NP-hardness of k-sparsest subgraph in Chordal Graphs

Rémi Watrigant, Marin Bougeret, Rodolphe Giroudeau

$\boldsymbol{NP}$-hardness of $\boldsymbol{k}$-sparsest subgraph in Chordal Graphs*

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 k-sparsest 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 k-sparsest subgraph remains NP-hard in this graph class.

1 Introduction and Preliminaries

In this report we study the following decision problem:
k-sparsest 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), k-sparsest subgraph is NP-hard [7] as well as $W[1]$-hard [6] and $O(n^{1-\epsilon})$-inapproximable (unless $P = NP$) [12] in general graphs.
Its maximization version, namely the k-densest subgraph or the k-cluster problem, has been extensively studied in the last three decades: it remains NP-hard in chordal graphs, bipartite graphs and comparability graphs, whereas it is polynomial-time solvable in trees, cographs, bounded treewidth and in split graphs [5]. Notice that several exact or approximation algorithms have been designed for this problem [3,4,8,9,10,11]. In addition, it appears that some interesting open problems exists around k-densest subgraph: in particular its complexity status (polynomial *vs* NP-hardness) in interval (and even proper interval) as well as its approximability status (APX or not) in chordal graphs are unknown. Unfortunately, most of these results seem useless for k-sparsest subgraph, as we apparently need to complement the input graph to apply them. Nevertheless we can deduce that k-sparsest 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, remains NP-hard in line graphs [1], and seems to remain NP-hard in bipartite graphs [2].
In this report we study the complexity status of k-sparsest subgraph in chordal graphs. Whereas the independent set problem is polynomial-time solvable in perfect graphs (and thus in chordal graphs), we show that k-sparsest subgraph remain 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:

* This work has been funded by grant ANR **2010** BLAN **021902**

- A graph is chordal if every cycle of length at least 4 has a chord.
- A vertex v of G is called simplicial if $N(v)$ is a clique. The ordering $v_1, ..., v_n$ of the vertices of G is a simplicial elimination scheme if for all i, v_i is simplicial in $G[v_{i+1}, ..., v_n]$. A graph is chordal if it has a simplicial elimination scheme.

2 The Main Result

Theorem 1. k-SPARSEST 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 $|V| = n$, $V = \{v_1, ..., v_n\}$, $|E| = m$ 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 define $A = \{a_i^j : i, j \in \{1, ..., n\}\}$. Thus, A is a clique of size n^2. Moreover, for all $j \in \{1, ..., n\}$, we note $A_j = \{a_1^j, ..., a_n^j\}$.
- For all $e \in E$, we construct a graph with F_e as vertex set, composed of three sets of T vertices, namely $X_e = \{x_1^e, ..., x_T^e\}$, $Y_e = \{y_1^e, ..., y_T^e\}$ and $Z_e = \{z_1^e, ..., z_T^e\}$. 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 $\{y_2^e, ..., y_T^e\}$ (thus, y_1^e is not connected to vertices of Y^e). Then, for all $j \in \{1, ..., T\}$, x_j^e is connected to y_j^e, and y_j^e is connected to all vertices of Z^e. Finally, we add to F_e a pending vertex α^e connected to y_1^e. An example of such a gadget is represented in Figure 1.
- We define $F = \bigcup_{e \in E} F_e$.
- For all $e = \{v_p, v_q\} \in E$, all vertices of Z^e are connected to $\{a_p^j : j \in \{1, ..., n\}\}$ and $\{a_q^j : j \in \{1, ..., n\}\}$.
- We define $k' = m(2T+1) + T$ and $C' = m\binom{T}{2} + \binom{T}{2} + (m - \binom{k}{2})$.

The above construction can clearly be performed in polynomial time. The following lemma proves that the constructed graph is chordal:

Lemma 1. G' *is chordal.*

Proof. We have the following simplicial elimination scheme:
- For all $e \in E$, we can remove X^e since for all $j \in \{1, ..., T\}$, x_j^e is only connected to y_j^e. Similarly, we can remove α_e since it is connected to y_1^e only.
- for all $e \in E$, we can remove Y^e. Indeed, we remove y_1^e first, since it is connected to Z^e only which induces a clique, and then successively for $j = 2, ..., T$ we can remove y_j^e, since it is connected to $\{y_{j+1}^e, ..., y_T^e\}$ and Z^e which form a clique.
- for all $e \in E$ and $j = 1, ..., T$ successively, we can remove z_j^e since it is connected to $\{z_{j+1}^e, ..., z_T^e\}$ and some vertices of A which induce a clique.
- it now remains A which is a clique and can 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.

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

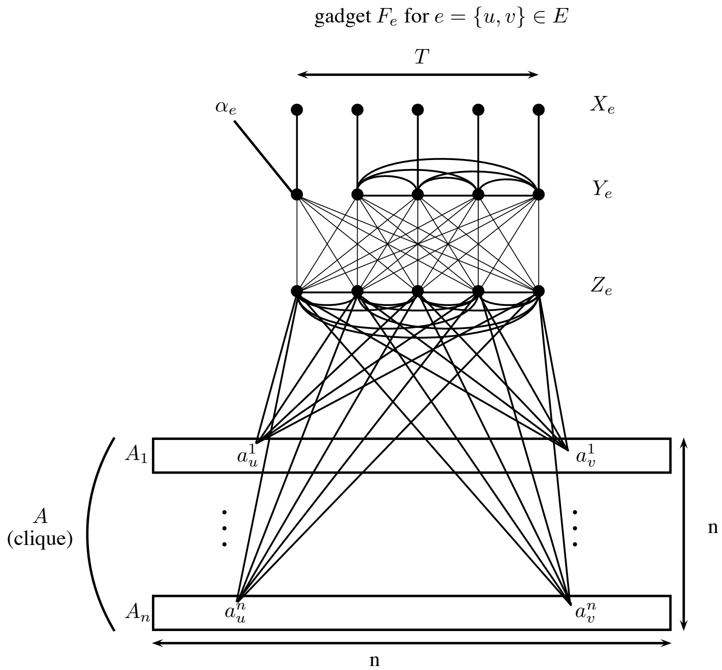

Figure1: Example of a gadget F_e (with $T = 5$) and its relations to A.

- For all $i \in \{(k+1), ..., n\}$ and all $j = \{1, ..., n\}$, we add a_i^j to K'.
- For all $e \in E$, we add α_e and all vertices of X_e in 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' vertices inducing exactly C' 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 $(2T+1)$ vertices, which induce $\binom{T}{2}$ edges if $e \in E_0$, and $(\binom{T}{2}+1)$ edges if $e \in E_1$ (because of the adjacency between α_e and y_1^e). Since $|E_0| = \binom{k}{2}$ (and thus $|E_1| = m - \binom{k}{2}$), we have the desired number of edges.

$\Leftarrow$ Suppose now that K' is a set of k' vertices of G' which induce C' edges or less. We re-define the sets E_0 and E_1 as follows: $E_0 = \{\{v_p, v_q\} \in E$ such that for all $j \in \{1, ..., n\}$ we have $a_p^j \notin K'$ and $a_q^j \notin K'\}$, and $E_1 = E \backslash E_0$. 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 neighbours of v belonging to K'. The proof consists in replacing some vertices of K' by other vertices not in K' without increasing the number of induced edges. We call such a replacement a *safe modification* or a *safe replacement*. Let $u \in K'$ and $v \in V' \backslash K'$. It is clear that $K' \backslash \{u\} \cup \{v\}$ 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 $\{u, v\} \in E'$. Remark that performing replacements on vertices not in A does not change the definition of sets E_0 nor E_1.

We first prove that K' can be safely modified such that for all $e \in E$, $\alpha_e \in K'$ and $X_e \subset K'$. Let $S = \bigcup_{e \in E} (\{\alpha_e\} \cup X_e)$. Since we have $k' > |S|$, there always exists $u \in K' \backslash S$. Suppose that there exists $e \in E$ such that $\alpha_e \notin K'$. If $y_1^e \notin K'$, then we have $\mu(\alpha_e) = 0$ and we can thus safely replace any other vertex of $K' \backslash S$ by α_e. Now, if $y_1^e \in K'$, then $\mu(\alpha_e) = 1 \leq 1 = \mu(y_1^e)$, and $K' \backslash \{y_1^e\} \cup \{\alpha_e\}$ is a safe replacement. Using the same arguments, K' can be safely modified such that $X_e \subset K'$.

In the following, we suppose that for all $e \in E$, $\alpha_e \in K'$ and $X_e \subset K'$.

Restructuration of gadgets F_e, $e \in E_0$.
We first restructure each gadget separately: for all $e \in E_0$, if $tr(Y_e) \neq \emptyset$ and $tr(Z_e) \neq Z_e$, let $j_0 = \max\{j \in \{1, ..., n\} : y_j^e \in tr(Y_e)\}$ and let j_1 be such that $z_{j_1}^e \notin tr(Z_e)$. 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 $\mu(z_{j_1}^e) \leq y + z$ (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 $y_1^e \notin K'$). Hence $\mu(z_{j_1}^e) \leq \mu(y_{j_0}^e)$ and $K' \backslash \{y_{j_0}^e\} \cup \{z_{j_1}^e\}$ 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 $\mu(y_1^e) = z + 2$ where $z = |N(y_1^e) \cap tr(Z_e)|$, and $\mu(z_{j_1}^e) = z$. Here again $K' \backslash \{y_1^e\} \cup \{z_{j_1}^e\}$ 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, $K' \backslash \{y_{j_0}^a\} \cup \{z_{j_1}^b\}$ is a safe replacement.

After all these replacements, one of the two following cases must happen:

case A1: For all $e \in E_0$, $\alpha_e \in K'$, $X_e \subset K'$, $Z_e \subset K'$ and there exists $D_e \subseteq Y_e$ such that $D_e \subset K'$. Notice that we may have $D_e = \emptyset$ or $D_e = Y_e$ for some e.

case A2: For all $e \in E_0$, $\alpha_e \in K'$, $X_e \subset K'$, $Y_e \not\subset K'$ and there exists $D_e \subseteq Z_e$ such that $D_e \not\subset K'$. Notice that we may have $D_e = \emptyset$ or $D_e = Z_e$ for some e.

Notice that if $D_e = \emptyset$ for all $e \in E_0$, then cases A1 and A2 collapse. These cases are depicted in Figure 2.

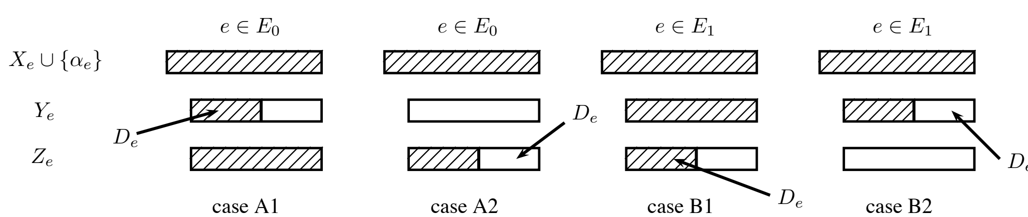

Figure2: Schema of different cases. Shaded rectangles represent part of K'.

Restructuration of gadgets F_e, $e \in E_1$.
Similarly to the previous paragraph, we first restructure each gadget separately: for all $e \in E_1$, if $tr(Z_e) \neq \emptyset$ and $tr(Y_e) \neq Y_e$, let $j_0 = \max\{j \in \{1, ..., n\} : y_j^e \notin K'\}$ and let j_1 be such that $z_{j_1}^e \in tr(Z_e)$ (hence we have $j_0 > 1$). Recall that by definition of E_1, there exists $i, j \in \{1, ..., n\}$

such that $z^e_{j_1}$ is adjacent to a^j_i. We have $\mu(z^e_{j_1}) = y + z + 1$, where $y = |N(z^e_{j_1}) \cap Y_e|$ and $z = |N(z^e_{j_1}) \cap Z_e|$. On the other side, since y_{j_0} is connected to $(y-1)$ vertices of Y^e and to one vertex of X^e (namely $x^e_{j_0}$), we have $\mu(y^e_{j_0}) \leq z + y$. Thus $K' \backslash \{z_{j_1}\} \cup \{y_{j_1}\}$ 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 $y^b_{j_0} \notin tr(Y_b)$ and let j_1 be such that $z^a_{j_1} \in tr(Z_a)$. We have $\mu(z^a_{j_1}) \geq T + 1$ and $\mu(y^b_{j_0}) \leq T - 1$. Thus $K' \backslash \{z_{j_1}\} \cup \{y_{j_1}\}$ is a safe replacement.

After all these replacements, one of the two following cases must happen:

case B1: For all $e \in E_1$, $\alpha_e \in K'$, $X_e \subset K'$, $Y_e \subset K'$ and there exists $D_e \subseteq Z_e$ such that $D_e \subset K'$. Notice that we may have $D_e = \emptyset$ or $D_e = Z_e$ for some e.

case B2: For all $e \in E_1$, $\alpha_e \in K'$, $X_e \subset K'$, $Z_e \not\subset K'$, and there exists $D_e \subseteq Y_e$ such that $D_e \not\subset K'$. Notice that we may have $D_e = \emptyset$ or $D_e = Y_e$ for some e.

Notice that if $D_e = \emptyset$ for all $e \in E_1$, then cases B1 and B2 collapse. These cases are depicted in Figure 2. We now prove the following:

Lemma 2. *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 $|tr(F_e)| = 2T + 1$ for all $e \in E$. Thus, $E(tr(A)) = \binom{T}{2}$ and $E(tr(F_e)) = \binom{T}{2} + 1$ if $y^e_1 \in K'$, and $E(tr(F_e)) = \binom{T}{2}$ otherwise. By construction, $y^e_1 \in K'$ if and only if $e \in E_1$. Thus, since $E(K') \leq \binom{T}{2} + m\binom{T}{2} + m - \binom{k}{2}$, 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.

□

Combining the four cases.
We suppose in the following that $D_e \neq \emptyset$ for some $e \in E$. Combining the previous cases, we have four cases to analyse:

- Case $A1$ and $B1$: 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, $D_e \subset K'$ for all $e \in E$). If $\Delta = 0$, then by Lemma 2 G contains a clique of size k. Thus we suppose in the following that $\Delta > 0$. It is clear that $|tr(A)| = T - \Delta$. Moreover:

$$E(K') \geq m\binom{T}{2} + \Delta T + \binom{T}{2} - \binom{\Delta}{2} - \Delta(T - \Delta) + \Delta_1 + |E_1| \qquad (1)$$

Indeed, for all $e \in E$, $tr(F_e)$ contains at least $(2T + D_e + 1)$ vertices, and thus $\binom{T}{2} + |D_e|T$ edges. In addition, for all $e \in E_1$ we have $Y_e \subset K'$, and in particular y^e_1 which adds another edge (and explains the term $|E_1|$). Then, $|tr(A)| = T - \Delta$, which implies that $tr(A)$ induces $\binom{T}{2} - \binom{\Delta}{2} - \Delta(T - \Delta)$ edges. Finally, by definition of E_1, for all $e \in E_1$ and all $j \in \{1, ..., T\}$, z^e_j must be adjacent to some vertex of $tr(A)$, which adds at least Δ_1 edges. Hence,

$$\begin{aligned} E(K') - C' &\geq \Delta T - \binom{\Delta}{2} - \Delta(T - \Delta) + \Delta_1 + |E_1| - m + \binom{k}{2} \\ &= \frac{1}{2}\Delta(\Delta + 1) + \Delta_1 + |E_1| + \binom{k}{2} - m \\ &= \frac{1}{2}\Delta(\Delta + 1) + \Delta_1 + \binom{k}{2} - |E_0| \quad (\text{since } |E_0| + |E_1| = m) \end{aligned}$$

Since K' is supposed to be a set of k' vertices inducing at most C' edges, we must have $E(K') - C' \leq 0$, *i.e.* $|E_0| \geq \frac{1}{2}\Delta(\Delta+1) + \Delta_1 + \binom{k}{2}$. Let $\overline{tr(A)} = A \backslash tr(A)$. It is clear that $|\overline{tr(A)}| = kn + \Delta$.
Recall that for all $e = \{v_p, v_q\} \in E_0$ we have for all $j \in \{1, ..., n\}$ $a_p^j, a_q^j \notin K'$. Thus, if there exists $i_0 \in \{1, ...n\}$ and $j_0 \in \{1, ..., n\}$ such that $a_i^{j_0} \in \overline{tr(A)}$, then we must have $a_{i_0}^j \in \overline{tr(A)}$ for all $j \in \{1, ..., n\}$. Thus the number of vertices inducing all edges of E_0 is at most $\lfloor \frac{nk+\Delta}{n} \rfloor = k + \lfloor \frac{\Delta}{n} \rfloor$, *i.e.* there exists at most $(k + \lfloor \frac{\Delta}{n} \rfloor)$ vertices in G which induce at least $(\frac{1}{2}\Delta(\Delta+1) + \Delta_1 + \binom{k}{2})$ edges. If $\Delta < n$, then it means that k vertices induce strictly more than $\binom{k}{2}$ edges, which is impossible. If $\Delta \geq n$, then $|E_0| \geq \frac{n(n+1)}{2} + \binom{k}{2} > m$ which is also impossible. Thus it implies that $E(K') - k' > 0$, and K' must induce more than C' edges which contradicts the hypothesis and implies that this case cannot happen.

- 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, $D_e \not\subset K'$ for all $e \in E$). Here again we suppose $\Delta > 0$. Let us notice that for all $u \in tr(A)$, $\mu(u) \geq T$. 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 $tr(A)$ by definition of E_0). Thus $K' \backslash \{u\} \cup \{v\}$ 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.* $K' \backslash \{u\} \cup \{v\}$ where $u \in tr(A)$ and $v \in D_e$ for some $e \in E$) Δ times safely. At this point, the updated value of Δ is 0, *i.e.* $D_e = \emptyset$ for all $e \in E$. By Lemma 2, we must have a clique of size k in G.

- Case $A2$ and $B1$: 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, $D_e \subset K'$ for all $e \in E_1$ and $D_e \not\subset K'$ for all $e \in E_0$). If one can find $e_0 \in E_0$ and $e_1 \in E_1$ such that there exists $u \in D_{e_1}$ and $v \in D_{e_0}$, then one can observe that $\mu(u) \geq T$ and $\mu(v) \leq T$ (notice that u is adjacent to every vertex of $Y_{e_1} \subset K'$ and that by definition of E_0, v is not adjacent to any vertex of $tr(A)$). Thus, $K' \backslash \{u\} \cup \{v\}$ is a safe replacement, and this replacement can be made $\min\{\Delta_0, \Delta_1\}$ times. If before replacements we had $\Delta_0 = \Delta_1$, then we must now have $D_e = \emptyset$ for all $e \in E$, and by Lemma 2 G contains a clique of size k. Thus we suppose that we had $\Delta_0 \neq \Delta_1$. Depending on the sign of $\Delta_0 - \Delta_1$, we have two sub-cases, depicted in Figure 3:
  - If before replacements $\Delta_0 - \Delta_1 > 0$, then we now have $D_e = \emptyset$ for all $e \in E_1$, and there exists $e \in E_0$ such that $D_e \neq \emptyset$ (and $\Delta_0 = \sum_{e \in E_0} |D_e|$). Let us count the number of edge in such a situation:

$$
\begin{aligned}
E(K') &\geq \overbrace{m\binom{T}{2} + |E_1| - \binom{\Delta_0}{2} - \Delta_0(T - \Delta_0)}^{\text{edges in } \bigcup_{e \in E} tr(F_e)} + \overbrace{\binom{T}{2} + \binom{\Delta_0}{2} + \Delta_0 T}^{\text{edges in } tr(A)} \\
&= m\binom{T}{2} + \binom{T}{2} + \Delta_0^2 + m - |E_0|
\end{aligned}
$$

  Hence,

$$
E(K') - C' \geq \binom{k}{2} + \Delta_0^2 - |E_0|
$$

  Since K' is supposed to be a set of K' vertices inducing at most C' edges, we must have $E(K') - C' \leq 0$, *i.e.* that $|E_0| \geq \binom{k}{2} + \Delta_0^2$, which implies that there exists at most $(k + \lfloor \frac{\Delta_0}{n} \rfloor)$ vertices in G inducing at most $(\binom{k}{2} + \Delta_0^2)$ edges. If $\Delta_0 < n$, then it means that k vertices induce strictly more than $\binom{k}{2}$ edges, which is impossible. Thus $\Delta_0 \geq n$ which implies $|E_0| \geq \binom{k}{2} + n^2 > m$ which is also impossible. Thus it implies that $E(K') - C' > 0$ and K' must induce more than C' edges which contracts the hypothesis and implies that this case cannot happen.

- If before replacements $\Delta_0 - \Delta_1 < 0$, then we now have $D_e = \emptyset$ for all $e \in E_0$, and there exists $e \in E_1$ such that $D_e \neq \emptyset$ (and $\Delta_1 = \sum_{e \in E_1} |D_e|$). Let us count the number of edges in such a situation:

$$E(K') \geq \overbrace{m\binom{T}{2} + \Delta_1 T + |E_1|}^{\text{edges in } \bigcup_{e\in E} tr(F_e)} + \overbrace{\binom{T}{2} - \binom{\Delta_1}{2} - \Delta_1(T - \Delta_1)}^{\text{edges in } tr(A)}$$
$$= m\binom{T}{2} + \binom{T}{2} + m - |E_0| + \frac{\Delta_1(\Delta_1+1)}{2}$$

Hence,

$$E(K') - C' \geq \binom{k}{2} + \frac{\Delta_1(\Delta_1+1)}{2} - |E_0|$$

Using the same arguments as in the previous case, we conclude that this case cannot happen either.

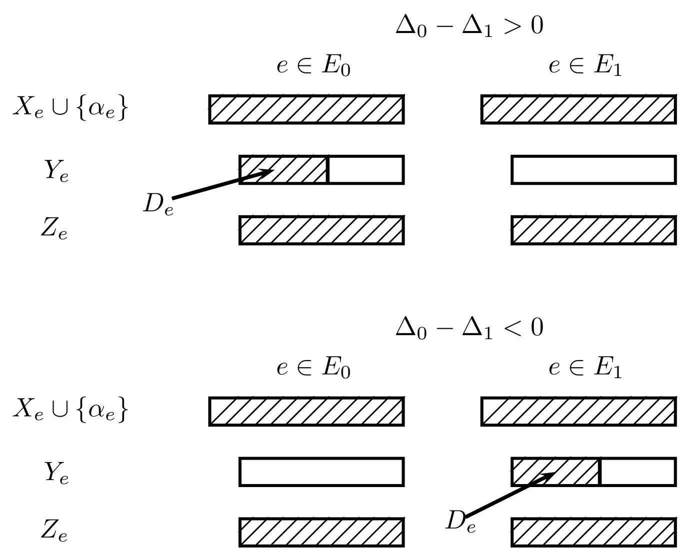

Figure3: Subcases of case A2 and B1.

- Case $A1$ 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, $D_e \not\subset K'$ for all $e \in E_1$ and $D_e \subset K'$ for all $e \in E_0$). If one can find $e_0 \in E_0$ and $e_1 \in E_1$ such that there exists $u \in D_{e_0}$ and $v \in D_{e_1}$, then one can observe that

$\mu(u) \geq T$ and $\mu(v) \leq T$ (notice that u is adjacent to every vertex of $Z_{e_0} \subset K'$, and that $Z_{e_1} \not\subseteq K'$). Thus, $K' \backslash \{u\} \cup \{v\}$ is a safe replacement, and this replacement can be made $\min\{\Delta_0, \Delta_1\}$ times. If before replacements we had $\Delta_0 = \Delta_1$, then we must now have $D_e = \emptyset$ for all $e \in E$, and by Lemma 2 G contains a clique of size k. Thus we suppose that we had $\Delta_0 \neq \Delta_1$. Depending on the sign of $\Delta_0 - \Delta_1$, we have two sub-cases, depicted in Figure 4:

- If before replacements $\Delta_0 - \Delta_1 > 0$, then we now have $D_e = \emptyset$ for all $e \in E_1$, and there exists $e \in E_0$ such that $D_e \neq \emptyset$ (and $\Delta_0 = \sum_{e \in E_0} |D_e|$). Let us count the number of edges in such a situation:

$$E(K') \geq \overbrace{m\binom{T}{2} + \binom{\Delta_0}{2} + \Delta_0(T+1) + |E_1|}^{\text{edges in } \bigcup_{e\in E} tr(F_e)} + \overbrace{\binom{T}{2} - \binom{\Delta_0}{2} - \Delta_0(T - \Delta_0)}^{\text{edges in } tr(A)}$$
$$= m\binom{T}{2} + \binom{T}{2} + \Delta_0^2 + \Delta_0 + m - |E_1|$$

  Hence,

$$E(K') - C' \geq \binom{k}{2} + \Delta_0^2 + \Delta_0 - |E_0|$$

  As previously, $E(K') - C' \leq 0$ would imply that there exists in G at most $(k + \lfloor \frac{\Delta_0}{n})$ vertices inducing at least $\binom{k}{2} + \Delta_0^2 + \Delta_0$ edges. If $\Delta_0 < n$, then we have k vertices inducing strictly more than $\binom{k}{2}$ edges, which is impossible. If $\Delta_0 \geq n$, then $|E_0| \geq \binom{k}{2} + n^2 > m$ which is also impossible. Thus we must have $E(K') - C' > 0$ which is impossible, as K' is supposed to induce at most C' edges. Thus, this case cannot happen.
- If before replacements $\Delta_0 - \Delta_1 < 0$, then we now have $D_e = \emptyset$ for all $e \in E_0$, and there exists $e \in E_1$ such that $D_e \neq \emptyset$ (and $\Delta_1 = \sum_{e \in E_1} |D_e|$). Let us notice that for all $u \in tr(A)$, $\mu(u) \leq T$. On the other hand, for all $e \in E_1$ such that there exists $v \in D_e$, we have $\mu(v) \leq T$ (notice that in this case $Z_e \not\subset K'$). Thus, $K' \backslash \{u\} \cup \{v\}$ is a safe replacement. Since before this replacement we had $tr(A) = T + \Delta_1$, it is clear that we can repeat this replacement (*i.e.* $K' \backslash \{u\} \cup \{v\}$ where $u \in tr(A)$ and $v \in D_e$ for some $e \in E_1$) Δ_1 times safely. At this point, the updated value of Δ_1 is 0, *i.e.* $D_e = \emptyset$ for all $e \in E$. By Lemma 2, G must contain a clique of size k.

□

References

1. N. Apollonio and A. Sebő. Minconvex factors of prescribed size in graphs. SIAM Journal of Discrete Mathematics, 23(3):1297–1310, 2009.
2. Nicola Apollonio. Private communication, 2012.
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. D. Chen, R. Fleischer, and J. Li. Densest k-subgraph approximation on intersection graphs. In Proceedings of the 8th international conference on Approximation and online algorithms, pages 83–93. Springer, 2011.
5. D.G. Corneil and Y. Perl. Clustering and domination in perfect graphs. Discrete Applied Mathematics, 9(1):27 – 39, 1984.

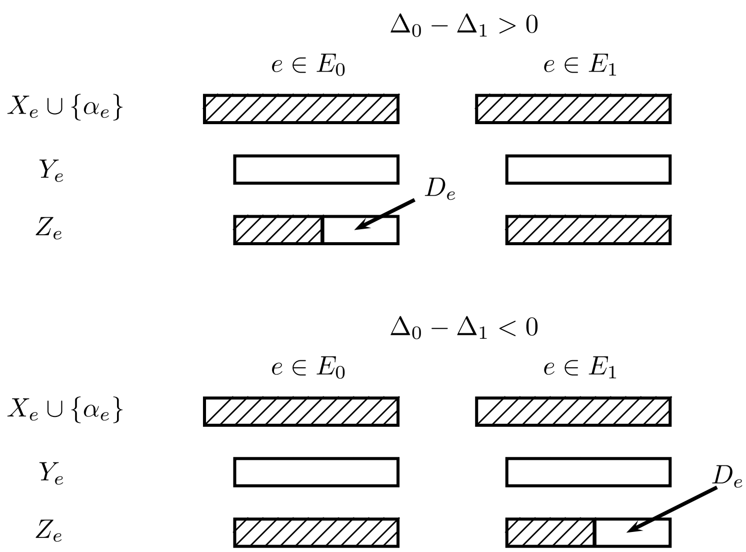

Figure4: Subcases of case A1 and B2.

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. 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.
9. M. Liazi, I. Milis, and V. Zissimopoulos. Polynomial variants of the densest/heaviest k-subgraph problem. In Proceedings of the 20th British Combinatorial Conference, Durham, 2005.
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.
12. L. Trevisan. Inapproximability of combinatorial optimization problems. Electronic Colloquium on Computational Complexity (ECCC), (065), 2004.