Complexity and approximation results for the connected vertex cover problem in graphs and hypergraphs

Bruno Escoffier, Laurent Gourvès, Jérôme Monnot

Laboratoire d'Analyse et Modélisation de Systèmes pour l'Aide à la Décision
CNRS UMR 7024

CAHIER DU LAMSADE 262

Juillet 2007

Complexity and approximation results for the connected vertex cover problem in graphs and hypergraphs

Bruno Escoffier, Laurent Gourvès, Jérôme Monnot

Complexity and approximation results for the connected vertex cover problem in graphs and hypergraphs

Bruno Escoffier[*] *Laurent Gourvès*[*] *Jérôme Monnot*[*]

Abstract

We study a variation of the vertex cover problem where it is required that the graph induced by the vertex cover is connected. We prove that this problem is polynomial in chordal graphs, has a PTAS in planar graphs, is **APX**-hard in bipartite graphs and is 5/3-approximable in any class of graphs where the vertex cover problem is polynomial (in particular in bipartite graphs). Finally, dealing with hypergraphs, we study the complexity and the approximability of two natural generalizations.

Keywords: Connected vertex cover, chordal graphs, bipartite graphs, planar graphs, hypergraphs, **APX**-complete, approximation algorithm.

1 Introduction

In this paper, we study a variation of the vertex cover problem where the subgraph induced by any feasible solution must be connected. Formally, a vertex cover of a simple graph $G = (V, E)$ is a subset of vertices $S \subseteq V$ which covers all edges, *i.e.* which satisfies: $\forall e = \{x, y\} \in E$, $x \in S$ or $y \in S$. The vertex cover problem (MinVC in short) consists in finding a vertex cover of minimum size. MinVC is known to be **APX**-complete in cubic graphs [1] and **NP**-hard in planar graphs, [17]. MinVC is 2-approximable in general graphs, [3] and admits a polynomial approximation scheme in planar graphs, [5]. On the other hand, MinVC is polynomial for several classes of graphs such as bipartite graphs, chordal graphs, graphs with bounded treewidth, etc. [18, 7].

The connected vertex cover problem, denoted by MinCVC, is the variation of the vertex cover problem where, given a connected graph $G = (V, E)$, we seek a vertex cover S^* of minimum size such that the subgraph induced by S^* is connected. This problem has been introduced by Garey and Johnson [16], where it is proved to be **NP**-hard in planar graphs of maximum degree 4. As indicated in [25], this problem has some applications in the domain of wireless network design. In such a model, the vertices of the network are connected by transmission links. We want to place a minimum number of relay stations on vertices such that any pair of relay stations are connected (by a path which uses only relay stations) and every transmission link is incident to a relay station. This is exactly the connected vertex cover problem.

[*]CNRS-LAMSADE, Université Paris-Dauphine, Place du Maréchal De Lattre de Tassigny, F-75775 Paris Cedex 16, France. E-mail: {escoffier,laurent.gourves,monnot}@lamsade.dauphine.fr

1.1 Previous related works

The main complexity and approximability results known on this problem are the following: in [29], it is shown that MinCVC is polynomially solvable when the maximum degree of the input graph is at most 3. However, it is **NP**-hard in planar bipartite graphs of maximum degree 4, [14], as well as in 3-connected graphs, [30]. Concerning the positive and negative results of the approximability of this problem, MinCVC is 2-approximable in general graphs, [26, 2] but it is **NP**-hard to approximate within ratio $10\sqrt{5}-21$, [14]. Finally, recently the fixed-parameter tractability of MinCVC with respect to the vertex cover size or to the treewidth of the input graph has been studied in [14, 19, 23, 24, 25]. More precisely, in [14] a parameterized algorithm for MinCVC with complexity $O^*(2.9316^k)$ is presented improving the previous algorithm with complexity $O^*(6^k)$ given in [19] where k is the size of an optimal connected vertex cover. Independently, the authors of [23, 24] have also obtained FPT algorithms for MinCVC and they obtain in [24] an algorithm with complexity $O^*(2.7606^k)$. In [25], the author gives a parameterized algorithm for MinCVC with complexity $O^*(2^t \cdot t^{3t+2}n)$ where t is the treewidth of the graph and n the number of vertices.

MinCVC is related to the unweighted version of tree cover. The tree cover problem has been introduced in [2] and consists, given a connected graph $G=(V,E)$ with non-negative weights w on the edges, in finding a tree $T=(S,E')$ of G with $S \subseteq V$ and $E' \subseteq E$ which spans all edges of G and such that $w(T)=\sum_{e\in E'} w(e)$ is minimum. In [2], the authors prove that the tree cover problem is approximable within factor 3.55 (this ratio has been improved to 3 in [22]) and the unweighted version is 2-approximable. Recently, (weighted) tree cover has been shown to be approximable within a factor of 3 in [22], and a 2-approximation algorithm is proposed in [15]. Clearly, the unweighted version of tree cover is (asymptotically) equivalent to the connected version since S is a connected vertex cover of G iff there exists a tree cover $T'=(S,E')$ for some subset E' of edges. Since in this latter case, the weight of T' is $|S|-1$, the result follows.

1.2 Our contribution

In this article, we mainly deal with complexity and approximability issues for MinCVC in particular classes of graphs. More precisely, we first present some structural properties on connected vertex covers (Section 2). Using these properties, we show that MinCVC is polynomial in chordal graphs (Section 3). Then, in Section 4, we prove that MinCVC is **APX**-complete in bipartite graphs of maximum degree 4, even if each vertex of one block of the bipartition has a degree at most 3. On the other hand, if each vertex of this block of the bipartition has a degree at most 2 and the vertices of the other part have an arbitrary degree, then MinCVC is polynomial. Section 5 deals with the approximability of MinCVC. We first show that MinCVC is 5/3-approximable in any class of graphs where MinVC is polynomial (in particular in bipartite graphs, or more generally in perfect graphs). Then, we present a polynomial approximation scheme for MinVC in planar graphs. Section 6 concerns two natural generalization of the connected vertex cover problem in hypergraphs. We mainly prove that the first generalization, called the weak connected vertex cover problem, is polynomial in hypergraphs of maximum degree 3, and is $H(\Delta-1)-1/2$-approximable. Finally, we prove that the other generalization, called the strong connected vertex cover problem, is **APX**-hard, even in 2-regular hypergraphs.

Notation. All graphs considered are undirected, simple and without loops. Unless otherwise stated, n and m will denote the number of vertices and edges, respectively, of the graph $G = (V, E)$ considered. $N_G(v)$ denotes the *neighborhood* of v in G, ie., $N_G(v) = \{u \in V : \{u, v\} \in E\}$ and $d_G(v)$ its *degree* that is $d_G(v) = |N_G(v)|$. Finally, $G[S]$ denotes the subgraph of G induced by S.

2 Structural properties

We present in this subsection some properties on vertex covers or connected vertex covers. These properties will be useful in the rest of the article to devise polynomial algorithms that solve MinCVC either optimally (chordal graphs) or approximately (bipartite graphs,...).

2.1 Vertex cover and graph contraction

For a subset $A \subseteq V$ of a graph $G = (V, E)$, the *contraction* of G with respect to A is the simple graph $G_A = (V', E')$ where we replace A in V by a new vertex v_A (so, $V' = (V \setminus A) \cup \{v_A\}$) and $\{x, y\} \in E'$ iff either $x, y \notin A$ and $\{x, y\} \in E$ or $x = v_A$, $y \neq v_A$ and there exists $v \in A$ such that $\{v, y\} \in E$. The *connected contraction* of G following $V' \subseteq V$ is the graph $G^c_{V'}$ corresponding to the iterated contractions of G with respect to the connected components of V' (note that contraction is associative and commutative). Formally, $G^c_{V'}$ is constructed in the following way: let $A_1, \cdots, A_q$ be the connected components of the subgraph induced by V'. Then, we inductively apply the contraction with respect to A_i for $i = 1, \cdots, q$. Thus, $G^c_{V'} = G_{A_1 \circ \cdots \circ A_q}$. Finally, let $New(G^c_{V'}) = \{v_{A_1}, \cdots, v_{A_q}\}$ be the new vertices of $G^c_{V'}$ (those resulting from the contraction). The following Lemma concerns contraction properties that will, in particular, be the basis of the approximation algorithm presented in Subsection 5.1.

Lemma 1. *Let $G = (V, E)$ be a connected graph and let $S \subseteq V$ be a vertex cover of G. Let $G_0 = (V_0, E_0) = G^c_S$ be the connected contraction of G following S where $A_1, \cdots, A_q$ are the connected components of the subgraph induced by S. The following assertions hold:*

- (i) *G_0 is connected and bipartite.*
- (ii) *If $S = S^*$ is an optimal vertex cover of G, then $New(G_0)$ is an optimal vertex cover of G_0.*
- (iii) *If $S = S^*$ is an optimal vertex cover of G and $v \in V \setminus S^*$ with $d_{G^c_{S^*}}(v) \geq 2$, then $New(G_0)$ is an optimal vertex cover of $G_0 = G^c_{S^* \cup \{v\}}$.*

Proof. For (i), G_0 is connected since the contraction preserves the connectivity. Let $New(G_0)$ be the new vertices resulting from the connected contraction of G following S. By construction of the connected contraction, $New(G_0)$ is an independent set of G_0. Now, the remaining vertices of G_0 also forms an independent set since S is a vertex cover of G.

For (ii), since the contraction is associative, we only prove the result when $|A_1| = r \geq 2$ and $|A_2| = \cdots = |A_q| = 1$. By construction, $New(G_0)$ is obviously a vertex cover of G_0; thus $opt(G_0) \leq opt(G) - r + 1$. Conversely, Let S^*_0 be an optimal vertex cover of G_0. If $v_{A_1} \notin S^*_0$, then the neighborhood $N_{G_0}(v_{A_1})$ of v_{A_1} in G_0 verifies $N_{G_0}(v_{A_1}) \subseteq S^*_0$. So, $N_G(A_1) \setminus A_1 \subseteq S^*_0$, and if $v \in A_1$, then $S' = S^*_0 \cup (A_1 \setminus \{v\})$ is a vertex cover of G, hence $opt(G) \leq opt(G_0) + r - 1$. Otherwise, $v_{A_1} \in S^*_0$, and $S' = (S^*_0 \setminus \{v_{A_1}\}) \cup A_1$ is a vertex cover of G. Thus, $opt(G) \leq opt(G_0) + r - 1$. We conclude that $opt(G) = opt(G_0) + r - 1$ and the result follows.

For (iii), using (ii) and the associativity of the contraction, we only prove the result when S^* is also an independent set of G (in other words, we first apply the connected contraction following S^*); then, the connected components of the subgraph induced by $S^* \cup \{v\}$ satisfy $|A_1| = r \geq 3$ and $|A_2| = \cdots = |A_q| = 1$. Using the same argument as previously, on the one hand, we get $opt(G_0) \leq opt(G) - (r-1) + 1$ where $G_0 = G^c_{S^* \cup \{v\}}$ since $New(G_0)$ is a vertex cover of G_0; on the other hand, if $v_{A_1} \notin S_0^*$ (where S_0^* is an optimal vertex cover of G_0) then $S_0^* \cup \{v\}$ is a vertex cover of G, hence $opt(G) \leq opt(G_0) + 1 \leq opt(G_0) + (r-2)$. If $v_{A_1} \in S_0^*$, $(S_0^* \setminus \{v_{A_1}\}) \cup (A_1 \setminus \{v\})$ is a vertex cover of G and then $opt(G) \leq opt(G_0) + r - 2$. The proof is now complete.

□

2.2 Connected vertex covers and biconnectivity

Now, we deal with connected vertex covers. It is easy to see that if the removal of a vertex v disconnects the input graph (v is called a *cut-vertex*, or an *articulation point*), then v has to be in any connected vertex covers. In this section we show that, informally, solving MinCVC in a graph is equivalent to solve it on the biconnected components of the graph, under the constraint of including all cut vertices.

Formally, a connected graph $G = (V, E)$ with $|V| \geq 3$ is *biconnected* if for any two vertices x, y there exists a simple cycle in G containing both x and y. A *biconnected component* (also called *block*) $G_i = (V_i, E_i)$ is a maximal connected subgraph of G that is biconnected. For a connected graph $G = (V, E)$, V_c denotes the set of cut-vertices of G and $V_{i,c}$ its restriction to V_i.

Lemma 2. *Let $G = (V, E)$ be a connected graph. $S \subseteq V$ is a connected vertex cover of G iff for each biconnected component $G_i = (V_i, E_i)$, $i = 1, \cdots, p$, $S_i = S \cap V_i$ is a connected vertex cover of G_i containing $V_{i,c}$.*

Proof. Let $S \subseteq V$ be a connected vertex cover of a connected graph G. Obviously, $V_c \subseteq S$ since on the one hand, each biconnected component contains at least one edge, and on the other hand, the only vertices linking two distinct biconnected components are the cut-vertices. Moreover, trivially the restriction of S to V_i (ie., S_i) is a vertex cover of G_i containing $V_{c,i}$. Finally, if S_i is not connected in G_i, then there is two connected components $S_{i,1}$ and $S_{i,2}$ in the subgraph of G_i induced by S_i. By construction, there is a path μ which connects a vertex of $S_{i,1}$ to a vertex of $S_{i,2}$ and which only contains vertices of S (since S is connected). Thus, all vertices of μ (except its endpoints) are outside G_i. In this case, the subgraph $G_i + \mu$ would be biconnected, contradiction since G_i is assumed to be maximal.

Conversely, let S_i be a connected vertex cover of $G_i = (V_i, E_i)$ containing $V_{c,i}$ for $i = 1, \cdots, p$. Let us prove that $S = \cup_{i=1}^p S_i$ is a connected vertex cover of G. Obviously, S is a vertex cover of G since $E_1, \cdots, E_p$ is a partition of E. Moreover, since $S = \cup_{i=1}^p S_i$ contains V_c, the solution is connected. □

Lemma 2 allows us to characterize the optimal connected vertex covers of G.

Corollary 3. *Let $G = (V, E)$ be a connected graph. $S^* \subseteq V$ is an optimal connected vertex cover of G iff for each biconnected component $G_i = (V_i, E_i)$, $i = 1, \cdots, p$, $S_i^* = S^* \cap V_i$ is an* *optimal connected vertex cover of G_i among the connected vertex covers of G_i containing $V_{i,c}$.*

Proof. Let $S^* \subseteq V$ be an optimal connected vertex cover of G. If for some $i_0 \in \{1, \cdots, p\}$, $S^* \cap V_{i_0}$ is not an optimal connected vertex cover of G_{i_0} among the connected vertex covers of G_{i_0} containing $V_{i_0,c}$, then we deduce that there exists a vertex cover $S^*_{i_0}$ of G_{i_0} with $V_{i_0,c} \subseteq S^*_{i_0}$ and $|S^*_{i_0}| < |S^* \cap V_{i_0}|$ (since from Lemma 2, we know that $V_{i_0,c}$ is included in $S^* \cap V_{i_0}$). In this case, using one more time Lemma 2, we obtain that $S = (\cup_{j \neq i_0} S^* \cap V_j) \cup S^*_{i_0}$ is also a connected vertex cover of G with $|S| < |S^*|$, contradiction.

Conversely, let S^*_i be an optimal connected vertex cover of $G_i = (V_i, E_i)$ among the connected vertex covers of G_i containing $V_{i,c}$ for any $i = 1, \cdots, p$. if $S = \cup_{i=1}^{p} S^*_i$ is not an optimal connected vertex cover of G, then there exists another connected vertex cover S^* of G with $|S^*| < |S|$. Thus, we deduce that there exists at least one index $i_0 \in \{1, \cdots, p\}$, such that $|S^* \cap V_{i_0}| < |S^*_i|$. However, using Lemma 2, we know that $S^* \cap V_{i_0}$ is a connected vertex cover of G_{i_0} containing $V_{i_0,c}$, contradiction. □

For instance, using Corollary 3, we deduce that for the class of *trees* or *split graphs* MinCVC is polynomial. More generally, we will see in Section 3 that this result holds in chordal graphs. If we denote by MinPrextCVC (by analogy with the well known PreExtension Coloring problem) the variation of MinCVC where given $G = (V, E)$ and $V_0 \subseteq V$, we seek a connected vertex cover S of G containing V_0 and of minimal size, we obtain the following result:

Lemma 4. *Let $\mathcal{G}$ be a class of connected graphs defined by a hereditary property. Solving* MinCVC *in $\mathcal{G}$ polynomially reduces to solve* MinPrextCVC *in the biconnected graphs of $\mathcal{G}$. Moreover, if $\mathcal{G}$ is closed by pendent addition (ie., is closed under addition of a new vertex v and a new edge $\{u, v\}$ where $u \in V$), then they are polynomially equivalent.*

Proof. Let $G = (V, E) \in \mathcal{G}$ be a biconnected graph and $V_0 \subseteq V$, an instance of MinPrextCVC. By adding a new pendent edge for each vertex $v \in V_0$ (*i.e.*, a new vertex $v' \notin V$ and an edge $\{v, v'\}$), we obtain a new graph G' such that any connected vertex cover S' of G' contains V_0. Since $\mathcal{G}$ is assumed to be closed by pendent addition, then $G' \in \mathcal{G}$ and MinCVC is **NP**-hard in $\mathcal{G}$ if MinPrextCVC is **NP**-hard in the subclass of biconnected graphs of $\mathcal{G}$.

Conversely, given a graph $G \in \mathcal{G}$, we can compute the biconnected components G_i and the cut-vertices V_c of G in $O(n + m)$ time, see [27] for instance. Since the graph property is hereditary, we deduce $G_i \in \mathcal{G}$. Using Corollary 3, we deduce that if we had a polynomial algorithm which solves MinPrextCVC in the subclass of biconnected graphs of $\mathcal{G}$, then we could solve MinCVC in $\mathcal{G}$ in polynomial time. □

3 Chordal graphs

The class of *chordal graphs* is a very well known class of graphs which arises in many practical situations. A graph G is chordal if any cycle of G of size at least 4 has a chord (*i.e.*, an edge linking two non-consecutive vertices of the cycle). There are many characterizations of chordal graphs, see for instance [7]. It is well known that the vertex cover problem is polynomial in this class, [18].

In this section, we devise a polynomial time algorithm to compute an optimal connected vertex cover in chordal graphs. To achieve this, we need the following lemma.

Lemma 5. *Let $G = (V, E)$ be a connected chordal graph and let S be a vertex cover of G. The following properties hold:*

(i) The connected contraction $G_0 = (V_0, E_0) = G_S^c$ of G following S is a tree.

(ii) If G is biconnected, then S is a connected vertex cover of G.

Proof. Let S be a vertex cover of G.

For (i): from Lemma 1, we know that $G_0 = (V_0, E_0) = G_S^c$ is bipartite and connected. Assume that G_0 is not a tree, and let Γ be a cycle of G_0 with a minimal size. By construction, Γ is chordless, has a size at least 4 and alternates vertices of $New(G_0)$ and vertices of $V \setminus S$. From Γ, we can build a cycle Γ' of G using the following rule: if $\{x, v_{A_i}\} \in \Gamma$ and $\{v_{A_i}, y\} \in \Gamma$ where $x, y \notin S$ and $v_{A_i} \in New(G_0)$ (where we recall that A_i is some connected component of $G[S]$), then we replace these two edges by a shortest path $\mu_{x,y}$ from x to y in G among the paths from x to y in G which only use vertices of A_i (such a path exists since A_i is connected and is linked to x and y); by repeating this operation, we obtain a cycle Γ' of G with $|\Gamma'| \geq |\Gamma| \geq 4$. Let us prove that Γ' is chordless which will lead to a contradiction since G is assumed to be chordal. Let v_1, v_2 be two non consecutive vertices of Γ'. If $v_1 \notin S$ and $v_2 \notin S$, then $\{v_1, v_2\} \notin E$ since otherwise Γ would have a chord in G_0. So, we can assume that $v_1 \in (\mu_{x,y} \setminus \{x, y\})$ and $v_2 \in \mu_{x,y}$ (since there is no edge linking two vertices of disjoint paths $\mu_{x,y}$ and $\mu_{x',y'}$); in this case, using edge $\{v_1, v_2\}$, we could obtain a path which uses strictly less edges than $\mu_{x,y}$.

For (ii): Suppose that S is not connected. Then, from (i) we deduce that G_0 is not a star and thus, there are two edges $\{v_{A_i}, x\}$ and $\{x, v_{A_j}\}$ in G_0 where A_i and A_j are two connected components of S. We deduce that x would be a cut-vertex of G, contradiction since G is assumed to be biconnected. □

In particular, using (ii) of Lemma 5, we deduce that any optimal vertex cover S^* of a biconnected chordal graph G is also an optimal connected vertex cover.

Now, we give a simple linear algorithm for computing an optimal connected vertex cover of a chordal graph.

Theorem 6. MinCVC *is polynomial in chordal graphs. Moreover, an optimal solution can be found in linear time.*

Proof. Following Lemma 4, solving MinCVC in a chordal graph $G = (V, E)$ can be done by solving MinPrextCVC in each of the biconnected components $G_i = (V_i, E_i)$ of G. Since G_i is both biconnected and chordal, by Lemma 5, MinPrextCVC is the same problem as MinPrextVC (in G_i). But, by adding a pendent edge to vertices required to be taken in the vertex cover, we can easily reduce MinPrextVC to MinVC (note that the graph remains chordal). Since computing the biconnected components and solving MinVC in a chordal graph can be done in linear time (see [7]), the result follows. □

4 Bipartite graphs

A bipartite graph $G = (V, E)$ is a graph where the vertex set is partitioned into two independent sets L and R. Using the result of [14], we already know that MinCVC is **NP**-hard in planar bipartite graphs of maximum 4. Using Lemma 4, we can strengthen this result:

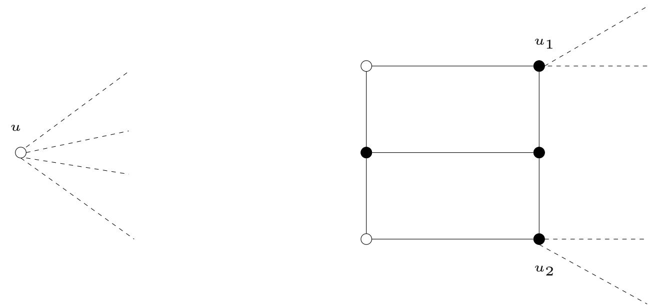

Figure 1: Local replacement of a vertex $u \in V_0$ by gadget $H(u)$.

Lemma 7. MinCVC *is* ***NP****-hard in biconnected planar bipartite graphs of maximum degree 4.*

Proof. Using the **NP**-hardness of MinCVC in bipartite planar graphs of maximum degree 4, given in [14], we only prove that MinPrextCVC in the subclass of biconnected bipartite graphs of maximum degree 4 can be polynomially reduced to MinCVC in the subclass of biconnected bipartite graphs of maximum degree 4. Note that the simple reduction given in Lemma 4 does not preserve the biconnectivity.

Let $G = (V, E)$ be a planar biconnected bipartite graph of maximum degree 4 and let V_0 an instance of MinPrextCVC. We replace each vertex $u \in V_0$ by the gadget $H(u)$ depicted in Figure 1. Actually, if the neighborhood of u is $N = \{v_1, \cdots, v_p\}$ with $2 \leq p \leq 4$ (since G is biconnected of maximum degree 4), then we link u_1 to some vertices of $v_1, \cdots, v_p$ and u_2 to the remaining vertices in such a way that on the one hand u_1 and u_2 have at least one neighbor in N and at most 2 neighbors in N, and on the other hand, the new graph remains planar. Let G' be the new graph. It is easy to see that G' is planar, bipartite, biconnected and of maximum degree 4.

Let S^* containing V_0 be an optimal connected vertex cover of G. Then, by deleting V_0 and by adding the vertices drawn in black for each gadget $H(u)$ (see Figure 1), we obtain a connected vertex cover of G'. Thus,

$$opt(G') \leq opt(G) + 3|V_0| \tag{1}$$

Conversely, let S' be a connected vertex cover of G'. It is easy to see that S' takes at least 4 vertices for each gadget $H(u)$. Thus, wlog., we can assume that S' only takes the black vertices for each gadget $H(u)$. By deleting these black vertices and by adding V_0, we obtain a solution S of G satisfying

$$|S| = |S'| - 3|V_0| \tag{2}$$

Using inequality (1) and equality (2), the expected result follows. □

Now, one can show that MinCVC has no PTAS in bipartite graphs of maximum degree 4.

Theorem 8. MinCVC *is not* 1.001031*-approximable in connected bipartite graphs* $G = (L, R; E)$ *where* $\forall l \in L$, $d_G(l) \leq 4$ *and* $\forall r \in R$, $d_G(r) \leq 3$, *unless* ***P****=****NP****.*

Proof. We give a reduction from the vertex cover problem in cubic graphs. In [10] it is proved that, given a connected cubic graph $G = (V, E)$ of n vertices, it is **NP**-hard to

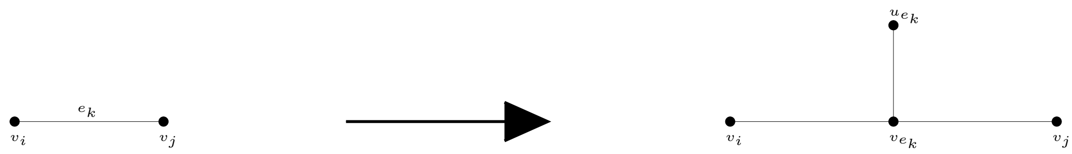

Figure 2: Local replacement of edge $e_k = \{v_i, v_j\}$ using gadget $H(e_k)$.

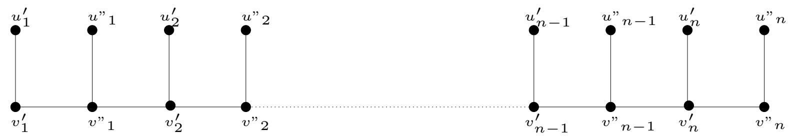

Figure 3: The graph H''.

decide whether $opt(G) \leq 0.5103305n$ or $opt(G) \geq 0.5154986n$ where $opt(G)$ is the value of an optimal vertex cover of G.

Given a cubic connected graph $G = (V, E)$ where $V = \{v_1, \cdots, v_n\}$ and $E = \{e_1, \cdots, e_m\}$ instance of MinVC, we build an instance $H = (V', E')$ of MinCVC in the following way.

- We start from G and each edge $e_k = \{v_i, v_j\}$ is replaced by the gadget $H(e_k)$ described in Figure 2. Let H' be this graph.
- We add the graph H'' depicted in Figure 3.
- Finally, we connect the graph H' to the graph H''. For each $i = 1, \cdots, n$, we link v_i to v'_i by using a gadget isomorphic to $H(e_k)$ (we denote by w_i the vertex of degree 3 in the gadget, ie the vertex v_{e_k} in Figure 2).

Clearly H is of maximum degree 4 and bipartite. Finally, we can easily observe that any vertex of this graph has degree at most 4 for one part of the bipartition and at most 3 for the other part.

Let S^* be an optimal vertex cover of G with value $opt(G)$. Clearly, $S^* \cup \{v_{e_k} : k = 1, \cdots, m\} \cup \{v'_i, v"_i, w_i : i = 1, \cdots, n\}$ is a connected vertex cover of H. Conversely, let S^* be an optimal connected vertex cover of H with value $opt(H)$. Wlog, we can assume that S^* contains $\{v_{e_k} : k = 1, \cdots, m\} \cup \{v'_i, v"_i, w_i : i = 1, \cdots, n\}$ since these vertices are cut vertices of H. Thus, $S = S^* \setminus (\{v_{e_k} : k = 1, \cdots, m\} \cup \{v'_i, v"_i, w_i : i = 1, \cdots, n\})$ is a vertex cover of G. Indeed, if an edge $e_k = \{v_i, v_j\}$ is not covered by S, then the vertex v_{e_k} will not be connected to the other vertices of S^*, which is impossible. Thus, we deduce:

$$opt(H) = opt(G) + m + 3n \tag{3}$$

Using the **NP**-hard gap of [10], the fact that G is cubic and equality (3), we deduce that it is **NP**-hard to decide whether $opt(H) \leq 5.0103305n$ or $opt(H) \geq 5.0154986n$. $\square$

In Theorem 8, we proved in particular that MinCVC is **NP**-hard when all the vertices of one part of the bipartition have a degree at most 3. It turns out that if all the vertices of one part of this bipartition have a degree at most 2, the problem becomes easy. This property will be very useful to devise our approximation algorithm in Subsection 5.1.

Lemma 9. MinCVC *is polynomial in bipartite graphs* $G = (L, R; E)$ *such that* $\forall r \in R$, $d_G(r) \leq 2$. *Moreover, if* $L_2 = \{l \in L : d_G(l) \geq 2\}$, *then* $opt(G) = |L| + |L_2| - 1$.

Proof. Let $G = (L, R; E)$ be a bipartite graph such that $\forall r \in R$, $d_G(r) \leq 2$ and assume that $|L| \geq 3$ and G is connected. Let $L_1 = L \setminus L_2$ and let $R_1 = N_G(L_1)$ be the neighbors of L_1. Let $G' = (L \setminus L_1, R \setminus R_1; E')$ be the bipartite subgraph of G induced by $(L \setminus L_1) \cup (R \setminus R_1)$ and let $G_{L_2} = (L_2, E_{L_2})$ where $e_r = \{l, l'\} \in E_{L_2}$ iff $\exists r \in R \setminus R_1$ with $\{l, r\} \in E'$ and $\{r, l'\} \in E'$. Finally, let T be a spanning tree of G_{L_2}.

We claim that $S_T = L_2 \cup R_1 \cup \{r \in R \setminus R_1 : e_r \in T\}$ is an optimal connected vertex cover of G.

Let S^* be an optimal connected vertex cover of G and let $L'_2 = N_G(R_1) \cap L_2$ be the neighbors of R_1 in G not in L_1. Clearly $R_1 \subseteq S^*$, since $|L| \geq 3$ and each vertex of L_1 has degree 1. Moreover, since each vertex of R has a maximum degree 2, then $L'_2 \subseteq S^*$. Now, let us prove that we can assume that $L_2 \subseteq S^*$. Assume the reverse and let $l_0 \in L_2 \setminus S^*$. Using the previous remark, we know that $l_0 \in L_2 \setminus L'_2$. Let $r_1, \cdots, r_q$ be the neighbors of l_0 in G. By construction, $q \geq 2$ and $r_i \in S^*$ since S^* is a vertex cover. Moreover, $\forall i = 1, \cdots, q$, $d_G(r_i) = 2$ since S^* must induce a connected subgraph and if l_i is the other neighbor of r_i, then $l_i \in S^*$. Let us prove that $S^* \setminus \{r_1\} \cup \{l_0\}$ is a connected vertex cover of G. First, $S^* \setminus \{r_1\}$ is a connected vertex cover in the subgraph $(L, R; E \setminus \{l_0, r_1\})$ since $S^* \setminus \{r_1\}$ is connected (r_1 is a leaf of the subgraph induced by S^*) and r_1 only covers edges $\{l_0, r_1\}, \{r_1, l_1\}$, but the edge $\{r_1, l_1\}$ is also covered by $l_1 \in S^*$. Then, by adding l_0, we now cover the missing edge $\{l_0, r_1\}$ and since $q \geq 2$, l_0 is linked to r_2 in $S^* \setminus \{r_1\} \cup \{l_0\}$. By repeating this operation, we obtain another optimal solution with $L_2 \subseteq S^*$. Thus, in S^*, we need to connect together the vertices of L_2 by using some vertices of R. Since the vertices of R_1 cannot link together vertices of L_2 (we recall that the degree of each vertex of R is at most 2), the vertices of $S^* \setminus L_2 \setminus R_1$ correspond to a set of edges $E^*_{L_2}$ in G_{L_2} such that the subgraph $(L_2, E^*_{L_2})$ of G_{L_2} is connected. Hence $|E^*_{L_2}| \geq |T|$ or equivalently $|S^* \setminus L_2 \setminus R_1| \geq |S_T \setminus L_2 \setminus R_1|$. In conclusion, S_T is an optimal connected vertex cover of G with value $opt(G) = |L_2| + |T| + |R_1| = 2|L_2| - 1 + |R_1|$ since T is a spanning tree of G_{L_2}. Now, observe that $|R_1| = |L_1|$ since otherwise G would not be connected, and the proof is complete. □

5 Approximation results

MinCVC is trivially **APX**-complete in k-connected graphs for any $k \geq 2$ since starting from graph $G = (V, E)$, instance of MinVC, we can add a clique K_k of size k and link each vertex of G to each vertex of K_k. This new graph G' is obviously k-connected and S is a vertex cover of G iff S union the k vertices of K_k (we can always assumed that $S \neq V$) is a connected vertex cover of G'. Thus, using the negative result of [21] it is quite improbable that one can improve the approximation ratio of 2 for MinCVC, even k-connected graphs. Thus, in this subsection we deal with the approximability of MinCVC in particular classes of graphs.

In Subsection 5.1, we devise a 5/3-approximation algorithm for any class of graphs where the classical vertex cover problem is polynomial. In Subsection 5.2, we show that MinCVC admits a PTAS in planar graphs.

5.1 When MinVC is polynomial

Let $\mathcal{G}$ be a class of connected graphs where MINVC is polynomial (for instance, the connected bipartite graphs). The underlying idea of the algorithm is simple: we first compute an optimal vertex cover, and then try to connect it by adding vertices (either using high degree vertices or Lemma 9). The analysis leading to the ratio 5/3 is based on Lemma 1 which deals with graph contraction.

Now, let us formally describe the algorithm. Recall that given a vertex set V', $G^c_{V'}$ denotes the connected contraction of V following V', and $New(G^c_{V'})$ denotes the set of new vertices (one for each connected component of $G[V']$).

`algo`$_{CVC}$ input: A graph $G = (V, E)$ of $\mathcal{G}$ with at least 3 vertices.

1 Find an optimal vertex cover S^* of G such that in $G^c_{S^*}$, $\forall v \in New(G^c_{S^*})$, $d_{G^c_{S^*}}(v) \geq 2$;

2 Set $G_1 = G^c_{S^*}$, $N_1 = New(G^c_{S^*})$, $S = S^*$ and $i = 1$;

3 While $|N_i| \geq 2$ and there exists $v \notin N_i$ such that v is linked in G_i to at least 3 vertices of N_i do

  3.1 Set $S := S \cup \{v\}$ and $i := i + 1$;
  3.2 Set $G_i := G^c_S$ and $N_i = New(G^c_S)$;

4 If $|N_i| \geq 2$, apply the polynomial algorithm of Lemma 9 on G_i (let S' be the produced solution) and set $S := S \cup (V \cap S')$;

5 Output S;

Now, we show that `algo`$_{CVC}$ outputs a connected vertex cover of G in polynomial time. First of all, given an optimal vertex cover S^* of a graph G (assumed here to be computable in polynomial time), we can always transform it in such a way that $\forall v \in New(G^c_{S^*})$, $d_{G^c_{S^*}}(v) \geq 2$. Indeed, if a vertex of $G^c_{S^*}$ corresponding to a connected component of S^* has only one neighbor in $G^c_{S^*}$, then we can take this neighbor in S^* and remove one vertex on this connected component (and the number of such 'leaf' connected components decreases, as soon as $G^c_{S^*}$ has at least 3 vertices). Now, using (ii) of Lemma 1, we know that $New(G^c_{S^*})$ is an optimal vertex cover of $G^c_{S^*}$. Then, from $New(G^c_{S^*})$, we can find such a solution within polynomial time.

Moreover, using (i) of Lemma 1 with S^*, we deduce that the graph G_i is bipartite, for each possible value of i. Assume that $G_i = (N_i; R_i, E_i)$ for iteration i where N_i is the left set corresponding to the contracted vertices and R_i is the right set corresponding to the remaining vertices and let p be the last iteration. Clearly, if $|N_p| = 1$, the the output solution S is connected. Otherwise, the algorithm uses step 4; we know that G_p is bipartite and by construction $\forall r \in R_p$, $d_{G_p}(r) \leq 2$. Thus, we can apply Lemma 9 on G_p. Moreover, a simple proof also gives that $\forall l \in N_p$, $d_{G_p}(l) \geq 2$. Indeed, otherwise there exists $l \in N_p$ such that l has a unique neighbor $r_0 \in R_p$. Let $\{x_1, \cdots, x_j\} \subseteq N_{p-1}$ with $j \geq 3$ and r_1 be the vertices contracted in G_{p-1} in order to obtain G_p. We conclude that the neighborhood

of $\{x_1, \cdots, x_j\}$ is $\{r_0, r_1\}$ in G_{p-1} which is impossible since on the one hand, N_{p-1} is an optimal vertex cover of G_{p-1} (using (iii) of Lemma 1), and on the other hand, by flipping $\{x_1, \cdots, x_j\}$ with $\{r_0, r_1\}$, we obtain another vertex cover of G_{p-1} with smaller size than N_{p-1}! Finally, using Lemma 9, an optimal connected vertex cover of G_p consists of taking N_p and $|N_p| - 1$ of R_p. In conclusion, S is a connected vertex cover of G.
We now prove that this algorithm improves the ratio 2.

Theorem 10. *Let $\mathcal{G}$ be a class of connected graphs where* MinVC *is polynomial. Then,* $\mathtt{algo}_{CVC}$ *is a* $5/3$*-approximation for* MinCVC *in* $\mathcal{G}$.

Proof. Let $G = (V, E) \in \mathcal{G}$. Let S be the approximate solution produced by $\mathtt{algo}_{CVC}$ on G. Using the previous notations and Lemma 9, the solution S has a value $apx(G)$ satisfying:

$$apx(G) = |S^*| + p - 1 + |N_p| - 1 \tag{4}$$

where p is the number of iterations of step 3. Obviously, we have:

$$opt(G) \geq |S^*| \tag{5}$$

Now let us prove that for any $i = 1, \cdots, p-1$, we also have $opt(G_i) \geq opt(G_{i+1}) + 1$. Let S_i^* be an optimal connected vertex cover of G_i. Let $r_i \in R_i$ be the vertex added to S during iteration i and let $N_{G_i}(r_i)$ be the neighbors of r_i in G_i. The graph G_{i+1} is obtained from the contraction of G_i with respect to the subset $S_i = \{r_i\} \cup N_{G_i}(r_i)$. Thus, if v_{S_i} denotes the new vertex resulting from the contraction of S_i, then $(S_i^* \setminus S_i) \cup \{v_{S_i}\}$ is a connected vertex cover of G_{i+1}. Moreover, $|S_i^* \cap S_i| \geq 2$ since either $r_i \in S_i^*$ and at least one of these neighbors must belong to S_i^* (S_i^* is connected and $i < p$) or $N_{G_i}(r_i) \subseteq S_i^*$ since S_i^* is a vertex cover. Thus $opt(G_{i+1}) \leq |S_i^* \setminus S_i| + 1 = opt(G_i) - |S_i^* \cap S_i| + 1 \leq opt(G_i) - 1$. Summing up these inequalities for $i = 1$ to $p - 1$, and using that $opt(G) \geq opt(G_1)$, we obtain:

$$opt(G) \geq opt(G_p) + p - 1 \tag{6}$$

Moreover, thanks to Lemma 9, we know that $opt(G_p) = 2|N_p| - 1$. Together with equation (6), we get:

$$opt(G) \geq 2|N_p| - 1 + p - 1 \tag{7}$$

Finally, since each vertex chosen in step 3 has degree at least 3, we get $|N_{i+1}| \leq |N_i| - 2$. This immediately leads to $|N_1| \geq |N_p| + 2(p-1)$. Since $|S^*| \geq |N_1|$, we get:

$$|S^*| \geq |N_p| + 2(p-1) \tag{8}$$

Combination of equations (5), (7) and (8) with coefficients 4, 1 and 1 (respectively) gives:

$$5opt(G) \geq 3|S^*| + 3|N_p| - 1 + 3(p-1) \tag{9}$$

Then, equation (4) allows to conclude.

□

5.2 Planar graphs

Voire si ref [12, 11] donne un PTAS generique pour CVC

Given a planar embedding of a planar graph $G = (V, E)$, the level of a vertex is defined as follows (see for instance [4]): the vertices on the exterior face are at level 1. Given vertices at level i, let f be an interior face of the subgraph induced by vertices at level i. If G_f denotes the subgraph induced by vertices included in f, then the vertices on the exterior face of G_f are at level $i + 1$. The set of vertices at level i is called the layer L_i.

This is illustrated on Figure 4. The dashed ellipse represents an interior face on level $i - 1$. Depicted vertices are at level i. There are 3 interior faces (constituted respectively by the u_i's, by $\{v_1, v_2, t\}$ and $\{t, w_1, w_2\}$).

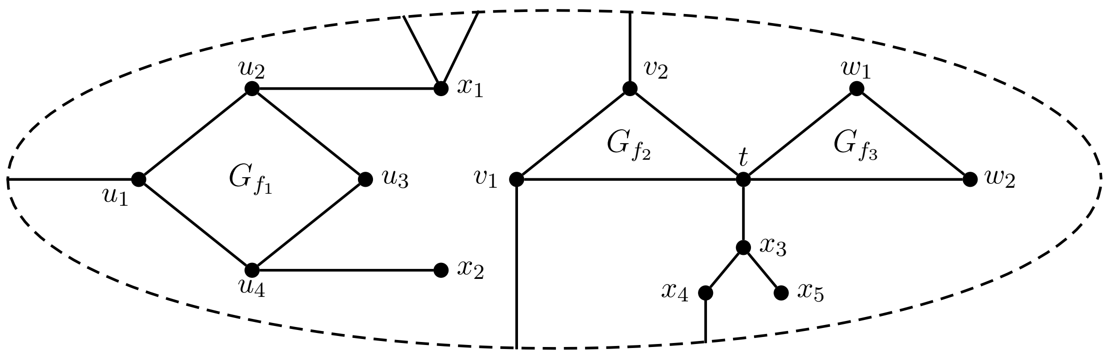

Figure 4: Level of a planar graph

Baker gave in [4] a polynomial time approximation scheme for several problems including vertex cover in planar graphs. The underlying idea is to consider k-outerplanar subgraphs of G constituted by k consecutive layers. The polynomiality of vertex cover in k-outerplanar graphs (for a fixed k) allows to achieve a $(k + 1)/k$ approximation ratio.

We adapt this technique in order to achieve an approximation scheme for MinCVC (MinCVC is **NP**-hard in planar graphs, see [16]). First of all, note that k-outerplanar graphs have treewidth bounded above by $3k - 1$, [6]. Since MinCVC is polynomially solvable for graphs with bounded treewidth, [25], MinCVC is polynomial for k-outerplanar graphs.

Theorem 11. MinCVC *admits an approximation scheme in planar graphs.*

Proof. Given an embedding of a planar (connected) graph G, we define, as previously, the layers $L_1, \cdots, L_q$ of G. For each layer L_i, we define F_i as the set of vertices of L_i that are in an interior face of L_i. For instance, in Figure 4, all vertices but the x_i's are in F_i.

Following the principle of the approximation scheme for vertex cover, we define an algorithm for any integer $k > 0$. Let $V_i = F_i \cup L_{i+1} \cup L_{i+2} \cup \ldots \cup L_{i+k}$, and G_i be the graph induced by V_i. Note that G_i is not necessarily connected since for example there can be several disjoint faces in F_i (there are two connected components in Figure 4).

Let S^* be an optimum connected vertex cover on G with value $opt(G)$, and $S^*_i = S^* \cap V_i$. Then of course S^*_i is a vertex cover of G_i. However, even restricted to a connected component of G_i, it is not necessarily connected. Indeed, S^* is connected but the path(s) connecting two vertices of S^* in a connected component of G_i may use vertices out of this connected component. To overcome this problem, notice that only vertices in F_i or in F_{i+k} connect V_i to $V \setminus V_i$. Hence, $S_i^* \cup F_i \cup F_{i+k}$ can be partitioned into a set of connected vertex covers on each of the connected components of G_i (since F_i and F_{i+k} are made of cycles). Now, take an optimum connected vertex cover on each of these connected components, and define S_i as the union of these optimum solutions. Then, we have :

$$|S_i^* \cup F_i \cup F_{i+k}| \geq |S_i| \tag{10}$$

Now, let $p \in \{1, \ldots, k\}$. Let $V_0 = L_1 \cup L_2 \cup \ldots \cup L_p$, G_0 be the subgraph of G induced by V_0, $S_0^* = S^* \cap V_0$, and S_0 be an optimum vertex cover on G_0. With similar arguments as previously, we have:

$$|S_0^* \cup F_p| \geq |S_0| \tag{11}$$

We build a solution S^p on the whole graph G as follows. S^p is the union of S_0 and of all S_i's for $i = p \ mod \ k$. Of course, S^p is a vertex cover of G, since any edge of G appears in at least one G_i (or G_0). Moreover, it is connected since:

- S_0 is connected, and each S_i is made of connected vertex covers on the connected components of G_i;
- each of these connected vertex covers in S_i is connected to S_{i-k} (or to S_0 if $i = p$): this is due to the fact that F_i belongs to V_i and to V_{i-k} (or V_0). Hence, a level i interior face f is common to S_{i-k} (or S_0) and to the connected vertex cover of S_i we are dealing with. Both partial solutions cover all the edges of this face f. Since f is a cycle, the two solutions are necessarily connected. In other words, each connected component of S_i is connected to S_{i-k} (or S_0) and, by recurrence, to S_0. Consequently, the whole solution S^p is connected.

Summing up equation (10) for each $i = p \bmod k$ and equation (11), we get:

$$|S_0^* \cup F_p| + \sum_{i=p \bmod k} |S_i^* \cup F_i \cup F_{i+k}| \geq |S_0| + \sum_{i=p \bmod k} |S_i| \tag{12}$$

By definition of S^p, we have $|S^p| \leq |S_0| + \sum_{i=p \bmod k} |S_i|$. On the other hand, since only vertices in F_i ($i = p \bmod k$) appear in two different V_i's ($i = 0$ or $i = p \ mod \ k$), we get that $|S_0^* \cup F_p| + \sum_{i=p \bmod k} |S_i^* \cup F_i \cup F_{i+k}| \leq |S^*| + 2\sum_{i=p \bmod k} |F_i|$. This leads to:

$$opt(G) + 2 \sum_{i=p \bmod k} |F_i| \geq |S^p| \tag{13}$$

If we consider the best solution S with value $apx(G)$ among the S^p's ($p \in \{1, \ldots, k\}$), we get :

$$opt(G) + \frac{2}{k} \sum_{i=1}^{q} |F_i| \geq apx(G) \tag{14}$$

To conclude, we observe that the following property holds:

Property 12. *S^* takes at least one fourth of the vertices of each F_i.*

To see this property of $S^* \cap F_i$, consider F_i and the set E_i of edges of G that belong to one and only one interior face of F_i (for example, in Figure 4, if there were edges $\{u_2, u_4\}$ and $\{u_3, v_1\}$, they would not be in E_i). Let n_i be the number of vertices in F_i, and m_i the number of edges in E_i. This graph is a collection of edge-disjoint (but not vertex-disjoint, as one can see in Figure 4) interior faces (cycles). Of course, $S^* \cap F_i$ is a vertex cover of this graph. Since this graph is a collection of interior faces (cycles), on each of these faces f $S^* \cap F_i$ cannot reject more than $|f|/2$ vertices. In all,

$$|S^* \cap F_i| \geq n_i - \sum_{f \in F_i} \frac{|f|}{2}$$

But since faces are edge-disjoint, $\sum_{f \in F_i} |f| = m_i$. On the other hand, if N_f denotes the number of interior faces in F_i, since each face contains at least 3 vertices, $m_i = \sum_{f \in F_i} |f| \geq 3N_f$. Since the graph is planar, using Euler formula we get $1 + m_i = n_i + N_f \leq n_i + m_i/3$. Hence, $m_i \leq 3n_i/2$. Finally, $|S^* \cap F_i| \geq n_i - m_i/2 \geq n_i/4$. Based on this property, we get:

$$opt(G)\left(1 + \frac{8}{k}\right) \geq apx(G) \tag{15}$$

Taking k sufficiently large leads to a $1 + \varepsilon$ approximation. The polynomiality of this algorithm follows from the fact that each subgraph we deal with is (at most) $k+1$-outerplanar, hence for a fixed k we can find an optimum solution in polynomial time. □

6 Connected vertex cover in hypergraphs

Here, we extend the notions of vertex cover and connected vertex cover to hypergraphs. Whereas the generalization of the vertex cover problem to hypergraphs is quite natural, it turns out that the generalization of the connected vertex cover problem is a task much harder due to the notion of connected hypergraphs. Actually, we will give two generalizations: the *weak connected vertex cover* problem and the *strong connected vertex cover* problem.

Before establishing a definition of these two problems, we recall some definitions on hypergraphs. A simple hypergraph $\mathcal{H}$ is a pair $(V, \mathcal{E})$ where $V = \{v_1, \cdots, v_n\}$ is the vertex set and $\mathcal{E} = \{e_1, \cdots, e_m\} \subseteq 2^V$ is the hyperedge set. Given a hypergraph $\mathcal{H} = (V, \mathcal{E})$, $d_{\mathcal{H}}(v)$, $N_{\mathcal{H}}(v)$ and $s_{\mathcal{H}}(e)$ denote respectively the degree, the neighborhood of a vertex $v \in V$ and the size of an hyperedge $e \in \mathcal{E}$, that is $d_{\mathcal{H}}(v) = |\{e \in \mathcal{E} : v \in e\}|$, $N_{\mathcal{H}}(v) = \{u \in V \setminus \{v\} : \exists e \in \mathcal{E}$ containing vertices $u, v\}$ and $s_{\mathcal{H}}(e) = |\{v : v \in e\}|$. $\Delta(\mathcal{H})$ and $s(\mathcal{H})$ denote respectively the *maximum degree* of a vertex and the *maximum size* of a hyperedge in $\mathcal{H}$. The following definition are introduced in [7]: $\mathcal{H}' = (V', \mathcal{E}')$ is a *partial hypergraph* of $\mathcal{H} = (V, \mathcal{E})$ if $\mathcal{E}' \subseteq \mathcal{E}$ and V' is the union of the hyperedges in $\mathcal{E}'$. The *restriction* of a hypergraph $\mathcal{H} = (V, \mathcal{E})$ to $V' \subseteq V$ is the partial hypergraph $\mathcal{H}' = (V', \mathcal{E}')$ (that is satisfying $\mathcal{E}' = \{e \in \mathcal{E} : e \cap V' = e\}$). The subhypergraph of $\mathcal{H} = (V, \mathcal{E})$ *induced by* V' is the hypergraph $\mathcal{H}' = (V', \mathcal{E}')$ where $\mathcal{E}' = \{e \cap V' : e \in \mathcal{E}\}$. A hypergraph is *simple* if no hyperedge is a subset of any other hyperedge. A hypergraph is r-*uniform* if each hyperedge has a size r and r-*regular* if each vertex has a degree r. A *path* of length k from v_1 to v_k in a hypergraph $\mathcal{H} = (V, \mathcal{E})$ is a sequence $(v_1, e_1, v_2, \cdots, e_k, v_{k+1})$ with $k \geq 1$ such that $e_1, \cdots, e_k$ and $v_1, \cdots, v_{k+1}$ are sets of distinct hyperedges and vertices respectively, and $\forall i = 2, \cdots, k-1$, $v_i \in e_{i-1} \cap e_i$ and $v_1 \in e_1, v_{k+1} \in e_k$. A hypergraph $\mathcal{H}$ is *connected* if between every pair (v_i, v_j) of disjoint vertices, there is path from v_i to v_j.

The dual hypergraph of an hypergraph $\mathcal{H} = (V, \mathcal{E})$ is the hypergraph $\mathcal{H}^* = (V_\mathcal{E}, E^*)$ such that the vertices of $V_\mathcal{E} = \{v_e : e \in \mathcal{E}\}$ correspond to the hyperedges of $\mathcal{E}$ and the hyperedge set is $E^* = \{\mathcal{E}_v, v \in V\}$ where $\mathcal{E}_v = \{v_e : v \in e\}$.

The generalization of MinVC and MinCVC in graphs to hypergraphs can be defined as follows. Given a hypergraph $\mathcal{H} = (V, \mathcal{E})$, a vertex cover of $\mathcal{H}$ is a subset of vertices $S \subseteq V$ such that for any hyperedge $e \in \mathcal{E}$, we have $S \cap e \neq \emptyset$; the vertex cover problem in hypergraphs is the problem of determining a vertex cover S^* of $\mathcal{H}$ minimizing $|S^*|$. It is well known that this problem is equivalent to the *set cover problem* (in short MinSC) by considering the dual hypergraphs, (see for instance [9]). Thus, the vertex cover problem in hypergraphs is not approximable within performance ratio $(1-\varepsilon)\ln m$ for all $\varepsilon > 0$ unless **NP⊂DTIME**$(m^{loglogm})$. **Moreover, it is not** $\ln(\Delta) - c\ln(\ln(\Delta))$**-approximable (for some constant** c**), for any constant** Δ**, in hypergraphs of degree** Δ**, [28]**. Recently, new inapproximation results have been given. In [13], the authors prove that the vertex cover problem in k-uniform hypergraphs is not $(k-1-\varepsilon)$-approximable unless **P**=**NP** for any $k \geq 3$ and $\varepsilon > 0$. At the same time, based on **the so-called *unique games conjecture***, it is shown that $(k-\varepsilon)$ is a lower bound of the approximation of vertex cover in k-uniform hypergraphs for any $k \geq 2$ and $\varepsilon > 0$, [21].

We consider two versions of the connected vertex cover problem in hypergraphs, namely a weak and strong one. Given a connected hypergraph $\mathcal{H} = (V, \mathcal{E})$, the weak (resp., strong) connected vertex cover problem, denoted by MinWCVC (resp., MinSCVC) consists in finding a minimum size vertex cover S^* of $\mathcal{H}$ such that the subhypergraph induced by S^* (resp., the restriction of $\mathcal{H}$ to S^*) is connected. Obviously, when we restrict these problems to graphs, we again find the connected vertex cover problem.

6.1 The weak connected vertex cover problem

The weak connected vertex cover problem is as hard as the vertex cover problem in hypergraphs since starting from any hypergraph $\mathcal{H} = (V, \mathcal{E})$ and by adding a new hyperedge e containing the entire vertex set (ie., $e = V$), any connected vertex cover of the new hypergraph is a vertex set of the initial hypergraph. Thus, we deduce that on the one hand MinWCVC is **NP**-hard in connected hypergraphs of maximum degree 4 and is **not** $c\ln m$ **approximable, for some constant** c, unless **P**=**NP**, [9]. Moreover, using another simple reduction, the negative approximation results established in [13, 21] also hold for MinWCVC. Actually, starting with a k-uniform hypergraph $\mathcal{H} = (V, \mathcal{E})$ where $V = \{v_1, \cdots, v_n\}$ and $\mathcal{E} = \{e_1, \cdots, e_m\}$, we first add a new vertex v_0 connected to each vertex v_i by edges e'_i for any $i = 1, \cdots, n$. Then, we replace each edge e'_i by a hyperedge by introducing $k-2$ new vertices. Obviously, this new hypergraph $\mathcal{H}'$ is connected and k-uniform, and it is easy to see that S is a vertex cover of $\mathcal{H}$ iff $S \cup \{v_0\}$ is a weak connected vertex cover of $\mathcal{H}'$. In conclusion, for k-uniform hypergraphs, MinWCVC is not $(k-\varepsilon)$ (or $(k-1-\varepsilon)$) -approximable under the same hypothesis as those given in [13, 21]. **We now present a simple approximation algorithm which shows** that the previous bound is sharp.

For a connected hypergraph $\mathcal{H} = (V, \mathcal{E})$ and a hyperedge $e \in \mathcal{E}$, we set $N_\mathcal{H}(e) = \cup_{v \in e} N_\mathcal{H}(v)$; remark that $e \subseteq N_\mathcal{H}(e)$ **(assuming wlog. that there is no edge of size 1)**. The following greedy algorithm is a generalization of the classical 2-approximation algorithm for the vertex cover problem.

`Greedy2HCVC` input: A connected hypergraph $\mathcal{H} = (V, \mathcal{E})$.

1 Set $S = \emptyset$ and $Label = \{v\}$ where v is a vertex of $\mathcal{H}$;

2 While there exists a hyperedge $e \in \mathcal{E}$ with $e \cap Label \neq \emptyset$ do

2.1 $S := S \cup e$ and $Label := Label \cup N_{\mathcal{H}}(e)$;

2.2 Delete from $\mathcal{H}$ all the hyperedges adjacent to e and all the vertices in e. Let $\mathcal{H}$ be the resulting hypergraph;

3 Output S;

Let us prove that S is a weak vertex cover of the initial hypergraph $\mathcal{H}$. Otherwise, we have $Label \neq V$ and let $\mathcal{H}'' = (V \setminus Label, \mathcal{E}'')$ be the subhypergraph of $\mathcal{H}$ induced by $V \setminus Label$. By assumption, $\mathcal{H}''$ contains some hyperedges of $\mathcal{E}$; actually, it is easy to prove that each vertex $v \notin Label$ is not isolated in $\mathcal{H}''$ and each hyperedge of $\mathcal{H}''$ is a hyperedge of $\mathcal{H}$ with the same size (thus, $\mathcal{H}''$ is also the restriction of $\mathcal{H}$ to $V \setminus Label$). Since $\mathcal{H}$ is connected, there is a hyperedge $e \in \mathcal{E} \setminus \mathcal{E}''$ such that $v \in e \cap Label$ and $w \in e \cap (V \setminus Label)$. This hyperedge e has been deleted by `Greedy2HCVC` because either e has been added to S or e is adjacent to a hyperedge $e' \notin \mathcal{E}''$ with $e' \subset S$. In any case, w would have been added to $Label$, contradiction. Finally, we can easily prove that at each iteration of `Greedy2HCVC`, the current set S induces a connected subhypergraph and then, the solution output by this algorithm is a weak connected vertex cover.

The following result is an obvious generalization of the analysis of the classical matching algorithm for MinVC.

Theorem 13. *`Greedy2HCVC` is a $s(\mathcal{H})$-approximation of* MinWCVC.

We now establish some connection between the weak connected vertex cover problem in hypergraphs and the minimum labeled spanning tree in multigraphs. In the minimum labeled spanning tree problem (MinLST in short) in multigraphs, we are given a connected, undirected multigraph $G = (V, E)$ on n vertices. Each edge e in E is colored (or labeled) with the color $\mathcal{L}(e) \in \{c_1, c_2, \ldots, c_q\}$ and for $E' \subseteq E$, we denote $\mathcal{L}(E') = \cup_{e \in E'} \mathcal{L}(e)$ the set of colors used by E'. The goal of the minimum labeled spanning tree problem is to find, given $I = (G, \mathcal{L})$ an instance of MinLST, a spanning tree T in G that uses the minimum number of colors, that is minimizing $|\mathcal{L}(T)|$. Equivalently, if $\mathcal{L}^{-1}(\mathcal{C}) \subseteq E$ denotes the set of edges with color $c_i \in \mathcal{C}$ for any set $\mathcal{C} \subseteq \{c_1, c_2, \ldots, c_q\}$, then another formulation of MinLST asks to find a smallest cardinality subset $\mathcal{C} \subseteq \{c_1, c_2, \ldots, c_q\}$ of the colors, such that the subgraph induced by the edge sets $\mathcal{L}^{-1}(\mathcal{C})$ is connected and touches all vertices in V. The minimum labeled spanning tree problem has been studied in the context of simple graphs for instance in [8], but it is easy to see that all the obtained results also hold in multigraphs, [20]. The *color frequency* of $I = (G, \mathcal{L})$ denoted by r, is the maximum number of times that a color appears, that is $r = \max\{|\mathcal{L}^{-1}(c_i)| : i = 1, \cdots, q\}$.

Theorem 14. A $\rho(r)$-approximation of MinLST can be polynomially converted into a $\rho(r)$-approximation of MinWCVC in hypergraph of maximum degree $r+1$.

Proof. Let $\mathcal{H} = (V, \mathcal{E})$ be a connected hypergraph with maximum degree Δ, instance of MinWCVC. We build the multigraph $G = (V_{\mathcal{E}}, E)$ where the vertex set is given by $V_{\mathcal{E}} = \{v_e : e \in \mathcal{E}\}$; the edge set is $E = \cup_{v \in V} T_v$ where T_v is an arbitrary spanning tree on the subset of vertices $\{v_e \in V_{\mathcal{E}} : v \in e\}$. Finally, the color set is $\{c_v : v \in V\}$ and if $e \in T_v$, then the edge e is colored with color c_v, that is $\mathcal{L}(e) = c_v$. It is easy to observe that color c_v appears exactly $d_{\mathcal{H}}(v) - 1$ times. In conclusion $I = (G, \mathcal{L})$ is an instance of MinLST with color frequency $r = \Delta - 1$.

We claim that $S \subseteq V$ is a weak connected vertex cover of $\mathcal{H}$ iff the subgraph $G' = (V_{\mathcal{E}}, E')$ where $E' = \cup_{v \in S} T_v$ is connected.

Assume that $G' = (V_{\mathcal{E}}, \cup_{v \in S} T_v)$ is a connected subgraph of $G = (V_{\mathcal{E}}, \cup_{v \in V} T_v)$. Let $e \in \mathcal{E}$; since G' spans all the vertices of $V_{\mathcal{E}}$, there exists $v \in S$ such that $v_e \in T_v$ (formally, v_e is adjacent to e' with $e' \in T_v$). Thus, v covers the hyperedge e in H and more generally S is a vertex cover of $\mathcal{H}$. Let us prove that the subhypergraph induced by S is a connected hypergraph. Let $s, t \in S$; since G' is connected, there is a shortest path μ in G' linking a vertex of T_s to a vertex of T_t. Assume that this path μ uses edges colored with colors $c_{v_1}, \cdots, c_{v_p}$. By construction, $\{v_1, \cdots, v_p\} \subseteq S$ and since μ is a shortest path, we can assume, wlog., that the colors met in μ are $c_{v_1}, \cdots, c_{v_p}$ in this order. Let v_{e_j} for $j = 1, \cdots, p-1$ be the vertex adjacent to colors c_{v_j} and $c_{v_{j+1}}$ in μ. By construction, $\{v_j, v_{j+1}\} \subseteq e_j$ in hypergraph $\mathcal{H}$. Moreover, for the same reasons, there is also two hyperedges e_0 and e_p such that $\{s, v_1\} \subseteq e_0$ and $\{v_p, t\} \subseteq e_p$. In conclusion, $(s, e_0, v_1, e_1, v_2, \cdots, e_p, t)$ is a path from s to t in $\mathcal{H}'$ and $\mathcal{H}'$ is connected.

Conversely, let $S \subseteq V$ be a weak vertex cover of $\mathcal{H}$. Obviously, $\cup_{v \in S} T_v$ spans all the vertices of $V_{\mathcal{E}}$ since S is a vertex cover of $\mathcal{H}$. Besides, it turns out that any path $(v_1, e_1, v_2, \cdots, e_{p-1}, v_p)$ from v_1 to v_p in the restriction $\mathcal{H}'$ of $\mathcal{H}$ to S can be transformed into a path going through edges from $\cup_{i=1}^{p} T_{v_i}$. In conclusion, $G' = (V_{\mathcal{E}}, \cup_{v \in S} T_v)$ is a connected subgraph.

Now, since the number of colors used by $G' = (V_{\mathcal{E}}, E')$ where $E' = \cup_{v \in S} T_v$ is exactly $|S|$, the result follows. In particular, any ρ-approximation for MinLST can be polynomially converted into a ρ-approximation for MinWCVC. If ρ depends on parameter r, the final performance ratio **is valid in hypergraphs of degree Δ.** □

In [8], it is proved that the restriction of MinLST to the instances $I = (G, \mathcal{L})$ where each color appears at most twice (ie, $r \leq 2$) is polynomial, even if G is a multigraph. Thus, using Theorem 14, we strengthen the result of [29], establishing that the connected vertex cover problem is polynomial in simple graphs with maximum degree 3.

Corollary 15. MinWCVC *is polynomial in hypergraphs with maximum degree 3.*

On the other hand, using the $(H(r) - 1/2)$-approximation for MinLST where $H(r) = \sum_{i=1}^{r} \frac{1}{i}$ is the r-th harmonic number given in [20], we deduce:

Corollary 16. MinWCVC *is $(H(\Delta - 1) - 1/2)$-approximable in hypergraphs of maximum degree Δ.*

Note that this result is very close to the lower bound of $\ln(\Delta) - c\ln(\ln(\Delta))$ already mentioned ([28]).

6.2 The strong connected vertex cover problem

It turns out that the complexity of the strong connected vertex cover problem is much harder than the one of the weak connected vertex cover problem. Actually, in contrast

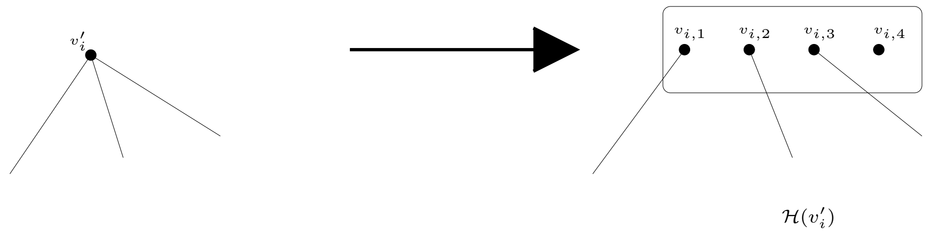

Figure 5: The gadget $\mathcal{H}(v'_i)$.

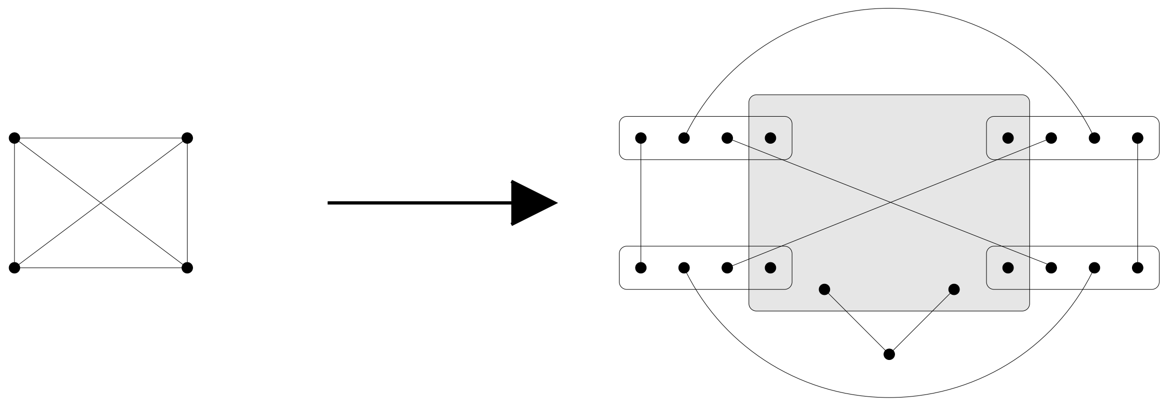

Figure 6: An example of the construction from a K_4.

to Corollary 15, we now prove that MinSCVC has no approximation scheme in 2-regular hypergraphs.

Theorem 17. MinSCVC *is* ***APX****-complete in connected 2-regular hypergraphs.*

Proof. We give an approximation preserving L-reduction from the vertex cover problem in cubic graphs. This restriction has been proved **APX**-complete in [1].

Let $G' = (V', E')$ be a cubic graph with $V' = \{v'_1, \cdots, v'_n\}$ and $E' = \{e'_1, \cdots, e'_m\}$, instance of MinVC. We build the connected 2 regular hypergraph $\mathcal{H} = (V, \mathcal{E})$ containing vertices $v_{i,j}$ for $i = 1, \cdots, n$, $j = 1, \cdots, 4$ and u_j for $j = 1, 2, 3$. Moreover,

- Each vertex v'_i of G' with $i = 1, \cdots, n$, is split into $d_{G'}(v'_i) + 1$ (=4 since G' is cubic) vertices $v_{i,1}, \cdots, v_{i,4}$ such that the edges of G' become a matching in the hypergraph $\mathcal{H}$ saturating vertices $v_{i,j}$ for $i = 1, \cdots, n$, $j = 1, \cdots, 3$. Moreover, we add the hyperedge $e_i = \{v_{i,1}, \cdots, v_{i,4}\}$. This gadget $\mathcal{H}(v'_i)$ is described in Figure 5.
- We add the path μ of length 2 $\mu = \{\{u_1, u_2\}, \{u_2, u_3\}\}$ and the hyperedge $e_0 = \{v_{i,4} : i = 1, \cdots, n\} \cup \{u_1, u_3\}$.

Clearly, $\mathcal{H} = (V, \mathcal{E})$ is a connected hypergraph where each vertex has a degree 2. Figure 6 gives a simple illustration of this construction when G' is a K_4.

If S^* is an optimal vertex cover of G' with value $opt(G')$, then by taking $\{e_i : v'_i \in S^*\} \cup e_0$, we obtain a strong connected vertex cover of $\mathcal{H}$. Thus,

$$opt(\mathcal{H}) \leq 3opt(G') + n + 2 \tag{16}$$

Conversely, let V_0 be a strong connected vertex cover of $\mathcal{H}$ with value $apx(\mathcal{H})$. By construction, V_0 contains e_0 (i.e., the vertices of this hyperedges) since it is the only way to connect the edges of the path μ to the rest of the solution. Moreover, for each edge $e'_k = \{v_{i,i_1}, v_{j,j_1}\}$ where $i_1, j_1 \in \{1,2,3\}$ of $\mathcal{H}$ we have $e_i \subseteq V_0$ or $e_j \subseteq V_0$ since on the one hand, V_0 is a vertex cover of $\mathcal{H}$ and on the other hand, as previously the only way to connect the hyperedge e_0 to v_{i,i_1} or v_{j,j_1} consists of taking the whole hyperedge e_i or e_j. Finally, wlog. we may assume that $e_i \cap V_0 = e_i$ or $e_i \cap V_0 = \{v_{i,4}\}$. Thus, $\{v'_i : e_i \in V_0\}$ is a vertex cover of G', with value:

$$apx(G') \leq \frac{apx(\mathcal{H}) - n - 2}{3} \tag{17}$$

Using inequalities (16) and (17), we obtain $3opt(G') = opt(\mathcal{H}) - n - 2$. Thus, on the one hand we have $apx(G') - opt(G') \leq apx(\mathcal{H}) - opt(\mathcal{H})$ and on the other hand, $opt(\mathcal{H}) = 3opt(G') + n + 2 \leq (5 + \varepsilon)opt(G')$ since G' is an instance of MinVC and cubic.

□

References

[1] P. Alimonti and V. Kann. Some APX-completeness results for cubic graphs. *Theor. Comput. Sci.*, 237(1-2):123–134, 2000.

[2] E. M. Arkin, M. M. Halldórsson, and R. Hassin. Approximating the tree and tour covers of a graph. *Inf. Process. Lett.*, 47(6):275–282, 1993.

[3] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protasi. *Complexity and Approximation (Combinatorial Optimization Problems and Their Approximability Properties).* Springer, Berlin, 1999.

[4] B. S. Baker. Approximation algorithms for NP-complete problems on planar graphs. *J. ACM*, 41(1):153–180, 1994.

[5] R. Bar-Yehuda and S. Even. On approximating a vertex cover for planar graphs. In *STOC*, pages 303–309, 1982.

[6] H. L. Bodlaender. A partial k-arboretum of graphs with bounded treewidth. *Theor. Comput. Sci.*, 209(1-2):1–45, 1998.

[7] A. Brandstadt, V. B. Le, and J. Spinrad. *Graph classes : a survey.* Philadelphia : Society for Industrial and Applied Mathematic, 1999.

[8] T. Brüggemann, J. Monnot, and G. J. Woeginger. Local search for the minimum label spanning tree problem with bounded color classes. *Op. Res. Lett.*, 31(3):195–201, 2003.

[9] M. Chlebík and J. Chlebíková. Approximation hardness of dominating set problems. In *ESA*, pages 192–203, 2004.

[10] M. Chlebík and J. Chlebíková. Complexity of approximating bounded variants of optimization problems. *Theor. Comput. Sci.*, 354(3):320–338, 2006.

[11] A. Dawar, M. Grohe, S. Kreutzer, and N. Schweikardt. Approximation schemes for first-order definable optimisation problems. In *LICS*, pages 411–420, 2006.

[12] E. D. Demaine and M. T. Hajiaghayi. Bidimensionality: new connections between fpt algorithms and ptass. In *SODA*, pages 590–601, 2005.

[13] I. Dinur, V. Guruswami, S. Khot, and O. Regev. A new multilayered PCP and the hardness of hypergraph vertex cover. In *STOC*, pages 595–601, 2003.

[14] H. Fernau and D. Manlove. Vertex and edge covers with clustering properties: Complexity and algorithms. In *Algorithms and Complexity in Durham*, pages 69–84, 2006.

[15] T. Fujito. How to trim an mst: A 2-approximation algorithm for minimum cost tree cover. In *ICALP (1)*, pages 431–442, 2006.

[16] M. R. Garey and D. S. Johnson. The rectilinear steiner tree problem is NP complete. *SIAM Journal of Applied Mathematics*, 32:826–834, 1977.

[17] M. R. Garey and D. S. Johnson. *Computer and Intractability: A Guide to the Theory of NP-Completeness.* W. H. Freeman, 1979.

[18] F. Gavril. Algorithms for minimum coloring, maximum clique, minimum covering by cliques, and maximum independent set of a chordal graph. *SIAM J. Comput.*, 1(2):180–187, 1972.

[19] J. Guo, R. Niedermeier, and S. Wernicke. Parameterized complexity of generalized vertex cover problems. In *WADS*, pages 36–48, 2005.

[20] R. Hassin, J. Monnot, and D. Segev. Approximation algorithms and hardness results for labeled connectivity problems. In *Journal of Combinatorial Optimization (article in press)*, 2007.

[21] S. Khot and O. Regev. Vertex cover might be hard to approximate to within $2 - \varepsilon$. In *IEEE Conference on Computational Complexity*, pages 379–386, 2003.

[22] J. Könemann, G. Konjevod, O. Parekh, and A. Sinha. Improved approximations for tour and tree covers. *Algorithmica*, 38(3):441–449, 2003.

[23] D. Mölle, S. Richter, and P. Rossmanith. Enumerate and expand: Improved algorithms for connected vertex cover and tree cover. In *CSR*, pages 270–280, 2006.

[24] D. Mölle, S. Richter, and P. Rossmanith. Enumerate and expand: New runtime bounds for vertex cover variants. In *COCOON*, pages 265–273, 2006.

[25] H. Moser. Exact algorithms for generalizations of vertex cover. Master's thesis, Institut für Informatik, Friedrich-Schiller-Universität Jena, 2005.

[26] C. D. Savage. Depth-first search and the vertex cover problem. *Inf. Process. Lett.*, 14(5):233–237, 1982.

[27] A. Schrijver. *Combinatorial Optimization - Polyhedra and Efficiency.* Springer-Verlag, Berlin, 2003.

[28] L. Trevisan. Non-approximability results for optimization problems on bounded degree instances. In *STOC*, pages 453–461, 2001.

[29] S. Ueno, Y. Kajitani, and S. Gotoh. On the nonseparating independent set problem and feedback set problem for graphs with no vertex degree exceeding three. *Discrete Mathematics*, 72:355–360, 1988.

[30] T. Wanatabe, S. Kajita, and K. Onaga. Vertex covers and connected vertex covers in 3-connected graphs. In *IEEE International Sympoisum on Circuits and Systems*, pages 1017–1020, 1991.