HAL
archives-ouvertes.fr

Algorithmic Aspects of the Intersection and Overlap Numbers of a Graph

Danny Hermelin, Romeo Rizzi, Stéphane Vialette

Algorithmic Aspects of the Intersection and Overlap Numbers of a Graph

Danny Hermelin[1], Romeo Rizzi[2], and Stéphane Vialette[3]

[1] Max-Planck-Institut für Informatik, Saarbrücken, Germany
`hermelin@mpi-inf.mpg.de`
[2] Dipartimento di Matematica ed Informatica, Universit degli Studi di Udine, Italy
`rrizzi@dimi.uniud.it`
[3] LIGM CNRS UMR 8049, Université Paris-Est, France
`vialette@univ-mlv.fr`

Abstract. The intersection number of a graph G is the minimum size of a ground set S such that G is an intersection graph of some family of subsets $\mathcal{F} \subseteq 2^S$. The overlap number of G is defined similarly, except that G is required to be an overlap graph of $\mathcal{F}$. In this paper we show two algorithmic aspects concerning both these graph invariants. On the one hand, we show that the corresponding optimization problems associated with these numbers are both **APX**-hard, where for the intersection number our results hold even for biconnected graphs of maximum degree 7, strengthening the previously known hardness result. On the other hand, we show that the recognition problem for any specific intersection graph class (*e.g.* interval, unit disc, string, ...) is easy when restricted to graphs of fixed bounded intersection or overlap number.

1 Introduction

An *intersection graph* is a graph that represents the pattern of intersections of a family of sets. Any undirected graph G may be represented as an intersection graph: For each vertex of G, form a set consisting of the edges incident to this vertex; then two such sets have a nonempty intersection if and only if the corresponding vertices share an edge []. Erdös, Goodman, and Pósa [5] provided a construction that is more efficient in which the total number of set elements is at most $n^2/4$, where n is the number of vertices in the graph. Many important graph families can be described as intersection graphs of more restricted types of set families, in particular sets corresponding to geometric objects. Examples of such graph classes are *interval graphs* (intersection graphs of intervals on the real line), *circle graphs* (intersection graphs of chords in a circle), *unit disc graphs* (intersection graphs of unit discs in the plane), and *string graphs* (intersection graphs of simple curves in the plane).

The *intersection number* of G, denoted $i(G)$, is defined to be the minimum cardinality of a (ground) set S such that G is an intersection graph of a family of subsets $\mathcal{F} \subseteq 2^S$ of S. In [5], it was shown that $i(G)$ also equals the minimum number of complete subgraphs needed to cover the edges of G. This latter number is known as the *edge-clique cover* number of G, and is denoted $\theta(G)$. (The best general reference is [14].) Computing $\theta(G)$ and $i(G)$ is **NP**-hard [11,16], even when restricted to planar graphs [3] or graphs with maximum degree 6 [9]. It is polynomial-time solvable for chordal graphs [13], graphs with maximum degree 5 [9], line graphs [16], and circular-arc graphs [10]. By way of contrast, it is not approximable within ratio n^ε for some $\varepsilon > 0$ unless $\mathbf{P} = \mathbf{NP}$ [12], and so far nothing better than a polynomial ratio of $O(n^2 \frac{(\log \log n)^2}{(\log n)^3})$ is known [2]. As for its parameterized complexity, computing $\theta(G)$

is fixed-parameter tractable under the standard parameterization [6]. Guillaume and Latapy [7] argue that computing $\theta(G)$ is very widely applicable to discover underlying structure in complex real-world networks, while [15] give a bioinformatic application for this problem.

The overlap model for graph representations arose much later and is not as well studied [4]. The *overlap graph* of a family of sets $\mathcal{F} = \{S_1, S_2, \ldots, S_n\}$, denoted $O(\mathcal{F})$, is the graph having $\mathcal{F}$ as vertex set with S_i adjacent to S_j if and only if S_i and S_j intersect and neither set is contained in the other, *i.e.*, $S_i \cap S_j \neq \emptyset$, $S_i \setminus S_j \neq \emptyset$, and $S_j \setminus S_i \neq \emptyset$. Notice that some graph classes can play it both ways: A graph is an intersection graph of chords in a circle (*i.e.* circle graph) iff it is has an overlap representation using intervals on a line. The *overlap number* of a graph G, denoted $\varphi(G)$, is the minimum size of the ground set in any overlap representation of G. The following upper bounds for the overlap number of a n-vertex graph are known [18,19]: $n + 1$ for trees, $2n$ for chordal graphs, $\frac{10}{3}n - 6$ for planar graphs, and $\lfloor n^2/4 \rfloor + n$ for general graphs. It was also observed that the overlap number of K_n is the minimum ℓ such that a ℓ-set contains n pairwise incomparable sets, that the overlap number of C_n is $n - 1$, and that the overlap number of any caterpillar with at least 2 vertices if the number of vertices in the longest path. In [4], it is shown among other results that an optimal overlap representation of a tree can be produced in linear time, and its size is the number of vertices in the largest subtree in which the neighbor of any leaf has degree 2.

The results in this paper are of twofold. In the first part of the paper, we consider the INTERSECTION NUMBER and OVERLAP NUMBER problems, the optimization problems that ask to respectively determine the intersection and overlap number of a given input graph. We show that both problems are **APX**-hard. While for INTERSECTION NUMBER this was already known for general graphs [12], our result proves this is the case also for graphs of maximum degree 7. Moreover, this result is used to show the **APX**-hardness of OVERLAP NUMBER. In the second part of the paper, we show that for any intersection graph class $\mathcal{G}$, *i.e.* any graph class defined by specifying the allowed intersection model, the recognition problem associated with $\mathcal{G}$ is linear-time solvable when restricted to graphs with bounded fixed intersection or overlap number.

2 Notations

Let G be a graph. We write $\mathbf{V}(G)$ for the set of vertices and $\mathbf{E}(G)$ for the set of edges of G. An *adjacent vertex* of a vertex u in a graph is a vertex that is connected to u by an edge. The *neighborhood* of a vertex u, denoted $N_G(u)$ or (when the graph is unambiguous) $N(u)$, is the set of adjacent vertices to u. We let $N[v]$ denote the set $\{v\} \cup N(v)$. The *degree* of a vertex $u \in \mathbf{V}(G)$, denoted $d(u)$, is the number of vertices adjacent to u. The *maximum degree* of G, denoted by $\Delta(G)$, is the maximum degree of its vertices. A *biconnected graph* is a connected graph that is not broken into disconnected pieces by deleting any single vertex (and its incident edges). An *edge-clique cover* of G is any family $\mathcal{E} = \{Q_1, Q_2, \ldots, Q_k\}$ of complete subgraphs of G such that every edge of G is in at least one of $Q_1, Q_2, \ldots, Q_k$. The minimum cardinality of an edge-clique cover of G is denoted $\theta(G)$, and we write EDGE-CLIQUE COVER for the combinatorial problem of computing $\theta(G)$.

The *Cartesian product* $G \times H$ of graphs G and H is the graph such that the vertex set of $G \times H$ is the Cartesian product $\mathbf{V}(G) \times \mathbf{V}(H)$, and any two vertices (u, u') and (v, v') are adjacent in $G \times H$ if and only if either $u = v$ and u' is adjacent with v' in H, or $u' = v'$ and u is adjacent with v in G. A *column* of $G \times H$ is the set of vertices $\{(u, u') : u \in \mathbf{V}(G)\}$ for some vertex $u' \in \mathbf{V}(H)$, and a *row* of $G \times H$ is the set of vertices $\{(u, u') : u' \in \mathbf{V}(H)\}$ for some vertex $u \in \mathbf{V}(G)$. Observe that each row induces a copy of H, and each column induces a copy of G (see Figure 1). This terminology is consistent with a representation of $G \times H$ by the points of the $|\mathbf{V}(G)| \times |\mathbf{V}(H)|$ grid. (See Figure 1 for an illustration.)
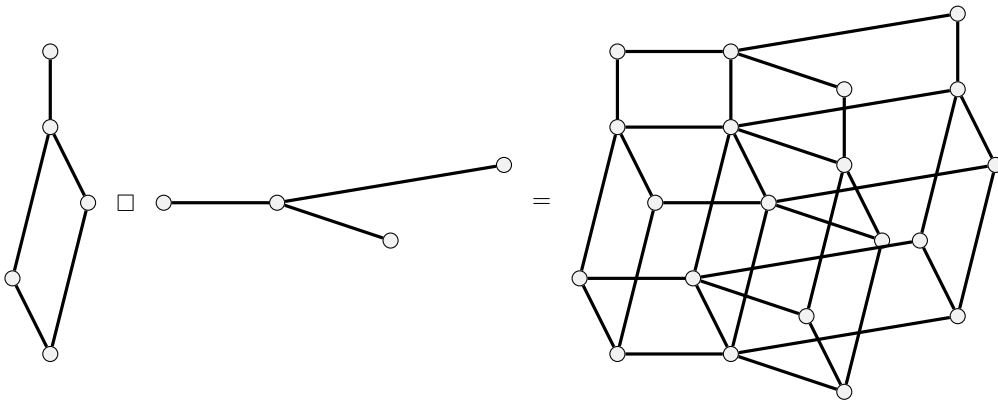

Fig. 1. The cartesian products of two graphs.

3 Hardness of approximating $i(G)$

Since $i(G) = \theta(G)$ [5], we prove hardness of approximation for bounded degree graphs in terms of edge-clique covers. Notice that this result will be the main ingredient of upcoming Proposition 2.

Proposition 1. EDGE-CLIQUE COVER *is* **APX**-*hard for biconnected graphs with maximum degree* 7.

Proof. The reduction is from VERTEX COVER in cubic graphs, which is known to be APX-hard [17,1]. Let G be a cubic n-vertex graph. We construct a new graph H as follows. We represent each vertex $u \in \mathbf{V}(G)$ by a triangle T_u with vertices u_0, u_1 and u_2 in the new graph H. These n triangles are all vertex disjoint in H, and each of them can offer a different edge for three connections. Let us be more specific. For each vertex $u \in \mathbf{V}(G)$ with incident edges $e_u[0]$, $e_u[1]$ and $e_u[2]$ (the order is arbitrary), the edge $\{u_i, u_{i+1 \pmod 3}\} \in T_u$, $0 \le i \le 2$, is devoted to the edge $e_u[i] \in \mathbf{E}(G)$. Now, to represent an edge $\{u, v\} \in \mathbf{E}(G)$ in H, where $\{u_i, u_{i+1 \pmod 3}\}$, $0 \le i \le 2$, is the edge of the triangle T_u devoted to representing $\{u, v\}$, and $\{v_j, v_{j+1 \pmod 3}\}$, $0 \le j \le 2$, is the edge of the triangle T_v devoted to representing $\{u, v\}$, we introduce two new vertices $A_{u,v}$ and $B_{u,v}$ and the 6 edges $\{A_{u,v}, u_i\}$, $\{A_{u,v}, v_j\}$, $\{B_{u,v}, u_{i+1 \pmod 3}\}$, $\{B_{u,v}, v_{j+1 \pmod 3}\}$, and $\{u_i, v_j\}$. What is left is to add m non-incident edges to H (one additional edge for each edge of G): For each edge $\{u, v\} \in \mathbf{E}(G)$ in H, where
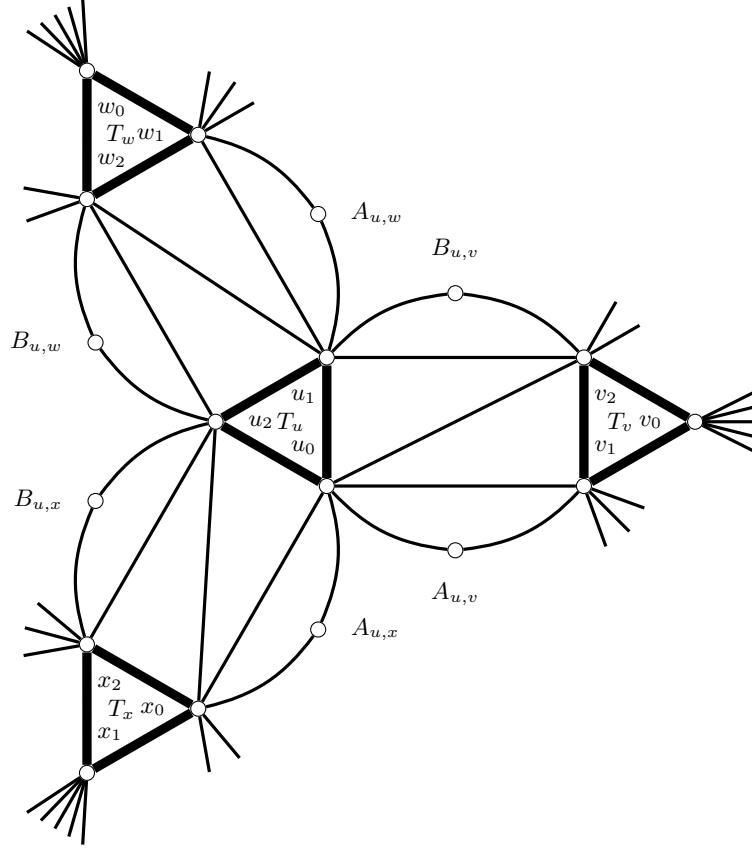
Fig. 2. The three edge-gadgets used in the proof of Proposition 1 for vertex $u \in \mathbf{V}(G)$ with edges $e_u[0] = \{u, v\}$, $e_u[1] = \{u, w\}$, and $e_u[2] = \{u, x\}$.

$\{u_i, u_{i+1 \pmod 3}\}$, $0 \leq i \leq 2$, is the edge of the triangle T_u devoted to representing $\{u, v\}$, and $\{v_j, v_{j+1 \pmod 3}\}$, $0 \leq j \leq 2$, is the edge of the triangle T_v devoted to representing $\{u, v\}$, we add the edge $\{u_i, v_{j+1 \pmod 3}\}$ or $\{u_{i+1 \pmod 3}, v_j\}$ (the choice is made so that these m additional edges form a matching). We refer the reader to Figure 2 for an illustration. Clearly $|\mathbf{V}(H)| = \frac{9}{2}n$ and $|\mathbf{E}(H)| = \frac{27}{2}n$. Moreover, it follows from the construction that H is a biconnected graph with maximum degree 7.

We claim that G has a vertex cover of size k if and only if $\theta(H) \leq k + 3m$, where $m = \frac{3}{2}n$ is the number of edges of G.

Suppose G has a vertex cover $V' \subseteq \mathbf{V}(G)$ of size k. Construct an edge-clique cover $\mathcal{E}$ of H as follows. For each $u \in V'$, add T_u to $\mathcal{E}$. For each edge $\{u, v\} \in \mathbf{E}(G)$, let $\{u_i, u_{i+1 \pmod 3}\}$, $i \in \{0, 1, 2\}$, be the edge of triangle T_u devoted to representing edge $\{u, v\}$, and $\{v_j, v_{j+1 \pmod 3}\}$, $j \in \{0, 1, 2\}$, be the edge of triangle T_v devoted to representing edge $\{u, v\}$. Without loss of generality, assume $\{u_i, v_{j+1 \pmod 3}\} \in \mathbf{E}(H)$. Add the two cliques $\{u_i, v_j, A_{u,v}\}$ and $\{u_{i+1 \pmod 3}, v_{j+1 \pmod 3}, B_{u,v}\}$ to $\mathcal{E}$. Furthermore, if $u \in V'$, add the clique $\{u_i, v_j, v_{j+1 \pmod 3}\}$ to $\mathcal{E}$, and $\{u_i, u_{i+1 \pmod 3}, v_{j+1 \pmod 3}\}$ otherwise. Since V' is a vertex cover of G, it follows that $\mathcal{E}$ is an edge-clique cover of H of cardinality $k + 3m$.

For the reverse direction, let $\mathcal{E}$ be an edge-clique cover of H. Let $\{u, v\}$ be any edge of G, and let $\{u_i, u_{i+1 \pmod 3}\}$, $i \in \{0, 1, 2\}$, be the edge of triangle T_u de-

voted to representing edge $\{u, v\}$, and $\{v_j, v_{j+1 \ (\text{mod } 3)}\}$, $j \in \{0, 1, 2\}$, be the edge of triangle T_v devoted to representing edge $\{u, v\}$. Without loss of generality, assume $\{u_i, v_{j+1 \ (\text{mod } 3)}\} \in \mathbf{E}(H)$. If we let $H_{u,v}$ stand for be the subgraph of H induced by the subset $\{u_i, u_{i+1 \ (\text{mod } 3)}, A_{u,v}, B_{u,v}, v_j, v_{j+1 \ (\text{mod } 3)}\}$, we make the easy observations (see Figure 2) that (i) 4 cliques are needed to cover the edges of $H_{u,v}$, and (ii) 3 cliques are needed to cover the edges of $H_{u,v}$ if $\{u_i, u_{i+1 \ (\text{mod } 3)}\}$ or $\{v_j, v_{j+1 \ (\text{mod } 3)}\}$ (possibly both) is removed. Therefore, $|\mathcal{E}| = 3m + k$ for some non-negative integer $k \leq m$. But each triangle T_w, $w \in \mathbf{V}(G)$, can be covered by 1 clique, and hence there is no loss of generality in assuming $k \leq n$. Furthermore, there is no loss go generality in assuming that $\mathcal{E}$ satisfies the following property: for every edge $\{u, v\} \in \mathbf{E}(G)$, either T_u or T_v (possibly both) is in $\mathcal{E}$. Let $V' \subseteq \mathbf{V}(G)$ be the subset defined as follows: $u \in V'$ if $T_u \in \mathcal{E}$. According to the above, if $|\mathcal{E}| = 3m + k$ for some non-negative integer $k \leq n$, then $|V'| = k$ is a vertex cover of G. $\qquad \square$

4 Hardness of approximating $\varphi(G)$

This section is devoted to proving that there exists a constant $c > 1$ such that computing the overlap number of a graph is hard to approximate to within c.

Proposition 2. OVERLAP NUMBER *is* **APX**-*hard.*

Proof. According to Proposition 1, there exists a constant $c > 0$ such that $\theta(G)$ cannot be approximated to within c (unless $\mathbf{P} = \mathbf{NP}$). We shall prove that a $\sqrt{c}$-approximation algorithm for OVERLAP NUMBER would yield a c-approximation algorithm for EDGE-CLIQUE COVER

Let G be a n-vertex graph for which we are asked to c-approximate $\theta(G)$. Without loss of generality, we assume that G has no isolated vertices and is biconnected (see Proposition 1). Let m be the smallest integer such that $m \geq n$ and $\frac{m}{m-1} < c$, and let K_m be the complete graph on m vertices. Let $H = K_m \times G$ be the cartesian product of K_m by G. For the sake of simplicity, write $\mathbf{V}(K_m) = \{u_1, u_2, \ldots, u_m\}$ and $\mathbf{V}(G) = \{v_1, v_2, \ldots, v_n\}$. We have divided the proof into a sequence of claims.

Claim 1. $\varphi(H) \leq n + m \, \theta(G)$.

Proof (Of Claim 1). Let $k = \theta(G)$ and let $\mathcal{E} = \{Q_1, Q_2, \ldots, Q_k\}$ be a size-k edge-clique cover of G. To every node $(u_i, v_j) \in \mathbf{V}(H)$, we associate the set $S_{(u_i, v_j)}$ defined as follows:

$$S_{(u_i, v_j)} = \{v_j\} \cup \{(u_i, p) : v_j \in Q_p\}.$$

Consider the family $\mathcal{F} = \{S_{(u_i, v_j)} : (u_i, v_j) \in \mathbf{V}(H)\}$ defined over the ground set

$$X = \bigcup_{(u_i, v_j) \in \mathbf{V}(H)} S_{(u_i, v_j)} = \mathbf{V}(G) \cup (\mathbf{V}(K_m) \times [k]),$$

where $[k]$ is the set of the first k integers $\{1, 2, \ldots, k\}$. Notice that $|X| = n + km$. We prove that $O(\mathcal{F})$ and H are isomorphic graphs, thereby proving the claim. Indeed, let $S_{(u_i, v_j)}$ and $S_{(u_r, v_s)}$ be two subsets of $\mathcal{F}$. We need to consider 3 cases.

- If $u_i \neq u_r$ and $v_j \neq v_s$, then (u_i, v_j) and (u_r, v_s) are not adjacent vertices in H. It can be easily verified that $S_{(u_i, v_j)}$ and $S_{(u_r, v_s)}$ are disjoint subsets, and hence $S_{(u_i, v_j)}$ and $S_{(u_r, v_s)}$ are not adjacent vertices in $O(\mathcal{F})$.

– If $u_i \neq u_r$ and $v_j = v_s$, then (u_i, v_j) and (u_r, v_s) are adjacent vertices in H since K_m is a clique. Firstly, $v_j \in S_{(u_i,v_j)}$ and $v_j \in S_{(u_r,v_s)}$ since $v_j = v_s$, and hence $S_{(u_i,v_j)} \cap S_{(u_r,v_s)} \neq \emptyset$. Secondly, both $v_j \in S_{(u_i,v_j)} \setminus S_{(u_r,v_s)}$ and $v_s \in S_{(u_r,v_s)} \setminus S_{(u_i,v_j)}$ are non-empty (*i.e.*, the two sets have some private element) since $u_i \neq u_r$ and v_j is not an isolated vertex of G. Therefore, $S_{(u_i,v_j)}$ and $S_{(u_r,v_s)}$ overlap, and hence $S_{(u_i,v_j)}$ and $S_{(u_r,v_s)}$ are adjacent vertices in $O(\mathcal{F})$.

– If $u_i = u_r$ and $v_j \neq v_s$, then (u_i, v_j) and (u_r, v_s) are adjacent vertices in H if and only if $\{v_i, v_j\} \in \mathbf{E}(G)$. We have $v_j \in S_{(u_i,v_j)} \setminus S_{(u_r,v_s)}$ and $v_s \in S_{(u_r,v_s)} \setminus S_{(u_i,v_j)}$ (*i.e.*, the two sets have some private element) Therefore, the two sets overlap if and only if v_j and v_j belong to a same Q_p for some $1 \leq p \leq k$, which amounts to saying that $\{v_i, v_j\} \in \mathbf{E}(G)$. Hence, $S_{(u_i,v_j)}$ and $S_{(u_r,v_s)}$ are adjacent vertices in $O(\mathcal{F})$ if and only if $\{v_i, v_j\} \in \mathbf{E}(G)$.

□

For the reverse direction, we need the following technical claim that deals with containment in overlap representations of H.

Claim 2. Let $(\mathcal{F} = \{S_{(u_i,v_j)} : (u_i, v_j) \in \mathbf{V}(H)\}, X)$ be an overlap representation of H. If $S_{(u_r,v_s)} \subset S_{(u_i,v_j)}$ for some vertices (u_i, v_j) and (u_r, v_s) of H, then $S_{(u_p,v_q)} \subset S_{(u_i,v_j)}$ for every vertex (u_p, v_q) of H which is not adjacent to vertex (u_i, v_j).

Proof (Of claim 2). First, if $S_{(u_r,v_s)} \subset S_{(u_i,v_j)}$ then vertices (u_r, v_s) and (u_i, v_j) are not adjacent in H since $(\mathcal{F}, X)$ is an overlap representation of H. Now, let (u_p, v_q) be any vertex of H distinct from (u_r, v_s) that is not adjacent to (u_i, v_j). Also, let H' be the graph obtained from H by deleting every vertex in the close neighborhood of vertex (u_i, v_j). Notice that, since (u_r, v_s) and (u_p, v_q) are not adjacent to (u_i, v_j) in H, they are both vertices of H'. We claim that there exists a path between vertices (u_r, v_s) and (u_p, v_q) in H'. Indeed, since G is biconnected there exists a path in G between vertices v_s and v_q that does not go through vertex v_j, and hence there exists a path in H' between vertices (u_r, v_s) and (u_r, v_q). If $u_r = u_p$ we are done. Otherwise, since each column of H' is a clique then the two vertices (u_r, v_q) and (u_p, v_q) are connected by an edge in H'.

To prove the claim it is now enough to show that $S_{(u_p,v_q)} \subset S_{(u_i,v_j)}$ for any vertex (u_p, v_q) of H that is adjacent to (u_r, v_s) but not to (u_i, v_j). The proof follows from an easy contradiction. Suppose $S_{(u_p,v_q)} \not\subset S_{(u_i,v_j)}$. Since $S_{(u_p,v_q)} \neq \emptyset$ (H does not contain any isolated vertex), then there exists $x \in X$ such that $x \in S_{(u_p,v_q)}$ and $x \notin S_{(u_i,v_j)}$, and hence $S_{(u_p,v_q)} \setminus S_{(u_i,v_j)} \neq \emptyset$. Furthermore, (u_p, v_q) and (u_r, v_s) are adjacent vertices in H, and hence (since $S_{(u_p,v_q)}$ and $S_{(u_r,v_s)}$ have to overlap) there exist $x', x'' \in X$ such that (i) $x' \in S_{(u_p,v_q)}$ and $x' \in S_{(u_r,v_s)}$, and (ii) $x'' \notin S_{(u_p,v_q)}$ and $x'' \in S_{(u_r,v_s)}$. But $S_{(u_r,v_s)} \subset S_{(u_i,v_j)}$, and hence $x' \in S_{(u_i,v_j)}$ and $x'' \in S_{(u_i,v_j)}$. Then it follows that $S_{(u_i,v_j)} \setminus S_{(u_p,v_q)} \neq \emptyset$ and $S_{(u_p,v_q)} \cap S_{(u_i,v_j)} \neq \emptyset$, and hence $S_{(u_p,v_q)}$ and $S_{(u_i,v_j)}$ overlap. This is the sought contradiction since this would result in (u_p, v_q) and (u_i, v_j) being adjacent in H. □

We are now ready for the reverse direction of the reduction.

Claim 3. $\theta(G) \leq \dfrac{\varphi(H) - n - 1}{m - 1} + 7.$

Proof (Of claim 3). Let $(\mathcal{F} = \{S_{(u_i,v_j)} : (u_i,v_j) \in \mathbf{V}(H)\}, X)$ be an overlap representation of H. Suppose that there exists some subset $S_{(u_i,v_j)} \in \mathcal{F}$ that strictly contains at least one set of $\mathcal{F}$. Then, according to Claim 2, $S_{(u_i,v_j)}$ contains all subsets $S_{(u_r,v_s)} \in \mathcal{F}$ such that $u_i \neq u_r$ and $v_j \neq v_s$ (*i.e.*, $S_{(u_i,v_j)}$ contains all those subsets of $\mathcal{F}$ that are associated to vertices of H that are not in the same row nor column of vertex (u_i,v_j)). Furthermore, if there exist subsets $S_{(u_r,v_s)}, S_{(u_p,v_q)} \in \mathcal{F}$ distinct from $S_{(u_i,v_j)}$ such that $S_{(u_r,v_s)} \subset S_{(u_p,v_q)}$, then $u_i = u_p$ or $v_j = v_q$ (*i.e.*, vertex (u_p,v_q) is on the same row or on the same column of vertex (u_i,u_j)). Indeed, assuming $u_i \neq u_p$ and $v_j \neq v_q$, Claim 2 would yield to $S_{(u_i,v_j)} \subset S_{(u_p,v_q)}$ and $S_{(u_i,v_j)} \subset S_{(u_p,v_q)}$, a contradiction. Now, let H' be the graph obtained from H by deleting all vertices (u_r,v_s) such that $u_r = u_i$ or $v_s = v_j$ (*i.e.*, deleting all vertices that are in the same row or column of vertex (u_i,v_j)). Also, let $\mathcal{F}' \subseteq \mathcal{F}$ be those subsets of $\mathcal{F}$ that correspond to vertices of H', and $X' \subseteq X$ be the union of the subsets in $\mathcal{F}'$ (X' is the ground set of $\mathcal{F}'$). Notice that $\mathcal{F}'$ is an overlap representation of H' where no subset being a subset of another, and that $|X'| \leq |S_{(u_i,v_j)}|$ since every subset of $\mathcal{F}'$ is strictly contained in $S_{(u_i,v_j)}$. Moreover, if we let G' stand for the graph obtained from G by deleting vertex v_j we have $H' = K_{m-1} \times G'$. We now claim that $\theta(G') \leq \frac{|X|-n-1}{m-1}$. Indeed, consider the "*edge-multi-coloring*" procedure of H' defined by assigning to every edge $e = \{(u_r,v_s), (u_p,v_q)\}$ of H' the "*colors*" $\mathrm{col}(e) = S_{(u_r,v_s)} \cap S_{(u_p,v_q)}$. Since $\mathcal{F}'$ is an overlap representation of H', it follows that at least one color is assigned to every edge of H'. Furthermore, since no subset being a subset of another in $\mathcal{F}'$, it follows that for every color c, $\{e \in \mathbf{E}(H') : c \in \mathrm{col}(e)\}$ induces a clique in G', and hence H' can be covered with at most $|X|$ cliques. But the maximal cliques of H' are either columns (there are $n-1$ of these and at least $n-1$ vertical edges must have received a different color), or are contained in a single row and correspond to maximal clique of G. Therefore, $m-1$ disjoint copies of G' can be covered with at most $|X|-n-1$ cliques. This proves $\theta(G') \leq \frac{|X|-n-1}{m-1}$. What is left is to prove $\theta(G) \leq \dfrac{\varphi(H)-n-1}{m-1} + 7$. This follows from $\theta(G') \leq \frac{|X|-n-1}{m-1}$ and $\theta(G) \leq \theta(G') + \Delta(G)$. $\qquad\square$

Suppose, aiming at a contradiction, that there exists a $\sqrt{c}$-approximation algorithm B for EDGE-CLIQUE COVER. Then, we have $\mathsf{B}(H) \leq \sqrt{c}\,\varphi(H)$. Combining this inequality with Claim 1 yield $\mathsf{B}(H) \leq \sqrt{c}\,(n+m\,\theta(G))$. We now apply the constructive proof of Claim 3 to obtain an approximate $\mathsf{A}(G)$ of $\theta(G)$. We have

$$
\begin{aligned}
\mathsf{A}(G) &\leq \frac{\mathsf{B}(H)-n-1}{m-1} + 7 \\
&= \frac{\mathsf{B}(H)}{m-1} - \frac{n+1}{m-1} + 7 \\
&\leq \frac{\sqrt{c}\,(n+m\,\theta(G))}{m-1} - \frac{n+1}{m-1} + 7 \\
&\leq \frac{n\,\sqrt{c}}{m-1} + \frac{(\sqrt{c})^2\,\theta(G)}{m-1} - \frac{n+1}{m-1} + 7 \\
&\leq (\sqrt{c})^2\,\theta(G) + \frac{n\,\sqrt{c}}{m-1} - \frac{n+1}{m-1} + 7 \\
&= c\,\theta(G) + O(1).
\end{aligned}
$$

The constant makes no problem since $\theta(G)$ is bound to grow with n since we assume $\Delta(G)$ is bounded. $\qquad\square$

5 Recognizing graphs with fixed intersection or overlap number

A central algorithmic problem corresponding to an intersection (*resp.* overlap) graph class $\mathcal{G}$ is the so called $\mathcal{G}$-RECOGNITION problem: Given a graph G, is $G \in \mathcal{G}$? In this section, we show that this problem is linear-time solvable for graphs with fixed intersection or overlap numbers.

Proposition 3. *Let $\mathcal{G}$ be any intersection (resp. overlap) graph class, and let $k \in \mathbb{N}$. The $\mathcal{G}$-RECOGNITION problem can be solved in linear-time when restricted to graphs with intersection (resp. overlap) number at most k.*

For ease of presentation, we will prove this result for graphs with small intersection numbers. The proof for overlap numbers is along the same lines.

The proof of proposition 3 heavily relies on the notion of well quasi orders. A quasi order (*i.e.*, a binary reflexive transitive relation) is a *well quasi order* (or *wqo* for short) if it does not contain infinitely descending sequences nor infinite antichains. For example, the standard order $\leq$ of the natural numbers $\mathbb{N}$ is a well quasi order. Another less obvious example is given by considering vectors in $\mathbb{N}^d$. For two vectors $\overrightarrow{x} \in \mathbb{N}^{K_1}$ and $\overrightarrow{y} \in \mathbb{N}^{K_2}$, let us write $\overrightarrow{x} \leq \overrightarrow{y}$ if $K_1 \leq K_2$ and $x_i \leq y_i$ for all $i \in \{1, \ldots, K_1\}$. The following lemma follows directly from a classical result known in the literature as Higman's Lemma [8].

Lemma 1. *The set $\mathbb{N}^K$ is well quasi ordered by $\leq$ for any fixed $K \in \mathbb{N}$.*

A *characteristic vector* of a graph G is a vector $\overrightarrow{c} \in \mathbb{N}^K$ such that there exists a partitioning $\{V_1, V2, \ldots, V_k\}$ of $\mathbf{V}(G)$ satisfying the two following properties for each $i \in \{1, 2, \ldots, K\}$: (i) $|V_i| = c_i$, and (ii) $N[u] = N[v]$ for all $u, v \in V_i$. Define the *dimension* of a graph G to be the minimum number K such that G has a characteristic vector of dimension K.

Lemma 2. *A graph G with $i(G) \leq k$ has dimension at most $K = 2^k$.*

Proof. Since $i(G) \leq k$, G has a representation $\mathcal{F}$ with $|\bigcup F| \leq k$, which implies that there are at most 2^k distinct sets in $\mathcal{F}$. Since vertices of G corresponding to identical sets in $\mathcal{F}$ have identical neighborhoods, the lemma follows. □

According to the above definition, a graph G of dimension K may have several characteristic vectors $\overrightarrow{c} \in \mathbb{N}^K$. However, all of these must be isomorphic. We say that two characteristic vectors $\overrightarrow{c}, \overrightarrow{d} \in \mathbb{N}^K$ are isomorphic if there is a permutation $\pi \in S_K$ such that $c_i = d_{\pi(i)}$ for all $i \in \{1, 2, \ldots, K\}$. It is not difficult to see that two graphs are isomorphic if and only if they both have the same dimension K, and any pair of characteristic vectors of dimensions K for these graphs are isomorphic. The immediately implies the following lemma whose proof is omitted.

Lemma 3. *A graph H is an induced subgraph of a graph G if and only if H has a characteristic vector $\overrightarrow{c}(H)$ with $\overrightarrow{c}(H) \leq \overrightarrow{c}(G)$ for some characteristic vector $\overrightarrow{c}(G)$ of G.*

We next show two applications of Lemma 3. The first application allows us to show that graphs of bounded intersection number are wqo by the induced subgraph order.

Lemma 4. *Let $k \in \mathbb{N}$. The set of all graphs G with $i(G) \leq k$ is wqo by the induced subgraph order.*

Proof. Let $\mathcal{G}$ denote the set of all finite graphs G with $i(G) \leq k$. Clearly there are no infinite descending sequences in $\mathcal{G}$ w.r.t. the induced subgraph order, and so to prove the lemma it suffices to show that there are also no infinite antichains in G. Consider a infinite subset of graphs $\mathcal{G}' \subseteq \mathcal{G}$, and let $\overrightarrow{C}(\mathcal{G}')$ denote a set of characteristic vectors of minimum dimension for each graph in $\mathcal{G}'$. By Lemma 2 all vectors in $\overrightarrow{C}(\mathcal{G}')$ have dimension at most $K = 2^k$, and so Lemma 1 implies that $\overrightarrow{C}(\mathcal{G}')$ is wqo by $\leq$. But then $\mathcal{G}'$ cannot be an antichain w.r.t. to the induced subgraph order according to Lemma 3. □

INDUCED SUBGRAPH problem asks to determine given a pair of input graph H and G whether H is an induced subgraph of G. The second application of Lemma 3 is that INDUCED SUBGRAPH can be solved in linear-time when restricted to graphs of bounded intersection number.

Lemma 5. *Let $k \in \mathbb{N}$. There is a linear time algorithm for* INDUCED SUBGRAPH *when restricted to graphs H and G with intersection number at most k.*

Proof. Let N_H and N_G denote the total number of vertices and edges in H and G, respectively, and let K_H and K_G respectively denote the dimensions of H and G. By Lemma 2 we have $K_H, K_G \leq 2^k$. We first compute a characteristic vector $\overrightarrow{c}(H) \in \mathbb{N}^{K_H}$ for H, and a characteristic vector $\overrightarrow{c}(G) \in \mathbb{N}^{K_G}$ for H. This can be done in $O(K_H \cdot N_H + K_G \cdot N_G)$ time by a simple linear scan of the neighborhoods in H and G. Next we check for each permutation $f : \{1, \ldots, K_H\} \to \{1, \ldots, K_H\}$, whether $\overrightarrow{c}_f(H) \leq \overrightarrow{c}(G)$, where $\overrightarrow{c}_f(H)$ is the vector $\overrightarrow{c}(H)$ whose entries are permutated by f. By Lemma 3 this check suffices to determine whether H is an induced subgraph of G. Since this check requires $O(K_H!)$ time, the lemma follows. □

We are now ready to give the proof of Proposition 3.

Proof (of proposition 3). Let $\mathcal{G}$ any intersection graph class, and let $\overline{\mathcal{G}}$ denote the set of all finite graphs not in $\mathcal{G}$. Also, let $\mathcal{H}$ denote the set of all minimal graphs in $\overline{\mathcal{G}}$ w.r.t. the induced subgraph order. That is,

$$\mathcal{H} := \{H \in \overline{\mathcal{G}} : \nexists H' \in \overline{\mathcal{G}} \text{ such that } H \text{ is an induced subgraph of } H\}.$$

Observe that $\mathcal{G}$ is closed under induced subgraphs (*i.e.*, $H \in \mathcal{G}$ whenever H is an induced subgraph of some graph $G \in \mathcal{G}$). This implies that a graph G belongs to $\mathcal{G}$ if and only if no graph $H \in \mathcal{H}$ is an induced subgraph of G.

Now by Lemma 4, the set $\mathcal{H}$ is finite, and its size depends only on $\mathcal{G}$. Thus our recognition algorithm for $\mathcal{G}$ has the set of graphs $\mathcal{H}$ "hard-wired" into it, and on given input graph G, it simply checks whether any $H \in \mathcal{H}$ is an induced subgraph of G, determining that $G \notin \mathcal{G}$ if and only if any of these checks turns out positive. The running-time of this algorithm is linear by Lemma 5, and since the number and sizes of graphs in $\mathcal{H}$ is constant w.r.t. the size of G. □

References

1. P. Alimonti and V. Kann, *Some APX-completeness results for cubic graphs*, **237** (2000), no. 1-2, 123–134.
2. G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protasi, *Complexity and Approximation: Cominatorial optimization problems and their approximability properties*, Springer-Verlag, 1999.
3. M.-S. Chang and H. Müller, *On the tree-degree of graphs*, Proc. 27st International Workshop International Workshop on Graph-Theoretic Concepts in Computer Science (WG), Boltenhagen, Germany (A. Brandstädt and V. Bang Le, eds.), Lecture Notes in Computer Science, vol. 2204, 2001, pp. 44–54.
4. D.W. Cranston, N. Korula, T.D. LeSaulnier, K. Milans, C. Stocker, J. Vandenbussche, and D.B. West, *Overlap number of graphs*, Journal of Graph Theory (2011), (in press, 21 pages, posted online 18 February 2011), DOI: 10.1002/jgt.20596).
5. P. Erdös, A.W. Goodman, and L. Pósa, *The intersection of a graph by set intersections*, Canad. J. Math. **18** (1966), 106–112.
6. J. Gramm, J. Guo, F. Hüffner, and R. Niedermeier, *Data reduction and exact algorithms for clique cover*, ACM Journal of Experimental Algorithmics **13** (2008).
7. J.-L. Guillaume and M. Latapy, *Bipartite structure of all complex networks*, Information Processing Letters **90** (2004), no. 5, 215–221.
8. G. Higman, *Ordering by divisibility in abstract algebras*, Proceedings of the London Mathematical Society III **2** (1952), no. 7, 326–336.
9. D.N. Hoover, *Complexity of graph covering problems for graphs of low degree*, Journal of Combinatorial Mathematics and Combinatorial Computing **11** (1992), 187–200.
10. W.-L. Hsu and K.-H. Tsai, *Linear time algorithms on circular-arc graphs*, Information Processing Letters **40** (1991), no. 3, 123–129.
11. L.T. Kou, L.J. Stockmeyer, and C.K. Wong, *Covering graphs by cliques with regard to keyword conflicts and intersection graphs*, Comm. ACM **21** (1978), 135–139.
12. C. Lund and M. Yannakakis, *The approximation of maximum subgraph problems*, Proc. 20th Int. Colloqium on Automata, Languages and Programming (ICALP), Jerusalem, Israel (S. Abiteboul and E. Shamir, eds.), Lecture Notes in Computer Science, vol. 820, Springer, 1994, pp. 40–51.
13. S. Ma, W.D. Wallis, and J.Wu, *Clique covering of chordal graphs*, Utilitas Mathematica **36** (1989), 151–152.
14. T.A. McKee and F.R. McMorris, *Topics in intersection graph theory*, SIAM monographs on discrete mathematics and applications, 1999.
15. I. Nor, D. Hermelin, S. Charlat, J. Engelstadter, M. Reuter, O. Duron, and M.F. Sagot, *Mod/resc parsimony inference*, Proc. 21st Annual Symposium on Combinatorial Pattern Matching (CPM) (A. Amir and L. Parida, eds.), Lecture Notes in Computer Science, vol. 6129, Springer, 2010, pp. 202–213.
16. J.B. Orlin, *Contentment in graph theory: Covering graphs with cliques*, Indagationes Mathematicae (Proceedings) **80** (1977), no. 5, 406–424.
17. C.H. Papadimitriou and M. Yannakakis, *Optimization, approximation and complexity classes*, Journal of Computer and System Sciences **43** (1991), 425–440.
18. B. Rosgen, *Set representations of graphs*, Master's thesis, University of Alberta, 2005.
19. B. Rosgen and L. Stewart, *The overlap number of a graph*, (2010), Submitted, (20 pages, 2 figures, arXiv:1008.2170v2).