,t,\alpha,\beta} && \text{(7c)} \\
&= 2|\Omega|k(\lambda+1)n^2 \log n \sum_{j=1}^{n} f_{j,t} \\
&\leq 2|\Omega|k(\lambda+1)n^2 \log n \sum_{j=1}^{n} \pi(\cdot)^2 |\mathcal{C}_p| \cdot \left((1+\epsilon)\epsilon^{-1}\right)^{2|S_j|} && \text{(7d)} \\
&= \frac{2|\mathcal{C}_p|k(\lambda+1)n^2\log n}{|\Omega|} \sum_{j=1}^{n} \left((1+\epsilon)\epsilon^{-1}\right)^{2|S_j|} \\
&\leq 2k(\lambda+1)n^3 \log n \left((1+\epsilon)\epsilon^{-1}\right)^{2\,\mathrm{pw}(G)}.
\end{aligned}$$

Inequality (7a) is because the length of any hybrid path is at most $(\lambda+1)n\log n$; Equality (7b) is due to the definition of q: $q(\sigma,\sigma') = \pi(\sigma)P(\sigma,\sigma')$, where $\pi(\sigma) = |\Omega|^{-1}$ and $P(\sigma,\sigma') = (2kn)^{-1}$; Equality (7c) is simply a rephrasing that holds because

$$\sum_{\alpha,\beta\in\Omega} \sum_{p:t\in p\in\mathcal{P}(\alpha,\beta)} f(p)$$

is the flow through t under f; Inequality (7d) is due to Lemma 3.6. The final inequality is due to Theorem 2.6, as the pathwidth of a graph equals its vertex separation number. ☐

3.3 Mixing time

Lemma 3.7 applies to the congestion from flow between proper colorings only. We extend this result to all of f. We rephrase our main lemma:

Lemma 3.1. *The congestion of any transition t under f satisfies*

$$\rho_f(t) \leq 8k^3(\lambda+1)n^5 \log n \left((1+\epsilon)\epsilon^{-1}\right)^{2\,\mathrm{pw}(G)}.$$

Proof. We use the function g' from singly-flawed to proper colorings described in Corollary 2.3 to define the flows that have a singly-flawed coloring as (at least) one of their endpoints. For every α, β such that $\alpha \in \mathcal{C}_{sf}$ and $\beta \in \mathcal{C}_p$, we route the entire flow on the transition $(\alpha, g'(\alpha))$ and then proceed using the canonical paths described above for routing the flow from $g'(\alpha)$ to β. If $\alpha \in \mathcal{C}_p$, and $\beta \in \mathcal{C}_{sf}$, we route the flow from α to $g'(\beta)$ using the canonical paths above and then on the edge $(g'(\beta), \beta)$. Finally, if $\alpha, \beta \in \mathcal{C}_{sf}$, we route the entire flow on $(\alpha, g'(\alpha))$, use the canonical paths above to route from $g'(\alpha)$ to $g'(\beta)$ and finally route the entire flow on $(g'(\beta), \beta)$. For every state $\sigma \in \mathcal{C}_p$, there are at most kn states $\sigma' \in \mathcal{C}_{sf} : g'(\sigma') = \sigma$. Therefore, we have multiplied the flow on every edge by at most

$$k^2n^2 + 2kn + 1 < 4k^2n^2, \tag{8}$$

where the first term is for pairs $\alpha, \beta \in \mathcal{C}_{sf}$, the third is for $\alpha, \beta \in \mathcal{C}_p$, and the second term on the left hand side is for mixed pairs. We added a further

$$kn\pi(\cdot)^2|\Omega| = \frac{2kn}{|\Omega|} < 1 \tag{9}$$

to each edge $(\sigma, g'(\sigma))$ and $(g'(\sigma), \sigma)$: there are at most $|\Omega|$ paths from a state $\sigma \in \mathcal{C}_p$ (to any other state), hence at most $kn|\Omega|$ paths from any $\sigma' \in \mathcal{C}_{sf}$. We absorb Inequality (9) and the fact that $\mathrm{pw}(G) + \log n = (\lambda+1)\log n$ into Inequality (8). Multiplying the bound of Lemma 3.7 by $4k^2n^2$ gives the required bound. □

4 The sampling algorithm

In order to sample a proper coloring, we need to execute the Markov chain sufficiently many times to guarantee that w.h.p. it outputs a proper coloring, and when it does, return that coloring. The pseudo code is given as Algorithm 1. For graphs of pathwidth bounded by $O(\log n)$, the algorithm runs time polynomial in n and $\log \delta$, where δ is the required bias parameter.

Algorithm 1: An almost-uniform sampler for proper colorings

Input : $G = (V, E)$ with maximal degree Δ, a number of colors $k \geq \Delta + 2$, a bias parameter $\delta > 0$

Output: a proper k-coloring of G

Set $\epsilon = \lceil \frac{k}{\Delta} \rceil$;
Set $\delta_1 = \delta/(kn+1)^2$;
Set $T = \lceil \ln(3/\delta)(kn+2)^2 \rceil$;
for $t = 1$ *to* T **do**
  Simulate $\mathcal{MC}(G, \epsilon)$ for $\tau(\delta_1)$ steps, starting from an arbitrary proper coloring;
  If the final state σ is a proper coloring, return σ;
Return an arbitrary proper coloring;

Theorem 4.1. *Algorithm 1 is a fully polynomial almost uniform sampler for proper colorings with bias parameter δ.*

Proof. We denote by $\hat{\pi}$ the distribution reached by $\mathcal{MC}$ after $\tau(\delta_1)$ steps. By definition, the total variation distance between and π and $\hat{\pi}$ is at most δ_1, hence for any $S \subset \Omega$, it holds that

$$|\pi(S) - \hat{\pi}(S)| \leq \delta_1. \tag{10}$$

Choosing $S = \mathcal{C}_p$ and applying Corollary 2.2 gives that the probability of the final state being a proper coloring is at least $\frac{1}{kn+1} - \delta_1$. Our choice of T is so that Hoeffding's bound guarantees that Algorithm 1 will output a proper coloring during the **for** loop (i.e., a final state of the Markov chain and not an arbitrary coloring), with probability at least $1 - \delta_H$, where $\delta_H \leq \frac{\delta}{3}$.

To show that Algorithm 1 is an almost uniform sampler for proper colorings, we need to show that the sampled coloring is drawn from a distribution that is close to uniform. In other words, if σ is the state output by Algorithm 1, then for any $S \subseteq \mathcal{C}_p$

$$\frac{\pi(S)}{\pi(\mathcal{C}_p)} - \delta \leq \Pr[\sigma \in S] \leq \frac{\pi(S)}{\pi(\mathcal{C}_p)} + \delta.$$

$$\begin{aligned}
\Pr[\sigma \in S] &\geq \frac{\hat{\pi}(S)}{\hat{\pi}(\mathcal{C}_p)}(1 - \delta_H) && \text{(11a)} \\
&\geq \frac{\hat{\pi}(S)}{\hat{\pi}(\mathcal{C}_p)} - \delta_H \\
&\geq \frac{\pi(S) - \delta_1}{\pi(\mathcal{C}_p) + \delta_1} - \delta_H && \text{(11b)} \\
&\geq \frac{\pi(S) - 2\delta_1}{\pi(\mathcal{C}_p)} - \delta_H && \text{(11c)} \\
&= \frac{\pi(S)}{\pi(\mathcal{C}_p)} - \frac{2\delta_1}{\pi(\mathcal{C}_p)} - \delta_H \\
&\geq \frac{\pi(S)}{\pi(\mathcal{C}_p)} - \frac{2\delta}{3} - \frac{\delta}{3},
\end{aligned}$$

where (11a) is the probability that the Algorithm outputs a proper coloring and it is in S; (11b) is due to Equation (10); (11c) is because $\frac{a-1}{b+1} \geq \frac{a+2}{b}$ for $b \geq a$.

The complementary $\Pr[\sigma \in S] \leq \frac{\pi(S)}{\pi(\mathcal{C}_p)} + \delta$ is immediate by considering the set $\mathcal{C}_p \setminus S$: if this were not the case then it would hold that $\Pr[\sigma \in S] + \Pr[\sigma \in \mathcal{C}_p \setminus S] > 1$.

□

Acknowledgements We thank Leonard Schulman for valuable discussions, and Adam Wierman, Leonard Schulman and the anonymous reviewers for comments on earlier versions of this draft.

References

[1] David J. Aldous. Random walks on finite groups and rapidly mixing markov chains. *Sminaire de probabilits de Strasbourg*, 17:243–297, 1983. 2.2, 2.2

[2] Noam Berger, Claire Kenyon, Elchanan Mossel, and Yuval Peres. Glauber dynamics on trees and hyperbolic graphs. *Probability Theory and Related Fields*, 131(3):311–340, 2005. 1, 1, 1.1, 1.2

[3] Umberto Bertelè and Francesco Brioschi. *Nonserial Dynamic Programming*. Academic Press, Inc., Orlando, FL, USA, 1972. 1.2

[4] H.L. Bodlaender, J.R. Gilbert, H. Hafsteinsson, and T. Kloks. Approximating treewidth, pathwidth, frontsize, and shortest elimination tree. *Journal of Algorithms*, 18(2):238 – 255, 1995. 1.1

[5] G.R. Brightwell and P. Winkler. Random coloring of a Cayley tree. *Contemporary combintorics*, 10:247 – 276, 2002. 1

[6] Graham Brightwell and Peter Winkler. Counting linear extensions is #P-complete. In *Proceedings of the Twenty-third Annual ACM Symposium on Theory of Computing*, STOC '91, pages 175–181, 1991. 1

[7] Andrei Z. Broder. How hard is to marry at random? (on the approximation of the permanent). In *Proceedings of the 18th Annual ACM Symposium on Theory of Computing, STOC '86*, pages 50–58, 1986. 1.2

[8] Chandra Chekuri, Sanjeev Khanna, and F. Bruce Shepherd. A note on multiflows and treewidth. *Algorithmica*, 54(3):400–412, 2009. 1.2

[9] Persi Diaconis and Daniel Stroock. Geometric bounds for eigenvalues of markov chains. *Ann. Appl. Probab.*, 1(1):36–61, 02 1991. 1.1, 1.1, 2.2

[10] Martin E. Dyer, Abraham D. Flaxman, Alan M. Frieze, and Eric Vigoda. Randomly coloring sparse random graphs with fewer colors than the maximum degree. *Random Struct. Algorithms*, 29(4):450–465, 2006. 1.1

[11] Martin E. Dyer and Alan M. Frieze. Randomly coloring graphs with lower bounds on girth and maximum degree. *Random Struct. Algorithms*, 23(2):167–179, 2003. 1, 1, 1.1

[12] Martin E. Dyer and Alan M. Frieze. Randomly coloring random graphs. *Random Struct. Algorithms*, 36(3):251–272, 2010. 1, 1.1, 1.2

[13] Martin E. Dyer, Alan M. Frieze, Thomas P. Hayes, and Eric Vigoda. Randomly coloring constant degree graphs. *Random Struct. Algorithms*, 43(2):181–200, 2013. 1, 1, 1.1, 1.2

[14] Charilaos Efthymiou. A simple algorithm for sampling colorings of g(n, d/n) up to the gibbs uniqueness threshold. *SIAM J. Comput.*, 45(6):2087–2116, 2016. 1.2

[15] J. A. Ellis, I. H. Sudborough, and J. S. Turner. The vertex separation and search number of a graph. *Inf. Comput.*, 113(1):50–79, August 1994. 5

[16] Alan Frieze and Eric Vigoda. *A survey on the use of Markov chains to randomly sample colorings*. Oxford University Press, 2007. 1

[17] Leslie Ann Goldberg, Mark Jerrum, and Marek Karpinski. The mixing time of Glauber dynamics for coloring regular trees. *Random Struct. Algorithms*, 36(4):464–476, 2010. 1, 1.2

[18] Heng Guo and Mark Jerrum. Random cluster dynamics for the ising model is rapidly mixing. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '17*, pages 1818–1827, 2017. 1.2

[19] Venkatesan Guruswami. Rapidly mixing markov chains: A comparison of techniques (A survey). *CoRR*, abs/1603.01512, 2016. 1.1

[20] Thomas P. Hayes. Randomly coloring graphs of girth at least five. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing, June 9-11, 2003, San Diego, CA, USA*, pages 269–278, 2003. 1, 1.1

[21] Thomas P. Hayes. A simple condition implying rapid mixing of single-site dynamics on spin systems. In *47th Annual IEEE Symposium on Foundations of Computer Science (FOCS*, pages 39–46, 2006. 1, 1

[22] Thomas P. Hayes, Juan Carlos Vera, and Eric Vigoda. Randomly coloring planar graphs with fewer colors than the maximum degree. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing, '07*, pages 450–458, 2007. 1, 1.1

[23] Thomas P. Hayes, Juan Carlos Vera, and Eric Vigoda. Randomly coloring planar graphs with fewer colors than the maximum degree. *Random Struct. Algorithms*, 47(4):731–759, 2015. 1.1, 1.2

[24] Thomas P. Hayes and Eric Vigoda. A non-markovian coupling for randomly sampling colorings. In *44th Symposium on Foundations of Computer Science (FOCS '03), Proceedings*, pages 618–627, 2003. 1, 1, 1.1

[25] Thomas P. Hayes and Eric Vigoda. Coupling with the stationary distribution and improved sampling for colorings and independent sets. *Ann. Appl. Probab.*, 16(3):1297–1318, 2006. 1, 1.1

[26] Mark Jerrum. A very simple algorithm for estimating the number of k-colorings of a low-degree graph. *Random Struct. Algorithms*, 7(2):157–166, 1995. 1, 1, 1.1, 1.2

[27] Mark Jerrum and Alistair Sinclair. Approximating the permanent. *SIAM J. Comput.*, 18(6):1149–1178, 1989. 1, 1.1, 1.1, 1.2

[28] Mark Jerrum, Alistair Sinclair, and Eric Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *J. ACM*, 51(4):671–697, 2004. 1.1, 1.2

[29] Nancy G. Kinnersley. The vertex separation number of a graph equals its path-width. *Information Processing Letters*, 42(6):345 – 350, 1992. 2.6

[30] V. S. Anil Kumar and H. Ramesh. Coupling vs. conductance for the jerrum-sinclair chain. *Random Struct. Algorithms*, 18(1):1–17, 2001. 1.1

[31] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times.* American Mathematical Society, 2006. 2.2

[32] Jun S. Liu. *Monte Carlo Strategies in Scientific Computing.* Springer Publishing Company, Incorporated, 2008. 1.2

[33] Pinyan Lu, Kuan Yang, Chihao Zhang, and Minshen Zhu. An FPTAS for counting proper four-colorings on cubic graphs. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2017*, pages 1798–1817, 2017. 1, 1.2

[34] Brendan Lucier and Michael Molloy. The Glauber dynamics for colorings of bounded degree trees. *SIAM J. Discrete Math.*, 25(2):827–853, 2011. 1, 1.2

[35] Fabio Martinelli, Alistair Sinclair, and Dror Weitz. Fast mixing for independent sets, colorings and other models on trees. In *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '04*, pages 456–465, 2004. 1

[36] Michael Molloy. The Glauber dynamics on colorings of a graph with high girth and maximum degree. *SIAM J. Comput.*, 33(3):721–737, 2004. 1, 1.1, 1.2

[37] Ben Morris and Alistair Sinclair. Random walks on truncated cubes and sampling 0-1 knapsack solutions. *SIAM J. Comput.*, 34(1):195–226, 2004. 1, 1.1, 1.2

[38] Elchanan Mossel and Allan Sly. Gibbs rapidly samples colorings of g(n, d/n). *Probability Theory and Related Fields*, 148(1):37–69, Sep 2010. 1.2

[39] R. B. Potts. Some generalized order-disorder transformations. *Mathematical Proceedings of the Cambridge Philosophical Society*, 48(1):106–109, 1952. 1.2

[40] Neil Robertson and P.D. Seymour. Graph minors. i. excluding a forest. *Journal of Combinatorial Theory, Series B*, 35(1):39 – 61, 1983. 1.2

[41] Neil Robertson and P.D Seymour. Graph minors. iii. planar tree-width. *Journal of Combinatorial Theory, Series B*, 36(1):49 – 64, 1984. 1.2

[42] J. Salas and A.D. Sokal. Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem. *J Stat Phys*, 86:551–579, 1997. 1

[43] Alistair Sinclair. Improved bounds for mixing rates of marcov chains and multicommodity flow. *Combinatorics, Probability & Computing*, 1:351–370, 1992. 1.1, 1.1, 2.2, 2.4

[44] Prasad Tetali, Juan Carlos Vera, Eric Vigoda, and Linji Yang. Phase transition for the mixing time of the Glauber dynamics for coloring regular trees. In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '10*, pages 1646–1656, 2010. 1, 1.2

[45] Eric Vigoda. Improved bounds for sampling colorings. In *40th Annual Symposium on Foundations of Computer Science, FOCS '99*, pages 51–59, 1999. 1, 1, 1.1, 1.2

[46] Jian-Sheng Wang, Robert H. Swendsen, and Roman Kotecký. Antiferromagnetic potts models. *Phys. Rev. Lett.*, 63:109–112, Jul 1989. 1.2