On approximating the achromatic number *

Guy Kortsarz† Robert Krauthgamer‡

April 16, 2001

Abstract

The achromatic number problem is to legally color the vertices of an input graph with the maximum number of colors, denoted ψ^*, so that every two color classes share at least one edge. This problem is known to be NP-hard.

For general graphs we give an algorithm that approximates the achromatic number within ratio of $O(n \cdot \log\log n/\log n)$. This improves over the previously known approximation ratio of $O(n/\sqrt{\log n})$, due to Chaudhary and Vishwanathan [CV97].

For graphs of girth at least 5 we give an algorithm with approximation ratio $O(\min\{n^{1/3}, \sqrt{\psi^*}\})$. This improves over an approximation ratio $O(\sqrt{\psi^*}) = O(n^{3/8})$ for the more restricted case of graphs with girth at least 6, due to Krysta and Loryś [KL99].

We also give the first hardness result for approximating the achromatic number. We show that for every fixed $\epsilon > 0$ there is no $2 - \epsilon$ approximation algorithm, unless $P = NP$.

*A preliminary version of this paper appeared in SODA 2001

†Open University of Israel, 16 Klauzner St., Ramat-Aviv, Israel. Email: `guyk@oumail.openu.ac.il`.

‡Department of computer science and applied mathematics, Weizmann Institute of Science, Rehovot 76100, Israel. Email: `robi@wisdom.weizmann.ac.il`. Work supported in part by a Dora Ostre memorial scholarship.

1 Introduction

A *coloring* (a.k.a. *legal coloring*) of a graph $G(V, E)$ is an assignment of colors to the graph vertices, such that whenever two vertices are adjacent, they are colored differently. A k-coloring is one that uses k colors. A coloring is called *complete* (or *achromatic*), if for every pair of distinct colors, there exist two adjacent vertices which are assigned these two colors. The *achromatic number* $\psi^*(G)$ of a graph G is the largest number k such that G has a complete k-coloring. (For known results on $\psi^*(G)$ see the surveys of Edwards [Edw97] and of Hughes and MacGillivray [HM97]).

Yannakakis and Gavril [YG80] proved that the problem of computing the achromatic number of a graph, called *the achromatic number problem*, is NP-hard. We are therefore interested in algorithms that find approximate solutions. An *approximation algorithm* with ratio $\alpha \geq 1$ (called in short an approximation ratio α) for the achromatic number problem is an algorithm that runs in polynomial time and finds for an input graph G a complete coloring with at least $\psi^*(G)/\alpha$ colors. Throughout, let n denote the number of vertices in the input graph G, let m denote the number of edges in it, and let $\psi^* = \psi^*(G)$ be its achromatic number.

1.1 Previous work

Yannakakis and Gavril [YG80] proved that the achromatic number problem is NP-hard. In [FHHM86], Farber et al. show that the problem is NP-hard on bipartite graphs. In [Bod89], Bodlaender shows that the problem is NP-hard on graphs that are simultaneously co-graphs and interval graphs. In [CE97], Cairnie and Edwards show that the problem is NP-hard on trees.

Chaudhary and Vishwanathan [CV97] give the first sublinear approximation algorithm for the achromatic problem, with an approximation ratio $O(n/\sqrt{\log n})$. They conjecture that it can be approximated within a much better ratio of $O(\sqrt{\psi^*})$, and give an algorithm with this approximation ratio of $O(\sqrt{\psi^*}) = O(n^{7/20})$ for graphs with *girth* (length of the shortest simple cycle) at least 7. For trees they give an algorithm with approximation ratio 7.

Krysta and Loryś [KL99] give an algorithm with approximation ratio $O(\sqrt{\psi^*}) = O(n^{3/8})$ for graphs with girth at least 6, proving for these graphs the conjecture of [CV97]. For trees, they improve the approximation ratio to 2.

1.2 Related problems

An *independent set* in a graph is a subset of the vertices, every two of which are non-adjacent. The size of the largest independent set in G is denoted by $\alpha(G)$.

A coloring of a graph is a partition of the vertex set into color classes, each of which is an independent set. A possible "greedy" approach for obtaining a complete coloring is to iteratively remove from the graph maximal independent sets of small size (maximality here is with respect to containment). However, the problem of finding a *minimum maximal independent set* cannot be approximated within ratio $n^{1-\epsilon}$ for any $\epsilon > 0$, unless P=NP [Hal93a].

A related notion is the *chromatic number* $\chi(G)$ of a graph G, which is the smallest number

k such that G has a (complete) k-coloring. Clearly, $\psi^*(G) \geq \chi(G)$. The chromatic number problem, i.e. computing $\chi(G)$, is also NP-hard.

In many respects, the achromatic number problem differs from the chromatic number problem. For example, when $\psi^*(G) = O(1)$, a complete coloring with $\psi^*(G)$ colors can be found in polynomial time, e.g. by guessing $\binom{\psi^*(G)}{2}$ "critical" edges (see [FHHM86] for a more efficient algorithm). In contrast, even when $\chi(G) = 3$, it is NP-hard to find a 3-coloring.

In terms of approximability, the chromatic number $\chi(G)$ appears to be understood relatively well: an algorithm with approximation ratio of $O\left(n\frac{(\log\log n)^2}{(\log n)^3}\right)$ is given in [Hal93b], and an $n^{1-\epsilon}$ hardness of approximation (for every fixed $\epsilon > 0$, assuming NP$\not\subseteq$ZPP) is shown in [FK98]. The achromatic number $\psi^*(G)$ is conjectured [CV97] to be approximable better than that of the chromatic number, namely $O(\sqrt{\psi^*})$.

1.3 Our results

Our results narrow, in three different ways, the (currently huge) gap in the approximation ratio of the achromatic number problem. We improve the known approximation ratio for general graphs (upper bound); we give the first hardness of approximation result (lower bound); and we extend the family of graphs for which the $O(\sqrt{\psi^*})$ approximation ratio (that is conjectured in [CV97]) is known to hold.

In Section 2 we give an algorithm with approximation ratio $O(n \log\log n / \log n)$ for general graphs, improving over the previous $O(n/\sqrt{\log n})$ ratio of [CV97]. The algorithm actually produces a complete $\Omega(\log\psi^* / \log\log\psi^*)$-coloring. We remark that a similar result is not known for some related problems such as the maximum independent set and the chromatic number. Namely, it is not known whether an independent set of size roughly $\log\alpha(G)$ can be found in polynomial time, or whether G can be colored in polynomial time with $2^{O(\chi(G))}$ colors.

Our approximation algorithm for general graphs uses a result of Máté [Mát81] for finding a complete $\Omega(\log\psi^* / \log\log\psi^*)$-coloring in a restricted family of graphs called irreducible graphs. (see Section 2 for the definition of reducibility). For certain bipartite graphs (those that can be significantly reduced in a sense that is described in Section 2) we devise another algorithm whose approximation ratio is significantly better.

In Section 3 we give an approximation algorithm with ratio $O(\min\{\sqrt{\psi^*}, n^{1/3}\})$ for graphs of girth at least 5. In terms of ψ^*, our ratio of $O(\sqrt{\psi^*})$ proves the conjecture of [CV97] for the special case of graphs with girth at least 5, extending the previously known result of [KL99] for (the more restricted case of) graphs with girth at least 6. In terms of n, our ratio of $O(n^{1/3})$ for graphs with girth at least 5 improves over the previously known ratios for (the more restricted cases of) graphs with girth at least 6 and 7, namely the $O(n^{3/8})$ ratio of [KL99] and the $O(n^{7/20})$ ratio of [CV97], respectively.

Our proof also gives a lower bound $\psi^* \geq m/n$ in graphs of girth at least 5 (recall that m is the number of edges in the graph), which may be of independent interest in the context of graph theory. This lower bound does not hold for graphs of girth 4, as demonstrated by a complete bipartite graph whose achromatic number is 2.

In Section 4 we show that for any fixed $\epsilon > 0$, the achromatic number problem cannot be approximated within ratio $2-\epsilon$, unless $P = NP$. No hardness of approximation was previously known for this problem.

In particular, our reduction shows that it is NP-hard to decide whether a graph has a complete coloring with all color classes of size exactly 2. In contrast, for a complete coloring with all color classes of size exactly 1, the graph has to be a complete graph (i.e. a clique), and so the corresponding decision problem is in P.

1.4 Preliminaries

For the algorithms, we may assume that $\psi^*(G)$ is known, e.g. by exhaustively searching over the n possible values or by using binary search. Throughout, an algorithm is considered efficient if it runs in polynomial time.

For a subset U of the vertices, let $G[U]$ be the subgraph of G induced on U. We say that a vertex v is *adjacent* to U if v is adjacent to at least one vertex in U. Otherwise, we say that v is *independent* of U. Two subsets of vertices U, W are said to be *adjacent* if they contain a pair of vertices $u \in U, w \in W$ which are adjacent in G. We also say that U, W *share* an edge in G. In a complete coloring, every two color classes are adjacent to each other. The *girth* of a graph is the length of its shortest simple cycle.

A *partial complete coloring* of G is a complete coloring of an induced subgraph $G[U]$, namely a coloring of a subset of the vertices such that each color class is adjacent to every other color class. The next straightforward lemma is well known [Mát81, Edw97, CV97, KL99].

Lemma 1 *A partial complete coloring can be extended greedily to a complete coloring of the entire graph.*

Proof. Consider a yet uncolored vertex. It can be added to one of the independent sets unless it is adjacent to all color classes, in which case this vertex can form a new color class. $\square$

Let $G \setminus v$ denote the graph resulting from removing a vertex v (and its incident edges) from G. The next two lemmas follow from the definition of the achromatic number.

Lemma 2 $\psi^*(G) - 1 \le \psi^*(G \setminus v) \le \psi^*(G)$ *for any vertex v in the graph G.*

Lemma 3 $\binom{\psi^*}{2} \le m$ *and hence* $\psi^* \le O(\sqrt{m})$.

Semi-independent matchings

A collection M of edges of a graph G is called a *matching* if no two edges in M have a common endpoint. A matching $M = \{(x_1, y_1), \ldots, (x_k, y_k)\}$ is called *independent* if no edge of $G \setminus M$ connects two matched vertices, i.e. both $X = \{x_1, \ldots, x_k\}$ and $\{y_1, \ldots, y_k\}$ are independent sets, and for all $i \neq j$, it holds that x_i is not adjacent to y_j.

A matching $M = \{(x_1, y_1), \ldots, (x_k, y_k)\}$ is called *semi-independent* if both $X = \{x_1, \ldots, x_k\}$ and $\{y_1, \ldots, y_k\}$ are independent sets, and for all $j > i$, it holds that x_i is not adjacent to y_j. (Here we assume that the edges of M are ordered from 1 to k, and that the endpoints of each edge are also ordered. Note that x_i may be adjacent to y_j for $j < i$, which makes the difference between an independent and a semi-independent matching).

A semi-independent matching can be used to obtain a partial complete coloring, as demonstrated in the next lemma. This lemma is also used in [Mát81], and a weaker version, based on an independent matching, is used in [CV97].

Lemma 4 *Given a semi-independent matching of size $\binom{t}{2}$ in a graph, a partial complete t-coloring of the graph can be computed efficiently.*

Proof. First color $x_1, \ldots, x_{t-1}$ with colors $2, \ldots, t$, respectively, and color all their matched vertices $y_1, \ldots, y_{t-1}$ with color 1. Next color $x_t, \ldots, x_{2t-3}$ with colors $3, 4, \ldots, t$, respectively, and color all their matched vertices $y_t, \ldots, y_{2t-3}$ with color 2. Proceed similarly until for some i, x_i is colored with color t and its matched vertex y_i is colored with color $t-1$. This happens for $i = \binom{t}{2}$, as each pair of distinct colors appears in exactly one edge (x_j, y_j). The lemma follows. □

2 Algorithm for general graphs

Our main result in this section is an algorithm with approximation ratio $O(n \cdot \log\log n / \log n)$ for the achromatic number problem on general graphs. For bipartite graphs we give another algorithm whose approximation ratio is better in certain cases.

2.1 The equivalence graph

Hell and Miller [HM76] define the following equivalence relation (called the reducing congruence) on the vertex set of a graph G (see also [Edw97, HM97]). Two vertices in G are *equivalent* if they have the same set of neighbors in the graph.

Let $r_1, \ldots, r_q$ denote the different equivalence classes of G, where q is the number of equivalence classes. Let $S(r_i)$ denote the vertices that belong to the class r_i. Each vertex in $S(r_i)$ is called a *copy* of r_i. Note that two equivalent vertices cannot be adjacent to each other in G, so $S(r_i)$ forms an independent set in G.

The *equivalence graph* (also called the *reduced graph*) of G, denoted Q, is a graph whose vertices are the equivalence classes r_i, and whose edges connect two vertices r_i, r_j whenever a vertex of $S(r_i)$ is adjacent in G to a vertex in $S(r_j)$. Note that if r_i, r_j are adjacent in Q, the subgraph induced by G on $S(r_i) \cup S(r_j)$ is a complete bipartite graph. The equivalence graph can be used to obtain a complete coloring of G, as shown in the following lemma.

Lemma 5 *A partial complete t-coloring of the equivalence graph Q implies a partial complete t-coloring of G (which can be computed efficiently). Hence, $\psi^*(G) \geq \psi^*(Q)$.*

Proof. Follows easily by replacing each vertex of Q with an arbitrary copy of it from G. □

2.2 Irreducible graphs

A graph is called *irreducible* if all its equivalence classes are singletons, i.e. consist of a single vertex.

Máté [Mát81] shows that the achromatic number of an irreducible graph on n vertices is lower bounded by $\Omega(\log n / \log\log n)$. The proof of [Mát81] essentially gives an efficient algorithm for finding such a coloring, as stated in the next theorem. For completeness, we review the algorithm in Appendix A. We remark that Erdős constructed a family of graphs that nearly match this $\Omega(\log n/\log\log n)$ lower bound, as their achromatic number is $O(\log n)$, see [Mát81, Section 2] for more details.

Theorem 1 (Máte [Mát81]) *There exists a polynomial time algorithm that computes for an irreducible graph on n vertices a partial complete $\Omega(\log n/\log\log n)$-coloring.*

Corollary 6 *There exists a polynomial time algorithm that computes for a graph with q equivalence classes a partial complete $\Omega(\log q/\log\log q)$-coloring.*

Proof. Note that the equivalence graph Q is always irreducible, and hence Theorem 1 yields an algorithm that finds a partial complete $\Omega(\log q/\log\log q)$-coloring of Q. By Lemma 5 this coloring implies a partial complete $\Omega(\log q/\log\log q)$-coloring of G. □

For $q \geq n^{\Omega(1)}$, this corollary clearly yields an approximation ratio $O(n \cdot \log\log n/\log n)$. To obtain such an approximation ratio for general graphs we next deal with the case that q is small. Note that in general, q might be as small as $\log_2 \psi^*$.

2.3 Algorithm for general graphs

The next theorem is the main result of this section.

Theorem 2 *A complete $\Omega(\log \psi^*/\log\log \psi^*)$-coloring can be computed efficiently.*

To prove the theorem, it suffices to assume that $q < (\psi^*)^\alpha$ for a constant $0 < \alpha < 1$, as otherwise the theorem follows from Corollary 6 since $\log q/\log\log q = \Omega(\log \psi^*/\log\log \psi^*)$. (This assumption on q will be used only at the final step of the algorithm).

The first step of the algorithm is to remove from G the vertices whose equivalence class is of size smaller than $\psi^*/2q$. Let $\hat{V}$ denote these vertices, and then $|\hat{V}| \leq q \cdot (\psi^*/2q) = \psi^*/2$. Let $G' = G \setminus \hat{V}$ denote the resulting graph, let Q' be the equivalence graph of G', and let q' be the number of vertices in Q'.

It suffices to find a partial complete $\Omega(\log \psi^*/\log\log \psi^*)$-coloring of G', since G' is an induced subgraph of G. By considering G' instead of G, we only lose a constant factor in the achromatic number because by Lemma 2, $\psi^*(G') \geq \psi^*(G) - |\hat{V}| \geq \psi^*/2$. The advantage of G' is that all its equivalence classes are relatively large. Indeed, when $\hat{V}$ is removed from G, some equivalence classes of G are removed, and the rest either remain intact or merge with each other. Hence any equivalence class of G' is of size at least $\psi^*/2q$.

The next goal of the algorithm will be to find as many as possible independent sets of Q' (not necessarily disjoint, i.e. a vertex in Q' may belong to more than one independent set) such that there is an edge (in the graph Q') between every two of these sets. The reason is that we can later replace the vertices from Q' with their copies from G'. We can actually replace each of these vertices with a distinct copy from G' (and obtain a coloring of G'), since for each vertex of Q' there are sufficiently many copies in G'.

For a set A of vertices of Q', let $N_{Q'}(A)$ denote the neighbors of A in the equivalence graph Q', i.e. all the vertices of Q' that are adjacent to A in the graph Q'. Call two (not necessarily disjoint) subsets A, B of vertices in Q' *equivalent* if $N_{Q'}(A) = N_{Q'}(B)$.

The algorithm iteratively constructs a collection $\mathcal{S}_i$ of (not necessarily disjoint) independent sets in Q', with the property that no two sets in $\mathcal{S}_i$ are equivalent, i.e. $N_{Q'}(A) \neq N_{Q'}(B)$ for all $A \neq B \in \mathcal{S}_i$. The initial collection $\mathcal{S}_1$ consists of all the vertex subsets of size one in Q', i.e. $\mathcal{S}_1$ consists of all the sets $\{u\}$ where u is a vertex in Q'. Clearly, each set in $\mathcal{S}_1$ is an independent set and no two sets are equivalent. The next collection $\mathcal{S}_{i+1}$ is constructed from $\mathcal{S}_i$ as follows. Start with $\mathcal{S}_{i+1} = \mathcal{S}_i$ and go over all $A \in \mathcal{S}_i$ and all vertices u of Q' which are independent of A. For every combination of A and u, insert the set $A \cup \{u\}$ to $\mathcal{S}_{i+1}$, unless $\mathcal{S}_{i+1}$ already contains a set that is equivalent to it. It follows that no two sets in $\mathcal{S}_{i+1}$ are equivalent, i.e. $N_{Q'}(A) \neq N_{Q'}(B)$ for all $A \neq B \in S_{i+1}$.

We next claim that $\mathcal{S}_i$ contains an equivalent for every independent set in Q' whose size is at most i.

Lemma 7 *For any independent set A in Q' with $|A| \leq i$, there exists a set $B \in \mathcal{S}_i$ such that $N_{Q'}(A) = N_{Q'}(B)$.*

Proof. Proceed by induction on i. For $i = 1$ the lemma follows from the fact that $\mathcal{S}_1$ contains all the subsets of one vertex of Q'.

Assuming that the lemma holds for $i \geq 1$, we show that it holds for $i + 1$. Let A be an independent set of size $i + 1$ in Q'. Then $A = \hat{A} \cup \{u\}$ for some $\hat{A}$ and u with $|\hat{A}| = i$. By the induction hypothesis, there exists a set $B' \in \mathcal{S}_i$ which is equivalent to $\hat{A}$, i.e. $N_{Q'}(\hat{A}) = N_{Q'}(B')$. The set $B' \cup \{u\}$ is then equivalent to A because $N_{Q'}(B' \cup \{u\}) = N_{Q'}(\hat{A} \cup \{u\}) = N_{Q'}(A)$. Since $B' \in \mathcal{S}_i$ we conclude that $\mathcal{S}_{i+1}$ contains a set that is equivalent to $B' \cup \{u\}$ and thus to A, as desired. □

The algorithm iteratively computes the collection $\mathcal{S}_i$ for $i = 1, 2, \ldots$ until the first time that $|\mathcal{S}_i| \geq \psi^*/2$. The following lemma guarantees that this happens within q' iterations, where clearly $q' \leq n$.

Lemma 8 $|\mathcal{S}_{q'}| \geq \psi^*(G') \geq \psi^*/2$.

Proof. An optimal complete coloring of G' uses at least $\psi^*/2$ colors. Each color class in this coloring is an independent set of G', and hence defines an independent set in Q'. No two of these independent sets of Q' are equivalent because their corresponding color classes in G' share at least one edge. Each of these independent sets is of size at most q' because there are q' vertices in the graph Q' and there is no need to take a vertex of Q' more than once. So by Lemma 7, $\mathcal{S}_{q'}$ contains an equivalent for each of these independent sets. It follows that $|\mathcal{S}_{q'}| \geq \psi^*(G') \geq \psi^*/2$. □

Given the collection $\mathcal{S}_i$ with $|\mathcal{S}_i| \geq \psi^*/2$, define the following graph $\mathcal{G}_i$ whose vertex set is $\mathcal{S}_i$ (i.e., each vertex in $\mathcal{G}_i$ is an independent set in Q'). Two vertices are joined by an edge in the graph $\mathcal{G}_i$ if the two corresponding independent sets share an edge in Q'. It is not difficult to see that the graph $\mathcal{G}_i$ is irreducible. Indeed, every two of its vertices correspond to two sets $A, B \in \mathcal{S}_i$ that are not equivalent in Q'. Thus, there exists a vertex u_j of Q' which is independent of A but adjacent to B or vice-versa. Since the set $\{u_j\}$ (or a set equivalent to it) belongs to $\mathcal{S}_i$, it corresponds to a vertex of $\mathcal{G}_i$ which is adjacent to exactly one of A and B.

Finally, apply Theorem 1 on $\mathcal{G}_i$, and compute for it a partial complete coloring with $\Omega(\log|\mathcal{S}_i|/\log\log|\mathcal{S}_i|)$ colors. Since $|\mathcal{S}_i| \geq \psi^*/2$, the number of colors is $\Omega(\log\psi^*/\log\log\psi^*)$. Each color class is an independent set in $\mathcal{G}_i$ and thus corresponds to an independent set of Q'. Every two color classes in $\mathcal{G}_i$ share an edge (in $\mathcal{G}_i$), and thus their corresponding independent sets in Q' share an edge (in Q'). Now replace each vertex of these $\Omega(\log\psi^*/\log\log\psi^*)$ independent sets of Q' with a distinct copy of it from G'. Recall that each vertex of Q' has at least $\psi^*/2q > (\psi^*)^{1-\alpha}/2 \gg \log\psi^*/\log\log\psi^*$ copies in G'. By replacing each vertex of Q' with a distinct copy of it from G' we therefore obtain a partial complete $\Omega(\log\psi^*/\log\log\psi^*)$-coloring of G', which concludes the proof of Theorem 2.

Theorem 3 *The achromatic number problem can be approximated within ratio of* $O(n\cdot\log\log n/\log n)$.

Proof. Follows from the $O(\psi^*\cdot\log\log\psi^*/\log\psi^*)$ approximation ratio of Theorem 2 since $\psi^* \leq n$. □

2.4 Algorithm for bipartite graphs

We give an algorithm for bipartite graphs that is based on the algorithm for general graphs, but does not use the result of Máté. This algorithm obtains an improved approximation ratio compared to Corollary 6 when q, the number of equivalence classes of G, is not too large.

Theorem 4 *The achromatic number of a bipartite graph can be approximated within ratio of* $O(\max\{q,\sqrt{\psi^*}\})$.

Proof. As in the algorithm of Theorem 2 for general graphs, we first find a collection $\mathcal{S}_i$ with $|\mathcal{S}_i| \geq \psi^*/2$ (note that no restriction on q is needed for this part). Instead of the final step which constructs the graph $\mathcal{G}_i$ and applies on it Theorem 1, we exploit the bipartiteness of G, as follows.

G is bipartite, and hence G' is bipartite. Q' is also bipartite, since each equivalence class of G' is a subset of vertices all of which are from the same side of G'.

Let W^1, W^2 denote the two sides of Q', and "project" $\mathcal{S}_i$ on each side W^j, as follows. Start with $\mathcal{T}^j = \emptyset$ and go over all $A \in \mathcal{S}_i$. For every such A, insert $A \cap W^j$ to $\mathcal{T}^j$, unless $\mathcal{T}^j$ already contains a set that is equivalent to it. It follows that no two sets in $\mathcal{T}^j$ are equivalent.

By definition, $\mathcal{S}_i$ is closed, up to equivalence, under taking non-empty subsets, i.e. if $\mathcal{S}_i$ contains a set A then for every non-empty $B \subset A$, there is some set equivalent to B in $\mathcal{S}_i$. Therefore, every set in $\mathcal{S}_i$ is equivalent to either a set in $\mathcal{T}^1$, or a set in $\mathcal{T}^2$, or the union of a set from $\mathcal{T}^1$ and a set from $\mathcal{T}^2$. Since no two sets in $\mathcal{S}_i$ are equivalent, we conclude that $|\mathcal{S}_i| \leq |\mathcal{T}^1| + |\mathcal{T}^2| + |\mathcal{T}^1|\cdot|\mathcal{T}^2|$, and hence either $|\mathcal{T}^1|$ or $|\mathcal{T}^2|$ is at least $\Omega(\sqrt{|\mathcal{S}_i|})$.

Let $\mathcal{A}$ be the larger of the two sets $\mathcal{T}^1, \mathcal{T}^2$, and thus $|\mathcal{A}| = \Omega(\sqrt{|\mathcal{S}_i|}) = \Omega(\sqrt{\psi^*})$. Without loss of generality, we assume $\mathcal{A} = \mathcal{T}^1$. Observe that every set in $\mathcal{A}$ is an independent set in Q'. We will now apply on $\mathcal{A}$ a divide-and-conquer technique so that every two sets in it will share an edge.

Similarly to the well-known Quicksort algorithm, choose a vertex $p \in W^2$ to be a *pivot* element, and separate the collection $\mathcal{A}$ into non-empty $\mathcal{A}^+$ and $\mathcal{A}^-$, where $\mathcal{A}^+$ consists of the sets that are adjacent to p (in Q'), and $\mathcal{A}^-$ consists of the sets that are independent of p (in

Q'). (We show below that there always exists such a pivot $p \in W^2$, which can be found using exhaustive search). Add p to every set $B \in \mathcal{A}^-$, i.e. those sets that are not adjacent to p. Possibly, a set B already contains p in which case the operation has no effect. Observe that now every set $B \in \mathcal{A}^-$ shares an edge with every set $C \in \mathcal{A}^+$, because $p \in B$ and C is adjacent to p. For each of $\mathcal{A}^+, \mathcal{A}^-$ apply the same procedure recursively, unless every two subsets in it already share an edge.

The resulting sets are independent sets, since a vertex p is added to a set B only if B is independent p, and initially every set in $\mathcal{T}^1 \subseteq \mathcal{S}_i$ is an independent set. In addition, every two of the resulting sets share an edge, since either (i) the two sets are separated at some point in the recursion, and then the pivot guarantees that they share an edge; or (ii) the two sets are never separated in the recursion, and then the stopping condition for the recursion guarantees that they share an edge.

We now show that if a collection $\hat{\mathcal{A}}$ contains two sets $\hat{B}$ and $\hat{C}$ which share no edge, then there exists a pivot $p \in W^2$ that separates $\hat{\mathcal{A}}$ into two non-empty parts. In the beginning of the recursive process the two sets $\hat{B}$ and $\hat{C}$ were $B \in \mathcal{T}^1$ and $C \in \mathcal{T}^1$, respectively. The recursive process may only add vertices to a set, and so $B \subseteq \hat{B}$ and $C \subseteq \hat{C}$. Since B and C are two different sets in $\mathcal{T}^1$, they are not equivalent (in Q'), and hence there is a vertex p (in Q') that is adjacent to exactly one of B, C. But Q' is bipartite and both $B, C \in \mathcal{T}^1$ contain only vertices from W^1, so $p \in W^2$. We claim that p is also adjacent to exactly one of $\mathcal{B}, \mathcal{C}$, and is hence a pivot that separates $\hat{\mathcal{A}}$ into two non-empty parts. Indeed, the vertices of $\hat{B} \setminus B$ and of $\hat{C} \setminus C$ are added as pivots at some point in the recursion. They must be from W^2 and are thus independent of $p \in W^2$. We conclude p is adjacent to exactly one of $\hat{B}, \hat{C}$, as claimed.

We thus find $\Omega(\sqrt{\psi^*})$ independent sets in Q', every two of which share an edge. Now replace each vertex of these independent sets with a distinct copy of it from G'. Recall that each vertex of Q' has at least $\psi^*/2q$ copies in G', so we can obtain a partial complete $\min\{\psi^*/2q, \Omega(\sqrt{\psi^*})\}$-coloring of G', which concludes the proof of Theorem 4. $\square$

3 Algorithms for graphs with girth at least 5

In this section we give an algorithm with approximation ratio $O(\min\{n^{1/3}, \sqrt{\psi^*}\})$ for the achromatic number problem on graphs with girth at least 5. We note in passing that for every such graph, $\psi^* \geq m/n$.

A *star* is a graph in which one of its vertices, called the *head* of the star, is connected to all other vertices, called the leaves of the star. Usually, it is also required that the leaves of the star are not connected to each other, but for graphs of girth larger than 3, this follows immediately.

3.1 Partial complete m/n-coloring

Theorem 5 *For any graph of girth at least 5, a partial complete m/n-coloring can be computed in polynomial time, and in particular, $\psi^*(G) \geq m/n$.*

Proof. Iteratively remove from the graph any vertex whose degree is smaller than m/n. The graph does not turn empty as less than m edges are removed. In the remaining graph G', all degrees are at least m/n.

Consider in G' an arbitrary vertex v, its neighbors $N(v)$ and the set of vertices of distance 2 from v, denoted $N_2(v)$. Let $N(v) = \{x_1, \ldots, x_k\}$ with $k \geq m/n$. Let R_i denote the star connecting x_i to its neighbors in $N_2(v)$. Note that the stars R_i are vertex disjoint, or otherwise a cycle of length 4 (or less) is formed. In particular, for any $i \neq j$ there is no edge between the head of R_i and a leaf of R_j. Since all degrees in G' are at least m/n, each R_i has at least $m/n - 1$ leaves.

We now use the stars $R_1, \ldots, R_{\lceil m/n \rceil}$ to obtain a partial complete $\lceil m/n \rceil$-coloring (we ignore vertices outside these stars). Iteratively, for $i = 1, \ldots, \lceil m/n \rceil$, color with color i the head of the star R_i and one leaf from each of the stars R_j with $i < j \leq \lceil m/n \rceil$, so that a total of $\lceil m/n \rceil - i + 1$ vertices are colored with color i. The one leaf from each star R_j is chosen by going iteratively over $j = i+1, i+2, \ldots, \lceil m/n \rceil$, each time choosing (and coloring with color i) an arbitrary leaf of R_j which was not colored yet and is independent of all the vertices that were previously colored with color i. To see that such a leaf of R_j always exists, observe that (i) R_j contains at least $(\lceil m/n \rceil - 1) - (i-1)$ uncolored leaves; (ii) the head of R_i is independent of all the leaves of R_j; and (iii) a leaf of $R_{j'}$ for $i < j' < j$ is adjacent to at most one leaf of R_j, or otherwise a cycle of length 4 is formed. Hence, at least $(\lceil m/n \rceil - 1) - (i-1) - (j-i-1) = \lceil m/n \rceil - j + 1 \geq 1$ leaves of R_j can be chosen.

It follows that the vertices that are colored with each color i form an independent set. Moreover, each color class i contains a leaf from R_j for $i < j \leq \lceil m/n \rceil$. This leaf is adjacent to the head of R_j, which is colored by j, and thus color class i is adjacent to every color class $j > i$. Hence, this coloring is a partial complete $\lceil m/n \rceil$-coloring. By Lemma 1 this coloring can be extended to a complete coloring of the entire graph G. $\square$

Theorem 5 implies a relatively simple $O(\sqrt{n})$ approximation ratio. We further improve this ratio in the next subsections.

Theorem 6 *There is an approximation algorithm with ratio $O(\sqrt{n})$ for the achromatic problem on graph with girth at least 5.*

Proof. $\psi^* \leq O(\sqrt{m})$ by Lemma 3. If $m < n$ then $\psi^* = O(\sqrt{n})$, and any greedy complete coloring will have at least one color and hence ratio $O(\sqrt{n})$. If $m \geq n$, the algorithm of Theorem 5 finds a complete coloring with at least m/n colors, and its ratio is thus $O(n/\sqrt{m}) = O(\sqrt{n})$. $\square$

3.2 An $O(\sqrt{\psi^*})$ approximation algorithm

Theorem 7 *For every graph of girth at least 5, a partial complete $\Omega(\sqrt{\psi^*})$-coloring can be computed in polynomial time, and hence the achromatic number can be approximated within a ratio of $O(\sqrt{\psi^*})$.*

Proof. In the sequel, we describe the algorithm together with its analysis. The algorithm is also depicted more schematically in Fig. 1.

Algorithm Girth.

(1). Let $\rho \leftarrow \lfloor\sqrt{\psi^*}\rfloor$

(2). $V_h \leftarrow \{v \in V : deg(v) \geq \rho\}$

(3). If $|V_h| \geq \rho$

then return partial complete ρ-coloring of $V_h \cup N(V_h)$ using Lemma 9

(4). $G'' \leftarrow G \setminus V_h, \mathcal{R} \leftarrow \emptyset, r \leftarrow 0$

(5). While G'' contains edges

(a) Pick a non-isolated vertex u in G''

(b) Remove from G'' the vertex u and all its neighbors, and insert them to $\mathcal{R}$

(c) $r \leftarrow r + 1$

(6). If $|\mathcal{R}| \geq \rho^3/16$

then return a partial complete $\Omega(\sqrt{r})$-coloring of $G[\mathcal{R}]$ using Lemma 10.

(7). Return a partial complete $m_{\mathcal{R}}/n_{\mathcal{R}}$-coloring of $G[\mathcal{R}]$ using Theorem 5, where $m_{\mathcal{R}}$ and $n_{\mathcal{R}}$ are the number of edges and vertices in $G[\mathcal{R}]$.

Figure 1: Algorithm for graphs of girth at least 5

Let $\rho = \lfloor\sqrt{\psi^*}\rfloor$ (recall we assumed that ψ^* is known). Let V_h be the set of vertices whose degree in G is at least ρ, and let $N(V_h)$ be the set of the neighbors of V_h in G. If V_h contains at least ρ vertices, then we use Lemma 9 (below) to compute a partial complete ρ-coloring of (the subgraph induced on) $V_h \cup N(V_h)$ and the desired $\Omega(\sqrt{\psi^*})$-coloring follows.

So from now on we assume that $|V_h| \leq \rho - 1$, and let $G' \leftarrow G \setminus V_h$. By Lemma 2 we know that $\psi^*(G') \geq \psi^* - \rho + 1 \geq \psi^*/2 + 1$. G' is an induced subgraph of G, so it suffices to compute a partial complete $\Omega(\sqrt{\psi^*})$-coloring of G'.

We partition the vertices of G' into two disjoint sets $\mathcal{R}$ and $\mathcal{I}$, as follows. Initially, let $G'' \leftarrow G', \mathcal{R} \leftarrow \emptyset$, and while G'' contains edges iteratively perform the following 2 operations: (a) pick a non-isolated vertex u in G''; (b) remove from G'' the vertex u and all its neighbors and insert them to $\mathcal{R}$. Note that when removing this star, some vertices of G'' which are not removed may become isolated vertices and remain in G'' in all the iterations. Let $\mathcal{I}$ be the (possibly empty) set of vertices that remain in G'' at the end of this process (i.e. when G'' has no edges). If $\mathcal{I}$ is not empty then it is an independent set in G. Indeed, G'' is obtained from G by operations of removing vertices, so G and G'' have exactly the same edges inside $\mathcal{I}$.

Let r be the number of iterations in the above process. Note that in each iteration, we insert to $\mathcal{R}$ a star (consisting of a vertex u and its neighbors in the corresponding iteration), and hence $\mathcal{R}$ can be partitioned to r disjoint subsets, each corresponding to a star in G.

Consider the case that $|\mathcal{R}| \geq \rho^3/16$. Since all vertices in G' have degree less than ρ (in G and hence also in G'), each of the r stars (of $\mathcal{R}$) is of size at most ρ, and hence $r \geq |\mathcal{R}|/\rho \geq \rho^2/16$. Each of these r stars contains at least one leaf. We can thus use Lemma 10 (below) to compute a partial complete $\Omega(\sqrt{r})$-coloring of $G[\mathcal{R}]$, the subgraph of G that is induced on these r stars. Since $\Omega(\sqrt{r}) = \Omega(\rho) = \Omega(\sqrt{\psi^*})$, this coloring is as desired.

In the case that $|\mathcal{R}| < \rho^3/16$, we apply the algorithm of Theorem 5 on the graph $G[\mathcal{R}] = G' \setminus \mathcal{I}$. Let $n_{\mathcal{R}}$ and $m_{\mathcal{R}}$ be the number of vertices and edges, respectively, in $G[\mathcal{R}]$. Clearly, $n_{\mathcal{R}} = |\mathcal{R}| < \rho^3/16 = O((\psi^*)^{3/2})$, so let us give a lower bound on $m_{\mathcal{R}}$. The number of edges in G' is at least $\binom{\psi^*/2+1}{2} \geq (\psi^*)^2/8$ by Lemma 3. Since $\mathcal{I}$ is an independent set also in G', all the edges in G' which are adjacent to $\mathcal{I}$ are also adjacent to $\mathcal{R}$. By the upper bound ρ on the degrees in G', we get that the number of these edges is at most $|\mathcal{R}| \cdot \rho \leq \rho^4/16 \leq (\psi^*)^2/16$. It follows that the number of edges in $G' \setminus \mathcal{I}$ is $m_{\mathcal{R}} \geq \psi^{*2}/8 - \psi^{*2}/16 = \Omega(\psi^{*2})$. Hence, $m_{\mathcal{R}}/n_{\mathcal{R}} = \Omega(\sqrt{\psi^*})$, and Theorem 5 produces a partial complete $\Omega(\sqrt{\psi^*})$-coloring. □

To finish the proof of Theorem 7, we need to prove Lemma 9 and Lemma 10.

Lemma 9 *If $|V_h| \geq \rho$, a partial complete ρ-coloring of (the subgraph of G induced on) $V_h \cup N(V_h)$ can be computed in polynomial time.*

Proof. Consider the stars that each vertex of V_h defines (together with its neighbors) in G, denoted $Q_1, \ldots, Q_{|V_h|}$. Each star Q_i contains at least ρ leaves, since the degree of each vertex of V_h is at least ρ. Note that the stars $Q_1, \ldots, Q_k$ need not be disjoint. However, the heads of the stars are distinct, since they are distinct vertices in V_h.

We now proceed with coloring the stars $Q_1, \ldots, Q_\rho$ similarly to the coloring of the stars $R_1, \ldots, R_{m/n}$ in the proof of Theorem 5. Iteratively, for $i = 1, \ldots, \rho$, color with the ith color the head of the star Q_i and one leaf from each of the stars Q_j with $i < j \leq \rho$, so a total of $\rho - i + 1$ vertices are colored with color i. The $\rho - i$ leaves of the stars Q_j for $i < j \leq \rho$ are chosen iteratively for $j = i+1, i+2, \ldots, \rho$, each iteration coloring with color i a leaf of Q_j which was not colored yet and is independent of all the vertices that were previously colored with color i. To see that such a leaf of Q_j always exists, observe that (i) Q_j contains at least $\rho - (i-1)$ uncolored leaves; (ii) each vertex which was previously colored with color i is adjacent to at most one leaf of Q_j, or otherwise a cycle of length 4 is formed. Since the number of vertices which were previously colored with color i is $j - i$, at least $\rho - (i-1) - (j-i) = \rho - j + 1 \geq 1$ leaves of Q_j can be chosen.

It follows that the vertices that are colored with each color i form an independent set. Moreover, each color class i contains a leaf from Q_j for $i < j \leq \rho$. This leaf is adjacent to the head of Q_j, which is colored by j, and thus the color class i is adjacent to every color class $j > i$. Hence, this coloring is a partial complete ρ-coloring. □

Lemma 10 *If $r \geq \rho^2/16$ then a partial complete $\Omega(\sqrt{r})$-coloring of the subgraph of G induced on the r stars of $\mathcal{R}$ can be computed in polynomial time.*

Proof. Let $x_1, \ldots, x_r$ be the heads of the r stars, and choose arbitrarily one edge (x_i, y_i) from each star. This gives a matching with r edges. (Observe that each star contains at least one leaf, since the vertex u chosen is not an isolated vertex.)

We now proceed with a partial $\sqrt{r}/8$-coloring for the vertices of the matching $(x_1, y_1), \ldots, (x_r, y_r)$, as follows. (In a sense, this extends the coloring from Lemma 4.) Iteratively, for $j = 1, 2, \ldots, \sqrt{r}/8$, perform the following four operations: (a) find *candidates* to be colored by color j, which are the vertices y_i that are yet uncolored and are independent of the vertices that were already colored by j; (b) find in the induced subgraph (of G) on these candidates y_i an independent set I_j of size $\sqrt{r}/8 - j$; (c) color all the vertices of I_j with color j; (d) color each of the vertices x_i which are matched to I_j with a distinct color from the set $\{j+1, j+2, \ldots, \sqrt{r}/8\}$.

To see that operation (b) is always feasible, consider an iteration j. Less than $r/64$ of the y_i vertices are colored, because each previous iteration colors less than $\sqrt{r}/8$ of them. In addition, $j - 1 \leq \sqrt{r}/8$ vertices were already colored with j (these are x_i vertices), and each of them has less than $\rho \leq 4\sqrt{r}$ neighbors in G, so the number of y_i candidates to be colored by color j is at least $r - r/64 - 4\sqrt{r} \cdot \sqrt{r}/8 > r/3$. By Lemma 11 (below) we know that one can efficiently find an independent set of size $\sqrt{r/3}/3 \geq \sqrt{r}/8$ in the subgraph induced on these candidates. □

Lemma 11 *Let G be a graph of girth 5 on n vertices. Then an independent set of G of size $\sqrt{n}/3$ can be found in polynomial time.*

Proof. The average degree $\bar{d}$ is at most $2\sqrt{n}$ as a graph with girth 5 has at most $n\sqrt{n}$ edges, cf. [Bol78]. Turan's theorem implies that G has an independent set of size $n/(\bar{d}+1) \geq \sqrt{n}/3$. Furthermore, a simple greedy algorithm finds such an independent set, see for example [HR97]. □

3.3 An $O(n^{1/3})$ approximation algorithm

In terms of n we have the following ratio.

Theorem 8 *The achromatic number problem can be approximated within ratio of $O(n^{1/3})$ in graphs with girth at least 5.*

Proof. If $m > n^{4/3}$ then Theorem 5 gives ratio of $O(n/\sqrt{m}) = O(n^{1/3})$, as desired. If $m \leq n^{4/3}$, then $\psi^* \leq O(\sqrt{m}) = O(n^{2/3})$ and the required ratio follows from Theorem 7. □

4 Hardness of approximation

In this section we prove NP-hardness for the problem approximating the achromatic number within ratio of $2 - \epsilon$. The same reduction shows that it is NP-hard to decide whether a graph has a complete coloring with all color classes of size exactly 2 (or, alternatively, at most 3).

Our reduction uses a graphs operation called the (disjoint) union of graphs. We remark that Hell and Miller [HM92] give a tight analysis of the effect of this operation on the achromatic number from an extremal point of view.

Theorem 9 *For every fixed $\epsilon > 0$, it is NP-hard to approximate the achromatic number within ratio of $2 - \epsilon$.*

Proof. Our starting point is a reduction that creates a gap in the chromatic number, in the following way. Given an input x for an NP-complete language L, one can efficiently produce a graph $G(V, E)$ which satisfies (let $n = |V|$):

(1) if $x \in L$, then for a certain $\hat{\chi}$, the graph G can be (legally) colored with $\hat{\chi}$ colors, i.e. $\chi(G) \leq \hat{\chi}$, so that all color classes will be of equal size $n/\hat{\chi}$.

(2) if $x \notin L$, then every independent set in G is of size at most $\hat{\alpha}$, i.e. $\alpha(G) \leq \hat{\alpha}$, for a certain $\hat{\alpha} < n/\hat{\chi}$. (Hence $\chi(G) \geq n/\hat{\alpha}$).

This reduction creates a gap of $n/\hat{\alpha}\hat{\chi} > 1$ in the chromatic number of G. Lund and Yannakakis [LY94] have shown such reductions (see also [FK98]). The gap that is created in these reductions is an arbitrarily large constant. (We remark that for every fixed $\delta > 0$, a larger gap of $n^{1-\delta}$ can be obtained using randomized reductions).

Lemma 12 *For every fixed $\epsilon > 0$, there exists a reduction as above with gap $n/\hat{\alpha}\hat{\chi} \geq 1/\epsilon$.*

To produce a gap for the achromatic number problem, start with the graph G from the reduction of Lemma 12, and construct from it a graph H on $n(\hat{\chi}+1)$ vertices, as follows. H is the union of $\bar{G}$, the edge complement of G, and a complete $\hat{\chi}$-partite graph on the vertex set W which consists of $\hat{\chi}$ parts $W_1, \cdots, W_{\hat{\chi}}$, each of size $n/\hat{\chi}$. In other words, the vertex set of H is $V \cup W_1 \cup \cdots \cup W_{\hat{\chi}}$; its edges inside V form the graph $\bar{G}$; a vertex from a set W_i is connected by an edge to a vertex of a set W_j if and only if $i \neq j$; and there are no other edges (e.g. between V and W).

We first show that if $x \in L$ then $\psi^*(H) \geq n$. If $x \in L$ then there is a coloring of G with $\hat{\chi}$ colors, so that all color classes are of equal size $n/\hat{\chi}$. Let V_i denote the ith color class in this coloring, for $i = 1, \ldots, \hat{\chi}$. Each color class V_i is an independent set in G and thus a clique in $\bar{G}$ and in H. Let us pair each vertex of V_i with a distinct vertex of W_i (observe that $|V_i| = n/\hat{\chi} = |W_i|$), so there would be exactly n pairs, one for each vertex of V.

We claim that considering each of the n pairs as a distinct color class gives a complete n-coloring of H, and hence $\psi^*(H) \geq n$. Indeed, each pair is an independent set because it contains one vertex from each of V, W and there are no edge between these two sets. To see that there is an edge between every two pairs, consider two arbitrary pairs. Suppose that one pair has a vertex from V_i and a vertex from W_i, and the other pair has a vertex from V_j and a vertex from W_j. If $i = j$, then the two vertices from V_i are connected, because V_i forms a clique in $\bar{G}$ and thus in H. If $i \neq j$, then the vertex from W_i and the vertex from W_j are connected. Each of the $2n$ vertices of H belongs to some pair, and hence the n pairs form a complete n-coloring of H, as claimed.

Consider now the case that $x \notin L$. Let $l = \psi^*(H)$ and let $C_1, \ldots, C_l$ be the corresponding color classes. Observe that a color class C_k is an independent set, and thus contains vertices from at most one set W_j. So every class can be identified by a particular W_j from which it contains vertices, or it may contain no vertices from $W = \cup_j W_j$.

Consider first the classes C_k which contain no vertices from V, and let l_0 denote the number of such classes. By the above observation, each of these l_0 classes is entirely contained in one set W_j. Furthermore, each of these l_0 color classes is contained in a different W_j, since each W_j is an independent set, and every two color classes share an edge. Therefore, $l_0 \leq \hat{\chi}$.

Consider next the color classes C_k which contain exactly one vertex from V and possibly some vertices from W, and let l_1 denote the number of these classes. We group these classes according to the observation above, as follows. Let I_j denote the classes which contain exactly one vertex from V and one or more vertices from W_j, and let I_0 denote the classes which consist of a single vertex of V and no vertices of W. Then $l_1 = |I_0 \cup I_1 \cup \cdots I_{\hat{\chi}}|$.

The classes in I_j, for $j \geq 1$, cannot be connected to each other using their W_j vertices, because W_j is an independent set. Since these color classes share an edge, they must be connected using their vertices from V. But each of these classes has exactly one vertex from V, so these vertices from V form a clique in $\bar{G}$. A clique in $\bar{G}$ is of size at most $\hat{\alpha}$ (since $x \notin L$), and hence $|I_j| \leq \hat{\alpha}$ for every $j \geq 1$. A similar argument applies for I_0 and also for $I_0 \cup I_1$ (i.e. their corresponding vertices in V must form a clique), and hence also $|I_0 \cup I_1| \leq \hat{\alpha}$. We conclude that

$$l_1 = |I_0 \cup I_1 \cup \cdots \cup I_{\hat{\chi}}| \leq \hat{\chi} \cdot \hat{\alpha}$$

Consider finally the remaining color classes C_k, i.e. those that contain two or more vertices from V and possibly some vertices from W, and let l_2 denote the number of such classes. These l_2 classes contain together the remaining $n - l_1$ vertices of V, and hence the number of such classes is $l_2 \leq (n - l_1)/2$.

The total number of classes C_k in a complete coloring is thus

$$l = l_0 + l_1 + l_2 \leq l_0 + l_1 + \frac{n - l_1}{2} \leq \hat{\chi} + \frac{n + \hat{\chi} \cdot \hat{\alpha}}{2}$$

Since $\hat{\alpha} \geq 1$, we can have in Lemma 12 $\hat{\chi} \leq \hat{\chi} \cdot \hat{\alpha} \leq \epsilon n$, and hence

$$\psi^*(H) = l \leq \frac{n(1 + 3\epsilon)}{2}$$

We conclude that for any fixed $\epsilon > 0$ there is a gap of $2/(1 + 3\epsilon)$ between the achromatic number of H in the cases $x \in L$ and $x \notin L$. Since $\epsilon > 0$ is arbitrarily small, we can obtain any gap of $2 - \epsilon$, as desired. □

Theorem 10 *It is NP-hard to decide whether an input graph has a complete coloring with all color classes of size exactly 2, or whether every complete coloring of the graph has a color class of size at least 4.*

Proof. Consider the reduction from the proof of Theorem 9. If $x \in L$, then there is a complete coloring of H with every color class of size exactly 2. If $x \notin L$, then a complete coloring of H can use at most $n(1 + 3\epsilon)/2$ colors; since H has $2n$ vertices, there must be a color class of size at least $\frac{2n}{n(1+3\epsilon)/2} > 3$. □

Acknowledgements

The second author thanks Michael Langberg for helpful discussions.

References

[Bod89] H. L. Bodlaender. Achromatic number is NP-complete for cographs and interval graphs. *Inform. Process. Lett.*, 31(3):135–138, 1989.

[Bol78] B. Bollobás. *Extremal graph theory.* Academic Press Inc., London, 1978.

[CE97] N. Cairnie and K. Edwards. Some results on the achromatic number. *J. Graph Theory*, 26(3):129–136, 1997.

[CV97] A. Chaudhary and S. Vishwanathan. Approximation algorithms for the achromatic number. In *Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 558–563, 1997.

[Edw97] K. Edwards. The harmonious chromatic number and the achromatic number. In *Surveys in combinatorics, 1997 (London)*, pages 13–47. Cambridge Univ. Press, 1997.

[FHHM86] M. Farber, G. Hahn, P. Hell, and D. Miller. Concerning the achromatic number of graphs. *J. Combin. Theory Ser. B*, 40(1):21–39, 1986.

[FK98] U. Feige and J. Kilian. Zero knowledge and the chromatic number. *J. Comput. System Sci.*, 57(2):187–199, 1998.

[Hal93a] M. M. Halldórsson. Approximating the minimum maximal independence number. *Inform. Process. Lett.*, 46(4):169–172, 1993.

[Hal93b] M. M. Halldórsson. A still better performance guarantee for approximate graph coloring. *Inform. Process. Lett.*, 45(1):19–23, 1993.

[HM76] P. Hell and D. J. Miller. On forbidden quotients and the achromatic number. In *Proceedings of the 5th British Combinatorial Conference (1975)*, pages 283–292. Congressus Numerantium, No. XV. Utilitas Math., 1976.

[HM92] P. Hell and D. J. Miller. Achromatic numbers and graph operations. *Discrete Math.*, 108(1-3):297–305, 1992.

[HM97] F. Hughes and G. MacGillivray. The achromatic number of graphs: a survey and some new results. *Bull. Inst. Combin. Appl.*, 19:27–56, 1997.

[HR97] M. M. Halldórsson and J. Radhakrishnan. Greed is good: approximating independent sets in sparse and bounded-degree graphs. *Algorithmica*, 18(1):145–163, 1997.

[KL99] P. Krysta and K. Loryś. Efficient approximation algorithms for the achromatic number. In *ESA '99 (Prague)*, pages 402–413. Springer, 1999.

[LY94] C. Lund and M. Yannakakis. On the hardness of approximating minimization problems. *J. Assoc. Comput. Mach.*, 41(5):960–981, 1994.

[Mát81] A. Máté. A lower estimate for the achromatic number of irreducible graphs. *Discrete Math.*, 33(2):171–183, 1981.

[YG80] M. Yannakakis and F. Gavril. Edge dominating sets in graphs. *SIAM J. Appl. Math.*, 38(3):364–372, 1980.

A An algorithm for irreducible graphs

In this appendix we describe the algorithm of Máté [Mát81] for finding a partial complete $\Omega(\log n/\log\log n)$-coloring of an irreducible graph G.

1. Find greedily a (legal) coloring of G, by iteratively removing from the graph a maximal independent set. Let $I_1, \ldots, I_p$ be the color classes, i.e. the independent sets that are removed in the iterations.

2. If the greedy coloring uses more than $\log n$ colors then *return* this coloring. (Note that the greedy coloring is complete).

3. Assume without loss of generality that $I = I_1$ is the largest of the independent sets. (Note that $|I| \geq n/\log n$).

4. Fix an arbitrary total order $\prec$ on the vertices of I.

5. For each unordered pair $x, y \in I$ fix a *distinguisher* $d_{x,y}$, which is a vertex that is adjacent to exactly one of x, y. Let $f(\{x, y\})$ be the color of $d_{x,y}$ in the greedy coloring of step 1. (Note that $d_{x,y} \notin I$ and thus $f(x, y) \in \{2, \ldots, k\}$).

6. For each unordered triple $x, y, z \in I$ let $g(\{x, y, z\})$ be the following boolean value. Assuming that $x \prec y \prec z$, let $g(\{x, y, z\})$ be 0 if $d_{x,y}$ and z are adjacent in G, and 1 otherwise.

7. Iteratively construct $I = S_0 \supset S_1 \supset \cdots \supset S_l$, and a set $Z = \{z_0, \ldots, z_{2l-1}\}$, as follows. Start with $S_0 \leftarrow I$ and let $M \leftarrow \log^{-2}|I|$. While $|S_i| \geq 50M^{-3}$, construct S_{i+1}:

   (a) let $z_{2i} \leftarrow \min_{\prec} S_i$ and $z_{2i+1} \leftarrow \min_{\prec} S_i \setminus \{z_{2i}\}$;

   (b) partition $S_i \setminus \{z_{2i}, z_{2i+1}\}$ into S^0_{i+1} and S^1_{i+1}, where S^r_{i+1} consists of the vertices $x \in S_i \setminus \{z_{2i}, z_{2i+1}\}$ with $g(\{z_{2i}, z_{2i+1}, x\}) = r$;

   (c) let S_{i+1} be one of S^0_{i+1} and S^1_{i+1}, as follows:
   if $|S^0_{i+1}| > Me^{-M/2}|S_i|$ then $S_{i+1} \leftarrow S^0_{i+1}$;
   otherwise (it must hold that $|S^1_{i+1}| > e^{-M}|S_i|$) let $S_i \leftarrow S^1_{i+1}$.

   (Note that $z_0 \prec z_1 \prec \ldots \prec z_{2l-1}$).

8. For $r = 0, 1$

   (a) let Q_r be the set of iterations i in which $S_{i+1} \leftarrow S^r_{i+1}$ is taken in (7c);

   (b) let $U_r \leftarrow \{(z_{2i}, z_{2i+1}) : i \in Q_r\}$.

9. If $|Q_0| \geq \Omega(\log|I|/\log\log|I|)$
   Return the following partial complete $|Q_0|$-coloring. For each pair of vertices $(z_{2i}, z_{2i+1}) \in U_0$ color with a fresh color its distinguisher $d_{z_{2i},z_{2i+1}}$ and one of the pair z_{2i}, z_{2i+1} which is not adjacent to the distinguisher $d_{z_{2i},z_{2i+1}}$.

10. Otherwise (it must hold that $|Q_1| \geq \Omega(\log^3 |I|)$)
Find the f value that is most frequent among the pairs $z_{2i}, z_{2i+1} \in U_1$, and let U_1' consist of the pairs with this f value. (Since f accepts at most $\log n$ values, $|U_1'| \geq |U_1| / \log n = \Omega(\log^2 n)$).

11. Select from each pair of vertices $(z_{2i}, z_{2i+1}) \in U_1'$ the one which is adjacent to its distinguisher $d_{z_{2i}, z_{2i+1}}$. These selected vertices form a matching to their corresponding distinguishers $d_{z_{2i}, z_{2i+1}}$. This matching of size $\Omega(\log^2 n)$ can be shown to be semi-independent. We can thus use Lemma 4 to *return* a partial complete $\Omega(\log n)$-coloring of G.