An Optimal Decentralized $(\Delta + 1)$-Coloring Algorithm

Daniel Bertschinger
Department of Computer Science,
ETH Zürich, Switzerland
daniel.bertschinger@inf.ethz.ch

Johannes Lengler
Department of Computer Science,
ETH Zürich, Switzerland
johannes.lengler@inf.ethz.ch

Anders Martinsson
Department of Computer Science,
ETH Zürich, Switzerland
anders.martinsson@inf.ethz.ch

Robert Meier
Department of Computer Science,
ETH Zürich, Switzerland
robert.meier@inf.ethz.ch

Angelika Steger
Department of Computer Science,
ETH Zürich, Switzerland
steger@inf.ethz.ch

Miloš Trujić
Department of Computer Science,
ETH Zürich, Switzerland
mtrujic@inf.ethz.ch

Emo Welzl
Department of Computer Science,
ETH Zürich, Switzerland
emo@inf.ethz.ch

Abstract

Consider the following simple coloring algorithm for a graph on n vertices. Each vertex chooses a color from $\{1, \ldots, \Delta(G) + 1\}$ uniformly at random. While there exists a conflicted vertex choose one such vertex uniformly at random and recolor it with a randomly chosen color. This algorithm was introduced by Bhartia et al. [MOBIHOC'16] for channel selection in WIFI-networks. We show that this algorithm always converges to a proper coloring in expected $O(n \log \Delta)$ steps, which is optimal and proves a conjecture of Chakrabarty and de Supinski [SOSA'20].

2012 ACM Subject Classification Theory of computation → Distributed algorithms; Mathematics of computing → Graph algorithms; Mathematics of computing → Graph coloring; Mathematics of computing → Probabilistic algorithms

Keywords and phrases Decentralized Algorithm, Distributed Computing, Graph Coloring, Randomized Algorithms

Digital Object Identifier 10.4230/LIPIcs.ESA.2020.17

Funding *Miloš Trujić*: author was supported by grant no. 200021 169242 of the Swiss National Science Foundation.

1 Introduction

It is well known that an undirected graph $G = (V, E)$ with maximum degree $\Delta = \Delta(G)$ can be properly colored by using $\Delta + 1$ colors. In fact, a simple greedy algorithm which assigns the colors successively achieves this bound by just touching each vertex once. Note that the bound $\Delta + 1$ is tight, as cliques and odd cycles require this number of colors.

In [1] Bhartia et al. introduced the use of a simple decentralized coloring algorithm as an efficient solution to the channel selection problem in wireless networks. Their algorithm can be formulated as follows.

28th Annual European Symposium on Algorithms (ESA 2020).
Editors: Fabrizio Grandoni, Grzegorz Herman, and Peter Sanders; Article No. 17; pp. 17:1–17:12
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

DECENTRALIZED GRAPH COLORING

For a graph $G = (V, E)$
1. choose for each vertex $v \in V$ a color from $\{1, \ldots, \Delta + 1\}$ independently and uniformly at random;
2. choose a vertex $v \in V$ uniformly at random among all vertices which have a neighbor in the same color;
3. recolor v into a color chosen from $\{1, \ldots, \Delta + 1\}$ uniformly at random;
4. repeat steps 2 and 3 until a proper coloring of G is found.

They showed that this algorithm finds a proper coloring in $O(n\Delta)$ rounds in expectation. Chakrabarty and de Supinski [2] introduced a variant of the coloring algorithm: instead of recoloring a vertex v once as above, in their "Persistent Decentralized Coloring Algorithm" such a vertex v persistently (hence the name) gets recolored until it has no neighbor in the same color. They showed that this modified algorithm only requires $O(n \log \Delta)$ recolorings and conjectured that the same bound also holds for the original algorithm. In this paper we prove their conjecture.

▶ **Theorem 1.1.** *The decentralized coloring algorithm converges in expectation to a proper $(\Delta + 1)$-coloring in $O(n \log \Delta)$ recoloring steps.*

In fact, our argument shows that the same runtime bound holds true if the initial coloring is chosen adversarially. This is in contrast with the persistent version of the algorithm mentioned above, as that one takes $\Theta(n\Delta)$ recolorings in expectation when starting with an adversarial coloring (see [1, Theorem 3]). However, the question raised in [2] of 'which algorithm is faster in the random setting' remains open.

We note that the bound in Theorem 1.1 is best possible, as for the complete graph K_n the decentralized coloring algorithm essentially performs a *Coupon Collector* process. Indeed, once a color (coupon) has been acquired it remains in the graph until the end of the process and we need to see all colors. The claim thus follows from the well known fact that in expectation the coupon collector process with n coupons requires $nH_n = \Theta(n \log n)$ rounds, where $H_k = \sum_{i=1}^{k} \frac{1}{i}$ is the k-th Harmonic number. Moreover, the result is tight for every combination of n and Δ. Namely, consider a vertex-disjoint union of n/Δ complete graphs K_Δ and by the same argument the process requires $\Theta(n/\Delta \cdot \Delta \log \Delta) = \Theta(n \log \Delta)$ rounds.

Our proof of Theorem 1.1 is short and elegant, and is based on *drift analysis* [8]. It is presented in an expository way and provides insight in why our potential function is appropriate for the analysis. We complement the analyis by tail bounds in Section 3.

Finally, we conclude the paper by a brief discussion of the parallelized version of this algorithm, where all "conflicted" vertices get recolored simultaneously (instead of Step 3), and we prove that this variant takes exponential time.

2 Proof of Theorem 1.1

We start with introducing some notation. We use c_t to denote the coloring of the graph after t recoloring steps, that is c_t is a function $c_t \colon V(G) \to \{1, \ldots, \Delta + 1\}$. With $M_t \subseteq E$ we denote the set of *monochromatic edges* in c_t. Observe that c_t is a proper coloring of G if and only if $|M_t| = 0$. Our main goal is thus to establish good bounds on (the reduction of) the size of the sets M_t. In order to do so it is helpful to view the recoloring step(s) (i.e. Step 2 and Step 3) as a (slightly different) three step process:

RECOLORING STEP

For every $t \geqslant 1$,

S1 choose a monochromatic connected component $C \subseteq M_{t-1}$ at random proportional to the number of vertices in C;

S2 choose a vertex $v \in V(C)$ uniformly at random;

S3 let $c_t(v)$ be a uniformly at random chosen color from $\{1, \ldots, \Delta + 1\}$ and set $c_t(u) = c_{t-1}(u)$ for all $u \neq v$.

As a main tool in bounding the expected number of recoloring steps we use a so-called drift theorem (see [8, Theorem 2.3.1]).

▶ **Theorem 2.1** (Additive Drift Theorem [6]). *Let $(X_t)_{t\geqslant 0}$ be a sequence of non-negative random variables with a finite state space $\mathcal{S} \subset \mathbb{R}_0^+$ such that $0 \in \mathcal{S}$. Let $T := \inf\{t \geqslant 0 \mid X_t = 0\}$. If there exists $\delta > 0$ such that for all $s \in \mathcal{S} \setminus \{0\}$ and for all $t > 0$,*

$$\mathbb{E}[X_t - X_{t-1} \mid X_t = s] \leqslant -\delta, \tag{1}$$

then

$$\mathbb{E}[T] \leqslant \mathbb{E}[X_0] \,/\, \delta.$$

By *drift* we refer to the expectation $\mathbb{E}[X_t - X_{t-1} \mid X_t = s]$. (For a more extensive introduction to drift analyis, we refer the reader to [8].) Our goal is to apply Theorem 2.1 by assigning to each coloring c_t a real value $\Phi(t)$ (which we plug in for X_t) so that $\Phi(t) = 0$ if and only if c_t is a proper coloring. The *potential function* $\Phi(\cdot)$ we eventually use to prove Theorem 1.1 consists of several terms (see equation (3) below) and in order to motivate each of the terms we introduce them one by one. The simplest and most natural choice is to consider just the number of monochromatic edges, i.e. $\Phi(t) := |M_t|$. (Mind that this is only for explanatory purposes and will not be the final definition of Φ.) To apply Theorem 2.1 we need to estimate the drift from a single recoloring step. In the following claim (and in fact all similar ones in this section), the expectation is always taken with respect to a single recoloring step. That is, we (implicitly) condition on the coloring c_{t-1} without stating it every time. Note that this formulation implies what is required by equation (1).

As it turns out, in the case of $\Phi(t) := |M_t|$ we do not need to make use of the fact that the component C is chosen randomly, we may assume that the component C is given arbitrarily or even by an adversary.

▷ Claim 2.2. For all $t \geqslant 1$ and any connected component C in M_{t-1} we have

$$\mathbb{E}\left[|M_t| \;\middle|\; C\right] \leqslant |M_{t-1}| - \bar{d}(C) + 1 - \frac{1}{\Delta + 1},$$

where $\bar{d}(C)$ denotes the average degree of the graph induced by $V(C)$.

Proof. The claim follows easily from the following two observations. As v is chosen uniformly at random within C (as in Step 2), we decrease the number of monochromatic edges within C by $\bar{d}(C)$ whenever the newly chosen color is different from the current color of C, which happens with probability $\Delta/(\Delta + 1)$. All edges incident to v that do not belong to C become monochromatic with probability $1/(\Delta + 1)$. Thus we have

$$\mathbb{E}\left[|M_t| \;\middle|\; C\right] \leqslant |M_{t-1}| - \bar{d}(C) \cdot \frac{\Delta}{\Delta + 1} + \frac{\Delta - \bar{d}(C)}{\Delta + 1} = |M_{t-1}| - \bar{d}(C) + 1 - \frac{1}{\Delta + 1},$$

as claimed. ◁

As the average degree of every monochromatic component in M_t is at least one, Claim 2.2 implies $\mathbb{E}[|M_t|] \leqslant |M_{t-1}| - 1/(\Delta + 1)$ whenever $|M_{t-1}| > 0$. The following proposition then easily follows from Theorem 2.1.

▶ **Proposition 2.3.** *Let $D > 0$ be any fixed constant. For every graph G and every coloring c_0 of G such that $|M_0| \leqslant Dn/\Delta$ the decentralized coloring algorithm reaches a proper $(\Delta + 1)$-coloring in expectation after $O(n)$ recoloring steps.*

Unfortunately, a random coloring of a graph G with $\Delta + 1$ colors has in expectation $\Theta(n)$ monochromatic edges, so Proposition 2.3 is not immediately applicable. Instead, Claim 2.2 together with Theorem 2.1 only provide us with the bound of $O(n\Delta)$ (see Bhartia et al. [1]). In order to go beyond this, observe that Claim 2.2 actually gives a drift of $-1/3$ whenever $|V(C)| \geqslant 3$, as the average degree of a connected graph on $s \geqslant 3$ vertices is at least $4/3$. Thus, the only critical case are components C that consist of only one edge. To handle these we introduce some more notation.

We denote by $I_t \subseteq M_t$ the set of *isolated edges*, that is all edges which are monochromatic components of size two. We also let $P_t \subseteq V$ stand for the set of all properly colored vertices, i.e. the vertices that are not incident to any edge in M_t. Akin to Claim 2.2, the next claim gives a bound on the expected change in the number of isolated edges in one recoloring step.

▷ Claim 2.4. For all $t \geqslant 1$ and any connected component C in M_{t-1} we have

$$\mathbb{E}\left[|I_t| \;\middle|\; C\right] \leqslant |I_{t-1}| + \bar{d}(C) + 1.$$

For components C that form an isolated edge, we have in addition

$$\mathbb{E}\left[|I_t| \;\middle|\; C = uw\right] \leqslant |I_{t-1}| - \frac{\Delta}{\Delta + 1} + \frac{|N(u) \cap P_{t-1}| + |N(w) \cap P_{t-1}|}{2(\Delta + 1)}.$$

Proof. By recoloring a vertex v, the only isolated edges that can be created are edges that are incident to neighbors of v within C (at most one isolated edge per neighbor of v) and edges between v and P_{t-1} (naturally at most one edge incident to v can be isolated and monochromatic). This proves the first inequality. For the second assume that $C = uw$. Clearly, after recoloring one of u and w with a different color (which happens with probability $\Delta/(\Delta + 1)$), the isolated edge $C = uw$ disappears. Observe, also that a new isolated edge can only be generated if we choose as a new color for u (or w) a color of a vertex in $N(u) \cap P_{t-1}$ (or $N(w) \cap P_{t-1}$) respectively. This, together with the fact that each of u or w is chosen in Step 2 with probability $1/2$, proves the second inequality. ◁

We pause for a moment from the proof of Theorem 1.1 to showcase the use of previous claims for proving a positive result about complete bipartite graphs.

▶ **Proposition 2.5.** *For complete bipartite graphs $G = K_{n,m}$ the decentralized coloring algorithm reaches a proper $(\Delta + 1)$-coloring in expectation after $O(\min\{n, m\})$ recoloring steps.*

Proof. Observe that for complete bipartite graphs vertices of $P_{t-1} \cap A$ and $P_{t-1} \cap B$ need to be colored with *different* colors (here A and B denote the two parts of the bipartite graph). Also note that an isolated edge can only be generated if a color appears only once in $P_{t-1} \cap A$ (and $P_{t-1} \cap B$). Therefore, for a monochromatic edge uv we have $|N(u) \cap P_{t-1}| + |N(w) \cap P_{t-1}| \leqslant \Delta$. We can thus replace the bound in the second inequality of Claim 2.4 by

$$\mathbb{E}\left[|I_t| \;\middle|\; C = uw\right] \leqslant |I_{t-1}| - \frac{\Delta}{\Delta + 1} + \frac{\Delta}{2(\Delta + 1)} \leqslant |I_{t-1}| - \frac{\Delta}{2(\Delta + 1)}. \tag{2}$$

Consider now the potential function $\Phi(t) := |M_t| + \frac{1}{10}|I_t|$. In case $|V(C)| \geqslant 3$, from Claim 2.2 and Claim 2.4, we get (with room to spare)

$$\begin{aligned}\mathbb{E}\left[\Phi(t) \mid C, |V(C)| \geqslant 3\right] &\leqslant |M_{t-1}| - \bar{d}(C) + 1 - \frac{1}{\Delta+1} + \frac{1}{10}|I_{t-1}| + \frac{1}{10}\bar{d}(C) + \frac{1}{10}\\ &\leqslant \Phi(t-1) - \frac{9}{10}\bar{d}(C) + \frac{11}{10} \leqslant \Phi(t-1) - \frac{1}{20},\end{aligned}$$

where we used the fact that $\bar{d}(C) \geqslant 4/3$. On the other hand, if $C = uw$ then by Claim 2.2 and (2) we have

$$\begin{aligned}\mathbb{E}\left[\Phi(t) \mid C = uw\right] &\leqslant |M_{t-1}| - \bar{d}(C) + 1 - \frac{1}{\Delta+1} + \frac{1}{10}|I_{t-1}| - \frac{1}{10}\frac{\Delta}{2(\Delta+1)}\\ &\leqslant \Phi(t-1) - \frac{1}{20}\Big(\frac{20}{\Delta+1} + \frac{\Delta}{\Delta+1}\Big) \leqslant \Phi(t-1) - \frac{1}{20}.\end{aligned}$$

In conclusion,

$$\mathbb{E}\left[\Phi(t) \mid C\right] \leqslant \Phi(t-1) - \frac{1}{20},$$

for every component C. The proposition now follows from Theorem 2.1 together with the fact that in a random x-coloring an edge is monochromatic with probability $1/x$ and thus

$$\mathbb{E}[\Phi(0)] \leqslant \mathbb{E}[|M_0|] + \mathbb{E}[|I_0|] \leqslant 2 \cdot \frac{n \cdot m}{\max\{n,m\}+1} \leqslant 4\min\{n,m\},$$

with room to spare. ◀

We note that this proof actually shows that the assertion of Proposition 2.5 remains true, for sufficiently small $\varepsilon > 0$, if we reduce the number of colors to be used by the algorithm to $(1-\varepsilon)\Delta$, that is $(1-\varepsilon)\max\{n,m\}$. We do not elaborate further on this.

After this short detour we come back to the proof of Theorem 1.1. What one could conclude from the two claims above is that if we were to choose a component C in Step 1 which is of size at least three throughout the process, then the drift obtained (Claim 2.2) would always be less than $-1/3$. However, this is far too optimistic to hope for.

Consider an isolated edge uv and assume we recolor v. If the new color chosen *does not* belong to its properly-colored neighborhood $N(v) \cap P_{t-1}$, then the number of monochromatic isolated edges decreases by one. This happens with constant probability unless the size of $N(v) \cap P_{t-1}$ is close to Δ.

Since in Step 1 we choose C *randomly*, we expect a strong drift "towards the target" as long as we are in one of the situations from the paragraphs above. In other words, we have a desired drift unless M_{t-1} comprises mostly of isolated edges *and* most vertices $u \in V(I_{t-1})$ have almost Δ neighbors in $N(u) \cap P_{t-1}$.

Let us hence analyze what happens if in such a case we recolor a vertex v belonging to an isolated edge uv. Suppose we set $c_t(v) := c_{t-1}(x)$ for some $x \in N(v) \cap P_{t-1}$. If the color $c_{t-1}(x)$ appears multiple times in $N(v) \cap P_{t-1}$, we do not create a new isolated edge. Otherwise, the edge xv becomes isolated and $P_t := (P_{t-1} \setminus \{x\}) \cup \{u\}$. However, crucially, as we assumed that every vertex $u \in V(I_{t-1})$ had roughly Δ neighbors in P_{t-1}, we conversely have that an average vertex in P_{t-1} has roughly $\Delta|V(I_{t-1})|/|P_{t-1}|$ neighbors in $V(I_{t-1})$. Thus, we may expect that $N(x) \cap P_t$ is *smaller* than Δ. In other words, we expect that $e(V(I_t), P_t)$ is smaller than $e(V(I_{t-1}), P_{t-1})$. Here and throughout we use $e(X,Y)$ to denote the number of edges between two disjoint vertex sets X and Y.

Previous considerations motivate keeping track of $e(V(I_t), P_t)$ as well and lead us to formulate the following potential function:

$$\Phi(t) := |M_t| + \frac{|I_t|}{10} + \frac{e(V(I_t), P_t)}{100\Delta}. \tag{3}$$

Note that the value of $\Phi(t)$ is always proportional to the number of monochromatic edges.

▷ Claim 2.6. For all $t \geqslant 1$ we have

$$|M_t| \leqslant \Phi(t) \leqslant 2|M_t|.$$

Proof. The first inequality is trivial. The second follows, with room to spare, as $I_t \subseteq M_t$ and $e(V(I_t), P_t) \leqslant 2|I_t| \cdot \Delta$. ◁

With Claim 2.6 at hand we deduce from Proposition 2.3 that in order to complete the proof of Theorem 1.1 it suffices to show that the algorithm reduces the potential Φ to a value of Dn/Δ in $O(n \log \Delta)$ steps, for some arbitrarily large but fixed constant $D > 0$. This is what we do in the remainder of this section.

Note also that there is no hope to always get a *constant* drift, as by Theorem 2.1 this would then lead to a bound of $O(n)$ recoloring steps, which would contradict the bound of $\Omega(n \log n)$ for K_n. Instead we show a *multiplicative* drift.

▷ Claim 2.7. For any $t \geqslant 1$ with $\Phi(t-1) > 0$, we have

$$\mathbb{E}[\Phi(t)] \leqslant \Phi(t-1)\Big(1 - \frac{1}{1000n}\Big).$$

Proof. By linearity of expectation we can consider each term of $\Phi(\cdot)$ in (3) independently. The first two terms are handled by Claim 2.2 and Claim 2.4, so we first establish some bounds on the third. Observe that in order for an edge to be counted in $e(V(I_t), P_t)$ but not in $e(V(I_{t-1}), P_{t-1})$ it must be incident to a vertex in either $V(I_t) \setminus V(I_{t-1})$ or $P_t \setminus P_{t-1}$. Let C be a component chosen in Step 1 and v a vertex chosen in Step 2. For any vertex in $\{v\} \cup (N(v) \cap V(C))$, we either get one new isolated edge or one new properly colored vertex (or neither). In the former, the other endpoint of that edge potentially contributes by Δ to $e(V(I_t), P_t)$, and in the latter each monochromatic edge with one endpoint in $N(v) \cap V(C)$ potentially contribute by Δ to $e(V(I_t), P_t)$ for each of its endpoints. Thus we have

$$\mathbb{E}[e(V(I_t), P_t) \mid C] \leqslant e(V(I_{t-1}), P_{t-1}) + (\bar{d}(C) + 1) \cdot 2\Delta,$$

where as before $\bar{d}(C)$ denotes the average degree of the component C. Together with Claim 2.2 and Claim 2.4, for all components C on at least three vertices we get

$$\begin{aligned} \mathbb{E}\left[\Phi(t) \mid C, |V(C)| \geqslant 3\right] &\leqslant \Phi(t-1) - \Big(1 - \frac{1}{10} - \frac{2}{100}\Big)\bar{d}(C) + 1 + \frac{1}{10} + \frac{2}{100} \\ &\leqslant \Phi(t-1) - \frac{1}{25}\bar{d}(C), \end{aligned} \tag{4}$$

where the last inequality follows from $\bar{d}(C) \geqslant 4/3$.

Next we consider the third term of $\Phi(\cdot)$ conditioned on choosing a component $C \subseteq I_{t-1}$, i.e. C is an isolated edge. We first let $d(v, X) := |N(v) \cap X|$ for all $v \in V$ and sets $X \subseteq V$ and denote by

$$\bar{d}_{IP} := \frac{1}{|V(I_{t-1})|} \sum_{u \in V(I_{t-1})} d(u, P_{t-1}) \qquad \text{and} \qquad \bar{d}_{PI} := \frac{1}{|P_{t-1}|} \sum_{u \in P_{t-1}} d(u, V(I_{t-1}))$$

the average degree of vertices in $V(I_{t-1})$ into P_{t-1}, and the average degree of vertices in P_{t-1} into $V(I_{t-1})$, respectively. Note that, of course, we have $\sum_{u \in V(I_{t-1})} d(u, P_{t-1}) = \sum_{u \in P_{t-1}} d(u, V(I_{t-1}))$, and hence $\bar{d}_{IP}|V(I_{t-1})| = \bar{d}_{PI}|P_{t-1}|$.

Consider an isolated edge wv and assume v gets recolored with a new color. Then, since w is now properly colored, all $d(w, P_{t-1})$ edges incident to w which contributed to $e(V(I_{t-1}), P_{t-1})$ are not counted in $e(V(I_t), P_t)$, except possibly one in case v forms a new isolated edge with a neighbor of w. Moreover, any new edge counted in $e(V(I_t), P_t)$ must be incident to either v, w, or a vertex $x \in P_{t-1}$ for which $vx \in I_t$. There are at most $\Delta - d(v, P_{t-1})$, $\Delta - d(w, P_{t-1})$, and $\Delta - d(x, V(I_{t-1}))$ such edges respectively not already counted in $e(V(I_{t-1}), P_{t-1})$. Combining all this we get

$$\begin{aligned} e(V(I_t), P_t) \leqslant e(V(I_{t-1}), P_{t-1}) &- d(w, P_{t-1}) + 1 + \Delta - d(v, P_{t-1}) \\ &+ \Delta - d(w, P_{t-1}) + \sum_{x \in P_{t-1}} \mathbb{1}_{vx \in I_t} \big(\Delta - d(x, V(I_{t-1}))\big), \end{aligned}$$

if $c_t(v) \neq c_{t-1}(v)$, and of course $e(V(I_t), P_t) = e(V(I_{t-1}), P_{t-1})$ if $c_t(v) = c_{t-1}(v)$.

We conclude that

$$\begin{aligned} \mathbb{E}\left[e(V(I_t), P_t) \mid C \subseteq I_{t-1}\right] &\leqslant e(V(I_{t-1}), P_{t-1}) + \frac{\Delta}{\Delta+1}(2\Delta + 1 - 3\bar{d}_{IP}) \\ &+ \frac{1}{|V(I_{t-1})|} \sum_{x \in P_{t-1}} \sum_{v \in V(I_{t-1}) \cap N(x)} \frac{1}{\Delta+1} \big(\Delta - d(x, V(I_{t-1}))\big), \end{aligned}$$

where the last term can be rewritten as

$$\frac{1}{|V(I_{t-1})|} \sum_{x \in P_{t-1}} \frac{d(x, V(I_{t-1}))\big(\Delta - d(x, V(I_{t-1}))\big)}{\Delta + 1}.$$

We note that the summand above can be written as $f(d(x, V(I_{t-1}))$ where $f(y) := y(\Delta - y)/(\Delta + 1)$ is a concave function. Hence, by Jensen's inequality, we can upper bound the expression by $|P_{t-1}|f(\bar{d}_{PI})/|V(I_{t-1})| = \bar{d}_{IP}(\Delta - \bar{d}_{PI})/(\Delta + 1)$. Altogether we get

$$\begin{aligned} \mathbb{E}\left[e(V(I_t), P_t) \mid C \subseteq I_{t-1}\right] &\leqslant e(V(I_{t-1}), P_{t-1}) + \frac{\Delta(2\Delta + 1 - 3\bar{d}_{IP})}{\Delta+1} + \frac{\bar{d}_{IP}(\Delta - \bar{d}_{PI})}{\Delta + 1} \\ &\leqslant e(V(I_{t-1}), P_{t-1}) + \frac{\Delta}{\Delta+1}(2\Delta + 1 - 2\bar{d}_{IP} - \bar{d}_{IP}\bar{d}_{PI}/\Delta). \end{aligned}$$

Finally, by combining this with Claim 2.2 and Claim 2.4 (and some tedious calculation) we deduce

$$\begin{aligned} \mathbb{E}\left[\Phi(t) \mid C \subseteq I_{t-1}\right] &\leqslant \Phi(t-1) - \frac{1}{\Delta+1} - \frac{1}{10}\frac{\Delta - \bar{d}_{IP}}{\Delta+1} + \frac{1}{100}\frac{2\Delta + 1 - 2\bar{d}_{IP} - \bar{d}_{IP}\bar{d}_{PI}/\Delta}{\Delta+1} \\ &\leqslant \Phi(t-1) - \frac{2}{25}\frac{\Delta - \bar{d}_{IP}}{\Delta+1} - \frac{1}{100}\frac{\bar{d}_{IP}\bar{d}_{PI}}{\Delta(\Delta+1)}. \end{aligned} \tag{5}$$

With all these preparations we are now in a position to bound $\mathbb{E}[\Phi(t)]$. As seen in (4) and (5), both conditioning on components of size at least three or on vertices in isolated edges lead to a non-positive contribution to the drift. In order to derive an upper bound on $\mathbb{E}[\Phi(t)]$ we may thus ignore one of the terms for convenience. If we assume $|I_{t-1}| \leqslant |M_{t-1}|/2$ one would expect that the larger contribution to the change of $\Phi(t-1)$ comes from components which are not isolated edges. Indeed, in that case we may ignore the term from (5) and use (4) only to get

$$\begin{aligned}\mathbb{E}[\Phi(t)] &\overset{(4)}{\leqslant} \Phi(t-1) - \sum_{C,\ |V(C)|\geqslant 3} \frac{|V(C)|}{|V(M_{t-1})|}\frac{\bar{d}(C)}{25} = \Phi(t-1) - \frac{2|M_{t-1} \smallsetminus I_{t-1}|}{25|V(M_{t-1})|}\\ &\leqslant \Phi(t-1) - \frac{|M_{t-1}|}{25n} \leqslant \Phi(t-1)\Big(1 - \frac{1}{50n}\Big),\end{aligned}$$

where the last inequality follows from Claim 2.6.

On the other hand, suppose $|I_{t-1}| \geqslant |M_{t-1}|/2$ and observe that this implies $|V(I_{t-1})| \geqslant |V(M_{t-1})|/2$. This means that the probability of picking a vertex in $V(I_{t-1})$ to recolor is at least 1/2 and one may hope that the larger contribution to the change of $\Phi(t-1)$ comes from the isolated edges. Indeed, similarly as above, we now ignore the contribution from components of size at least three to get:

$$\mathbb{E}[\Phi(t)] \overset{(5)}{\leqslant} \Phi(t-1) - \frac{1}{25}\frac{\Delta - \bar{d}_{IP}}{\Delta+1} - \frac{1}{200}\frac{\bar{d}_{IP}\bar{d}_{PI}}{\Delta(\Delta+1)} \leqslant \Phi(t-1) - \frac{1}{50}\frac{\Delta-\bar{d}_{IP}}{\Delta} - \frac{1}{400}\frac{\bar{d}_{IP}\bar{d}_{PI}}{\Delta^2}. \quad (6)$$

If $\bar{d}_{IP} \leqslant \Delta - \Delta\Phi(t-1)/(30n)$, then the claim follows just from the first term. Otherwise, by Claim 2.6

$$\Phi(t-1) \leqslant 2|M_{t-1}| \leqslant 4|I_{t-1}| \leqslant 2n,$$

which in turn implies $\bar{d}_{IP} \geqslant \Delta(1 - \Phi(t-1)/(30n)) \geqslant 14\Delta/15$. Recall, $\bar{d}_{PI}|P_{t-1}| = \bar{d}_{IP}|V(I_{t-1})|$, and note that $|V(I_{t-1})|/|P_{t-1}| \geqslant 2|I_{t-1}|/n \geqslant \Phi(t-1)/(2n)$. Therefore,

$$\frac{1}{400}\frac{\bar{d}_{IP}\bar{d}_{PI}}{\Delta^2} = \frac{1}{400}\frac{|V(I_{t-1})| \cdot \bar{d}_{IP}\bar{d}_{IP}}{|P_{t-1}|\Delta^2} \geqslant \Phi(t-1)\frac{1}{800n}\cdot\Big(\frac{14}{15}\Big)^2$$

and the second term in (6) is enough to conclude the proof of Claim 2.7. ◁

As mentioned in the paragraph before Claim 2.7, in order to make use of the assertion of Claim 2.7, we need a slightly different drift theorem, one for multiplicative drift.

▶ **Theorem 2.8** (Multiplicative Drift Theorem [4]). *Let $(X_t)_{t\geqslant 0}$ be a sequence of non-negative random variables with a finite state space $\mathcal{S} \subset \mathbb{R}_0^+$ such that $0 \in \mathcal{S}$. Let $s_{\min} := \min\{\mathcal{S} \smallsetminus \{0\}\}$, let $s_0 \in \mathcal{S} \smallsetminus \{0\}$, and let $T := \inf\{t \geqslant 0 \mid X_t = 0\}$. If there exists $\delta > 0$ such that for all $s \in \mathcal{S} \smallsetminus \{0\}$ and for all $t > 0$,the case of*

$$\mathbb{E}[X_t - X_{t-1} \mid X_{t-1} = s] \leqslant -\delta s,$$

then

$$\mathbb{E}[T \mid X_0 = s_0] \leqslant \frac{1 + \ln(s_0/s_{\min})}{\delta}.$$

Now we are ready to put things together to prove Theorem 1.1.

Proof of Theorem 1.1. For every $t \geqslant 0$, we define

$$\Phi'(t) = \begin{cases} \Phi(t), & \text{if } \Phi(t) \geqslant n/\Delta,\\ 0, & \text{otherwise.}\end{cases}$$

Note that, as long as $\Phi(t-1) \geqslant n/\Delta$, we have $\Phi'(t-1) = \Phi(t-1)$ and $\Phi'(t) \leqslant \Phi(t)$, so the deduced bound on $\Phi(t)$ in Claim 2.7 is also a bound for $\Phi'(t)$. Using Theorem 2.8 with Claim 2.7 for $T' := \inf\{t \geqslant 0 \mid \Phi'(t) = 0\} = \inf\{t \geqslant 0 \mid \Phi(t) < n/\Delta\}$, we get for all $s_0 > 0$

$$\mathbb{E}[T' \mid \Phi'(0) = s_0] \leqslant \frac{1 + \ln\left(\frac{s_0}{n/\Delta}\right)}{(1000n)^{-1}}.$$

By Claim 2.6 we have $\Phi'(0) \leqslant 2|M_0| \leqslant n\Delta$, and therefore

$$\mathbb{E}[T'] \leqslant 1000n\big(1 + 2\ln\Delta\big) = O(n\log\Delta).$$

Finally, as by Claim 2.6 we then have $|M_{T'}| = O(n/\Delta)$, we conclude from Proposition 2.3 that the expected number of steps after T' to reach a legal coloring is $O(n)$. Therefore, the total number of required steps to reach a legal coloring is $O(n\log\Delta)$, which finishes the proof of Theorem 1.1. ◀

3 Tail Bounds

In this section, we prove that the runtime of the decentralized coloring algorithm is of order $O(n\log\Delta)$ not only in expectation, but also with high probability. It turns out that this does not require much additional work, as the drift theorems are accompanied with suitable tail bounds. In many situations, concentration bounds require conditions beyond the drift, for example bounds on the step size. Notably, for multiplicative drift such additional conditions are not necessary, as the following theorem holds.

▶ **Theorem 3.1** (Multiplicative Drift Tail Bound [3]). *Let $(X_t)_{t\geqslant 0}$ be a sequence of non-negative random variables with a finite state space $\mathcal{S} \subset \mathbb{R}_0^+$ such that $0 \in \mathcal{S}$. Let $s_{\min} := \min\{\mathcal{S} \setminus \{0\}\}$, let $s_0 \in \mathcal{S} \setminus \{0\}$, and let $T := \inf\{t \geqslant 0 \mid X_t = 0\}$. Suppose that $X_0 = s_0$, and that there exists $\delta > 0$ such that for all $s \in \mathcal{S} \setminus \{0\}$ and for all $t > 0$,*

$$\mathbb{E}[X_t - X_{t-1} \mid X_{t-1} = s] \leqslant -\delta s.$$

Then, for all $r \geqslant 0$

$$\Pr\left[T > \left\lceil \tfrac{r+\ln(s_0/s_{\min})}{\delta} \right\rceil\right] < e^{-r}.$$

The following proposition is a straightforward application of this theorem.

▶ **Proposition 3.2.** *Suppose $\Delta = \Omega(n^c)$ for some constant $c > 0$. Then, the decentralized coloring algorithm terminates after $O(n\log\Delta)$ steps with high probability.*

Proof. By Claim 2.6 we know that $s_0 \leqslant n\Delta$. In the proof of Theorem 1.1 we analyzed the process with multiplicative drift until the potential Φ hits n/Δ. Here, we track this potential until the end of the algorithm. Note that the smallest nonzero value Φ can attain is at least 1. Thus, under the assumption $\Delta = \Omega(n^c)$, we also have $\ln(s_0/s_{\min}) \leqslant C\log\Delta$ for large enough n, even if we do not truncate Φ. Here, $C > 0$ is a suitable constant, for example $C = 2(1 + 1/c)$. Setting $r = \log\Delta$, we can apply Theorem 3.1 to get

$$\Pr\big[T > \lceil(1 + C)1000n\log\Delta\rceil\big] \leqslant e^{-\log\Delta} = O(n^{-c}) = o(1),$$

where we used $\delta = (1000n)^{-1}$ as before, due to Claim 2.7. ◀

The proof of Proposition 3.2 fails when $\log n = \omega(\log\Delta)$. In the proof of Theorem 1.1 we switched to additive drift to analyze the second phase of the process. We use this approach again. The following tail bound for additive drift will be useful. It is a rather straightforward consequence of Azuma's inequality. Note that there is an additional assumption, namely that we have bounded step size.

▶ **Theorem 3.3** (Additive Drift Tail Bound [7]). *Let $(X_t)_{t\geqslant 0}$ be a sequence of non-negative random variables with a finite state space $\mathcal{S} \subset \mathbb{R}_0^+$ such that $0 \in \mathcal{S}$. Let $T := \inf\{t \geqslant 0 \mid X_t = 0\}$. Suppose there are $c, \delta > 0$ such that for all $s \in \mathcal{S} \setminus \{0\}$ and for all $t > 0$, we have both $\mathbb{E}[X_t - X_{t-1} \mid X_t = s] \leqslant -\delta$ and $|X_{t+1} - X_t| < c$. Then, for all $r \geqslant 2X_0/\delta$,*

$$\Pr[T \geqslant r] \leqslant \exp\Big(-\frac{r\delta^2}{8c^2}\Big).$$

The smaller Δ is, the less the potential changes at each step. Using this fact, the theorem above allows us to prove the next proposition. It gives us that the runtime of the algorithm is $O(n \log \Delta)$ for smaller Δ.

▶ **Proposition 3.4.** *If $\Delta = O(n^{1/4})$, the decentralized coloring algorithm terminates after $O(n \log \Delta)$ steps with high probability*

Proof. We go back to splitting the process in two phases as in the proof of Theorem 1.1. Let T_1 and T_2 be the duration of Phase 1 and 2 respectively. We consider Phase 1 first. To be able to apply Theorem 3.3, we use the potential function $\Psi(t) := \max\{\log(\Delta\Phi(t)/n), 0\}$. As we will see, the logarithm converts multiplicative drift into additive drift. Note that $T_1 = \inf\{t \geqslant 0 \mid \Psi(t) = 0\}$. Using Jensen's inequality and Claim 2.7, we get for all $s \in S \setminus \{0\}$ and $t \geqslant 0$

$$\begin{aligned}
\mathbb{E}[\Psi(t+1) - \Psi(t) \mid \Psi(t) = s] &= \mathbb{E}\left[\log\left(\frac{\Delta\Phi(t+1)}{n}\right) - \log\left(\frac{\Delta\Phi(t)}{n}\right)\middle|\Psi(t) = s\right] \\
&= \mathbb{E}\left[\log\left(\frac{\Phi(t+1)}{\Phi(t)}\right)\middle|\Psi(t) = s\right] \\
&\leqslant \log \mathbb{E}\left[\frac{\Phi(t+1)}{\Phi(t)}\middle|\Psi(t) = s\right] \\
&\leqslant \log\left(1 - \frac{1}{1000n}\right) \\
&\leqslant -\frac{1}{1000n}.
\end{aligned}$$

The last inequality follows as $\log x \leqslant x - 1$ for all $x > 0$. Now that we have determined the drift, we need to bound the step size. $|M_t|$ and $|I_t|$ can change by at most Δ at each step and $e(V(I_t, P_t)$ by at most Δ^2. Thus, the step size of Φ is bounded by 2Δ. As the logarithm is concave, the largest effect of such a change in Φ on the value of Ψ is when Φ is as small as possible. In particular, this is the case if Φ goes from $2\Delta + n/\Delta$ to n/Δ. Thus we have

$$|\Psi(t+1) - \Psi(t)| \leqslant \left|\log\left(\frac{2\Delta + \frac{n}{\Delta}}{\frac{n}{\Delta}}\right)\right| = \log\left(1 + \frac{2\Delta^2}{n}\right) < \frac{2\Delta^2}{n}.$$

Therefore, $2\Delta^2/n$ is a bound on the step size of Ψ. As $\Phi(0) \leqslant n\Delta$ we have $\Psi(0) \leqslant 2 \log \Delta$. Hence we can use Theorem 3.3 with $r = 4000n(1 + \log \Delta)$. We get

$$\Pr[T_1 \geqslant r] \leqslant \exp\left(-\frac{4000n(1+\log\Delta)\cdot\left(\frac{1}{1000n}\right)^2}{8\left(\frac{2\Delta^2}{n}\right)^2}\right) = \exp\left(-\frac{(1+\log\Delta)n}{8000\Delta^4}\right) = o(1).$$

We turn our attention to Phase 2. Here, we already have additive drift for $|M_t|$, so we can use Theorem 3.3 immediately. Claim 2.2 gives us

$$\mathbb{E}\big[|M_t| - |M_{t-1}| \;\big|\; |M_{t-1}| = s\big] \leqslant -\frac{1}{\Delta+1} \leqslant -\frac{1}{2\Delta},$$

for all $s,t > 0$. We also have $\big||M_t| - |M_{t-1}|\big| \leqslant \Delta$ for all $t > 0$. By Claim 2.6 we get $|M_t| \leqslant 2n/\Delta$ at the start of Phase 2. Hence we can apply Theorem 3.3 with $r = (4+\log\Delta)n$. We get

$$\Pr[T_2 \geqslant r] \leqslant \exp\left(-\frac{(4+\log\Delta)n\cdot\left(\frac{1}{2\Delta}\right)^2}{8\Delta^2}\right) = \exp\left(-\frac{(4+\log\Delta)n}{32\Delta^4}\right) = o(1).$$

Using a union bound gives us that the algorithm terminates in $O(n\log\Delta)$ round with high probability. ◀

Proposition 3.2 and 3.4 cover all possible values of Δ. Therefore, the algorithm has runtime $O(n\log\Delta)$ with high probability for any Δ.

4 A simultaneous-recoloring variant of the algorithm

A natural question is whether the original algorithm can be parallelized. So what if instead of choosing one conflicted vertex at a time in Step 2 *all* conflicted vertices would simultaneously want to change their color? It turns out that this process does not even have polynomial runtime on the complete graph K_n.

▶ **Proposition 4.1.** *The Decentralized Graph Colouring algorithm in which all conflicted vertices choose a new color uniformly at random needs* $e^{\Omega(n)}$ *rounds in expectation to terminate on a complete graph on* n *vertices* K_n.

Proof. Fix a sufficiently small constant $\varepsilon > 0$, e.g. $\varepsilon = 0.1$. For a round t, let X_t be the number of conflicted vertices, i.e., the number of vertices whose color is not unique. Due to symmetry, X_t is a Markov process. Let T_ε be the first round in which $X_t \leqslant \varepsilon n$. We show that T_ε has exponentially large expectation. Consider any round t with $X_t = x > \varepsilon n$. Then we show that

$$\Pr[X_{t+1} \leqslant \varepsilon n \mid X_t = x] = e^{-\Omega(n)}, \tag{7}$$

where the hidden constant is independent of x. We remark that the same argument also shows that with high probability $X_1 > \varepsilon n$, since the initial round is formally equivalent to the hypothetical case $X_0 = n$. So the proposition follows if we can show (7).

To show (7), we uncover the new colors in two batches. In the first batch, we uncover the colors of all but εn vertices. If there are more than εn vertices in conflicts from the first batch, then there is nothing to show. So in the following we may assume (and implicitly condition) on the opposite event that uncovering the first batch creates at most εn conflicted vertices. This implies that the set C_1 of colors appearing among the $(1-\varepsilon)n$ uncovered vertices, has size at least $|C_1| \geqslant (1-2\varepsilon)n$. Let $C_2 \subseteq C_1$ be the set of colors in C_1 that also appear in the second batch, i.e., for which a conflict is created by the second batch. The probability that a fixed color in C_1 does *not* occur in C_2 is $(1-1/n)^{\varepsilon n} = e^{-\varepsilon + O(1/n)} \leqslant 1 - 7\varepsilon/8$ for sufficiently large n, where we use that $e^{-\varepsilon} < 1 - 7\varepsilon/8$ for $\varepsilon < 0.2$. Hence, $\mathbb{E}[|C_2|] \geqslant 7\varepsilon/8 \cdot (1-2\varepsilon)n \geqslant 5\varepsilon/8 \cdot n$ for $\varepsilon \leqslant 1/7$.

The size of C_2 is given by the number of non-empty bins in a *Balls-and-Bins* problem, and this number is known to be concentrated around its expectation since the number of empty bins is negatively associated, and thus the Chernoff bounds are applicable. Since this is a well-known argument, we refrain from spelling out the details and refer the reader to the standard exposition [5, Proposition 29 and Section 3.3]. The result is that $\Pr[|C_2| \leqslant \varepsilon/2 \cdot n] = e^{-\Omega(n)}$.

It remains to observe that $X_{t+1} > 2|C_2|$, since every color in C_2 causes at least two conflicted vertices (one from the second batch and one from the rest). Hence, $\Pr[X_{t+1} \leqslant \varepsilon n] \leqslant \Pr[|C_2| \leqslant \varepsilon/2 \cdot n] = e^{-\Omega(n)}$, as required. ◀

References

1 Apurv Bhartia, Deeparnab Chakrabarty, Krishna Chintalapudi, Lili Qiu, Bozidar Radunovic, and Ramachandran Ramjee. IQ-Hopping: distributed oblivious channel selection for wireless networks. In *Proceedings of the 17th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pages 81–90. ACM, 2016. doi:10.1145/2942358.2942376.

2 Deeparnab Chakrabarty and Paul de Supinski. On a Decentralized $(\Delta + 1)$-Graph Coloring Algorithm. In *Symposium on Simplicity in Algorithms*, pages 91–98. SIAM, 2020. doi:10.1137/1.9781611976014.13.

3 Benjamin Doerr and Leslie Ann Goldberg. Adaptive drift analysis. *Algorithmica*, 65(1):224–250, 2013. doi:10.1007/s00453-011-9585-3.

4 Benjamin Doerr, Daniel Johannsen, and Carola Winzen. Multiplicative drift analysis. *Algorithmica*, 64(4):673–697, 2012. doi:10.1007/s00453-012-9622-x.

5 Devdatt P Dubhashi and Desh Ranjan. Balls and bins: A study in negative dependence. *BRICS Report Series*, 3(25), 1996. doi:10.7146/brics.v3i25.20006.

6 Jun He and Xin Yao. A study of drift analysis for estimating computation time of evolutionary algorithms. *Natural Computing*, 3(1):21–35, 2004. doi:10.1023/b:naco.0000023417.31393.c7.

7 Timo Kötzing. Concentration of first hitting times under additive drift. *Algorithmica*, 75(3):490–506, 2016. doi:10.1007/s00453-015-0048-0.

8 Johannes Lengler. Drift analysis. In Benjamin Doerr and Frank Neumann, editors, *In: Theory of Evolutionary Computation: Recent Developments in Discrete Optimization*, pages 89–131. Springer International Publishing, Cham, 2020. doi:10.1007/978-3-030-29414-4_2.