# DISTRIBUTED $(\Delta+1)$-COLORING IN SUBLOGARITHMIC ROUNDS

DAVID G. HARRIS[1], JOHANNES SCHNEIDER[2], AND HSIN-HAO SU[3]


## Abstract

The $(\Delta+1)$-coloring problem is a fundamental symmetry breaking problem in distributed computing. We give a new randomized coloring algorithm for $(\Delta+1)$-coloring running in $O(\sqrt{\log \Delta}) + 2^{O(\sqrt{\log \log n})}$ rounds with probability $1 - 1/n^{\Omega(1)}$ in a graph with $n$ nodes and maximum degree $\Delta$. This implies that the $(\Delta+1)$-coloring problem is easier than the maximal independent set problem and the maximal matching problem, due to their lower bounds by Kuhn, Moscibroda, and Wattenhofer [PODC'04]. Our algorithm also extends to the list-coloring problem where the palette of each node contains $\Delta+1$ colors.


## 1. Introduction

Given a graph $G = (V, E)$, let $n = |V|$ denote the number of vertices and let $\Delta$ denote the maximum degree. The $k$-coloring problem is to assign each vertex a color from $\{1, 2, \ldots, k\}$ such that no two neighbors are assigned with the same color.

In this paper, we study the $(\Delta+1)$-coloring problem in the distributed LOCAL model. In this model, vertices host processors and operate in synchronized rounds. In each round, each vertex sends one message of arbitrary size to each of its neighbors, receives messages from its neighbors, and performs (unbounded) local computations. The time complexity of an algorithm is measured by the number of rounds until every vertex commits its output, in our case, its color.

The distributed coloring problem, and variants, have a long history dating back to the 1980's. We consider the most common form: the $(\Delta+1)$-coloring problem. Table 1 summarizes the results for this problem. Two major branches of study have been developed, the deterministic approach and the randomized approach. For the deterministic approach, several algorithms with running time of $O(f(\Delta) + \log^* n)$ have been developed [4, 8, 26, 27, 17, 16]. The latter term is necessary as Linial showed that 3-coloring a ring requires $\Omega(\log^* n)$ rounds [27]. It had been questioned whether an algorithm with a sublinear function $f(\Delta)$ exists, since there are $\Omega(\Delta)$ lower bounds for related problems, in more restrictive settings [18, 21, 26, 42]. A breakthrough by Barenboim [4] first gave an algorithm running in $O(\Delta^{3/4} \log \Delta + \log^* n)$ rounds, which is notably sublinear in $\Delta$. Subsequently, the bound was improved to $O(\sqrt{\Delta} \log^{2.5} \Delta + \log^* n)$ by Fraigniaud et al. [13].

The randomized approach can be traced back to the $O(\log n)$ rounds maximal independent set (MIS) algorithm of Alon, Babai, and Itai [1] and Luby [29], where the latter showed that the $(\Delta+1)$-coloring problem can be reduced to the MIS problem. The $O(\log n)$ upper bound lasted until Schneider and Wattenhofer gave an algorithm of running time $O(\log \Delta + \sqrt{\log n})$ [41]. Then, Barenboim et al. [9] improved the dependence on $n$ to $2^{O(\sqrt{\log \log n})}$ by a graph shattering technique. All the algorithms require $\Omega(\log n)$ rounds when $\Delta = n^c$ for some constant $0 < c \leq 1$.

[1]Department of Applied Mathematics, University of Maryland, College Park, MD 20742. Research supported in part by NSF Awards CNS-1010789 and CCF-1422569. Email: davidgharris29@hotmail.com.

[2]ABB Corporate Research, Segelhofstr. 1K, Baden Switzerland. Email: johannes.schneider@ch.abb.com.

[3]Massachusetts Institute of Technology, supported by NSF Awards CCF-0939370, BIO-1455983, and AFOSR FA9550-13-1-0042. Part of the work was done while at University of Michigan, supported by NSF grant CCF-1217338. Email: hsinhao@csail.mit.edu.

| Bounds | Randomized | | Deterministic | |
|---|---|---|---|---|
| Upper | $O(\sqrt{\log \Delta}) + 2^{O(\sqrt{\log \log n})}$ | [This paper] | $O(\sqrt{\Delta} \log^{2.5} \Delta + \log^* n)$ | [13] |
| | $O(\log \Delta) + 2^{O(\sqrt{\log \log n})}$ | [9] | $O(\Delta^{3/4} \log \Delta + \log^* n)$ | [4] |
| | $O(\log \Delta + \sqrt{\log n})$ | [41] | $O(\Delta + \log^* n)$ | [8] |
| | $O(\log n)$ | [29, 1, 22] | $O(\Delta \log \Delta + \log^* n)$ | [26] |
| | | | $O(\Delta^2 + \log^* n)$ | [27, 17] |
| | | | $O(\Delta \log n)$ | [17] |
| | | | $\Delta^{O(\Delta)} + O(\log^* n)$ | [16] |
| | | | $2^{O(\sqrt{\log n})}$ | [34] |
| Lower | $\Omega(\log^* n)$ | [27] | $\Omega(\log^* n)$ | [27] |

TABLE 1. Comparison of $(\Delta + 1)$-coloring algorithms and lower bounds

We give an algorithm that runs in $O(\sqrt{\log \Delta}) + 2^{O(\sqrt{\log \log n})}$ time, which is the first algorithm that runs in $o(\log n)$ rounds for any graph. Moreover, this implies a separation between the $(\Delta+1)$-coloring and the MIS problem. The coloring problem and the MIS problem are related; for example, given a $(\Delta+1)$-coloring one can compute a MIS in $\Delta+1$ rounds by letting a node with color $i$ join the MIS in round $i$ (if no neighbor joined previously). Conversely, Luby [29] showed that any MIS algorithm can be used for $(\Delta + 1)$-coloring in the same running time by simulating it on a blow-up graph. Kuhn, Moscibroda, and Wattenhofer [24] showed that there exists a family of graphs with $\Delta = 2^{O(\sqrt{\log n \log \log n})}$ such that computing an MIS or a maximal matching requires at least $\Omega(\sqrt{\log n / \log \log n})$ rounds[1]. To this date, it has been unclear whether $(\Delta + 1)$-coloring, MIS and maximal matching are equally hard problems. (A separation was known between $(2\Delta - 1)$-edge coloring problem and the maximal matching problem [12]). Our algorithm computes $(\Delta + 1)$-coloring in the above graphs in $O((\log n \log \log n)^{1/4})$ rounds. Thus, it implies $(\Delta + 1)$-coloring is an easier problem.

In addition, our algorithm extends to a closely related generalization of the vertex-coloring problem known as *list-coloring*. Here, each vertex is equipped with a palette containing $(\Delta + 1)$ colors; each vertex selects one color from its palette, and no two neighbors can be assigned the same color. $(\Delta + 1)$-coloring is a special case, in which every vertex has the same palette of size $\Delta + 1$.

1.1. **Technical Summary.** We begin by observing that if we use more colors than are needed, then it is possible to color the graph faster. For example, graphs can be colored very fast using $(1+\Omega(1))\Delta$ colors [41, 12]. Similar ideas apply to sparse graphs, whose chromatic number is known to be smaller than $\Delta+1$. Elkin, Pettie, and Su [12] showed that if a graph is $(1-\epsilon)$-locally-sparse, then it is possible to obtain a $(\Delta + 1)$-coloring in $O(\log(1/\epsilon)) + 2^{O(\sqrt{\log \log n})}$ rounds.

It is thus the dense parts of the graph that become a bottleneck. On the other hand, if a graph is dense, then it is likely to have short diameter. Since computation is free in the distributed setting, a single vertex in the graph can read in all the information in diameter time, make a decision, and broadcast it to the whole graph.

We develop a network decomposition procedure based on *local sparseness*. Our decomposition algorithm is targeted towards identifying dense components of constant weak diameter and sparse components in a constant number of rounds. Roughly speaking, a sparse vertex is one which has $\leq (1 - \epsilon)\Delta^2$ edges in its neighborhood, where $\epsilon > 0$ is a parameter that we will carefully choose. At the same time, we would also like to bound the number of neighbors of a dense component that are not members of the dense component itself, called *external neighbors*. This step is necessary to bound the influence of color choices of nodes in one component on other components. This

[1]The same authors claimed the bound can be improved to $\Omega(\sqrt{\log n})$ with graphs of $\Delta = 2^{O(\sqrt{\log n})}$ in [25]. However, recently Bar-Yehuda, Censor-Hillel, and Schwartzman [3] pointed our an error in their proof.

mechanism may help to leverage algorithms for other distributed problems that can handle either dense or sparse graphs well.

First, we ignore the sparse vertices. Since each dense component has constant weak diameter, it can elect a leader to assign a color to every member so that no intra-component conflicts occur (i.e. the endpoints of the edges inside the same component are always assigned different colors). Meanwhile, we hope that the assignments are random enough so that the chance of inter-component conflicts will be small. Combined with the property of the decomposition that the number of external neighbors is bounded, we show that the probability that a vertex remains uncolored is roughly $O(\epsilon)$ in each round. After $O(\log_{1/\epsilon}\Delta)$ rounds, the degree of each vertex becomes sufficiently small so that the algorithm of Barenboim [4] can handle the residual graph efficiently.

For the sparse vertices we analyze a preprocessing *initial coloring step* of the algorithm. We show that there will be an $\Omega(\epsilon^2\Delta)$ gap between the palette size and the degree due to the sparsity. The gap remains while the dense vertices are colored. So, we will be able to color the sparse vertices by using the algorithm of Elkin et al. [12], which requires $O(\log(1/\epsilon))+\exp(O(\sqrt{\log\log n}))$ rounds. In contrast to [12], our analysis generalizes to the list-coloring problem. By setting $\epsilon = 2^{-\Theta(\sqrt{\log\Delta})}$, we balance the round complexity between the dense part and the sparse part, yielding the desired running time.

The main technical challenge lies in the dense components. In each component, we need to generate a random proper coloring so that each vertex has a small probability of receiving the same color as one of its external neighbors. We give a process for generating a proper coloring where the probability that a vertex gets any color from its palette is close to uniform. Additionally, we will need to show that the structure of the decomposition is maintained so that a vertex remains to have a small fraction of external neighbors in the next round. This requires showing tight concentration bounds on certain quantities. However, the process of generating a random proper coloring creates a cascade of dependence on the colors received by the vertices. Standard concentration inequality arguments based on bounded differences such as Azuma's inequality do not apply. We use instead a novel argument based on the rank statistics of the random permutations, which are independent between vertices in the same component.

1.2. **Overview.** In Section 2, we review related algorithms for network decomposition and coloring. In Section 3, we state our network decomposition. In Section 4, we outline the full algorithm for list-coloring. It consists of two steps: an initial coloring step applied to all vertices, and multiple rounds of dense coloring. In Section 5, we describe the initial coloring step for creating the gap between the palette size and the degree for sparse vertices. In Section 6, we describe a single round of the dense coloring procedure and analyze the behavior of the graph structure. In Section 7, we finish our analysis by solving recurrence relations for dense components which yield the overall algorithm run time. In Section 8, we apply the initial coloring step to give a full algorithm for locally-sparse graphs; this extends the algorithm of [12] to list-coloring.

## 2. Related Work

Various network decompositions have been developed to solve distributed computing problems. Awerbuch et al. [2] introduced the notion of $(d,c)$-decompositions where each component has a diameter $d$ and the contracted graph can be colored in $c$ colors. They give a deterministic procedure for obtaining a $(2^{O(\sqrt{\log n\log\log n})}, 2^{O(\sqrt{\log n\log\log n})})$-decomposition, which can be used to compute a $(\Delta+1)$-coloring and MIS in $2^{O(\sqrt{\log n\log\log n})}$ rounds deterministically. Panconesi and Srinivasan [34] showed how to obtain a $(2^{O(\sqrt{\log n})}, 2^{O(\sqrt{\log n})})$-decomposition, yielding $2^{O(\sqrt{\log n})}$-time algorithms for $(\Delta+1)$-coloring and MIS. Linial and Saks [28] gave a randomized algorithm for obtaining a $(O(\log n), O(\log n))$-decomposition in $O(\log^2 n)$ rounds. Barenboim [7] gave a randomized algorithm for obtaining $(O(1), O_\epsilon(1))$-decompositions in $O(n^\epsilon)$ rounds.

Reed [37] introduced the structural decomposition to study the chromatic number of graphs of bounded clique size (see [31] for a detailed exposition). It was later used for various applications including total coloring, frugal coloring, and computation of the chromatic number [38, 33, 30, 32]. Our network decomposition method is inspired by theirs in the sense that they showed a graph can be decomposed into a sparse component and a number of dense components. However, as their main goal was to study the existential bounds, the properties of the decomposition between our needs are different. For example, the diameter is an important constraint in our case. Also, our decomposition must be computable in parallel, while theirs is obtained sequentially.

The $(\Delta+1)$-coloring algorithms are briefly summarized in Table 1. Barenboim and Elkin's monograph [6] contains an extensive survey of coloring algorithms. Faster algorithms also exist if we use more than $(\Delta+1)$ colors. For deterministic algorithms, Linial [27] and Szegedy and Vishwanathan [42] gave algorithms for obtaining a $O(\Delta^2)$-coloring running in $O(\log^* n)$ rounds. Barenboim and Elkin [5] showed how to obtain an $O(\Delta^{1+\epsilon})$-coloring in $O(\log\Delta\cdot\log n)$ rounds. For randomized algorithms, Schneider and Wattenhofer [41] showed that an $O(\Delta\log^{(k)} n+\log^{1+1/k} n)$-coloring can be obtained in $O(k)$ rounds. Combining the results in [41] with Kothapalli et al. [23], an $O(\Delta)$-coloring can be obtained in $O(\sqrt{\log n})$ rounds. Barenboim et al. [9] showed it can be improved to $2^{O(\sqrt{\log\log n})}$ rounds.

On the other hand, there are algorithms for coloring the graph using less than $(\Delta+1)$ colors for sparse-type graphs. Panconesi and his co-authors [20, 11, 19, 35] developed a line of randomized algorithms for edge coloring (the line graph is sparse) and Brook-Vizing colorings in the distributed setting. For example, in [20], they showed that an $O(\frac{\Delta}{\log\Delta})$-coloring for girth-5 graphs can be obtained in $O(\log n)$ rounds, provided $\Delta=(\log n)^{1+\Omega(1)}$. Pettie and Su [36] generalized it to triangle-free graphs. The restriction on $\Delta$ can be removed by applying the constructive Lovász Local Lemma in the distributed setting [10].

Distributed coloring using less than $\Delta+1$ depending on the chromatic numbers $\chi$ has been investigated by Schneider et al. [39]. They require $(1-1/O(\chi))\cdot(\Delta+1)$ colors for a running time of $O(\log\chi+\log^* n)$ for graphs with $\Delta\in\Omega(\log^{1+1/\log^* n} n)$ and $\chi\in O(\Delta/\log^{1+1/\log^* n} n)$.

More efficient algorithms for $(\Delta+1)$-coloring exist for very dense graphs, e.g. a deterministic $O(\log^* n)$ algorithm for growth bounded graphs (e.g. unit disk graphs) [40], as well as for many types of sparse graphs [9, 12, 36], e.g. for graphs of low arboricity. The arboricity of a graph is the minimum number of edge-disjoint forests, whose union contains all edges of the graph. A graph is $(1-\epsilon)$-locally sparse, if for every vertex $v\in V$, its neighborhood induces at most $(1-\epsilon)\binom{\Delta}{2}$ edges. In [12], a distributed $(\Delta+1)$-coloring algorithm was given for locally-sparse graphs, which we expand to cover list-colorings as well.

As we have discussed, the MIS problem and the coloring problems are related. The MIS can be computed deterministically in $O(\Delta+\log^* n)$ rounds [8] and in $2^{O(\sqrt{\log n})}$ rounds [34]. Very recently, Ghaffari [15] reduced the randomized complexity of MIS to $O(\log\Delta)+2^{O(\sqrt{\log\log n})}$. Whether an MIS can be obtained in polylogarithmic deterministic time or sublogarithmic randomized time remain interesting open problems.

A generalization of MIS, known as an *ruling set*, has also been considered. A $(\alpha,\beta)$-ruling set $U\subseteq V$ [2] is a set of vertices such that two nodes $u,u'\in U$ have distance at least $\alpha$ and for any node $v\in V\setminus U$ there exists a node $u\in U$ with distance at most $\beta$. MIS is a special case, namely a $(2,1)$-ruling set. A number of papers [14, 39, 2] use ruling sets to compute colorings in different kinds of graphs. A ruling set can be viewed as defining a network decomposition, such that any component has diameter at least $\alpha$ and at most $2\beta$.

## 3. Network decomposition and sparsity

In this section, we define a structural decomposition of the graph $G$ into *sparse* and *dense* vertices. We measure these notions with respect to a parameter $\epsilon\in[0,1]$.

**Definition 3.1** (Friend edge)**.** *We say an edge $uv$ is a* friend *edge if $u$ and $v$ share at least $(1-\epsilon)\Delta$ neighbors (i.e. $|N(u) \cap N(v)| \geq (1-\epsilon)\Delta$). We define $F \subseteq E$ to be the set of friend edges.*

*For any vertex $u$, we say $v$ is a friend of $u$ if $uv \in F$; we denote the friends of $u$ by $F(u)$.*

**Definition 3.2** (Dense and sparse vertices)**.** *A vertex $v \in V$ is* dense *if it has at least $(1-\epsilon)\Delta$ friends. Otherwise, it is* sparse*.*

*We write $V^{dense} \subseteq V$ for the set of dense vertices in $G$, and $V^{sparse}$ for the set of sparse vertices in $G$.*

**Definition 3.3** (Weak diameter)**.** *Let $H \subseteq G$ be a subgraph of $G$. For vertices $u, v \in V$, let $d(u,v)$ denote the distance between $u$ and $v$ in $G$. The weak diameter of $H$ is defined to be $\max_{u,v \in H} d(u,v)$.*

Let $C_1, \ldots, C_k$ be the connected components of the subgraph $H = (V^{\text{dense}}, E_H) \subseteq G$, where $E_H = \{uv \mid u, v \in V^{\text{dense}} \text{ and } uv \in F\}$. That is, they are the connected components induced by friend edges and dense vertices. The vertices of $G$ are partitioned disjointly as $V = V^{\text{sparse}} \sqcup V^{\text{dense}} = V^{\text{sparse}} \sqcup C_1 \sqcup \cdots \sqcup C_k$. We refer to each component $C_j$ as an *almost-clique*.

**Lemma 3.4.** *Suppose $\epsilon < 1/5$. Then, for any vertices $x, y \in C_j$, we have $|N(x) \cap N(y)| \geq (1-2\epsilon)\Delta$.*

*Proof.* As $x, y$ are in the same component $C_j$, there is a path of friend edges $x = u_0, \ldots, u_t = y$ connecting them. We claim that $|N(x) \cap N(u_i)| \geq (1-2\epsilon)\Delta$ for all $i \geq 1$. We will show this by induction on $i$. The base case $i = 1$ follows as $xu_1$ is a friend edge.

Now, consider the induction step. As $u_{i-1}u_i$ is a friend, $|N(u_i) \cap N(u_{i-1})| \geq (1-\epsilon)\Delta$. By the induction hypothesis, $|N(x) \cap N(u_{i-1})| \geq (1-2\epsilon)\Delta$.

We thus have:

$$\begin{aligned}
|N(x) \cap N(u_i)| &\geq |N(x) \cap N(u_{i-1}) \cap N(u_i)| \\
&= |N(u_{i-1}) \cap N(u_i)| + |N(u_{i-1}) \cap N(x)| - |(N(u_{i-1}) \cap N(u_i)) \cup (N(u_{i-1}) \cap N(x))| \\
&\qquad \text{(inclusion-exclusion)} \\
&\geq |N(u_{i-1}) \cap N(u_i)| + |N(u_{i-1}) \cap N(x)| - |N(u_{i-1})| \\
&\geq (1-\epsilon)\Delta + (1-2\epsilon)\Delta - \Delta \\
&= (1-3\epsilon)\Delta
\end{aligned}$$

Since $x$ and $u_i$ are dense, we have $|N(x) \setminus F(x)| \leq \epsilon\Delta$ and $|N(u_i) \setminus F(u_i)| \leq \epsilon\Delta$. Therefore, $|F(x) \cap F(u_i)| = |(N(x) \cap N(u_i)) \setminus (N(x) \setminus F(x)) \setminus (N(u_i) \setminus F(u_i))| \geq (1-3\epsilon)\Delta - \epsilon\Delta - \epsilon\Delta \geq (1-5\epsilon)\Delta > 0$.

So $x$ and $u_i$ have a common friend $w$. This implies that $|N(x) \cap N(w)| \geq (1-\epsilon)\Delta$ and $|N(u_i) \cap N(w)| \geq (1-\epsilon)\Delta$. By a similar inclusion-exclusion argument, this implies $|N(x) \cap N(u_i)| \geq (1-2\epsilon)\Delta$. □

**Corollary 3.5.** *Suppose $\epsilon < 1/5$. Then all almost-cliques have weak diameter at most 2.*

*Proof.* By Lemma 3.4, any vertices $x, y \in C_j$ have $|N(x) \cap N(y)| \geq (1-2\epsilon)\Delta > 0$. In particular, they have a common neighbor. □

A vertex $v$ in $C_j$ can identify all other members of $C_j$ in $O(1)$ rounds by the following: Initially, each vertex $u \in G$ broadcasts the edges incident to $u$ to all nodes within distance 3. In this way, every vertex $v$ learns the graph topology of all nodes up to distance 3, which is sufficient to determine whether an edge (both of whose endpoints are within distance 2 of $v$) is a friend edge and whether a vertex (within distance 2) is dense. Since by Corollary 3.5, all members of $C_j$ are within distance 2 to $v$, all the members can be identified. Also, the leader of $C_j$ can be elected as the member with the smallest ID.

**Definition 3.6** (External degree)**.** *For any dense vertex $v \in C_j$, we define $\overline{d}(v)$, the* external degree *of $v$, as the number of* dense *neighbors of $v$ outside $C_j$. (Sparse neighbors are not counted.)*

**Lemma 3.7.** *For any dense vertex $v$, we have $\overline{d}(v) \leq \epsilon\Delta$.*

*Proof.* Let $v \in C_j$ be dense. As $v$ is dense, it has at least $(1-\epsilon)\Delta$ friends. So it has at most $\epsilon\Delta$ dense vertices which are not friends. If any dense vertex $w$ is a friend of $v$, then by definition $w \in C_j$. So $v$ has at most $\epsilon\Delta$ dense neighbors outside $C_j$. □

**Definition 3.8** (Anti-degree)**.** *For any dense vertex $v \in C_j$, we define the* anti-degree *of $v$ to be $a(v) = |C_j \setminus N(v)|$.*

**Lemma 3.9.** *Suppose $\epsilon < 1/5$. Then for any $C_j$ and $v \in C_j$, we have $a(v) \leq 3\epsilon\Delta$.*

*Proof.* We will show this by counting in two ways the number of length-2 paths of the form $v, x, u$ where $x \in G$ and $u \in C_j \setminus N(v)$. First, observe that for any $u \in C_j \setminus N(v)$, there are precisely $|N(v) \cap N(u)|$ possibilities for the middle vertex $x$. Thus we have

$$R = \sum_{u \in C_j \setminus N(v)} |N(v) \cap N(u)|;$$

by Lemma 3.4, for any $u \in C_j \setminus N(v)$ we have $|N(v) \cap N(u)| \geq (1-2\epsilon)\Delta$. So $R \geq a(v)(1-2\epsilon)\Delta$.

We can also count $R$ by summing over the middle vertex $x$:

$$\begin{aligned}
R &= \sum_{x \in N(v)} |N(x) \cap (C_j \setminus N(v))| \\
&\leq \sum_{x \in N(v)} |N(x) \setminus N(v)| \\
&= \sum_{x \in F(v)} |N(x) \setminus N(v)| + \sum_{x \in N(v) \setminus F(v)} |N(x) \setminus N(v)| \\
&\leq \sum_{x \in F(v)} \epsilon\Delta + \sum_{x \in N(v) \setminus F(v)} \Delta \\
&\leq \sum_{x \in N(v)} \Delta - \sum_{x \in F(v)} (\Delta - \epsilon\Delta) \\
&\leq \Delta^2 - (1-\epsilon)(\Delta - \epsilon\Delta) \qquad |F(v)| \geq (1-\epsilon)\Delta \text{ as } v \text{ is dense} \\
&= \Delta^2(2-\epsilon)\epsilon
\end{aligned}$$

Thus, we must have

$$a(v)(1-2\epsilon)\Delta \leq R \leq \Delta^2(2-\epsilon)\epsilon$$

So $a(v) \leq \Delta\frac{\epsilon(2-\epsilon)}{1-2\epsilon}$; this is at most $3\epsilon\Delta$ for $\epsilon < 1/5$. □

**Corollary 3.10.** *For $\epsilon < 1/5$, all almost-cliques have size at most $(1+3\epsilon)\Delta$.*

*Proof.* Let $v \in C_j$. Then $|C_j| = |C_j \setminus N(v)| + |C_j \cap N(v)| \leq a(v) + |N(v)| \leq (1+3\epsilon)\Delta$. □

## 4. Full algorithm outline

We can now describe our complete algorithm for list-coloring graphs, whether sparse or dense. We set the density parameter $\epsilon = C \cdot 100^{-\lceil\sqrt{\ln\Delta}\rceil}$, where $C > 0$ is a small constant (to be specified later). Also, we assume that $\epsilon^4\Delta \geq K \ln n$ for sufficiently large constant $K$; if $\epsilon^4\Delta < K \ln n$, then $\Delta < \text{polylog}(n)$, and so then the coloring procedure of [9] will already color the graph in $O(\log\Delta) + 2^{O(\sqrt{\log\log n})} = 2^{O(\sqrt{\log\log n})}$ rounds.

1. Decompose $G$ into $V^{sparse}, C_1, \ldots, C_k$.
2. Execute *the initial coloring step* for all vertices.
3. For $i = 1, \ldots, \lceil\sqrt{\ln\Delta}\rceil$, execute the $i^{\text{th}}$ *dense coloring step* on the dense vertices.

4. Run the algorithm of [12] to color the sparse vertices.
5. Run the algorithm of [9] to color the residual graph.

We note that the decomposition of $G$ in step (1) remains fixed for the entire algorithm. Although in later steps vertices become colored and are removed from $G$, we always define the decomposition in terms of the *original* graph $G$, not the residual graph. However, we abuse notation so that when we refer to a component $C_j$ during an intermediate step, we mean the intersection of $C_j$ with the residual (uncolored) vertices.

The initial coloring step is the following: With probability 99/100, each vertex does nothing; otherwise it selects a color from its palette. All these choices are made uniformly and independently. If two adjacent vertices select the same color, then the vertex with higher ID discards its choice of color (i.e. it becomes decolored).

We assume at the end of the initial coloring step and the end of each dense coloring step, if a vertex gets colored, it will be removed from the graph (as well as from the vertex set $V^{\text{sparse}}, C_1, \ldots, C_k$ it belongs to). Also, the color used by it will be removed from the palettes of its neighbors. For any vertex $v$, let $\text{Pal}_0(v), d_0(v)$ denote, respectively, the palette and degree of $v$ after the initial coloring step (note that the 0 does not denote time 0, but the time immediately after the initial coloring step), and we let $Q_0(v) = |\text{Pal}_0(v)|$. We will show in Lemma 5.8 that whp for every sparse vertex $v$ we have

$$Q_0(v) \geq d_0(v) + \Omega(\epsilon^2\Delta).$$

and that for every vertex $v$ (sparse or dense) we have $Q_0(v) \geq \Delta/2$.

Then we turn our attention to the dense vertices and we will show that they can be colored efficiently. For a dense vertex $x \in C_j$, we let $\overline{d}_0(x)$ and $a_0(x)$ denote its external degree and anti-degree after the initial coloring step. Let $\text{Pal}_i(x), d_i(x), \overline{d}_i(x), a_i(x), Q_i(x)$ denote the quantities at the end of the $i^{\text{th}}$ dense coloring step. As we color the graph, we maintain two key parameters, $D_i, Z_i$ which bound the external degree, anti-degree, and palette size for dense vertices after the $i^{\text{th}}$ dense coloring step. Namely, we ensure the invariant that for all dense vertices $v$ we have

$$a_i(v) \leq D_i, \overline{d}_i(v) \leq D_i, Q_i(v) \geq Z_i$$

Initially,

$$D_0 = 3\epsilon\Delta \qquad Z_0 = \Delta/2$$

The $i^{\text{th}}$ dense coloring step is as follows:

1. For each $C_j$, elect a leader to simulate the following process to color $C_j$:
    2. Generate a random permutation $\pi_j$ on $1, \ldots, |C_j|$
    3. For $k = 1, \ldots, L = \lceil |C_j|(1 - 2(D_{i-1}/Z_{i-1})\ln(Z_{i-1}/D_{i-1}))\rceil$:
        4. $v_{\pi(k)} \in C_j$ randomly selects a color in its palette not selected by its neighbors in $C_j$.
5. If some vertex $v \in C_j$ has a color conflict with a vertex in $C_{j'}$ for $j' < j$, then decolor $v$. Note that we can simply use the ID of the leader of the almost-clique as the index of the almost-clique for this.

It may seem more natural to attempt to color all the vertices of $C_j$ in step 3, as opposed to only $L$ of them. However, this would cause the palette sizes to shrink too quickly. We discuss this issue in more detail in Section 6. (As we will show in Lemma 7.2, the choice of $L$ in step (3) is meaningful, that is, we have $(1 - 2D_i/Z_i \ln(Z_i/D_i)) \in [0, 1]$ for all $i \leq \lceil\sqrt{\ln\Delta}\rceil$. )

Note that the excess of palette size over degree can only increase during the course of this algorithm (every time we color a neighbor of $v$, we delete at most one color from $\text{Pal}(v)$. So at the end of the dense coloring steps, we have that for every sparse vertex $v$

$$Q_{\lceil\sqrt{\ln\Delta}\rceil}(v) - d_{\lceil\sqrt{\ln\Delta}\rceil}(v) \geq Q_0(v) - d_0(v) \geq \Omega(\epsilon^2\Delta)$$

The algorithm of Elkin et al. [12] is designed for list-coloring in which the palette sizes significantly exceed the degree. This indeed holds for the sparse vertices: their palette sizes exceeds their degree by $\Omega(\epsilon^2\Delta)$. Thus they can be colored in $O(\log(1/\epsilon^2)) + 2^{O(\sqrt{\log\log n})} = O(\sqrt{\log\Delta}) + 2^{O(\sqrt{\log\log n})}$ rounds. This removes the sparse vertices from the graph, leaving only the dense vertices behind.

After the sparse vertices are removed, Theorem 7.5 shows that each remaining dense vertex is connected to $\Delta' \leq O(\log n)\cdot 2^{O(\sqrt{\log\Delta})}$ other vertices. The algorithm of [9] then takes $O(\log\Delta') + 2^{O(\sqrt{\log\log n})} = O(\sqrt{\log\Delta}) + 2^{O(\sqrt{\log\log n})}$ steps.

## 5. The Initial Coloring Step

We assume that the vertices initially have a palette containing exactly $\Delta+1$ colors. The key to analyzing the local situation, as in [12], is to show that after the initial coloring step, every sparse vertex has significantly more colors in its palette than it has neighbors.

Recall the procedure is that with probability 99/100, each vertex does nothing; otherwise it selects a color $A(v)$ from its palette uniformly at random and discards its choice of color if a neighbor with lower ID chooses the same color. For each vertex, we let $B(v)$ denote the final choice of color; we say $A(v)=0$ if vertex $v$ chose not to select a color initially and we say $B(v)=0$ if $v$ is uncolored (either because it did not select an initial color, or it became decolored). One simple property of this process is that any vertex receives any potential color with probability $\Omega(1/\Delta)$:

**Lemma 5.1.** *For any vertex $v$ and any color $c \in \mathrm{Pal}(v)$, we have*

$$P(B(v)=c) \geq \frac{0.009}{\Delta+1}$$

*Proof.* We have $B(v)=c$ if $A(v)=c$ and there is no $w\in N(v)$ with $A(w)=c$ and $\mathrm{ID}(w)<\mathrm{ID}(v)$.

$$\begin{aligned} P(B(v)=c) &\geq P(A(v)=c)\prod_{w\in N(v)} P(A(w)\neq c) \\ &\geq 1/100\cdot\frac{1}{\Delta+1}\cdot\left(1-\frac{1}{100(\Delta+1)}\right)^{|N(v)|} \\ &\geq \frac{0.009}{\Delta+1} \end{aligned}$$

□

One crucial property of this process is that each $B(v)$ is completely determined by the random variables $A(u)$, where $\mathrm{ID}(u)\leq\mathrm{ID}(v)$. We can think of this coloring procedure as a stochastic process (in which the vertex ID plays the role of time).

Let us now fix some sparse vertex $v$, and show that after the initial coloring step $v$ has more colors in its palette than it has neighbors. For each color $c$, let $X_c$ denote the number of neighbors $w\in N(v)$ which have $B(w)=c$ at the end of the coloring process.

**Lemma 5.2.** *Let $d(v), \mathrm{Pal}(v)$ denote the initial degree and palette of vertex $v$. Then we have*

$$Q_0(v)-d_0(v) \geq \Delta+1-d(v)+\sum_{c\in\mathrm{Pal}(v)}\max(0,X_c-1)+\sum_{c\notin\mathrm{Pal}(v)}X_c$$

*where recall that $d_0(v), Q_0(v)$ are the degree and palette size of $v$ after the initial coloring step.*

*Proof.* Suppose we go through the vertices in an increasing order of their ID, $v_1, v_2, \ldots, v_n$, where $\mathrm{ID}(v_1)<\mathrm{ID}(v_2)<\ldots<\mathrm{ID}(v_n)$. At stage $i$ we fix the color of vertex $v_i$ to $B(v_i)$. If $B(v_i)\neq 0$, we remove $v_i$ from the graph and remove $B(v_i)$ from the palette of all its neighbors. Given the vertex $v$, let $Q(v)$ denote the palette size and let $\phi_i$ denote the value of $Q(v)-d(v)$ after processing vertices

$v_1, \ldots, v_i$ in this manner. In this case, $\phi_0 = Q(v) - d(v) = \Delta + 1 - d(v)$ and $\phi_n = Q_0(v) - d_0(v)$, which is the quantity we are trying to estimate.

Suppose that $v_i$ is a neighbor of $v$. Let us now examine how $\phi_i$ changes depending on $B(v_i)$. If $B(v_i) = 0$, then $v_i$ remains in the residual graph and neither $\mathrm{Pal}(v)$ nor $d(v)$ are affected, so $\phi_{i+1} = \phi_i$.

Suppose that $B(v_i) = c$ and $c \notin \mathrm{Pal}(v)$. This means that $v_i$ has selected a color not appearing in $\mathrm{Pal}(v)$. Thus, the degree of $v$ decreases by one while its palette is unaffected, so $\phi_{i+1} = \phi_i + 1$. Note the case that $c \notin \mathrm{Pal}(v)$ can only occur in the list-coloring problem.

Suppose that $B(v_i) = c$ and $c \in \mathrm{Pal}(v)$. If $v_i$ is the first vertex such that $B(v_i) = c$, we have $\phi_{i+1} = \phi_i$ and color $c$ is removed from $\mathrm{Pal}(v)$. Otherwise, there exists $j < i$ with $B(v_j) = c$, then $c$ is no longer in $\mathrm{Pal}(v)$ and thus we have $\phi_{i+1} = \phi_i + 1$.

Thus, we have identified $\max(0, X_c - 1)$ (for $c \in \mathrm{Pal}(v)$) and $X_c$ vertices (for $c \notin \mathrm{Pal}(v)$) which select color $c$ and cause $\phi_{i+1} = \phi_i + 1$. These vertices all must be disjoint, so there are at least $\sum_{c \in \mathrm{Pal}(v)} \max(0, X_c - 1) + \sum_{c \notin \mathrm{Pal}(v)} X_c$ vertices for which $\phi_{i+1} = \phi_i + 1$.

This implies that $\phi_n \geq \phi_0 + \sum_{c \in \mathrm{Pal}(v)} \max(0, X_c - 1) + \sum_{c \notin \mathrm{Pal}(v)} X_c$, which is what we claim. □

For any vertex $v$ and color $c$, we say that color $c$ is *good* for $v$ if the following occurs. For $c \in \mathrm{Pal}(v)$, then $c$ is good for $v$ if $X_c \geq 2$; if $c \notin \mathrm{Pal}(v)$, then $c$ is good for $v$ if $X_c \geq 1$. Let $J(v)$ denote the set of colors that are good for $v$. By Lemma 5.2, we have that

$$Q_0(v) - d_0(v) \geq \Delta + 1 - d(v) + |J(v)|$$

We will next show that $|J(v)|$ is large with high probability. For each vertex $v$ and color $c$, we define $N_c(v)$ to be the set of neighbors whose palette contains $c$; that is,

$$N_c(v) = \{w \in N(v) \mid c \in \mathrm{Pal}(w)\}$$

For colors $c \notin \mathrm{Pal}(v)$, it is easy to show that $c$ has a good probability of going into $J(v)$:

**Lemma 5.3.** *Suppose that $c \notin \mathrm{Pal}(v)$. Then*

$$P(c \in J(v)) \geq \Omega\Big(\frac{|N_c(v)|}{\Delta}\Big)$$

*Proof.* For $c \notin \mathrm{Pal}(v)$, we have $c \in J(v)$ iff $X_c \geq 1$. By inclusion-exclusion, we have

$$\begin{aligned}
P(X_c \geq 1) &\geq \sum_{w \in N_c(v)} P(B(w) = c) - \sum_{w, w' \in N_c(v)} P(B(w) = B(w') = c) \\
&\geq \sum_{w \in N_c(v)} P(B(w) = c) - \sum_{w, w' \in N_c(v)} P(A(w) = A(w') = c) \\
&\geq \frac{0.009|N_c(v)|}{\Delta + 1} - \frac{|N_c(v)|^2}{100^2(\Delta+1)^2} \qquad \text{by Lemma 5.1} \\
&\geq \frac{0.009|N_c(v)|}{\Delta + 1} - \frac{|N_c(v)|\Delta}{100^2(\Delta+1)^2} \\
&\geq 0.004|N_c|/\Delta
\end{aligned}$$

□

For colors $c \in \mathrm{Pal}(v)$, it is harder to bound the probability that $c \in J(v)$.

**Lemma 5.4.** *Suppose that $\epsilon\Delta \geq 3$ and $\epsilon < 1/5$. If $v$ is a sparse vertex, $c \in \mathrm{Pal}(v)$, and $|N_c(v)| \geq (1 - 0.01\epsilon)\Delta$, then $P(c \in J(v)) \geq \Omega(\epsilon^2)$.*

*Proof.* Let $S$ denote the set of all neighbors of $v$ which contain color $c$ and are not friends of $v$;

$$S = \{w \in N_c(v) \mid wv \notin F\}$$

By definition of sparsity, $v$ has at most $(1-\epsilon)\Delta$ friends. Thus, $|S| \geq |N_c(v)| - (1-\epsilon)\Delta \geq 0.99\epsilon\Delta$.

For each $w \in S$, we have $|N(w) \cap N(v)| < (1-\epsilon)\Delta$. So $|N_c(v) - N(w) - \{w\}| \geq |N_c(v)| - 1 - |N(v) \cap N(w)| > 0.99\epsilon\Delta - 1$; by our assumption on the size of $\epsilon\Delta$, this is at least $\frac{1}{2}\epsilon\Delta$. Thus, for each $w \in S$, one can identify a subset of vertices $H_w$ with the following properties:

(1) $H_w \subseteq N_c(v) - N(w) - \{w\}$
(2) $|H_w| = \lceil \epsilon\Delta/2 \rceil$

Let us fix some subset $S' \subseteq S$, of cardinality exactly $|S'| = \lceil 0.01\epsilon\Delta \rceil$. Now, note that a sufficient condition to have $X_c \geq 2$ is that there is some $w \in S'$ and $u \in H_w$ with $B(w) = B(u) = c$. This happens with probability at least:

$$\begin{aligned} &P(B(w) = c \wedge B(u) = c \text{ for some } w \in S', u \in H_w) \geq \\ &\sum_{w \in S', u \in H_w} P(B(w) = B(u) = c) - \sum_{w \in S', u, u' \in H_w, u \neq u'} P(B(w) = B(u) = B(u') = c) \\ &\quad - \sum_{w, w' \in S', w \neq w'} P(B(w) = B(w') = c) \end{aligned}$$

Notice that the inequality holds by considering the number of times the event "there is some $w \in S'$ and $u \in H_w$ with $B(w) = B(u) = c$" is counted on both sides. Let $A = \{w \in S' \mid B(w) = c\}$ and $B_a = \{u \in H_a \mid B(u) = c\}$. It suffices to show that $1 \geq \sum_{a \in A} \sum_{b \in B_a} 1 - \sum_{a \in A} \sum_{b, b' \in B_a, b \neq b'} 1 - \sum_{a, a' \in A, a \neq a'} 1$.

$$\begin{aligned} \sum_{a \in A} \sum_{b \in B_a} 1 - \sum_{a \in A} \sum_{b, b' \in B_a, b \neq b'} 1 - \sum_{a, a' \in A, a \neq a'} 1 &= \sum_{a \in A} (|B_a| - |B_a| \cdot (|B_a| - 1)) - \sum_{a, a' \in A, a \neq a'} 1 \\ &\leq \sum_{a \in A} 1 - \sum_{a, a' \in A, a \neq a'} 1 \\ &= |A| - |A| \cdot (|A| - 1) \leq 1 \end{aligned}$$

We can derive an upper bound on $P(B(w) = B(u) = B(u') = c)$ by noting that a necessary condition for this event is that $A(w) = A(u) = A(u') = c$, and this occurs with probability exactly $\frac{1}{100^3(\Delta+1)^3}$. There are at most $(0.01\epsilon\Delta + 1) \cdot (\epsilon\Delta/2 + 1)^2$ choices for $w, u, u'$ so this term is at most $2 \cdot 10^{-8}\epsilon^3$.

Similarly, we have that $P(B(w) = B(w') = c) \leq \frac{1}{100^2(\Delta+1)^2}$. There are at most $(0.01\epsilon\Delta + 1)^2$ choices for $w, w'$ so this term is at most $4 \cdot 10^{-8}\epsilon^2$.

Next, consider some $u \in H_w$. A sufficient condition to have $B(w) = B(u) = c$ is if $A(w) = A(u) = c$ and there is no $z \in N(u) \cup N(w)$ with $A(z) = c$. Furthermore, any such $z$ cannot itself be equal to $u$ or $w$ as $u$ and $w$ are non-neighbors. Thus,

$$\begin{aligned} P(B(w) = B(u) = c) &\geq P(A(w) = A(u) = c) \prod_{z \in N(u) \cup N(w)} P(A(z) \neq c) \\ &\geq \frac{1}{100^2(\Delta+1)^2} \left(1 - \frac{1}{100(\Delta+1)}\right)^{|N(u) \cup N(w)|} \\ &\geq 9.8 \cdot 10^{-5} \cdot (\Delta+1)^{-2} \qquad \text{as } |N(u) \cup N(w)| \leq 2\Delta \end{aligned}$$

Thus, we have

$$\sum_{w \in S', u \in H_w} P(B(w) = B(u) = c) \geq 0.01\epsilon\Delta \cdot \epsilon\Delta/2 \cdot 9.8 \cdot 10^{-5} \cdot (\Delta+1)^{-2} \geq 10^{-7}\epsilon^2$$

And overall we thus have

$$P(X_c \geq 2) \geq 10^{-7}\epsilon^2 - 4 \cdot 10^{-8}\epsilon^2 - 2 \cdot 10^{-8}\epsilon^3 \geq \Omega(\epsilon^2)$$

□

**Lemma 5.5.** *Suppose that $\epsilon\Delta \geq 3$ and $\epsilon < 1/5$. For any sparse vertex $v$ with $d(v) \geq (1-0.005\epsilon)\Delta$, we have $\mathbf{E}[|J(v)|] \geq \Omega(\epsilon^2\Delta)$.*

*Proof.* Observe that for each $w \in N(v)$, there are exactly $\Delta+1$ values of $c$ for which $w \in N_c(v)$. Hence, by double counting, we have

$$\sum_c |N_c(v)| = (\Delta+1)|N(v)| \tag{1}$$

Let us divide the set of colors into three disjoint sets

$$\begin{aligned} A_1 &= \{c \mid c \notin \mathrm{Pal}(v)\} \\ A_2 &= \{c \mid c \in \mathrm{Pal}(v), |N_c(v)| \geq (1-0.01\epsilon)\Delta\} \\ A_3 &= \{c \mid c \in \mathrm{Pal}(v), |N_c(v)| < (1-0.01\epsilon)\Delta\} \end{aligned}$$

We can decompose the sum (1) as:

$$\begin{aligned} (\Delta+1)|N(v)| = \sum_c |N_c(v)| &= \sum_{c\in A_1} |N_c(v)| + \sum_{c\in A_2} |N_c(v)| + \sum_{c\in A_3} |N_c(v)| \\ &\leq \sum_{c\notin \mathrm{Pal}(v)} |N_c(v)| + |A_2|\Delta + |A_3|(1-0.01\epsilon)\Delta \end{aligned}$$

Rearranging, and using the fact that $|A_2| + |A_3| = \Delta+1$, we have

$$\begin{aligned} \sum_{c\notin \mathrm{Pal}(v)} |N_c(v)| &\geq (\Delta+1)(1-0.005\epsilon)\Delta - |A_2|\Delta - |A_3|(1-0.01\epsilon)\Delta \\ &\geq (\Delta+1)(1-0.005\epsilon)\Delta - |A_2|\Delta - (\Delta+1-|A_2|)(1-0.01\epsilon)\Delta \\ &= (\Delta+1)((1-0.005\epsilon)-(1-0.01\epsilon))\Delta - |A_2|(1-(1-0.01\epsilon))\Delta \\ &= (\Delta+1)(0.005\epsilon)\Delta - |A_2|(0.01\epsilon)\Delta \\ &= 0.005\epsilon\Delta(\Delta+1-2|A_2|) \\ &= \Omega(\epsilon\Delta(\Delta-2|A_2|)) \end{aligned}$$

Thus, we have

$$\begin{aligned} \mathbf{E}\Big[\sum_{c\notin \mathrm{Pal}(v)} [c\in J(v)]\Big] &\geq \sum_{c\notin \mathrm{Pal}(v)} |N_c(v)|\Omega(1/\Delta) \qquad \text{by Lemma 5.3} \\ &\geq \max(0, \epsilon\Delta(\Delta-2|A_2|)\cdot\Omega(1/\Delta)) \\ &\geq \max(0, \Omega(\epsilon(\Delta-2|A_2|))) \end{aligned}$$

Also, by Lemma 5.4, for each $c \in A_2$ we have $P(c \in J(v)) \geq \Omega(\epsilon^2)$. So, summing over all $c \in \mathrm{Pal}(v)$ we have

$$\mathbf{E}[\sum_{c\in \mathrm{Pal}(v)} [c\in J(v)] + \sum_{c\notin \mathrm{Pal}(v)} [c\in J(v)]] \geq |A_2|\Omega(\epsilon^2) + \Omega(\max(0, \epsilon(\Delta-2|A_2|)))$$

(Here and in the remainder of the paper, we use the Iverson notation so that for any predicate $\mathcal{P}$, $[\mathcal{P}]$ is equal to 1 if $\mathcal{P}$ is true and zero otherwise.)

This expression is piecewise-linear in $|A_2|$, so it must achieve its minimum value at one of its corner points $|A_2| = 0, \Delta/2, \Delta+1$. At these points, it takes on the expressions respectively $\Omega(\epsilon\Delta), \Delta/2\cdot\Omega(\epsilon^2)$, and $(\Delta+1)\Omega(\epsilon^2)$. Hence, in all three cases it is at least $\Omega(\epsilon^2\Delta)$. □

We next show that there is a concentration phenomenon for the number of good colors.

**Lemma 5.6.** *Suppose $\epsilon < 1/5$. Let $v$ be a sparse vertex. With probability at least $1 - e^{-\Omega(\epsilon^4 \Delta)}$, we have $d_0(v) - Q_0(v) \geq \Omega(\epsilon^2 \Delta)$.*

*Proof.* If $d(v) \leq (1 - 0.005\epsilon\Delta)$ then $Q_0(v) - d_0(v) \geq 0.005\epsilon\Delta \geq \Omega(\epsilon^2\Delta)$ with certainty. So, we may assume $d(v) > (1 - 0.005\epsilon\Delta)$.

Also, suppose that $\epsilon\Delta < 3$. In this case, we need to show that $Q_0(v) - d_0(v) \geq \Omega(\epsilon)$. But, in the initial graph, we have $Q(v) - d(v) = 1 \geq \Omega(\epsilon)$, so again in this case the event holds with certainty. So, we may assume that $\epsilon\Delta \geq 3$.

If none of these occur, then we will show that $|J(v)| \geq \Omega(\epsilon^2\Delta)$ with probability at least $1 - e^{-\Omega(\epsilon^4\Delta)}$ which suffices to show this claim.

Let $W = \{v\} \cup N(v)$ and let $\bar{W}$ denote the set of vertices whose distance to $v$ are exactly 2. Suppose we examine the values of $A(u)$, where $u \in W$. Some of these vertices may decolor others; other vertices may or may not become decolored, based on $A(w)$ where $w \in \bar{W}$. Based only on the colors of the vertices in $W$, we may derive a set "pre-good" colors $J'(v)$; that is, colors $c$ which will go into $J(v)$ unless they become de-colored due to vertices in $\bar{W}$.

Observe that $J(v) \subseteq J'(v)$, so that $\mathbf{E}[|J'(v)|] \geq \Omega(\epsilon^2\Delta)$. Let $\phi > 0$ be a constant such that $\mathbf{E}[|J'(v)|] \geq \phi\epsilon^2\Delta$ for all $\Delta \geq 1$. Also, observe that for $u \in N(v)$ changing the value of $A(u)$ may only change $|J'(v)|$ by at most 2; (the value of $A(u)$ can only affect $A(u) \in J(v)$; colors $c' \neq A(u)$ are not affected). Hence, by the bounded differences inequality, the probability that $|J'(v)|$ is smaller than $\phi\epsilon^2\Delta$ by an amount of $\frac{\phi}{2}\epsilon^2\Delta$ is at most

$$\exp\left(-\frac{(\phi\epsilon^2\Delta)^2}{2 \cdot \sum_{i \in \{v\} \cup N(v)} 2^2}\right) \leq \exp(-\Omega(\epsilon^4\Delta))$$

Now, let us condition on the event $|J'(v)| \geq \frac{\phi}{2}\epsilon^2\Delta$. This event depends only on the values of $A(u)$ for $u \in W$. So, the values of $A(w)$ for $w \in \bar{W}$ are still independent and uniform. Each such vertex has the possibility of decoloring a vertex in $W$, possibly causing a color in $J'(v)$ to not occur in $J(v)$.

For each color $c \in J'(v), c \notin \mathrm{Pal}(v)$, let $y_c$ denote the vertex with the smallest ID in the neighborhood of $v$ with $A(y_c) = c$ and not decolored by any vertices in $W$. Similarly, if $c \in J'(v), c \in \mathrm{Pal}(v)$, let $y_c, y'_c$ denote the two vertices with smallest IDs in the neighborhood of $v$ with $A(y_c) = A(y'_c) = c$ and not decolored by any vertices in $W$ (so $y_c$ and $y'_c$ cannot be neighbors). Such colors will go into $J(v)$ unless a vertex in $N(y_c)$ selects $c$ (respectively, in $N(y_c) \cup N(y'_c)$ selects color $c$.)

If a vertex $w \in \bar{W}$ selects $A(w) = c$ for such a color $c$, causing color $c$ to not appear in $J(v)$, we say that $w$ *disqualifies* color $c$. Observe that

$$|J(v)| \geq |J'(v)| - \sum_{\substack{c \in J'(v) \\ w \in (N(y_c) \cup N(y'_c)) \cap \bar{W}}} [w \text{ disqualifies color } c].$$

So it suffices to show that

$$\sum_{\substack{c \in J'(v) \\ w \in (N(y_c) \cup N(y'_c)) \cap \bar{W}}} [w \text{ disqualifies color } c] < \tfrac{\phi}{4}\epsilon^2\Delta$$

with good probability. Now, observe that each event that $w$ disqualifies color $c$ occurs with probability at most $1/(100(\Delta + 1))$. Furthermore, for each color $c$, there are at most $2\Delta$ choices of $w$ that can disqualify it. Hence, the expected number of such disqualifications is at most $\frac{\phi}{2}\epsilon^2\Delta \cdot (2\Delta) \cdot 1/(100(\Delta + 1)) \leq 0.01\phi\epsilon^2\Delta$.

Furthermore, all such disqualification events are negatively correlated (they are not necessarily independent; a vertex $w$ may possibly disqualify multiple colors). Hence, Chernoff's bound applies, and the probability that the number of disqualifications exceeds $0.02\phi\epsilon^2\Delta$ is at most $\exp(-\Omega(0.01\phi\epsilon^2\Delta)) = \exp(-\Omega(\epsilon^2\Delta))$.

Overall, we have that $|J(v)| \geq \Omega(\epsilon^2\Delta)$ with probability $1-\exp(-\Omega(\epsilon^4\Delta))$. □

We also note a useful property of this coloring procedure: vertices, whether sparse or dense, retain most of their palette:

**Lemma 5.7.** *At the end of this procedure, for any vertex $v$ we have*

$$P(Q_0(v) \geq \Delta/2) \geq 1 - e^{-\Omega(\Delta)}$$

*Proof.* Each vertex $w \in N(v)$ chooses an initial color with probability at most $1/100$, independently of any other vertices. Thus, a simple Chernoff bound shows that there is a probability of $e^{-\Omega(\Delta)}$ that there are no more than $\Delta/2$ neighbors of $v$ which are colored. So with probability $1-e^{-\Omega(\Delta)}$, vertex $v$ loses $\leq \Delta/2$ colors from its original palette size of $\Delta+1$. □

**Lemma 5.8.** *For $K$ a sufficiently large constant, the following events occur whp:*

(1) *For every sparse vertex $v$ we have $d_0(v) - Q_0(v) \geq \Omega(\epsilon^2\Delta)$*
(2) *For every vertex $v$ we have $Q_0(v) \geq \Delta/2$*

*Proof.* By Lemma 5.6, for any individual sparse vertex $v$ the probability that (1) fails is at most $e^{-\Omega(\epsilon^4\Delta)}$. Since $\epsilon^4\Delta \geq K\ln n$, then for $K$ sufficiently large this is $< n^{-100}$. We take a union bound over all sparse vertices and the overall probability that there is some vertex $v$ violating (1) is also $\leq n^{-\Omega(1)}$.

By Lemma 5.7, for any individual vertex $v$ the probability that (2) fails is at most $e^{-\Omega(\Delta)}$. Again, for $K$ sufficiently large, this is $\leq n^{-100}$ and we have that (2) holds whp for all vertices. □

## 6. Coloring the dense vertices

We suppose that we are at the beginning of the $i^{\text{th}}$ dense coloring step. We assume that there are parameters $D_{i-1}, Z_{i-1}$ with the following properties such that for all dense vertices $v$ we have:

(1) $a_{i-1}(v) \leq D_{i-1}$
(2) $\bar{d}_{i-1}(v) \leq D_{i-1}$
(3) $Q_{i-1}(v) \geq Z_{i-1}$

Henceforth we will suppress the dependence on $i$ and write simply $D, Z, \delta = D/Z, a(v), \bar{d}(v), \mathrm{Pal}(v)$, and $Q(v)$. Recall that the dense coloring step (see Section 4) is that every $C_j$ generates a rank for its members. Starting from the vertex with rank 1 to rank $\lceil |C_j|(1-2\delta\ln(1/\delta))\rceil$, each vertex selects a color from its palette excluding the colors selected by lower rank vertices uniformly at random. This is done by having a leader in $C_j$ simulating the process. Then, a vertex becomes decolored if there is an external neighbor from a lower indexed component choosing the same color .

Our goal is to show for some new parameters $D'$ and $Z'$ that we have at the end of the round $a_i(v) \leq D'$, $\overline{d}_i(v) \leq D'$ and $Q_i(v) \geq Z'$. To do this, we will show that most vertices are colored in round $i$.

We require throughout this section the following conditions on $D$ and $Z$, which we will refer to as the *regularity conditions*:

(1) $D\delta \geq K\ln n$ for some sufficiently large constant $K$;
(2) $\delta \leq 1/K$ for some sufficiently large constant $K$;
(3) $Z \geq 1$.

Here $K$ is some universal constant that we will not explicitly compute; at several places we will assume it is sufficiently large. In Section 7 we will discuss how to satisfy these regularity conditions (or how our algorithm can succeed when they become false).

Consider some almost-clique $C_j$, with $M = |C_j|$ vertices. For any vertex $v \in C_j$, we define $N_j(v) = N(v) \cap C_j$. The dense coloring step operates by selecting a random permutation $\pi$ to order the vertices to be colored. Then, the first $L = \lceil M(1-2\delta\ln(1/\delta))\rceil$ vertices in this ordering select their color $\chi(v)$ uniformly from their palettes.

6.1. **Overview.** We first contrast our dense coloring procedure with a naive one, which assigns each vertex a random color and decolors a vertex if there is a conflict. It is not hard to show that such a procedure successfully assigns a color to a vertex with constant probability. Thus, in each round, the degrees are shrinking by a constant factor in expectation. So it takes $\Omega(\log \Delta)$ rounds to reduce to a low (near-constant) degree.

In order to get a faster running time, we need to color much more than a constant fraction of the vertices. Here, the network decomposition plays the decisive role as only *external* neighbors of a vertex $v$ can decolor $v$. To illustrate, suppose that each vertex $v$ selects a color from its palette uniformly at random. (That is, suppose we ignore the interaction between $v$ and the other vertices in $C_j$). Since the external neighbors are upper bounded by $D$ and the palette size is at least $Z$, even if the external neighbors of $v$ choose distinct colors, the probability that $v$ has any conflicts with its neighbors is upper bounded by $D/Z = \delta = O(\epsilon)$. Ideally, we would like to show that each cluster shrinks by a factor of $\delta$ in each round. Moreover, one would also need to prove that the ratio $D'/Z'$ in the next round remains approximately $\delta$, so that the almost-cliques continue to shrink by the same factor.

In our coloring procedure, a vertex does not really get a color from its original palette uniformly at random, but close to uniform. For example, if $v$ comes late in the ordering of $\pi$, then its palette may have been reduced to just a single color. We prove that any vertex $v$ has, in expectation, a large palette to draw from at the time when it is colored. Usually, the vertex $v$ comes roughly in the middle of the permutation, and its effective palette is still relatively large. In expectation, the probability a vertex remains uncolored is $O(\delta \ln(1/\delta))$.

The reason why we only attempt to color the first $L$ vertices rather than the entire almost-clique is that we cannot afford the palette size to shrink too fast. A "controlled" uniform shrinking process maintains the overall ratio between palette size, external neighbors, and internal neighbors. This prevents us to go into hard case scenarios.

While it is not difficult to bound the expected palette size for a single vertex, this is not sufficient for the proof. We will use a concentration inequality, which boils down to showing that multiple vertices *simultaneously* have a large effective palette size (Lemma 6.4). This does not follow from standard concentration arguments based on bounded differences such as Azuma's inequality; the reason for this is that changing the color of a single vertex could cause of cascade of color changes in other vertices. Instead, we will use a novel argument which estimates the palette size of a vertex in terms of certain rank statistics of the random permutations, and then we will show that these rank statistics are independent. In particular, the probability that a vertex $v_i$ selects a certain color is $\frac{1}{Q(v_i)-\mathrm{Rank}(v_i,N_j(v_i))}$, where $\mathrm{Rank}(v_i, N_j(v_i))$ is the rank of $v_i$ among its neighbors in $C_j$. $\mathrm{Rank}(v_i, N_j(v_i))$ is a random variable that may depend on $\mathrm{Rank}(v_{i'}, N_j(v_{i'}))$ for some other $i' \neq i$. To achieve the independence among vertices in $C_j$, we introduce global ranks, $R_i$, defined in Lemma 6.2. Our upper bound on the anti-degree, $a(v)$, guarantees that we do not lose too much when using the global rank to approximate the local rank.

One useful lemma will be the following, which uses the regularity conditions to show bounds on the relative magnitudes of several parameters of the almost-clique.

**Lemma 6.1.** *Suppose the regularity conditions are satisfied. Then*

$$Q(v_i) - L \geq Q(v_i)\delta \ln(1/\delta) + D \geq D(\ln(1/\delta) + 1)$$

*Proof.* First since all the vertices in the almost-clique must be a neighbor of $v_i$ or contribute to $v_i$'s anti-degree, we must have $M = \deg(v_i) + 1 + a(v_i) \leq Q(v_i) + a(v_i) \leq Q(v_i) + D$. Therefore,

$$\begin{aligned} Q(v_i) - L &\geq Q(v_i) - \lceil M(1 - 2\delta \ln(1/\delta))\rceil \\ &\geq Q(v_i) - M(1 - 2\delta \ln(1/\delta)) - 1 \\ &\geq Q(v_i) - (Q(v_i) + D)(1 - 2\delta \ln(1/\delta)) - 1 \end{aligned}$$

$$\begin{aligned}
&\geq 2Q(v_i)\delta\ln(1/\delta) - D + 2D\delta\ln(1/\delta) - 1 \\
&\geq 2Q(v_i)\delta\ln(1/\delta) - D \qquad \text{as } D\delta \geq K\ln n \text{ and } (1/\delta) \geq K \text{ by regularity conditions} \\
&\geq 2Q(v_i)\delta\ln(1/\delta) - \delta Q(v_i) \qquad \text{as } D = \delta Z \text{ by definition and } Z \leq Q(v_i) \\
&= Q(v_i)(2\delta\ln(1/\delta) - \delta) \\
&\geq Q(v_i)(\delta\ln(1/\delta) + \delta) \qquad \text{for } K \text{ sufficiently large} \\
&\geq Q(v_i)\delta\ln(1/\delta) + D \qquad \text{as } \mathrm{Pal}(v_i) \geq Z \text{ and } D = \delta Z
\end{aligned}$$

To show the second inequality in the lemma, observe that $Q(v_i)\delta\ln(1/\delta) \geq Z\delta\ln(1/\delta) = D\ln(1/\delta)$. □

6.2. **Rank statistics of the random permutation.** For any vertex $v \in C_j$ and any subset of vertices $X \subseteq C_j$, we define $\mathrm{Rank}(v, X)$ to be the number of vertices $x \in X$ such that $\pi^{-1}(x) < \pi^{-1}(v)$; that is, such that $x$ comes before $v$ in the ordering $\pi$ (and hence such that $x$ will choose its color before $v$).

**Lemma 6.2.** *Let $v_1, \ldots, v_M$ be an* arbitrary *ordering of the vertices of $C_j$.*

*Let $R_i = Rank(v_i, \{v_1, \ldots, v_{i-1}\})$. The random variables $R_1, \ldots, R_M$ are independent; each $R_i$ is a uniform discrete random variable in the set $[i] = \{0, 1, \ldots, i-1\}$.*

*Proof.* Note that each $R_i$ can only take values in $[i]$. So, the total number of possibilities for $\langle R_1, \ldots, R_M\rangle$ is $1 \cdot 2 \cdot \cdots \cdot M = M!$

Next, note that $\langle R_1, \ldots, R_M\rangle$ can be determined uniquely from $\pi$, and there are exactly $M!$ possibilities for $\pi$. As the mapping from $\langle R_1, \ldots, R_M\rangle$ to $\pi$ is injective and the two spaces have equal cardinality, it must be a bijective mapping as well. That is, $\langle R_1, \ldots, R_M\rangle$ must be uniformly selected from $[1] \cdot [2] \cdot \cdots \cdot [M]$. This implies that each $R_i$ is independently chosen from $[i]$. □

**Lemma 6.3.** *Let $T = \{v_1, \ldots, v_t\} \subseteq C_j$ and let $\chi(v_i)$ denote the color assigned to $v_i$. Let $c_1, \ldots, c_t$ be an arbitrary sequence of colors. Then, conditioned on the random ordering $\pi$, we have*

$$P(\chi(v_1) = c_1 \wedge \cdots \wedge \chi(v_t) = c_t \mid \pi) \leq \prod_{i=1}^{t} \frac{[\pi^{-1}(v_i) \leq L]}{Q(v_i) - \mathrm{Rank}(v_i, N_j(v_i))}$$

*(Recall that $[\pi^{-1}(v_i) \leq L]$ is the Iverson notation.)*

*Proof.* Suppose without loss of generality that $\pi^{-1}(v_1) < \pi^{-1}(v_2) < \cdots < \pi^{-1}(v_t)$. Now, the coloring procedure will color $v_1, v_2, \ldots, v_t$ in that order (as well as coloring some additional vertices in between.)

Now suppose we come to $v_i$, and we select a color remaining from the palette of $v_i$. If $\pi^{-1}(v_i) > L$, then vertex $v_i$ will not receive any color. Otherwise, at this point the vertices in the neighborhood of $v_i$ which appear earlier in $\pi$ have already chosen their color. There now remain at least $Q(v_i) - \mathrm{Rank}(v_i, N_j(v_i))$ colors. As we choose any color uniformly at random, the probability of choosing a particular color $c_i$ is at most $\frac{1}{Q(v_i) - \mathrm{Rank}(v_i, N_j(v_i))}$. □

**Lemma 6.4.** *Suppose the regularity conditions are satisfied. Let $T = \{v_1, \ldots, v_t\} \subseteq C_j$ and let $\chi(v_i)$ denote the color assigned to $v_i$. Let $c_1, \ldots, c_t$ be an arbitrary sequence of colors. Suppose $t \leq D$. Then we have*

$$P(\chi(v_1) = c_1 \wedge \cdots \wedge \chi(v_t) = c_t) \leq (2\ln(Z/D)/Z)^t$$

*Proof.* Define the event $\mathcal{E} = (\chi(v_1) = c_1 \wedge \cdots \wedge \chi(v_t) = c_t)$. If $t > L$, this statement is vacuously true as $\mathcal{E}$ has probability zero. So, we assume $t \leq L$ for the remainder of the proof.

We break the probabilistic process into two parts. First, we select $\pi$, and then the colors for the first $L$ vertices in the ordering $\pi$. For each $\pi$, we apply Lemma 6.3. Integrating over $\pi$ then gives

us the bound

$$P(\mathcal{E}) \leq \mathbf{E}_\pi\Big[\prod_{i=1}^{t} \frac{[\pi^{-1}(v_i) \leq L]}{Q(v_i) - \mathrm{Rank}(v_i, N_j(v_i))}\Big]$$

For each $i = 1, \dots, t$, define $S_i = C_j - \{v_i, \dots, v_t\}$, and define $R_i = \mathrm{Rank}(v_i, S_i)$. Now consider the ordering $\rho$ of the vertices in which all the vertices in $C_j - T$ come first (in an arbitrary order) followed by $v_1, \dots, v_t$ in that order. Applying Lemma 6.2 to $\rho$, we see that each $R_i$ is independent and uniform on the range $\{0, \dots, M - t + i - 1\}$.

We claim that if $\mathcal{E}$ occurs, then we must have $R_i < L - (t - i)$ for all $i = 1, \dots, t$. Suppose that we have $R_i \geq L - (t - i)$ for some $i = 1, \dots, t$. Then $\mathcal{E}$ is impossible, since there are $L - (t - i)$ vertices in $S_i$ which all come before $v_i$, and the $t - i + 1$ vertices $v_i, \dots, v_t$ must come before $L$ as well. Hence, we may assume for the remainder of this proof that $R_i < L - (t - i)$ for all $i = 1, \dots, t$.

We will lower-bound the denominator, $Q(v_i) - \mathrm{Rank}(v_i, N_j(v_i))$. We decompose the term $\mathrm{Rank}(v_i, N_j(v_i))$ as:

$$\begin{aligned}\mathrm{Rank}(v_i, N_j(v_i)) &= \mathrm{Rank}(v_i, N_j(v_i) \cap S_i) + \mathrm{Rank}(v_i, N_j(v_i) \cap \{v_{i+1}, \dots, v_t\}) \\ &\leq \mathrm{Rank}(v_i, S_i) + (t - i) = R_i + (t - i)\end{aligned}$$

Thus,

$$Q(v_i) - \mathrm{Rank}(v_i, N_j(v_i)) \geq Q(v_i) - (t - i) - R_i \tag{2}$$

We claim that the RHS of (2) is positive. For, when $R_i < L - (t - i)$ then $Q(v_i) - (t - i) - R_i \geq Q(v_i) - L$, and by Lemma 6.1 this is at least $D(\ln(1/\delta) + 1) > 0$.

Putting this together with our bound $R_i < L - (t - i)$, we have that

$$P(\mathcal{E}) \leq \mathbf{E}_\pi\Big[\prod_{i=1}^{t} \frac{[R_i < L - (t - i)]}{Q(v_i) - (t - i) - R_i}\Big] \tag{3}$$

Note that the $i^{\text{th}}$ term in (3) depends only on the random variable $R_i$. As these are all independent, we have

$$P(\mathcal{E}) \leq \prod_{i=1}^{t} \mathbf{E}\Big[\frac{[R_i < L - (t - i)]}{Q(v_i) - (t - i) - R_i}\Big]$$

Let us fix $i$ and compute the corresponding expectation. As $R_i$ is uniform in the range $\{0, \dots, M - t + i - 1\}$ we have

$$\begin{aligned}\mathbf{E}\Big[\frac{[R_i < L - (t - i)]}{Q(v_i) - (t - i) - R_i}\Big] &= \frac{1}{M - (t - i)} \sum_{r=0}^{L-t+i-1} \frac{1}{Q(v_i) - (t - i) - r} \\ &\leq \frac{1}{M - t + i}\Big(\ln \frac{Q(v_i) - t + i}{Q(v_i) - L + 1}\Big)\end{aligned}$$

Now there are two cases depending on the size of $M$. First, suppose that $M \leq \frac{3}{4}Q(v_i)$. Then

$$\begin{aligned}
\mathbf{E}\Big[\frac{[R_i < L-(t-i)]}{Q(v_i)-(t-i)-R_i}\Big] &\leq \frac{1}{M-t+i}\Big(\ln\frac{Q(v_i)-t+i}{Q(v_i)-L+1}\Big) \\
&\leq \frac{1}{M-t+i}\Big(\frac{L-t+i-1}{Q(v_i)-L+1}\Big) \qquad \ln(1+x)\leq x \text{ for } x\geq 0 \\
&\leq \frac{1}{Q(v_i)-L} \qquad \text{as } L\leq M \\
&\leq \frac{1}{Q(v_i)-(3/4)\cdot Q(v_i)} \qquad \text{as } L\leq M\leq \frac{3}{4}Q(v_i) \\
&\leq \frac{4}{Z} \qquad \text{as } Q(v_i)\geq Z \\
&\leq 2\ln(1/\delta)/Z \qquad \text{for } \delta \text{ sufficiently small (and hence for } K \text{ sufficiently large)}
\end{aligned}$$

Otherwise, if $M > \frac{3}{4}Q(v_i)$

$$\begin{aligned}
\mathbf{E}\left[\frac{[R_i < L-(t-i)]}{Q(v_i)-(t-i)-R_i}\right] &\leq \frac{1}{M-t+i}\left(\ln\frac{Q(v_i)-t+i}{Q(v_i)-L+1}\right) \\
&\leq \frac{1}{3Z/4-D}\left(\ln\frac{Q(v_i)}{Q(v_i)-L+1}\right) \qquad \text{as } t-i\in[0,D], M\geq 3Z/4 \\
&\leq \frac{1}{3Z/4-D}\ln\left(\frac{Q(v_i)}{Q(v_i)\delta\ln(1/\delta)+D}\right) \qquad \text{by Lemma 6.1} \\
&\leq \frac{1}{Z(3/4-\delta)}\ln\left(\frac{1}{\delta\ln(1/\delta)}\right) \\
&\leq \frac{2}{Z}\ln(1/\delta) \qquad \text{as } \frac{3}{4}-\delta\geq 1/2 \text{ and } \ln(1/\delta)\geq 1 \text{ for } \delta \text{ sufficiently small}
\end{aligned}$$

Putting this together, we have

$$\mathbf{P}(\mathcal{E}) \leq \prod_{i=1}^{t}\mathbf{E}\Big[\frac{[R_i < L-(t-i)]}{Q(v_i)-(t-i)-R_i}\Big] \leq \left(\frac{2\ln(1/\delta)}{Z}\right)^t$$

□

6.3. **Concentration for the number of uncolored vertices.** We will now show that most vertices become colored at the end of this process. We distinguish two ways in which a vertex $v$ could fail to be colored: first, it may be *decolored* in the sense that it initially chose a color, but then had a conflict with an almost-clique of earlier index. Second, it may be *initially-uncolored* in the sense that $\pi^{-1}(v) > L$.

**Lemma 6.5.** *Suppose the regularity conditions are satisfied. Let $T \subseteq V^{dense}$ and $|T| \leq D$. The probability that all the vertices in $T$ become* decolored *is at most*

$$P(\textit{all vertices in } T \textit{ are decolored}) \leq (2\delta\ln(1/\delta))^{|T|}$$

*Proof.* Write $T_j = T \cap C_j$. For $j = 1,\ldots,n$ we compute the probability that the vertices in $T_j$ become decolored, conditioned on the event that the vertices in $T_1,\ldots,T_{j-1}$ become decolored. In fact, we will not just condition on the event that the vertices in $T_1,\ldots,T_{j-1}$ become decolored, but we will condition on the complete set of random variables involved in $C_1,\ldots,C_{j-1}$. (Observe that the event that $T_j$ becomes decolored is a function of only the colors involved in $C_1,\ldots,C_j$.) We claim that, conditioned on all such random variables, the event that $T_j$ becomes decolored has probability at most $(2\delta\ln(1/\delta))^{|T_j|}$.

For each $v \in T_j$, the event that $v$ becomes decolored is a union of at most $\bar{d}(v) \leq D$ events of the form $\chi(v) = c$, where $c$ enumerates the colors of the neighbors of $v$ in earlier almost-cliques. Hence, the event that all of the vertices in $T_j$ become decolored is a union of $D^{|T_j|}$ events of the form stated in Lemma 6.4, each of which has probability at most $(2\ln(Z/D)/Z)^{|T_j|}$. Therefore, the probability that all of them become decolored is $(2(D/Z)\ln(Z/D))^{|T_j|} = (2\delta\ln(1/\delta))^{|T_j|}$.

Now, multiplying all such probabilities, we get that the total probability that $T$ is decolored is at most $(2\delta\ln(1/\delta))^{|T|}$. □

**Lemma 6.6.** *Let $T \subseteq V^{dense}$. The probability that all of the vertices in $T$ are initially-uncolored is at most $(2\delta\ln(1/\delta))^{|T|}$.*

*Proof.* It suffices to show that for a particular $C_j$, the probability that all vertices in $T_j = T \cap C_j$ are initially-uncolored is bounded by $(2\delta\ln(1/\delta))^{|T_j|}$, since there are no dependencies between the almost-cliques.

Let $M = |C_j|$ and $L = \lceil M(1-2\delta\ln(1/\delta))\rceil$. We select a set of $L$ vertices to be colored, uniformly without replacement from $C_j$. Thus, the probability that all vertices in $T_j$ are uncolored is:

$$\begin{aligned}
P(\text{vertices in } T_j \text{ are all uncolored}) &= \frac{\binom{M-|T_j|}{L}}{\binom{M}{L}} = \frac{M-L}{M} \cdot \dots \cdot \frac{M-L-(|T_j|-1)}{M-(|T_j|-1)} \\
&\leq \left(\frac{M-L}{M}\right)^{|T_j|} \\
&= \left(\frac{M-\lceil M(1-2\delta\ln(1/\delta))\rceil}{M}\right)^{|T_j|} \\
&\leq \left(\frac{M-M(1-2\delta\ln(1/\delta))}{M}\right)^{|T_j|} \\
&= (2\delta\ln(1/\delta))^{|T_j|}
\end{aligned}$$

□

**Lemma 6.7.** *Suppose the regularity conditions are satisfied. Let $T \subseteq V^{dense}$ with $|T| \leq D$. The probability that, at the end of round $i$, $T$ contains more than $12\delta\ln(1/\delta)D$ uncolored vertices, is at most $n^{-100}$.*

*Proof.* We separately show concentration bounds on the number of decolored and initially-uncolored vertices in $T$. Let $x = 6\delta\ln(1/\delta)D$, we claim that each quantity has an $1/n^{\Omega(1)}$ probability of exceeding $x$, and this shows the claim.

By union bound over all possible sets of size $x$, the probability that a quantity exceeds $x$ is bounded by

$$\begin{aligned}
\binom{|T|}{x}(2\delta\ln(1/\delta))^x &\leq \left(\frac{e|T|}{x}\right)^x \cdot (2\delta\ln(1/\delta))^x \\
&\leq \left(\frac{eD(2\delta\ln(1/\delta)}{6D\delta\ln(1/\delta)}\right)^x \qquad \text{as } x = 6D\delta\ln(1/\delta) \text{ and } |T| \leq D \\
&\leq \left(\frac{2e}{6}\right)^{K\ln n} \qquad \text{as } x \geq D\delta \geq K\ln n
\end{aligned}$$

This is $< n^{-100}$ for $K$ being a sufficiently large constant. As the total number of initially-uncolored vertices and decolored vertices are separately bounded by $x$, it follows that their sum (the number of uncolored vertices of any kind) is at most $2x$. □

**Lemma 6.8.** *Suppose the regularity conditions are satisfied at the beginning of round $i$. Then whp at the end of round $i$, all the vertices satisfy the bounds*

$$a_i(v) \le D', \quad \overline{d}_i(v) \le D', \quad Q_i(v) \ge Z'$$

*for the parameters*

$$D' = 12D\delta \ln(1/\delta) \qquad Z' = D\ln(1/\delta)$$

*Proof.* By Lemma 6.7 with $T$ being the set of external neighbors of $v$, we have that $\overline{d}_i(v) \le D'$ holds with probability $\ge 1 - n^{-100}$. Similarly, by Lemma 6.7 with $T = C_j \setminus N(v)$, we have that $a_i(v) \le D'$ holds with probability $\ge 1 - n^{-100}$. Thus by taking the union bound of both events over each vertex, the probability any of them fails is at most $2n \cdot n^{-100} = n^{-\Omega(1)}$.

Finally, we bound $Q_i(v)$. We color at most $D$ external neighbors and at most $L$ internal neighbors. Thus, at the end, its palette has size at least $Q_i(v) - L - D$. By Lemma 6.1, this is at least $D\ln(1/\delta)$ for $K$ sufficiently large. □

We next show how to bound the size of an almost-clique.

**Lemma 6.9.** *Suppose that at the beginning of round $i$, each almost-clique has size $M = |C_j|$ and that the regularity conditions are satisfied. Then, whp at the end of round $i$ all almost-cliques have size at most*

$$M'_j \le \max(2K\ln n, 12M\delta\ln(1/\delta))$$

*Proof.* Let us consider some fixed almost-clique $C_j$. A vertex in $C_j$ survives only if it is initially uncolored, or decolored. There are exactly $M - L \le 2M\delta\ln(1/\delta)$ initially-uncolored vertices at the end of round $i$, so we only need to bound the number of decolored vertices.

Let $w = K\ln n$ and let $x = \max(w, 10M\delta\ln(1/\delta))$. Note that $w \le D$, since $D \ge K\ln n$. If component $C_j$ contains more than $x$ decolored vertices, then $C_j$ has at least $\binom{x}{w}$ sets of $w$-tuples of decolored vertices. On the other hand, by Lemma 6.5, each $w$-tuple of vertices in $C_j$ is decolored with probability at most $(2\delta\ln(1/\delta))^w$ (as $w \le D$). Thus, the expected number of $w$-tuples of decolored vertices is at most $\binom{M}{w}(2\delta\ln(1/\delta))^w$. By Markov's inequality, the probability that there are more than $x$ decolored vertices is bounded by

$$\begin{aligned}
P(C_j \text{ has} > x \text{ decolored vertices}) &\le \frac{\binom{M}{w}(2\delta\ln(1/\delta))^w}{\binom{x}{w}} \\
&\le \frac{(\frac{eM}{w})^w(2\delta\ln(1/\delta))^w}{(\frac{x}{w})^w} \qquad \text{as } \left(\frac{n}{k}\right)^k \le \binom{n}{k} \le \left(\frac{en}{k}\right)^k \\
&\le \left(\frac{eM}{x}\right)^w (2\delta\ln(1/\delta))^w \\
&\le \left(\frac{2eM\delta\ln(1/\delta)}{10M\delta\ln(1/\delta)}\right)^{K\ln n} \\
&\le n^{-100} \qquad \text{for } K \text{ sufficiently large}
\end{aligned}$$

By taking a union bound over all almost-cliques, we see that whp every almost-clique has at most $x$ decolored vertices. Thus, whp, each almost-clique has at most $2M\delta\ln(1/\delta) + x \le \max(2K\ln n, 12M\delta\ln(1/\delta))$ uncolored vertices. □

## 7. Solving the recurrence

In light of Lemma 6.8, we can explicitly derive a recurrence relation for the parameters $D_i, Z_i$.

**Lemma 7.1.** *Define the recurrence relation with initial conditions*

$$D_0 = 3\epsilon\Delta \qquad Z_0 = \Delta/2$$

*and recurrence*

$$D_{i+1} = 12 D_i \delta_i \ln(1/\delta_i) \qquad Z_{i+1} = D_i \ln(1/\delta_i)$$

*where we define* $\delta_i = D_i / Z_i$.

*Let* $i \leq \lceil \sqrt{\ln \Delta} \rceil$. *Assuming that the regularity conditions* $\delta_j < 1/K, D_j > K \ln n, Z \geq 1$ *are satisfied for* $j = 0, \ldots, i-1$, *we have that whp*

$$a_i(v) \leq D_i, \overline{d}_i(v) \leq D_i, Q_i(v) \geq Z_i$$

*Proof.* The bound on $Z_0$ is given in Lemma 5.7. By Lemma 3.7, 3.9, we have $a(v) \leq 3\epsilon\Delta, \overline{d}(v) \leq 3\epsilon\Delta$ in the initial graph. The initial coloring step cannot increase these parameters, so we have $a_0(v) \leq 3\epsilon\Delta, \overline{d}_0(v) \leq 3\epsilon\Delta$ as well; this shows the bound on $D_0$.

A simple induction, using Lemma 6.8, shows that for all $i = 1, \ldots, n$ we have the following:

$$a_i(v) \leq D_i, \overline{d}_i(v) \leq D_i, Q_i(v) \geq Z_i \qquad \text{with probability} \geq 1 - in^{-\Omega(1)}$$

Thus, for any fixed $i \leq \lceil \sqrt{\ln \Delta} \rceil$, the probability that this fails to hold is at most $(1+\sqrt{\ln \Delta}) n^{-\Omega(1)} = n^{-\Omega(1)}$. □

We will now show how to solve this recurrence.

**Lemma 7.2.** *Recall that we set* $\epsilon = C \cdot 100^{-\lceil \sqrt{\ln \Delta} \rceil}$. *We can choose the constant term* $C$ *sufficiently small so that for all* $i \leq \lceil \sqrt{\ln \Delta} \rceil$ *we have*

$$\delta_i = 6\epsilon \cdot 12^i < 1/K$$

*Proof.* For each $i > 0$, we may compute $\delta_i$ as

$$\delta_i = \frac{D_i}{Z_i} = \frac{12 D_{i-1} \delta_{i-1} \ln(1/\delta_{i-1})}{D_{i-1} \ln(1/\delta_{i-1})} = 12 \delta_{i-1}$$

As $\delta_0 = D_0 / Z_0 = 6\epsilon$, we have $\delta_i = 6\epsilon \cdot 12^i$ as claimed.

Thus, we have $\delta_i \leq 6\epsilon 12^{\lceil \sqrt{\ln \Delta} \rceil} = 6C \cdot 100^{-\lceil \sqrt{\ln \Delta} \rceil} 12^{\lceil \sqrt{\ln \Delta} \rceil} \leq 6C$. By selecting $C$ sufficiently small, we can ensure that this is at most $1/K$ as desired.

□

**Lemma 7.3.** *For all* $5 \leq i \leq \lceil \sqrt{\ln \Delta} \rceil$, *we have*

$$D_i \leq 12^{i^2/2} \cdot 100^{-i \lceil \sqrt{\ln \Delta} \rceil / 2} \cdot \Delta$$

*Proof.* We can recursively compute $D_i$ from $D_0$ as:

$$D_i = D_0 \cdot \prod_{j=0}^{i-1} 12 \delta_j \ln(1/\delta_j) \tag{4}$$

Thus we can estimate:

$$\begin{aligned}
D_i &\leq 3\epsilon\Delta \prod_{j=0}^{i-1} \left( 12 \delta_j^{1/2} \right) \qquad \ln x \leq x^{1/2} \text{ for } x > 0 \\
&\leq 3\epsilon\Delta \prod_{j=0}^{i-1} \left( 12 (6\epsilon \cdot 12^j)^{1/2} \right) \qquad \text{by Lemma 7.2} \\
&= (3\epsilon\Delta) \cdot \left( 12^i (6\epsilon)^{i/2} \cdot 12^{i(i-1)/4} \right) \\
&\leq \Delta \cdot \left( 12^i (6 \cdot C \cdot 100^{-\lceil \sqrt{\ln \Delta} \rceil})^{i/2} \cdot 12^{i(i-1)/4} \right) \\
&\leq \Delta \cdot 12^{i(i+5)/4} \cdot 100^{-i \lceil \sqrt{\ln \Delta} \rceil / 2} \qquad 6C \leq 12 \text{ for } C \leq 2 \\
&\leq \Delta \cdot 12^{i^2/2} \cdot 100^{-i \lceil \sqrt{\ln \Delta} \rceil / 2} \qquad (i+5)/4 \leq i/2 \text{ for } i \geq 5
\end{aligned}$$

□

**Corollary 7.4.** *We have the bound* $D_{\lceil\sqrt{\ln\Delta}\rceil} = O(1)$.

*Proof.* We apply Lemma 7.3:

$$\begin{aligned} D_{\lceil\sqrt{\ln\Delta}\rceil} &\le \Delta \cdot 12^{\lceil\sqrt{\ln\Delta}\rceil^2/2} \cdot 100^{-\lceil\sqrt{\ln\Delta}\rceil^2/2} \\ &\le \Delta \cdot \left(\frac{12}{100}\right)^{(\ln\Delta)/2} \le \Delta \cdot e^{-\ln\Delta} = 1 \end{aligned}$$

□

**Expository remark:** Corollary 7.4 explains why we selected $\epsilon = \exp(-\Theta(\sqrt{\log\Delta}))$ and ran our coloring steps for $O(\sqrt{\log\Delta})$ rounds. Suppose instead we set $\epsilon = \exp(-\ln^a\Delta)$ and ran $\ln^a\Delta$ dense coloring steps, for some $a < 1/2$. At the end of these steps, we would have $D_{\ln^a\Delta} = \Delta\exp(-\ln^{2a}\Delta)$. This is close to $\Delta$ (differing in only a sub-polynomial term), which implies that we have hardly made any progress in reducing the number of uncolored vertices.

**Theorem 7.5.** *At the end of the dense coloring steps, whp every dense vertex is connected to at most* $O(\log n)\cdot 2^{O(\sqrt{\log\Delta})}$ *other dense vertices.*

*Proof.* By Lemma 7.2, we have $\delta_i < 1/K$ for all dense coloring rounds. Let $i^* \le \lceil\sqrt{\ln\Delta}\rceil$ be minimal such that $D_{i^*}\delta_{i^*} \le K\ln n$; by Corollary 7.4 such an $i^*$ exists. Also, observe that for each $i \le i^*$ we have

$$Z_i = D_{i-1}\ln(1/\delta_{i-1}) \ge D_{i-1}\delta_{i-1} \ge K\ln n \ge 1$$

So the regularity conditions are satisfied up to round $i^*$, and hence by Lemma 7.1, we have at the end of round $i^*$:

$$\overline{d}_{i^*}(v) \le D_{i^*} \le (K\ln n)/\delta_{i^*} \le (K\ln n)\cdot 2^{O(\sqrt{\log\Delta})} \qquad \text{as } \delta_{i^*} \ge \epsilon \ge 100^{-\sqrt{\log\Delta}} \tag{5}$$

This shows that the external degree of each dense vertex is at most $O(\log n)\cdot 2^{O(\sqrt{\log\Delta})}$ at the end of the dense coloring steps.

Next, we bound the size of each almost-clique. As we satisfy the conditions $\delta_j < 1/K, D_j\delta_j \ge K\ln n$ for $j < i^*$, we can apply Lemma 6.9 repeatedly to deduce that the size of any almost-clique has been reduced from its initial size (at most $(1+3\epsilon)\Delta$) to $O(\max(2K\ln n, (1+3\epsilon)\Delta\cdot(\prod_{j=0}^{i^*-1} 12\delta_j\ln(1/\delta_j))))$; the probability this fails in an individual round is $n^{-\Omega(1)}$ and thus the total failure probability over each clique and over $i^* \le \lceil\sqrt{\ln\Delta}\rceil$ rounds is also $n^{-\Omega(1)}$.

Thus, we need to bound the term $(1+3\epsilon)\Delta\cdot(\prod_{j=0}^{i^*-1} 12\delta_j\ln(1/\delta_j))$ which we do as follows:

$$\begin{aligned} (1+3\epsilon)\Delta\cdot\prod_{j=0}^{i^*-1}(12\delta_j\ln(1/\delta_j)) &= (1+3\epsilon)\Delta\cdot\frac{1}{D_0}\left(\prod_{j=0}^{i^*-1}12\delta_j\ln(1/\delta_j)\right)\cdot D_0 \\ &= (1+3\epsilon)\Delta\cdot\frac{D_{i*}}{3\epsilon\Delta} \qquad \text{by (4) and } D_0 = 3\epsilon\Delta \\ &\le \frac{(1+3\epsilon)K\ln n}{3\epsilon}\cdot 2^{O(\sqrt{\log\Delta})} \qquad \text{by (5)} \\ &= O(\log n)\cdot 2^{O(\sqrt{\log\Delta})} \qquad \text{as } \epsilon = C\cdot 100^{-\lceil\sqrt{\log\Delta}\rceil} \end{aligned}$$

Since a dense vertex $v$ has at most $O(\log n)\cdot 2^{O(\sqrt{\log\Delta})}$ external neighbors and the clique size is also bounded by $O(\log n)\cdot 2^{O(\sqrt{\log\Delta})}$, it can have $O(\log n)\cdot 2^{O(\sqrt{\log\Delta})}$ neighbors.

□

We have shown that after the $\lceil\sqrt{\ln\Delta}\rceil$ dense coloring steps, the number of dense neighbors of each dense vertex shrinks to $O(\log n)\cdot 2^{O(\sqrt{\log\Delta})}$. Also, for each sparse vertex $x$, we have $Q_{\lceil\sqrt{\ln\Delta}\rceil}(x) \ge \deg_{\lceil\sqrt{\ln\Delta}\rceil}(x) + \Omega(\epsilon^2\Delta)$ due to the initial coloring step. By applying the algorithm of

Elkin, Pettie, and Su [12, Section 4] on the sparse component, it can be colored in $O(\log(1/\epsilon)) + 2^{O(\sqrt{\log\log n})} = O(\sqrt{\log \Delta}) + 2^{O(\sqrt{\log\log n})}$ rounds. Then, we apply the algorithm of Barenboim et al. [9] to color the remaining vertices whose degree are bounded by $\Delta' = O(\log n) \cdot 2^{O(\sqrt{\log \Delta})}$. It then runs in $O(\log \Delta') + 2^{O(\sqrt{\log\log n})} = O(\sqrt{\log \Delta}) + 2^{O(\sqrt{\log\log n})}$ rounds. The total number of rounds is $O(\sqrt{\log \Delta}) + 2^{O(\sqrt{\log\log n})}$.

## 8. List-coloring locally-sparse graphs

Although the overall focus of this paper is an algorithm for coloring arbitrary graphs in time $O(\sqrt{\log \Delta}) + 2^{O(\sqrt{\log\log n})}$, we note that our initial coloring step may also be used to obtain a faster list-coloring algorithm for graphs which are sparse. This result extends the work of [12], which showed a similar type of $(\Delta+1)$-coloring algorithm for sparse graphs.

In [12], a slightly different definition of sparsity was introduced, known as *local sparsity.* We define this and show that it is essentially equivalent to the definition of sparsity defined in Section 3.

**Definition 8.1.** *We say that a graph $G$ is* $(1-\delta)$ locally sparse *if very vertex contains at most $(1-\delta)\binom{\Delta}{2}$ edges in its neighborhood, for some parameter $\delta \in [0,1]$. (That is, the induced subgraph $G[N(v)]$ contains $\leq (1-\delta)\binom{\Delta}{2}$ edges).*

**Lemma 8.2.** *Suppose that $G$ is $(1-\delta)$-locally sparse. Then if we apply the network decomposition of Section 3 with parameter $\epsilon = \delta/2$, then every vertex is sparse, i.e. $V^{sparse} = V$.*

*Proof.* Suppose that $v \in V$ is dense with respect to $\epsilon$. Then $v$ has at least $(1-\epsilon)\Delta$ friends. Each such friend $u$ corresponds to at least $(1-\epsilon)\Delta$ edges between $u$ and another neighbor of $v$, that is, an edge in $G[N(v)]$. Furthermore, any such edge in $G[N(v)]$ is counted at most twice, so $G[N(v)]$ must contain $(1-\epsilon)^2\Delta^2/2 \geq (1-2\epsilon)\binom{\Delta}{2}$ edges, which contradicts our hypothesis for $\epsilon \geq \delta/2$. □

**Corollary 8.3.** *Suppose that $G$ is $(1-\delta)$-locally-sparse and that every vertex has a palette of size exactly $\Delta+1$. Then $G$ can be list-colored whp in $O(\log(1/\delta)) + 2^{O(\sqrt{\log\log n})}$ rounds.*

*Proof.* By Proposition 8.2, every vertex in $G$ is sparse with respect to parameter $\epsilon = \delta/2$. Suppose that $\delta^4\Delta \geq K \ln n$, where $K$ is a sufficiently large constant. Then, by Lemma 5.8, each vertex satisfies at the end of the initial coloring step $d_0(v) \geq Q_0(v) + \Omega(\epsilon^2\Delta)$ whp. Now, apply the algorithm of [12] to the residual graph; this runs in $O(\log(1/\delta^2)) + 2^{O(\sqrt{\log\log n})}$ rounds.

Next, suppose that $\delta^4\Delta \leq K \ln n$. So $\Delta \leq K\delta^{-4}\ln n$. Then run the coloring algorithm of [9], which runs in $O(\log \Delta) + 2^{O(\sqrt{\log\log n})} = O(\log(1/\delta)) + 2^{O(\sqrt{\log\log n})}$ rounds. □

**Acknowledgements:** We would like to thank Seth Pettie for his valuable comments. We thank the anonymous reviewers for many helpful suggestions and comments.

## References

[1] N. Alon, L. Babai, and A. Itai. A fast and simple randomized parallel algorithm for the maximal independent set problem. *J. Algorithms*, 7(4):567–583, 1986.

[2] B. Awerbuch, A. V. Goldberg, M. Luby, and S. A. Plotkin. Network decomposition and locality in distributed computation. In *Symp. on Foundations of Computer Science (FOCS)*, 1989.

[3] R. Bar-Yehuda, K. Censor-Hillel, and G. Schwartzman. A distributed $(2+\epsilon)$-approximation for vertex cover in $O(\log \Delta/\epsilon \log\log \Delta)$ rounds. *arXiv preprint arXiv:1602.03713*, 2016.

[4] L. Barenboim. Deterministic $(\Delta+1)$-coloring in sublinear (in $\Delta$) time in static, dynamic and faulty networks. In *Symp. on Principles of Distributed Computing(PODC)*, pages 345–354, 2015.

[5] L. Barenboim and M. Elkin. Deterministic distributed vertex coloring in polylogarithmic time. *J. ACM*, 58(5):23, 2011.

[6] L. Barenboim and M. Elkin. Distributed graph coloring: Fundamentals and recent developments. *Synthesis Lectures on Distributed Computing Theory*, 4(1):1–171, 2013.

[7] L. Barenboim, M. Elkin, and C. Gavoille. A fast network-decomposition algorithm and its applications to constant-time distributed computation. *CoRR*, abs/1505.05697, 2015.

[8] L. Barenboim, M. Elkin, and F. Kuhn. Distributed $(\Delta+1)$-coloring in linear (in $\Delta$) time. *SIAM Journal on Computing*, 43(1):72–95, 2014.

[9] L. Barenboim, M. Elkin, S. Pettie, and J. Schneider. The locality of distributed symmetry breaking. In *Symp. on Foundations of Computer Science (FOCS)*, pages 321–330. IEEE, 2012.

[10] K.-M. Chung, S. Pettie, and H.-H. Su. Distributed algorithms for Lovász local lemma and graph coloring. In *Proc. 33rd ACM Symposium on Principles of Distributed Computing (PODC)*, pages 134–143, 2014.

[11] D. Dubhashi, D. A. Grable, and A. Panconesi. Near-optimal, distributed edge colouring via the nibble method. *Theor. Comput. Sci.*, 203(2):225–251, Aug. 1998.

[12] M. Elkin, S. Pettie, and H.-H. Su. $(2\Delta-1)$-edge-coloring is much easier than maximal matching in the distributed setting. In *Symp. on Discrete Algorithms (SODA)*, pages 355–370. SIAM, 2015.

[13] P. Fraigniaud, M. Heinrich, and A. Kosowski. Local conflict coloring. *arXiv preprint arXiv:1511.01287*, 2015.

[14] B. Gfeller and E. Vicari. A randomized distributed algorithm for the maximal independent set problem in growth-bounded graphs. In *Symp. on Principles of Distributed Computing*, 2007.

[15] M. Ghaffari. An improved distributed algorithm for maximal independent set. In *Proc. 27th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2016. to appear.

[16] A. V. Goldberg and S. A. Plotkin. Parallel $(\Delta+1)$-coloring of constant-degree graphs. *Information Processing Letters*, 25(4):241 – 245, 1987.

[17] A. V. Goldberg, S. A. Plotkin, and G. E. Shannon. Parallel symmetry-breaking in sparse graphs. *SIAM J. Discrete Math.*, 1(4):434–446, 1988.

[18] M. Göös, J. Hirvonen, and J. Suomela. Linear-in-$\Delta$ lower bounds in the local model. In *Proceedings of the 2014 ACM Symposium on Principles of Distributed Computing(PODC)*, pages 86–95, 2014.

[19] D. A. Grable and A. Panconesi. Nearly optimal distributed edge coloring in $O(\log\log n)$ rounds. *Random Structures & Algorithms*, 10(3):385–405, 1997.

[20] D. A. Grable and A. Panconesi. Fast distributed algorithms for Brooks–Vizing colorings. *Journal of Algorithms*, 37(1):85 – 120, 2000.

[21] J. Hirvonen and J. Suomela. Distributed maximal matching: Greedy is optimal. In *Proceedings of the 2012 ACM Symposium on Principles of Distributed Computing(PODC)*, pages 165–174, 2012.

[22] Ö. Johansson. Simple distributed $\Delta+1$-coloring of graphs. *Inf. Process. Lett.*, 70, 1999.

[23] K. Kothapalli, C. Scheideler, M. Onus, and C. Schindelhauer. Distributed coloring in $\tilde{O}(\sqrt{\log n})$ bit rounds. In *Proc. 20th International Parallel and Distributed Processing Symposium (IPDPS)*, 2006.

[24] F. Kuhn, T. Moscibroda, and R. Wattenhofer. What cannot be computed locally! In *Symp. on Principles of Distributed Computing (PODC)*, 2005.

[25] F. Kuhn, T. Moscibroda, and R. Wattenhofer. Local computation: Lower and upper bounds. *arXiv preprint arXiv:1011.5470*, 2010.

[26] F. Kuhn and R. Wattenhofer. On the complexity of distributed graph coloring. In *Symp. on Principles of Distributed Computing (PODC)*, 2006.

[27] N. Linial. Locality in Distributed Graph Algorithms. *SIAM Journal on Computing*, 21(1):193–201, 1992.

[28] N. Linial and M. Saks. Low diameter graph decompositions. *Combinatorica*, 13(4):441–454, 1993.

[29] M. Luby. A Simple Parallel Algorithm for the Maximal Independent Set Problem. *SIAM Journal on Computing*, 15:1036–1053, 1986.

[30] M. Molloy and B. Reed. A bound on the total chromatic number. *Combinatorica*, 18(2):241–280, 1998.

[31] M. Molloy and B. Reed. *Graph Colouring and the Probabilistic Method*. Algorithms and Combinatorics. Springer, 2001.

[32] M. Molloy and B. Reed. Asymptotically optimal frugal colouring. *J. Comb. Theory, Ser. B*, 100(2):226–246, 2010.

[33] M. Molloy and B. Reed. Colouring graphs when the number of colours is almost the maximum degree. *Journal of Combinatorial Theory, Series B*, 109:134 – 195, 2014.

[34] A. Panconesi and A. Srinivasan. Improved distributed algorithms for coloring and network decomposition problems. In *Symp. on Theory of computing (STOC)*, 1992.

[35] A. Panconesi and A. Srinivasan. Randomized distributed edge coloring via an extension of the Chernoff–Hoeffding bounds. *SIAM Journal on Computing*, 26(2):350–368, 1997.

[36] S. Pettie and H.-H. Su. Distributed coloring algorithms for triangle-free graphs. *Information and Computation*, 243(0):263 – 280, 2015. Preliminary version appeared in *40th Intl. Colloq. on Automata, Languages and Programming (ICALP)*, pages 681–693, 2013.

[37] B. Reed. $\omega$, $\Delta$, and $\chi$. *Journal of Graph Theory*, 27(4):177–212, 1998.

[38] B. Reed. A strengthening of brooks' theorem. *Journal of Combinatorial Theory, Series B*, 76(2):136 – 149, 1999.

[39] J. Schneider, M. Elkin, and R. Wattenhofer. Symmetry breaking depending on the chromatic number or the neighborhood growth. *Theoretical Computer Science*, 509:40–50, 2013.

[40] J. Schneider and R. Wattenhofer. A log-star distributed maximal independent set algorithm for growth-bounded graphs. In *Symp. on Principles of Distributed Computing(PODC)*, 2008.
[41] J. Schneider and R. Wattenhofer. A new technique for distributed symmetry breaking. In *Symp. on Principles of Distributed Computing(PODC)*, 2010.
[42] M. Szegedy and S. Vishwanathan. Locality based graph coloring. In *Proc. 25th ACM Symposium on Theory of Computing (STOC)*, pages 201–207, 1993.