



**HAL**  
open science

# NP-hardness of the Sparsest $k$ -Subgraph Problem in Chordal Graphs

Rémi Watrigant, Marin Bougeret, Rodolphe Giroudeau

► **To cite this version:**

Rémi Watrigant, Marin Bougeret, Rodolphe Giroudeau. NP-hardness of the Sparsest  $k$ -Subgraph Problem in Chordal Graphs. RR-12026, 2012. lirmm-00744655v3

**HAL Id: lirmm-00744655**

**<https://hal-lirmm.ccsd.cnrs.fr/lirmm-00744655v3>**

Submitted on 7 Jun 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# *NP*-hardness of the SPARSEST $k$ -SUBGRAPH Problem in Chordal Graphs<sup>\*</sup>

R. Watrigant, M. Bougeret, and R. Giroudeau

LIRMM-CNRS-UMR 5506 - 161, rue Ada 34090 Montpellier, France

**Abstract** Given a simple undirected graph  $G = (V, E)$  and an integer  $k \leq |V|$ , the SPARSEST  $k$ -SUBGRAPH problem asks for a set of  $k$  vertices which induce the minimum number of edges. Whereas its special case INDEPENDENT SET and many other optimization problems become polynomial-time solvable in chordal graphs, we show that SPARSEST  $k$ -SUBGRAPH remains *NP*-hard in this graph class.

## 1 Introduction and Preliminaries

In this report we study the following decision problem:

**SPARSEST  $k$ -SUBGRAPH**

- Input: a simple undirected graph  $G = (V, E)$ ,  $k \in \mathbb{N}$ ,  $C \in \mathbb{N}$
- Question: is there a subset  $S \subseteq V$  such that  $|S| = k$  and  $E(S) \leq C$ ? Where  $E(S)$  is the number of edges induced by  $S$ .

As a generalization of the classical INDEPENDENT SET problem (for which we have  $C = 0$  in the input), SPARSEST  $k$ -SUBGRAPH is *NP*-hard [7] and even not approximable unless  $P = NP$ . Moreover, it is  $W[1]$ -hard (parameterized by  $k$ ) [6].

Its maximization version, namely the  $k$ -DENSEST SUBGRAPH (or the  $k$ -CLUSTER problem), has been extensively studied in the last three decades: in [5], the authors show that  $k$ -DENSEST SUBGRAPH is  $\mathcal{NP}$ -hard in bipartite, comparability and chordal graphs, and is polynomial-time solvable in trees, cographs, bounded treewidth graphs and split graphs. The question of the complexity status of  $k$ -DENSEST SUBGRAPH in interval graphs (and even in proper interval graphs) is stated by the authors as an open problem, and is still not answered yet. In addition, [4] shows that both SPARSEST  $k$ -SUBGRAPH and  $k$ -DENSEST SUBGRAPH are polynomial time solvable in bounded cliquewidth graphs. Notice that several exact or approximation algorithm exists for  $k$ -DENSEST SUBGRAPH in subclasses of perfect graphs: among others, constant approximation algorithms are known for chordal graphs [10], bipartite permutation graphs [3] and *PTAS* are known for interval graphs [11] and for chordal graphs having a special clique tree [9]. Unfortunately, most of these results seem useless for SPARSEST  $k$ -SUBGRAPH, as we apparently need to complement the input graph to apply them. Nevertheless we can deduce that SPARSEST  $k$ -SUBGRAPH remains *NP*-hard in co-chordal (which is a subclass of perfect graphs) and is polynomial-time solvable in split graphs.

On the other side, its dual version, namely the MAXIMUM PARTIAL VERTEX COVER problem, for which we are looking for  $k$  vertices in the input graph which *cover* the maximum number of edges, is polynomial-time solvable in line graphs [2], and remains *NP*-hard in bipartite graphs [1,8].

In this report we study the complexity status of SPARSEST  $k$ -SUBGRAPH in chordal graphs.

---

<sup>\*</sup> This work has been funded by grant ANR 2010 BLAN 021902

Whereas the INDEPENDENT SET problem is polynomial-time solvable in perfect graphs (and thus in chordal graphs), we show that SPARSEST  $k$ -SUBGRAPH remains  $NP$ -hard in chordal graphs. Obviously, the same result holds for the MAXIMUM PARTIAL VERTEX COVER problem.

The two following definitions of chordal graphs are equivalent:

- A graph is chordal if there it does not contain any cycle of length four or more as an induced subgraph.
- A vertex  $v$  of  $G$  is called simplicial if its neighbourhood  $N(v)$  is a clique. The ordering  $v_1, \dots, v_n$  of the vertices of  $G$  is a simplicial elimination scheme if for all  $i$ ,  $v_i$  is simplicial in  $G[v_i, \dots, v_n]$ . A graph is chordal if it has a simplicial elimination scheme.

## 2 The Main Result

### 2.1 Idea of the Proof

The following  $\mathcal{NP}$ -hardness proof is a reduction from the  $k$ -CLIQUE problem in general graphs. Roughly speaking, given an input instance  $G = (V, E)$  together with  $k \in \mathbb{N}$ , we construct the split graph of adjacencies of  $G$ , *i.e.* we build a clique on a set  $A$  representing the vertices of  $G$ , and an independent on a set  $F$  representing the edges of  $G$ , connecting  $A$  and  $F$  with respect to the adjacencies of the graph. Then, we duplicate each vertex of  $A$   $n$  times, creating thus a clique of size  $n^2$ . On the other hand, we replace each vertex of the independent set by a gadget. If  $G$  contains a clique of size  $k$ , that is a set of  $k$  vertices inducing  $\binom{k}{2}$  edges, then the solution will take vertices *not* corresponding to vertices of the clique. Hence, there will be  $\binom{k}{2}$  gadgets *not* adjacent to the solution. Finally, we will force the solution to take the same number of vertices among each gadget.

### 2.2 $\mathcal{NP}$ -hardness

**Theorem 1.** SPARSEST  $k$ -SUBGRAPH remains  $\mathcal{NP}$ -hard in chordal graphs.

*Proof.* We reduce from the classical  $k$ -CLIQUE problem in general graphs. Let  $G = (V, E)$  and  $k \in \mathbb{N}$ . We note  $|V| = n$ ,  $V = \{v_1, \dots, v_n\}$ ,  $|E| = m$  and  $T = n(n - k)$ . In the following we will define  $G' = (V', E')$  together with  $k', C' \in \mathbb{N}$  such that:

- $G', k', C'$  can be constructed in polynomial time
- $G'$  is a chordal graph
- $G$  contains a clique of size  $k$  if and only if one can find  $k'$  vertices in  $G'$  which induce  $C'$  edges or less.

#### The construction

$V'$  is composed of two parts  $A$  and  $F$ .

- We first define a clique over  $A = \{a_i^j : i, j \in \{1, \dots, n\}\}$ . Thus,  $A$  is a clique of size  $n^2$ . Moreover, for all  $j \in \{1, \dots, n\}$ , we note  $A_j = \{a_1^j, \dots, a_n^j\}$ .
- For all  $e \in E$ , we construct a graph with  $F_e$  as vertex set, where  $F_e$  is composed of three sets of  $T$  vertices:  $X_e = \{x_1^e, \dots, x_T^e\}$ ,  $Y_e = \{y_1^e, \dots, y_T^e\}$  and  $Z_e = \{z_1^e, \dots, z_T^e\}$ . The set  $X^e$  induces a stable set,  $Z^e$  induces a clique, and  $Y^e$  contains a clique of size  $T - 1$  on vertices  $\{y_2^e, \dots, y_T^e\}$  (thus,  $y_1^e$  is not connected to the other vertices of  $Y^e$ ). Then, for all  $j \in \{1, \dots, T\}$ ,  $x_j^e$  is connected to  $y_j^e$ , and  $y_j^e$  is connected to all vertices of  $Z^e$ . An example of such a gadget is represented in Figure 1. We define  $F = \bigcup_{e \in E} F_e$ .

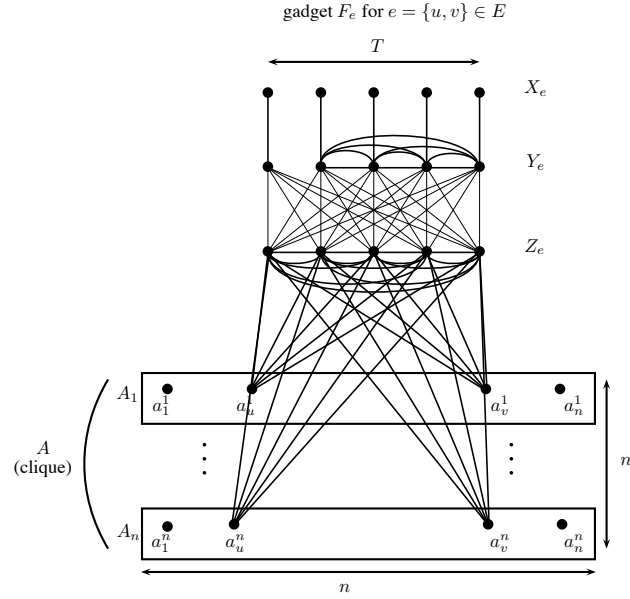


Figure1: Example of a gadget  $F_e$  (with  $T = 5$ ) and its relations to  $A$ .

- For all  $e = \{v_p, v_q\} \in E$ , all vertices of  $Z^e$  are connected to  $\{a_p^j : j \in \{1, \dots, n\}\}$  and  $\{a_q^j : j \in \{1, \dots, n\}\}$ .
- We define  $k' = m2T + T$  and  $C' = m\binom{T}{2} + \binom{T}{2} + (m - \binom{k}{2})$ .

**Lemma 1.**  $G'$  is a chordal graph

*Proof.* We have the following simplicial elimination scheme:

- For all  $e \in E$ , we can remove  $X^e$  since for all  $j \in \{1, \dots, T\}$ ,  $x_j^e$  is only connected to  $y_j^e$ .
- For all  $e \in E$ , we can remove  $Y^e$ . Indeed the remaining neighbourhood of  $y_1^e$  is  $Z^e$  which is a clique. And the remaining neighbourhood of  $y_j^e$  with  $j \geq 2$  is a subset  $Y^e \cup Z^e \setminus \{y_1^e\}$  which induces a clique.
- For all  $e \in E$ , we can remove  $Z^e$  since the remaining neighbourhood of  $z_j^e$  is a subset of  $Z^e$  and vertices of  $A$  which induce a clique.
- The remaining vertices induces a clique on  $A$  and thus be eliminated.

□

Now we prove that  $G$  contains a clique of size  $k$  if and only if  $G'$  contains  $k'$  vertices inducing at most  $C'$  edges.

**$G$  contains a  $k$ -clique  $\Rightarrow G'$  contains  $k'$  vertices inducing at most  $C'$  edges.**

Let us suppose that  $K \subseteq V$  is a clique of size  $k$  in  $G$ . Without loss of generality we suppose  $K = \{v_1, \dots, v_k\}$ . Moreover, we note  $E_0 = \{\{v_p, v_q\} \in E \text{ such that } v_p, v_q \in K\}$  and  $E_1 = \{\{v_p, v_q\} \in E \text{ such that } v_p \notin K \text{ or } v_q \notin K\}$ . We construct  $K' \subseteq V'$  as follows:

- For all  $i \in \{(k+1), \dots, n\}$  and all  $j = \{1, \dots, n\}$ , we add  $a_i^j$  to  $K'$ .
- For all  $e \in E$ , we add all vertices of  $X_e$  to  $K'$ .

- For all  $e \in E_0$ , we add all vertices of  $Z_e$  to  $K'$ .
- For all  $e \in E_1$ , we add all vertices of  $Y_e$  to  $K'$ .

One can verify that  $K'$  is a set of  $k' = 2mT + T$  vertices inducing exactly  $C' = m\binom{T}{2} + \binom{T}{2} + (m - \binom{k}{2})$  edges. Indeed, we picked  $T = n(n - k)$  vertices from  $A$  which is a clique and thus induce  $\binom{T}{2}$  edges. Then, for all  $e \in E$ , we picked  $2T$  vertices, which induce  $\binom{T}{2}$  edges if  $e \in E_0$ , and  $(\binom{T}{2} + 1)$  edges if  $e \in E_1$ . Since  $|E_0| = \binom{k}{2}$  (and thus  $|E_1| = m - \binom{k}{2}$ ), we have the desired number of edges.

**$G$  contains a  $k$ -clique  $\Leftrightarrow G'$  contains  $k'$  vertices inducing at most  $C'$  edges.**

Suppose now that  $K'$  is a set of  $k'$  vertices of  $G'$  which induces at most  $C'$  edges. We redefine the sets  $E_0$  and  $E_1$  as follows:  $E_0 = \{\{v_p, v_q\} \in E \text{ such that for all } j \in \{1, \dots, n\} \text{ we have } a_p^j \notin K' \text{ and } a_q^j \notin K'\}$ , and  $E_1 = E \setminus E_0$ .

For all  $R \subseteq V'$ , let  $tr(R) = K' \cap R$  be the trace of  $K'$  on  $R$ , and for all  $v \in V'$ , let  $\mu(v) = |tr(N(v))|$  be the number of neighbors of  $v$  belonging to  $K'$ .

Let  $u \in K'$  and  $v \in V' \setminus K'$ . We say that  $(K' \setminus \{u\}) \cup \{v\}$  is a safe replacement if and only if we have  $\mu(v) \leq \mu(u)$  if  $\{u, v\} \notin E'$  and  $\mu(v) - 1 \leq \mu(u)$  if  $\{u, v\} \in E'$ . For sake of readability, we will keep and update the definitions of  $E_0$  and  $E_1$  when replacing vertices of  $A$  (e.g. if we remove a vertex  $u \in A$  from  $K'$  and that there exists  $e \in E_1$  such that vertices of  $Z_e$  were only adjacent to  $u$  among all vertices of  $A$ , then  $e$  now belongs to  $E_0$ ).

The proof consists in replacing some vertices of  $K'$  by other vertices not in  $K'$  without increasing the number of induced edges, in order to obtain a solution that has the same structure as previously. We call such a replacement a *safe modification* or a *safe replacement*. The core of the proof is based on the three following lemmas.

**Lemma 2.** *Without loss of generality (and optimality of  $K'$ ), we can suppose that for all  $e \in E$  we have  $X_e \subseteq K'$ .*

*Proof.* Let  $S = \bigcup_{e \in E} X_e$ . Since we have  $k' > |S|$ , there always exists  $u \in K' \setminus S$ . Suppose that there exists  $e \in E$  and  $i \in \{1, \dots, T\}$  such that  $x_i^e \notin K'$ . If  $y_i^e \notin K'$ , then we have  $\mu(x_i^e) = 0$  and we can thus safely replace any other vertex of  $K' \setminus S$  by  $x_i^e$ . Now, if  $y_i^e \in K'$ , then  $\mu(x_i^e) = 1$ . Since  $\{x_i^e, y_i^e\} \in E'$ ,  $(K' \setminus \{y_i^e\}) \cup \{x_i^e\}$  is a safe replacement.  $\square$

**Lemma 3.**  *$K'$  can be safely modified such that one of the two following holds:*

*Case A1: for all  $e \in E_0$  we have  $tr(Z_e) = Z_e$ .*

*Case A2: for all  $e \in E_0$  we have  $tr(Y_e) = \emptyset$ .*

*Proof.* Let us first restructure each gadget of  $E_0$  separately. For all  $e \in E_0$  such that  $tr(Y_e) \neq \emptyset$  and  $tr(Z_e) \neq Z_e$ , let  $j_0 = \max\{j \in \{1, \dots, T\} : y_j^e \in tr(Y_e)\}$  and let  $j_1$  be such that  $z_{j_1}^e \notin tr(Z_e)$ . Recall that Lemma 2 ensures that  $x_{j_0}^e$  is in  $K'$ . If  $j_0 \neq 1$ , then  $\mu(y_{j_0}^e) = y + z + 1$ , where  $y = |N(y_{j_0}^e) \cap tr(Y_e)|$  and  $z = |N(y_{j_0}^e) \cap tr(Z_e)|$ . On the other side, we have  $\mu(z_{j_1}^e) \leq y + z + 1$  (more precisely,  $\mu(z_{j_1}^e) = y + z + 1$  if  $y_1^e \in K'$ , and  $\mu(z_{j_1}^e) = y + z$  if  $y_1^e \notin K'$ ). Roughly speaking, this switch ensures that we necessarily “loose” the edge due to the vertex of  $X^e$  and we gain at most one edge due to  $y_1^e$ . Hence  $\mu(z_{j_1}^e) \leq \mu(y_{j_0}^e)$  and  $(K' \setminus \{y_{j_0}^e\}) \cup \{z_{j_1}^e\}$  is a safe replacement. If  $j_0 = 1$ , then it means that  $tr(Y_e) = \{y_1^e\}$ . Suppose that there exists  $j_1$  such that  $z_{j_1}^e \notin tr(Z_e)$ . We have  $\mu(y_1^e) = z + 1$  where  $z = |N(y_1^e) \cap tr(Z_e)|$ , and  $\mu(z_{j_1}^e) = z + 1$ . Here again  $(K' \setminus \{y_1^e\}) \cup \{z_{j_1}^e\}$  is a safe replacement. After all these replacements, given any  $e \in E_0$ ,  $tr(Y_e) \neq \emptyset$  implies that  $tr(Z_e) = Z_e$ .

Then, we proceed to replacements between gadgets  $F_e$ ,  $e \in E_0$ . If one can find  $a, b \in E_0$  such that  $tr(Y_a) \neq \emptyset$  and  $tr(Z_b) \neq Z_b$ , then let  $j_0$  be such that  $y_{j_0}^a \in tr(Y_a)$  and let  $j_1$  be such that  $z_{j_1}^b \notin tr(Z_b)$ . We have  $\mu(y_{j_0}^a) \geq T + 1$  and  $\mu(z_{j_1}^b) \leq T - 1$ . Thus,  $(K' \setminus \{y_{j_0}^a\}) \cup \{z_{j_1}^b\}$  is a safe replacement.

Theses replacements end either when all the  $Y_e$  are empty for all  $e \in E_0$  or when all the  $Z_e$  are full for all  $e \in E_0$ , which achieves the proof of Lemma 3.  $\square$

**Lemma 4.**  $K'$  can be safely modified such that one of the two following holds:

Case B1: for all  $e \in E_1$  we have  $tr(Y_e) = Y_e$ .

Case B2: for all  $e \in E_1$  we have  $tr(Z_e) = \emptyset$ .

*Proof.* The proof is roughly based on the fact that replacing a vertex of  $Z_e$  by a vertex of  $Y_e$  permits to “lose” at least one edge with vertices  $A$  and “gain” one edge with a vertex of  $X_e$ . Let us formally prove Lemma 4. Similarly to the proof of Lemma 3, we first restructure each gadget of  $E_1$  separately: for all  $e \in E_1$  such that  $tr(Z_e) \neq \emptyset$  and  $tr(Y_e) \neq Y_e$ , let  $j_0 = \max\{j \in \{1, \dots, T\} : y_j^e \notin K'\}$  and let  $j_1$  be such that  $z_{j_1}^e \in tr(Z_e)$ . Recall that by definition of  $E_1$ , there exists  $i, j \in \{1, \dots, n\}$  such that  $z_{j_1}^e$  is adjacent to  $a_i^j$ . We have  $\mu(z_{j_1}^e) \geq y + z + 1$ , where  $y = |N(z_{j_1}^e) \cap Y_e|$  and  $z = |N(z_{j_1}^e) \cap Z_e|$ . On the other side, we have  $\mu(y_{j_0}^e) \leq z + y + 2$  (indeed,  $|N(y_{j_0}^e) \cap Z_e| = z + 1$ ,  $|N(y_{j_0}^e) \cap Y_e| \leq y$  and  $|N(y_{j_0}^e) \cap X_e| = 1$ ). Since  $\{y_{j_0}^e, z_{j_1}^e\} \in E'$ , it holds that  $(K' \setminus \{z_{j_1}^e\}) \cup \{y_{j_0}^e\}$  is a safe replacement. After all these replacements, given any  $e \in E_1$ ,  $tr(Z_e) \neq \emptyset$  implies that  $tr(Y_e) = Y_e$ .

We now proceed to replacements between gadgets  $F_e$ ,  $e \in E_1$ . If one can find  $a, b \in E_1$  such that  $tr(Z_a) \neq \emptyset$  and  $tr(Y_b) \neq Y_b$ , then let  $j_0$  be such that  $y_{j_0}^b \notin tr(Y_b)$  and let  $j_1$  be such that  $z_{j_1}^a \in tr(Z_a)$ . We have  $\mu(z_{j_1}^a) \geq T + 1$  and  $\mu(y_{j_0}^b) \leq T - 1$ . Thus  $(K' \setminus \{z_{j_1}^a\}) \cup \{y_{j_0}^b\}$  is a safe replacement.  $\square$

Let us now define for each case and each  $e \in E$  the set of vertices  $D_e \subseteq Y_e \cup Z_e$  that have to be replaced (see Figure 2):

- case A1: for all  $e \in E_0$ ,  $D_e = Y_e \cap K'$
- case A2: for all  $e \in E_0$ ,  $D_e = Z_e \setminus K'$
- case B1: for all  $e \in E_1$ ,  $D_e = Z_e \cap K'$
- case B2: for all  $e \in E_1$ ,  $D_e = Y_e \setminus K'$

Notice that if  $D_e = \emptyset$  for all  $e \in E_0$  (resp.  $e \in E_1$ ), then cases A1 and A2 (resp. B1 and B2) collapse. If such a case happen for all  $e \in E$ , we can immediately conclude, as shown by the following lemma:

**Lemma 5.** If  $D_e = \emptyset$  for all  $e \in E$ , then  $G$  contains a clique of size  $k$ .

*Proof.* By construction, we have  $|tr(A)| = T$  and  $|tr(F_e)| = 2T$  for all  $e \in E$ . Thus,  $cost^*(tr(A)) = \binom{T}{2}$  and  $cost^*(tr(F_e)) = \binom{T}{2} + 1$  if  $Y_e \subseteq K'$ , and  $cost^*(tr(F_e)) = \binom{T}{2}$  if  $Z_e \subseteq K'$ . By construction,  $Y_e \subseteq K'$  if and only if  $e \in E_1$ . Thus, since  $cost^*(K') \leq \binom{T}{2} + m \binom{T}{2} + m - \binom{k}{2}$ , we must have  $|E_1| \leq m - \binom{k}{2}$  which is equivalent to  $|E_0| \geq \binom{k}{2}$ . Hence, there exists at most  $\lfloor \frac{|A| - T}{n} \rfloor = k$  vertices in  $G$  inducing at least  $\binom{k}{2}$  edges, *i.e.*  $G$  contains a clique of size  $k$ .  $\square$

We now have to analyse the four cases of Lemma 3 and 4 (see Figure 2).

*Case A1 and B1*

To summarize the situation, the solution  $K'$  can be partitionned in  $K'_A = K' \cap A$ , and  $K'_F = K' \setminus K'_A$ , the vertices selected in the gadgets. Let  $\Delta_0 = \sum_{e \in E_0} |D_e|$  be the number of extra vertices allocated in all the gadgets  $F_e$ ,  $e \in E_0$ , and  $\Delta_1 = \sum_{e \in E_1} |D_e|$  be the number of extra vertices allocated in all the gadgets  $F_e$ ,  $e \in E_1$ . Let  $\Delta = \Delta_0 + \Delta_1$ . Notice that we have  $|K'_A| = T - \Delta$ , as a "regular" solution that does not select any extra vertex in a gadget has to pick  $T$  vertices in  $A$ . Moreover,

- vertices of  $K'$  selected in gadgets of  $E_0$  are not adjacent to  $K'_A$  (by definition of  $E_0$ )
- each gadget of  $E_0$  induces at least  $\binom{T}{2}$  edges (as we are in case A1)
- each gadget of  $E_1$  induces at least  $\binom{T}{2} + 1$  edges (as we are in case B1)

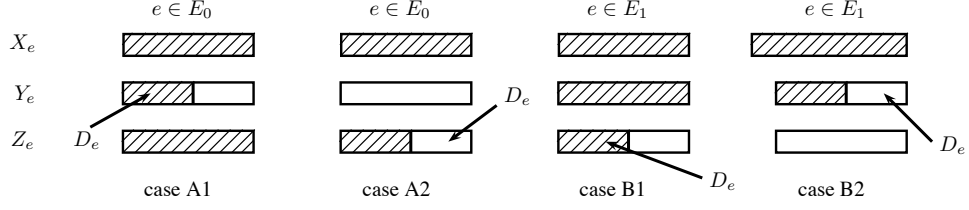


Figure2: Schema of different cases. Shaded rectangles represent part of  $K'$ .

- each of the  $\Delta_0$  vertices is adjacent to at least  $T$  vertices in  $K'$  (such a vertex is in a set  $Y_e$ , and thus is connected to the  $T$  vertices of  $Z_e$ )
- each of the  $\Delta_1$  vertices is adjacent to at least  $T + 1$  vertices in  $K'$  (such a vertex is in a set  $Z_e$ , and thus is connected to at least 1 vertex of  $K'_A$  and to the  $T$  vertices of  $Y_e$ )

Let us now lower bound the total cost of  $K'$ . We have

$$\begin{aligned}
\text{cost}^*(K') &\geq |E_0| \binom{T}{2} + |E_1| \left( \binom{T}{2} + 1 \right) + \Delta_0 T + \Delta_1 (T + 1) + \binom{T - \Delta}{2} \\
&\geq |E_0| \binom{T}{2} + |E_1| \left( \binom{T}{2} + 1 \right) + \Delta T + \binom{T - \Delta}{2} \\
&\geq |E_0| \binom{T}{2} + |E_1| \left( \binom{T}{2} + 1 \right) + \binom{T}{2} + \frac{\Delta^2}{2}
\end{aligned}$$

Notice that in a bad structured solution, a large  $\Delta$  allows to select only a few vertices in  $A$  ( $T - \Delta$  instead of  $T$ ), and thus to have many gadgets (more than  $\binom{k}{2}$ ) in  $E_0$ . Let us now consider the contrapositive, *i.e.* we consider that  $G$  does not contain a  $k$ -clique, and show that  $K'$  induces more than  $C'$  edges.

Let  $q$  and  $r$  such that  $\Delta = qn + r$ ,  $r < n$ . Let us upper bound  $|E_0|$ . As there is  $T - \Delta$  vertices in  $A$ , the number of empty "columns" (column  $u$  is empty iff none of the  $a_u^t$  is selected) is at most  $n - \frac{T - \Delta}{n} \leq k + q$ .

As  $G$  does not contain a  $k$ -clique, the  $k + q$  vertices corresponding to these  $k + q$  columns cannot induce a clique of size  $k + q$ , and thus  $|E_0| < \binom{k + q}{2}$ . Thus, we get

$$\begin{aligned}
\text{cost}^*(K') &> \binom{k + q}{2} \binom{T}{2} + \left( m - \binom{k + q}{2} \right) \left( \binom{T}{2} + 1 \right) + \binom{T}{2} + \frac{\Delta^2}{2} \\
&= C' - \left( \binom{q}{2} + kq \right) + \frac{\Delta^2}{2}
\end{aligned}$$

Thus, as  $\frac{\Delta^2}{2} > \binom{q}{2} + kq$ , we get the desired inequality.

#### Case A2 and B2

Let  $\Delta_0 = \sum_{e \in E_0} |D_e|$ ,  $\Delta_1 = \sum_{e \in E_1} |D_e|$  and  $\Delta = \Delta_0 + \Delta_1$  (recall that in this case,  $D_e \not\subseteq K'$  for all  $e \in E$ ). Here again we suppose  $\Delta > 0$ . Let us notice that for all  $u \in \text{tr}(A)$ ,  $\mu(u) \geq T$ . On the other hand, for all  $e \in E$  such that there exists  $v \in D_e$ , we have  $\mu(v) \leq T$  (remark that if  $e \in E_1$ , then  $D_e \subseteq Y_e$ , and if  $e \in E_0$ , then  $v$  is not adjacent to  $\text{tr}(A)$  by definition of  $E_0$ ). Thus  $(K' \setminus \{u\}) \cup \{v\}$  is a safe replacement. Since before this replacement we had

$tr(A) = T + \Delta$ , it is clear that we can repeat this replacement (*i.e.*  $K' \setminus \{u\} \cup \{v\}$  where  $u \in tr(A)$  and  $v \in D_e$  for some  $e \in E$ )  $\Delta$  times safely. At this point, the updated value of  $\Delta$  is 0, *i.e.*  $D_e = \emptyset$  for all  $e \in E$ . By Lemma 5, we must have a clique of size  $k$  in  $G$ .

#### Case A2 and B1

If there exists  $e \in E_0$  such that there exists  $u \in D_e$ , then  $\mu(u) < T$ . If such a vertex exists, then either  $|tr(A)| > T$  or there exists  $e' \in E_1$  such that there exists  $v \in D_{e'}$ . In the first case for all  $x \in tr(A)$  we have  $\mu(x) \geq T$ , and  $(K' \setminus \{x\}) \cup \{u\}$  is a safe replacement. In the second case we have  $\mu(v) > T$  and here again  $(K' \setminus \{v\}) \cup \{u\}$  is a safe replacement. After these replacements we must have  $D_e = \emptyset$  for all  $e \in E_0$ , and we can apply the same arguments as for case A1 and B1.

#### Case A1 and B2

If there exists  $e \in E_1$  such that there exists  $u \in D_e$ , then  $\mu(u) < T$ . If such a vertex exists, then either  $|tr(A)| > T$  or there exists  $e' \in E_0$  such that there exists  $v \in D_{e'}$ . In the first case for all  $x \in tr(A)$  we have  $\mu(x) \geq T$ , and  $(K' \setminus \{x\}) \cup \{u\}$  is a safe replacement. In the second case we have  $\mu(v) > T$  and here again  $(K' \setminus \{v\}) \cup \{u\}$  is a safe replacement. After these replacements we must have  $D_e = \emptyset$  for all  $e \in E_1$ , and we can apply the same arguments as for case A1 and B1. □

## References

1. N. Apollonio. Private communication, 2012.
2. N. Apollonio and A. Sebő. Minconvex factors of prescribed size in graphs. *SIAM Journal of Discrete Mathematics*, 23(3):1297–1310, 2009.
3. J. Backer and J.M. Keil. Constant factor approximation algorithms for the densest  $k$ -subgraph problem on proper interval graphs and bipartite permutation graphs. *Information Processing Letters*, 110(16):635–638, 2010.
4. H. Broersma, P. A. Golovach, and V. Patel. Tight complexity bounds for FPT subgraph problems parameterized by clique-width. In *Proceedings of the 6th international conference on Parameterized and Exact Computation, IPEC'11*, pages 207–218, Berlin, Heidelberg, 2012. Springer-Verlag.
5. D.G. Corneil and Y. Perl. Clustering and domination in perfect graphs. *Discrete Applied Mathematics*, 9(1):27 – 39, 1984.
6. J. Flum and M. Grohe. *Parameterized Complexity Theory*. Springer, 2006.
7. M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.
8. G. Joret and A. Vetta. Reducing the rank of a matroid. *CoRR*, abs/1211.4853, 2012.
9. M. Liazi, I. Milis, F. Pascual, and V. Zissimopoulos. The densest  $k$ -subgraph problem on clique graphs. *Journal of Combinatorial Optimization*, 14(4):465–474, 2007.
10. M. Liazi, I. Milis, and V. Zissimopoulos. A constant approximation algorithm for the densest  $k$ -subgraph problem on chordal graphs. *Information Processing Letters*, 108(1):29–32, 2008.
11. T. Nonner. PTAS for densest  $k$ -subgraph in interval graphs. In *Proceedings of the 12th international conference on Algorithms and Data Structures*, pages 631–641. Springer, 2011.