Information and Computation 222 (2013) 265–277

Contents lists available at SciVerse ScienceDirect

Information and Computation

www.elsevier.com/locate/yinco

Vertex cover in graphs with locally few colors ☆

Fabian Kuhn [a], Monaldo Mastrolilli [b,*]

[a] Faculty of Informatics, University of Lugano (USI), Via G. Buffi 13, 6904 Lugano, Switzerland
[b] Dalle Molle Institute for Artificial Intelligence (IDSIA), 6928 Manno, Switzerland

ARTICLE INFO

Article history:
Available online 2 November 2012

ABSTRACT

Erdős et al. defined the local chromatic number of a graph as the minimum number of colors that must appear within distance 1 of a vertex. For any $\Delta \geqslant 2$, there are graphs with arbitrarily large chromatic number that can be colored so that (i) no vertex neighborhood contains more than Δ different colors (*bounded local colorability*), and (ii) adjacent vertices from two color classes induce a complete bipartite graph (*biclique coloring*).

We investigate the weighted vertex cover problem in graphs when a locally bounded coloring is given. This generalizes the vertex cover problem in bounded degree graphs to a class of graphs with arbitrarily large chromatic number. Assuming the Unique Game Conjecture (UGC), we provide a tight characterization. We prove that it is UGC-hard to improve the approximation ratio of $2 - 2/(\Delta + 1)$ on $(\Delta + 1)$-locally (but not necessarily biclique) colorable graphs. A matching upper bound is also provided. Vice versa, when properties (i) and (ii) hold, we present a randomized algorithm with approximation ratio of $2 - \Omega(1)\frac{\ln\ln\Delta}{\ln\Delta}$. This matches known inapproximability results for the special case of bounded degree graphs.

Moreover, we show that when both the above two properties (i) and (ii) hold, the obtained result finds a natural application in a classical scheduling problem, namely the precedence constrained single machine scheduling problem to minimize the total weighted completion time, denoted as $1|prec|\sum w_j C_j$ in standard scheduling notation. In a series of recent papers it was established that this scheduling problem is a special case of the minimum weighted vertex cover in graphs $G_\mathbf{P}$ of incomparable pairs defined in the dimension theory of partial orders. We show that $G_\mathbf{P}$ satisfies properties (i) and (ii) where $\Delta - 1$ is the maximum number of predecessors (or successors) of each job.

1. Introduction

Vertex cover is one of the most studied problems in combinatorial optimization: Given a graph $G = (V, E)$ with weights w_i on the vertices, find a subset $V' \subseteq V$, minimizing the objective function $\sum_{i \in V'} w_i$, such that for each edge $\{u, v\} \in E$, at least one of u and v belongs to V'.

The related bibliography is vast and cannot be covered in this introductory note. We mention here that vertex cover cannot be approximated within a factor of 1.3606 [2], unless P = NP. Moreover, if the Unique Game Conjecture (UGC) [3] holds, Khot and Regev [4] show that vertex cover is hard to approximate within a factor of $2 - \varepsilon$ for any constant $\varepsilon > 0$. On the other side several simple 2-approximation algorithms are known (see e.g. [5,6]). Hochbaum [6] uses the natural

☆ A preliminary version of this paper appeared in the Proceedings of the 38th Int. Colloquium on Automata, Languages and Programming (ICALP) 2011.

* Corresponding author.
E-mail addresses: fabian.kuhn@usi.ch (F. Kuhn), monaldo@idsia.ch (M. Mastrolilli).

0890-5401/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ic.2011.12.008

linear program (LP) relaxation and a threshold rounding approach to obtain better than 2 approximation algorithms when a k-coloring of the graph is given as input. An optimal solution to the LP assigns a non-negative real value to each vertex of the input graph G. It is well known that such a solution is half-integral [5]. Let S_1 be the vertices of the input graph which attain value 1 in this solution; and let S_2 be the vertices which attain value 1/2. The vertices S_1 together with a cover of the subgraph induced by S_2 are sufficient to cover the whole graph G. A k-coloring of the subgraph induced by S_2 gives an independent set of value at least $w(S_2)/k$, where $w(S_2)$ is the sum of the vertex weights in S_2. This yields a $(2-2/k)$-approximation for the minimum weighted vertex cover problem. For graphs with degree bounded by d, this leads to a $(2-2/d)$-approximation [6] by Brooks' theorem (see e.g. [7]).

This basic approach has been considerably improved (for sufficiently large d) by Halperin [8]. The improvement is obtained by replacing the LP relaxation with a stronger semidefinite program (SDP) relaxation, and using a fundamental result by Karger, Motwani and Sudan [9]. The algorithm in [8] achieves a performance ratio of $2-(1-o(1))\frac{2\ln\ln d}{\ln d}$, which improves the previously known [10] ratio of $2-\frac{\ln d+O(1)}{d}$. Under the UGC [3], Austrin, Khot and Safra [11] have recently proved that it is NP-hard to approximate vertex cover in bounded degree graphs to within a factor $2-(1+o(1))\frac{2\ln\ln d}{\ln d}$. This exactly matches the algorithmic result of Halperin [8] up to the $o(1)$ term. For general vertex cover, the currently best approximation ratio is due to Karakostas [12], achieving a performance of $2-\Theta(1/\sqrt{\ln n})$.

The famous Brooks' theorem states that for a graph in which every vertex has at most d neighbors, the vertices may be colored with only d colors, except for two cases, complete graphs and cycle graphs of odd length, which require $d+1$ colors. In this paper we consider the vertex cover problem in graphs with bounded local chromatic number, a generalization of the bounded degree case with arbitrarily large chromatic number.

For an undirected graph $G=(V,E)$ and a vertex $u\in V$, we use $N(u):=\{v\in V\colon\ \{u,v\}\in E\}$ to denote the set of neighbors of u. Given a graph G, a valid vertex coloring of G is a function $\varphi: V\to\mathbb{N}$ such that $\varphi(u)\neq\varphi(v)$ whenever $\{u,v\}\in E$. In [1], Erdős et al. introduced the notion of local colorings.

Definition 1 *(Local Coloring).* Let k be a positive integer. A k-local coloring of a graph G is a valid vertex coloring φ such that for every vertex $u\in V$, $|\{\varphi(v)\colon\ v\in N(u)\}|<k$.

Note that the definition requires the number of colors in each neighborhood $N(u)$ to be strictly less than k. This guarantees that in each closed neighborhood $\{u\}\cup N(u)$, the number of different colors is bounded by k. The *local chromatic number* $\psi(G)$ of a graph G is the minimum k such that G admits a k-local coloring [1]. Since any valid coloring with k colors also is a local k-coloring, clearly, $\psi(G)\leqslant\chi(G)$. Interestingly, it can also be shown that the local chromatic number is always at least as large as the fractional chromatic number, i.e., $\psi(G)\geqslant\chi_f(G)$ [13]. For the definition and basic properties of the fractional chromatic number we refer to [14].

Given a valid vertex coloring φ and an integer i, let $C_i:=\{v\in V\colon\ \varphi(v)=i\}$ be the set of vertices with color i. Further, for integers $i\neq j$, we use $N_j(C_i):=C_j\cap\bigcup_{v\in C_i}N(v)$ to denote the vertices with color j that have a neighbor with color i. We consider colorings with the following density condition.

Definition 2 *(Biclique Coloring).* A coloring φ of a graph G is called a biclique coloring if for any two colors i and j, the subgraph induced by $N_i(C_j)$ (i.e. the i-colored set of neighbors of C_j) and $N_j(C_i)$ (i.e. the j-colored set of neighbors of C_i) is either empty or a complete bipartite graph.

(Note that in a biclique coloring if $\{x,y\}\in E$ and $\{w,z\}\in E$ with $\{x,w\}\in C_i$ and $\{y,z\}\in C_j$ then $\{x,z\}\in E$ and $\{w,y\}\in E$ as well; analogously, if you consider the graph induced by only the nodes in $C_i\cup C_j$ and remove isolated vertices, then the resulting graph is either empty or a complete bipartite graph.)

We will consider vertex cover in graphs for which a local biclique coloring φ is given. For any fixed $k\geqslant 3$, it is shown in [1] that there are n-vertex graphs with local chromatic number k and chromatic number $\Theta(\log\log n)$ and also that the gap between the local chromatic number and the chromatic number cannot be larger than this. This is shown using biclique colorings (cf. Definition 1.3 and Lemma 1.1 in [1]). Consequently, there are graphs that have chromatic number $\Theta(\log\log n)$ and admit a 3-local biclique coloring.

1.1. Our contribution

In this paper we study the vertex cover problem in graphs with bounded local colorings. Assuming the UGC [3], the provided results give a tight characterization of the problem. The two main results are summarized as follows.

Theorem 1. *The vertex cover problem in graphs $G=(V,E)$ for which a $(\Delta+1)$-local biclique coloring φ of G is given as input admits a randomized polynomial-time algorithm with approximation ratio $2-\Omega(1)\frac{\ln\ln\Delta}{\ln\Delta}$.*

Theorem 2. *Assuming the UGC, it is NP-hard to approximate the vertex cover problem in graphs for which a $(\Delta+1)$-local (but not necessarily biclique) coloring is given as input, within any constant factor better than $2-2/(\Delta+1)$.*

The result stated in Theorem 1 matches (up to the constant factor in the lower order term) a known inapproximability result [11]. In Section 3 we provide a matching upper bound for the inapproximability result of Theorem 2.

Besides generalizing the bounded degree case to a class of graphs with arbitrarily large chromatic number, we show that Theorem 1 finds a natural application in a classical scheduling problem with precedence constraints, known as $1|prec|\sum w_jC_j$ in standard scheduling notation (see Section 1.3 for the definition and a review of the literature). This is a very basic and fundamental scheduling problem that asks for computing a total ordering of the jobs that complies with the precedence constraints (or poset $\mathbf{P}$) among the jobs, while minimizing the weighted sum of jobs completion times. In a series of papers [15–17] it was established that this scheduling problem is a special case of minimum weighted vertex cover in graphs $G_{\mathbf{P}}$ of *incomparable pairs*[1] defined in the dimension theory of partial orders [18,19]. We prove the following in this paper.

Theorem 3. *For any graph of incomparable pairs $G_{\mathbf{P}}$, a $(\Delta+1)$-local biclique coloring of $G_{\mathbf{P}}$ can be computed in polynomial time, where $\Delta-1$ is the maximum number of predecessors (or successors) of each job.*

Together with Theorem 1, this result improves the previously best $(2-2/\max\{\Delta,2\})$-approximation algorithm described in [20].

Corollary 4. *If the maximum number of predecessors (or successors) of each job is at most Δ, there is a randomized approximation algorithm for the precedence constrained single machine scheduling problem $1|prec|\sum w_jC_j$ with approximation ratio*

$$2-\Omega(1)\frac{\log\log\Delta}{\log\Delta}.$$

1.2. Review of the SDP approach for bounded degree graphs

In [9] the authors consider the problem of coloring k-colorable graphs as a function of the graph maximum degree using semidefinite programming. Given a graph $G=(V,E)$ on n vertices, and a real number $k\geqslant 2$, a *vector k-coloring* [9] of G is an assignment of unit vectors $v_i\in\mathbb{R}^n$ to each vertex $i\in V$, such that for any two adjacent vertices i and j the dot product of their vectors satisfies the inequality $v_i\cdot v_j\leqslant -1/(k-1)$. They show that it is possible to check if a graph admits a vector k-coloring by using semidefinite programming. Moreover, they prove that vector k-colorable graphs have a "large" independent set when k is "small".

Theorem 5. *(See [9].) For every integral $k\geqslant 2$, a vector k-colorable graph $G=(V,E)$ with maximum degree d has an independent set I of value*

$$\Omega\left(\frac{w(V)}{d^{1-2/k}\sqrt{\ln d}}\right),$$

where $w(V)$ is the sum of the vertex weights.

The result in Theorem 5 has been obtained by generalizing the technique of *Rounding via Hyperplane Partitions* [21] to the technique of *Rounding via Vector Projections* [9]. More precisely, the independent set I claimed in Theorem 5 is obtained as follows. Suppose we have a vector k-coloring assigning unit vectors v_i to the vertices. Fix a parameter $c=c(k,d)$ and choose a random n-dimensional vector r from the n-dimensional standard normal distribution. Consider the set I' of vertices i with $v_i\cdot r\geqslant c$. The independent set I consists of all the isolated vertices in the subgraph induced by the vertex set I'. Intuitively, for small k the vector k-colorability implies that endpoints of an edge have vectors pointing away from each other. If the vector associated with a vertex has a large dot product with r, then the vector corresponding to an adjacent vertex will not have such a large dot product with r and hence will not be selected. Thus, only a relatively small fraction of the edges is likely to be in the induced subgraph on the selected set of vertices. The independent set I is large iff in the induced subgraph the expected number of edges is a small fraction of n. The latter is guaranteed by the rounding technique and analysis in [9] assuming that the graph has "few", i.e., $nd/2$, edges (a property that does not hold for the problem addressed in the present paper).

In [22], it is proved that the following program is a semidefinite relaxation of the vertex cover problem. Moreover, it can be solved within an additive error of $\varepsilon>0$ in polynomial time in $\ln\frac{1}{\varepsilon}$ and n using the ellipsoid method.

[1] $G_{\mathbf{P}}$ is the graph that has a node for each pair (i,j) of jobs with $(i,j)\notin\mathbf{P}$. Two nodes (i,j) and (k,ℓ) are adjacent in $G_{\mathbf{P}}$ if $(i,\ell),(k,j)\in P$ (see Section 4 for more details).

$$\begin{aligned} \min \quad & \sum_{u=1}^{n} w_u \frac{1+v_0\cdot v_u}{2} \\ \text{s.t.} \quad & (v_i - v_0)(v_j - v_0) = 0, \quad \{i,j\} \in E \\ & \|v_u\| = 1, \quad u \in V \cup \{0\},\ v_u \in \mathbb{R}^{n+1}. \end{aligned} \tag{1}$$

Note that in an "integral" solution of (1) (corresponding to a vertex cover), vectors for vertices that are picked coincide with v_0, while the other vectors coincide with $-v_0$. For general graphs, it is shown in [22] that the integrality gap is $2-\varepsilon$, for any $\varepsilon > 0$, i.e., for every $\varepsilon > 0$ there is a graph G_ε such that $vc(G_\varepsilon)/sd(G_\varepsilon)$ is at least $2-\varepsilon$, where $vc(G_\varepsilon)$ and $sd(G_\varepsilon)$ denote, respectively, the minimum vertex cover value and the optimum value of (1).

Halperin [8] uses (1) to provide an efficient randomized algorithm that approximates vertex cover in graphs with maximum degree d. The improvement is obtained as follows by using the following threshold rounding approach. Solve relaxation (1) and let

$$S_1 = \{u \in V \mid v_0 \cdot v_u \geqslant x\}, \tag{2}$$

$$S_2 = \{u \in V \mid -x \leqslant v_0 \cdot v_u < x\}, \tag{3}$$

where x is a positive number. As in Hochbaum's approach [6], it holds that the vertices S_1 together with a cover of the subgraph induced by S_2 are sufficient to cover the whole graph G (and this holds for any x). Moreover, for any two adjacent vertices i and j in the graph $G[S_2]$ induced by S_2, the edge constraint of (1) implies that $v_i \cdot v_j \leqslant -1 + 2x$. This has the important consequence that $G[S_2]$ is a vector k-colorable graph, where $k = \frac{2-2x}{1-2x}$ is close to 2 for small x. We can then use[2] Theorem 5 to obtain a large valued independent set I of $G[S_2]$. The returned vertex cover is $S_1 \cup (S_2 \setminus I)$ and the result of [8] follows by choosing a suitable value for x.

1.3. Literature review of the scheduling problem

The problem we consider in this paper is a classical problem in scheduling theory, known as $1|prec|\sum w_j C_j$ in standard scheduling notation (see e.g. Graham et al. [23]). It is defined as the problem of scheduling a set $N = \{1,\dots,n\}$ of n jobs on a single machine, which can process at most one job at a time. Each job j has a processing time p_j and a weight w_j, where p_j and w_j are non-negative integers. Jobs also have precedence constraints between them that are specified in the form of a *partially ordered set* (*poset*) $\mathbf{P} = (N, P)$, consisting of the set of jobs N and a partial order, i.e., a reflexive, antisymmetric, and transitive binary relation P on N, where $(i,j) \in P$, whenever $i \neq j$, implies that job i must be completed before job j can be started. The goal is to find a non-preemptive schedule which minimizes $\sum_{j=1}^{n} w_j C_j$, where C_j is the time at which job j completes in the given schedule.

The described problem was shown to be strongly NP-hard already in 1978 [24,25] by Lawler [24] and Lenstra and Rinnooy Kan [25]. For the general version of $1|prec|\sum w_j C_j$, several polynomial time 2-approximation algorithms are known [26,27,16,28,29]. Until recently, no inapproximability results were known, and closing the approximability gap has been listed as one of ten outstanding open problems in scheduling theory (e.g., [30]). Ambühl, Mastrolilli and Svensson [31] proved that the problem does not admit a PTAS, assuming that NP-complete problems cannot be solved in randomized subexponential time. Moreover, if a fixed cost present in all feasible schedules is ignored then the problem is as hard to approximate as vertex cover [31]. Recently, Bansal and Khot [32] showed that the gap for the general problem indeed closes assuming a variant of the UGC [3], by providing a $(2-\varepsilon)$-inapproximability result based on that assumption.

In a series of papers [15–17] it was proved that $1|prec|\sum w_j C_j$ is a special case of minimum weighted vertex cover in some special graphs $G_{\mathbf{P}}$ that depend on the input poset $\mathbf{P}$. More precisely, it is shown that any feasible solution to the vertex cover problem in graphs $G_{\mathbf{P}}$ can be turned in polynomial time into a feasible solution to $1|prec|\sum w_j C_j$ without deteriorating the objective value. This result was achieved by investigating different integer LP formulations and relaxations [33,16,17] of $1|prec|\sum w_j C_j$, using linear ordering variables δ_{ij} such that the variable δ_{ij} has value 1 if job i precedes job j in the corresponding schedule, and 0 otherwise.

Dushnik and Miller [34] introduced dimension as a parameter of partial orders in 1941. The dimension of a poset $\mathbf{P}$ is the minimum number of total orders such that their intersection is $\mathbb{P}$. The dimension is one of the most heavily studied parameters of partial orders, and many beautiful results have been obtained (see e.g. [19]).

There is a natural way to associate with a poset $\mathbf{P}$ a hypergraph $\mathbf{H_P}$, called the *hypergraph of incomparable pairs*, so that the dimension of $\mathbf{P}$ is the chromatic number of $\mathbf{H_P}$ [18]. Furthermore, the fractional dimension of $\mathbf{P}$, a generalization due to Brightwell and Scheinerman [35] is equal to the fractional chromatic number of $\mathbf{H_P}$. It turns out [36] that graph $G_{\mathbf{P}}$ is the (ordinary) graph obtained by removing from $\mathbf{H_P}$ all edges of cardinality larger than two. This allows to apply the rich vertex cover theory to $1|prec|\sum w_j C_j$ together with the dimension theory of partial orders. One can, e.g., conclude that the scheduling problem with two-dimensional precedence constraints is solvable in polynomial time, as $G_{\mathbf{P}}$ is bipartite in

[2] In Theorem 5 the integrality assumption on k is not technically necessary, and it can be easily generalized to fractional k. As remarked in [8], by using exactly the same analysis and rounding technique as in [9], it is possible to compute an independent set of value at least $\Omega(\frac{w(S_2)}{d^{x/(1-x)}\sqrt{x \ln d}})$ for $k = \frac{2-2x}{1-2x}$.

this case [18,17], and the vertex cover problem is well-known to be solvable in polynomial time on bipartite graphs. This considerably extends Lawler's result [24] from 1978 for series-parallel precedence constraints. Further, these connections between the $1|prec|\sum w_j C_j$ and the vertex cover problem on $G_{\mathbf{P}}$, and between dimension and coloring, yield a framework for obtaining $(2-2/f)$-approximation algorithms for classes of precedence constraints with bounded (fractional) dimension f [20,15]. The framework is inspired by Hochbaum's approach [6] for the vertex cover problem on "easily" colorable graphs. It yields the best known approximation ratios for all previously considered special classes of precedence constraints, like semi-orders, convex bipartite orders, interval orders, interval dimension 2, bounded in-degree posets.

2. Vertex cover using bounded local biclique colorings

In this section we provide and analyze an approximation algorithm for the vertex cover problem in graphs $G=(V,E)$ for which a $(\Delta+1)$-local biclique coloring φ of G is given.

The presented approximation algorithm follows the threshold rounding approach used for the bounded-degree vertex cover problem [8]: first solve the SDP (1) and, based on this solution and a parameter x that will be determined later, two vertex sets, S_1 and S_2, are computed as follows.

$$S_1 = \{u \in V \mid v_0 \cdot v_u \geqslant x\}, \tag{4}$$

$$S_2 = \{u \in V \mid -x \leqslant v_0 \cdot v_u < x\}. \tag{5}$$

The cover is obtained by picking the vertices in S_1 together with a cover of the subgraph induced by S_2.

However, unlike in [8], Theorem 5 does not apply to graph $G[S_2]$. Indeed, the graph degree is not bounded and the theorem does not generalize to these graphs since the rounding procedure and analysis in [9] are strongly based on the assumption that the graph has "few", i.e. $O(n)$ edges.

We show that graph $G[S_2]$ has a "large" independent set by using a new rounding procedure to compute it that works as follows. The vertices in S_2 are first grouped into overlapping *clusters*. For every two colors i and j (of the given coloring φ) such that the number of edges connecting vertices with the two colors is non-zero, there are two clusters $N_i(C_j)$ and $N_j(C_i)$. Note that $N_j(C_i) \neq \emptyset$ iff $N_i(C_j) \neq \emptyset$ in $G[S_2]$. Because we assume that the coloring φ is a local $(\Delta+1)$-coloring, each vertex $u \in V$ belongs to at most Δ clusters. Furthermore, for all i and j, clusters $N_i(C_j)$ and $N_j(C_i)$ are connected by complete bipartite sub-graphs (note that this property also holds when restricting the graph G to the vertex set S_2). It can be easily proved, that all vectors v_i corresponding to vertices from the same cluster almost point in the same direction. This essentially follows from having a biclique coloring and because, by the definition of the set S_2, vectors corresponding to adjacent vertices in S_2 are almost antipodal. For each cluster $N_i(C_j)$, we arbitrarily choose one *representative* vertex. Let R be the set of representatives of all clusters $N_i(C_j)$. Further, for a vertex u, let R_u be the set of representatives of clusters $N_i(C_j)$ for which $u \in N_i(C_j)$. Note that because every node u belongs to at most Δ clusters, we have $|R_u| \leqslant \Delta$. By only using the vectors of the representatives, we compute a subset $I' \subseteq S_2$ as follows: vertex $u \in S_2$ belongs to I' if and only if every representative $a \in R_u$ of u satisfies $v_a \cdot r \geqslant c$, where c is a parameter and r is a random $(n+1)$-dimensional vector r from the $(n+1)$-dimensional standard normal distribution, i.e., the components of r are independent Gaussian random variables with mean 0 and variance 1. We show that the set I' has a large independent set. Formally, we have

$$I' = \left\{ u \in S_2 \colon \bigwedge_{a \in R_u} v_a \cdot r \geqslant c \right\}. \tag{6}$$

The complete procedure is summarized in Algorithm 1.

Algorithm 1 Vertex Cover Approximation Algorithm.

---

1. Set $x = \frac{1-o(1)}{25} \cdot \frac{\ln\ln\Delta}{\ln\Delta}$, $c = (1+o(1)) \cdot \sqrt{\frac{2x}{1-25x} \ln\Delta}$.
2. Solve SDP (1).
3. Let $S_1 = \{u \in [n] \mid v_0 \cdot v_u \geqslant x\}$, $S_2 = \{u \in [n] \mid -x \leqslant v_0 \cdot v_u < x\}$.
4. Find an IS I in $G[S_2]$ as follows:
   (a) Compute the set R of representatives.
   (b) Let $R_u = \{a \in R \colon \{u,a\} \subseteq N_i(C_j)$ for any two colors i and j of the given coloring $\varphi\}$.
   (c) Choose a random vector r and define
   $$I' = \left\{ u \in S_2 \;\middle|\; \bigwedge_{a \in R_u} v_a \cdot r \geqslant c \right\}.$$
   (d) Get I by removing one vertex of every edge in $G[I']$ from I'.
5. Output the constructed vertex cover $S_1 \cup (S_2 \setminus I)$.

2.1. Analysis

We first show that for all $u \in S_2$, the vectors corresponding to representatives of u's clusters point in almost the same direction as the vector v_u corresponding to vertex u.

Lemma 6. *Let $u \in S_2$ be a vertex, $a \in R_u$ be the representative vertex of any cluster to which u belongs, and $x \in (0,1)$. We have $v_u \cdot v_a \geqslant 1 - 8x + o(x)$.*

Proof. From the biclique coloring condition, it follows that for some i and j, there is a vertex $b \in S_2 \cap N_i(C_j)$ that is a common neighbor of u and a. From the definition of the SDP (1) and the definition of S_2 in (5), we have $v_a \cdot v_b \leqslant -1 + 2x$ and $v_b \cdot v_u \leqslant -1 + 2x$. Hence, the angle between v_a and v_b and the angle between v_b and v_u are both not smaller than $\theta := \arccos(-1 + 2x)$. The angle between v_a and v_u is therefore at most $2\pi - 2\theta$, that is, $\arccos(v_u \cdot v_a) \leqslant 2\pi - 2\theta$. The claim now follows by observing that $v_u \cdot v_a \geqslant \cos(2\pi - 2\theta) = \cos(2\theta) = 1 - 8x - o(x)$ by using the well known trigonometry identity $\cos(2\theta) = 2\cos^2(\theta) - 1$. □

The expected size of the set I' is $\sum_{u \in S_2} \Pr[u \in I']$. Using (6), we have

$$\Pr\big[u \in I'\big] = \Pr\Bigg[\bigwedge_{a \in R_u} v_a \cdot r \geqslant c\Bigg]. \tag{7}$$

In the following we compute a lower bound on the above probability. The subsequent analysis uses some basic properties of the normal distribution. Let X be a standard normal random variable (X has mean 0 and variance 1). We use $\mathcal{N}(x)$ to denote the probability that X is at least x. Hence, $\mathcal{N}(x) := \Pr(X \geqslant x) = \int_x^\infty \phi(t)\,dt$, where $\phi(t) = e^{-t^2/2}/\sqrt{2\pi}$ is the density of X (see e.g. [37]).

Lemma 7. *For any $u \in S_2$ and any constant $\gamma > 0$*

$$\Pr\Bigg[\bigwedge_{a \in R_u} v_a \cdot r \geqslant c\Bigg] \geqslant \mathcal{N}(\alpha c) - \Delta \cdot \mathcal{N}\left(\frac{c}{\beta\sqrt{x}}\right),$$

where $\alpha = \gamma + 1 + o(1)$ and $\beta = \frac{4+o(1)}{\gamma}$.

Proof. Assume, without loss of generality, that $v_u = (1, 0, \dots, 0)$. By Lemma 6, for every $a \in R_u$, $v_u \cdot v_a \geqslant 1 - \delta$, where $\delta = 8x + o(x)$. Then the first component $v_{a,1}$ of $v_a = (v_{a,1}, v_{a,2}, \dots, v_{a,n+1})$ must be at least $1 - \delta$. Let $A = (1 - \delta, 0, \dots, 0)$ and $B_a = (0, v_{a,2}, \dots, v_{a,n+1})$ for any $a \in R_u$.

Since v_a is a unit length vector, we have that $(1-\delta)^2 + \sum_{j=2}^{n+1} v_{a,j}^2 \leqslant 1$, which gives an upper bound on the length of B_a:

$$\|B_a\|^2 = \sum_{j=2}^{n+1} v_{a,j}^2 \leqslant 2\delta - \delta^2 < 2\delta.$$

We then get

$$\begin{aligned}
\Pr\Bigg[\bigwedge_{a \in R_u} v_a \cdot r \geqslant c\Bigg] &\geqslant \Pr\Bigg[\bigwedge_{a \in R_u} \big(A \cdot r \geqslant (\gamma+1)c \wedge B_a \cdot r \geqslant -\gamma c\big)\Bigg] \\
&= \Pr\Bigg[A \cdot r \geqslant (\gamma+1)c \wedge \bigwedge_{a \in R_u} B_a \cdot r \geqslant -\gamma c\Bigg] \\
&= 1 - \Pr\Bigg[A \cdot r < (\gamma+1)c \vee \bigvee_{a \in R_u} B_a \cdot r < -\gamma c\Bigg] \\
&\geqslant 1 - \Pr\big[A \cdot r < (\gamma+1)c\big] - \sum_{a \in R_u} \Pr[B_a \cdot r < -\gamma c] \\
&= \Pr\Bigg[\frac{A}{\|A\|} \cdot r \geqslant \frac{(\gamma+1)c}{\|A\|}\Bigg] - \sum_{a \in R_u} \Pr\Bigg[-\frac{B_a}{\|B_a\|} \cdot r > \frac{\gamma c}{\|B_a\|}\Bigg] \\
&\geqslant \mathcal{N}\left(\frac{(\gamma+1)c}{1-\delta}\right) - \Delta \cdot \mathcal{N}\left(\frac{\gamma c}{\sqrt{2\delta}}\right) \\
&= \mathcal{N}(\alpha c) - \Delta \cdot \mathcal{N}\left(\frac{c}{\beta\sqrt{x}}\right).
\end{aligned}$$

The last inequality follows because $|R_u| \leqslant \Delta$ and because the sum of k independent Gaussian random variables with variances $\sigma_1^2, \dots, \sigma_k^2$ is a Gaussian random variable with variance $\sum_{i=1}^k \sigma_i^2$ (see e.g. [37]). Consequently, the dot product of a unit vector with r is a standard normal random variable. □

The total weight of the vertices that are removed from I' in step (d) of Algorithm 1 is upper bounded by the weight of vertices of the edges in the graph $G[I']$ induced by I'. The next lemma bounds the probability that a vertex $u \in S_2$ is in I' and that u has a neighbor $u' \in S_2$ that is also in I'.

Lemma 8. *Consider a vertex $u \in S_2$. The probability that u as well as some neighbor u' of u are in I' is upper bounded by*

$$\Pr\bigl[u \in I' \wedge \exists u' \in S_2\colon \{u, u'\} \in E \wedge u' \in I'\bigr] \leqslant \Delta \cdot \mathcal{N}\left(\frac{c}{\sqrt{x}}\right).$$

Proof. Recall that R_u is the set of representative vertices for u's clusters. Because we assume that we are given a local $(\Delta + 1)$-coloring, u can only have neighbors in at most Δ other clusters (note that u has a neighbor in cluster $N_j(C_i)$ if and only if $u \in N_i(C_j)$), and sets $N_j(C_i)$ and $N_i(C_j)$ form a biclique. We call this at most Δ other clusters where u can have neighbors, *neighboring clusters of u*. Let Q_u be the representative vertices of the at most Δ neighboring clusters of u.

Let A be the event that $u \in I'$ and let B be the event that there is at least one neighbor u' of u that is in I'. For the event $A \wedge B$ to happen we need that all the representative vertices for u are in I', namely it must hold that $\bigwedge_{a \in R_u} v_a \cdot r \geqslant c$, and if a neighbor u' of u is in I', with $u \in C_i$ and $u' \in N_j(C_i)$ for some j, it is necessary that the representative $b \in Q_u$ of $N_j(C_i)$ is also in I', i.e., $v_b \cdot r \geqslant c$. Given some vertex $b \in Q_u$, we define $a_b \in R_u$ such that $\{b, a_b\} \in E$. Note that such a vertex a_b exists because we are given a biclique coloring. Such a vertex exists because if b is the representative of a cluster $N_j(C_i)$, we have $u \in N_i(C_j)$ and therefore we can choose a_b to be the representative of cluster $N_i(C_j)$. It follows that the probability of event $A \wedge B$ can be bounded as follows.

$$\begin{aligned}
\Pr[A \wedge B] &\leqslant \Pr\left[\left(\bigwedge_{a \in R_u} v_a \cdot r \geqslant c\right) \wedge \left(\bigvee_{b \in Q_u} v_b \cdot r \geqslant c\right)\right] \\
&\leqslant \sum_{b \in Q_u} \Pr\bigl[(v_{a_b} \cdot r \geqslant c) \wedge (v_b \cdot r \geqslant c)\bigr] \\
&\leqslant \sum_{b \in Q_u} \Pr\left[\frac{v_{a_b} + v_b}{\|v_{a_b} + v_b\|} \cdot r \geqslant \frac{2c}{\|v_{a_b} + v_b\|}\right] \\
&\leqslant \Delta \cdot \mathcal{N}\left(\frac{c}{\sqrt{x}}\right).
\end{aligned}$$

The last inequality follows because the definition of SDP (1), $a_b, b \in S_2$, and $\{a_b, b\} \in E$ together imply that $\|v_{a_b} + v_b\| \leqslant 2\sqrt{x}$. $\square$

Based on Lemmas 7 and 8, we can now lower bound the expected size of the computed independent set I and we can thus obtain a bound on the expected approximation ratio of Algorithm 1.

Theorem 9. *Choosing $x = \frac{1-o(1)}{25} \cdot \frac{\ln\ln\Delta}{\ln\Delta}$, $c = (1 + o(1)) \cdot \sqrt{\frac{2x}{1-25x}\ln\Delta}$, and $\gamma = 4$, Algorithm 1 has an expected approximation ratio of $2 - \frac{2-o(1)}{25} \cdot \frac{\ln\ln\Delta}{\ln\Delta}$.*

Proof. For a set of vertices $U \subseteq V$, let $w(U)$ be the sum of the weights of the vertices in U. We start the proof by lower bounding the expected size of the computed independent set I of $G[S_2]$. By linearity of expectation and Lemmas 7 and 8, the expected size of I can be bounded as

$$\mathbb{E}\bigl[w(I)\bigr] \geqslant w(S_2) \cdot \left(\mathcal{N}(\alpha c) - \Delta \cdot \mathcal{N}\left(\frac{c}{\beta\sqrt{x}}\right) - \Delta \cdot \mathcal{N}\left(\frac{c}{\sqrt{x}}\right)\right), \tag{8}$$

where $\alpha = \gamma + 1 + o(1)$, $\beta = (4 + o(1))/\gamma$ and $\gamma > 0$ is some constant. We choose $\gamma = 4$, yielding $\alpha = 5 + o(1)$ and $\beta = 1 + o(1)$. Using Lemma 18 in Appendix A and the choice of x and c, the right-hand side of (8) can then be lower bounded by $w(S_2) \cdot \mu x$, for any constant $\mu > 0$.

For a set $U \subseteq V$ of vertices, we define $sd(U) := \sum_{u \in U} w_u \cdot \frac{1 + v_0 \cdot v_u}{2}$ to be the contribution of U to the objective value of SDP (1). For $u \in S_1$, we have $v_0 \cdot v_u \geqslant x$ and thus, following the analysis in [8],

$$sd(S_1) \geqslant \sum_{u \in S_1} w_u \frac{1 + x}{2}. \tag{9}$$

Similarly, for $u \in S_2$, it holds that $v_0 \cdot v_u \geqslant -x$ and therefore

$$sd(S_2) \geqslant \sum_{u \in S_2} w_u \frac{1 - x}{2}. \tag{10}$$

The expected approximation ratio of the algorithm can be upper bounded by

$$\begin{aligned}\max\left\{\frac{w(S_1)}{sd(S_1)},\frac{w(S_2)-w(I)}{sd(S_2)}\right\} &\overset{(9),(10)}{\leqslant} \max\left\{\frac{2}{1+x},\frac{w(S_2)-w(I)}{w(S_2)(1-x)/2}\right\}\\ &= \max\left\{2-2x+o(1),\left(2-\frac{2w(I)}{w(S_2)}\right)\bigl(1+x+o(1)\bigr)\right\}\\ &\overset{(8)}{\leqslant} 2-2x+o(1).\end{aligned}$$

The claim now follows from the choice of the parameter x. □

Remark. Repeating the construction of the independent set I (i.e., choosing the random vector r) sufficiently often, we can obtain the same approximation ratio (up to a slightly worse $o(1)$ term) with high probability.

3. Vertex cover in graphs with bounded local chromatic number

Consider the vertex cover problem in graphs G_Δ for which a local $(\Delta+1)$-coloring is given. Theorem 10 shows that the approximation ratio achievable from relaxation (1) is no better than $2-\frac{2}{\Delta+1}$ if the bounded local colorability property holds, but the biclique condition does not necessarily hold.

Theorem 10. *For any fixed $\Delta \geqslant 2$ and $\varepsilon > 0$, there is a local $(\Delta+1)$-colorable graph $G_{\Delta,\varepsilon}$ for which*

$$\frac{vc(G_{\Delta,\varepsilon})}{sd(G_{\Delta,\varepsilon})} \geqslant 2-\frac{2}{\Delta+1}-\varepsilon,$$

where $vc(G)$ and $sd(G)$ denote the size of a minimum weighted vertex cover of G and the solution value for the corresponding SDP relaxation (1), respectively.

Proof. The proof follows by reducing the graphs used by Kleinberg and Goemans [22] to show an integrality gap of $2-\varepsilon$ of relaxation (1) for general vertex cover and any $\varepsilon > 0$, to local $(\Delta+1)$-colorable graphs $G_{\Delta,\varepsilon} = (V,E)$.

Let γ and m be two parameters that depend on ε and Δ. The graph consists of $\Delta+1$ *classes* $C_1,\dots,C_{\Delta+1}$ of vertices. Every class is an independent set of 2^m vertices corresponding to the 2^m many m bit sequences of zeroes and ones. Therefore, $|V| = (\Delta+1)\cdot 2^m$. Two vertices from two different classes are joined by an edge if they differ in exactly $(1-\gamma)m$ bits. Every vertex has unit weight. This defines a $(\Delta+1)$-colorable graph $G_{\Delta,\varepsilon}$.

For any $i \in V \cup \{0\}$, a feasible solution of relaxation (1) for graph $G_{\Delta,\varepsilon}$ is obtained by setting vector v_i as in [22]. Let

$$\alpha = \frac{\varepsilon'}{2}, \qquad \beta = \sqrt{1-\alpha^2}, \quad \text{and} \quad \gamma = \frac{1}{2}-\frac{(1-\alpha)^2}{2\beta^2},$$

where $\varepsilon' \leqslant \varepsilon$ is a rational number. Note that this also implies that γ is a rational number and we can therefore choose m such that $(1-\gamma)m$ is an integer. The corresponding unit vectors in $\mathbb{R}^{m+1}$, for $i \in V \cup \{0\}$, are as follows: v_0 has all zeros, but the $(m+1)$st bit that is 1; For $i \in V$, $v_i^{(p)} = \beta/\sqrt{m}$ if the pth bit of i is 1, $v_i^{(p)} = -\beta/\sqrt{m}$ if it is zero, and $v_i^{(m+1)} = \alpha$.

It is easy to check that the defined vectors satisfy the constraints of (1) and

$$sd(G_{\Delta,\varepsilon}) \leqslant \frac{1}{2}(1+\alpha)\cdot(\Delta+1)\cdot 2^m. \tag{11}$$

As in [22], we make use of the following theorem by Frankl and Rödl [38].

Theorem 11. *(See [38].) Let $\mathcal{C}$ be a collection of m-bit strings, ξ a constant satisfying $0 < \xi < 1/2$, and d an even integer satisfying $\xi m < d < (1-\xi)m$. Then for some constant δ depending only on ξ, if $|\mathcal{C}| > (2-\delta)^m$, then $\mathcal{C}$ contains two strings with Hamming distance exactly d.*

Choose $\xi < \gamma$ and let δ denote the constant obtained by applying the above theorem. Let $d = (1-\gamma)m$, and choose m large enough so that d is an even integer and

$$(2-\delta)^m \leqslant \frac{\alpha}{\Delta+1}\cdot 2^m.$$

Now, consider two different classes, C_i and C_j, of vertices. We say that two vertices x and y, with $x \in C_i$ and $y \in C_j$, are *twins* if they have associated the same string of bits. Let $I = I_1 \cup I_2 \cup \dots \cup I_{\Delta+1}$ be some maximum independent set of $G_{\Delta,\varepsilon}$, where $I_i \subseteq C_i$, for $i \in [\Delta+1]$. For each $i \in [\Delta+1]$, let T_i denote the set of nodes from I_i that have at least one twin in a set I_j with $j \neq i$. By Theorem 11, in $G_{\Delta,\varepsilon}$ if for some i, set T_i is larger than $\frac{\alpha}{\Delta+1}\cdot 2^m$, then there must be two vertices, say

a and b, in T_i whose corresponding vertices in G are not independent; the latter (and by the definition of set T_i) would imply that there is a twin of a that is in I, which is therefore connected by an edge with b, a contradiction. Therefore, since I is an independent set, we obtain that $|T_i| \leqslant \frac{\alpha}{(\Delta+1)} \cdot 2^m$ for all $i \in [\Delta+1]$. Let $S_i = I_i \setminus T_i$. We have that

$$|I| = \sum_{i \in [\Delta+1]} |S_i| + \sum_{i \in [\Delta+1]} |T_i| \leqslant 2^m + \alpha \cdot 2^m,$$

where $\sum_{i \in [\Delta+1]} |S_i| \leqslant 2^m$ follows by the definition of S_i and observing that the largest set with no twins has size at most 2^m. This implies that the following bound on the optimal vertex cover holds.

$$vc(G_{\Delta,\varepsilon}) = |V| - |I| \geqslant (\Delta+1) \cdot 2^m \left(1 - \frac{1+\alpha}{\Delta+1}\right).$$

Therefore,

$$\frac{vc(G_{\Delta,\varepsilon})}{sd(G_{\Delta,\varepsilon})} \geqslant \frac{2}{1+\alpha} - \frac{2}{\Delta+1} \geqslant 2 - \frac{2}{\Delta+1} - \varepsilon. \qquad \square$$

Under the Unique Game Conjecture, Khot and Regev [4] proved that vertex cover is NP-hard to approximate better than $2 - \varepsilon$, for any $\varepsilon > 0$.

Theorem 12. *(See [4].) Assuming the Unique Game Conjecture, for arbitrarily small constants $\varepsilon, \delta > 0$, there is a polynomial time reduction mapping a SAT formula ϕ to an n-vertex graph G such that if ϕ is satisfiable then G has an independent set of size $(\frac{1}{2} - \varepsilon)n$ and if ϕ is unsatisfiable, then G has no independent set of size δn.*

By using Theorem 12, the reduction and the analysis in the proof of Theorem 10 can be easily adapted to obtain the following conditional hardness.

Theorem 13. *Assuming the Unique Game Conjecture, it is NP-hard to approximate the vertex cover problem in graphs for which a local $(\Delta+1)$-coloring is given as input, within any constant factor better than $2 - 2/(\Delta+1)$.*

Proof. Consider the graph G of Theorem 12 and let n denote the number of vertices of G. We show that there is a polynomial time reduction mapping G to a new graph $G_{\Delta,\varepsilon}$ that admits an easy $(\Delta+1)$-local coloring; Moreover, assuming the Unique Game Conjecture, if G has an independent set of size $n(\frac{1}{2} - \varepsilon)$ then there is a vertex cover in $G_{\Delta,\varepsilon}$ of size $(\Delta+1) \cdot n(1 - \frac{1}{2} + \varepsilon)$. Otherwise if the largest independent set in G has size at most δn then any vertex cover has size at least $(\Delta+1) \cdot n(1 - \frac{1}{\Delta+1} - \delta)$, and the claim follows by choosing δ and ε sufficiently small.

The graph construction is similar to the one given in the proof of Theorem 10. The graph $G_{\Delta,\varepsilon}$ consists of $\Delta+1$ *classes* $C_1, \ldots, C_{\Delta+1}$ of vertices. Every class is an independent set of n vertices, in each class each vertex corresponds to a vertex in G. Therefore, $|V| = (\Delta+1) \cdot n$ where V is the set of vertices in $G_{\Delta,\varepsilon}$. Two vertices from two different classes are joined by an edge in $G_{\Delta,\varepsilon}$ if the corresponding nodes in G are connected in G. Every vertex has unit weight. This defines a $(\Delta+1)$-colorable graph $G_{\Delta,\varepsilon}$.

Now, consider two different classes, C_i and C_j, of vertices. As in the proof of Theorem 10, we say that two vertices x and y, with $x \in C_i$ and $y \in C_j$, are *twins* if they correspond to the same original vertex in G. Let $I = I_1 \cup I_2 \cup \cdots \cup I_{\Delta+1}$ be some maximum independent set of $G_{\Delta,\varepsilon}$, where $I_i \subseteq C_i$, for $i \in [\Delta+1]$. For each $i \in [\Delta+1]$, let T_i denote the set of nodes from I_i that have at least one twin in a set I_j with $j \neq i$.

By Theorem 12, if the SAT formula ϕ is not satisfiable then the largest independent set in G has size δn. So, in this case, since I is an independent set, we obtain that $|T_i| \leqslant \delta n$ for all $i \in [\Delta+1]$. Let $S_i = I_i \setminus T_i$. We have that

$$|I| = \sum_{i \in [\Delta+1]} |S_i| + \sum_{i \in [\Delta+1]} |T_i| \leqslant n + \delta(\Delta+1)n,$$

where $\sum_{i \in [\Delta+1]} |S_i| \leqslant n$ follows by the definition of S_i and observing that the largest set with no twins has size at most n. This implies that the following bound on the optimal vertex cover holds.

$$vc^{UNSAT} = |V| - |I| \geqslant (\Delta+1) \cdot n\left(1 - \frac{1}{\Delta+1} - \delta\right).$$

If the sat formula is satisfiable then $I \geqslant (\Delta+1)(\frac{1}{2} - \varepsilon)n$ and

$$vc^{SAT} = |V| - |I| \leqslant (\Delta+1) \cdot n\left(1 - \frac{1}{2} + \varepsilon\right).$$

Therefore, by setting $\varepsilon = \delta$, we have

$$\frac{vc^{UNSAT}}{vc^{SAT}} \geqslant 2 - \frac{2}{\Delta + 1} - O(\varepsilon). \qquad \square$$

A matching upper bound can be obtained by showing that an independent set of value at least $\frac{\sum_{i=1}^{n} w_i}{\Delta+1}$ is computable in polynomial time. Indeed, the following result shows that Turan's theorem (as proved by Caro and Wei, see e.g. [39]) for bounded degree graphs can be generalized to graphs with bounded local colorings.

Theorem 14. *For any weighted graph $G = (V, E)$, there exists an independent set of value at least $\sum_{v\in V} \frac{w_v}{\Delta_v+1}$, where $\Delta_v = |\{c(u): u \in N(v)\}|$ is the number of different colors in the neighborhood of $v \in V$ of any proper coloring c of G.*

Proof. Consider a uniformly chosen total ordering $<$ of the color classes $\{C_1, \ldots, C_k\}$, where $C_i = \{v \in V: c(v) = i\}$. Define $I := \{v \in V: \{v, w\} \in E \wedge v \in C_i \wedge w \in C_j \rightarrow C_i < C_j\}$. Let X_v be the indicator random variable for $v \in I$ and $X = \sum_{v\in V} X_v = |I|$. For each v, $E[X_v] = Pr[v \in I] = \frac{1}{\Delta_v+1}$. Hence, there exists a specific ordering of the classes so that the corresponding independent set I has value at least $\sum_{v\in V} \frac{w_v}{\Delta_v+1}$. $\square$

By using standard techniques (see e.g. [40]), an upper bound that matches the lower bound of Theorem 2 can be obtained in deterministic polynomial time.

Corollary 15. *There exists a $(2 - \frac{2}{\Delta+1})$-approximation algorithm for the vertex cover problem in graphs for which a local $(\Delta + 1)$-coloring is given as input.*

4. The scheduling application

Problem $1|prec|\sum w_j C_j$ is a classical and fundamental problem in scheduling theory [41,30]. Its complexity certainly depends on the poset complexity: indeed, it can be efficiently solved when there are no precedence constraints or when the precedence constraints are not "very complicated", namely when the dimension of the poset is at most two. More generally, the dimension of the input poset has been established to be an important parameter for the approximability of the problem [20,15], with the lower the (fractional) dimension the better is the approximation ratio. Unfortunately, in the general case, recognizing the (fractional) dimension of a poset is hard even to approximate [42].

Another natural parameter of partial orders is given by the poset in- or out-degree [43], namely the job maximum number of predecessors or successors, respectively. One of the first NP-complete proofs [25] for $1|prec|\sum w_j C_j$ shows that the problem remains strongly NP-hard even if every job has at most two predecessors (or successors) in the poset. In [20], the authors present a $(2 - \frac{2}{\max\{\Delta,2\}})$-approximation algorithm, where $\Delta - 1$ is the minimum between the in- and the out-degree of the input poset. This gives a "good" approximation algorithm when the poset has a "small" Δ.

In this section we show two important properties of $G_{\mathbf{P}}$: (i) graph $G_{\mathbf{P}}$ can be efficiently colored so that no vertex neighborhood contains more than Δ different colors (*bounded local colorability*); and (ii) the coloring is a *biclique coloring*. This result together with Theorem 1 improves the previously known $(2 - 2/\max\{\Delta, 2\})$-approximation algorithm in [20].

A *partially ordered set* (*poset*) $\mathbf{P} = (N, P)$, consists of a set N and a partial order P on N, i.e., a reflexive, antisymmetric, and transitive binary relation P on N. For $x, y \in N$, we write $x \leqslant y$ when $(x, y) \in P$, and $x < y$ when $(x, y) \in P$ and $x \neq y$. When neither $(x, y) \in P$ nor $(y, x) \in P$, we say that x and y are incomparable, denoted by $x||y$. We call $\mathrm{inc}(\mathbf{P}) = \{(x, y) \in N \times N: x||y \text{ in } P\}$ the set of *incomparable pairs* of $\mathbf{P}$.

For any integer $k \geqslant 2$, a subset $S = \{(x_i, y_i): 1 \leqslant i \leqslant k\} \subset \mathrm{inc}(\mathbf{P})$ is called an *alternating cycle* when $x_i \leqslant y_{i+1}$ in P, for all $i = 1, 2, \ldots, k$, and where $y_{k+1} = y_1$. An alternating cycle $S = \{(x_i, y_i): 1 \leqslant i \leqslant k\}$ is *strict* if $x_i \leqslant y_j$ in P if and only if $j = i + 1$, for all $i, j = 1, 2, \ldots, k$.

Definition 3 *(Hypergraph and Graph of Incomparable Pairs).* (See [18].) For any given poset $\mathbf{P}$, the *hypergraph of incomparable pairs of* $\mathbf{P}$, denoted $\mathbf{H}_{\mathbf{P}} = (V, E)$, is the hypergraph that satisfies the following conditions:

1. The vertex set V is the set $\mathrm{inc}(\mathbf{P})$ of incomparable pairs of $\mathbf{P}$.
2. The edge set E consists of those subsets of V that form strict alternating cycles.

The *graph* $G_{\mathbf{P}}$ *of incomparable pairs of* $\mathbf{P}$ is the ordinary graph determined by all edges of size 2 in $\mathbf{H}_{\mathbf{P}}$.

Hence, in $G_{\mathbf{P}}$, there is an edge between incomparable pairs (i, j) and (k, ℓ) if and only if $(i, \ell), (k, j) \in P$.

In a series of papers [15,17,16], it was proved that $1|prec|\sum w_j C_j$ is equivalent to a weighted vertex cover problem on the graph of incomparable pairs $G_{\mathbf{P}}$ of the poset $\mathbf{P}$ characterizing the precedence constraints of the scheduling problem. More precisely, given a scheduling instance S with precedence constraints $\mathbf{P}$, let $G_{\mathbf{P}}^S$ be the *weighted* version of graph $G_{\mathbf{P}}$

with the following weights. For all incomparable pairs $(i, j) \in \mathrm{inc}(\mathbf{P})$, the weight of vertex (i, j) in $G_{\mathbf{P}}$ is $p_i \cdot w_j$, where p_i is the processing time of job i and w_j is the weight of process j.

Theorem 16. *(See [15,17,16].) For any given instance S of problem $1|prec|\sum w_j C_j$ with precedence constraints $\mathbf{P}$, an α-approximate solution to the weighted vertex cover in graphs $G^S_{\mathbf{P}}$ can be turned in polynomial time into an α-approximate solution to S.*

Let $\mathbf{P} = (N, P)$ be a poset. For any $j \in N$, define the *degree of* j, denoted $\deg(j)$, as the number of elements comparable (but not equal) to j in $\mathbf{P}$. Given $j \in N$, let $D(j)$ denote the set of all elements which are less than j, and $U(j)$ those which are greater than j in P. Let $\deg_D(j) := |D(j)|$ be the *in-degree* of j and the *maximum in-degree* $\Delta_D(\mathbf{P}) := \max\{\deg_D(j)\colon\ j \in N\}$. The *out-degree* of j $\deg_U(j)$ and the *maximum out-degree* $\Delta_U(\mathbf{P})$ are defined analogously (see also [43]). The maximum vertex degree in the graph of incomparable pairs $G_{\mathbf{P}}$ is bounded by $(\Delta_D(\mathbf{P}) + 1) \cdot (\Delta_U(\mathbf{P}) + 1)$. Hence, if both $\Delta_D(\mathbf{P})$ and $\Delta_U(\mathbf{P})$ are bounded, $G_{\mathbf{P}}$ has bounded degree and therefore, the bounded degree vertex cover approximation of [8] can be used to approximate the scheduling problem $1|prec|\sum w_j C_j$ with precedence constraints $\mathbf{P}$. If only either the in-degree or the out-degree of $\mathbf{P}$ is bounded, $G_{\mathbf{P}}$ does not have bounded degree. However, we will now show that in this case $G_{\mathbf{P}}$ has a good local biclique coloring.

Theorem 17. *Let $\mathbf{P} = (N, P)$ be any poset and let $\Delta = 1 + \min\{\Delta_D(\mathbf{P}), \Delta_U(\mathbf{P})\}$. Then, we can efficiently compute a $(\Delta + 1)$-local biclique coloring of $G_{\mathbf{P}} = (V, E)$.*

Proof. We assume that $\Delta - 1 = \Delta_D(\mathbf{P})$ is the largest in-degree. The case $\Delta - 1 = \Delta_U(\mathbf{P})$ can be proven analogously. We first show how to compute a $(\Delta + 1)$-local coloring of $G_{\mathbf{P}}$ (cf. Definition 1). Partition the incomparable pairs into $|N|$ color classes: $C_i = \{(i, j) \in \mathrm{inc}(\mathbf{P})\}$ for $i \in [|N|]$. It is easy to check that every C_i forms an independent set. Moreover $(k.j) \in P$ is a necessary condition for any incomparable pair $(i, j) \in \mathrm{inc}(\mathbf{P})$ to be adjacent to $(k, \ell) \in \mathrm{inc}(\mathbf{P})$. Since the in-degree of j is bounded by $\Delta - 1$, it follows that the number of distinct pairs (k, j) such that $(k, j) \in P$ is bounded by Δ (it is Δ and not $\Delta - 1$ because we also have to consider the pair $(j, j) \in P$). Therefore any incomparable pair (i, j) has neighbors in at most Δ clusters and thus the coloring is $(\Delta + 1)$-local.

In order to show that the coloring is a biclique coloring (cf. Definition 2), assume that $\{(i, a), (j, c)\} \in E$ and $\{(i, b), (j, d)\} \in E$. The claim follows by proving that $\{(i, a), (j, d)\} \in E$ and $\{(i, b), (j, c)\} \in E$. By the assumption we have: $(i, c) \in P$, $(j, a) \in P$, $(i, d) \in P$ and $(j, b) \in P$. Since $(j, a) \in P \wedge (i, d) \in P$ we have $\{(i, a), (j, d)\} \in E$, and $(i, c) \in P \wedge (j, b) \in P$ implies $\{(i, b), (j, c)\} \in E$. $\square$

5. Final remarks

This paper is about one of the very classic problem in combinatorial optimization. More precisely, we consider the minimum weighted vertex cover problem in graphs with bounded local chromatic number, a natural generalization of the bounded degree case with arbitrarily large chromatic number that naturally applies to scheduling with precedence constraints. Assuming the Unique Game Conjecture, a tight characterization of the weighted vertex cover problem in graphs with bounded local colorings is provided. We show that one needs precisely the two defined properties, namely bounded local colorability and biclique coloring, in order to obtain Halperin's bound.

Acknowledgments

The second author is indebted to Nikos Mutsanas and George Karakostas for many interesting discussions on the subject. We would also like to thank Jiři Sgall for suggesting the term "biclique coloring". The research is supported by the Swiss National Science Foundation project 200020-144491 "Approximation Algorithms for Machine Scheduling Through Theory and Experiments" and by Hasler Foundation Grant 11099.

Appendix A. On the independent set size

Let α, β, λ and μ be positive parameters. The goal of this section is to compute suitable values for x and c such that the following is guaranteed:

$$\left(\mathcal{N}(\alpha \cdot c) - \lambda \cdot \Delta \cdot \mathcal{N}\left(\frac{c}{\beta \cdot \sqrt{x}}\right)\right) \geqslant \mu x. \tag{A.1}$$

The proof of the following lemma is a rewriting of the analysis that appeared in [8,9] where the parameters α, β, μ and λ are all equal to one.

Lemma 18. *For any fixed $\alpha, \beta, \mu, \lambda > 0$ and sufficiently large Δ, it is possible to satisfy (A.1) by setting:*

$$c = \sqrt{\frac{2x\beta^2}{1-(\alpha\beta)^2 x} \ln \Delta}, \tag{A.2}$$

$$x = \frac{1-o(1)}{(\alpha\beta)^2} \cdot \frac{\ln\ln\Delta}{\ln\Delta}. \tag{A.3}$$

Proof. The claim follows by verifying that the following sufficient conditions hold when c and x are chosen as in (A.2) and (A.3), respectively:

$$\mathcal{N}(\alpha \cdot c) \geqslant 2\mu x, \tag{A.4}$$

$$\Delta \cdot \mathcal{N}\left(\frac{c}{\beta \cdot \sqrt{x}}\right) \leqslant \frac{\mu x}{\lambda}. \tag{A.5}$$

We will use the following bounds [37]:

$$\phi(c)\left(\frac{1}{c} - \frac{1}{c^3}\right) < \mathcal{N}(c) < \frac{\phi(c)}{c} \tag{A.6}$$

where

$$\phi(c) = \frac{1}{\sqrt{2\pi}} e^{-\frac{c^2}{2}}. \tag{A.7}$$

Condition (A.4): By (A.6) and since, for large Δ, we have $\alpha^2 c^2 > 2$, it follows that

$$\mathcal{N}(\alpha c) > \frac{\phi(\alpha c)}{\alpha c}\left(1 - \frac{1}{\alpha^2 c^2}\right) \geqslant \frac{\phi(\alpha c)}{2\alpha c}.$$

Therefore a sufficient condition to satisfy condition (A.4) is

$$\frac{\phi(\alpha c)}{\alpha c} = 4\mu x. \tag{A.8}$$

Condition (A.5): We can guarantee condition (A.5) by satisfying the following condition:

$$\Delta \cdot 4\alpha\beta\lambda\sqrt{x} \leqslant \frac{\phi(\alpha c)}{\phi(\frac{c}{\beta\sqrt{x}})}. \tag{A.9}$$

For large Δ we have $x \leqslant 1/(4\alpha\beta\lambda)^2$ and by (A.9) it is sufficient that

$$\Delta \leqslant \frac{\phi(\alpha c)}{\phi(\frac{c}{\beta\sqrt{x}})} = e^{-\frac{(\alpha c)^2}{2} + \frac{(c/\beta)^2}{2x}}$$

and condition (A.5) (using $x \leqslant 1/(\alpha\beta)^2$) is satisfied by setting c as in (A.2). By (A.2), it follows that condition (A.8) holds by choosing a value for x that satisfies the following equality:

$$\frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{\frac{2(\alpha\beta)^2 x}{1-(\alpha\beta)^2 x}\ln\Delta}{2}}}{\sqrt{\frac{2(\alpha\beta)^2 x}{1-(\alpha\beta)^2 x}\ln\Delta}} = 4\mu x. \tag{A.10}$$

By using $1/(1-(\alpha\beta)^2 x) = 1 + o(1)$ we have:

$$\frac{e^{-x(\alpha\beta)^2(1+o(1))\ln\Delta}}{\sqrt{4\pi x(\alpha\beta)^2(1+o(1))\ln\Delta}} = 4\mu x.$$

Let $y := x\ln\Delta$ and rewrite the equality:

$$e^{(\alpha\beta)^2(1+o(1))} y^{3/2} = \frac{\ln\Delta}{C}$$

where $C = 4\mu\sqrt{4\pi(\alpha\beta)^2(1+o(1))}$. By applying the logarithm:

$$(\alpha\beta)^2\big(1+o(1)\big)y + \frac{3}{2}\ln y = \ln\frac{\ln\Delta}{C}.$$

Note that for sufficiently large Δ, it is $(\alpha\beta)^2(1+o(1))y + \frac{3}{2}\ln y = (\alpha\beta)^2(1+o(1))y$. Therefore, there is an x that satisfies (A.10) such that (A.3) holds. $\square$

References

[1] P. Erdös, Z. Füredi, A. Hajnal, P. Komjáth, V. Rödl, Á. Seress, Coloring graphs with locally few colors, Discrete Mathematics 59 (1–2) (1986) 21–34.
[2] I. Dinur, S. Safra, On the hardness of approximating minimum vertex cover, Annals of Mathematics 162 (1) (2005) 439–485.
[3] S. Khot, On the power of unique 2-prover 1-round games, in: STOC, 2002, pp. 767–775.
[4] S. Khot, O. Regev, Vertex cover might be hard to approximate to within 2-epsilon, Journal of Computer and System Sciencies 74 (3) (2008) 335–349.
[5] G.L. Nemhauser, L.E. Trotter, Vertex packings: Structural properties and algorithms, Mathematical Programming 8 (1975) 232–248.
[6] D.S. Hochbaum, Efficient bounds for the stable set, vertex cover and set packing problems, Discrete Applied Mathematics 6 (1983) 243–254.
[7] R. Diestel, Graph Theory, 3rd edition, Graduate Texts in Mathematics, vol. 173, Springer-Verlag, Heidelberg, 2005. URL http://www.math.uni-hamburg.de/home/diestel/books/graph.theory/GraphTheoryIII.pdf.
[8] E. Halperin, Improved approximation algorithms for the vertex cover problem in graphs and hypergraphs, SIAM Journal on Computing 31 (5) (2002) 1608–1623.
[9] D.R. Karger, R. Motwani, M. Sudan, Approximate graph coloring by semidefinite programming, Journal of the ACM 45 (2) (1998) 246–265.
[10] M.M. Halldórsson, J. Radhakrishnan, Greed is good: Approximating independent sets in sparse and bounded-degree graphs, Algorithmica 18 (1) (1997) 145–163.
[11] P. Austrin, S. Khot, M. Safra, Inapproximability of vertex cover and independent set in bounded degree graphs, in: IEEE Conference on Computational Complexity, 2009, pp. 74–80.
[12] G. Karakostas, A better approximation ratio for the vertex cover problem, ACM Transactions on Algorithms 5 (4) (2009).
[13] J. Körner, C. Pilotto, G. Simonyi, Local chromatic number and Sperner capacity, Journal on Combinatorial Theory, Series B 95 (1) (2005) 101–117.
[14] E.R. Scheinerman, D.H. Ullman, Fractional Graph Theory, John Wiley and Sons Inc., 1997.
[15] C. Ambühl, M. Mastrolilli, Single machine precedence constrained scheduling is a vertex cover problem, Algorithmica 53 (4) (2009) 488–503.
[16] F.A. Chudak, D.S. Hochbaum, A half-integral linear programming relaxation for scheduling precedence-constrained jobs on a single machine, Operations Research Letters 25 (1999) 199–204.
[17] J.R. Correa, A.S. Schulz, Single machine scheduling with precedence constraints, Mathematics of Operations Research 30 (4) (2005) 1005–1021.
[18] S. Felsner, W.T. Trotter, Dimension, graph and hypergraph coloring, Order 17 (2) (2000) 167–177.
[19] W.T. Trotter, Combinatorics and Partially Ordered Sets: Dimension Theory, Johns Hopkins Series in the Mathematical Sciences, The Johns Hopkins University Press, 1992.
[20] C. Ambühl, M. Mastrolilli, N. Mutsanas, O. Svensson, Scheduling with precedence constraints of low fractional dimension, in: IPCO, 2007, pp. 130–144.
[21] M.X. Goemans, D.P. Williamson, Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, Journal of the ACM 42 (6) (1995) 1115–1145.
[22] J.M. Kleinberg, M.X. Goemans, The lovász theta function and a semidefinite programming relaxation of vertex cover, SIAM Journal on Discrete Mathematics 11 (2) (1998) 196–204.
[23] R. Graham, E. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, Optimization and approximation in deterministic sequencing and scheduling: A survey, in: Annals of Discrete Mathematics, vol. 5, North-Holland, 1979, pp. 287–326.
[24] E.L. Lawler, Sequencing jobs to minimize total weighted completion time subject to precedence constraints, in: Annals of Discrete Mathematics, vol. 2, 1978, pp. 75–90.
[25] J.K. Lenstra, A.H.G. Rinnooy Kan, The complexity of scheduling under precedence constraints, Operations Research 26 (1978) 22–35.
[26] A.S. Schulz, Scheduling to minimize total weighted completion time: Performance guarantees of LP-based heuristics and lower bounds, in: Proceedings of the 5th Conference on Integer Programming and Combinatorial Optimization (IPCO), 1996, pp. 301–315.
[27] L.A. Hall, A.S. Schulz, D.B. Shmoys, J. Wein, Scheduling to minimize average completion time: off-line and on-line algorithms, Mathematics of Operations Research 22 (1997) 513–544.
[28] C. Chekuri, R. Motwani, Precedence constrained scheduling to minimize sum of weighted completion times on a single machine, Discrete Applied Mathematics 98 (1–2) (1999) 29–38.
[29] F. Margot, M. Queyranne, Y. Wang, Decompositions, network flows and a precedence constrained single machine scheduling problem, Operations Research 51 (6) (2003) 981–992.
[30] P. Schuurman, G.J. Woeginger, Polynomial time approximation algorithms for machine scheduling: ten open problems, Journal of Scheduling 2 (5) (1999) 203–213.
[31] C. Ambühl, M. Mastrolilli, O. Svensson, Inapproximability results for sparsest cut, optimal linear arrangement, and precedence constrained scheduling, in: FOCS, 2007, pp. 329–337.
[32] N. Bansal, S. Khot, Optimal Long-Code test with one free bit, in: Foundations of Computer Science (FOCS), 2009, pp. 453–462.
[33] C.N. Potts, An algorithm for the single machine sequencing problem with precedence constraints, Mathematical Programming Study 13 (1980) 78–87.
[34] B. Dushnik, E. Miller, Partially ordered sets, American Journal of Mathematics 63 (1941) 600–610.
[35] G.R. Brightwell, E.R. Scheinerman, Fractional dimension of partial orders, Order 9 (1992) 139–158.
[36] C. Ambühl, M. Mastrolilli, O. Svensson, Approximating precedence-constrained single machine scheduling by coloring, in: Proceedings of APPROX + RANDOM, in: Lecture Notes in Computer Science, vol. 4110, Springer, 2006, pp. 15–26.
[37] W. Feller, An Introduction to Probability Theory and Its Applications, Wiley, 1968. URL http://www.worldcat.org/isbn/0471257087.
[38] P. Frankl, V. Rödl, Forbidden intersections, Transactions of the American Mathematical Society 300 (1) (1987) 259–286.
[39] N. Alon, J.H. Spencer, The Probabilistic Method, Wiley, New York, 1992.
[40] N. Alon, J.H. Spencer, The Probabilistic Method, Wiley–Interscience Series in Discrete Mathematics and Optimization, Wiley–Interscience, 2000. URL http://www.amazon.com/exec/obidos/redirect?tag=citeulike07-20&path=ASIN/0471370460.
[41] E.L. Lawler, J.K. Lenstra, A.H.G. Rinnooy Kan, D.B. Shmoys, Sequencing and scheduling: Algorithms and complexity, in: S.C. Graves, A.H.G. Rinnooy Kan, P. Zipkin (Eds.), Handbooks in Operations Research and Management Science, vol. 4, North-Holland, 1993, pp. 445–552.
[42] R. Hegde, K. Jain, The hardness of approximating poset dimension, Electronic Notes in Discrete Mathematics 29 (2007) 435–443.
[43] Felsner Trotter, On the fractional dimension of partially ordered sets, DMATH: Discrete Mathematics 136 (1994) 101–117. URL citeseer.ist.psu.edu/felsner93fractional.html.