Noname manuscript No.
(will be inserted by the editor)

Jean Cardinal, Samuel Fiorini, Gwenaël Joret

Minimum Entropy Coloring

the date of receipt and acceptance should be inserted later

Abstract We study an information-theoretic variant of the graph coloring problem in which the objective function to minimize is the entropy of the coloring. The minimum entropy of a coloring is called the chromatic entropy and was shown by Alon and Orlitsky (1996) to play a fundamental role in the problem of coding with side information. In this paper, we consider the minimum entropy coloring problem from a computational point of view. We first prove that this problem is NP-hard on interval graphs. We then show that it is NP-hard to find a coloring whose entropy is within $(\frac{1}{7}-\varepsilon)\log n$ of the chromatic entropy for any $\varepsilon > 0$, where n is the number of vertices of the graph. A simple polynomial case is also identified. It is known that the graph entropy is a lower bound for the chromatic entropy. We prove that this bound can be arbitrarily bad, even for chordal graphs. Finally, we consider the minimum number of colors required to achieve minimum entropy and prove a Brooks-type theorem.

1 Introduction

The minimum graph coloring problem asks to color the vertices of a given graph with a minimum number of colors so that no two adjacent vertices have the same color. The minimum number of colors in a coloring of G is the *chromatic number* of G, denoted by $\chi(G)$. Numerous variants of this problem have been studied, with

Computer Science Dept., CP 212, jcardin@ulb.ac.be

Mathematics Dept., CP 216, sfiorini@ulb.ac.be, support from the Fonds National de la Recherche Scientifique and GERAD - HEC Montréal is acknowledged

Computer Science Dept., CP 212, gjoret@ulb.ac.be, Research Fellow of the Fonds National de la Recherche Scientifique

Université Libre de Bruxelles
Boulevard du Triomphe
B-1050, Brussels, Belgium

different objective functions and constraints (Jensen and Toft, 1995). An example of such alternative graph coloring problem is the *optimum cost chromatic partition problem* (Jansen, 2000), in which the cost of a coloring is the sum over all colors of the size of the corresponding color class multiplied by some coefficient. In the problem we consider, the cost of each color class is a concave function of its size.

The problem is actually defined on specific vertex-weighted graphs called probabilistic graphs. A *probabilistic graph* is a graph equipped with a probability distribution on its vertices. Let (G,P) be a probabilistic graph, and let X be any random variable over the vertex set $V(G)$ of G with distribution P. We define the *entropy* $H(\phi)$ of a coloring ϕ as the entropy of the random variable $\phi(X)$. In other words, the entropy of ϕ is:

$$H(\phi) = -\sum_i c_i \log c_i,$$

where $c_i = \sum_{x:\phi(x)=i} P(x)$ is the probability that X has color i. Throughout this paper, we always use base-2 logarithms. The *chromatic entropy* $H_\chi(G,P)$ of the probabilistic graph (G,P) is the minimum entropy of any of its colorings. We consider the problem of finding a minimum entropy coloring of a probabilistic graph.

The notion of chromatic entropy was first proposed in an information-theoretic context by Alon and Orlitsky (1996). They considered the problem of *(zero-error) coding with side information*, in which a random variable X must be transmitted to a receiver having already some partial information about X. Witsenhausen (1976) showed how this transmission scenario could be encoded in a *characteristic graph* G, the set of vertices of which is the set of possible values of X. Alon and Orlitsky (1996) proved that the minimum achievable rate for coding with side information is between $H_\chi(G,P)$ and $H_\chi(G,P)+1$ where P the probability distribution of X.

Given a minimum entropy coloring of (G,P), a Huffman code computed from the color probabilities will provide a suitable code with average length at most $H_\chi(G,P)+1$. So minimum entropy colorings directly yield good codes for the problem of coding with side information. Heuristic algorithms for practical coding with side information based on minimum entropy colorings have been proposed by Zhao and Effros (2003).

Minimum entropy coloring also applies to the compression of digital image partitions created by segmentation algorithms (Accame, Natale, and Granelli, 2000; Agarwal and Belongie, 2002).

While the problem has received attention in the information theory and data compression community, it has not yet been studied thoroughly from a computational and combinatorial point of view. Our contribution aims at filling this gap. Preliminary results have already been presented in Cardinal, Fiorini, and Van Assche (2004). Note that another combinatorial optimization problem with an entropy-like objective function has been recently studied by Halperin and Karp (2004).

We first prove in Section 2 some useful lemmas concerning the structure of minimum entropy colorings, and introduce the definition of maximal color-feasible sequences.

In Section 3, we consider the computational complexity of the minimum entropy coloring problem. We show that the problem is NP-hard even if the input graph G is an interval graph, a class of graphs on which many classic NP-hard problems become polynomial. We also study the approximability of the problem. Since the chromatic entropy takes value in the interval $[0, \log n]$, where n is the number of vertices of the graph, it is natural to consider additive approximations, i.e. approximations within an additive term. This translates to a multiplicative factor if we consider $2^{H(\phi)}$ instead of $H(\phi)$ as the objective function. Coloring each vertex with a different color trivially yields a coloring whose entropy is at most $\log n$. On the other hand, we show that, unless P=NP, there is no polynomial algorithm finding a coloring of entropy at most $H_\chi(G,P) + (1/7 - \varepsilon)\log n$ For any $\varepsilon > 0$. With the stronger assumption that ZPP≠NP, we show that the problem cannot be approximated to within a $(1-\varepsilon)\log n$ term. These results hold even if P is the uniform distribution. We end the section by giving a simple polynomial case, namely when the input graph G satisfies $\alpha(G) \leq 2$.

Alon and Orlitsky showed that the chromatic entropy was bounded from below by a well-known quantity called graph entropy (Simonyi, 2001), also known as *Körner entropy*. They left open the question of how tight a lower bound the Körner entropy is. In Section 4, we first note that the ratio between those two quantities is unbounded and then prove that the difference between them can be made arbitrarily large, even if the graph is chordal and the probability distribution is uniform.

Finally, we provide results on the minimum number $\chi_H(G,P)$ of colors required to achieve minimum entropy in Section 5. We first relate minimum entropy colorings to *Grundy colorings*, the family of graph colorings obtained by iteratively removing maximal stable sets. It is simple to show that all minimum entropy colorings are Grundy colorings, hence that $\chi_H(G,P)$ is bounded by the Grundy number of G, defined as the maximum number of colors in a Grundy coloring. We also show a converse: any Grundy coloring of a graph G is a minimum entropy coloring of a probabilistic graph (G,P) for some probability distribution P. Then we prove that if P is uniform a Brooks-type theorem holds: $\chi_H(G,P)$ is at most the maximum degree of G, provided G is connected and different from an odd cycle or a complete graph.

2 Preliminaries

Consider a probabilistic graph (G,P), where G is a graph and P a probability distribution defined on $V(G)$. For simplicity, we denote by $P(S)$ the sum $\sum_{x\in S} P(x)$, where $S \subseteq V(G)$. Formally, a *coloring* of G is a map ϕ from the vertex set $V(G)$ of G to the set of positive integers $\mathbb{N}^+$. We also use $\phi^{-1}(i)$ for the set of vertices colored with color i. As above, let c_i be the probability mass of the i-th color class. Hence we have $c_i = P(\phi^{-1}(i)) = Pr[\phi(X) = i]$, where $X \sim P(x)$ is a random vertex with distribution P. The *color sequence* of ϕ with respect to P is the infinite vector $c = (c_i)$.

A sequence c is said to be *color-feasible* for a given probabilistic graph (G,P) if there exists a coloring ϕ of G having c as color sequence. Most of the time, we will restrict to nonincreasing color sequences, that is, color sequences c such that $c_i \geq c_{i+1}$ for all i. This can be easily achieved for a given color sequence by renaming the colors. Note that color sequences define discrete probability distributions on $\mathbb{N}^+$. The entropy of a coloring is the entropy of the discrete random variable having its color sequence as distribution. In other words, we have $H(\phi) = H(c)$ whenever c is the color sequence of ϕ, where (with a slight abuse of terminology) $H(c)$ is the *entropy of color sequence* c, that is, $H(c) = -\sum_{i\in\mathbb{N}^+} c_i \log c_i$.

The following lemma is of fundamental importance for the remaining proofs and was noted by Alon and Orlitsky (1996). The proof is straightforward and only relies on the concavity of the function $p \mapsto -p\log p$.

Lemma 1 (Alon and Orlitsky, 1996) *Let c be a nonincreasing color sequence, let i, j be two indices such that $i < j$ and let α a real number such that $0 < \alpha \le c_j$. Then we have $H(c) > H(c_1,\ldots,c_{i-1},c_i+\alpha,c_{i+1},\ldots,c_{j-1},c_j-\alpha,c_{j+1},\ldots)$.*

We now examine the consequences of this lemma. We say that a color sequence c *dominates* another color sequence d if $\sum_{i=1}^{j} c_i \ge \sum_{i=1}^{j} d_i$ holds for all j. We denote this by $c \succeq d$. The partial order $\succeq$ is known as the *dominance ordering*. It is often restricted to nonincreasing color sequences in order to avoid unwanted incomparabilities. We also let $\succ$ denote the strict part of $\succeq$. A nonincreasing color sequence is said to be *maximal color-feasible* when it is not dominated by any other (nonincreasing) color sequence of the considered probabilistic graph. The next lemma indicates that color sequences of minimum entropy colorings are always maximal color-feasible.

Lemma 2 *Let c and d be two nonincreasing rational color sequences such that $c \succ d$. Then we have $H(c) < H(d)$.*

Proof Let i be the smallest index such that $c_i > d_i$ and let j be the smallest index such that $c_j < d_j$. Because c and d are distinct, such indices i and j exist. We have $i < j$ because $c \succ d$. Moreover, we have $c_k = d_k$ for $1 \le k < i$ and $c_k \ge d_k$ for $i \le k < j$. Let K denote the smallest common denominator of the components of c and d, and let $\alpha = 1/K$. We define e as the nonincreasing sequence obtained from d by incrementing d_i by α, decrementing d_j by α and then sorting the resulting sequence. Note that e is not necessarily color-feasible. By Lemma 1, we have $H(e) < H(d)$. We claim that we have $c \succeq e \succ d$.

Let m be an index such that $e = (d_1,\ldots,d_{i-1},d_i+\alpha,d_{i+1},\ldots,d_{j-1},d_{j+1}, d_{j+2},\ldots,d_m,d_j-\alpha,d_{m+1},d_{m+2},\ldots)$. Proving that we have $e \succ d$ is easy and left to the reader. Now consider the sum $\sum_{k=1}^{l}(c_k - e_k)$ for some index l. If we have $l \le j-1$ or $l \ge m$, then the sum is clearly nonnegative. Otherwise, we have $j \le l \le m-1$ and

$$
\begin{aligned}
\sum_{k=1}^{l}(c_k - e_k) &= -\alpha + \sum_{k=1}^{j-1}(c_k - d_k) + c_j - d_{j+1} + \ldots + c_l - d_{l+1} \\
&\ge c_j - d_j + \sum_{k=1}^{j-1}(c_k - d_k) + c_j - d_{j+1} + \ldots + c_l - d_{l+1} \\
&= \sum_{k=1}^{l}(c_k - d_k) + c_j - d_{l+1} \ge \sum_{k=1}^{l+1}(c_k - d_k) \ge 0.
\end{aligned}
$$

The claim follows. Because there is a finite number of nonincreasing sequences whose components are integral multiples of $1/K$ and sandwiched between c and d

in the dominance ordering, we reach the desired conclusion by iterating the above arguments a finite number of times. □

A property similar to that of Lemma 2 was observed for other coloring problems, in particular by de Werra, Glover, and Silver (1995); de Werra, Hertz, Kobler, and Mahadev (2000) for minimum cost edge colorings. A further consequence of Lemma 1 is that any minimum entropy coloring can be constructed by iteratively removing maximal stable sets, i.e., subsets of pairwise nonadjacent vertices that are inclusionwise maximal.

Lemma 3 *Assume that $P(x) > 0$ holds for all vertices of a probabilistic graph (G,P). Let ϕ be a minimum entropy coloring of G with respect to P. If the color sequence of ϕ is nonincreasing, then the i-th color class of ϕ is a maximal stable set in the subgraph of G induced by the vertices with colors $j \geq i$.*

Proof If the i-th color class is not maximal, we can recolor a vertex of the jth color class with color i, for some $j > i$. Because P is positive, Lemma 1 implies that this operation decreases the entropy, a contradiction. □

3 Complexity and approximability

We study in this section the complexity of the minimum entropy coloring problem and its approximability. We first note that the minimum entropy coloring problem has already been shown to be NP-hard on planar graphs with the uniform distribution (Cardinal et al, 2004).

An *interval graph* is the intersection graph of a set of open intervals on the real line: vertices correspond to intervals and two distinct vertices are adjacent if the corresponding intervals overlap. Our first result shows that finding a minimum entropy coloring of a probabilistic interval graph is NP-hard. Since the numerators and denominators of the probabilities that are used in our reduction are polynomial in the size of the input, the proof also shows that NP-hardness holds in the strong sense.

Theorem 1 *Finding a minimum entropy coloring of a probabilistic interval graph is strongly NP-hard.*

Proof Our reduction is from the problem of deciding if a circular arc graph G is k-colorable, which is NP-complete (Garey, Johnson, Miller, and Papadimitriou, 1980). *Circular arc graphs* are defined similarly as interval graphs, except that vertices corresponds to open arcs on a circle. Given a circular arc graph G, one can construct a circular representation for G in polynomial time (Tucker, 1980). The basic idea of the proof is to start with a circular-arc graph and cut it open somewhere to obtain an interval graph. The same idea is used in Marx (2005), where it is proved that finding a minimum sum coloring of an interval graph is NP-hard.

Let y be an arbitrary point on the circle that is not the endpoint of any arc in the considered representation of G. Let k' be the number of arcs in which y is included. If $k' > k$, then G is not k-colorable. If $k' < k$, we add to the representation $k - k'$ sufficiently small arcs that intersect only arcs including y. This clearly does

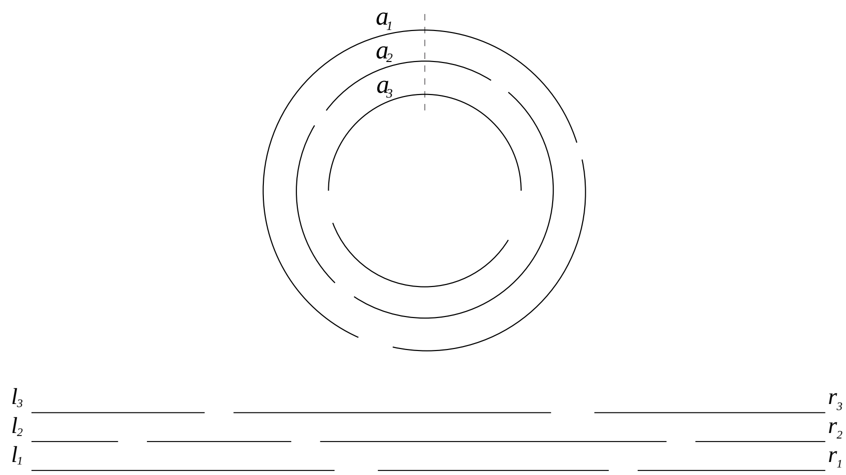

Fig. 1 Splitting of the circular arc graph.

not increase the chromatic number of G above k. Thus, it can be assumed that y is contained in exactly k arcs.

Denote $a_1,\ldots,a_k$ the arcs that contain y. By splitting each arc a_i into two parts l_i and r_i at point y we obtain an interval representation of some interval graph G' (see Figure 1 for an illustration). As is easily checked, G is k-colorable if and only if there is a k-coloring of G' in which l_j and r_j receive the same color for $1 \leq j \leq k$.

Since interval graphs are chordal, we can use an algorithm designed for chordal graphs (Golumbic, 2004) to list in linear time all maximal cliques of G'. For each such clique K, we do the following. If $|K| > k$ then we reject the input because in this case G is not k-colorable. Assume now $|K| \leq k$. By the Helly property for intervals, there exists a point z of the real line contained in the intervals of K and in no other. We extend the clique K by adding $k-|K|$ sufficiently small intervals centered at z in the interval representation. This is done in such a way that the new intervals intersect only intervals corresponding to vertices of K. As before, this operation does not increase the chromatic number of G' above k. Let H denote the resulting interval graph. By construction, all maximal cliques of H are also maximum.

Let $\mathscr{K}$ denote the collection of maximum cliques of H and let $C = |\mathscr{K}|$. Consider the auxiliary bipartite graph B having $V(H)$ and $\mathscr{K}$ as color classes in which $x \in V(H)$ is adjacent to $K \in \mathscr{K}$ whenever $x \in K$. We define a probability distribution P on the vertices of H as follows:

$$P(x) = \lambda \deg_B(x) + \begin{cases} \lambda j & \text{if } x = l_j \text{ or } r_j, \\ 0 & \text{otherwise,} \end{cases}$$

where $\lambda > 0$ is chosen such that the sum of $P(x)$ over all vertices x of H equals 1, and $\deg_B(x)$ denotes the degree of x in the auxiliary graph B.

Letting c^* denote the sequence $\lambda(2k+C, 2(k-1)+C, \ldots, 2+C, 0, \ldots)$, we claim that the following two assertions hold for the probabilistic graph (H,P):

(i) if G is k-colorable then c^* is color-feasible;

(ii) c^* dominates all color-feasible sequences and every coloring whose color sequence equals c^* assigns the same color to vertices l_j and r_j for all j.

First assume that G is k-colorable. Then there exists a k-coloring ϕ of (H,P) assigning the same color to l_j and r_j for all j. Let c denote the color sequence of ϕ. Without loss of generality, we assume that vertices l_{k-i+1} and r_{k-i+1} belong to the i-th color class. For every color $i \in \{1,\ldots,k\}$, we have

$$\begin{aligned} c_i &= \sum_{x\in\phi^{-1}(i)} P(x) \\ &= 2\lambda(k-i+1) + \lambda \sum_{x\in\phi^{-1}(i)} \deg_B(x) \\ &= \lambda(2(k-i+1)+C). \end{aligned}$$

The third equality holds for the following reasons. Because ϕ is proper, no two vertices of H with color i are contained in the same maximum clique. Moreover, if some maximum clique is disjoint from the i-th color class then ϕ cannot possibly be a k-coloring. It follows that every maximum clique of H contains exactly one vertex of H with color i. Hence the third equality holds and we have $c = c^*$. Claim (i) follows.

Now consider any stable set S in H. Clearly, no two vertices of S are contained in the same clique. For the auxiliary graph B, this means that no $K \in \mathscr{K}$ is adjacent to two distinct elements of S. It follows that the sum $\sum_{x\in S} \deg_B(x)$ is at most C. Moreover, S contains at most one vertex in $\{l_1,\ldots,l_k\}$ and at most one vertex in $\{r_1,\ldots,r_k\}$. Let j and j' be indices such that $S\cap\{l_1,\ldots,l_k\} \subseteq \{l_j\}$ and $S\cap\{r_1,\ldots,r_k\} \subseteq \{r_{j'}\}$. Then the total probability mass of S in (H,P) is at most $\lambda(j+j'+C)$ with equality only if S contains both l_j and $r_{j'}$. Let now ψ be any coloring of H and let c denote the color sequence of ψ. Again, we assume that c is nonincreasing. By what precedes, we have

$$\begin{aligned} c_1 &\le c_1^*, \\ c_1 + c_2 &\le c_1^* + c_2^*, \\ &\vdots \\ c_1 + c_2 + \ldots + c_k &\le c_1^* + c_2^* + \ldots + c_k^*. \end{aligned}$$

Hence c^* dominates c. Moreover, if $c = c^*$ then the i-th color class of ψ contains both l_{k-i+1} and r_{k-i+1} for $1 \le i \le k$. Claim (ii) follows.

From Lemma 2, we then infer that G is k-colorable if and only if every minimum entropy coloring of (H,P) has c^* as color sequence. We conclude that finding a minimum entropy coloring of a probabilistic interval graph is (strongly) NP-hard. □

We now consider the approximability of the minimum entropy coloring problem. Since the objective function takes values in the interval $[0,\log n]$, it makes sense to look for polynomial time algorithms that find a coloring whose entropy is within an additive term of the chromatic entropy, i.e. is at most $H_\chi(G,P)+\delta$ for some positive real number δ. We call such an algorithm a *δ-approximation algorithm*. Note that coloring each vertex with a different color gives a trivial $\log n$-approximation algorithm for the minimum entropy coloring problem.

Theorem 2 *Let c be a real such that* $0 < c \leq 1$ *and assume that for some positive real* ε *there exists a* $(c-\varepsilon)\log n$*-approximation algorithm A for the minimum entropy coloring problem. Then there exists a polynomial time algorithm coloring G with at most* $n^{c-\varepsilon/2}\chi(G)$ *colors.*

Proof Without loss of generality, we may suppose $\varepsilon < c$. Let U denote the uniform distribution on $V(G)$, $n = |G|$ denote the order of G and $\chi = \chi(G)$ denote the chromatic number of G. We claim that some color class in the coloring ϕ output by algorithm A on the input (G,U) contains at least $n^{1-c+\varepsilon}/\chi$ vertices. In order to show this, list the color classes of ϕ in nonincreasing cardinalities as S_1, S_2, ..., S_k. Letting $H(\phi)$ denote the entropy of ϕ with respect to U, we have

$$-\log\frac{|S_1|}{n} \leq H(\phi) \leq H_\chi(G,U) + (c-\varepsilon)\log n \leq \log\chi + (c-\varepsilon)\log n.$$

The first inequality follows from the fact that $\sum_i -P(i)\log P(i) \geq -\log P_{max}$ holds for all probability distributions P whose maximum is P_{max}. The middle one holds by hypothesis and the last one comes from the fact that the entropy of a minimum cardinality coloring is at most $\log\chi$. Hence the size of S_1 is at least $n^{1-c+\varepsilon}/\chi$, so our claim holds.

Let A' denote the polynomial time algorithm that uses A as a subroutine to find in any graph G with n vertices and chromatic number χ a stable set of size at least $n^{1-c+\varepsilon}/\chi$. Now we iteratively use A' to color any graph by coloring with the same color all the vertices in the stable set output by A' and removing these vertices from the graph. Let $G_0 = G$, $G_1 = G_0 - A'(G_0)$, $G_2 = G_1 - A'(G_1)$, ..., $G_\ell = G_{\ell-1} - A'(G_{\ell-1})$ be the sequence of graphs considered, and let $t = (\chi n^{c-\varepsilon})/(\chi n^{c-\varepsilon} - 1)$. For each i between 1 and ℓ, we have

$$|G_i| \leq |G_{i-1}| - \frac{|G_{i-1}|^{1-c+\varepsilon}}{\chi(G_{i-1})} \leq |G_{i-1}| - \frac{|G_{i-1}|}{\chi n^{c-\varepsilon}} = \frac{|G_{i-1}|}{t}.$$

It follows that $|G_i| \leq n/t^i$ for all i. Because G_ℓ is nonempty, we have $n/t^\ell \geq 1$ and hence $\ell \leq \log_t n = \ln n/\ln t$. The number of colors in the resulting coloring of G equals $\ell+1$. By what precedes, we have

$$\ell+1 \leq \frac{\ln n}{\ln t} + 1 \leq \frac{\ln n}{(t-1) - \frac{(t-1)^2}{2}} + 1 = (n^{c-\varepsilon}\chi - 1)\frac{\ln n}{1 - \frac{1}{2(\chi n^{c-\varepsilon} - 1)}} + 1.$$

In the second inequality we used that $\ln(x+1) \geq x - \frac{x^2}{2}$ for $x \geq 0$. Because the case $\chi = 1$ is trivial, we can assume that $\chi \geq 2$. It follows that

$$\ell+1 \leq 2(n^{c-\varepsilon}\chi - 1)\ln n + 1 = 2n^{c-\varepsilon}\ln n \cdot \chi - 2\ln n + 1.$$

If n is large enough, that is, greater or equal to some constant depending on ε, we find $\ell+1 \leq n^{c-\varepsilon/2}\chi$. Consequently, we can in polynomial time find a coloring of a graph G with at most $n^{c-\varepsilon/2}\chi$ colors. (Indeed, if n is small we use a brute force algorithm to color the graph exactly and we can easily detect if $\chi = 1$ in polynomial time.) □

It is known that the existence of a polynomial time algorithm coloring any graph G with at most $n^{1-\varepsilon}\chi(G)$ colors for some positive real ε implies ZPP=NP (Feige and Kilian, 1998). Moreover, if the number of colors used by such an algorithm is bounded by $n^{1/7-\varepsilon}\chi(G)$ then it implies P=NP (Bellare, Goldreich, and Sudan, 1995). Combining these results with Theorem 2 we obtain the following corollaries.

Corollary 1 *Let ε be any positive real. There is no $(1-\varepsilon)\log n$-approximation for the minimum entropy coloring problem, unless ZPP=NP.*

Corollary 2 *Let ε be any positive real. There is no $(1/7-\varepsilon)\log n$-approximation for the minimum entropy coloring problem, unless P=NP.*

In view of the proof of Theorem 2, the theorem and its corollaries remain true when the problem is restricted to probabilistic graphs equipped with the uniform distribution. We end this section by identifying an easy polynomial case for the minimum entropy coloring problem.

Theorem 3 *There exists a polynomial time algorithm for the minimum entropy coloring problem restricted to graphs G satisfying $\alpha(G) \leq 2$.*

Proof Let (G,P) be a probabilistic graph such that $\alpha(G) \leq 2$. A color class in any coloring of G is composed of one or two vertices. Each color class of size two corresponds to an edge in the complement $\bar{G}$ of G. In fact, the set of edges of $\bar{G}$ corresponding to the color classes of size two forms a matching.

Let $f(p) = -p\log p$. If a color class is composed of a single vertex x, then the contribution of this color to the overall entropy is $f(P(x))$. Otherwise, the contribution of a color is $f(P(x)+P(y))$, where x and y are the only two vertices in the color class. So if we denote by M the matching in $\bar{G}$ induced by a coloring of G, the entropy of this coloring with respect to P is

$$
\begin{aligned}
&\sum_{x\notin V(M)} f(P(x)) + \sum_{xy\in E(M)} f(P(x)+P(y)) \\
&= \sum_{x\in V(G)} f(P(x)) + \left(\sum_{xy\in E(M)} f(P(x)+P(y)) - f(P(x)) - f(P(y))\right) \\
&= H(X) + \sum_{e\in E(M)} \rho(e),
\end{aligned}
$$

where $X \sim P(x)$ and $\rho(xy) = f(P(x)+P(y)) - f(P(x)) - f(P(y))$. Hence finding a minimum entropy coloring amounts to finding a maximum weight matching in $\bar{G}$, each edge e of which has nonnegative weight $-\rho(e)$. This can be done in $O(|V(\bar{G})||E(\bar{G})| + |V(\bar{G})|^2 \log|V(\bar{G})|)$ time (Gabow, 1990). □

4 Chromatic vs. Körner entropy

We first give a definition of a previously known quantity that is often referred to as graph entropy. Following Alon and Orlitsky (1996) and to avoid ambiguities, we call it Körner entropy.

The *Körner entropy* $H_\kappa(G,P)$ of a probabilistic graph (G,P) can be defined by

$$H_\kappa(G,P) = \min_{\substack{a\in\mathrm{STAB}(G)\\ a>0}} - \sum_{x\in V(G)} P(x)\log a_x, \tag{1}$$

where $\mathrm{STAB}(G)$ is the stable set polytope of G, defined in $\mathbb{R}^{V(G)}$ as the convex hull of the characteristic vectors of the stable sets of G. The Körner entropy has a number of applications, the most prominent being the problem of sorting with partial information studied by Kahn and Kim (1995) in their celebrated paper.

We also define $\alpha(G,P)$ which is simply the maximum weight $P(S)$ of a stable set S of (G,P).

Lemma 4 *For any probabilistic graph (G,P), we have*

$$-\log\alpha(G,P) \le H_\kappa(G,P) \le H_\chi(G,P) \le \log\chi(G).$$

Proof The last inequality comes from the fact that in the worst case, the distribution of the colors in a minimum cardinality coloring is uniform, hence its entropy is at most $\log\chi(G)$. The second inequality is proved in Alon and Orlitsky (1996). We here give a shorter proof based on (1). Consider a minimum entropy coloring ϕ of (G,P), and let $a_x = P(\phi^{-1}(\phi(x))$ for each vertex $x \in V(G)$. Since each color class of ϕ is a stable set and the probability masses of the color classes sum up to one, the vector a is a convex combination of characteristic vectors of stable sets, hence we have $a \in \mathrm{STAB}(G)$. Furthermore we can check that for this value of a we have

$$\begin{aligned} -\sum_{x\in V(G)} P(x)\log a_x &= -\sum_{x\in V(G)} P(x)\log P(\phi^{-1}(\phi(x))) && (2)\\ &= -\sum_i P(\phi^{-1}(i))\log P(\phi^{-1}(i)) && (3)\\ &= H_\chi(G,P), && (4)\end{aligned}$$

thus the minimum defining $H_\kappa(G,P)$ is at most $H_\chi(G,P)$.

The first inequality is derived as follows. Let $a \in \mathrm{STAB}(G)$. A stable set has weight at most $\alpha(G,P)$, so we have $\sum_{x\in V(G)} P(x)a_x \le \alpha(G,P)$. Combining this with the concavity of $x \mapsto \log(x)$ yields

$$-\sum_{x\in V(G)} P(x)\log a_x \ge -\log\sum_{x\in V(G)} P(x)a_x \ge -\log\alpha(G,P).$$

□

The bounds on the chromatic entropy given in Lemma 4 can be computed in polynomial time only for certain classes of probabilistic graphs. In particular, when G is a perfect graph, the two lower bounds can be computed (to any fixed accuracy) in polynomial time, as follows from Grötschel, Lovász, and Schrijver (1993). The chromatic number can also be computed in polynomial time on these graphs.

The question of the quality of the lower bound given by the Körner entropy on the chromatic entropy was raised by Alon and Orlitsky (1996). The next two results provide an answer to this question.

Proposition 1 *The ratio $H_\chi(G,P)/H_\kappa(G,P)$ can be arbitrarily large.*

Proof Let G_n be the graph consisting of a matching of size $n+1$ and let $\varepsilon = 1/n^2$. Choose an edge $xy \in E(G_n)$ and let P_n be the probability distribution such that $P_n(x) = (1-n\varepsilon)(1-\varepsilon)$, $P_n(y) = (1-n\varepsilon)\varepsilon$ and $P_n(z) = \varepsilon/2$ for $z \in V(G_n), z \neq x,y$.

Define $h(p) : (0,1) \to \mathbb{R}$ as the function $p \mapsto -p\log p - (1-p)\log(1-p)$. The chromatic and Körner entropy of G_n are easily obtained:

$$\begin{aligned} H_\chi(G_n,P_n) =& h(n\varepsilon(1/2-\varepsilon)+\varepsilon), \\ H_\kappa(G_n,P_n) =& n\varepsilon + (1-n\varepsilon)h(\varepsilon). \end{aligned}$$

Now, it can be checked that

$$\lim_{n\to\infty} \frac{H_\chi(G_n,P_n)}{H_\kappa(G_n,P_n)} = \infty.$$

□

The above proof does not show that the difference $H_\chi(G,P) - H_\kappa(G,P)$ is not bounded. We now prove this in the next lemma.

Proposition 2 *The difference $H_\chi(G,P) - H_\kappa(G,P)$ can be arbitrarily large, even if G is chordal and P is the uniform distribution.*

In order to prove this result, we define a graph $G_k(n)$ ($n \geq 2, k \geq 1$) inductively on k. The graph $G_1(n)$ is the single vertex graph K_1, and for $k \geq 2$ the graph $G_k(n)$ is obtained as follows:

- start with the complete graph $K_{n^{k-1}}$ on n^{k-1} vertices,
- partition its vertex set $V(K_{n^{k-1}})$ in n sets $V_1, V_2, \ldots, V_n$ of equal sizes,
- for each set V_i ($1 \leq i \leq k$) add a disjoint copy of $G_{k-1}(n)$ and all edges with one endpoint in V_i and the other in the vertex set of the i-th copy of $G_{k-1}(n)$.

A drawing of the graph $G_3(3)$ is given in Figure 2. It can easily be checked that $G_k(n)$ is chordal. We study in the next two lemmas the behavior of $H_\kappa(G_k(n),U)$ and $H_\chi(G_k(n),U)$ when k is fixed and n goes to infinity, where U is the uniform distribution. Note that $G_k(n)$ has kn^{k-1} vertices in total, so we have $U(x) = \frac{1}{kn^{k-1}}$ for all vertices x.

Lemma 5 $H_\kappa(G_k(n),U) \leq \frac{(k-1)}{2}\log n + o(1)$.

Proof We first associate to each vertex of $G_k(n)$ a *level* between 1 and k: the only vertex of $G_1(n)$ has level 1, and for $k \geq 2$ the level of a vertex x is either k if it belongs to the *central clique* $K_{n^{k-1}}$ arising in the definition of $G_k(n)$ above, or its level in the i-th copy of the graph $G_{k-1}(n)$.

Now consider the point $\tilde{a}$ of $\mathbb{R}^{V(G_k(n))}$ defined by

$$\tilde{a}_x = \frac{(n-1)^{k-i}}{n^{k-1}},$$

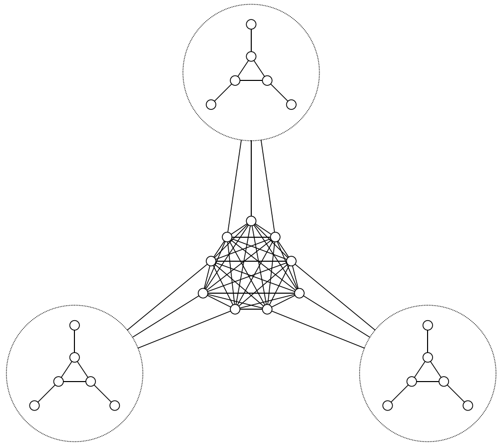

Fig. 2 The graph $G_3(3)$.

where i denotes the level of vertex x. As chordal graphs are perfect, it follows from a classic theorem of Chvátal (1975) that $\mathrm{STAB}(G_k(n))$ is described by the trivial inequalities $a_x \geq 0$ for all $x \in V(G_k(n))$, and the clique inequalities $\sum_{x \in K} a_x \leq 1$ for all cliques K of $G_k(n)$. Using this, it can be checked that $\tilde{a} \in \mathrm{STAB}(G_k(n))$ (for instance, by induction on k). This point yields the desired upper bound on $H_\kappa(G_k(n),U)$:

$$\begin{aligned} H_\kappa(G_k(n),U) &\leq \sum_{x \in V(G_k(n))} -\frac{1}{kn^{k-1}} \log \tilde{a}_x \\ &= \frac{1}{kn^{k-1}} \sum_{1 \leq i \leq k} n^{k-1} \log \frac{n^{k-1}}{(n-1)^{k-i}} \\ &= (k-1)\log n - \frac{k-1}{2} \log(n-1) \\ &= \frac{(k-1)}{2} \log n + o(1). \end{aligned}$$

The second equation above holds because each level of $G_k(n)$ contains exactly n^{k-1} vertices. □

Lemma 6 $H_\chi(G_k(n),U) \geq \log k + \frac{(k-1)}{2} \log n - o(1)$.

Proof We first note that the unique maximum clique of $G_k(n)$ is its central clique $K_{n^{k-1}}$, implying that $\chi(G_k(n)) = n^{k-1}$ as chordal graphs are perfect.

Let c be a maximal color-feasible sequence of $(G_k(n),U)$. We prove the following by induction on k:

$$\begin{aligned} c_1 &\geq \frac{(n-1)^{k-1}}{kn^{k-1}}, \\ c_j &\geq \frac{(n-1)^{k-i}}{kn^{k-1}} \qquad \text{for } 1 \leq i \leq k \text{ and } n^{i-2}+1 \leq j \leq n^{i-1}. \end{aligned} \tag{5}$$

This is clearly true for $k=1$. From now on, we assume $k \geq 2$. Because $\chi(G_k(n)) = n^{k-1}$, we have $c_j \geq \frac{1}{kn^{k-1}}$ for $j \in \{n^{k-2}+1,\ldots,n^{k-1}\}$. Thus the case $i=k$ in (5) is settled. Now decompose c as $c = c^0 + c^1 + \cdots + c^n$ where c^0 counts the probability masses coming from the central clique and $c^1,\ldots,c^n$ count those coming from the n copies of $G_{k-1}(n)$ for each color class. Let d^l ($1 \leq l \leq n$) denote the sequence c^l with all zero entries removed, except the trailing zeroes. Because c is maximal, the sequence $\frac{kn}{k-1}d^l$ is a maximal color-feasible sequence of $G_{k-1}(n)$ for all l. Moreover, the maximality of c and the structure of $G_k(n)$ imply that at most one of the n components $c^1_j,\ldots,c^n_j$ is zero for any j between 1 and n^{k-2}. Indeed, a stable set of $G_k(n)$ either contains no vertex of the central clique and is thus composed of n stable sets coming from each of the n copies of $G_{k-1}(n)$, or includes exactly one vertex of the central clique, say x, and is thus composed of x and $n-1$ stable sets coming from $n-1$ copies of $G_{k-1}(n)$ (that is, all the copies except the one which is totally adjacent to x). In particular, if two of the n components $c^1_j,\ldots,c^n_j$ equalled zero we would be able to switch the entries c^l_j and $c^l_{j'}$ for some l and some $j' > j$ such that $c^l_j = 0$ and $c^l_{j'} > 0$, while keeping a color-feasible sequence c, contradicting the maximality of c. Let $l \in \{1,\ldots,n\}$. Because d^l is maximal color-feasible, it is nonincreasing. So if $c^l_j \neq 0$, we have $c^l_j \geq d^l_j$. From the induction hypothesis applied to d^l and the latter observations, we infer the following inequalities:

$$\begin{aligned} c_1 &\geq (n-1)\tfrac{(n-1)^{k-2}}{kn^{k-1}} = \tfrac{(n-1)^{k-1}}{kn^{k-1}} \\ c_j &\geq (n-1)\tfrac{(n-1)^{k-i-1}}{kn^{k-1}} = \tfrac{(n-1)^{k-i}}{kn^{k-1}} \quad \text{for } 1 \leq i \leq k-1 \text{ and } n^{i-2}+1 \leq j \leq n^{i-1}. \end{aligned}$$

Let $\widehat{c}$ be the sequence such that $\widehat{c}_j = 0$ for $j > n^{k-1}$, $\widehat{c}_j$ is the minimum possible value of c_j allowed by (5) for $2 \leq j \leq n^{k-1}$ and $\widehat{c}_1 = 1 - \sum_{j\geq 2}\widehat{c}_j$. As is easily verified, $\widehat{c}$ dominates c. Using Lemma 2 we get

$$H_\chi(G_k(n),U) \geq H(\widehat{c}) = \log k + \frac{(k-1)}{2}\log n - o(1).$$

□

Proposition 2 follows from Lemmas 5 and 6. Before turning to the next section, we mention that we can show that Proposition 2 also holds when G is an interval graph and P is arbitrary by adapting the construction used above.

5 Number of Colors

We consider in this section the number of colors used in a minimum entropy coloring. We denote by $\chi_H(G,P)$ the minimum number of colors in a minimum entropy coloring of the probabilistic graph (G,P). We mention that the upper bounds on the number of colors used in a minimum entropy coloring given in this section hold for all minimum entropy colorings whenever we have $P(x) > 0$ for all vertices x.

We first relate minimum entropy colorings to another kind of colorings studied in the literature. A *Grundy coloring* of a graph is a coloring such that for any color i, if a vertex has color i then it is adjacent to at least one vertex of color j for all $j < i$. The *Grundy number* $\Gamma(G)$, also called the *first-fit online coloring number* (Pemmaraju, Raman, and Varadarajan, 2004), of a graph G is the maximum number of colors in a Grundy coloring of G. See Erdös, Hedetniemi, Laskar, and Prins (2003) for a recent survey of this topic. Equivalently, Grundy colorings are colorings that can be obtained by iteratively removing maximal stable sets.

Proposition 3 *Any minimum entropy coloring of a graph G equipped with a probability distribution on its vertices is a Grundy coloring. Moreover, for any Grundy coloring ϕ of G, there exists a probability mass function P over $V(G)$ such that ϕ is the unique minimum entropy coloring of (G,P).*

Proof The first part of the claim is given by Lemma 3. We prove the second part by induction on the number k of colors used in ϕ. It is trivially true for $k = 1$, since in that case G has no edge and the unique minimum entropy coloring with respect to any positive distribution P assigns the same color to every vertex. Now assume that the proposition holds for colorings with less than k colors. We call V' the set of vertices of G having a color different from 1, and G' the corresponding induced subgraph. By the induction hypothesis, there exists a distribution P' such that ϕ restricted to G' is the unique minimum entropy coloring of (G',P').

We define a probability distribution P for G as follows. For each $x \in V(G)$ we set $P(x) = P'(x)/t$ if $x \in V'$ and $P(x) = (1-1/t)/|\phi^{-1}(1)|$ otherwise. The entropy of ϕ with respect to P equals $-\sum_{i\geq 1}((P'(\phi^{-1}(i))/t)\log(P'(\phi^{-1}(i))/t)) - (1-1/t)\log(1-1/t)$.

We first show that in a minimum entropy coloring ψ of (G,P), all vertices in $S = \phi^{-1}(1)$ must have the same color. Let us assume otherwise, that is, vertices in S do not all have the same color under coloring ψ. The maximum probability of a color class in ψ then satisfies $P_{\max} \leq 1-(1-1/t)/|S| = (|S|-1+1/t)/|S|$ for t large enough. Furthermore, the entropy of ψ with respect to P is at least $-\log P_{\max} \geq \log(|S|/(|S|-1+1/t))$. As t tends to infinity, the difference between this lower bound on the entropy of ψ and the entropy of ϕ tends to $\log(|S|/(|S|-1))$. Thus ψ cannot have minimum entropy if t is large enough.

Because S is a maximal stable set in G, if S is contained in a color class of a coloring ψ of G then S *is* a color class of ψ. The entropy of any such coloring ψ can be written as $H'/t + h(1/t)$, where H' denotes the entropy with respect to P' of the restriction of ψ to G', and $h(1/t)$ is the entropy of a Bernoulli random variable with parameter $1/t$. This shows that minimizing the entropy of any coloring ψ assigning the same color to all vertices of S amounts to minimizing the entropy of the same coloring restricted to (G',P'). Now it follows from the induction hypothesis that ψ and ϕ have the same color classes, which concludes the proof. □

It is easy to see that the Grundy number of a tree T can be arbitrarily large, thus giving the same property for $\chi_H(T,P)$ by the previous proposition. We note that this observation can be strengthened to the case of P uniform, as shown in Cardinal et al (2004).

Proposition 4 (Cardinal et al, 2004) *$\chi_H(G,P)$ is not bounded by any function of $\chi(G)$, even if P is uniform and G is a tree.*

We now consider upper bounds on $\chi_H(G,P)$ in terms of the maximum degree $\Delta(G)$ of a vertex in G. We first note the following easy lemma, which was already noted in Cardinal et al (2004).

Lemma 7 (Cardinal et al, 2004) *For any probabilistic graph* (G,P)*, we have* $\chi_H(G,P) \leq \Delta(G)+1$.

Brooks (1941) showed a classic result stating that if G is a connected graph different from a complete graph or an odd cycle, then $\chi(G) \leq \Delta(G)$, where $\Delta(G)$ is the maximum degree of a vertex in G. There are graphs G other than those cited above that have a Grundy coloring using $\Delta(G)+1$ colors, thus by Proposition 3 we cannot extend Brooks' theorem by substituting $\chi_H(G,P)$ to $\chi(G)$ without making any assumption on P. In the next theorem we prove that such an extension holds when P is the uniform distribution.

Theorem 4 *If G is a connected graph different from a complete graph or an odd cycle, then* $\chi_H(G,U) \leq \Delta(G)$*, where U is the uniform distribution over* $V(G)$.

Proof The proof closely follows the one given in Diestel (2000) for Brooks' theorem. We borrow the path notation in which xPy is a path between vertices x and y, and $\dot{x}$ means that vertex x is not included. The considered path should be clear from the context. For simplicity, we use the shorthand notation $P(i) = P(\phi^{-1}(i))$ for a color i.

Let $\Delta = \Delta(G)$ and $n = |G|$. If $\Delta \leq 2$, either G is an odd cycle and we have nothing to show, or G is a path or an even cycle and the proposition is trivial. Hence we assume $\Delta \geq 3$. Let us consider a minimum entropy coloring ϕ of (G,U) with colors in the set $\{1,2,\dots,\Delta+1\}$. We show that if the $\Delta+1$ colors are used, then ϕ cannot have minimum entropy with respect to U, unless G is the complete graph.

Without loss of generality, we consider that color $\Delta+1$ has minimum weight: for all $1 \leq i \leq \Delta$, we have $P(i) \geq P(\Delta+1)$. Let us choose a vertex $x \in V(G)$ such that $\phi(x) = \Delta+1$. The vertex x must be adjacent to Δ vertices colored with colors 1 to Δ, otherwise it could be recolored and, from Lemma 1, ϕ would not have minimum entropy. We denote by x_i the vertex adjacent to x and such that $\phi(x_i) = i$. Let F be defined as the graph induced by the vertices colored with colors $\{1,2,\dots,\Delta\}$, $F_{i,j}$ as the graph induced on F by the vertices colored with colors i or j, and C_{ij} (respectively C_{ji}) as the component of $F_{i,j}$ containing x_i (respectively x_j).

We first show the following:

$$C_{ij} \text{ is a path.} \tag{6}$$

First, x_i must have a single neighbor in C_{ij}. Otherwise, it could be recolored, with a color k different of i and j and x could in turn be recolored with color i. The probability mass $P(\Delta+1)$ would decrease by $1/n$, and $P(k)$ would increase by $1/n$. From Lemma 1 and the fact that $P(\Delta+1) \leq P(k)$, the entropy would decrease, and ϕ would not be optimal.

Let us assume that C_{ij} is not a path. Then there must be an inner vertex of C_{ij} having three identically colored neighbors. Let us define y as the first such vertex on the path from x_i in C_{ij}. The vertex y can be recolored with a color k different from i or j, since its neighbors cannot have more than $\Delta - 2$ distinct colors. We can then perform the following steps:

1. recolor y with color k,
2. interchange colors i and j on the path $x_i P \dot{y}$,
3. recolor x with color i.

We show that they can only decrease the entropy, and therefore that ϕ cannot be optimal unless C_{ij} is a path. We have to consider two cases: either $\phi(y) = i$ or $\phi(y) = j$.

If $\phi(y) = i$, then the path $x_i P \dot{y}$ is even. Interchanging the colors i and j on this path does not change the probability masses $P(i)$ and $P(j)$. The probability $P(i)$ is decreased by $1/n$ when y is recolored, but increased by $1/n$ when x is recolored. Hence the overall sequence of changes leaves the probability $P(i)$ unchanged, while $P(\Delta+1)$ is decreased by $1/n$, and $P(k)$ is increased by $1/n$. Since $P(\Delta+1) \leq P(k)$, we have the conditions of Lemma 1 and the entropy can only decrease.

If $\phi(y) = j$, then the path $x_i P \dot{y}$ is odd. Interchanging colors i and j on this path decreases $P(i)$ by $1/n$ and increases $P(j)$ by $1/n$. Recoloring y decreases $P(j)$ by $1/n$ and recoloring x increases $P(i)$ by $1/n$. So the probability masses $P(i)$ and $P(j)$ are left unchanged by the operations. Overall, we only have that $P(\Delta+1)$ is decreased by $1/n$, and $P(k)$ is increased by $1/n$. Again, Lemma 1 holds and the entropy decreases.

$$C_{ij} = C_{ji} \text{ is a } x_i - x_j \text{ path.} \tag{7}$$

To show this, we assume that C_{ij} and C_{ji} are disjoint components of $F_{i,j}$. Then we can interchange colors i and j in C_{ij} and recolor x with color i. We again have two cases: either C_{ij} is an even path, and $P(\Delta+1)$ is decreased by $1/n$ and $P(i)$ is increased by $1/n$, or C_{ij} is an odd path, and $P(\Delta+1)$ is decreased by $1/n$ and $P(j)$ is increased by $1/n$. In both cases, Lemma 1 holds and the entropy decreases.

$$\text{For distinct } i,j,k \text{ we have } C_{ij} \cap C_{jk} = \{x_j\}. \tag{8}$$

Otherwise there would be a vertex $x_j \neq y \in C_{ij} \cap C_{jk}$ with $\phi(y) = j$ and two pairs of neighbors colored with i and k respectively. We could then apply the same three steps as in the proof of point (6).

Now if the neighbors x_i of x are pairwise adjacent, then G can only be the graph induced by x and $\{x_1, x_2, \ldots, x_\Delta\}$, because all vertices have maximum degree Δ. Hence G is the complete graph, and we do not have to show anything.

We may thus assume without loss of generality that $x_1 x_2 \notin E(G)$. Let y be the neighbor of x_1 in C_{12}, with $\phi(y) = 2$. Interchanging colors 1 and 3 in C_{13} we obtain a new coloring ϕ' of F. This coloring has the same entropy as ϕ, since C_{13} is an even path. We define x'_i and C'_{ij} with respect to the new coloring ϕ'. As a neighbor of $x_1 = x'_3$, the vertex y now lies in C'_{23}, for $\phi(y) = \phi'(y) = 2$. By (8), however, the path $\dot{x}_1 C_{12}$ retained its original coloring, so $y \in \dot{x}_1 C_{12} \subset C'_{12}$. Hence $y \in C'_{23} \cap C'_{12}$, contradicting (8). □

6 Acknowledgements

We thank Stefan Langerman and Gilles Van Assche for their valuable comments on an earlier version of this paper.

References

Accame M, Natale FD, Granelli F (2000) Efficient labeling procedures for image partition encoding. Signal Processing 80(6):1127–1131

Agarwal S, Belongie S (2002) On the non-optimality of four color coding of image partitions. In: Proc. IEEE Int. Conf. Image Processing

Alon N, Orlitsky A (1996) Source coding and graph entropies. IEEE Trans Inform Theory 42(5):1329–1339

Bellare M, Goldreich O, Sudan M (1995) Free bits, PCPs and non-approximability—towards tight results. In: 36th Annual Symposium on Foundations of Computer Science (Milwaukee, WI, 1995), IEEE Comput. Soc. Press, Los Alamitos, CA, pp 422–431

Brooks RL (1941) On colouring the nodes of a network. Proc Cambridge Philos Soc 37:194–197

Cardinal J, Fiorini S, Van Assche G (2004) On minimum entropy graph colorings. In: Proc. IEEE Int. Symposium on Information Theory, p 43

Chvátal V (1975) On certain polytopes associated with graphs. J Combinatorial Theory Ser B 18:138–154

Diestel R (2000) Graph theory, Graduate Texts in Mathematics, vol 173, 2nd edn. Springer-Verlag, New York

Erdös P, Hedetniemi ST, Laskar RC, Prins GCE (2003) On the equality of the partial Grundy and upper ochromatic numbers of graphs. Discrete Math 272(1):53–64, in honor of Frank Harary

Feige U, Kilian J (1998) Zero knowledge and the chromatic number. J Comput Syst Sci 57(2):187–199

Gabow HN (1990) Data structures for weighted matching and nearest common ancestors with linking. In: *1st* Annual ACM-SIAM Symposium on Discrete Algorithms, pp 434–443

Garey MR, Johnson DS, Miller GL, Papadimitriou CH (1980) The complexity of coloring circular arcs and chords. SIAM J Algebraic Discrete Methods 1(2):216–227

Golumbic MC (2004) Algorithmic graph theory and perfect graphs. Academic Press

Grötschel M, Lovász L, Schrijver A (1993) Geometric algorithms and combinatorial optimization, Algorithms and Combinatorics, vol 2, 2nd edn. Springer-Verlag, Berlin

Halperin E, Karp RM (2004) The minimum-entropy set cover problem. In: Automata, languages and programming, Lecture Notes in Comput. Sci., vol 3142, Springer, Berlin, pp 733–744

Jansen K (2000) Approximation results for the optimum cost chromatic partition problem. Journal of Algorithms 34:54–89

Jensen T, Toft B (1995) Graph Coloring Problems. Wiley Interscience

Kahn J, Kim J (1995) Entropy and sorting. J Comput Syst Sci 51(3):390–399

Marx D (2005) A short proof of the NP-completeness of minimum sum interval coloring. Oper Res Lett 33(4):382–384

Pemmaraju SV, Raman R, Varadarajan K (2004) Buffer minimization using max-coloring. In: SODA '04: Proceedings of the fifteenth annual ACM-SIAM symposium on discrete algorithms, Society for Industrial and Applied Mathematics, pp 562–571

Simonyi G (2001) Perfect graphs and graph entropy. An updated survey. In: Perfect graphs, Wiley-Intersci. Ser. Discrete Math. Optim., Wiley, Chichester, pp 293–328

Tucker A (1980) An efficient test for circular-arc graphs. SIAM J Comput 9(1):1–24

de Werra D, Glover F, Silver EA (1995) A chromatic scheduling model with costs. IIE Trans 27:181–189

de Werra D, Hertz A, Kobler D, Mahadev NVR (2000) Feasible edge colorings of trees with cardinality constraints. Discrete Math 222(1-3):61–72

Witsenhausen HS (1976) The zero-error side information problem and chromatic numbers. IEEE Trans Inform Theory 22(5):592–593

Zhao Q, Effros M (2003) Low complexity code design for lossless and near-lossless side information source codes. In: Proc. IEEE Data Compression Conf.