Fast Distributed Coloring Algorithms for Triangle-Free Graphs*

Seth Pettie and Hsin-Hao Su

University of Michigan

Abstract. *Vertex coloring* is a central concept in graph theory and an important symmetry-breaking primitive in distributed computing. Whereas degree-Δ graphs may require palettes of $\Delta+1$ colors in the worst case, it is well known that the chromatic number of many natural graph classes can be much smaller. In this paper we give new distributed algorithms to find (Δ/k)-coloring in graphs of girth 4 (triangle-free graphs), girth 5, and trees, where k is at most $(\frac{1}{4} - o(1)) \ln \Delta$ in triangle-free graphs and at most $(1 - o(1)) \ln \Delta$ in girth-5 graphs and trees, and $o(1)$ is a function of Δ. Specifically, for Δ sufficiently large we can find such a coloring in $O(k + \log^* n)$ time. Moreover, for *any* Δ we can compute such colorings in roughly logarithmic time for triangle-free and girth-5 graphs, and in $O(\log \Delta + \log_\Delta \log n)$ time on trees. As a byproduct, our algorithm shows that the chromatic number of triangle-free graphs is at most $(4 + o(1))\frac{\Delta}{\ln \Delta}$, which improves on Jamall's recent bound of $(67 + o(1))\frac{\Delta}{\ln \Delta}$. Also, we show that $(\Delta + 1)$-coloring for triangle-free graphs can be obtained in sublogarithmic time for any Δ.

1 Introduction

A proper t-coloring of a graph $G = (V, E)$ is an assignment from V to $\{1, \ldots, t\}$ (colors) such that no edge is monochromatic, or equivalently, each color class is an independent set. The *chromatic number* $\chi(G)$ is the minimum number of colors needed to properly color G. Let Δ be the maximum degree of the graph. It is easy to see that sometimes $\Delta + 1$ colors are necessary, e.g., on an odd cycle or a $(\Delta + 1)$-clique. Brooks' celebrated theorem [9] states that these are the *only* such examples and that every other graph can be Δ-colored. Vizing [31] asked whether Brooks' Theorem can be improved for triangle-free graphs. In the 1970s Borodin and Kostochka [8], Catlin [10], and Lawrence [21] independently proved that $\chi(G) \leq \frac{3}{4}(\Delta + 2)$ for triangle-free G, and Kostochka (see [17]) improved this bound to $\chi(G) \leq \frac{2}{3}(\Delta + 2)$.

Existential Bounds. Better asymptotic bounds were achieved in the 1990s by using an iterated approach, often called the "Rödl Nibble". The idea is to color a very small fraction of the graph in a sequence of rounds, where after each

* This work is supported by NSF CAREER grant no. CCF-0746673, NSF grant no. CCF-1217338, and a grant from the US-Israel Binational Science Foundation.

F.V. Fomin et al. (Eds.): ICALP 2013, Part II, LNCS 7966, pp. 687–699, 2013.

round some property is guaranteed to hold with some small non-zero probability. Kim [18] proved that in any girth-5 graph G, $\chi(G) \leq (1+o(1))\frac{\Delta}{\ln \Delta}$. This bound is optimal to within a factor-2 under *any* lower bound on girth. (Constructions of Kostochka and Masurova [19] and Bollobás [7] show that there is a graph G of arbitrarily large girth and $\chi(G) > \frac{\Delta}{2\ln \Delta}$.) Building on [18], Johansson (see [23]) proved that $\chi(G) = O(\frac{\Delta}{\ln \Delta})$ for any triangle-free (girth-4) graph G.[1] In relatively recent work Jamall [14] proved that the chromatic number of triangle-free graphs is at most $(67+o(1))\frac{\Delta}{\ln \Delta}$.

Algorithms. We assume the $\mathcal{LOCAL}$ model [26] of distributed computation.[2] Grable and Panconesi [12] gave a distributed algorithm that Δ/k-colors a girth-5 graph in $O(\log n)$ time, where $\Delta > \log^{1+\epsilon'} n$ and $k \leq \epsilon \ln \Delta$ for any $\epsilon' > 0$ and some $\epsilon < 1$ depending on ϵ'.[3] Jamall [15] showed a sequential algorithm for $O(\Delta/\ln \Delta)$-coloring a triangle-free graph in $O(n\Delta^2 \ln \Delta)$ time, for any $\epsilon' > 0$ and $\Delta > \log^{1+\epsilon'} n$.

Note that there are *two* gaps between the existential [14,18,23] and algorithmic results [12, 15]. The algorithmic results use a constant factor more colors than necessary (compared to the existential bounds) and they only work when $\Delta \geq \log^{1+\Omega(1)} n$ is sufficiently large, whereas the existential bounds hold for all Δ.

New Results. We give new distributed algorithms for (Δ/k)-coloring triangle-free graphs that simultaneously improve on both the existential and algorithmic results of [12, 14, 15, 23]. Our algorithms run in $\log^{1+o(1)} n$ time for *all* Δ and in $O(k+\log^* n)$ time for Δ sufficiently large. Moreover, we prove that the chromatic number of triangle-free graphs is $(4+o(1))\frac{\Delta}{\ln \Delta}$.

Theorem 1. *Fix a constant $\epsilon' > 0$. Let Δ be the maximum degree of a triangle-free graph G, assumed to be at least some $\Delta_{\epsilon'}$ depending on ϵ'. Let $k \geq 1$ be a parameter such that $2\epsilon' \leq 1 - \frac{4k}{\ln \Delta}$. Then G can be (Δ/k)-colored, in time $O(k + \log^* \Delta)$ if $\Delta^{1-\frac{4k}{\ln \Delta}-\epsilon'} = \Omega(\ln n)$, and, for any Δ, in time on the order of*

$$\min\left(e^{O(\sqrt{\ln \ln n})}, \Delta + \log^* n\right) \cdot (k + \log^* \Delta) \cdot \frac{\log n}{\Delta^{1-\frac{4k}{\ln \Delta}-\epsilon'}} \;=\; \log^{1+o(1)} n$$

The first time bound comes from an $O(k + \log^* \Delta)$-round procedure, each round of which succeeds with probability $1 - 1/\operatorname{poly}(n)$. However, as Δ decreases the probability of failure tends to 1. To enforce that each step succeeds with high

[1] We are not aware of any extant copy of Johansson's manuscript. It is often cited as a DIMACS Technical Report, though no such report exists. Molloy and Reed [23] reproduced a variant of Johansson's proof showing that $\chi(G) \leq 160\frac{\Delta}{\ln \Delta}$ for triangle-free G.

[2] In short, vertices host processors which operate is synchronized rounds; vertices can communicate one arbitrarily large message across each edge in each round; local computation is free; *time* is measured by the number of rounds.

[3] They claimed that their algorithm could also be extended to triangle-free graphs. Jamall [15] pointed out a flaw in their argument.

probability we use a version of the Local Lemma algorithm of Moser and Tardos [24] optimized for the parameters of our problem.[4]

By choosing $k = \ln \Delta/(4+\epsilon)$ and $\epsilon' = \epsilon/(2(4+\epsilon))$, we obtain new bounds on the chromatic number of triangle-free graphs.

Corollary 1. *For any $\epsilon > 0$ and Δ sufficiently large (as a function of ϵ), $\chi(G) \le (4+\epsilon)\frac{\Delta}{\ln \Delta}$. Consequently, the chromatic number of triangle-free graphs is $(4+o(1))\frac{\Delta}{\ln \Delta}$, where the $o(1)$ is a function of Δ.*

Our result also extends to girth-5 graphs with $\Delta^{1-\frac{4k}{\ln \Delta}-\epsilon'}$ replaced by $\Delta^{1-\frac{k}{\ln \Delta}-\epsilon'}$, which allows us to $(1+\epsilon)\Delta/\ln \Delta$-color such graphs. Our algorithm can clearly be applied to trees (girth ∞). Elkin [11] noted that with Bollobás's construction [7], Linial's lower bound [22] on coloring trees can be strengthened to show that it is impossible to $o(\Delta/\ln \Delta)$-color a tree in $o(\log_\Delta n)$ time. We prove that it *is* possible to $(1+o(1))\Delta/\ln \Delta$-color a tree in $O(\log \Delta + \log_\Delta \log n)$ time. Also, we show that $(\Delta+1)$-coloring for triangle-free graphs can be obtained in $\exp(O(\sqrt{\log \log n}))$ time.

Technical Overview. In the iterated approaches of [12, 14, 18, 23] each vertex u maintains a *palette*, which consists of the colors that have not been selected by its neighbors. To obtain a t-coloring, each palette consists of colors $\{1, \ldots, t\}$ initially. In each round, each u tries to assign itself a color (or colors) from its palette, using randomization to resolve the conflicts between itself and the neighbors. The *c-degree* of u is defined to be the number of its neighbors whose palettes contain c. In Kim's algorithm [18] for girth-5 graphs, the properties maintained for each round are that the c-degrees are upper bounded and the palette sizes are lower bounded. In girth-5 graphs the neighborhoods of the neighbors of u only intersect at u and therefore have a negligible influence on each other, that is, whether c remains in one neighbor's palette has little influence on a different neighbor of u. Due to this independence one can bound the c-degree after an iteration using standard concentration inequalities. In triangle-free graphs, however, there is no guarantee of independence. If two neighbors of u have identical neighborhoods, then after one iteration they will either both keep or both lose c from their palettes. In other words, the c-degree of u is a random variable that may not have any significant concentration around its mean. Rather than bound c-degrees, Johansson [23] bounded the entropy of the remaining palettes so that each color is picked nearly uniformly in each round. Jamall [14] claimed that although each c-degree does not concentrate, the *average* c-degree (over each c in the palette) does concentrate. Moreover, it suffices to consider only those colors within a constant factor of the average in subsequent iterations.

Our (Δ/k)-coloring algorithm performs the same coloring procedure in each round, though the behavior of the algorithm has two qualitatively distinct phases.

[4] Note that for many reasonable parameters (e.g., $k = O(1), \Delta = \log^{1-\delta} n$), the running time is *sub*logarithmic.

In the first $O(k)$ rounds the c-degrees, palette sizes, and probability of remaining uncolored are very well behaved. Once the available palette is close to the number of uncolored neighbors the probability of remaining uncolored begins to decrease drastically in each successive round, and after $O(\log^* n)$ rounds all vertices are colored, w.h.p.

Our analysis is similar to that of Jamall [14] in that we focus on bounding the average of the c-degrees. However, our proof needs to take a different approach, for two reasons. First, to obtain an efficient *distributed* algorithm we need to obtain a tighter bound on the probability of failure in the last $O(\log^* n)$ rounds, where the c-degrees shrink faster than a constant factor per round. Second, there is a small flaw in Jamall's application of Azuma's inequality in Lemma 12 in [14], the corresponding Lemma 17 in [15], and the corresponding lemmas in [16]. It is probably possible to correct the flaw, though we manage to circumvent this difficulty altogether. See the full version for a discussion of this issue.

The second phase presents different challenges. The natural way to bound c-degrees using Chernoff-type inequalities gives error probabilities that are exponential *in the c-degree*, which is fine if it is $\Omega(\log n)$ but becomes too large as the c-degrees are reduced in each coloring round. At a certain threshold we switch to a different analysis (along the lines of Schneider and Wattenhofer [30]) that allows us to bound c-degrees with high probability in the *palette* size, which, again, is fine if it is $\Omega(\log n)$.

In both phases, if we cannot obtain small error probabilities (via concentration inequalities and a union bound) we revert to a distributed implementation of the Moser-Tardos Lovász Local Lemma algorithm [24]. We show that for certain parameters the symmetric LLL can be made to run in *sub*logarithmic time. For the extensions to trees and the $(\Delta + 1)$-coloring algorithm for triangle-free graphs, we adopt the ideas from [5,6,29] to reduce the graph into several smaller components and color each of them separately by deterministic algorithms [4,25], which will run faster as the size of each subproblem is smaller.

Organization. Section 2 presents the general framework for the analysis. Section 3 describes the algorithms and discusses what parameters to plug into the framework. Section 4 describes the extension to graphs of girth 5, trees, and the $(\Delta + 1)$-coloring algorithm for triangle-free graphs.

2 The Framework

Every vertex maintains a *palette* that consists of all colors not previously chosen by its neighbors. The coloring is performed in rounds, where each vertex chooses zero or more colors in each round. Let G_i be the graph induced by the uncolored vertices after round i, so $G = G_0$. Let $N_i(u)$ be u's neighbors in G_i and let $P_i(u)$ be its palette after round i. The *c-neighbors* $N_{i,c}(u)$ consist of those $v \in N_i(u)$ with $c \in P_i(v)$. Call $|N_i(u)|$ the *degree* of u and $|N_{i,c}(u)|$ the *c-degree* of u after round i. This notation is extended to sets of vertices in a natural way, e.g., $N_i(N_i(u))$ is the set of neighbors of neighbors of u in G_i.

Algorithm 2 describes the iterative coloring procedure. In each round, each vertex u selects a set $S_i(u)$ of colors by including each $c \in P_{i-1}(u)$ independently with probability π_i to be determined later. If some $c \in S_i(u)$ is not selected by any neighbor of u then u can safely color itself c. In order to remove dependencies between various random variables we exclude colors from u's palette more aggressively than is necessary. First, we exclude any color *selected* by a neighbor, that is, $S_i(N_{i-1}(u))$ does not appear in $P_i(u)$. The probability that a color c is *not* selected by a neighbor is $(1-\pi_i)^{|N_{i-1,c}(u)|}$. Suppose that this quantity is at least some threshold β_i for all c. We force c to be kept with probability *precisely* β_i by putting c in a keep-set $K_i(u)$ with probability $\beta_i/(1-\pi_i)^{|N_{i-1,c}(u)|}$. The probability that $c \in K_i(u) \backslash S_i(N_{i-1}(u))$ is therefore β_i, assuming $\beta_i/(1-\pi_i)^{|N_{i-1,c}(u)|}$ is a valid probability; if it is not then c is *ignored*. Let $\widehat{P}_i(u)$ be what remains of u's palette. Algorithm 2 has two variants. In Variant B, $P_i(u)$ is exactly $\widehat{P}_i(u)$ whereas in Variant A $P_i(u)$ is the subset of $\widehat{P}_i(u)$ whose c-degrees are sufficiently low, less than $2t_i$, where t_i is a parameter that will be explained below.

Include each $c \in P_{i-1}(u)$ in $S_i(u)$ independently with probability π_i.
For each c, calculate $r_c = \beta_i/(1-\pi_i)^{|N_{i-1,c}(u)|}$.
If $r_c \leq 1$, include $c \in P_{i-1}(u)$ in $K_i(u)$ independently with probability r_c.
return $(S_i(u), K_i(u))$.

Algorithm 1. Select(u, π_i, β_i)

repeat
  Round $i = 1, 2, 3, \ldots$.
  for each $u \in G_{i-1}$ **do**
    $(S_i(u), K_i(u)) \leftarrow$ Select(u, π_i, β_i)
    Set $\widehat{P}_i(u) \leftarrow K_i(u) \setminus S_i(N_{i-1}(u))$
    if $S_i(u) \cap \widehat{P}_i(u) \neq \emptyset$ **then** color u with any color in $S_i(u) \cap \widehat{P}_i(u)$ **end if**
    (Variant A) $P_i(u) \leftarrow \{c \in \widehat{P}_i(u) \mid |N_{i,c}(u)| \leq 2t_i\}$
    (Variant B) $P_i(u) \leftarrow \widehat{P}_i(u)$
  end for
  $G_i \leftarrow G_{i-1} \setminus \{$colored vertices$\}$
until the termination condition occurs

Algorithm 2. Coloring-Algorithm$(G_0, \{\pi_i\}, \{\beta_i\})$

The algorithm is parameterized by the sampling probabilities $\{\pi_i\}$, the ideal c-degrees $\{t_i\}$ and the ideal probability $\{\beta_i\}$ of retaining a color. The $\{\beta_i\}$ define how the ideal palette sizes $\{p_i\}$ degrade. Of course, the *actual* palette sizes and c-degrees after i rounds will drift from their ideal values, so we will need to reason about approximations of these quantities. We will specify the initial parameters and the terminating conditions when applying both variants in Section 3.

2.1 Analysis A

Given $\{\pi_i\}$, p_0, t_0, and δ, the parameters for Variant A are derived below.

$$\begin{aligned} \beta_i &= (1-\pi_i)^{2t_{i-1}} & \alpha_i &= (1-\pi_i)^{(1-(1+\delta)^{i-1}/2)p'_i} \\ p_i &= \beta_i p_{i-1} & t_i &= \max(\alpha_i\beta_i t_{i-1}, T) \\ p'_i &= (1-\delta/8)^i p_i & t'_i &= (1+\delta)^i t_i \end{aligned} \tag{1}$$

Let us take a brief tour of the parameters. The sampling probability π_i will be inversely proportional to t_{i-1}, the ideal c-degree at end of round $i-1$. (The exact expression for π_i depends on ϵ'.) Since we filter out colors with more than twice the ideal c-degree, the probability that a color is not selected by any neighbor is at least $(1-\pi_i)^{2t_{i-1}} = \beta_i$. Note that since $\pi_i = \Theta(1/t_{i-1})$ we have $\beta_i = \Theta(1)$. Thus, we can force all colors to be retained in the palette with probability precisely β_i, making the ideal palette size $p_i = \beta_i p_{i-1}$. Remember that a c-neighbor stays a c-neighbor if it remains uncolored *and* it does not remove c from its palette. The latter event happens with probability β_i. We use α_i as an upper bound on the probability that a vertex remains uncolored, so the ideal c-degree should be $t_i = \alpha_i\beta_i t_{i-1}$. To account for deviations from the ideal we let p'_i and t'_i be approximate versions of p_i and t_i, defined in terms of a small error control parameter $\delta > 0$. Furthermore, certain high probability bounds will fail to hold if t_i becomes too small, so we will not let it go below a threshold T.

When the graph has girth 5, the concentration bounds allow us to show that $|P_i(u)| \geq p'_i$ and $|N_{i,c}(u)| \leq t'_i$ with certain probabilities. As pointed out by Jamall [14,15], $|N_{i,c}(u)|$ does not concentrate in triangle-free graphs. He showed that the average c-degree, $\overline{n}_i(u) = \sum_{c\in P_i(u)} |N_{i,c}(u)|/|P_i(u)|$, concentrates and will be bounded above by t'_i with a certain probability. Since $\overline{n}_i(u)$ concentrates, it is possible to bound the fraction of colors filtered for having c-degrees larger than $2t_i$.

Let $\lambda_i(u) = \min(1, |P_i(u)|/p'_i)$. Since $P_i(u)$ is supposed to be at least p'_i, if we do not filter out colors, $1-\lambda_i(u)$ can be viewed as the fraction that has been filtered. In the following we state an induction hypotheses equivalent to Jamall's [14].

$$D_i(u) \leq t'_i, \text{ where } D_i(u) = \lambda_i(u)\overline{n}_i(u) + (1-\lambda_i(u))2t_i$$

$D_i(u)$ can be interpreted as the average of the c-degrees of $P_i(u)$ with $p'_i - |P_i(u)|$ dummy colors whose c-degrees are exactly $2t_i$. Notice that $D_i(u) \leq t'_i$ also implies $1-\lambda_i(u) \leq (1+\delta)^i/2$, because $(1-\lambda_i(u))2t_i \leq D_i(u) \leq t'_i$. Therefore:

$$|P_i(u)| \geq (1-(1+\delta)^i/2)p'_i$$

Recall $P_i(u)$ is the palette consisting of colors c for which $|N_{i,c}(u)| \leq 2t_i$.

The main theorem for this section shows the inductive hypothesis holds with a certain probability. See the full version for the proof.

Theorem 2. *Suppose that* $D_{i-1}(x) \leq t'_{i-1}$ *for all* $x \in G_{i-1}$*, then for a given* $u \in G_{i-1}$*,* $D_i(u) \leq t'_i$ *holds with probability at least* $1-\Delta e^{-\Omega(\delta^2 T)} - (\Delta^2+2)e^{-\Omega(\delta^2 p'_i)}$.

2.2 Analysis B

Analysis A has a limitation for smaller c-degrees, since the probability guarantee becomes smaller as t_i goes down. Therefore, Analysis A only works well for $t_i \geq T$, where T is a threshold for certain probability guarantees. For example, if we want Theorem 2 to hold with high probability in n, then we must have $T \gg \log n$.

To get a good probability guarantee below T, we will use an idea by Schneider and Wattenhofer [30]. They took advantage of the trials done for each color inside the palette, rather than just considering the trials on whether each neighbor is colored or not. We demonstrate this idea in the proof of Theorem 3 in the full version. The probability guarantee in the analysis will not depend on the current c-degree but on the *initial* c-degree and the current palette size.

The parameters for Variant B are chosen based on an initial lower bound on the palette size p_0, upper bound on the c-degree t_0, and error control parameter δ. The selection probability is chosen to be $\pi_i = 1/(t_{i-1}+1)$ and the probability a color remains in a palette $\beta_i = (1-\pi_i)^{t_{i-1}}$. The ideal palette size and its relaxation are $p_i = \beta_i p_{i-1}$ and $p'_i = (1-\delta)^i p_i$, and the ideal c-degree $t_i = \max(\alpha_i t_{i-1}, 1)$. One can show the probability of remaining uncolored is upper bounded by $\alpha_i = 5t_0/p'_i$,

Let $E_i(u)$ denote the event that $|P_i(u)| \geq p'_i$ and $|N_{i,c}(u)| < t_i$ for all $c \in P_i(u)$. Although a vertex could lose its c-neighbor if the c-neighbor becomes colored or loses c in its palette, in this analysis, we only use the former to bound its c-degree. Also, if $E_{i-1}(u)$ is true, then $\Pr(c \notin S_i(N_{i-1}(u))) > \beta_i$ for all $c \in P_{i-1}(u)$. Thus in Select(u, π_i, β_i), we will not ignore any colors in the palette. Each color remains in the palette with probability exactly β_i.

The following theorem shows the inductive hypothesis holds with a certain probability. See the full version for the proof.

Theorem 3. *If $E_{i-1}(x)$ holds for all $x \in G_{i-1}$, then for a given $u \in G_{i-1}$, $E_i(u)$ holds with probability at least* $1 - \Delta e^{-\Omega(t_0)} - (\Delta^2+1)e^{-\Omega(\delta^2 p'_i)}$

3 The Coloring Algorithms

The algorithm in Theorem 1 consists of two phases. Phase I uses Analysis A and Phase II uses Analysis B. First, we will give the parameters for both phases. Then, we will present the distributed algorithm that makes the induction hypothesis in Theorem 2 ($D_i(u) \leq t'_i$) and Theorem 3 ($E_i(u)$) hold for all $u \in G_i$ with high probability in n for every round i.

Let $\epsilon_1 = 1 - \frac{4k}{\ln \Delta} - \frac{2\epsilon'}{3}$ and $\epsilon_2 = 1 - \frac{4k}{\ln \Delta} - \frac{\epsilon'}{3}$. We will show that upon reaching the terminating condition of Phase I (which will be defined later), we will have $|P_i(u)| \geq \Delta^{\epsilon_2}$ for all $u \in G_i$ and $|N_{i,c}(u)| < \Delta^{\epsilon_1}$ for all $u \in G_i$ and all $c \in P_i(u)$. At this point, for a non-constructive version, we can simply apply the results about list coloring constants [13, 27, 28] to get a proper coloring, since at this point there is an $\omega(1)$ gap between $|N_{i,c}(u)|$ and $|P_i(u)|$ for every $u \in G_i$. One can turn the result of [27] into a distributed algorithm with the aid of Moser-Tardos

Lovász Local Lemma algorithm to amplify the success probability. However, to obtain an efficient distributed algorithm we use Analysis B in Phase II.

Since our result holds for large enough Δ, we can assume whenever necessary that Δ is sufficiently large. The asymptotic notation will be with respect to Δ.

3.1 Parameters for Phase I

In this phase, we use Analysis A with the following parameters and terminating condition: $\pi_i = \frac{1}{2Kt_{i-1}+1}$, where $K = 4/\epsilon'$ is a constant, $p_0 = \Delta/k$, $t_0 = \Delta$ and $\delta = 1/\log^2 \Delta$. This phase ends after the round when $t_i \le T \overset{\text{def}}{=} \Delta^{\epsilon_1}/3$.

First, we consider the algorithm for at most the first $O(\log \Delta)$ rounds. For these rounds, we can assume the error $(1+\delta)^i \le \left(1 + \frac{1}{\log^2 \Delta}\right)^{O(\log \Delta)} \le e^{O(1/\log \Delta)} = 1 + o(1)$ and similarly $(1-\delta/8)^i \ge \left(1 - \frac{1}{\log^2 \Delta + 1}\right)^{O(\log \Delta)} \ge e^{-O(1/\log \Delta)} = 1 - o(1)$. We will show the algorithm reaches the terminating condition during these rounds, where the error is under control.

The probability a color is retained, $\beta_i = (1-\pi_i)^{2t_{i-1}} \ge e^{-1/K}$, is bounded below by a constant. The probability a vertex remains uncolored is at most $\alpha_i = (1-\pi_i)^{(1-(1+\delta)^{i-1}/2)p'_i} \le e^{-(1-o(1))Cp_{i-1}/t_{i-1}}$, where $C = 1/(4Ke^{1/K})$.

Let $s_i = t_i/p_i$ be the ratio between the ideal c-degree and the ideal palette size. Initially, $s_0 = k$ and $s_i = \alpha_i s_{i-1} \le s_{i-1} e^{-(1-o(1))(C/s_{i-1})}$. Initially, s_i decreases roughly linearly by C for each round until the ratio $s_i \approx C$ is a constant. Then, s_i decreases rapidly in the order of iterated exponentiation. Therefore, it takes roughly $O(k + \log^* \Delta)$ rounds to reach the terminating condition where $t_i \le T$. Our goal is to show upon reaching the terminating condition, the palette size bound p_i is greater than T by some amount, in particular, $p_i \ge 30e^{3/\epsilon'}\Delta^{\epsilon_2}$. See the full version for the proof of the following Lemma.

Lemma 1. *Phase I terminates in* $(4+o(1))Ke^{1/K}k + O(\log^* \Delta)$ *rounds, where* $K = 4/\epsilon'$. *Moreover,* $p_i \ge 30e^{3/\epsilon'}\Delta^{\epsilon_2}$ *for every round* i *in this phase.*

Thus, if the induction hypothesis $D_i(u) \le t'_i$ holds for every $u \in G_i$ for every round i during this phase, we will have $|P_i(u)| \ge (1-(1+\delta)^i/2)p'_i \ge 10e^{3/\epsilon'}\Delta^{\epsilon_2}$ for all $u \in G_i$ and $|N_{i,c}(u)| \le 2t_i < \Delta^{\epsilon_1}$ for all $u \in G_i$ and all $c \in P_i(u)$ in the end.

3.2 Parameters for Phase II

In Phase II, we will use Analysis B with the following parameters and terminating condition: $p_0 = 10e^{3/\epsilon'}\Delta^{\epsilon_2}$, $t_0 = \Delta^{\epsilon_1}$ and $\delta = 1/\log^2 \Delta$. This phase terminates after $\frac{3}{\epsilon'}$ rounds.

First note that the number of rounds $\frac{3}{\epsilon'}$ is a constant. We show $p'_i \ge 5\Delta^{\epsilon_2}$ for each round $1 \le i \le \frac{3}{\epsilon'}$, so there is always a sufficient large gap between the current palette size and the initial c-degree, which implies the shrinking factor of the c-degrees is $\alpha_i = 5t_0/p'_i \le \Delta^{-\epsilon'/3}$. Since p_i shrinks by at most a $\beta_i \ge e^{-1}$ factor every round, $p'_i \ge (1-\delta)^i \prod_{j=1}^{i} \beta_j p_0 \ge ((1-\delta)e^{-1})^i 10e^{3/\epsilon'}\Delta^{\epsilon_2} \ge 5\Delta^{\epsilon_2}$.

Now since $\alpha_i \leq \Delta^{-\epsilon'/3}$, after $\frac{3}{\epsilon'}$ rounds, $t_i \leq t_0 \prod_{j=1}^{i} \alpha_j \leq \Delta \left(\Delta^{-\epsilon'/3}\right)^{\frac{3}{\epsilon'}} \leq 1$. The c-degree bound, $t_{\epsilon'/3}$, becomes 1. Recall that the induction hypothesis $E_i(u)$ is the event that $|P_i(u)| \geq p'_i$ and $|N_{i,c}(u)| < t_i$. If $E_i(u)$ holds for every $u \in G_i$ for every round i during this phase, then in the end, every uncolored vertex has no c-neighbors, as implied by $|N_{i,c}(u)| < t_i \leq 1$. This means these vertices can be colored with anything in their palettes, which are non-empty.

3.3 The Distributed Coloring Algorithm

We will show a distributed algorithm that makes the induction hypothesis in Phase I and Phase II hold with high probability in n.

Fix the round i and assume the induction hypothesis holds for all $x \in G_{i-1}$. For $u \in G_{i-1}$, define $A(u)$ to be the bad event that the induction hypothesis fails at u (i.e. $D_i(u) > t'_i$ in Phase I or $E_i(u)$ fails in Phase II). Let $p = e^{-\Delta^{1-\frac{4k}{\ln \Delta}-\epsilon'}}/(e\Delta^4)$. By Theorem 2 and Theorem 3, $\Pr(A(u))$ is at most:

$$\Delta e^{-\Omega\left(\delta^2 T\right)} + (\Delta^2+2)e^{-\Omega\left(\delta^2 p'_i\right)} \quad \text{or} \quad \Delta e^{-\Omega(t_0)} + (\Delta^2+1)e^{-\Omega\left(\delta^2 p'_i\right)}$$

Since $T = \Delta^{\epsilon_1}/3, t_0 = \Delta^{\epsilon_1}, p'_i \geq \Delta^{\epsilon_2}$, $\Pr(A(u)) \leq p$ for large enough Δ.

If $\Delta^{1-\frac{4k}{\ln \Delta}-\epsilon'} > c\log n$, then $p < 1/n^c$. By the union bound over all $u \in G_{i-1}$, the probability that any of the $A(u)$ fails is at most $1/n^{c-1}$. The induction hypothesis holds for all $u \in G_i \subseteq G_{i-1}$ with high probability. In this case, $O(k + \log^* \Delta)$ rounds suffice, because each round succeeds with high probability.

On the other hand, if $\Delta^{1-\frac{4k}{\ln \Delta}-\epsilon'} < c\log n$, then we apply Moser and Tardos' resampling algorithm to make $A(u)$ simultaneously hold for all u with high probability. At round i, the bad event $A(u)$ depends on the random variables which are generated by $\text{Select}(v, \pi_i, \beta_i)$ for v within distance 2 in G_{i-1}. Therefore, the dependency graph $G_{i-1}^{\leq 4}$ consists of edges (u,v) such that $dist_{G_{i-1}}(u,v) \leq 4$. Each event $A(u)$ shares variables with at most $d < \Delta^4$ other events. The Lovász Local Lemma [1] implies that if $ep(d+1) \leq 1$, then the probability that all $A(u)$ simultaneously hold is guaranteed to be non-zero. Moser and Tardos showed how to boost this probability by resampling. In each round of resampling, their algorithm finds an MIS I in the dependency graph induced by the set of bad events B and then resamples the random variables that I depends on. In our case, it corresponds to finding an MIS I in $G_{i-1}^{\leq 4}[B]$, where $B = \{u \in G_{i-1} \mid A(u) \text{ fails}\}$. Then, we redo $\text{Select}(v, \pi_i, \beta_i)$ for $v \in G$ within distance 2 from I to resample the random variables that I depends on. By plugging in the parameters for the symmetric case, their proof shows if $ep(d+1) \leq 1-\epsilon$, then the probability any of the bad events occur after t rounds of resampling is at most $(1-\epsilon)^t n/d$. Thus, $O(\log n / \log(\frac{1}{1-\epsilon}))$ rounds will be sufficient for all $A(u)$ to hold with high probability in n.[5]

[5] In the statement of Theorem 1.3 in [24], they used $1/\epsilon$ as an approximation for $\log(\frac{1}{1-\epsilon})$. However, this difference can be significant in our case, when $1-\epsilon$ is very small.

As shown in previous sections, $p \le e^{-\Delta^{1-\frac{4k}{\ln \Delta}-\epsilon'}}/(e\Delta^4)$. We can let $1-\epsilon = ep(d+1) \le e^{-\Delta^{1-\frac{4k}{\ln \Delta}-\epsilon'}}$. Therefore, $O(\log n/\Delta^{1-\frac{4k}{\ln \Delta}-\epsilon'})$ resampling rounds will be sufficient. Also, an MIS can be found in $O(\Delta + \log^* n)$ time [3, 20], or $e^{O(\sqrt{\log\log n})}$ since $\Delta \le (c\log n)^{1/(1-\frac{4k}{\ln \Delta}-\epsilon')} \le (c\log n)^{1/\epsilon'} \le \log^{O(1)} n$ [5]. Each of the $O(k+\log^* \Delta)$ rounds is delayed by $O(\log n/\Delta^{1-\frac{4k}{\ln \Delta}-\epsilon'})$ resampling rounds, which are futher delayed by the rounds needed to find an MIS. Therefore, the total number of rounds is

$$O\left((k+\log^* \Delta)\cdot \frac{\log n}{\Delta^{1-\frac{4k}{\ln \Delta}-\epsilon'}} \cdot \min\left(\exp\left(O\left(\sqrt{\log\log n}\right)\right), \Delta + \log^* n\right)\right)$$

Note that this is always at most $\log^{1+o(1)} n$.

4 Extensions

4.1 Graphs of Girth at Least 5

For graphs of girth at least 5, existential results [18, 23] show that there exists $(1+o(1))\Delta/\ln \Delta$-coloring. We state a matching algorithmic result. The proof will be included in the full version.

Theorem 4. *Fix a constant $\epsilon' > 0$. Let Δ be the maximum degree of a girth-5 graph G, assumed to be at least some $\Delta_{\epsilon'}$ depending on ϵ'. Let $k \ge 1$ be a parameter such that $2\epsilon' \le 1 - \frac{k}{\ln \Delta}$. Then G can be (Δ/k)-colored, in time $O(k+\log^* \Delta)$ if $\Delta^{1-\frac{k}{\ln \Delta}-\epsilon'} = \Omega(\ln n)$, and, for any Δ, in time on the order of*

$$\min\left(e^{O(\sqrt{\ln\ln n})}, \Delta + \log^* n\right)\cdot (k+\log^* \Delta)\cdot \frac{\log n}{\Delta^{1-\frac{k}{\ln \Delta}-\epsilon'}} \;=\; \log^{1+o(1)} n$$

4.2 Trees

Trees are graphs of infinity girth. According to Theorem 4, it is possible to get a (Δ/k)-coloring in $O(k+\log^* \Delta)$ time if $\Delta^{1-\frac{k}{\ln \Delta}-\epsilon'} = \Omega(\log n)$. If $\Delta^{1-\frac{k}{\ln \Delta}-\epsilon'} = O(\log n)$, we will show that using additional $O(q)$ colors, it is possible to get a $(\Delta/k + O(q))$-coloring in $O\left(k + \log^* n + \frac{\log\log n}{\log q}\right)$ time. By choosing $q = \sqrt{\Delta}$, we can find a $(1+o(1))\Delta/\ln \Delta$-coloring in $O(\log \Delta + \log_\Delta \log n)$ rounds.

The algorithm is the same with the framework, except that at the end of each round we delete the *bad* vertices, which are the vertices that fail to satisfy the induction hypothesis. The remaining vertices must satisfy the induction hypothesis, and then we will continue the next round on these vertices. Using the idea from [5, 6, 29], we can show that after $O(k+\log^* \Delta)$ rounds of the algorithm, the size of each component formed by the bad vertices is at most $O\left(\Delta^4 \log n\right)$ with high probability. See the full version for the proof.

Barenboim and Elkin's deterministic algorithm [4] obtains $O(q)$-coloring in $O\left(\frac{\log n}{\log q} + \log^* n\right)$ time for trees (arboricity = 1). We then apply their algorithm on each component formed by bad vertices. Since the size of each component is at most $O(\Delta^4 \log n)$, their algorithm will run in $O\left(\frac{\log\log n+\log \Delta}{\log q} + \log^* n\right)$ time, using the additional $O(q)$ colors. Note that this running time is actually $O\left(\frac{\log\log n}{\log q} + \log^* n\right)$, since $\Delta = \log^{O(1)} n$.

4.3 $(\Delta + 1)$-Color Triangle-Free Graphs in Sublogarithmic Time

We show that $(\Delta + 1)$-coloring in triangle-free graphs can be obtained in $\exp(O(\sqrt{\log\log n}))$ rounds for any Δ. Let $k = 1$ and $\epsilon' = 1/4$. By Theorem 1, there exists a constant Δ_0 such that for all $\Delta \geq \Delta_0$, if $\Delta^{1/2} \geq \log n$, then a $(\Delta+1)$-coloring can be found in $O(\log^* \Delta)$ time. If $\Delta < \Delta_0$, then $(\Delta+1)$ can be solved in $O(\Delta+\log^* n) = O(\log^* n)$ rounds [3,20]. Otherwise, if $\Delta_0 \leq \Delta < \log^2 n$, then we can apply the same technique for trees to bound the size of each bad component by $O(\Delta^4 \log n) = \text{polylog}(n)$, whose vertices failed to satisfy the induction hypothesis in the $O(\log^* \Delta)$ rounds. Panconesi and Srinivasan's deterministic network decomposition algorithm [25] obtains $(\Delta + 1)$-coloring in $\exp(O(\sqrt{\log n}))$ for graphs with n vertices. In fact, their decomposition can also obtain a proper coloring as long as the graph can be greedily colored (e.g. the palette size is more than the degree for each vertex). Therefore, by applying their algorithm, each bad component can be properly colored in $\exp(O(\sqrt{\log\log n}))$ rounds.

5 Conclusion

The time bounds of Theorem 1 show an interesting discontinuity. When Δ is large we can cap the error at $1/\text{poly}(n)$ by using standard concentration inequalities and a union bound. When Δ is small we can use the Moser-Tardos LLL algorithm to reduce the failure probability again to $1/\text{poly}(n)$. Thus, the distributed complexity of our coloring algorithm is tied to the distributed complexity of the constructive Lovász Local Lemma.

We showed that $\chi(G) \leq (4+o(1))\Delta/\ln\Delta$ for triangle-free graphs G. It would be interesting to see if it is possible to reduce the palette size to $(1+o(1))\Delta/\ln\Delta$, matching Kim's [18] bound for girth-5 graphs.

Alon et al. [2] and Vu [32] extended Johansson's result [23] for triangle-free graphs to obtain an $O(\Delta/\log f)$-coloring for locally sparse graphs (the latter also works for list coloring), in which no neighborhood of any vertex spans more than Δ^2/f edges. It would be interesting to extend our result to locally sparse graphs and other sparse graph classes.

References

1. Alon, N., Spencer, J.H.: The Probabilistic Method. Wiley Series in Discrete Mathematics and Optimization. Wiley (2011)
2. Alon, N., Krivelevich, M., Sudakov, B.: Coloring graphs with sparse neighborhoods. Journal of Combinatorial Theory, Series B 77(1), 73–82 (1999)

3. Barenboim, L., Elkin, M.: Distributed ($\Delta + 1$)-coloring in linear (in Δ) time. In: STOC 2009, pp. 111–120. ACM, New York (2009)
4. Barenboim, L., Elkin, M.: Sublogarithmic distributed MIS algorithm for sparse graphs using Nash-Williams decomposition. Distrib. Comput. 22, 363–379 (2010)
5. Barenboim, L., Elkin, M., Pettie, S., Schneider, J.: The locality of distributed symmetry breaking. In: FOCS 2012, pp. 321–330 (October 2012)
6. Beck, J.: An algorithmic approach to the lovász local lemma. Random Structures & Algorithms 2(4), 343–365 (1991)
7. Bollobás, B.: Chromatic number, girth and maximal degree. Discrete Mathematics 24(3), 311–314 (1978)
8. Borodin, O.V., Kostochka, A.V.: On an upper bound of a graph's chromatic number, depending on the graph's degree and density. Journal of Combinatorial Theory, Series B 23(2-3), 247–250 (1977)
9. Brooks, R.L.: On colouring the nodes of a network. Mathematical Proceedings of the Cambridge Philosophical Society 37(02), 194–197 (1941)
10. Catlin, P.A.: A bound on the chromatic number of a graph. Discrete Math. 22(1), 81–83 (1978)
11. Elkin, M.: Personal communication
12. Grable, D.A., Panconesi, A.: Fast distributed algorithms for Brooks-Vizing colorings. Journal of Algorithms 37(1), 85–120 (2000)
13. Haxell, P.E.: A note on vertex list colouring. Comb. Probab. Comput. 10(4), 345–347 (2001)
14. Jamall, M.S.: A Brooks' Theorem for Triangle-Free Graphs. ArXiv e-prints (2011)
15. Jamall, M.S.: A Coloring Algorithm for Triangle-Free Graphs. ArXiv e-prints (2011)
16. Jamall, M.S.: Coloring Triangle-Free Graphs and Network Games. Dissertation. University of California, San Diego (2011)
17. Jensen, T.R., Toft, B.: Graph coloring problems. Wiley-Interscience series in discrete mathematics and optimization. Wiley (1995)
18. Kim, J.H.: On brooks' theorem for sparse graphs. Combinatorics. Probability and Computing 4, 97–132 (1995)
19. Kostochka, A.V., Mazuronva, N.P.: An inequality in the theory of graph coloring. Metody Diskret. Analiz. 30, 23–29 (1977)
20. Kuhn, F.: Weak graph colorings: distributed algorithms and applications. In: SPAA 2009, pp. 138–144. ACM, New York (2009)
21. Lawrence, J.: Covering the vertex set of a graph with subgraphs of smaller degree. Discrete Mathematics 21(1), 61–68 (1978)
22. Linial, N.: Locality in distributed graph algorithms. SIAM J. Comput. 21(1), 193–201 (1992)
23. Molloy, M., Reed, B.: Graph Colouring and the Probabilistic Method. Algorithms and Combinatorics. Springer (2001)
24. Moser, R.A., Tardos, G.: A constructive proof of the general lovász local lemma. J. ACM 57(2), 11:1–11:15 (2010)
25. Panconesi, A., Srinivasan, A.: On the complexity of distributed network decomposition. Journal of Algorithms 20(2), 356–374 (1996)
26. Peleg, D.: Distributed Computing: A Locality-Sensitive Approach. Monographs on Discrete Mathematics and Applications. SIAM (2000)
27. Reed, B.: The list colouring constants. Journal of Graph Theory 31(2), 149–153 (1999)

28. Reed, B., Sudakov, B.: Asymptotically the list colouring constants are 1. J. Comb. Theory Ser. B 86(1), 27–37 (2002)
29. Rubinfeld, R., Tamir, G., Vardi, S., Xie, N.: Fast local computation algorithms. In: ICS 2011, pp. 223–238 (2011)
30. Schneider, J., Wattenhofer, R.: A new technique for distributed symmetry breaking. In: PODC 2010, pp. 257–266. ACM, New York (2010)
31. Vizing, V.G.: Some unsolved problems in graph theory. Uspekhi Mat. Nauk 23(6(144)), 117–134 (1968)
32. Van Vu, H.: A general upper bound on the list chromatic number of locally sparse graphs. Comb. Probab. Comput. 11(1), 103–111 (2002)