


Colorful Vertex Recoloring of Bipartite Graphs

Boaz Patt-Shamir ✉ 

School of Electrical Engineering, Tel Aviv University, Israel

Adi Rosén ✉

CNRS and Université Paris Cité, France

Seeun William Umboh ✉ 

School of Computing and Information Systems, The University of Melbourne, Australia

ARC Training Centre in Optimisation Technologies, Integrated Methodologies, and Applications (OPTIMA), Melbourne, Australia

Abstract

We consider the problem of vertex recoloring: we are given n vertices with their initial coloring, and edges arrive in an online fashion. The algorithm is required to maintain a valid coloring by means of vertex recoloring, where recoloring a vertex incurs a cost. The problem abstracts a scenario of job placement in machines (possibly in the cloud), where vertices represent jobs, colors represent machines, and edges represent “anti affinity” (disengagement) constraints. Online coloring in this setting is a hard problem, and only a few cases were analyzed. One family of instances which is fairly well-understood is bipartite graphs, i.e., instances in which two colors are sufficient to satisfy all constraints. In this case it is known that the competitive ratio of vertex recoloring is $\Theta(\log n)$.

In this paper we propose a generalization of the problem, which allows using additional colors (possibly at a higher cost), to improve overall performance. Concretely, we analyze the simple case of bipartite graphs of bounded largest *bond* (a bond of a connected graph is an edge-cut that partitions the graph into two connected components). From the upper bound perspective, we propose two algorithms. One algorithm exhibits a trade-off for the uniform-cost case: given $\Omega(\log \beta) \leq c \leq O(\log n)$ colors, the algorithm guarantees that its cost is at most $O(\frac{\log n}{c})$ times the optimal offline cost for two colors, where n is the number of vertices and β is the size of the largest bond of the graph. The other algorithm is designed for the case where the additional colors come at a higher cost, $D > 1$: given Δ additional colors, where Δ is the maximum degree in the graph, the algorithm guarantees a competitive ratio of $O(\log D)$. From the lower bounds viewpoint, we show that if the cost of the extra colors is $D > 1$, no algorithm (even randomized) can achieve a competitive ratio of $o(\log D)$. We also show that in the case of general bipartite graphs (i.e., of unbounded bond size), any deterministic online algorithm has competitive ratio $\Omega(\min(D, \log n))$.

2012 ACM Subject Classification Theory of computation → Online algorithms; Theory of computation → Graph algorithms analysis; Theory of computation → Adversary models

Keywords and phrases online algorithms, competitive analysis, resource augmentation, graph coloring

Digital Object Identifier 10.4230/LIPIcs.STACS.2025.70

Related Version *Full Version:* <https://arxiv.org/abs/2501.05796>

Funding *Boaz Patt-Shamir:* This research was supported by the Israel Science Foundation, grant No. 1948/21.

Adi Rosén: Research partially supported by CNRS grant IEA ALFRED and ANR grant AlgoriDAM.

Seeun William Umboh: Funded by CNRS grant IEA ALFRED and by the Australian Government through the Australian Research Council DP240101353.

1 Introduction

Motivation. In the cloud, jobs are placed on machines according to multiple criteria. Sometimes, during the execution of a job, it turns out that the job is in conflict with another job, in the sense that the two jobs must not run on the same machine. Such conflicts may



© Boaz Patt-Shamir, Adi Rosén, and Seeun William Umboh;
licensed under Creative Commons License CC-BY 4.0

42nd International Symposium on Theoretical Aspects of Computer Science (STACS 2025).

Editors: Olaf Beyersdorff, Michał Pilipczuk, Elaine Pimentel, and Nguyễn Kim Thăng;

Article No. 70; pp. 70:1–70:19



Leibniz International Proceedings in Informatics

LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



arise due to limited resources on a single machine, or due to security considerations, or other reasons. Whenever such a conflict between two jobs located in the same machine is detected, the system needs to migrate one of the conflicting jobs (or both) so as to separate them. This situation was abstracted in [1] as the “vertex recoloring” (or “disengagement”) problem. Generally speaking, the problem is stated as an online problem, where initially we are given n vertices (representing jobs) colored in k colors (representing machines). Edges (representing conflicts) arrive online, and the algorithm is asked to output a valid vertex coloring after the arrival of each edge. Recoloring a vertex incurs a cost, and the cost incurred by an algorithm is the total cost of vertex recoloring events by the algorithm. As usual with online problems, the performance measure is the competitive ratio, namely the worst-case ratio between the cost paid by the online algorithm in question and the best cost possible (offline).

In this paper we seek to generalize the problem so as to make it both richer theoretically and more realistic practically. Our starting point is the following. Recalling the motivating scenario of job migration as outlined above, we note that in many cases, it is possible to acquire more machines on which jobs may be executed. Therefore, we are interested in understanding what are the consequences of allowing the online algorithm to use more colors. Our reference cost is still the optimal cost when using the initial k colors, but now we allow the online algorithm to use more colors.

The idea is not new: in fact, it was used in the very first paper on competitive analysis [12] (for cache size). Later work called this type of comparison “resource augmentation” (see, e.g., [5]). However, in this paper, still motivated by the job migration scenario, we propose an additional generalization: we also consider the case in which the additional colors come at a greater cost. We are not aware of any previous study of such a model.¹ We believe that this approach, which we call *weighted* resource augmentation, may be appropriate in other scenarios as well.

Our results. The problem of vertex recoloring is natural and applicable in many situations, but one has to bear in mind that it entails the problem of graph coloring, which is notoriously hard, computationally speaking. Following the approach of [1], we focus on polynomially-solvable instances of coloring. Specifically, in this paper we consider bipartite graphs, i.e., we assume that after all edges have arrived, the graph is 2-colorable. In [1] it was shown that the competitive ratio of online vertex recoloring is $\Theta(\log n)$ in the case of bipartite graphs, where n is the number of vertices. We study the effect of using more colors in this case. In the following, the competitive ratio is w.r.t. the 2-color optimal solution. For general bipartite graphs and deterministic algorithms it turns out that extra colors do not really help:

► **Theorem 1.** *Let \mathcal{I} be the set of instances of recoloring with two basic colors and n special colors, where recoloring by a basic color costs 1 and recoloring by a special color costs D . Then for every deterministic online recoloring algorithm there is an instance in \mathcal{I} with competitive ratio $\Omega(\min\{\log n, D\})$.*

(We note that competitive ratio D is trivial to achieve when given n special colors. See Lemma 5.) We therefore restrict our attention to a subclass of bipartite graphs, namely bipartite graphs of bounded *bond* size. A bond of a connected graph is a set of edges whose removal partitions the graphs into two connected components (alternatively, a minimal edge cut, cf. [4], and see Definition 6). Our main result in this paper is the following:

¹ In paging, this could be interpreted as follows: The offline solution uses a cache of size k , while the online solution has the cache of size k and an additional cache whose *replacement cost may be greater*.

► **Theorem 2.** *Suppose that the final graph is bipartite, with all vertex degrees at most Δ , and with bond size at most β for each of its connected components. Then there is a deterministic algorithm that uses Δ special colors of cost at most D , whose competitive ratio is $O(\log D + \beta^2)$.*

We note that in acyclic graphs, the largest bond size is 1. The following result shows that the competitive ratio of Theorem 2 cannot be improved even if we restrict the instances to have largest bond size 1.

► **Theorem 3.** *Consider instances in which the final graph is a collection of paths, and there is an infinite set of special colors, each of cost $D > 1$. The competitive ratio of any (possibly randomized) recoloring algorithm in this class of instances is $\Omega(\log D)$.*

On the other hand, if $D = 1$ (i.e., the cost of all colors is 1), another algorithm provides a way to reduce the competitive ratio at the price of using more colors.

► **Theorem 4.** *Suppose that the final graph is bipartite with bond at most β . Then for any given $\frac{1}{2 \log \beta + 3} < \epsilon \leq 1$ there is a deterministic online algorithm that uses $O(\epsilon \log n)$ special colors of unit cost, and guarantees competitive ratio $O(\epsilon^{-1})$.*

Our techniques. Our algorithms use Algorithm 1 of [1], denoted \mathcal{A} in this paper, as a black box. To facilitate our algorithms, however, we need to develop a refined analysis of Algorithm \mathcal{A} . In Lemma 17 we prove a tighter upper bound on the cost of \mathcal{A} when applied to an input sequence that satisfies a certain condition (“moderate sequences,” cf. Definition 7). Using this bound, the main idea in our algorithms is to filter the input sequence so as to make it moderate, and to feed that sequence to \mathcal{A} for simulation (which we know to have good performance). The difference between our algorithms is what to do with the other edges, and how to determine the color of vertices which may be affected by the filtering.

The approach taken by our Algorithm \mathcal{B} (whose performance is stated Theorem 2), is to use special colors, essentially in greedy fashion, to resolve conflicts that involve edges that were rejected from simulation due to the moderate sequence condition. We show that this approach can guarantee $O(\log D)$ competitiveness while using Δ additional colors. The other approach, taken by our Algorithm \mathcal{C} , is to have a hierarchical set of instantiations (simulations) of \mathcal{A} . If an edge is rejected by one instantiation, one of its endpoints is sent to the next instantiation in the hierarchy, which uses a new set of colors; and if this fails, we send that vertex to the next instantiation etc. We show that the number of edges in this hierarchy of simulations decreases exponentially as a function of a parameter of the “moderation” of the input sequence, which allows us to prove Theorem 4.

We note that the largest bond size proves to be a very useful graph parameter, as it measures, in some precise sense, how far a graph is from a tree. Our Lemma 10 gives us a tool which may be handy in other contexts as well.

Related work. As mentioned above, the problem of vertex recoloring was introduced in [1], where vertices have weights, and the cost of recoloring a vertex is its weight. It is shown that for bipartite graphs, competitive ratio of $O(\log n)$ can be achieved by a deterministic online algorithm, and that no randomized algorithm has competitive ratio $o(\log n)$. Tight bounds are also presented for $(\Delta + 1)$ coloring for randomized and deterministic algorithms where Δ denotes the maximal degree in the graph. Recently, Rajaraman and Wasim [11] considered the capacitated setting where there is a bound B on the number or weight of vertices in each color. They give “traditional” resource-augmented algorithms: the algorithms are allowed to

violate the capacity bound by a $(1 + \epsilon)$ where ϵ is an arbitrarily small constant. They also study the $(1 + \epsilon)$ -overprovisioned setting where the algorithm is allowed Δ colors and the maximum degree of the graph is bounded by $(1 - \epsilon)\Delta$.

Recoloring (or coloring with recourse) has been considered previously in the context of dynamic data structures [6, 3, 13, 10]. In these papers, no initial coloring is given and the competitive ratio is not analyzed; their measure of performance is the absolute number of recolorings. Competitive analysis is implicit in [3, 13] which is bicriteria: the arrival model is similar to ours but the final graph may be arbitrary, and the goal is to minimize both the update time and the number of colors used. In this line of work, it is assumed that the algorithm has access to an oracle that can be queried about the chromatic number of a graph. The best general result is due to Solomon and Wein [13], who give a deterministic algorithm with $O(d)$ amortized running time using $O\left(\frac{\log^3 n}{d} \chi(G)\right)$ colors. Henzinger et al. [9] give better results for bounded arboricity graphs.

Although the largest bond and maximum cut of a graph may seem superficially similar, they are quite different, both in value and complexity. For example, finding the largest graph bond in bipartite graphs is NP-hard [7], but max-cut is trivial in bipartite graphs, and it is polynomially computable in planar graphs [8]; any tree has a max-cut of size $\Omega(n)$ but largest bond size 1.

Paper organization. The remainder of this paper is organized as follows. In Section 2, we formalize the problem and introduce some notation. In Section 3, we prove our main result Theorem 2. In Section 4, we consider the uniform case and prove Theorem 4. Additional material is included in the full version: we prove our lower bounds Theorem 3 and Theorem 1, and present a slightly improved version of Algorithm \mathcal{B} . A short conclusion is presented in Section 6.

2 Problem Statement and Notation

Problem statement. We consider the following model. Initially we are given a palette P of k basic colors, and a superset $P^* \supseteq P$ of colors. The extra colors in $P^* \setminus P$ are called *special* colors. Each color j is assigned a cost $\text{cost} : P^* \rightarrow \mathbb{R}^+$, such that $\text{cost}(j) = 1$ for all basic colors, and $1 \leq \text{cost}(j) \leq D$ for all special colors, where $D \geq 1$ is some given parameter. We are also given a set of n vertices V with an initial coloring $c_0 : V \rightarrow P$ using only the basic colors. The online input is a sequence of edges e_1, e_2, \dots , where each edge is an unordered pair of distinct vertices. We define the graph G_i by $G_i = (V, \{e_1, \dots, e_i\})$. In response to the arrival of each edge e_i , the algorithm outputs a coloring $c_i : V \rightarrow P^*$ such that none of the edges in G_i is monochromatic.

Given an algorithm A , an initial coloring c_0 and an edge sequence $\sigma = (e_1, \dots, e_\ell)$, $A(c_0, \sigma)$ is a sequence of colorings c_1, \dots, c_ℓ . Define the cost of A on an instance (c_0, σ) as

$$\text{cost}_A(c_0, \sigma) = \sum_{i=1}^{\ell} \sum_{v \in V} (1 - \delta_{c_{i-1}(v)c_i(v)}) \text{cost}(c_i(v)) ,$$

where δ is the Kronecker delta. That is, whenever a vertex is recolored, the algorithm pays the cost of its new color. In this paper we assume that the input is such that the final graph can be colored using P , i.e., G_ℓ is k -colorable. Given an initial coloring c_0 , the *offline cost* of an input sequence σ with respect to the basic colors P is

$$\text{OPT}_k(c_0, \sigma) = \min \left\{ \sum_{v \in V} (1 - \delta_{c_0(v)c^*(v)}) \mid c^* : c^* \text{ is a valid } k\text{-coloring of } G_i \text{ by } P \right\}.$$

In other words, the offline cost of σ w.r.t. k is the least cost of recoloring V using basic colors so that G_i has no monochromatic edges.

Let $I = (c_0, \sigma)$ denote an instance. The competitive ratio of an algorithm A with respect to k is $\sup_I \{\text{cost}_A(I)/\text{OPT}_k(I)\}$. Let us make a quick observation about the problem.

► **Lemma 5.** *Any instance I of recoloring with k basic colors can be solved by an online algorithm using $n - k$ special colors at cost at most $2D \cdot \text{OPT}_k(I)$, where n is the number of vertices and D is the cost of a special color.*

Proof. Consider the final graph, with vertices colored by the initial coloring. Any feasible solution must recolor a vertex cover of the monochromatic edges in this graph. Therefore, denoting the size of such a minimum vertex cover by q , we have that $\text{OPT}_k(I) \geq q$. On the other hand, we can maintain in an online manner a 2-approximate vertex cover (e.g., using an online version of the algorithm of Bar-Yehuda and Even [2]), and every time a vertex enters the online cover, we recolor it with a distinct new color. The special colors never collide because each is used at most once, and hence the total cost is at most $2qD$. ◀

Additional notation.

- Given a sequence σ , $\sigma[i]$ denotes the prefix of the first i elements of σ . Given sequences σ and σ' , $\sigma \circ \sigma'$ denotes the sequence obtained by concatenating σ and σ' .
- Given a graph $G = (V, E)$ and $V' \subseteq V$, $G[V']$ denotes the graph induced by V' , i.e., the graph with vertices V' whose edges are all edges of E with both endpoints in V' .

Graph bond. The largest bond size is a graph parameter related to max cut, but it is quite different. Intuitively, the graph bond is a measure of how close the graph is to a tree. We give below a definition that generalizes the standard one (e.g., [7]) to graphs with multiple connected components.

► **Definition 6.** *Let $G = (V, E)$ be a graph with $k \geq 1$ connected components. An edge set $B \subseteq E$ is a **bond** of G if $(V, E \setminus B)$ has exactly $k + 1$ connected components. The size of the largest bond of G is denoted $\beta(G)$.*

Note that $\beta(G) = 1$ if and only if G is a forest. Also note that the size of the largest bond of a graph is monotone in its edge set. More formally, if $G = (V, E)$ and $G' = (V, E')$ are two graphs with the same vertex set, then $E \subseteq E'$ implies $\beta(G) \leq \beta(G')$.

We remark that an alternative definition for bond is a *minimal edge cut* [4]: an edge cut of $G = (V, E)$ is an edge set $B \subseteq E$ such that $(V, E \setminus B)$ has more connected components than G , and an edge cut B is *minimal* if no proper subset of B is an edge cut of G .

3 Non-Uniform Cost

In this section we prove our main result. Intuitively, we show that while the competitive ratio for general bipartite graphs is known to be $\Theta(\log n)$, by using $\Delta + 1$ additional colors of cost (at most) D , one can reduce the competitive ratio to $\Theta(\log D)$ in graphs whose largest bond is small. In the full version, we show that we can achieve the same competitive ratio with only Δ additional colors, using a somewhat more complicated algorithm.

Our algorithm uses the algorithm of Azar *et al.* [1] for bipartite graphs as a black box; henceforth, we refer to Algorithm 1 of [1] as \mathcal{A} . The basic strategy of our algorithm is as follows. The input sequence of edges σ is split in two: a subsequence denoted σ^{sim} is sent to a simulation by algorithm \mathcal{A} (which uses only the two basic colors), and the remaining edges are sent to a procedure called **recx**, which uses the additional special colors. Determining which edge goes to σ^{sim} depends on the size of the connected components of its two endpoints, and by the number of vertices already recolored by \mathcal{A} in them. The idea is to control the simulation of \mathcal{A} so that its cost remains within the range of $O(\log D)$ competitiveness, and use the special colors only once we know that we can pay for them. The bond size affects the total cost due to the edges that are not sent to the simulation.

3.1 Algorithm \mathcal{B} : Specification

To describe the algorithm, we need some notation. Recall that $\sigma[i]$ is the sequence of the first i edges of σ . We use E_i to denote the set of edges in $\sigma[i]$: $E_i = \{e_1, e_2, \dots, e_i\}$. We use σ_i^{sim} to denote the subsequence of edges sent to the simulation of \mathcal{A} in steps $1, \dots, i$, and $E[i]_{\text{sim}}$ to denote the corresponding set of edges. R_i is used to denote the set of vertices ever recolored by algorithm \mathcal{A} when executed on input σ_i^{sim} . We also define R to be $R_{|\sigma|}$.

We now specify the algorithm, using the parameters D (the maximal cost of a color), and $\alpha \in (0, 1)$, a parameter of our choice which indirectly controls how many of the edges will be sent to the simulation: the smaller α is, less edges will be sent to the simulation.

We also define the following concepts.

► **Definition 7.** Consider a sequence of edges σ and a connected component C of the graph which is edge-induced by the set of edges in σ .

- C is **small** if $|C| \leq D$ and **large** otherwise, where $|C|$ is the number of vertices in C .
- C is σ -**light** if $|R \cap C|/|C| \leq \alpha$, and σ -**heavy** otherwise, where R is the set of vertices ever recolored by algorithm \mathcal{A} on input σ .

► **Definition 8.** Consider an input sequence σ , and two values D and α .

- An input sequence σ is called (D, α) -**moderate** if for every edge $e_i = (u, v)$ in σ , it holds that if u and v are in two distinct connected components C_u and C_v with respect to e_1, \dots, e_{i-1} , then either C_u or C_v is either small or $\sigma[i-1]$ -light (i.e., no edge in σ connects two distinct components which are large and $\sigma[i-1]$ -heavy).
- The two series of subsequences σ_i^{sim} and σ_i^{exc} of σ are defined inductively as follows.

$$\sigma_i^{\text{sim}} = \begin{cases} \perp & \text{if } i = 0 \\ \sigma_{i-1}^{\text{sim}} \circ e_i & \text{if } i > 0 \text{ and } \sigma_{i-1}^{\text{sim}} \circ e_i \text{ is } (D, \alpha)\text{-moderate} \\ \sigma_{i-1}^{\text{sim}} & \text{if } i > 0 \text{ and } \sigma_{i-1}^{\text{sim}} \circ e_i \text{ is not } (D, \alpha)\text{-moderate} \end{cases}$$

$$\sigma_i^{\text{exc}} = \begin{cases} \perp & \text{if } i = 0 \\ \sigma_{i-1}^{\text{exc}} & \text{if } i > 0 \text{ and } \sigma_{i-1}^{\text{sim}} \circ e_i \text{ is } (D, \alpha)\text{-moderate} \\ \sigma_{i-1}^{\text{exc}} \circ e_i & \text{if } i > 0 \text{ and } \sigma_{i-1}^{\text{sim}} \circ e_i \text{ is not } (D, \alpha)\text{-moderate} \end{cases}$$

(Informally, e_i is appended to $\sigma_{i-1}^{\text{sim}}$ if it keeps the moderation property, and to $\sigma_{i-1}^{\text{exc}}$ otherwise.)

The pseudocode for the algorithm appears in the following page.

3.2 Analysis

We now turn to analyze the algorithm.

The interesting part about the algorithm is that the simulation of \mathcal{A} is unaware of some of the edges. We start by showing that Algorithm \mathcal{B} produces a valid coloring.

► **Lemma 9.** *After every step i , c_i is a valid coloring of the graph $G_i = (V, E_i)$.*

Proof. First observe that once a vertex becomes special, then it is always colored by a special color. Likewise, only special vertices are colored by special colors.

Algorithm \mathcal{B}

State:

- Each vertex u has an *actual* color $c(u)$ and a *simulated* color $\bar{c}(u)$. The simulated coloring \bar{c} is the coloring maintained by the simulation of \mathcal{A} . The initial actual colors are given as input, and the initial simulated colors are the initial actual colors.
- Each vertex records whether its simulated color was ever changed by the simulation of \mathcal{A} . This allows the algorithm to maintain the set R_i .
- Each vertex has an indication whether it is “special” or not. Initially all vertices are not special.

Action:

Upon the arrival of edge $e_i = (u_i, v_i)$:

- a. If $\sigma_{i-1}^{\text{sim}} \circ e_i$ is (D, α) -moderate then
 - send e_i to \mathcal{A} (which updates $\bar{c}(\cdot)$); $// \sigma_i^{\text{sim}} = \sigma_{i-1}^{\text{sim}} \circ e_i$
 - set $c(w) := \bar{c}(w)$ for every non-special vertex w .
- b. Else
 - If both u_i and v_i are not special then mark u_i as *special*.
- c. Invoke **recx** (u_i, v_i)

Procedure **recx (u, v) :**

1. If u and v are special then
 - if $e = (u, v)$ is monochromatic then
 - recolor u with a free special color. $//$ there are $\Delta + 1$ special colors
2. Else
 - If u is special then
 - if u is colored by a basic color then
 - recolor u with a free special color. $//$ first time u is colored special
3. Else
 - If v is special then $//$ this case cannot happen because of the way **recx** is used.
 - if v is colored by a basic color then
 - recolor v with a free special color.

We prove the lemma by induction on i . The base case is $i = 0$, in which $E_0 = \emptyset$ and hence any coloring is valid. For the inductive step, we first consider all edges but the new edge e_i and for those we proceed in two sub-steps. Then we consider the new edge e_i .

For all edges but e_i , if e_i is not a simulated edge and is not sent to \mathcal{A} , then the first sub-step is empty. If e_i is a simulated edge and is sent to \mathcal{A} , then the colors of some non-special vertices may change. After this sub-step, for each (old) edge: (1) if its two endpoints are non-special then the edge is not monochromatic by the correctness of \mathcal{A} ; (2) if its two endpoints are special, then their color does not change in this sub-step and the edge is not monochromatic by the induction hypothesis; (3) if one endpoint is special and the other is not, then one and only one endpoint is colored by a special color, and hence the edge is not monochromatic.

The second sub-step is the invocation of **recx** on the input edge. Following this invocation one vertex might be recolored to a special color. Clearly any (old) edge with at least one node not special remains non-monochromatic by the induction hypothesis. For an (old) edge with two special endpoints, if their color is not changed, they remain non-monochromatic by the induction hypothesis. For such an edge for which one of its endpoints changed color by **recx**, the code ensures that it is non-monochromatic after the execution of **recx**.

70:8 Colorful Vertex Recoloring of Bipartite Graphs

Now as to the edge e_i itself, we have a number of cases depending whether it is a simulated edge and the status (special or not) of its two endpoints when step i begins.

1. If its two endpoints are non-special when step i starts:
 - if e_i is a simulated edge: at the end of the first sub-step e_i is not monochromatic by the correctness of \mathcal{A} ; none of its endpoints is marked special and hence recx does not recolor any node and e_i remains non-monochromatic.
 - if e_i is not a simulated edge: One of its endpoints is marked special; recx recolors that vertex by a special color, while the other endpoint remains colored by a basic color; hence e_i is not monochromatic.
2. If its two endpoints are special when step i starts, then regardless of whether e_i is a simulated edge or not the code of recx ensures that e_i is not monochromatic.
3. If one endpoint is special and the other is not when step i starts, then regardless of whether e_i is a simulated edge or not, the status of neither vertex changes, and one of them remains colored by a special color and the other by a basic color, hence e_i is not monochromatic. ◀

We now turn to analyze the competitiveness of our algorithm. To this end, we first prove the following property about graphs with bounded bond size.

► **Lemma 10.** *Let $G = (V, E)$ be a connected graph with largest bond size at most β , and let $\{V_1, \dots, V_k\}$ be a partition of V such that the induced subgraph $G[V_i]$ is connected, for every $1 \leq i \leq k$. Then the number of edges which are not contained in any of these induced subgraphs (i.e., the number of edges with endpoints in two different parts of the partition) is at most $(k - 1)\beta$.*

Proof. Consider the multi-graph $G' = (V', E')$ obtained from G by contracting each V_i to a single node and eliminating self-loops. We shall prove that $|E'| \leq (k - 1)\beta$.

First, we note that $|V'| = k$, and that the size of the largest bond of G' is at most β . We now prove, by induction on k , that if a loop-free multigraph $G' = (V', E')$ has k nodes and bond size at most β , then $|E'| \leq (k - 1)\beta$. The base case is $k = 1$; in this case, E' is empty since G' does not contain self loops.

For the inductive step, fix any $k \geq 2$. Let v be an arbitrary node in V' , and let $\bar{V} = V' \setminus \{v\}$. We proceed according to the connectivity of \bar{V} . If $G'[\bar{V}]$ is connected, then the degree of v is at most β . Since $|\bar{V}| = k - 1$, by the induction hypothesis and the fact that the size of the largest bond of $G'[\bar{V}]$ is at most the size of the largest bond of G' (which is at most β), we have that $|E'| \leq \beta + (k - 2)\beta = (k - 1)\beta$.

Otherwise, $G'[\bar{V}]$ is disconnected. Let C_1 be the set of vertices of an arbitrary connected component of \bar{V} , and denote $C_2 = \bar{V} \setminus C_1$. Let $C_1^v = C_1 \cup \{v\}$, and $C_2^v = C_2 \cup \{v\}$. Trivially $|C_1^v| + |C_2^v| = k + 1$, and $m_1 + m_2 = |E'|$, where m_1 and m_2 are the number of edges in $G'[C_1^v]$ and in $G'[C_2^v]$, respectively. Since the largest bond of a subgraph is no larger than the largest bond of the original graph, and since the number of nodes in each of $G'[C_1^v]$ and $G'[C_2^v]$ is strictly less than k , by the inductive hypothesis we have that $m_1 \leq (|C_1^v| - 1)\beta$ and $m_2 \leq (|C_2^v| - 1)\beta$. It follows that $|E'| = m_1 + m_2 \leq (k - 1)\beta$. ◀

► **Corollary 11.** *Let $G = (V, E)$ be a graph with $\beta(G) \leq \beta$, and let C_1, \dots, C_k be a collection of disjoint subsets of vertices in V such that $G[C_i]$ is connected for all $1 \leq i \leq k$. Then $|\{(v, u) \in E : v \in C_i, u \in C_j, i \neq j\}| \leq \beta(k - 1)$, i.e., the number of edges with endpoints in two different subsets is at most $\beta(k - 1)$.*

Proof. Let V_1, \dots, V_k, V_{k+1} be any partition of V such that:

- for each $1 \leq i \leq k$, $C_i \subseteq V_i$, and $G[C_i]$ is connected; and
- V_{k+1} is exactly the set of vertices which are not connected in G to any vertex of $\bigcup_{i=1}^k C_i$.

This can be achieved by associating each vertex of $V \setminus V_{k+1}$ with the set C_i closest to it (like in a Voronoi partition). Note that we may assume without loss of generality that V_{k+1} is empty: otherwise, consider the graph $G' \stackrel{\text{def}}{=} G[V \setminus V_{k+1}]$: Clearly, by the definition of V_{k+1} , the number of edges connecting vertices in different C_i in G and G' is the same.

Since every edge between two subsets C_i, C_j connects different parts of the partition V_i and V_j , the result follows from Lemma 10, when applied to each connected component of G separately. ◀

We proceed to bound from above the cost of Algorithm \mathcal{B} . The cost consists of two parts: the cost incurred by Procedure `recx` and the cost due to the simulation of \mathcal{A} . We start by stating a number of facts due to the definition of the algorithm.

► **Fact 12.** *A vertex which is colored by a special color after step i must belong to a connected component of $E[i - 1]_{\text{sim}}$ that has more than αD vertices of R .*

Proof. Observe that a node v is colored by a special color only in procedure `recx`, when it is a special vertex. Moreover, it is marked special only when an edge $e_j = (v, u)$ arrives, and that edge is in $E \setminus E_{\text{sim}}$. It follows that there is an edge e_j , for $j \leq i$, which connects two large and heavy connected components of $E[j - 1]_{\text{sim}}$. Hence, v is part of a large connected component of $E[j - 1]_{\text{sim}}$ which has more than αD vertices of R . ◀

Recalling the sequences σ^{exc} and σ^{sim} defined in Definition 8 as a function of any sequence σ , we prove the following lemma that allows us to (indirectly) give an upper bound on the number of vertices that require “special treatment”.

► **Lemma 13.** *For any given σ , D and α , it holds that $|\sigma^{\text{exc}}| \leq \frac{\beta|R|}{\alpha D}$, where R is the set of vertices whose colors are modified when σ^{sim} is given as input to \mathcal{A} .*

Proof. In order to prove the lemma we maintain, as the input sequence σ^{sim} proceeds, sets of vertices, B_1, B_2, \dots, B_s , which are pairwise disjoint; furthermore, all vertices in each B_i belong to the same connected component of the input graph (with respect to all edges, not only those in σ^{sim}). We call the B_i sets *witness sets*. We construct the witness sets online as follows. Initially, there are no witness sets. When an edge $e = (u, v) \in \sigma$ arrives, we modify the witness sets as follows.

- I If none of $\{u, v\}$ is already in one of the existing witness sets, and if the connected component that contains e (after e is added to the graph) has at most αD vertices from R : do nothing.
- II If none of $\{u, v\}$ is already in one of the existing witness sets, and if the connected component that contains e (after e is added to the graph) has more than αD vertices from R : create a witness set which contains all the vertices in that connected component.
- III If one of $\{u, v\}$, w.l.o.g. u , is already in an existing witness set, say B , and v is not: add all the vertices of the connected component of v to B . (As we prove below in Point 4, these vertices do not yet belong to any witness set.)
- IV If both $\{u, v\}$ are already in some witness sets (possibly the same one): do nothing.

Intuitively, witness sets are created when a connected component passes the “critical mass” threshold of αD vertices from R , and vertices that do not belong to any witness set are added to the witness set they encounter, i.e., the set of the first vertex that belongs to a witness set, to which they are connected.

The following properties follow from a straightforward induction on the input sequence.

70:10 Colorful Vertex Recoloring of Bipartite Graphs

1. All the vertices of a given witness set are in the same connected component of σ . This is because a new witness set is created from connected vertices. Moreover, additions to a witness set are always in the form of vertices connected to vertices already in that witness set.
2. All the vertices in a connected component of σ^{sim} with more than αD vertices of R are in some witness set (not necessarily all in the same witness set). This follows by induction on the arrival of new edges of σ^{sim} and the actions relative to witness sets taken when this happens.
3. Every two vertices u, v which are in the same connected component of σ are either both in some witness set (not necessarily the same), or both are not in any witness set. This follows from the fact that the actions taken regarding witness sets exclude the possibility of an edge with exactly one endpoint in a witness set.
4. If a vertex w belongs to some witness set B at some point of the execution, then w belongs to B in the remainder of the execution. This follows from the fact that a vertex w becomes a member of witness set B' only in Cases II or III above.
 In Case II, w did not belong before to any witness set because both u and v did not belong to any witness set and w is in the same connected component of either u or v . Hence, by Point 3, w is not in any witness set.
 In Case III, similarly, w is connected to v and since v does not belong to any witness set, then by Point 3 w is not in any witness set.

It follows from the above properties that the total number of witness sets is at most $\frac{|R|}{\alpha D}$, because each witness set that is created requires at least αD distinct vertices of R , vertices do not change their witness set, and the sets are pairwise disjoint.

Next, we claim that when an edge e is added to σ^{exc} (i.e., it connects two distinct large-and-heavy connected components C, C' of σ^{sim}), then the vertices of C and C' already belong to two distinct witness sets. By definition, each of C and C' contains more than αD vertices of R , and hence, by Point (2) above, the endpoints of e are already associated with witness sets. To show that the endpoints of e belong to distinct witness sets, assume, by way of contradiction, that the endpoints of e belong to the same witness set. Then from Point 1 we have that C and C' are already connected in σ (but, by assumption, they are not connected in σ^{sim}). Let e_0 be the first edge (according to the order of the arrival of the edges) in σ^{exc} that completes a path from some node in C to some node in C' . Since $e_0 \notin \sigma^{\text{sim}}$, it must be the case that when e_0 arrived, both of its endpoints belonged to two large and heavy connected components of σ^{sim} . Point 2 ensures then that the vertices of C and C' are already in witness sets. Point 1 implies that these witness sets must be distinct, because before the arrival of e_0 , C and C' are not connected to each other in σ . The sets to which the vertices of C and C' belong are the same, and hence distinct, when e arrives, by Point 4. This is in contradiction to the assumption that the witness sets are the same.

Thus, an edge is added to σ^{exc} only when its endpoints are in two *distinct* witness sets. Since there can be at most $\frac{|R|}{\alpha D}$ such sets, we get from Corollary 11 and the assumption that the graph has bond size at most β , that there are at most $\frac{\beta|R|}{\alpha D}$ such edges. ◀

► **Lemma 14.** *The total cost incurred in procedure `recx` is $O(\beta^2|R|/\alpha)$, where R is the set of vertices ever recolored by \mathcal{A} .*

Proof. Procedure `recx` only recolors vertices that are marked special. For each special vertex v , the procedure `recx` recolors v from a basic color to a special color exactly once in the step when v becomes special, and might recolor v again when a new edge (v, u) arrives, and also u is marked special. Let V' be the set of vertices ever marked special, and E' the set of edges such that when they arrive both their endpoints are marked special.

A node v is marked special only when an edge $e_i = (v, u)$ arrives and $e_i \in E \setminus E_{\text{sim}}$. By Lemma 13 $|E \setminus E_{\text{sim}}| \leq O(\frac{\beta|R|}{\alpha D})$, and hence $|V'| \leq O(\frac{\beta|R|}{\alpha D})$.

Let $|V'| = \ell$, and $V' = \{u_j : 1 \leq j \leq \ell\}$. By applying Corollary 11 with sets $C_1 = \{u_1\}, \dots, C_\ell = \{u_\ell\}$, we have that $|E'| \leq O(\frac{\beta^2|R|}{\alpha D})$.

The cost of `recx` is therefore $O(|V'| + |E'|) \cdot D = O(\frac{\beta|R|}{\alpha D} + \frac{\beta^2|R|}{\alpha D}) \cdot D = O(\frac{\beta^2|R|}{\alpha})$. ◀

Next, we bound from above the cost incurred by the simulation of \mathcal{A} . We need a refined analysis of Algorithm \mathcal{A} . In the following two lemmas, we recall properties of \mathcal{A} from [1]. The first lemma relates the number of vertices recolored by \mathcal{A} to the optimal cost.

► **Lemma 15** ([1], Lemma 3.4). *Let $\text{OPT}_2[i]$ denote the offline cost for input $\sigma[i]$ when using two colors of unit cost. Then $|R_i| \leq 3 \cdot \text{OPT}_2[i]$.*

The next lemma is used to upper-bound the number of times a vertex is recolored by \mathcal{A} .

► **Lemma 16** ([1], Proposition 3.5). *Let $\mathbf{r}(i)$ denote the number of vertices recolored by \mathcal{A} in step i . There exists a subset of steps $I^+ \subseteq \{1, \dots, |\sigma|\}$ such that $\sum_{i \in I^+} \mathbf{r}(i) \geq \frac{1}{7} \sum_{i=1}^{|\sigma|} \mathbf{r}(i)$. Moreover, for every $i \in I^+$, if the arriving edge connects two distinct connected components C and C' , and \mathcal{A} recolors C , then $|C' \cup C| \geq \frac{5}{4}|C|$.*

In [1], Lemma 15 and Lemma 16 are used to show an $O(\log n)$ bound on the competitive ratio of \mathcal{A} . Here, we prove a tighter bound on the competitive ratio when the input sequence is moderate (cf. Definition 7). The proof uses amortized analysis by designing an appropriate charging scheme.

► **Lemma 17**. *If σ is a (D, α) -moderate input sequence, then $\text{cost}_{\mathcal{A}}(\sigma) \leq O(|R| \log D / (1 - \alpha))$.*

Proof. Let R_i be the set of vertices recolored by \mathcal{A} so far after it processed the i th input edge. Let I^+ be the subset of steps given by Lemma 16; it suffices to bound from above the cost incurred by \mathcal{A} on steps in I^+ . We will do this via a charging scheme.

The high-level intuition behind the charging scheme is as follows. Let $i \in I^+$, and suppose e_i connects components C_i and C'_i , and that in response, \mathcal{A} recolors C_i (recall that we consider bipartite graphs and two colors, hence a monochromatic edge must connect two distinct connected components). Our goal is to charge a the load of 1 on certain vertices and show that (1) the total charge in each step i is at least a constant fraction of $|C_i \cap R|$ or $|C'_i \cap R|$ and (2) every vertex is charged $O(\log D)$ times in total over the course of the input sequence. There are *four* cases to consider and each case will use a different type of charging. (1) The first is when C_i is small, i.e. $|C_i| \leq D$. In this case, we charge every vertex in C_i . Lemma 16 implies that a vertex can only be charged by this charging type a total of $O(\log D)$ times. (2) The second case is when C_i is large and light. In this case, we charge to the vertices of C_i that are newly recolored by the algorithm. Since C_i is light, the number of vertices that are recolored for the first time by \mathcal{A} in this step is sufficiently large. Since this type of charging charges newly recolored vertices, a vertex can be charged by this charging type at most once.

The remaining possibilities concern case (3) when C_i is large and heavy. We consider two subcases. (3.1) The first subcase is when C_i is large and heavy and C'_i is small and heavy. In this case, we charge to $C'_i \cap R$. Lemma 16 implies that $|R \cap C'_i|$ is at least an α fraction of $|C'_i|$. Moreover, since the vertices in C'_i are part of a large component after this step, a vertex can be charged by this charging type at most once. The final and most interesting case is (3.2) when C'_i is light. To handle this case, we maintain, for each (maximal) connected component C that exists at a given time, a subset $X(C)$ of vertices. We will charge to the vertices of

70:12 Colorful Vertex Recoloring of Bipartite Graphs

$X(C_i)$ in this case. The sets X are defined inductively as described in the pseudocode of the charging scheme below. As we prove below, each vertex is charged by this type of charging only once, we the size of the sets X is large enough to compensate for the cost of recoloring.

Note that these cases are exhaustive as C_i and C'_i cannot be both large and heavy, otherwise the edge would not have been input to the simulation \mathcal{A} . The following summarizes in a formal manner the charging scheme described above. We then formally prove the desired properties of the charging scheme as stated informally above.

Charging scheme

- Initially, $X(\{v\}) = \{v\}$ for every vertex v .
- When a new edge $e_i = (u_i, v_i)$ arrives connecting components C_i and C'_i :
 1. Suppose \mathcal{A} recolors C_i
 - a. Charge vertices as follows:
 - i. If C_i is small, then charge 1 to each vertex in C_i
 - ii. If C_i is large and light, then charge 1 to each vertex in $C_i \setminus R_{i-1}$, i.e., the vertices of C_i that were never recolored before step i .
 - iii. If C_i is large and heavy, C'_i small and heavy, then charge 1 to each vertex in $R_{i-1}(C'_i)$
 - iv. If C_i is large and heavy, C'_i light, then charge 1 to each vertex in $X(C_i)$.
 - b. If C'_i is light, then $X(C_i \cup C'_i) = (C_i \cup C'_i) \setminus R_i$; else $X(C_i \cup C'_i) = X(C_i) \cup X(C'_i)$
 2. If \mathcal{A} does not recolor, then $X(C_i \cup C'_i) = X(C_i) \cup X(C'_i)$

Observe that every vertex charged is either in C_i , the component being recolored, or was previously recolored (case 1(a)iii). Thus every vertex that is charged is in R . It now remains to show that in each step, the number of vertices being charged is at least a constant fraction of the number of vertices being recolored – i.e. the total charge received by the vertices can pay for the recoloring cost – and that each vertex is not charged more than $O(\log D)$ times.

▷ **Claim 18.** Consider a connected component C that exists after an arbitrary step i , and the set $X(C)$ defined by the charging scheme. For every $C \in \mathcal{P}$, we have $|X(C)| \geq |C|(1 - \alpha)/5$.

Proof. We prove this claim by induction on $|C|$, the base case being $|C| = 1$. By definition for C , such that $|C| = 1$, $X(C) = C$ and hence the claim hold, since $0 < \alpha < 1$. For the inductive step, suppose that $C = A \cup A'$, and that $|X(A)| \geq |A|(1 - \alpha)/5$ and $|X(A')| \geq |A'|(1 - \alpha)/5$. If $X(C) = X(A) \cup X(A')$ (which occurs in two distinct cases in the specifications of the charging scheme) then, since $A \cap A' = \emptyset$,

$$|X(C)| = |X(A)| + |X(A')| \geq (|A| + |A'|)(1 - \alpha)/5 = |C|(1 - \alpha)/5.$$

Otherwise, $X(C) = (A \cup A') \setminus R_i$. Then, when e_i arrived, \mathcal{A} recolored one of the two connected components A, A' and the other connected component was light. Suppose w.l.o.g. that \mathcal{A} recolored A , and A' was light. Then, we have

$$C \setminus R_i \subseteq A' \setminus R_i = A' \setminus R_{i-1}.$$

By the lightness of A' , we have that $|A' \setminus R_{i-1}| \geq |A'|(1 - \alpha) \geq |C|(1 - \alpha)/5$ by Lemma 16. This concludes the induction argument. ◁

▷ **Claim 19.** Suppose \mathcal{A} recolors component C_i at step i . Then the number of vertices being charged is at least $\Omega(|C_i|(1 - \alpha))$.

Proof. We show that the number of vertices being charged is at least $|C|(1 - \alpha)/20$ in every case. In case 1(a)i, we charge C so this clearly holds. In case 1(a)ii, lightness of C implies that $|C \setminus R_{i-1}| \geq |C|(1 - \alpha)/5 \geq |C|(1 - \alpha)/20$. In case 1(a)iii, we have

$|R_{i-1}(C')| \geq |C'|(1 - \alpha)/5 \geq |C'|(1 - \alpha)/20$ where the first inequality is due to C' being heavy and the second due to Lemma 16. For case 1(a)iv, we get that $|X(C)| \geq |C|(1 - \alpha)/5$ from Claim 18. \triangleleft

\triangleright **Claim 20.** Every vertex is charged at most $O(\log D)$ times.

Proof. By Lemma 16, the number of times a vertex v is charged due to case 1(a)i is $O(\log D)$. The number of times a vertex v is charged due to line 1(a)ii is at most once since it is recolored and now belongs to R_i . The number of times it can be charged due to 1(a)iii is at most once since it now belongs to a large component. The number of times it can be charged due to 1(a)iv is at most once since afterwards it belongs to R_i and thus cannot be in $X(A)$ for any future component, due to the way the set X is modified in 1b. \triangleleft

Combining the fact that only vertices in R can be charged and Claims 19 and 20, we are done proving Lemma 17. \blacktriangleleft

We are now ready to conclude the analysis of Algorithm \mathcal{B} .

Proof of Theorem 2. By Lemma 14, the total cost incurred by recoloring using a special color is $O(|R| \cdot \frac{\beta^2}{\alpha})$. Next, note that by the code the sequence given as input to \mathcal{A} , σ^{sim} , is (D, α) -moderate. It therefore follows from Lemma 17 that the total cost due to the simulation of \mathcal{A} is $O(|R| \log D / (1 - \alpha))$. Since by Lemma 15 the optimal cost is $\Omega(|R|)$, we are done by picking, say, $\alpha = 1/2$. \blacktriangleleft

4 Uniform cost

In this section, we still consider vertex recoloring of bipartite graphs, but now in the setting where that all colors, basic and special, have the same cost. We thus present an algorithm which has better competitive ratio if more than two colors are available, covering the spectrum between an $O(\log n)$ competitive ratio with no special colors, and a $O(\log \beta)$ competitive ratio with $O(\log_\beta n)$ special colors, where β is an upper bound on the size of graph bonds. Specifically, we prove Theorem 4, reproduced below.

\blacktriangleright **Theorem 4.** *Suppose that the final graph is bipartite with bond at most β . Then for any given $\frac{1}{2 \log \beta + 3} < \epsilon \leq 1$ there is a deterministic online algorithm that uses $O(\epsilon \log n)$ special colors of unit cost, and guarantees competitive ratio $O(\epsilon^{-1})$.*

The idea is to use $O(\epsilon \log n)$ instances of Algorithm \mathcal{A} (Alg. 1 of [1]) for recoloring bipartite graphs.

High-Level Description. We push the ideas of simulation and promoting vertices to be “special”, used in algorithm \mathcal{B} , even further. Roughly speaking, the main idea of the algorithm is to create multiple (possibly overlapping) subinstances of the problem, such that the input sequence of each subinstance is moderate, as in Definition 7. Each subsequence is fed into a different instance of \mathcal{A} , which uses a distinct pair of colors. At every point in time, each vertex is associated with some instance j , which determines its actual color.

More specifically, denote the j th instance of \mathcal{A} by \mathcal{A}_j , its input sequence by σ^j , and the coloring that it maintains at any time by \mathcal{C}^j . Each vertex v is associated at any given time with one of the instances, denoted $\text{level}(v)$. Initially, $\text{level}(v) = 1$ for all $v \in V$. We define $V^j \stackrel{\text{def}}{=} \{v \in V \mid \text{level}(v) = j\}$. \mathcal{A}_j determines the coloring of V^j : $\mathcal{C}^j : V^j \rightarrow \{2j - 1, 2j\}$. Our

70:14 Colorful Vertex Recoloring of Bipartite Graphs

algorithm maintains the following invariants: (1) each subsequence σ^j is moderate; (2) for each level j , the coloring c^j is a valid coloring of V^j with respect to $G[V^j]$ where G is the graph consisting of every edge in the entire input sequence, not just those of σ^j .

Suppose that when an edge $e_i = (u_i, v_i)$ arrives, u_i and v_i have the same color. This can only happen if they have the same level j . We first check if appending e_i to σ^j is still moderate. If it is, then we append and send e_i to \mathcal{A}_j , and recolor vertices of level j accordingly. Otherwise, we choose one of them, arbitrarily, to promote to level $j + 1$.

When we promote a vertex u to level k , in order to maintain invariant (2), we attempt to append, one-by-one, each edge (u, v) in which $\text{level}(v) = k$ to σ^k , in arbitrary order. In this process, if for some edge (u, v) , we are unable to append it to σ^k without maintaining the moderate property, we promote u to the next level. Note a vertex cannot be promoted indefinitely as it eventually ends up at a level with no other vertices. Pseudocode of our algorithm is given below.

Algorithm C: Coloring a bipartite graph of largest bond size β using $O(\epsilon \log n)$ colors.

- a. Set $\text{level}(v) := 1$ for all $v \in V$.
- b. Set c^1 to be the initial vertex coloring.
- c. Fix an arbitrary initial coloring c^j for $j > 1$, say $c^j(v) = 2j$ for all $v \in V$.
- d. Let $\tau = \max\{2^{1/\epsilon}, 4\beta^2/\alpha\}$. *// α is a constant parameter in $(0, 1)$*
- e. When edge $e_i = (u_i, v_i)$ arrives:
 - (1) If $\text{level}(v_i) \neq \text{level}(u_i)$ then return; *// different colors*
 - (2) Let $j = \text{level}(v_i)$ *// $j = \text{level}(u_i)$ as well*
 - (3) If $\sigma^j \circ (e_i)$ is (τ, α) -moderate, then send e_i to \mathcal{A}_j *// e_i is appended to σ^j*
 - (4) Else invoke **promote** $(u_i, j + 1)$ *// e_i is j -excess; u_i is chosen arbitrarily*
- f. Output coloring $c(v) = c^{\text{level}(v)}(v)$ for all $v \in V$.

Procedure promote(u, k):

1. Let $V' := V^k \cup \{u\}$ *// u gets the initial color of c^k*
2. Let σ' be an arbitrary ordering of the edges $e = (u, v)$ arrived so far, such that $\text{level}(v) = k$.
3. For each edge e' in σ' :
 - a. If $\sigma^k \circ (e')$ is (τ, α) -moderate, send e' to \mathcal{A}_k *// e' is appended to σ^k*
 - b. Else invoke **promote** $(u, k + 1)$ and return *// e' is k -excess*
4. $\text{level}(u) := k$ *// promotion to level j was successful*

4.1 Analysis

▷ **Claim 21.** Let $e_i = (u, v)$. At any time $t \geq i$, if $\text{level}(u) = \text{level}(v) = k$, then $e_i \in \sigma^k$.

Proof. At any time $t \geq i$, if at the beginning of time step t , $\text{level}(u) \neq \text{level}(v)$, we are done as the levels do not change (step 5(a)). If at the beginning of time step t $\text{level}(u) = \text{level}(v) = k$ and $\sigma^k \circ e_i$ is moderate then e_i is added to σ^k (step 5(c)), and the claim holds. In the remaining case (Step 5(d)), u is promoted and v is not, hence their level will not be the same at the end of time step t . ◁

The correctness of the algorithm is now straightforward.

► **Lemma 22.** *The coloring produced by Algorithm C is correct.*

Proof. Follows from the fact that at any time (1) two vertices of different levels are not colored by the same color, (2) if an edge $e = (u, v)$ exists and $\text{level}(u) = \text{level}(v) = k$, then $e \in \sigma^k$ (3) the correctness of \mathcal{A}_k for all k . ◀

We now proceed to give an upper bound on the cost paid by Algorithm \mathcal{C} and on the number of colors it uses. Recoloring occurs in \mathcal{C} either by some \mathcal{A}_j or by procedure `promote` (which recolors to the initial color of the corresponding level). We analyze each separately.

Let us denote by $\text{OPT}_2(\sigma)$ the optimal cost of recoloring an input sequence σ , when only the two initial colors of unit cost are available; the initial coloring of the vertices is understood from the context. Regarding recoloring by some \mathcal{A}_j , we have the following. Let R_j denote the set of vertices whose color was altered by \mathcal{A}_j .

► **Lemma 23.** *For all j , $\text{cost}_{\mathcal{A}}^j(\sigma^j) = O(|R_j| \log \tau / (1 - \alpha))$.*

Proof. Follows from Lemma 17 and the fact that σ^j is (τ, α) -moderate by construction. ◀

We note that the recoloring done by the algorithm are either (1) changes of $c^k(v)$, for v such that $\text{level}(v) = k$, done by \mathcal{A}_k , or (2) change of the color of v due to the change of $\text{level}(v)$, in procedure `promote`. We now analyze the cost due to the latter; the former was analyzed in the above lemma.

We use the following definition.

► **Definition 24.** *An input edge $e = (u, v)$ is called j -excess if it caused `promote`($u, j + 1$) to be invoked, i.e., either of the two lines (1) Step e4 or (2) Step 3b of Procedure `promote`, was reached when e was processed. We denote the set of all j -excess edges by F_j .*

Clearly, the total cost due to recoloring by `promote` is at most $\sum_j |F_j|$, because each excess edge causes at most one vertex to be recolored (indirectly) when `promote` changes the level of a vertex. Then by Lemma 13 we have:

► **Lemma 25.** $|F_j| \leq \beta \frac{|R_j|}{\alpha\tau}$ for every $j \geq 1$.

We now have that the total cost of the algorithm is

$$\text{cost}_{\mathcal{C}}(\sigma) \leq \sum_{j \geq 2} |F_j| + \sum_{j \geq 1} \text{cost}_{\mathcal{A}}(\sigma^j) \leq O\left(\frac{\beta}{\alpha\tau} + \frac{\log \tau}{1 - \alpha}\right) \sum_{j \geq 1} |R_j|. \quad (1)$$

Next, we bound $\sum_{j \geq 1} |R_j|$. We will need the following lemma.

► **Lemma 26.** *Let E^j denote the set of edges in σ^j . Then for every $j > 1$, $|E^j| \leq \beta^2 \cdot \frac{|R_{j-1}|}{\alpha\tau}$.*

Proof. Let S_j be the set of vertices u to which `promote`(u, j) was applied at some point in the algorithm. Clearly, $|S_j| = |F_{j-1}|$, because each $(j - 1)$ -excess edge promotes a single vertex to level j . It follows that $|S_j| = |F_{j-1}| \leq \beta \cdot \frac{|R_{j-1}|}{\alpha\tau}$ by Lemma 25. Now, $E^j \subseteq G[S_j]$, because an edge is appended to σ^j only if both its endpoints are in V^j at that time. Since E^j is a subset of the edges of a graph with largest bond size β , we have from Corollary 11 that $|E^j| \leq \beta(|S_j| - 1)$. We therefore conclude that $|E^j| \leq \beta^2 \cdot \frac{|R_{j-1}|}{\alpha\tau}$. ◀

We can now prove that $|R_j|$ decreases exponentially with j , for large enough τ .

► **Proposition 27.** $|R_j| \leq 2\beta^2 \cdot \frac{|R_{j-1}|}{\alpha\tau}$ for every $j > 1$.

Proof. Since \mathcal{A}_j recolors a vertex only if it is incident on an input edge, we have that $|R_j| \leq 2|E^j|$. The result follows from Lemma 26. ◀

Proposition 27 immediately gives an upper bound on the number of colors used by Algorithm \mathcal{C} .

70:16 Colorful Vertex Recoloring of Bipartite Graphs

► **Corollary 28.** *With $\alpha = \frac{1}{2}$ and $1/\epsilon \geq 2 \log \beta + 3$, Algorithm \mathcal{C} uses at most $O(\epsilon^{-1} \log n)$ colors.*

Proof. Let $\gamma = \frac{2\beta^2}{\alpha\tau}$. By Step d of Algorithm \mathcal{C} , $\gamma < 1$. Clearly, $|R_1| \leq n$, and hence, by Proposition 27 we have that $|R_j| \leq n \cdot \gamma^j$. Hence, for $j > \log_{(1/\gamma)} n$ we have that $|R_j| < 1$. The result follows, since $\log(1/\gamma) = \log \tau - O(\log \beta) = \Theta(1/\epsilon)$. ◀

We can now conclude the analysis of Algorithm \mathcal{C} .

Proof of Theorem 4. Fix $\alpha = 1/2$. The number of colors is given by Corollary 28. By Inequality (1), the total cost of the algorithm is at most

$$O\left(\frac{\beta}{\alpha\tau} + \frac{\log \tau}{1-\alpha}\right) \sum_{j \geq 1} |R_j|.$$

Since $\epsilon > \frac{1}{2 \log \beta + 3}$ and $\alpha = 1/2$, we get that $\tau = 2^{1/\epsilon}$ and so $\frac{\beta}{\alpha\tau} + \frac{\log \tau}{1-\alpha} \leq O(1/\epsilon)$. Moreover, Proposition 27 implies that $\sum_{j \geq 1} |R_j| \leq O(1)|R_1|$. By Lemma 15, we have that $|R_1| \leq O(1)\text{OPT}_2(\sigma^1) \leq O(1)\text{OPT}_2(\sigma)$. Thus, $\sum_{j \geq 1} |R_j| \leq O(1)\text{OPT}_2(\sigma)$. Therefore, the total cost of the algorithm is at most $O(1/\epsilon)\text{OPT}_2(\sigma)$, as desired. ◀

5 Lower Bounds

In this section we prove two lower bounds on online recoloring. The first, Theorem 3, says that no matter how many extra colors we use, if they cost D , then the best competitive ratio one can hope for is $O(\log D)$ even if the underlying graph is acyclic. The second lower bound, Theorem 1, shows that if there is no upper bound on the bond size of the graph, the competitive ratio of any online recoloring algorithm is at least $\Omega(\min(\log n, D))$.

5.1 Acyclic Graphs

In this subsection we show that the competitive ratio of Algorithm \mathcal{B} cannot be improved, even by randomized algorithms, as stated by Theorem 3 reproduced below.

► **Theorem 3.** *Consider instances in which the final graph is a collection of paths, and there is an infinite set of special colors, each of cost $D > 1$. The competitive ratio of any (possibly randomized) recoloring algorithm in this class of instances is $\Omega(\log D)$.*

Proof. By Yao's Principle, it suffices to show the existence of an input distribution on which the expected cost of any deterministic algorithm is $\Omega(\log D)$ times the optimal cost. We describe such a distribution. Initially, there are n vertices, each colored by one of the two basic (unit cost) colors. Without loss of generality, let us assume that n is a multiple of $\geq D/\log D$ vertices. For ease of explanation, we shall assume that the vertices are arranged left-to-right on a line. The input sequence is constructed in H phases, where $H \stackrel{\text{def}}{=} \lceil \log D - \log \log D \rceil$. In phase $1 \leq h \leq H$, edges arrive so that at the end of the phase, each connected component is a path with 2^h vertices. Specifically, for each $1 \leq j \leq n/2^h$, an edge connecting random endpoints of the $(2j-1)$ -th leftmost path and the $2j$ -th leftmost path is inserted. Note that after the last phase, we have a set of paths of length 2^H vertices each. Note that for every phase h , the set of connected components at the end of the phase is predetermined; it is only the order in which the vertices in a connected component appear in the corresponding path that is randomized.

Fix an arbitrary deterministic recoloring algorithm A . We now bound from below the expected cost of A on the input distribution defined above. We first prove the following lemma.

► **Lemma 29.** *Let A^* be an arbitrary deterministic algorithm that uses only the two basic colors. Let k be any integer power of 2 not greater than 2^H . Consider a randomized input sequence as described above and a connected component consisting of k vertices after $\log k$ phases. Let σ be the restriction of the input sequence to the connected component. Then we have $\Pr[\text{cost}_{A^*}(\sigma) \geq k(\log k - 2)/8] \geq 1/2$.*

Proof. In phase $h > 1$, there are $k/2^h$ merges of paths each containing 2^{h-1} vertices. Observe that with probability exactly $1/2$ a given merge is monochromatic (i.e., the merging new edge connects two vertices of the same color) independently of all other merges (in the same and other phases). This is because for a given merge, each of the two merged paths have one end colored with each of the two basic colors. Further observe that a monochromatic merge at phase $h > 0$ results in recoloring of exactly 2^{h-1} vertices. Let Z_h be the indicator random variable indicating whether at least half of the merges in phase h are monochromatic. Clearly, if $Z_h = 1$ then the algorithm recolors at least $\frac{1}{2} \frac{k}{2^h} \cdot 2^{h-1} = k/4$ vertices in phase h . Moreover, for any $1 < h < \log k$, even when conditioned on all Z_j (except Z_h), it holds that $\Pr[Z_h = 1] = 1/2$.

It therefore follows that $\text{cost}_{A'} \geq \sum_{h=2}^{\log k} Z_h \cdot k/4$, and

$$\Pr\left[\text{cost}_{A^*} \geq \frac{k(\log k - 2)}{8}\right] \geq \Pr\left[\sum_{h=2}^{\log k} Z_h \cdot \frac{k}{4} \geq \frac{k(\log k - 2)}{8}\right] \geq \frac{1}{2},$$

where the last inequality follows from the fact that $\sum_{h=2}^{\log k-1} Z_h$ is a sum of $(\log k - 2)$ independent Bernoulli random variables with parameter $1/2$, and so their sum is at least $(\log k - 2)/2$ with probability $1/2$. ◀

We now continue with the proof of Theorem 3. Consider one of the paths of length 2^H that exist after the last phase. We give a lower bound on the expected cost of A for re-coloring the vertices of that path. Let A_D be the indicator random variable indicating if A uses at least once an extra color (of cost D) for recoloring the vertices of the given path. There are two cases to consider.

- (1) If $\Pr[A_D] \geq 1/4$, then $\Pr[\text{cost}_A \geq D] \geq 1/4$, and hence $E[\text{cost}_A] \geq D/4$.
- (2) Otherwise, $\Pr[A_D] < 1/4$. In this case, define a deterministic online recoloring algorithm A' that uses only the basic colors as follows. A' mimics A so long as A uses basic colors. Once A uses a special color (if it does), A' stops mimicking A and continues to recolor (properly), using only the basic colors.

We make the following observations.

- $\Pr[\text{cost}_{A'} \geq 2^H(H - 2)/8] \geq 1/2$ by Lemma 29.
- The probability that A' and A diverge in their colorings on the path under consideration (and hence *do not have* the same cost for that path) is less than $1/4$.

It follows that with probability more than $1/4$, $\text{cost}_A = \text{cost}_{A'}$, and that cost is at least

$$\begin{aligned} \text{cost}_A &\geq 2^H \cdot \frac{H - 2}{8} \\ &= (2^{\lceil \log D - \log \log D \rceil}) \cdot \frac{\lceil \log D - \log \log D \rceil - 2}{8} \\ &\geq \frac{D}{\log D} \cdot \frac{\log D - \log \log D - 2}{8} \\ &= \Omega(D). \end{aligned}$$

Therefore, in both cases the expected cost of A for recolorings of the vertices in the path that we consider is $\Omega(D)$.

The bound above is the cost for a single component of length 2^H . By linearity of expectation, the expected cost of A for all components is $\frac{n}{2^H} \cdot \Omega(D)$. On the other hand the cost of the optimal algorithms is $O(n)$ for any input sequence. Hence we have a lower bound on the competitive ratio of $\Omega\left(\frac{(n/2^H) \cdot D}{n}\right) = \Omega\left(\frac{D}{2^H}\right) = \Omega\left(\frac{D}{2^{\log D - \log \log D}}\right) = \Omega(\log D)$. ◀

5.2 Bipartite Graphs with Unbounded Bond Size

In this subsection we show that no deterministic online recoloring algorithm can have good competitive ratio for all bipartite graphs with large bond, even when special colors are available. In fact, using special colors cannot improve the competitive ratio beyond the easy $O(D)$ bound of Lemma 5 if the bond may be large.

► **Theorem 1.** *Let \mathcal{I} be the set of instances of recoloring with two basic colors and n special colors, where recoloring by a basic color costs 1 and recoloring by a special color costs D . Then for every deterministic online recoloring algorithm there is an instance in \mathcal{I} with competitive ratio $\Omega(\min\{\log n, D\})$.*

Proof. We construct an adversarial input sequence that consists of phases. At any point in time, a vertex is said to be *special* if the algorithm has recolored it at some earlier point to a special color; the vertex is called *basic* otherwise. Furthermore, at each phase we characterize the connected components as *active* or *inactive*. A component is said to be active if at least half of its vertices are basic, and it is said to be inactive otherwise. Let c_1, c_2 denote the two basic colors. Initially, each vertex is given either the color c_1 or c_2 . An active component is said to be *c-dominated*, for $c \in \{c_1, c_2\}$ if at least a quarter of its vertices are colored c (a component may be both c_1 -dominated and c_2 -dominated).

The construction maintains the following invariant:

■ At the start of phase $h > 0$, every active component has 2^{h-1} vertices.

In phase h , we match active components of the same dominating color in pairs. For each such pair (C, C') of connected components of the same dominating color, we add a perfect matching between their vertex sets such that at least a quarter of the matching edges (i.e., at least 2^{h-3} edges) are monochromatic. The process ends once there is at most one active component of dominating color c_1 and at most one active component of dominating color c_2 .

We claim that the cost paid by the algorithm for the above input sequence is at least $\Omega(\min\{n \log n, nD\})$. To prove the claim, let k be the number of special vertices at the end of the input sequence, and let M be the number of edges that were monochromatic when they arrived. Clearly, the cost for the algorithm is at least $\max\{kD, M\} \geq (kD + M)/2$. Now, if $k \geq n/4$, then the algorithm's cost is $\Omega(nD)$ and the claim holds. Otherwise, $k < n/4$, and hence, by the definition of active components, at the end of the execution, there are at most $2k < n/2$ vertices in inactive components. It follows that in each phase there are at least $n/2$ vertices in active components, and therefore at least $n/4$ vertices in active components of the same dominating color, which implies that in each phase, the number of introduced monochromatic edges is $\Omega(n)$. Further, note that if $k < n/4$ then the number of phases is at least $\log n - 3$: before every phase $h \leq \log n - 2$, each connected component contains less than $2^{\log n - 3} = n/8$ vertices, and hence there are at least 4 active connected components, so there are at least two connected components dominated by the same color. It follows that the total cost of the algorithm is $\Omega(n \log n)$.

In summary, we have shown that the cost of the online algorithm for the input sequence specified above is $\Omega(\min\{n \log n, nD\})$. As the optimal cost is $O(n)$ the theorem follows. ◀

We note that the largest bond size of any active connected component C in the construction above is $\Omega(|C|)$.

6 Conclusions

In this paper we studied competitive vertex recoloring with weighted augmentation. We have shown that the competitive ratio of recoloring bipartite graphs may be reduced if the online algorithm can use additional colors, even if they are more costly than the basic two colors. Beyond the direct technical contribution, we believe that we introduced some conceptual contributions:

- The approach of weighted resource augmentation is new, as far as we know.
- The concept of the largest graph bond as a way to generalize algorithmic results quantitatively.

It seems that these ideas can be useful in dealing with other problems of competitive analysis.

References

- 1 Yossi Azar, Chay Machluf, Boaz Patt-Shamir, and Noam Touitou. Competitive vertex recoloring. *Algorithmica*, 85(7):2001–2027, 2023. doi:10.1007/s00453-022-01076-x.
- 2 Reuven Bar-Yehuda and Shimon Even. A linear-time approximation algorithm for the weighted vertex cover problem. *Journal of Algorithms*, 2(2):198–203, 1981. doi:10.1016/0196-6774(81)90020-1.
- 3 Luis Barba, Jean Cardinal, Matias Korman, Stefan Langerman, André van Renssen, Marcel Roeloffzen, and Sander Verdonschot. Dynamic graph coloring. *Algorithmica*, 81(4):1319–1341, 2019. doi:10.1007/s00453-018-0473-y.
- 4 J.A. Bondy and U.S.R. Murty. *Graph Theory*. Springer Publishing Company, Incorporated, 1st edition, 2008.
- 5 Allan Borodin and Ran El-Yaniv. *Online Computation and Competitive Analysis*. Cambridge University Press, 1998.
- 6 Bartłomiej Bosek, Yann Disser, Andreas Emil Feldmann, Jakub Pawlewicz, and Anna Zych-Pawlewicz. Recoloring Interval Graphs with Limited Recourse Budget. In *Proc. 17th Scandinavian Symposium and Workshops on Algorithm Theory (SWAT 2020)*, volume 162 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 17:1–17:23, Dagstuhl, Germany, 2020. doi:10.4230/LIPIcs.SWAT.2020.17.
- 7 Gabriel L. Duarte, Hiroshi Eto, Tesshu Hanaka, Yasuaki Kobayashi, Yusuke Kobayashi, Daniel Lokshantov, Lehilton L. C. Pedrosa, Rafael C. S. Schouery, and Uéverton S. Souza. Computing the largest bond and the maximum connected cut of a graph. *Algorithmica*, 83:1421–1458, January 2021. doi:10.1007/S00453-020-00789-1.
- 8 F. Hadlock. Finding a maximum cut of a planar graph in polynomial time. *SIAM J. Comput.*, 4(3):221–225, 1975. doi:10.1137/0204019.
- 9 Monika Henzinger, Stefan Neumann, and Andreas Wiese. Explicit and implicit dynamic coloring of graphs with bounded arboricity. *CoRR*, abs/2002.10142, 2020. arXiv:2002.10142.
- 10 Manas Jyoti Kashyop, N. S. Narayanaswamy, Meghana Nasre, and Sai Mohith Potluri. Trade-offs in dynamic coloring for bipartite and general graphs. *Algorithmica*, 85(4):854–878, 2023. doi:10.1007/s00453-022-01050-7.
- 11 Rajmohan Rajaraman and Omer Wasim. Competitive capacitated online recoloring. In Timothy Chan, Johannes Fischer, John Iacono, and Grzegorz Herman, editors, *32nd Annual European Symposium on Algorithms, ESA 2024, September 2-4, 2024, Royal Holloway, London, United Kingdom*, volume 308 of *LIPIcs*, pages 95:1–95:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024. doi:10.4230/LIPIcs.ESA.2024.95.
- 12 Daniel M. Sleator and Robert E. Tarjan. Amortized efficiency of list update and paging rules. *Comm. ACM*, 28(2):202–208, 1985. doi:10.1145/2786.2793.
- 13 Shay Solomon and Nicole Wein. Improved dynamic graph coloring. *ACM Trans. Algorithms*, 16(3):1–24, June 2020. doi:10.1145/3392724.