



**HAL**  
open science

# Polynomial-time Recognition of 4-Steiner Powers

Guillaume Ducoffe

► **To cite this version:**

Guillaume Ducoffe. Polynomial-time Recognition of 4-Steiner Powers. [Research Report] University of Bucharest, Faculty of Mathematics and Computer Science; National Institute for Research and Development in Informatics, Romania. 2018. hal-01887266v2

**HAL Id: hal-01887266**

**<https://hal.science/hal-01887266v2>**

Submitted on 24 Dec 2018

**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.

# Polynomial-time Recognition of 4-Steiner Powers

Guillaume Ducoffe<sup>1,2,3</sup>

<sup>1</sup>National Institute for Research and Development in Informatics, Romania

<sup>2</sup>The Research Institute of the University of Bucharest ICUB, Romania

<sup>3</sup>University of Bucharest, Faculty of Mathematics and Computer Science, Romania

## Abstract

The  $k^{\text{th}}$ -power of a given graph  $G = (V, E)$  is obtained from  $G$  by adding an edge between every two distinct vertices at a distance  $\leq k$  in  $G$ . We call  $G$  a  $k$ -Steiner power if it is an induced subgraph of the  $k^{\text{th}}$ -power of some tree. Our main contribution is a polynomial-time recognition algorithm of 4-Steiner powers, thereby extending the decade-year-old results of (Lin, Kearney and Jiang, *ISAAC'00*) for  $k = 1, 2$  and (Chang and Ko, *WG'07*) for  $k = 3$ .

A graph  $G$  is termed  $k$ -leaf power if there is some tree  $T$  such that: all vertices in  $V(G)$  are leaf-nodes of  $T$ , and  $G$  is an induced subgraph of the  $k^{\text{th}}$ -power of  $T$ . As a byproduct of our main result, we give the first known polynomial-time recognition algorithm for 6-leaf powers.

## 1 Introduction

A basic problem in computational biology is, given some set of species and a dissimilarity measure in order to compare them, find a *phylogenetic tree* that explains their respective evolution. Namely, such a rooted tree starts from a common ancestor and branches every time there is a separation between at least two of the species we consider. In the end, the leaves of the phylogenetic tree should exactly represent our given set of species. This problem was brought to Graph theory under several disguises but, unfortunately, several of these formulations are NP-hard to solve [BFW92, Ste92]. We here study a related problem whose complexity status remains open. Specifically, a common assumption in the literature is that our dissimilarity measure can only tell us whether the separation between two given species has occurred quite recently. Let  $G = (V, E)$  be a graph whose vertices are the species we consider and such that an edge represents two species with a quite “close” common ancestor according to the dissimilarity measure. Given some fixed  $k \geq 1$ , we ask whether there exists some tree  $T$  whose leaf-nodes are exactly  $V$  and such that there is an edge  $uv$  in  $E$  if and only if the two corresponding nodes in  $T$  are at a distance  $\leq k$ . This is called the  $k$ -LEAF POWER problem [NRT02].

The structural properties of  $k$ -leaf powers (*i.e.*, graphs for which a tree as above exists) have been intensively studied [BPP10, BH08, BHMW10, BL06, BLS08, BLR09, WB09, CFM11, DGHN06, DGHN08, DGN05, KLY06, KKLY10, Laf17, NR16, Rau06]. From the algorithmic point of view,  $k$ -leaf powers are a subclass of bounded *clique-width* graphs, and many NP-hard problems can be solved efficiently for these graphs [FMR<sup>+</sup>08, GW07]. However, the computational complexity of recognizing  $k$ -leaf powers is an open problem. Very recently, parameterized algorithms were proposed for every fixed  $k$  on the graphs with degeneracy at most  $d$ , where the parameter is

$k + d$  [EH18]. Without this additional restriction on the degeneracy of the graphs, polynomial-time recognition algorithms are known only for  $k \leq 5$  [BL06, BLS08, CK07]. It is noteworthy that every algorithmic improvement for this problem has been incredibly hard to generalize to larger values of  $k$ . We contribute to this frustrating chain of improvements by providing the first known polynomial-time recognition algorithm for 6-leaf powers.

**Theorem A.** *There is a polynomial-time algorithm that given a graph  $G = (V, E)$ , correctly decides whether  $G$  is a 6-leaf power (and if so, outputs a corresponding tree  $T$ ).*

Several variations of  $k$ -leaf powers were introduced in the literature [BLR10, BW10, CK07, HT10, JKL00]. In this work, we consider  $k$ -Steiner powers: a natural relaxation of  $k$ -leaf powers where the vertices in the graph may also be internal nodes in the tree  $T$ . Interestingly, for every  $k \geq 3$ , the notions of  $k$ -leaf powers and  $(k - 2)$ -Steiner powers are equivalent for a *twin-free* graph. The latter implies a linear-time reduction from  $k$ -LEAF POWER to  $(k - 2)$ -STEINER POWER [BLS08]. Furthermore, there exist polynomial-time recognition algorithms for  $k$ -Steiner powers, for every  $k \leq 3$  [CK07, JKL00]. As our main contribution in the paper we obtain the first improvement on the recognition of  $k$ -Steiner powers in a decade. Specifically we prove that there is a polynomial-time recognition algorithm for the 4-Steiner powers.

**Theorem B.** *There is a polynomial-time algorithm that given a graph  $G = (V, E)$ , correctly decides whether  $G$  is a 4-Steiner power (and if so, outputs a corresponding tree  $T$ ).*

Note that Theorem A follows from the combination of Theorem B with the aforementioned reduction from  $k$ -LEAF POWER to  $(k - 2)$ -STEINER POWER [BLS08]. *Hence, the remaining of this paper is devoted to a polynomial-time solution for 4-STEINER POWER.* We think that our general approach (presented next) could be generalized to larger values of  $k$ . However, this would first require to strengthen the structure theorems we use in this paper (and probably to find less intricate proofs for some of our intermediate statements).

Before we can introduce our technical contributions in this paper, we must explain in Sec. 1.1 the dynamic programming approach behind the  $k$ -Steiner power recognition algorithms, for  $k \leq 3$ , and why it is so hard to apply this approach to 4-STEINER POWER. Doing so, we identify Obstacles O1 and O2 as the two main issues to fix in order to prove Theorem B. Then in Sec. 1.2, we briefly summarize our proposed solutions for Obstacles O1 and O2 while presenting the organization of the technical sections of this paper.

## 1.1 The difficulties of a dynamic programming approach

We refer to Sec. 2 for any undefined graph-theoretic terminology in this introduction. In what follows we give a high-level overview of our approach, that we compare to prior work on  $k$ -STEINER POWER and  $k$ -LEAF POWER. As our starting point we restrict our study to *chordal graphs* and *strongly chordal graphs*, that are two well-known classes in algorithmic graph theory of which  $k$ -Steiner powers form a particular subclass [ABNT16]. Doing so, we can use various properties of these classes of graphs such as: the existence of a tree-like representation of chordal graphs, that is called a *clique-tree* [BP93] and is commonly used in the design of dynamic programming algorithms on this class of graphs; and an auxiliary data structure which is called “*clique arrangement*” and is polynomial-time computable on strongly chordal graphs [NR16]. Roughly, this clique arrangement encodes all possible intersections of a subset of maximal cliques in a graph. It is worth noticing

that clique arrangements were introduced in the same paper as leaf powers, under the different name of “clique graph” [NRT02].

Our first result is that every maximal clique, minimal separator and, more generally, any intersection of maximal cliques in a  $k$ -Steiner power must induce a subtree with very specific properties – detailed next – of the tree  $T$  we aim at computing. This result – that extends to any  $k$  the structural results that were presented in [CK07] for  $k \leq 3$  – is a prerequisite for the design of a dynamic programming algorithm on a clique-tree. Unfortunately, as the value of  $k$  increases it becomes more and more difficult to derive from such structural results a polynomial-time recognition algorithm for  $k$ -Steiner powers. Our proposed solutions for  $k = 4$  are quite different from those used in the previous works on  $k$ -Steiner powers [CK07, JKL00] which results in an embarrassingly long and intricate proof.

To give a flavour of the difficulties we met, we consider the following common situation in a dynamic programming algorithm on chordal graphs. Given a graph  $G = (V, E)$ , let  $S$  be a minimal separator of  $G$  and  $C$  be a full component of  $G \setminus S$  (*i.e.*, such that every vertex in  $S$  has a neighbour in  $C$ ). If  $G$  is a  $k$ -Steiner power then, so must be the induced subgraph  $G[C \cup S]$ . We want to store the partial solutions obtained for  $G[C \cup S]$ , as at least one of them should be extendible to all of  $G$  (otherwise,  $G$  is not a  $k$ -Steiner power). However, for doing so efficiently we must overcome the following two obstacles:

- There may be exponentially many partial solutions already when  $G[C \cup S]$  is a complete subgraph and  $k \geq 3$ . Therefore, we cannot afford to store all possible solutions explicitly. Nevertheless it seems at the minimum we need to keep the part of these solutions which contains  $S$ : in order to be able to check later whether the solutions found for  $G[C \cup S]$  can be extended to all of  $G$ . We will prove in this paper that such a part  $T \langle S \rangle$  of the partial solutions is a subtree of diameter at most  $k - 1$ , and so, there may be exponentially many possibilities to store whenever  $k \geq 4$ .

**Obstacle O1.** *Decrease the number of possibilities to store for  $T \langle S \rangle$  to a polynomial.*

- Furthermore, since there can be no edge between  $C$  and  $V \setminus (C \cup S)$ , the tree  $T$  that we want to compute for  $G$  must satisfy that all vertices in  $C$  stay at a distance  $\geq k + 1$  from all vertices in  $V \setminus (S \cup C)$ . In order to ensure this will be the case, we wish to store a “distance profile”  $(dist_T(r, C))_{r \in V(T \langle S \rangle)}$  in the encoding of all the partial solutions found for  $G[C \cup S]$ . Storing this information would result in a combinatorial explosion of the number of possible encodings, even if there are only a few possibilities for the subtree  $T \langle S \rangle$ . Chang and Ko proposed two nice “heuristic rules” in order to overcome this distance issue for  $k = 3$  [CK07]. Unfortunately, these rules do not easily generalize to larger values of  $k$ .

**Obstacle O2.** *For any fixed  $T \langle S \rangle$ , decrease the number of possibilities to store for the “distance profiles” to a polynomial.*

In order to derive a polynomial-time algorithm for the case  $k = 4$ , we further restrict the structural properties of the “useful” partial solutions we need to store. This is done by carefully analysing the relationships between the structure of these solutions and the intersections between maximal cliques in the graph. Perhaps surprisingly, we need to combine these stronger properties on the partial solutions with several other tricks so as to bound the number of partial solutions that we need to store by a polynomial (*e.g.*, we also impose local properties on the clique-tree we use, and we introduce a new greedy selection procedure based on graph matchings).

## 1.2 Organization of the paper

We give the required graph-theoretic terminology for this paper in Section 2. We emphasize on Section 2.3: where we also provide a more detailed presentation of our algorithm, as a guideline for all the other sections.

Given a  $k$ -Steiner power  $G$ , let us call  $k$ -Steiner root a corresponding tree  $T$ . In Sections 3 and 4 we present new results on the structure of  $k$ -Steiner roots that we use in the analysis of our algorithm. Specifically, we show in Section 3 that any intersection of maximal cliques in a graph  $G$  must induce a particular subtree in any of its  $k$ -Steiner roots  $T$  where no other vertex of  $G$  can be present. Furthermore, the inclusion relationships between these “clique-intersections” in  $G$  are somewhat reflected by the diameter of their corresponding subtrees in  $T$ . An intriguing consequence of our results is that, in any  $k$ -Steiner power, there can be no chain of more than  $k$  minimal separators ordered by inclusion. This slightly generalizes a similar result obtained in [NRT02] for  $k$ -leaf powers.

Then, we partly complete this above picture in Section 4 for the case  $k = 4$ . For every clique-intersection  $X$  in a chordal graph  $G$ , we classify the vertices in  $X$  into two main categories: “free” and “constrained”, that depend on the other clique-intersections these vertices are contained into. Our study shows that “free” vertices cause a combinatorial explosion of the number of partial solutions we should store in a naive dynamic programming algorithm. However, on the positive side we prove that there always exists a “well-structured” 4-Steiner root where such free vertices are leaves with very special properties. This result will be instrumental in ruling out Obstacle O1.

Sections 5, 6, 7 and 8 are devoted to the main steps of the algorithm. We start presenting a constructive proof of a rooted clique-tree with quite constrained properties in Section 5. Roughly we carefully control the ancestor/descendant relationships between the edges that are labelled by different minimal separators of the graph. These technicalities are the cornerstone of our approach in Section 3 in order to bound the number of “distance profiles” which we need to account in our dynamic programming (cf. Obstacle O2).

Then in Section 6, we completely rule out Obstacle O1 — in fact, we solve a more general subproblem. For that, let  $T_G$  be the rooted clique-tree of Section 5. Recall that the maximal cliques and the minimal separators of  $G$  can be mapped to the nodes and edges of  $T_G$ , respectively. We precompute by dynamic programming, for every node and edge in  $T_G$ , a family of all the potential subtrees to which the corresponding clique-intersection of  $G$  could be mapped in some “well-structured” 4-Steiner root of  $G$ . Of particular importance is Section 6.1, where we give a polynomial-time algorithm in order to generate all the candidate subtrees any minimal separator of the graph can induce in its 4-Steiner roots. The result is then easily extended to the maximal cliques that appear as leaves in our clique-tree (Section 6.2). Correctness of these two first parts follows from Sec. 4. Finally, in Section 6.3 we give a more complicated representation of a family of potential subtrees  $T\langle K_i \rangle$  for all the other maximal cliques  $K_i$ . This part is based on a careful analysis of clique-intersections in  $K_i$  and several additional tricks. Roughly, our representation in Sec. 6.1 is composed of partially constructed subtrees and of “problematic” subsets that need to be inserted to these subtrees in order to complete the construction. The exact way these insertions must be done is postponed until the very end of the algorithm (Sec. 8).

Section 7 is devoted to the “distance profiles” of partial solutions and how to overcome Obstacle O2. Specifically, instead of computing partial solutions at each node of the clique-tree and storing their encodings, we rather pre-compute a polynomial-size subset of *imposed* encodings for

each node. Then, the problem becomes to decide whether given such an imposed encoding, there exists a corresponding partial solution. We formalize our approach by introducing an intermediate problem where the goal is to compute a 4-Steiner root with additional constraints on its structure and the distances between some sets of nodes.

Finally, we detail in Section 8 the resolution of our intermediate problem using an ad hoc data structure, thereby completing the presentation of our algorithm. An all new contribution in this part is a greedy procedure, based on MAXIMUM-WEIGHT MATCHING, in order to ensure some distances' constraints are satisfied by the solutions we generate during the algorithm. Interestingly, this procedure is very close in spirit to the implementation of the `alldifferent` constraint in constraint programming [Rég94].

Due to the intricacy of our proofs we gave up optimizing the runtime of our algorithm. We will only provide enough arguments in order to show it is polynomial.

We end up this paper in Section 9 with some ideas for future work.

## 2 Preliminaries

We refer to [BM08] for any undefined graph terminology. All graphs in this study are finite, simple (hence, with neither loops nor multiple edges), unweighted and connected – unless stated otherwise. Given a graph  $G = (V, E)$ , let  $n := |V|$  and  $m := |E|$ . The neighbourhood of a vertex  $v \in V$  is defined as  $N_G(v) := \{u \in V \mid uv \in E\}$ . By extension, we define the neighbourhood of a set  $S \subseteq V$  as  $N_G(S) := (\bigcup_{v \in S} N_G(v)) \setminus S$ . The subgraph induced by any subset  $U \subseteq V$  is denoted by  $G[U]$ .

For every  $u, v \in V$ , we denote by  $dist_G(u, v)$  the minimum length (number of edges) of a  $uv$ -path. The eccentricity of vertex  $v$  is defined as  $ecc_G(v) := \max_{u \in V} dist_G(u, v)$ . The radius and the diameter of  $G$  are defined, respectively, as  $rad(G) := \min_{v \in V} ecc_G(v)$  and  $diam(G) := \max_{v \in V} ecc_G(v)$ . We denote by  $\mathcal{C}(G)$  the center of  $G$ , *a.k.a.* the vertices with minimum eccentricity.

### 2.1 Problems considered

The  $k^{th}$ -power of  $G$ , denoted  $G^k$  has same vertex-set  $V$  as  $G$  and edge-set  $E_k := \{uv \mid 0 < dist_G(u, v) \leq k\}$ . We call  $G$  a  $k$ -Steiner power if there is some tree  $T$  such that  $G$  is an induced subgraph of  $T^k$ . Conversely,  $T$  is called a  $k$ -Steiner root of  $G$ . If in addition,  $G$  has a  $k$ -Steiner root where all vertices in  $V$  are leaves (degree-one nodes) then, we call  $G$  a  $k$ -leaf power.

**Problem 1** ( $k$ -STEINER POWER).

**Input:** A graph  $G = (V, E)$ .

**Output:** Is  $G$  a  $k$ -Steiner power?

**Problem 2** ( $k$ -LEAF POWER).

**Input:** A graph  $G = (V, E)$ .

**Question:** Is  $G$  a  $k$ -leaf power?

**Theorem 2.1** ([BLS08]). *There is a linear-time reduction from  $k$ -LEAF POWER to  $(k-2)$ -STEINER POWER for every  $k \geq 3$ .*

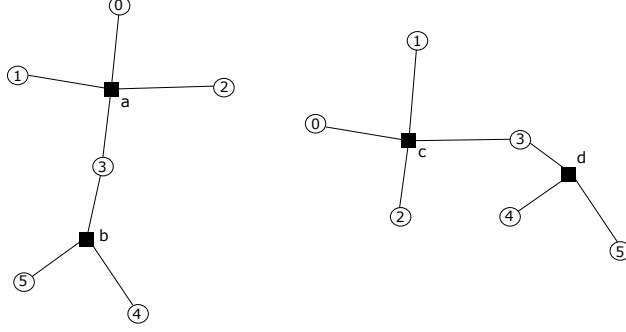


Figure 1: Two Steiner-equivalent trees. Cycles and rectangles represent real and Steiner nodes, respectively.

If  $T$  is any  $k$ -Steiner root of  $G$  then, nodes in  $V(G)$  are called *real*, whereas nodes in  $V(T) \setminus V(G)$  are called *Steiner*. We so define, for any  $S \subseteq V(T)$  (for any subtree  $T' \subseteq T$ , resp.):

$$Real(S) := S \cap V(G) \text{ and } Steiner(S) := S \setminus V(G)$$

(we define  $Real(T') := Real(V(T'))$  and  $Steiner(T') = Steiner(V(T'))$ , resp.).

Note that throughout all this paper we consider two (sub)trees being equivalent if they are equal up to an appropriate identification of their Steiner nodes, namely (see also Fig. 1):

**Definition 2.1.** Given  $G = (V, E)$ , we call any two trees  $T, T'$  *Steiner-equivalent*, denoted  $T \equiv_G T'$ , if and only if  $Real(T) = Real(T') = S$  and there exists an isomorphism  $\iota : V(T) \rightarrow V(T')$  such that  $\iota(v) = v$  for any  $v \in S$ .

Finally, given a node-subset  $X \subseteq V(T)$ ,  $T \langle X \rangle$  is the smallest subtree of  $T$  such that  $X \subseteq V(T \langle X \rangle)$ . Note that for a vertex-subset  $X \subseteq V$ , this is the smallest subtree of  $T$  such that  $X \subseteq Real(T \langle X \rangle)$ . Furthermore we observe  $T[X] \subseteq T \langle X \rangle$ , with equality if and only if  $T[X]$  is connected.

## 2.2 Algorithmic tool-kit: (Strongly) Chordal graphs

Given  $G = (V, E)$ , we call it a *chordal* graph if every induced cycle in  $G$  is a triangle. If in addition, for every cycle of even length in  $G$ , there exists a chord between two vertices at an odd distance ( $> 1$ ) apart from each other in the cycle then,  $G$  is termed *strongly chordal*. Chordal graphs and strongly chordal graphs can be recognized in  $\mathcal{O}(m)$ -time and  $\mathcal{O}(m \log n)$ -time, respectively [PT87, RTL76].

The following property is well-known:

**Theorem 2.2** ([ABNT16]). *For every  $k \geq 1$ , every  $k$ -Steiner power is a strongly chordal graph.*

**Minimal separators and Clique-tree.** Our main algorithmic tool in this paper is a *clique-tree* of  $G$ , defined as a tree  $T_G$  whose nodes are the maximal cliques of  $G$  and such that for every  $v \in V$ , the maximal cliques containing  $v$  induce a subtree of  $T_G$ .

**Theorem 2.3** ([BP93]). *A graph  $G = (V, E)$  is chordal if and only if it has a clique-tree. Moreover if  $G$  is chordal then, we can construct a clique-tree for  $G$  in  $\mathcal{O}(m)$ -time.*

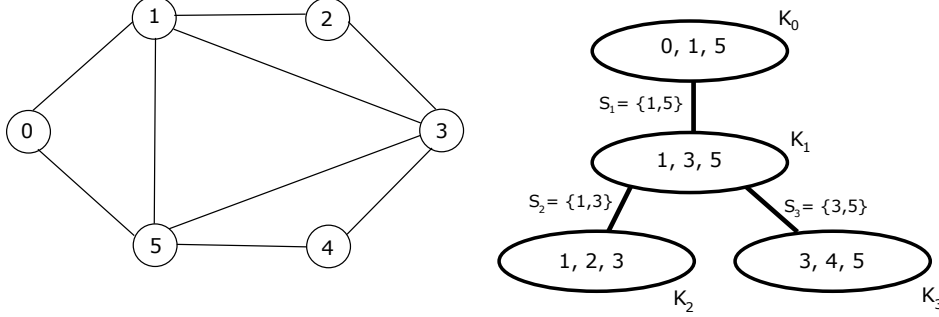


Figure 2: A chordal graph  $G$  (left) and a rooted clique-tree  $T_G$  (right).

An  $uv$ -separator is a subset  $S \subseteq V \setminus \{u, v\}$  such that  $u$  and  $v$  are disconnected in  $G \setminus S$ . If in addition, no strict subset of  $S$  is an  $uv$ -separator then,  $S$  is a *minimal  $uv$ -separator*. A *minimal separator* of  $G$  is a minimal  $uv$ -separator for some  $u, v \in V$ . It is known that any minimal separator in a chordal graph  $G$  is the intersection of two distinct maximal cliques of  $G$ . Specifically, the following stronger relationship holds between minimal separators and clique-trees:

**Theorem 2.4** ([BP93]). *Given  $G = (V, E)$  chordal, any of its clique-trees  $T_G$  satisfies the following properties:*

- For every edge  $K_i K_j \in E(T_G)$ ,  $K_i \cap K_j$  is a minimal separator;
- Conversely, for every minimal separator  $S$  of  $G$ , there exist two maximal cliques  $K_i, K_j$  such that  $K_i K_j \in E(T_G)$  and  $K_i \cap K_j = S$ .

Based on the above theorem, we can define  $E_S(T_G) := \{K_i K_j \in E(T_G) \mid K_i \cap K_j = S\}$ . The cardinality  $|E_S(T_G)|$  of this subset does not depend on  $T_G$  [BP93]. We sometimes say that edges in  $E_S(T_G)$  are *labeled* by  $S$ .

A *rooted* clique-tree of  $G$  is obtained from any clique-tree  $T_G$  by identifying an arbitrary maximal clique  $K_0$  as its root. Let  $(K_q, K_{q-1}, \dots, K_1, K_0)$  be a postordering of  $T_G$  (obtained by depth-first search). For any  $i > 0$ , we define  $K_{p(i)}$  as the father node of  $K_i$ . The common intersection of  $K_i$  with its father node is the minimal separator  $S_i := K_i \cap K_{p(i)}$ . By convention, we set  $S_0 := \emptyset$ . We refer to Fig. 2 for an illustration.

We define  $T_G^i$  as the subtree rooted at  $K_i$ , and let  $G_i$  be the subgraph induced by all the maximal cliques in  $V(T_G^i)$ . In particular, we have  $T_G^0 = T_G$  and  $G_0 = G$ . We also define  $V_i := V(G_i)$  and  $W_i := V_i \setminus S_i$  as shorthands. We will use these above notations for rooted clique-trees throughout the remaining of our paper.

**Clique arrangement.** We introduce a common generalization of both maximal cliques and minimal separators, that will play a key role in our analysis. Specifically, a *clique-intersection* in  $G$  is the intersection of a subset of maximal cliques in  $G$ . The family of all clique-intersections in  $G$  is denoted by  $\mathcal{X}(G)$ . For strongly chordal graphs, it is known [NR15] that every clique-intersection is the intersection of at most two maximal cliques. In particular, a (nonempty) clique-intersection of a given strongly chordal  $G$  is either: a maximal clique; or a minimal separator; or a *weak* minimal separator – *i.e.*, whose removal strictly increases the distance between two vertices that remain in



the graph (see [McK11]). We denote by  $\mathcal{K}(G)$ ,  $\mathcal{S}(G)$  and  $\mathcal{W}(G)$  the subfamilies of all the maximal cliques, minimal separators and weak minimal separators of  $G$ , respectively.

The *clique arrangement* of  $G$  is the inclusion (directed) graph of the clique-intersections of  $G$ . That is, there is a node for every clique-intersection, and there is an arc from  $X$  to  $X'$  if and only if we have  $X \subseteq X'$ .

**Theorem 2.5** ([NR15]). *Given  $G = (V, E)$  strongly chordal, the clique arrangement of  $G$  can be constructed in  $\mathcal{O}(m \log n)$ -time.*

### 2.3 Highlights of the algorithm

The remaining of the paper is devoted to the proof of Theorem B. By Theorem 2.1, this will also imply Theorem A. We start sketching our algorithm below in order to guide the readers throughout the next sections. Its analysis is based on the structure theorems in Sections 3 and 4. Perhaps surprisingly, we need several tricks in order to keep the running time of this algorithm polynomial.

**S.0** (*Initialization.*) Given  $G = (V, E)$ , we check whether  $G$  is strongly chordal. If this is not the case then, by Theorem 2.2,  $G$  cannot be a  $k$ -Steiner power for any  $k \geq 1$ , and we stop. Otherwise by Theorem 2.5 we can compute the clique-arrangement of  $G$  in polynomial time. Throughout all the remaining sections, we implicitly use the fact that we can access in polynomial time to the clique-arrangement of  $G$ . We will also assume in what follows that  $G$  is not a complete graph (otherwise,  $G$  is trivially a  $k$ -Steiner power for any  $k$ , and so we also stop in this case).

**S.1** (*Construction of the clique-tree.*) We construct a clique-tree  $T_G$  of  $G$  that we root in some  $K_0 \in \mathcal{K}(G)$ . This clique tree must satisfy very specific properties of which we postpone the precise statement until Section 5. In order to give the main intuition behind this construction, let us consider an arbitrary maximal clique  $K_i$  that is not the root (*i.e.*,  $i > 0$ ). Roughly, we would like to minimize the number of minimal separators in  $G_i$  to which the vertices in  $S_i := K_i \cap K_{p(i)}$  belong to. Indeed, doing so will help in bounding the number of possible partial solutions that we will need to consider for processing its father node  $K_{p(i)}$ . More specifically, for every descendant  $K_j$  of  $K_i$  in  $T_G$  we would like to impose  $S_j \not\subseteq S_i$  and  $S_i \not\subseteq S_j$ . However, both objectives are conflicting and so, we need to find a trade-off. The technical motivations behind our choices will be further explained in Sections 7 and 8.

**S.2** (*Candidate set generation.*) This step exploits a result of Section 3 which states that, for any 4-Steiner root  $T$  of  $G$  and for any clique-intersection  $X$ , we have  $Real(T\langle X \rangle) = X$ . Equivalently, this result says that in the smallest subtree containing  $X$  there can be no other real nodes than those in  $X$ . Then, our goal is, for every  $X \in \mathcal{X}(G)$ , to compute a polynomial-size family  $\mathcal{T}_X$  of “candidate subtrees” whose real nodes are exactly  $X$ . Intuitively,  $\mathcal{T}_X$  should contain all possibilities for  $T\langle X \rangle$  in a “well-structured” 4-Steiner root  $T$  (such a root must satisfy additional properties given in Sec. 4). Note that in practice, we only need to compute this above family for *minimal separators* and *maximal cliques*.

- In Section 6.1 we present an algorithm for computing the collection  $(\mathcal{T}_S)_{S \in \mathcal{S}(G)}$  for the minimal separators. This algorithm serves as a brick-basis construction for computing all the other families.

- Then, we consider in Section 6.2 the maximal cliques  $K_i$  that are leaf-nodes of  $T_G$ . We use the well-known property that all vertices in  $K_i \setminus S_i$  are simplicial in order to generalize the algorithm of the previous section to this new case.
- Finally, we consider the maximal cliques  $K_i$  that are internal nodes of  $T_G$  (Section 6.3). Unsurprisingly, several new difficulties arise in the construction of  $\mathcal{T}_{K_i}$ . Our bottleneck is solving the following subproblem: compute (up to Steiner equivalence) all possible central nodes and their neighbourhood in any subtree  $T\langle K_i \rangle$  of diameter four. We solved this subproblem in most situations, *e.g.*, when there is a minimal separator  $S \subseteq K_i$  such that  $T\langle S \rangle$  must be a bistar (diameter-three subtree). However in some other situations we failed to do so. That left us with some “problematic subsets” called *thin branches*: with exponentially many possible ways to include them in candidate subtrees. As a way to circumvent this combinatorial explosion, we also include in  $\mathcal{T}_{K_i}$  some partially constructed subtrees where the thin branches are omitted. We will greedily decide how to include the thin branches in these subtrees (*i.e.*, how to complete the construction) at Step S.4.

Correctness of this part mostly follows from our structure theorem of Section 4.

- S.3** (*Selection of the encodings.*) For the remaining of the algorithm, let  $(K_q, K_{q-1}, \dots, K_0)$  be a post-ordering of the maximal cliques (*i.e.*, obtained by depth-first-search traversal of  $T_G$ ). We consider the maximal cliques  $K_i \in \mathcal{K}(G)$  sequentially, from  $i = q$  down to  $i = 0$ . Indeed if  $G$  is a 4-Steiner power then (by hereditariness), so is the subgraph  $G_i = (V_i, E_i)$  that is induced by all the maximal cliques in the subtree  $T_G^i$  rooted at  $K_i$ . Steps S.3 and S.4 are devoted to the computation of a subset  $\mathcal{T}_i$  of 4-Steiner roots for  $G_i$ . Specifically, for any 4-Steiner root  $T_i$  of  $G_i$  we define the following encoding:

$$\text{encode}(T_i) := \left[ T_i\langle S_i \rangle \mid (dist_{T_i}(r, W_i))_{r \in V(T_i\langle S_i \rangle)} \right].$$

During Step S.3 we compute a polynomial-size subset of *allowed* encodings for the partial solutions in  $\mathcal{T}_i$ . That is, we only want to add in  $\mathcal{T}_i$  some partial solutions for which the encoding is in the list. Formally, we define an auxiliary problem called DISTANCE-CONSTRAINED ROOT, where given an encoding as input, we ask whether there exists a corresponding 4-Steiner root of  $G_i$ . Our set of allowed encodings for  $\mathcal{T}_i$  can be seen as a set of inputs for which we need to solve DISTANCE-CONSTRAINED ROOT.

We stress that in order to compute these encodings, we use: the families computed at Step S.2, some properties of the rooted clique-tree  $T_G$ , *and* some pre-computed subsets  $\mathcal{T}_j$  of partial solutions for the siblings  $K_j$  of  $K_i$ . Specifically, if  $K_{p(i)} \in \mathcal{K}(G)$  has children nodes  $K_{i_1}, K_{i_2}, \dots, K_{i_p}$ , where  $p(i) < i_1 < i_2 < \dots < i_p$  then, we impose that  $S_{i_1}, S_{i_2}, \dots, S_{i_p}$  are ordered by decreasing size. Step S.3 can start for  $K_{i_j}$  only after that Steps S.3 and S.4 are completed for all of  $K_{i_{j+1}}, K_{i_{j+2}}, \dots, K_{i_p}$ . This means in particular that executions of Steps S.3 and S.4 (for different maximal cliques) are intertwined.

- S.4** (*Greedy strategy.*) After Step S.3 is completed,  $K_i$  received a polynomial-size subset of constraints (*a.k.a.*, encodings) for the 4-Steiner roots of  $G_i$  we want to compute. For every such constraints, we are left to decide whether there exists a 4-Steiner root of  $G_i$  which satisfies all of them (*i.e.*, we must solve DISTANCE-CONSTRAINED ROOT).

- **Case  $K_i$  is a leaf-node.** In this situation,  $V_i = K_i$ . After Step S.2 is completed, we are given a family of all possible subtrees  $T\langle K_i \rangle$ . We are left verifying whether there exists a solution in this family which satisfies all of the constraints.
- **Case  $K_i$  is an internal node.** Let  $K_{i_1}, K_{i_2}, \dots, K_{i_p}$  be the children nodes of  $K_i$  in  $T_G$ . We will construct  $\mathcal{T}_i$  from the partial solutions in  $\mathcal{T}_{i_1}, \mathcal{T}_{i_2}, \dots, \mathcal{T}_{i_p}$ . For that, we try to combine all the possible subtrees  $T\langle K_i \rangle$  (computed during Step S.2) with the partial solutions stored in the sets  $\mathcal{T}_{i_j}$  by using a series of tests based on a maximum-weight matching algorithm (Section 8). – We use the same strategy in order to incorporate the so called thin branches, so as to complete the construction of the subtrees  $T\langle K_i \rangle$ . – We stress the intriguing relationship between our approach and the implementation of the `alldifferent` constraint in constraint programming [Rég94].

**S.5** (*Output.*) Overall since  $G_0 = G$ , we have  $G$  is a 4-Steiner power if and only if  $\mathcal{T}_0 \neq \emptyset$ . Furthermore, any tree  $T \in \mathcal{T}_0$  is a 4-Steiner root of  $G$ .

### 3 Playing with the root

Some general relationships between  $k$ -Steiner roots and clique-intersections are proved in Section 3.2, for any  $k$ . These structural results will be the cornerstone of our algorithm and its analysis. Before presenting all these properties, we establish several useful facts on trees in Section 3.1 (most of them being likely to be known).

#### 3.1 General results on trees

We first recall the *unimodality* property for the eccentricity function on trees, as well as some other related properties:

**Lemma 3.1** (folklore). *The following hold for any tree  $T$ :*

- (P-3.1.1.) *For every node  $v \in V(T)$  we have  $\text{ecc}_T(v) = \text{dist}_T(v, \mathcal{C}(T)) + \text{rad}(T)$ ;*
- (P-3.1.2.) *Every diametral path in  $T$  contains all the nodes in  $\mathcal{C}(T)$  (as its middle nodes);*
- (P-3.1.3.)  *$\mathcal{C}(T)$  is reduced to a node if  $\text{diam}(T)$  is even, and to an edge if  $\text{diam}(T)$  is odd;*
- (P-3.1.4.)  *$\text{rad}(T) = \lceil \text{diam}(T)/2 \rceil$ .*

Based on the above, the following properties on subtree intersections can be derived:

**Lemma 3.2.** *Given a tree  $T$  let  $T_1, T_2$  be two subtrees such that  $\text{diam}(T_1 \cap T_2) = \text{diam}(T_1)$ . Then, we have  $\text{diam}(T_1 \cup T_2) = \text{diam}(T_2)$ .*

*Proof.* First we claim that  $\mathcal{C}(T_1 \cap T_2) = \mathcal{C}(T_1)$ . Indeed, since  $T_1 \cap T_2$  and  $T_1$  are trees with equal diameter, and we have  $T_1 \cap T_2 \subseteq T_1$ , every diametral path for  $T_1 \cap T_2$  is also a diametral path for  $T_1$ . Furthermore, since on every diametral path in a tree, the middle nodes are exactly the center nodes (Prop. P-3.1.2), we obtain as claimed that  $\mathcal{C}(T_1 \cap T_2) = \mathcal{C}(T_1)$ .

Then, let  $x, y \in V(T_1 \cap T_2)$  be the two ends of a diametral path in the subtree  $T_1 \cap T_2$ . We set  $z \in \{x, y\}$  maximizing  $\text{dist}_T(z, \mathcal{C}(T_2))$  and we claim that, for every  $v_1 \in V(T_1)$ ,  $\text{dist}_T(v_1, \mathcal{C}(T_2)) \leq \text{dist}_T(z, \mathcal{C}(T_2))$ . Before we prove this claim, let us explain why this proves the lemma. Every node of

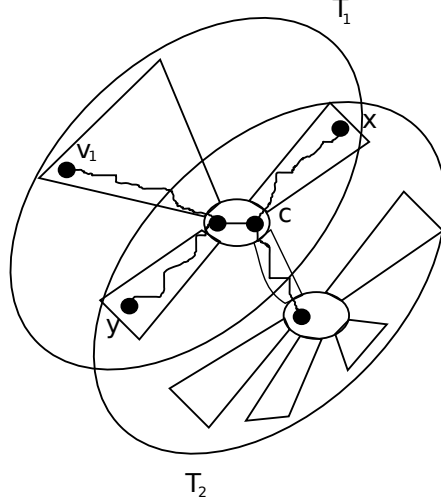


Figure 3: To the proof of Lemma 3.2.

$V(T_1)$  is at a distance  $\leq \text{dist}_T(z, \mathcal{C}(T_2)) + \text{rad}(T_2)$  from any node in  $V(T_2)$ . By unimodality (Prop. P-3.1.1),  $\text{ecc}_{T_2}(z) = \text{dist}_T(z, \mathcal{C}(T_2)) + \text{rad}(T_2) \leq \text{diam}(T_2)$ , and so,  $\text{diam}(T_1 \cup T_2) = \text{diam}(T_2)$ .

Finally, in order to prove this above claim there are two cases.

- First assume  $\mathcal{C}(T_1) \subseteq \mathcal{C}(T_2)$ . We recall that since the unique  $xy$ -path in  $T$  must contain all of  $\mathcal{C}(T_1)$  (Prop. P-3.1.2), there can be no component of  $T \setminus \mathcal{C}(T_1)$  that contains both  $x, y$ . In particular, there exists  $z \in \{x, y\}$  such that no component of  $T \setminus \mathcal{C}(T_1)$  can both contain  $z$  and intersect  $\mathcal{C}(T_2) \setminus \mathcal{C}(T_1)$ . Then,  $\text{dist}_T(z, \mathcal{C}(T_2)) = \text{dist}_T(z, \mathcal{C}(T_1))$ . Furthermore by unimodality (Prop. P-3.1.1) every node  $v_1 \in V(T_1)$  has eccentricity  $\text{dist}_T(v_1, \mathcal{C}(T_1)) + \text{rad}(T_1)$ . Since  $z$  is an end in a diametral path of  $T_1$  it maximizes  $\text{dist}_T(z, \mathcal{C}(T_1))$ , and so, for every  $v_1 \in V(T_1)$  we have  $\text{dist}_T(v_1, \mathcal{C}(T_2)) \leq \text{dist}_T(v_1, \mathcal{C}(T_1)) \leq \text{dist}_T(z, \mathcal{C}(T_1)) = \text{dist}_T(z, \mathcal{C}(T_2))$ .
- Otherwise, let  $c \in \mathcal{C}(T_1)$  minimize  $\text{dist}_T(c, \mathcal{C}(T_2))$ . Note that since we have  $\mathcal{C}(T_1) \not\subseteq \mathcal{C}(T_2)$ , there is a unique possible choice for  $c$ . Furthermore, every  $v_1 \in V(T_1)$  satisfies  $\text{dist}_T(v_1, \mathcal{C}(T_2)) \leq \text{dist}_T(v_1, c) + \text{dist}_T(c, \mathcal{C}(T_2)) \leq \text{rad}(T_1) + \text{dist}_T(c, \mathcal{C}(T_2))$ , and we will show this upper-bound is reached for at least one of  $x$  or  $y$ . Specifically, we can refine one observation from the previous case as follows: there exists  $z \in \{x, y\}$  such that no component of  $T \setminus \mathcal{C}(T_1)$  can both contain  $z$  and intersect  $\mathcal{C}(T_2) \setminus \mathcal{C}(T_1)$ ; and in the special case where  $\mathcal{C}(T_1)$  is an edge,  $c$  is not the closest central node to  $z$ . In this situation,  $\text{dist}_T(z, c) = \text{rad}(T_1)$  and the path between  $z$  and  $\mathcal{C}(T_2)$  goes by  $c$ . See Fig. 3 for an illustration.

In both cases we obtain, as claimed,  $\text{dist}_T(v_1, \mathcal{C}(T_2)) \leq \text{dist}_T(z, \mathcal{C}(T_2))$  for every  $v_1 \in V(T_1)$ .  $\square$

**Lemma 3.3.** *Given a tree  $T$  let  $T_1, T_2$  be two subtrees such that  $\mathcal{C}(T_1) \subseteq \mathcal{C}(T_2)$ . Then, we have that  $\text{diam}(T_1 \cup T_2) = \max\{\text{diam}(T_1), \text{diam}(T_2)\}$ .*

*Proof.* Since  $\mathcal{C}(T_1) \subseteq \mathcal{C}(T_2)$  we have for every  $v_1 \in V(T_1)$ :

$$\text{ecc}_{T_1 \cup T_2}(v_1) \leq \text{dist}_T(v_1, \mathcal{C}(T_1)) + \max\{\text{rad}(T_1), \text{rad}(T_2)\}.$$

By the unimodality property (Property P-3.1.1) we have that  $dist_T(v_1, \mathcal{C}(T_1)) = ecc_{T_1}(v_1) - rad(T_1) \leq diam(T_1) - rad(T_1)$ , and so by Property P-3.1.4:

$$dist_T(v_1, \mathcal{C}(T_1)) \leq \lfloor diam(T_1)/2 \rfloor \leq \max\{\lfloor diam(T_1)/2 \rfloor, \lfloor diam(T_2)/2 \rfloor\}.$$

In the same way, Property P-3.1.4 implies that:

$$\max\{rad(T_1), rad(T_2)\} = \max\{\lfloor diam(T_1)/2 \rfloor, \lfloor diam(T_2)/2 \rfloor\}.$$

We so obtain that  $ecc_{T_1 \cup T_2}(v_1) \leq \max\{diam(T_1), diam(T_2)\}$ .

Then, for every  $v_2 \in V(T_2)$ :

$$\begin{aligned} ecc_{T_1 \cup T_2}(v_2) &\leq dist_T(v_2, \mathcal{C}(T_2)) + \max\{rad(T_2), diam(\mathcal{C}(T_2)) + rad(T_1)\} \\ &\leq dist_T(v_2, \mathcal{C}(T_2)) + \max\{rad(T_2), 1 + rad(T_1)\}. \end{aligned}$$

We may assume  $rad(T_1) \geq rad(T_2)$  since otherwise,  $ecc_{T_1 \cup T_2}(v_2) \leq dist_T(v_2, \mathcal{C}(T_2)) + rad(T_2) = ecc_{T_2}(v_2) \leq diam(T_2)$  by unimodality (Property P-3.1.1). In particular we claim that it implies  $diam(T_1) \geq diam(T_2)$ . Indeed, by Property P-3.1.4 we must have  $\lfloor diam(T_1)/2 \rfloor \geq \lfloor diam(T_2)/2 \rfloor$ , and so  $diam(T_1) \geq diam(T_2) - 1$ . Moreover since by the hypothesis we also have  $\mathcal{C}(T_1) \subseteq \mathcal{C}(T_2)$ , by Property P-3.1.3 we obtain as claimed  $diam(T_1) \geq diam(T_2)$  (otherwise,  $diam(T_1)$  and  $diam(T_2)$  must be even and odd, respectively, and so we cannot have  $\lfloor diam(T_1)/2 \rfloor \geq \lfloor diam(T_2)/2 \rfloor$ , a contradiction). There are now two cases to consider:

- Case  $diam(T_1) = diam(T_2)$ . Then,  $\mathcal{C}(T_1) = \mathcal{C}(T_2)$  and we can strengthen our previous inequality as follows:  $ecc_{T_1 \cup T_2}(v_2) \leq dist_T(v_2, \mathcal{C}(T_2)) + \max\{rad(T_2), rad(T_1)\} \leq diam(T_2)$ .
- Case  $diam(T_1) > diam(T_2)$ . Recall that  $dist_T(v_2, \mathcal{C}(T_2)) \leq \lfloor diam(T_2)/2 \rfloor$ . In particular, either  $diam(T_1) \geq diam(T_2) + 2$ , and so,  $dist_T(v_2, \mathcal{C}(T_2)) \leq \lfloor diam(T_1)/2 \rfloor - 1$ ; or  $diam(T_1) = diam(T_2) + 1$  but then, since by the hypothesis we have  $\mathcal{C}(T_1) \subseteq \mathcal{C}(T_2)$ , by Property P-3.1.3  $diam(T_1)$  must be even, and so,  $dist_T(v_2, \mathcal{C}(T_2)) \leq \lfloor diam(T_1)/2 \rfloor - 1$  also in this case. Overall:

$$\begin{aligned} ecc_{T_1 \cup T_2}(v_2) &\leq dist_T(v_2, \mathcal{C}(T_2)) + \max\{rad(T_2), diam(\mathcal{C}(T_2)) + rad(T_1)\} \\ &\leq \lfloor diam(T_1)/2 \rfloor - 1 + rad(T_1) + 1 = diam(T_1). \end{aligned}$$

Therefore, in both cases we obtain  $diam(T_1 \cup T_2) \leq \max\{diam(T_1), diam(T_2)\}$ .  $\square$

### 3.2 A structure theorem

We are now ready to state the main result in this section:

**Theorem 3.4.** *Given  $G = (V, E)$  and  $T$  any  $k$ -Steiner root of  $G$ , the following properties hold for any clique-intersection  $X \in \mathcal{X}(G)$ :*

- (P-3.4.1.) *We have  $Real(T \langle X \rangle) = X$  and  $diam(T \langle X \rangle) \leq k$ ;*
- (P-3.4.2.) *If  $T'_X \supset T \langle X \rangle$  then, either  $X = Real(T'_X)$  or  $diam(T'_X) > diam(T \langle X \rangle)$ ;*
- (P-3.4.3.) *If  $\mathcal{C}(T \langle X \rangle) \subseteq \mathcal{C}(T \langle X' \rangle)$  then,  $X \cup X'$  is a clique of  $G$ .*

*Proof.* We prove each property separately.

(*Proof of Property P-3.4.1.*) First assume  $X \in \mathcal{K}(G)$  to be a maximal clique. Since all leaves of  $T\langle X \rangle$  are in  $X$ ,  $\text{diam}(T\langle X \rangle) = \max_{u,v \in X} \text{dist}_T(u,v)$ . By the hypothesis  $T$  is a  $k$ -Steiner root of  $G$ , and so, since  $X$  is a clique of  $G$ ,  $\max_{u,v \in X} \text{dist}_T(u,v) \leq k$ . In particular,  $\text{diam}(T\langle X \rangle) \leq k$ , that implies in turn the vertices of  $\text{Real}(T\langle X \rangle)$  must induce a clique of  $G$ . We can conclude that  $\text{Real}(T\langle X \rangle) = X$  by maximality of  $X$ . More generally, let  $X = \bigcap_{i=1}^{\ell} K_i$ , for some family  $K_1, K_2, \dots, K_{\ell} \in \mathcal{K}(G)$ . Clearly,  $T\langle X \rangle \subseteq \bigcap_{i=1}^{\ell} T\langle K_i \rangle$ , and so,  $X \subseteq \text{Real}(T\langle X \rangle) \subseteq \bigcap_{i=1}^{\ell} \text{Real}(T\langle K_i \rangle)$ . As we proved before,  $\text{Real}(T\langle K_i \rangle) = K_i$  for every  $1 \leq i \leq \ell$ , and so,  $\text{Real}(T\langle X \rangle) \subseteq \bigcap_{i=1}^{\ell} K_i = X$ . Altogether combined, we obtain that  $\text{Real}(T\langle X \rangle) = X$ .

(*Proof of Property P-3.4.2.*) Let  $T'_X \supset T\langle X \rangle$  be such that  $\text{diam}(T'_X) = \text{diam}(T\langle X \rangle)$ . We claim  $\text{Real}(T'_X) = X$ , that will prove the second part of the theorem. Indeed, for any maximal clique  $K_j$  that contains  $X$ , we have  $\text{diam}(T\langle K_j \rangle \cap T'_X) \geq \text{diam}(T\langle X \rangle) = \text{diam}(T'_X)$ , and so,  $\text{diam}(T\langle K_j \rangle \cup T'_X) = \text{diam}(T\langle K_j \rangle) \leq k$  by Lemma 3.2. It implies  $\text{Real}(T'_X) \subseteq K_j$ . Furthermore, since  $X \in \mathcal{X}(G)$ ,  $X$  is exactly the intersection of all the maximal cliques  $K_j$  that contains it, thereby proving the claim.

(*Proof of Property P-3.4.3.*) Finally, assume  $\mathcal{C}(T\langle X \rangle) \subseteq \mathcal{C}(T\langle X' \rangle)$ . By Lemma 3.3 we obtain that  $\text{diam}(T\langle X \rangle \cup T\langle X' \rangle) = \max\{\text{diam}(T\langle X \rangle), \text{diam}(T\langle X' \rangle)\} \leq k$ . In particular,  $X \cup X'$  must be a clique of  $G$ .  $\square$

Before ending this section, we slightly strengthen Property P-3.4.3 of Theorem 3.4, as follows:

**Lemma 3.5.** *Given  $G = (V, E)$  and  $T$  any  $2k$ -Steiner root of  $G$ , we have  $\mathcal{C}(T\langle K_i \rangle) \cap \mathcal{C}(T\langle K_j \rangle) = \emptyset$  for any two different maximal cliques  $K_i, K_j \in \mathcal{K}(G)$ .*

*Proof.* Suppose for the sake of contradiction  $\mathcal{C}(T\langle K_i \rangle) \cap \mathcal{C}(T\langle K_j \rangle) \neq \emptyset$ , and let  $r \in \mathcal{C}(T\langle K_i \rangle) \cap \mathcal{C}(T\langle K_j \rangle)$ . By Theorem 3.4 (Prop. P-3.4.1),  $\max\{\text{diam}(T\langle K_i \rangle), \text{diam}(T\langle K_j \rangle)\} \leq 2k$ . Thus, any vertex of  $T\langle K_i \rangle \cup T\langle K_j \rangle$  is at a distance  $\leq k$  from  $r$  in  $T$  (this follows from Prop. P-3.1.4 of Lemma 3.1). In particular,  $\text{diam}(T\langle K_i \rangle \cup T\langle K_j \rangle) \leq 2k$ , and so,  $K_i \cup K_j$  is a clique of  $G$ . The latter contradicts the fact that  $K_i, K_j$  are maximal cliques.  $\square$

## 4 Well-structured 4-Steiner roots

We refine our results in the previous Section when  $k = 4$ . Let us start motivating our approach. Given  $G = (V, E)$  we recall that one of our intermediate goals is to compute, by dynamic programming on a clique-tree, subsets  $\mathcal{T}_i$  of 4-Steiner roots for some collection of subgraphs  $G_i$ , with the following additional property: assuming  $G$  is a 4-Steiner power, there must be a partial solution in  $\mathcal{T}_i$  which can be extended to a 4-Steiner root for  $G$  (cf. Sec. 2.3, Step S.3). Ideally, one should store *all* the possible 4-Steiner roots for  $G_i$ , however this leads to a combinatorial explosion. In order to (partly) overcome this issue, we introduce the following important notion:

**Definition 4.1.** Given  $G = (V, E)$  and  $X \in \mathcal{X}(G)$ , a vertex  $v \in X$  is called  *$X$ -constrained* if it satisfies one of the following two conditions:

1. either there is another clique-intersection  $X' \subset X$  such that  $v \in X'$  and  $|X'| \geq 2$  (we call  $v$  *internally  $X$ -constrained*);
2. or there exist  $X_1, X_2 \in \mathcal{X}(G)$  such that  $X \subset X_1$  and  $X \cap X_2 = \{v\} \subset X_1 \cap X_2$  (we call  $v$   *$(X, X_1, X_2)$ -sandwiched*).

A vertex  $v \in X$  that is not  $X$ -constrained is called  $X$ -free.

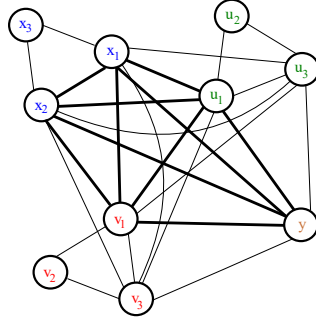


Figure 4: Five maximal cliques  $K_1 = \{x_1, x_2, x_3\}$ ,  $K_2 = \{u_1, u_2, u_3\}$ ,  $K_3 = \{v_1, v_2, v_3\}$ ,  $K_4 = \{y, x_1, x_2, u_1, v_1, u_3\}$  and  $K_5 = \{y, x_1, x_2, u_1, v_1, v_3\}$ . In the clique-intersection  $X = \{y, x_1, x_2, u_1, v_1\}$  both vertices  $x_1, x_2$  are internally  $X$ -constrained (for  $X' = X \cap K_1$ ), vertex  $u_1$  is  $(X, K_4, K_2)$ -sandwiched, vertex  $v_1$  is  $(X, K_5, K_3)$ -sandwiched and vertex  $y$  is  $X$ -free.

We refer to Fig. 4 for an illustration of all possibilities. Our study reveals on the one hand that  $X$ -constrained vertices have a very rigid structure. It seems on the other hand that  $X$ -free vertices are completely unstructured and mostly responsible for the combinatorial explosion of possibilities for  $T\langle X \rangle$ . As our main contribution in this section we prove that we can always force the  $X$ -free vertices to be *leaves* of this subtree  $T\langle X \rangle$ , thereby considerably reducing the number of possibilities for the latter. Specifically:

**Theorem 4.1.** *Let  $G = (V, E)$  be a 4-Steiner power. There always exists a 4-Steiner root  $T$  of  $G$  where, for any clique-intersection  $X \in \mathcal{X}(G)$ :*

- (P-4.1.1.) *all the  $X$ -free vertices are leaves of  $T\langle X \rangle$  with maximum eccentricity  $\text{diam}(T\langle X \rangle)$ ;*
- (P-4.1.2.) *there is a node  $c \in \mathcal{C}(T\langle X \rangle)$  such that for every  $X$ -free vertex  $v$ , except maybe one,  $\text{dist}_T(v, c) = \text{dist}_T(v, \mathcal{C}(T\langle X \rangle))$ ;*
- (P-4.1.3.) *all the internal nodes on a path between  $\mathcal{C}(T\langle X \rangle)$  and a  $X$ -free vertex are Steiner nodes of degree two;*
- (P-4.1.4.) *and if  $X \in \mathcal{K}(G)$  and it has a  $X$ -free vertex then,  $\text{diam}(T\langle X \rangle) = 4$ .*

Theorem 4.1 is proved by carefully applying a set of operations on an arbitrary 4-Steiner root until it satisfies all of the desired properties. We give two examples of such operations in Fig. 6 and 7 (see the proof of Theorem 4.1 in order to better understand these two examples). It is crucial for the proof that in any 4-Steiner root of  $G$  all the minimal separators yield subtrees of diameter at most three.

In the remaining of the paper, we call a 4-Steiner root with the above properties *well-structured*. It will appear in Sec. 6, 7, and 8 that we only add partial solutions in the subsets  $\mathcal{T}_i$  which are in some sense “close” to the subset of well-structured partial solutions. We first prove Theorem 4.1 for maximal cliques (Section 4.1). This first part of the proof looks easier to generalize to larger values of  $k$ . Then, we prove the result in its full generality in Section 4.2.

## 4.1 The case of (Almost) Simplicial vertices

Let  $K_i \in \mathcal{K}(G)$  be fixed. We start giving a simple characterization of  $K_i$ -free vertices in terms of simplicial vertices and cut-vertices. Then, we prove Theorem 4.1 in the special case when  $X$  is a maximal clique.

**Lemma 4.2.** *Given  $G = (V, E)$  and  $K_i \in \mathcal{K}(G)$ , a vertex  $v \in K_i$  is  $K_i$ -free if and only if:*

- *either it is simplicial;*
- *or it is a cut-vertex, and there is no other minimal separator of  $G$  contained into  $K_i$  that can also contain  $v$ .*

*Proof.* By maximality of  $K_i$ , a vertex  $v \in K_i$  can only be either  $K_i$ -free or *internally*  $K_i$ -constrained. In particular,  $v$  is  $K_i$ -free if and only if for any other  $K_j \in \mathcal{K}(G)$  we have either  $v \notin K_j$  or  $K_i \cap K_j = \{v\}$ . As an extremal case, a vertex  $v \in K_i$  is  $K_i$ -free if is not contained into any other maximal clique, and that is the case if and only if  $v$  is simplicial. Thus, from now on assume  $v$  is not simplicial. If  $v \in K_i \cap K_j$  then, in any clique-tree  $T_G$  of  $G$ , the vertex  $v$  and more generally, all of  $K_i \cap K_j$ , is contained into all the minimal separators that label an edge of the  $K_i K_j$ -path in  $T_G$ . This implies that there is always a largest clique-intersection  $X \subset K_i$  containing  $v$  that is a minimal separator. Hence a non simplicial  $v \in K_i$  is  $K_i$ -free if and only if it is a cut-vertex and there is no other minimal separator in  $K_i$  that contains this vertex.  $\square$

**Lemma 4.3.** *Let  $G = (V, E)$  be a 4-Steiner power. There exists a 4-Steiner root  $T$  of  $G$  such that the following hold for any maximal clique  $K_i$ :*

- *If there is at least one  $K_i$ -free vertex then,  $\text{diam}(T \langle K_i \rangle) = 4$ ;*
- *every  $K_i$ -free vertex  $v$  is a leaf of  $T \langle K_i \rangle$ . Moreover  $\text{dist}_T(v, \mathcal{C}(T \langle K_i \rangle)) = 2$  and the internal node onto the unique  $v\mathcal{C}(T \langle K_i \rangle)$ -path is a degree-two Steiner node.*

*Proof.* We give an illustration of the proof in Fig. 5. First we pick an arbitrary 4-Steiner root  $T$  of  $G$ , that exists by the hypothesis. Define  $S_1$  to be the set of all the cut-vertices in  $G$  that are  $K_i$ -free for some  $K_i \in \mathcal{K}(G)$ . We now proceed by induction on  $|S_1|$ .

Assume  $S_1 = \emptyset$  for the base case. For every maximal clique  $K_i$ , by Lemma 4.2, the  $K_i$ -free vertices are exactly the simplicial vertices in  $K_i$ . We consider all the simplicial vertices  $v \in K_i$  sequentially, and we proceed as follows. Let  $c_i \in \mathcal{C}(T \langle K_i \rangle)$  minimize  $\text{dist}_T(v, c_i)$  (possibly,  $v = c_i$ ). We first replace  $v$  by a Steiner node  $\alpha$ . In doing so, we get a 4-Steiner root  $T'$  for  $G \setminus v$ . Then, let  $c'_i$  be either  $c_i$  (if  $c_i \neq v$ ) or  $\alpha$  (if  $c_i = v$ ). We connect  $v$  to  $c'_i$  via a path of length exactly  $4 - \max_{u \in K_i \setminus \{v\}} \text{dist}_{T'}(c'_i, u)$  of which all internal nodes are Steiner. In doing so, we obtain a tree  $T''$  such that  $\text{Real}(T'') = V$ . By construction,  $\max_{u \in K_i \setminus \{v\}} \text{dist}_{T'}(c'_i, u) \leq \text{ecc}_{T \langle K_i \rangle}(c_i) \leq 2$  (since  $\text{diam}(T \langle K_i \rangle) \leq 4$ ). It implies:

$$\text{dist}_T(v, c_i) \leq \text{ecc}_{T \langle K_i \rangle}(c_i) \leq 4 - \max_{u \in K_i \setminus \{v\}} \text{dist}_{T'}(c'_i, u) = \text{dist}_{T''}(v, c'_i).$$

As a result, the distances between real nodes can only increase compared to  $T$ , and this new tree  $T''$  we get keeps the property of being a 4-Steiner root of  $G$ . Furthermore,  $\text{diam}(T'' \langle K_i \rangle) = 4$  and the unique central node in  $\mathcal{C}(T'' \langle K_i \rangle)$  is onto the  $v c'_i$ -path by construction. We observe that  $K_i$



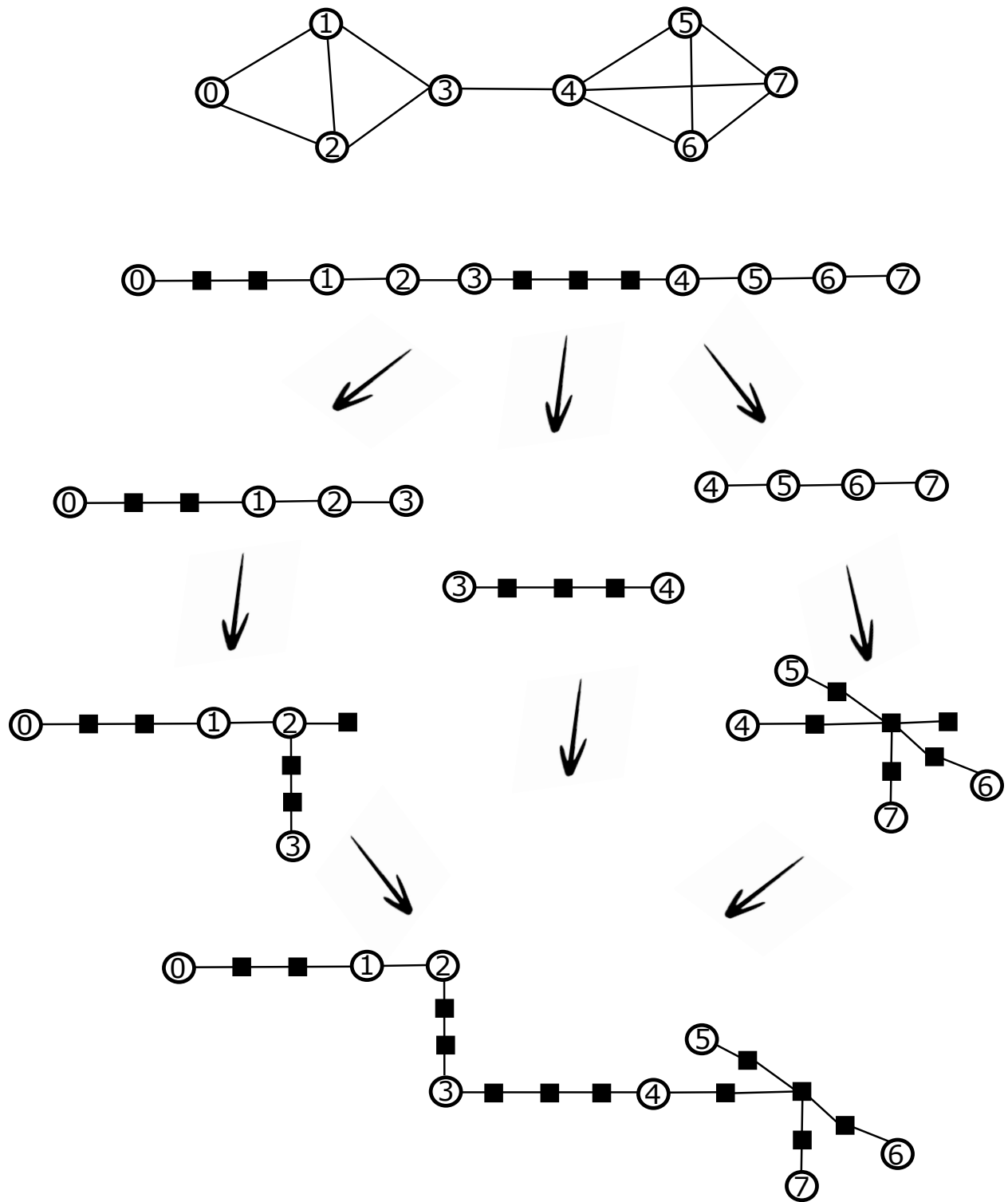


Figure 5: The transformation of Lemma 4.3 applied to an arbitrary 4-Steiner root.

and  $v$  cannot falsify the properties of the lemma at any further loop. Overall, after this first phase is done we may assume that all the simplicial vertices  $v$  are contained into some maximal clique  $K_i$  such that:  $\text{diam}(T\langle K_i \rangle) = 4$ ,  $v$  is a leaf of  $T\langle K_i \rangle$  such that  $\text{dist}_T(v, \mathcal{C}(T\langle K_i \rangle)) = 2$ , and the internal node onto the  $v\mathcal{C}(T\langle K_i \rangle)$ -path is Steiner.

From now on we assume  $S_1 \neq \emptyset$ . Let  $v \in S_1$  and let  $C_1, C_2, \dots, C_\ell$  be the connected components of  $G \setminus v$ . For every  $i \in \{1, 2, \dots, \ell\}$ , the graph  $G_i := G[C_i \cup \{v\}]$  is a 4-Steiner power as this is a hereditary property. Specifically, given a fixed 4-Steiner root  $T$  for  $G$ , we obtain a 4-Steiner root  $T^{(i)}$  for  $G_i$  by replacing every vertex in  $V(G) \setminus V(G_i)$  by a Steiner node. By induction, we can modify all the subtrees  $T^{(i)}$  into some new subtrees  $T^{(i,2)}$  that satisfy the properties of the lemma w.r.t.  $G_i$ . Overall, by identifying all the  $T^{(i,2)}$ 's at  $v$ , one obtains a tree  $T'$ . We claim that  $T'$  satisfies the two properties of the lemma. Indeed, it follows from the characterization of Lemma 4.2 that, for any maximal clique  $K_j \subseteq C_i \cup \{v\}$ , the  $K_j$ -free vertices in  $G$  are still  $K_j$ -free vertices in  $G_i$ . – Note that in particular, if  $v$  is  $K_j$ -free in  $G$  then,  $v$  is simplicial in  $G_i$ . – Therefore, the claim is proved. It remains now to show that  $T'$  is indeed a 4-Steiner root of  $G$ . This is not the case only if there exist  $x \in C_p$ ,  $y \in C_q$  such that  $p \neq q$  and  $\text{dist}_{T'}(x, v) + \text{dist}_{T'}(v, y) \leq 4$ . Our construction implies  $\text{dist}_{T'}(x, v) \geq \text{dist}_T(x, v)$  and  $\text{dist}_{T'}(y, v) \geq \text{dist}_T(y, v)$ . But then, we should already have  $\text{dist}_T(x, y) \leq 4$  in the original root  $T$ . Thus, since  $T$  is a 4-Steiner root of  $G$ , this case cannot happen and  $T'$  is also a 4-Steiner root of the graph  $G$ .  $\square$

## 4.2 The general case

We are now ready to prove Theorem 4.1 in its full generality:

*Proof of Theorem 4.1.* Let  $T$  be such that the result holds for maximal cliques (such a  $T$  exists by Lemma 4.3). For any  $X \in \mathcal{X}(G) \setminus \mathcal{K}(G)$  with at most two elements, the properties of the theorem always hold (for any  $T$ ). We so only consider the clique-intersections  $X \in \mathcal{X}(G) \setminus \mathcal{K}(G)$  with at least three elements. Then,  $\text{diam}(T\langle X \rangle) \leq 3$  by Theorem 3.4 – Prop P-3.4.2 (*i.e.*, because  $X$  is strictly contained into some maximal clique  $K$  and so,  $\text{diam}(T\langle X \rangle) < \text{diam}(T\langle K \rangle)$ ). In what follows, we will often use properties of the subtree  $T\langle X \rangle$  that only hold if  $\text{diam}(T\langle X \rangle) \leq 3$ ; namely:

- $T\langle X \rangle \setminus \mathcal{C}(T\langle X \rangle)$  is a collection of leaf-nodes;
- there are at least two leaf-nodes (*i.e.*, because  $|X| \geq 3$ ), and there is at least one leaf-node adjacent to every central node in  $\mathcal{C}(T\langle X \rangle)$ ;
- (as a direct consequence of the previous property) every central node in  $\mathcal{C}(T\langle X \rangle)$  has at least two neighbours in  $T\langle X \rangle$ .

In particular, Property P-4.1.3 is now implied by Property P-4.1.1 as there can be no internal node on the path between a leaf and the closest central node. In the same way, Property P-4.1.1 can be slightly simplified as *every* leaf has maximum eccentricity; hence, we only need to ensure that  $X$ -free vertices are leaves. Finally, this also implies that  $\text{Steiner}(T\langle X \rangle) \subseteq \mathcal{C}(T\langle X \rangle)$ , as all leaves of  $T\langle X \rangle$  must be in  $X$  (*i.e.*, by the very definition of  $T\langle X \rangle$ ).

The proof follows from different uses of a special operation on the tree  $T$  that we now introduce:

**Operation 1.** Let  $X \in \mathcal{X}(G) \setminus \mathcal{K}(G)$  have size at least three and let  $v \in X$ . We define  $R_v$  to be the forest of all the subtrees in  $T \setminus T\langle X \rangle$  that contain one node adjacent to  $v$  and have *no* real node at a distance  $\leq 4$  from  $X \setminus \{v\}$  in  $T$ . Let  $Q_v$  be the subtree of  $T$  that is induced by  $R_v \cup \{v\}$ .

We construct a new tree  $T'$  from  $T$  in two steps:

1. We remove  $R_v$  and we replace  $v$  by a Steiner node  $\alpha_v$ . In doing so, we obtain an intermediate tree denoted by  $T_v$ . Note that  $T_v$  contains the subtree  $T_v^X$ : that is obtained from  $T\langle X \rangle$  by replacing  $v$  with  $\alpha_v$  (and so, is isomorphic to  $T\langle X \rangle$ );
2. In order to obtain  $T'$  from  $T_v$ , we add a copy of  $Q_v$  and an edge  $vc$  between  $v$  and a center node of  $T_v^X$  (possibly,  $c = \alpha_v$ ).

We refer to Fig. 6 and 7 for some particular applications of Operation 1. Furthermore in what follows we prove that under some conditions of use, this above Operation 1 always outputs a 4-Steiner root  $T'$  that is closer to satisfy all the properties of the theorem than  $T$ . Specifically:

**Claim 4.3.1.** *Assume that  $v$  is  $X$ -free and that every center node of  $T\langle X \rangle$  is adjacent to a real node in  $X \setminus \{v\}$ . Then,  $T'$  keeps the property of being a 4-Steiner root of  $G$  if and only if the following conditions hold:*

- (Condition 4.3.1.a) either  $\text{dist}_T(\text{Real}(R_v), v) \geq 4$  or  $c$  is Steiner;
- (Condition 4.3.1.b) if  $c \neq \alpha_v$  then,  $\text{dist}_{T_v}(c, V \setminus N_G[v]) > 3$ .

Moreover, for any  $X' \in \mathcal{X}(G) \setminus \{X\}$ , if any of Properties P-4.1.1, P-4.1.2, P-4.1.3 or P-4.1.4 is satisfied for  $X'$  in  $T$  then, this stays so in  $T'$ .

*Proof.* First we prove that all the real vertices in  $R_v$  are at a distance  $> 4$  from  $V(G) \setminus V(Q_v)$  in the original tree  $T$  (Subclaim 4.3.1.1). This will help us in better understanding the structure of  $T$  in the remaining of the proof.

**Subclaim 4.3.1.1.**  $\text{dist}_T(\text{Real}(R_v), V(G) \setminus V(Q_v)) > 4$ .

*Proof.* Suppose for the sake of contradiction that there exist  $x \in \text{Real}(R_v), y \notin V(Q_v)$  such that  $\text{dist}_T(x, y) \leq 4$ . Then,  $v$  is onto the unique  $xy$ -path in  $T$ . Furthermore, recall that by definition of  $R_v$ ,  $\text{dist}_T(x, X \setminus \{v\}) > 4$ . By the hypothesis every central node of  $\mathcal{C}(T\langle X \rangle)$  is adjacent to a vertex of  $X \setminus \{v\}$ , and so we have  $\text{dist}_T(v, X \setminus \{v\}) \leq 2$ . As a result, we obtain  $\text{dist}_T(v, x) = 3$  and  $\text{dist}_T(v, y) = 1$ . But then,  $y$  is at a distance  $\leq \text{diam}(T\langle X \rangle) + 1 \leq 4$  from any vertex in  $X$ , that implies the existence of a clique-intersection  $X_1 \supseteq X \cup \{y\}$ . Let  $X_2$  be any clique-intersection that contains all of  $x, y, v$ . Since we assume  $\text{dist}_T(x, X \setminus \{v\}) > 4$ ,  $X \cap X_2 = \{v\}$  and in particular,  $X_1 \neq X_2$ . It implies that  $v$  is  $(X, X_1, X_2)$ -sandwiched, a contradiction.  $\circ$

It follows from Subclaim 4.3.1.1 that in order for  $T'$  to be a 4-Steiner root for  $G$ , one must ensure  $\text{dist}_{T'}(\text{Real}(R_v), V \setminus V(Q_v)) > 4$ . We then prove that this above condition is equivalent to Condition 4.3.1.a. Indeed, note that we always have  $\text{dist}_{T'}(\text{Real}(R_v), V \setminus V(Q_v)) > 4$  if  $\text{dist}_T(v, \text{Real}(R_v)) \geq 4$ . Otherwise, by the hypothesis every center node of  $\mathcal{C}(T\langle X \rangle)$  is adjacent to a real node in  $X \setminus \{v\}$ , thereby implying  $\text{dist}_T(v, V \setminus V(Q_v)) \leq \text{dist}_T(v, X \setminus \{v\}) \leq 2$ , and so,  $\text{dist}_T(v, \text{Real}(R_v)) = 3$ . Then, a necessary and sufficient condition for having that  $\text{dist}_{T'}(\text{Real}(R_v), V \setminus V(Q_v)) > 4$  is that  $c$  is Steiner.

In the same way, Condition 4.3.1.b implies the necessary condition  $\text{dist}_{T'}(u, v) > 4$  for every  $u \notin N_G(v)$ . However, the above does not prove that  $T'$  is a 4-Steiner root of  $G$  yet, as we also need to ensure  $\text{dist}_{T'}(u, v) \leq 4$  for every  $u \in N_G(v)$ . In order to prove this is the case, and to also prove the second part of the claim, we now consider the clique-intersections  $X' \in \mathcal{X}(G) \setminus \{X\}$  such that  $v \in X'$ . (Note that if  $v \notin X'$  then,  $T\langle X' \rangle = T'\langle X' \rangle$  and so, the result of our claim trivially holds for

such a  $X'$ ). Since we have  $dist_T(Real(R_v), V \setminus V(Q_v)) > 4$  (Subclaim 4.3.1.1), there are only two possibilities: either  $T\langle X' \rangle$  is fully contained into  $Q_v$  – in which case it is not modified –; or it does not intersect  $R_v$ . We then consider two different cases:

- Assume  $X \subset X'$ . In particular,  $T\langle X' \rangle \cap R_v = \emptyset$ . It implies that  $T'\langle X' \rangle$  is obtained from  $T\langle X' \rangle$  by replacing  $v$  by a Steiner node (only if it were an internal node of  $T\langle X' \rangle$ ) then, making of  $v$  a leaf.

**Subclaim 4.3.1.2.**  $diam(T'\langle X' \rangle) = diam(T\langle X' \rangle)$ .

*Proof.* We divide the proof into two parts:

- In order to prove  $diam(T'\langle X' \rangle) \leq diam(T\langle X' \rangle)$ , we consider an auxiliary subtree  $T''\langle X' \rangle$ : obtained from the original  $T\langle X' \rangle$  by adding a leaf  $v'$  to some arbitrary central node  $c'$  in  $\mathcal{C}(T\langle X' \rangle)$ . Note that  $T'\langle X' \rangle$  is isomorphic to a subtree of  $T''\langle X' \rangle$  for the choice of  $c' = c$ . Furthermore since  $c'$  has at least two neighbours in  $T\langle X' \rangle \subseteq T\langle X \rangle$  we have:

$$diam(T\langle X' \rangle \cap (N_{T\langle X' \rangle}[c'] \cup \{v'\})) = diam(N_{T\langle X' \rangle}[c']) = 2 = diam(N_{T\langle X' \rangle}[c'] \cup \{v'\}).$$

We so deduce from Lemma 3.2 that:

$$diam(T''\langle X' \rangle) = diam(T\langle X' \rangle \cup (N_{T\langle X' \rangle}[c'] \cup \{v'\})) = diam(T\langle X' \rangle).$$

In particular,  $diam(T'\langle X' \rangle) \leq diam(T''\langle X' \rangle) = diam(T\langle X' \rangle)$ .

- For the converse direction, it suffices to prove, if  $T\langle X' \rangle \neq T'\langle X' \rangle$ , the existence of a diametral path in  $T\langle X' \rangle$  of which  $v$  is not an end. Since the ends of a diametral path must be leaves, we can restrict our study to the case where:  $v$  is a leaf of  $T\langle X \rangle$ , and there is no subtree of  $T\langle X' \rangle \setminus T\langle X \rangle$  that contains a neighbour of  $v$  in  $T$ . In particular,  $T\langle X \rangle \neq T'\langle X \rangle$  (otherwise  $T\langle X' \rangle = T'\langle X' \rangle$ , and so we are done). We deduce from the above that  $v$  is adjacent to a central node  $c' \in \mathcal{C}(T\langle X \rangle)$ , and  $c' \neq c$  (otherwise  $T\langle X \rangle = T'\langle X \rangle$ , a contradiction). Then,  $\mathcal{C}(T\langle X \rangle) = \{c, c'\}$  and so,  $T\langle X \rangle$  is a bistar (diameter-three subtree). Recall that both  $c, c'$  have a neighbour in  $X \setminus \{v\}$  by the hypothesis. This implies that node  $c'$  must have at least two neighbours in  $T_v^X$ ; moreover each such neighbour must be on the path between  $c'$  and a real node in  $X \setminus \{v\}$ . But then, since there are at least three branches in  $T\langle X' \rangle \setminus \{c'\}$ , we obtain as desired the existence of a diametral path in  $T\langle X' \rangle$  of which  $v$  is not an end. As a result,  $diam(T'\langle X' \rangle) \geq diam(T\langle X' \rangle)$ .

Altogether combined, the above proves the subclaim that  $diam(T'\langle X' \rangle) = diam(T\langle X' \rangle)$ . ◻

By Subclaim 4.3.1.2, we cannot change  $diam(T\langle X' \rangle)$  and so, it immediately implies that we cannot break Property P-4.1.4. The proof of this subclaim is actually more precise as it also shows that if  $T'\langle X' \rangle \neq T\langle X' \rangle$  then, these two subtrees have a common diametral pair  $(x, y)$ , for some  $x, y \neq v$ . In particular (Prop. P-3.1.2 of Lemma 3.1), we also have that:

$$\mathcal{C}(T'\langle X' \rangle) = \begin{cases} \mathcal{C}(T\langle X' \rangle) & \text{if } v \notin \mathcal{C}(T\langle X' \rangle) \\ (\mathcal{C}(T\langle X' \rangle) \setminus \{v\}) \cup \{\alpha_v\} & \text{otherwise.} \end{cases}$$

Furthermore, for any  $u \in X' \setminus \{v\}$ , if  $P_u$  denotes the  $u\mathcal{C}(T\langle X' \rangle)$ -path in  $T$  then, the  $u\mathcal{C}(T'\langle X' \rangle)$ -path in  $T'$  can be obtained from  $P_u$  by replacing  $v$  with  $\alpha_v$ . – This above analysis excludes  $v$ ; however,  $v$  cannot be  $X'$ -free because  $v \in X \subset X'$  and  $|X| \geq 3$ . – Hence, this

above mapping implies that we cannot break Properties P-4.1.2 and P-4.1.3. It also implies  $dist_{T'}(u, \mathcal{C}(T'\langle X' \rangle)) = dist_T(u, \mathcal{C}(T\langle X' \rangle))$  and so, by unimodality (Prop. P-3.1.1 of Lemma 3.1) we did not change the eccentricity of any vertex  $u \in X' \setminus \{v\}$ . Altogether combined, every  $X'$ -free vertex that was a leaf of  $T\langle X' \rangle$  is also a leaf of  $T'\langle X' \rangle$  with same eccentricity (Property P-4.1.1 cannot be broken).

Finally, we also obtain  $dist_{T'}(u, v) \leq 4$  for every  $u \in X'$  (i.e.,  $T'$  is a 4-Steiner root of  $G$ ).

- Otherwise,  $X \not\subseteq X'$  and we prove  $T\langle X' \rangle = T'\langle X' \rangle$ . To see that, suppose for the sake of contradiction  $T\langle X' \rangle \neq T'\langle X' \rangle$ . We first note this may be the case only if  $T\langle X' \rangle$  is not fully contained into  $Q_v$ .

**Subclaim 4.3.1.3.**  $T\langle X' \rangle$  must intersect  $T\langle X \rangle \setminus \{v\}$ .

*Proof.* Suppose by contradiction that  $T\langle X' \rangle \cap T\langle X \rangle = \{v\}$ . As we also assume  $T\langle X' \rangle$  is not fully contained into  $Q_v$ , we have  $T\langle X' \rangle \cap R_v = \emptyset$ . Then,  $T\langle X' \rangle$  must intersect some subtree  $T^{(sub)}$  of  $T \setminus T\langle X \rangle$  that is not in  $R_v$ . Let  $x \in X' \cap T^{(sub)}$ . By the definition of  $R_v$ , there must exist  $x' \in Real(T^{(sub)})$  such that:  $dist_T(x', X \setminus \{v\}) \leq 4$  (possibly,  $x = x'$ ). In particular, note that there exists a  $xx'$ -path that does not go by  $v$  in  $T$  (otherwise,  $x, x'$  would be in different subtrees of  $T \setminus T\langle X \rangle$ ). On one hand as we assume  $v$  to be  $X$ -free, there must exist a clique-intersection  $X_1 \supseteq X \cup \{x'\}$ . It implies  $dist_T(x', v) \leq 2$  because  $ecc_{T\langle X \rangle}(v) \geq 2$ . On the other hand, we also know that  $dist_T(x, v) \leq 4$ . By considering the median node of the triple  $x, x', v$  we so obtain  $dist_T(x, x') \leq 4$ . Let  $X_2$  be a clique-intersection that contains all of  $x, x', v$  (possibly,  $X_2 = X'$ ). There are two cases:

- If  $X \not\subseteq X_2$  then,  $X \cap X_2 = \{v\}$  (otherwise,  $v$  would be internally  $X$ -constrained). But in this case,  $v$  is  $(X, X_1, X_2)$ -sandwiched, a contradiction.
- Otherwise,  $X \subseteq X_2$ . However by the hypothesis  $X \not\subseteq X'$ , and so  $X \cap X' = \{v\}$  (otherwise,  $v$  would be internally  $X$ -constrained). It implies  $v$  is  $(X, X_2, X')$ -sandwiched, a contradiction.

As a result,  $T\langle X' \rangle$  must intersect  $T\langle X \rangle \setminus \{v\}$ . ◊

Finally, since  $v$  is  $X$ -free any node  $\beta \in V(T\langle X' \rangle) \cap (V(T\langle X \rangle) \setminus \{v\})$  must be Steiner. This leaves  $\beta \in \mathcal{C}(T\langle X' \rangle) \setminus \{v\}$  as the only possibility. Furthermore, since  $\beta$  is Steiner there must exist  $y \in X'$  such that the unique  $vy$ -path in  $T$  goes by  $\beta$ . However, this implies  $diam(T\langle X' \rangle \cup N_T[\beta]) = diam(T\langle X' \rangle)$ . We recall that there exists at least one leaf node  $u \in Real(N_T(\beta)) \setminus \{v\}$  by the hypothesis. Thus, by Property P-3.4.2 of Theorem 3.4 we have  $u, v \in X \cap X'$ , thereby contradicting that  $v$  is  $X$ -free.

The claim directly follows from this above case analysis. ◊

The proof is now divided into two main phases.

**Phase 1: Transformation into leaves** (see Fig. 6 for an illustration). Let  $X \in \mathcal{X}(G) \setminus \mathcal{K}(G)$ ,  $|X| \geq 3$  be fixed. We first transform  $T$  so that all the  $X$ -free vertices are leaves in  $T\langle X \rangle$ . Assume the existence of a  $X$ -free vertex  $v \in X$  that is not a leaf. Note that we have  $v \in \mathcal{C}(T\langle X \rangle)$  because

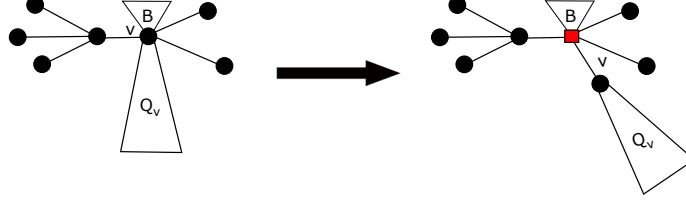


Figure 6: Forcing the  $X$ -free vertices as leaves.

$\text{diam}(T\langle X \rangle) \leq 3$ . In particular, every node in  $\mathcal{C}(T\langle X \rangle)$  is adjacent to a leaf in  $X \setminus \{v\}$ . We apply Operation 1 with  $c = \alpha_v$  (*i.e.*, the Steiner node substituting  $v$  in the intermediate tree  $T_v$ ). Since  $c = \alpha_v$  and so, in particular  $c$  is Steiner, both Conditions 4.3.1.a and 4.3.1.b of Claim 4.3.1 must hold. Overall, by Claim 4.3.1 we can repeat the above transformation until all the  $X$ -free vertices are leaves of  $T\langle X \rangle$ .

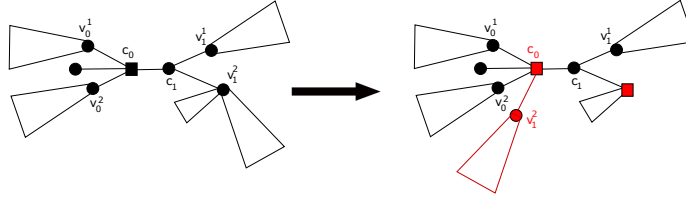


Figure 7: Grouping the  $X$ -free vertices on a same side.

**Phase 2: Grouping the  $X$ -free vertices** (see Fig. 7 for an illustration). After Phase 1, the properties of the theorem are true for any  $X \in \mathcal{X}(G)$  such that  $T\langle X \rangle$  is a star. Thus, from now on assume  $T\langle X \rangle$  is a bistar (diameter-three subtree). Write  $\mathcal{C}(T\langle X \rangle) = \{c_0, c_1\}$  and assume that each  $c_j$  is adjacent to two  $X$ -free vertices, denoted  $v_j^1, v_j^2$ . For any  $i \in \{1, 2\}$ , there is no vertex  $x_i \in \text{Real}(R_{v_j^i})$  such that  $\text{dist}_T(x_i, v_j^i) \leq 2$  (otherwise,  $\text{dist}_T(x_i, X \setminus \{v_j^i\}) \leq \text{dist}_T(x_i, v_j^{3-i}) \leq 4$ , a contradiction). More specifically, either  $\text{dist}_T(v_j^i, \text{Real}(R_{v_j^i})) \geq 4$ , or  $\text{dist}_T(v_j^i, \text{Real}(R_{v_j^i})) = 3$  but then  $c_j$  must be Steiner (otherwise,  $\text{dist}_T(x_i, X \setminus \{v_j^i\}) \leq \text{dist}_T(x_i, c_j) \leq 4$ , a contradiction). Furthermore, we can have  $\text{dist}_T(u, c_j) \leq 3$  only if  $X \subseteq N_G[u]$  (*i.e.*, because  $c_j$  is adjacent in  $T$  to at least two  $X$ -free vertices).

W.l.o.g., assume either  $\text{dist}_T(v_j^i, \text{Real}(R_{v_j^i})) \geq 4$  for any  $i, j$  or  $c_0$  is Steiner. If in addition both  $c_0, c_1$  are Steiner nodes (real nodes, resp.) then, we further assume w.l.o.g.  $c_0$  is adjacent to more  $X$ -free vertices than  $c_1$ . We apply Operation 1 for  $v = v_1^1$  and  $c = c_0$ . Note that by construction,  $c_0$  always satisfies Condition 4.3.1.a of Claim 4.3.1. Since we also have  $\text{dist}_T(u, c_0) \leq 3 \implies X \subseteq N_G[u]$ , Condition 4.3.1.b of Claim 4.3.1 is satisfied. Overall, we can repeat this transformation until  $X$  satisfies all the properties stated in the theorem, that does not impact the properties of the other clique-intersections  $X'$  by Claim 4.3.1.  $\square$

## 5 Step S.1: Construction of the clique-tree

Our main result in this section is that every chordal graph has a polynomial-time computable rooted clique-tree where some technical conditions must hold on the minimal separators that can

label an edge (Theorem 5.4). We will use these special properties of our rooted clique-tree in order to ensure that our main algorithm runs in polynomial time.

We start digressing on the motivations behind our construction. Given a rooted clique-tree  $T_G$  of  $G$  and an arbitrary maximal clique  $K_i$ , recall that  $G_i$  is the subgraph induced by all the maximal cliques in the subtree  $T_G^i$  rooted at  $K_i$ . Assume that  $T\langle S_i \rangle$ : the subtree induced by the minimal separator  $S_i := K_i \cap K_{p(i)}$ , is fixed. We ask for the number of possible distance profiles  $(\text{dist}_{T_i}(r, V_i \setminus S_i))_{r \in V(T\langle S_i \rangle)}$  over all possible *well-structured* 4-Steiner roots of  $G_i$  that contain  $T\langle S_i \rangle$ . By Theorem 4.1, one way to bound this number would be to force most vertices in  $S_i$  to be  $K_i$ -free in the subgraph  $G_i$ . Guided by this intuition, our construction aims at preventing  $S_i$  to contain, or to be contained, in a minimal separator of  $G_i$ . Since both objectives are conflicting, this results in a technical compromise.

We present a first, simpler construction in Section 5.1 that only reaches half of the goal. Then, we introduce the new notion of (weak) convergence, and we show its relationship with 4-Steiner powers (Section 5.2). We end up proving the main result of this part in Section 5.3.

## 5.1 A flat clique-tree

We start with an intermediate construction.

**Theorem 5.1.** *Given  $G = (V, E)$  chordal, we can compute in polynomial time a rooted clique-tree  $T_G$  such that, for any  $S_i = K_i \cap K_{p(i)}$  and for any child  $K_j$  of  $K_{p(i)}$ , there is no minimal separator of  $G_j$  that is contained into  $S_i$ .*

*Proof.* We modify an arbitrary clique-tree  $T_G$  of  $G$  until the property of the theorem is satisfied. Specifically, root  $T_G$  at some arbitrary maximal clique  $K_0$ . We consider all the minimal separators  $S \in \mathcal{S}(G)$  by decreasing size. Let  $K_S$  be incident to an edge in  $\bigcup_{S', S \subseteq S'} E_{S'}(T_G)$  and the closest possible to the root. We observe that  $K_S$  is the least common ancestor of all maximal cliques that are incident to an edge in  $\bigcup_{S', S \subseteq S'} E_{S'}(T_G)$ . Furthermore all edges in  $E_S(T_G)$  can be made incident to  $K_S$ , as follows. Assume there exists  $K_i K_{p(i)} \in E_S(T_G)$  such that  $K_S \notin \{K_i, K_{p(i)}\}$ . By the above observation,  $K_i, K_{p(i)}$  are into the subtree rooted at  $K_S$ . Since  $S \subseteq K_S \cap K_i$ ,  $S$  is contained into all the maximal cliques on the  $K_S K_i$ -path. In particular, we still obtain a clique-tree of  $G$  if we replace  $K_i K_{p(i)}$  by  $K_i K_S$  and in doing so,  $S \subseteq K_S \cap K_i \subseteq K_{p(i)} \cap K_i = S$ . Furthermore after this transformation,  $K_S$  became the new father node of the maximal clique  $K_i$  in  $T_G$ .

It now remains to prove that the gotten clique-tree  $T_G$  satisfies the conditions of the theorem. Suppose for the sake of contradiction that there exist  $i > 0$  and  $S_k \subseteq S_i$  a minimal separator of  $G_j$ , where  $K_j$  is a child of  $K_{p(i)}$  (possibly,  $i = j$ ). We observe that after we processed  $S_i$ , all edges in  $E_{S_i}(T_G)$  must label an edge between  $K_{S_i} = K_{p(i)}$  and its children nodes. Therefore, our transformation ensures  $S_k \neq S_i$ , *i.e.*,  $S_k \subset S_i$ . Since the subtree rooted at  $K_j$  is a rooted clique-tree of  $G_j$ , there must exist some edge  $K_k K_{p(k)}$  in this subtree such that  $K_k \cap K_{p(k)} = S_k$ . However, since we consider minimal separators by increasing size, the edge  $K_i K_{p(i)}$  should already exist when we process  $S_k$ . It implies that the maximal clique  $K_{S_k}$  to which we connected all edges in  $E_{S_k}(T_G)$  should be an ancestor of  $K_{p(i)}$ , that is a contradiction.  $\square$

## 5.2 Weak convergence

Unfortunately, the “flat” clique-tree of Theorem 5.1 does not prevent the case when a minimal separator  $S_i$  is contained into a minimal separator of  $G_i$ . As a new step toward our final construction,

we now introduce the following notions:

**Definition 5.1.** Given a clique-tree  $T_G$  of  $G = (V, E)$ , we say that a minimal separator  $S$  is *weakly  $T_G$ -convergent* if there exists some maximal clique  $K_S$  that is incident to all edges in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$ .  $S$  is termed  *$T_G$ -convergent* if it is weakly  $T_G$ -convergent and the maximal clique  $K_S$  is also incident to all edges in  $E_S(T_G)$ .

In order to motivate Definition 5.1, in what follows are two observations on the relationships between clique-trees, minimal separators and 4-Steiner roots:

**Lemma 5.2.** *Given  $G = (V, E)$  and  $T$  any 4-Steiner root of  $G$ , let  $K_i \in \mathcal{K}(G)$  and let  $X \subset K_i$  be a clique-intersection. If  $T\langle X \rangle$  is a bistar then,  $\mathcal{C}(T\langle K_i \rangle) \subset \mathcal{C}(T\langle X \rangle)$ .*

*In particular, there are exactly two maximal cliques that contain  $X$ .*

*Proof.* We have by Theorem 3.4  $\text{diam}(T\langle K_i \rangle) > \text{diam}(T\langle X \rangle)$ , and so,  $\text{diam}(T\langle K_i \rangle) = 4$ . In particular, write  $\mathcal{C}(T\langle K_i \rangle) = \{c_i\}$ . Every component in  $T\langle K_i \rangle \setminus \{c_i\}$  has diameter at most two, thereby implying  $c_i \in V(T\langle X \rangle)$ . Furthermore since  $\text{ecc}_{T\langle X \rangle}(c_i) \leq \text{rad}(T\langle K_i \rangle) = 2$ ,  $c_i$  cannot be a leaf of  $T\langle X \rangle$ . Equivalently,  $c_i \in \mathcal{C}(T\langle X \rangle)$ . By Lemma 3.5, there can be no two maximal cliques  $K_i, K_j \in \mathcal{K}(G)$  such that  $\mathcal{C}(T\langle K_i \rangle) = \mathcal{C}(T\langle K_j \rangle)$ . Therefore, the above implies that  $X$  can only be contained in at most two maximal cliques. Finally, since  $X$  is not a maximal clique, it is contained into exactly two maximal cliques.  $\square$

**Lemma 5.3.** *Given  $G = (V, E)$  and  $T$  any 4-Steiner root of  $G$ , let  $S \in \mathcal{S}(G)$ . If  $T\langle S \rangle$  is a non-edge star then,  $S$  is weakly  $T_G$ -convergent for any clique-tree  $T_G$  of  $G$ .*

*Proof.* We may assume that  $S$  is strictly contained into at least one minimal separator  $S'$  for otherwise there is nothing to prove. By Theorem 3.4,  $T\langle S' \rangle$  is a bistar and  $S'$  must be inclusion wise maximal in  $\mathcal{S}(G)$ . This implies  $\mathcal{C}(T\langle S \rangle) \subset \mathcal{C}(T\langle S' \rangle)$ . Furthermore, it follows from Lemma 5.2 that  $S'$  must be contained into exactly two maximal cliques  $K_i, K_j$  and  $\mathcal{C}(T\langle K_i \rangle) \cup \mathcal{C}(T\langle K_j \rangle) = \mathcal{C}(T\langle S' \rangle)$ . In particular, we may assume w.l.o.g. that  $\mathcal{C}(T\langle S \rangle) = \mathcal{C}(T\langle K_i \rangle)$ . But then, still by Lemma 5.2, any minimal separator  $S''$  that strictly contains  $S$  must be contained into  $K_i$  and exactly one other maximal clique  $K_{S''}$ . By Theorem 2.4, the latter implies  $K_i K_{S''} \in E(T_G)$  and  $K_i \cap K_{S''} = S''$ .  $\square$

Roughly, in order to prove Theorem 5.4 we will slightly modify the construction of Theorem 5.1 so as to force weak convergence to imply convergence. In doing so we will obtain that for a fixed minimal separator  $S$ , we may encounter some “inclusion issue” between  $S$  and another minimal separator  $S'$  at most *once*. We will show in Sec. 7 that this above “local” property of our clique-tree is enough in order to bound the number of possible distance profiles at each node of the clique-tree by a polynomial.

### 5.3 The final construction

The remaining of this section is now devoted to prove the following technical result:

**Theorem 5.4.** *Given  $G = (V, E)$  chordal, we can compute in polynomial time a rooted clique-tree  $T_G$  where the following conditions are true for any  $S_i := K_i \cap K_{p(i)}$ ,  $i > 0$ :*

- (Prop. P-5.4.1.) *If  $S_i$  is weakly  $T_G$ -convergent and  $|S_i| \geq 3$  then,  $S_i$  is  $T_G$ -convergent;*



- (Prop. P-5.4.2.) Any minimal separator of  $G_i$  that is contained into  $S_i$  is  $T_G$ -convergent, it has at least three vertices and it is strictly contained into at least one minimal separator of  $G_i$ .

We stress that in the above two statements, 3 can be replaced by any positive integer  $q$ . However, please note that such a change would affect both properties of the theorem.

*Proof.* We start describing the algorithm before proving its correctness. In what follows, the minimal separators of  $G$  are considered *in the same order* for every phase of the algorithm.

**Phase 1** Let  $T_G$  be an arbitrary *unrooted* clique-tree. We consider all the minimal separators  $S \in \mathcal{S}(G)$  by decreasing size. Let  $S^1, S^2, \dots, S^p$  be the list of all minimal separators containing  $S$  ordered by decreasing size. – In particular,  $S^p = S$ , and if several such minimal separators  $S^i$  have equal size then, we break ties according to the ordering in which they are considered. – For every maximal clique  $K$ , we can define a binary vector  $\vec{v}_S^K = (\delta_{S^1}^K, \delta_{S^2}^K, \dots, \delta_{S^p}^K)$ : where  $\delta_{S^i}^K = 1$  if and only if  $K$  is incident to an edge in  $E_{S^i}(T_G)$ . Then, we choose  $K_S^1 \in \mathcal{K}(G)$  such that  $\vec{v}_S^K(K_S^1)$  is lexicographically maximal. While there exists an edge  $KK' \in E_S(T_G)$  such that  $K_S^1 \notin \{K, K'\}$ , we do as follows. W.l.o.g.,  $K$  is on the  $K_S^1 K'$ -path. We replace the edge  $KK'$  by  $K_S^1 K'$ . Doing so, all edges in  $E_S(T_G)$  are now incident to  $K_S^1$ .

**Phase 2** We root  $T_G$  at some arbitrary maximal clique  $K_0$ . Then, we consider all the minimal separators  $S \in \mathcal{S}(G)$  by decreasing size. Let  $K_S^2$  be, under the following conditions, the closest possible to the root:

- $K_S^2$  is incident to an edge in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$ ;
- and if  $|S| \geq 3$  and  $S$  is  $T_G$ -convergent then,  $K_S^2$  is incident to all edges in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$ .

Note that  $K_S^2$  is always an ancestor of  $K_S^1$ . In particular if  $|S| \geq 3$  and  $S$  is  $T_G$ -convergent then,  $K_S^2$  is either  $K_S^1$  or its parent node (with the latter being possible only if  $|\bigcup_{S', S \subset S'} E_{S'}(T_G)| \leq 1$ ). Then, we use the same operation as in the proof of Theorem 5.1 in order to make all edges in  $E_S(T_G)$  incident to  $K_S^2$ .

**Phase 3** We end up considering all the minimal separators  $S \in \mathcal{S}(G)$  by decreasing size. If  $|S| \geq 3$  and  $S$  is *not*  $T_G$ -convergent then, we search for a child  $K_S^3$  of  $K_S^2$  that is incident to all edges in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$ . For simplicity of our analysis, we will also set  $K_S^3 = K_S^2$  when  $|S| \leq 2$ , or  $S$  is already  $T_G$ -convergent, or there is no child node of  $K_S^2$  which satisfies the desired property. Then, we consider all the children  $K$  of  $K_S^2$ ,  $K \neq K_S^3$ , such that  $K \cap K_S^2 = S$ . We replace the edge  $KK_S^2$  by  $KK_S^3$ .

Before proving correctness of this above algorithm, we want to emphasize two of its main invariants. In what follows, let  $S \in \mathcal{S}(G)$  be of size at least three.

- (Inv. I1) Assume that at the time we considered  $S$  during Phase 2,  $S$  was  $T_G$ -convergent. We observe that our choice for  $K_S^2$  preserved this property. In particular,  $S$  is  $T_G$ -convergent in the final clique-tree  $T_G$  that we output.
- (Inv. I2) If  $K_S^3 \neq K_S^2$  (or equivalently, we modify  $E_S(T_G)$  during Phase 3) then, we also claim that  $S$  is  $T_G$ -convergent in the final clique-tree  $T_G$  that we output. For that, by the definition of Phase 3 it suffices to prove that the edge between  $K_S^2$  and its parent node (if any) cannot

be labelled by  $S$ . According to Invariant I1,  $S$  was not  $T_G$ -convergent when it was considered during Phase 2 (otherwise, it should have stayed so until we considered  $S$  during Phase 3, and so we could have not modified  $E_S(T_G)$ ). Therefore by the definition of Phase 2,  $K_S^2$  is the least common ancestor in  $T_G$  of all the maximal cliques that are incident to an edge in  $\bigcup_{S', S \subseteq S'} E_{S'}(T_G)$ , thereby proving the claim.

We will often use these above two invariants in the remaining of the proof. We finally prove that both properties of the theorem are true for  $T_G$ .

**Proof of Property P-5.4.1.** Suppose for the sake of contradiction  $S \in \mathcal{S}(G)$  is weakly  $T_G$ -convergent but not  $T_G$ -convergent, and  $|S| \geq 3$ . For clarity, we divide the proof into small claims.

**Claim 5.4.1.** *After Phase 2 was completed,  $S$  was not  $T_G$ -convergent.*

*Proof.* Otherwise, every minimal separator  $S' \supset S$  should also be  $T_G$ -convergent after Phase 2. But then, the edge-set  $\bigcup_{S', S \subseteq S'} E_{S'}(T_G)$  could not be modified during Phase 3. This would imply  $S$  is  $T_G$ -convergent, a contradiction.  $\diamond$

**Claim 5.4.2.** *From the time we considered  $S$  during Phase 2 until the end of the algorithm,  $K_S^2$  is the least common ancestor of all maximal cliques incident to an edge in  $\bigcup_{S', S \subseteq S'} E_{S'}(T_G)$ .*

*Proof.* We recall that whenever we consider a minimal separator of size at least three during Phase 2, we preserve the property of being  $T_G$ -convergent (Invariant I1). Therefore, Claim 5.4.1 implies that at the time we considered  $S$  during Phase 2,  $S$  was not  $T_G$ -convergent. In particular, we chose  $K_S^2$  to be the least common ancestor of all maximal cliques incident to an edge in  $\bigcup_{S', S \subseteq S'} E_{S'}(T_G)$ . This must stay so until the end of the algorithm because during Phase 3, the edges in  $E_S(T_G)$  were not modified (Invariant I2) and all the edges in  $\bigcup_{S', S \subseteq S'} E_{S'}(T_G)$  could only be made incident to a descendant of  $K_S^2$ .  $\diamond$

**Claim 5.4.3.** *After Phase 2 was completed,  $K_S^2$  was not incident to any edge in  $\bigcup_{S', S \subseteq S'} E_{S'}(T_G)$ .*

*Proof.* Suppose for the sake of contradiction that  $K_S^2$  was incident to such an edge. We recall that  $K_S^2$  is the least common ancestor of all maximal cliques incident to an edge in  $\bigcup_{S', S \subseteq S'} E_{S'}(T_G)$  (Claim 5.4.2). Therefore, the only possibility left for  $S$  being weakly  $T_G$ -convergent but not  $T_G$ -convergent is that, during Phase 3, all edges in  $\bigcup_{S', S \subseteq S'} E_{S'}(T_G)$  that were incident to  $K_S^2$  were made incident to one of its children  $K$ . But then, during Phase 3 all edges in  $E_S(T_G)$  should have been made incident to  $K$ . This contradicts our assumption that  $S$  is not  $T_G$ -convergent.  $\diamond$

In what follows, recall that  $S^1$  is a largest minimal separator containing  $S$ . Since  $S$  is weakly  $T_G$ -convergent but not  $T_G$ -convergent, we have  $S \neq S^1$ .

**Claim 5.4.4.**  $K_{S^1}^3 = K_{S^1}^2 = K_S^1$ .

*Proof.* The proof that  $K_{S^1}^3 = K_{S^1}^2$  easily follows from the maximality of  $S^1$ . So, we will only prove  $K_{S^1}^2 = K_S^1$ . By maximality of  $\bar{v}_S^2(K_S^1)$ ,  $K_S^1$  was incident to an edge in  $E_{S^1}(T_G)$  after Phase 1 was completed. Therefore, at the time we considered  $S^1$  during Phase 2, either  $K_{S^1}^2 = K_S^1$  or  $K_{S^1}^2$  was a strict ancestor of  $K_S^1$ . But then in the second situation,  $K_S^2$  should have been incident to an

edge in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$  after Phase 2 was completed (*i.e.*, because  $K_S^2$  is an ancestor of  $K_{S^1}^2$  by Claim 5.4.2). The latter would contradict Claim 5.4.3.  $\diamond$

**Claim 5.4.5.** *From the time we considered  $S$  during Phase 2 until the end of the algorithm,  $K_S^2$  is the father node of  $K_S^1$ . Moreover,  $K_S^1 \cap K_S^2 = S$ .*

*Proof.* By Claim 5.4.3,  $K_S^2$  cannot be incident to any edge in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$ . We also know that all edges in  $E_S(T_G)$  are incident to  $K_S^1$ . Thus, either  $K_S^1 = K_S^2$  or  $K_S^2$  is the father node of  $K_S^1$ . But then in the former case,  $K_S^2$  would be incident to an edge in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$  (Claim 5.4.4), a contradiction. As a result,  $K_S^2$  is the father node of  $K_S^1$ , and then we must have  $K_S^1 \cap K_S^2 = S$ .  $\diamond$

Finally, since we did not modify  $E_S(T_G)$  during Phase 3 (cf. Invariant I2), not all edges in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$  can be incident to the same child node of  $K_S^2$ . Then, by Claim 5.4.5, not all edges in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$  can be incident to  $K_S^1$ . By Claim 5.4.4 we get that all the edges in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$  must be incident to some child  $K$  of  $K_S^1$ .

In particular, there must be a second largest minimal separator  $S^2$  that strictly contains  $S$  (otherwise it would follow from Claim 5.4.4 that at the time we considered  $S$  during Phase 2,  $S$  was  $T_G$ -convergent, thereby contradicting Claim 5.4.1). By maximality of  $|S^2|$  we did not modify the edges in  $E_{S^2}(T_G)$  during Phase 3, or equivalently  $K_{S^2}^2 \neq K_S^1$ . Therefore,  $K_{S^2}^2 = K$ . The above implies that at the time we considered  $S^2$  during Phase 2,  $K$  was already incident to an edge in  $E_{S^2}(T_G)$  (otherwise there would exist  $S' \supset S^2$  such that  $K$  was incident to an edge in  $E_{S'}(T_G)$ , and since we chose  $K$  instead of  $K_S^1$ , by Claim 5.4.4 we can always assume  $S' \neq S^1$ , a contradiction).

Overall, after Phase 1 was completed,  $K$  was incident to an edge in  $E_{S^1}(T_G)$  (consequence of Claim 5.4.4), and to an edge in  $E_{S^2}(T_G)$ . However, this should have already been the case when we considered  $S$  during Phase 1. Furthermore after Phase 1,  $K_S^1$  was not incident to any edge in  $E_{S^2}(T_G)$  (*i.e.*, because  $K_{S^2}^2$ , and so  $K_{S^2}^1$ , is a strict descendant of  $K_S^1$ , and we have  $K_{S^2}^2 \cap K_S^1 = S^1 \neq S^2$  by Claim 5.4.4). The latter contradicts the maximality of  $\vec{v}_S(K_S^1)$ .

**Proof of Property P-5.4.2.** Finally, for any  $i > 0$ , let us assume the existence of a minimal separator  $S$  of  $G_i$  that is contained into  $S_i$ . In particular,  $K_S^3$  is a descendant of  $K_i$ . There are two cases:

- Case  $K_S^3 \neq K_S^2$ . Then, as we modified  $E_S(T_G)$  during Phase 3,  $|S| \geq 3$  and  $S$  is  $T_G$ -convergent (cf. Invariant I2). It implies  $K_S^3 = K_i$  because we assume  $S \subseteq S_i$ . In this situation,  $K_S^2 = K_{p(i)}$ , and there is at least one edge in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$  to which  $K_{p(i)}$  is not incident (otherwise at the time we considered  $S$  during Phase 3,  $S$  would have already been  $T_G$ -convergent). In particular, since there is at least one edge in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$  that is different than  $K_i K_{p(i)}$ ,  $S$  is strictly contained into a minimal separator of  $G_i$ .
- Case  $K_S^3 = K_S^2$ . Note that  $K_{p(i)}$  is incident to an edge in  $E_{S_i}(T_G)$ , and so to an edge in  $\bigcup_{S', S \subseteq S'} E_{S'}(T_G)$ . However, at the time we considered  $S$  during Phase 2, we chose a  $K_S^2$  that was *not* an ancestor of  $K_{p(i)}$  (otherwise, this should have stayed so during Phase 3). It implies that  $|S| \geq 3$  and  $S$  was  $T_G$ -convergent. In particular, after Phase 2 was completed all the minimal separators  $S'$  containing  $S$  were  $T_G$ -convergent. Therefore the set  $\bigcup_{S', S \subseteq S'} E_{S'}(T_G)$  was not modified during Phase 3. Since  $S \subseteq S_i$  and  $K_S^2 \neq K_{p(i)}$ , we so obtain  $K_S^2 = K_i$ . Furthermore,  $K_{p(i)}$  was not incident to all edges in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$  (otherwise, we could

have chosen  $K_S^2 = K_{p(i)}$ ). Overall, it implies that  $S$  is  $T_G$ -convergent and, since  $K_{p(i)}$  is not incident to all edges in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$ ,  $S$  is strictly contained into a minimal separator of  $G_i$ .

□

## 6 Step S.2: A family of subtrees for the Clique-Intersections

For every clique-intersection  $X \in \mathcal{K}(G) \cup \mathcal{S}(G)$ , we aim at computing a polynomial-size representation of the family of all possible subtrees  $T\langle X \rangle$  that we could encounter in a well-structured 4-Steiner root of  $G$ . – Our approach also works for weak minimal separators, however this is not needed for our algorithm. – Correctness of this part mostly follows from Theorem 4.1.

In what follows, let  $T_G$  be an arbitrary rooted clique-tree of  $G$  (*i.e.*, not necessarily the one computed in Sec. 5). We first focus on the case of minimal separators in Section 6.1, before extending our results to the maximal cliques that are either leaf-nodes or internal nodes of  $T_G$  (Sections 6.2 and 6.3, respectively).

### 6.1 Case of Minimal Separators

We first prove that for any minimal separator  $S$ , the family  $\mathcal{T}_S$  of all the potential subtrees  $T\langle S \rangle$  has polynomial size. Moreover, we can enumerate all possible subtrees  $T\langle S \rangle$  in polynomial time. This result will be the cornerstone of all our other constructions in this part.

**Theorem 6.1.** *Let  $\mathcal{S}(G)$  be the set of all minimal separators in  $G = (V, E)$ . In  $\mathcal{O}(n^5 m)$ -time we can construct a collection  $(\mathcal{T}_S)_{S \in \mathcal{S}(G)}$  such that, for any well-structured 4-Steiner root  $T$  of  $G$ , and for any  $S \in \mathcal{S}(G)$ ,  $T\langle S \rangle$  is Steiner-equivalent to some subtree in  $\mathcal{T}_S$ .*

Before proving this above theorem, let us describe the main difficulty we had to face on. Roughly, given  $S \in \mathcal{S}(G)$  the difficulty in generating  $\mathcal{T}_S$  comes from the bisters, as a brute-force generation of all possibilities would take time exponential in  $|S|$ . In order to remedy to that issue we use the fact that in a well-structured 4-Steiner root of  $G$ ,  $S$ -free vertices are leaves of such a bistar with all of them, except maybe one, adjacent to the same central node. For a fixed placement of the  $S$ -constrained vertices, this only gives us  $\mathcal{O}(|S|)$  possibilities in order to place the  $S$ -free vertices. Overall, we reduce the number of possible bisters to an  $\mathcal{O}(|S|^5)$ .

*Proof.* Let  $S \in \mathcal{S}(G)$  be fixed. We will construct  $\mathcal{T}_S$  in  $\mathcal{O}(|S|^6)$ -time. Since  $\max\{|S| \mid S \in \mathcal{S}(G)\} = \mathcal{O}(n)$  and  $\sum_{S \in \mathcal{S}(G)} |S| = \mathcal{O}(m)$  [BP93], the latter will prove the result.

**Case  $\text{diam}(T\langle S \rangle) \leq 2$ .** Let us start with some easy cases. If  $|S| = 1$  then, it suffices to add a single-node tree to  $\mathcal{T}_S$ . Similarly, if  $|S| = 2$  then, by Theorem 3.4,  $S$  must induce a path of length at most  $k - 1 = 3$  in any 4-Steiner root of  $G$  with its two ends being the vertices of  $S$ . This gives only  $\mathcal{O}(1)$  possibilities to put into  $\mathcal{T}_S$ . Thus, from now on assume  $|S| \geq 3$ . Given any 4-Steiner root  $T$  of  $G$ , by Theorem 3.4 the subtree  $T\langle S \rangle$  can only be a star or a bistar (but the latter only if  $S$  is inclusion wise maximal in  $\mathcal{S}(G)$ ). Furthermore in the former case, all leaves in the star  $T\langle S \rangle$  must be in  $S$ , and the center node can either be in  $S$  or Steiner. Overall, this gives  $\mathcal{O}(|S|)$  possibilities of stars to put into  $\mathcal{T}_S$ , and so, this takes  $\mathcal{O}(|S|^2)$ -time.

**Case  $\text{diam}(T\langle S \rangle) = 3$ .** We end up focusing on the case where  $|S| \geq 3$  and  $T\langle S \rangle$  may be a bistar. In what follows, the two central nodes of such a bistar will be always denoted by  $\mathcal{C}(T\langle S \rangle) = \{c_0, c_1\}$ .

We will introduce the following additional terminology. A *heavy part* of  $S$  is any clique-intersection  $X \subset S$  such that  $|X| \geq 3$ . A *light part* of  $S$  is any clique-intersection  $X \subset S$  such that  $|X| = 2$ . We prove the following intermediate claim (also used in other parts of the paper):

**Claim 6.1.1.** *If  $T\langle S \rangle$  is a bistar and  $\mathcal{C}(T\langle S \rangle)$  is a light part then, there is a heavy part that strictly contains  $\mathcal{C}(T\langle S \rangle)$ .*

*Proof.* Suppose by contradiction  $X = \{c_0, c_1\}$  is a light part and no heavy part contains it. Let  $K_i$  be any maximal clique such that  $X \subseteq K_i$  but  $S \not\subseteq K_i$ . Such a  $K_i$  always exists since otherwise, taking the intersection of  $S$  with all the maximal cliques that contains  $X$ , one would obtain  $S = X$ , a contradiction. In this situation,  $X \subseteq K_i \cap S$ , and so  $X = K_i \cap S$  since there is no heavy part containing  $X$ . Furthermore we have  $K_i \not\subseteq S$ . Hence, there exists  $j \in \{0, 1\}$  such that  $c_j$  has a neighbour in  $T\langle K_i \rangle \setminus S$  (possibly, a Steiner node). By applying Lemma 3.2 to  $N_T[c_j]$  and  $T\langle K_i \rangle$ , we obtain:

$$\text{diam}(N_T[c_j] \cup T\langle K_i \rangle) = \text{diam}(T\langle K_i \rangle) \leq 4.$$

In particular,  $K_i \cup \text{Real}(N_T[c_j])$  is a clique of  $G$ . By maximality of  $K_i$ ,  $\text{Real}(N_T[c_j]) \subseteq K_i$ . However, there is at least one leaf in  $\text{Real}(N_T[c_j]) \setminus X$ , that implies  $X \subset K_i \cap S$ , a contradiction.  $\diamond$

Then, we divide the proof in two subcases. We stress that both subcases only depend on the clique-arrangement of  $G$ , and are independent of any 4-Steiner root  $T$  — we use the existence of such a root only for proving correctness of our construction.

- We first consider the particular subcase when there exists a heavy part  $X \subset S$ . In this situation,  $X \subseteq N_T[c_0]$  or  $X \subseteq N_T[c_1]$  for any well-structured 4-Steiner root  $T$  of  $G$  where  $T\langle S \rangle$  is a bistar. Moreover, by Property P-3.4.2 of Theorem 3.4 either  $\text{Real}(N_T[c_0]) = X$  or  $\text{Real}(N_T[c_1]) = X$ . Therefore, we can start choosing among  $\mathcal{O}(|X|)$  possibilities the star induced by  $X$  in  $T$ . W.l.o.g.,  $c_0$  is the center of this star. The other center  $c_1$  must be either a Steiner node adjacent to  $c_0$  (in  $T$ ) or any vertex in  $X \setminus \{c_0\}$ . Hence, there are also  $\mathcal{O}(|X|)$  possibilities for  $c_1$ . Finally, since we have  $\text{Real}(N_T[c_0]) = X$  all the nodes in  $S \setminus X$  must be leaves adjacent to  $c_1$ . Overall, this gives  $\mathcal{O}(|S|^2)$  possibilities of bistars to put into  $\mathcal{T}_S$ , and so, this takes  $\mathcal{O}(|S|^3)$ -time<sup>1</sup>.
- From now on we assume that there is no heavy part. Following Definition 4.1 we process the vertices in  $S$  according to the other clique-intersections in which they are contained.

**Processing the internally  $S$ -constrained vertices.** For any well-structured 4-Steiner root  $T$  where  $T\langle S \rangle$  is a bistar, by Claim 6.1.1,  $\mathcal{C}(T\langle S \rangle) = \{c_0, c_1\}$  is not a light part. Furthermore, given any light part  $X \subset S$ , we can prove that either  $X$  induces an edge of  $T\langle S \rangle$ , or  $T\langle X \rangle$  is a non-edge star and  $X = \text{Real}(N_T[c_j])$  for some  $j$  (this also follows from Property P-3.4.2 of Theorem 3.4). In this situation, we construct the subgraph  $I_2$ : of which the vertices are the internally  $S$ -constrained vertices, and the edges are the light parts.

**Claim 6.1.2.** *If  $G$  has a 4-Steiner root  $T$  where  $T\langle S \rangle$  is a bistar then,  $I_2$  has at most two connected components. Specifically:*

---

<sup>1</sup>We will actually show in Lemma 6.3 this number of potential bistars can be reduced. However, we choose not to include this improvement in this part of the proof in order to keep it as simple as possible.

1. Either  $I_2$  is connected, there is a unique light part  $X$  such that  $T\langle X \rangle$  is a non-edge star, and all other light parts must be edges incident to the unique central node in  $\mathcal{C}(T\langle S \rangle) \setminus \mathcal{C}(T\langle X \rangle)$ .
2. Or the connected components of  $I_2$  induce node-disjoint stars in  $T\langle S \rangle$ .

*Proof.* The result is a consequence of the following case analysis:

- Assume there is a light part  $X \subset S$  such that  $T\langle X \rangle$  is a non edge star and  $Real(N_T[c_j]) = X$ . Two different situations might occur:
  1. Situation # 1 (see Fig. 8): There is a light part  $X' \neq X$  intersecting  $X$ . Since the vertex in  $X' \cap X$  cannot be a leaf adjacent to  $c_j$  (otherwise, we should have  $X' \subseteq Real(N_T[c_j]) = X$ , a contradiction), this must be  $c_{1-j}$ . In particular, for any light part  $X' \neq X$ ,  $T\langle X' \rangle$  must be an edge between  $c_{1-j}$  and a leaf (thereby implying  $X' \cap X \neq \emptyset$ ).

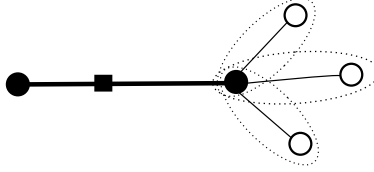


Figure 8: Situation 1: the subtree  $T\langle X \rangle$  is drawn in bold. There are 3 other light parts represented by dashed ellipses.

2. Situation # 2 (see Fig. 9): There is no other light part  $X' \neq X$  intersecting  $X$ . Since  $T\langle S \rangle \setminus N_T[c_j]$  is an independent set, any light part  $X' \subset S$  that does not intersect  $X$  cannot be an edge. We so deduce that if such a  $X'$  exists then,  $Real(N_T[c_{1-j}]) = X'$ , and so, there are no other light part in  $S$  than  $X$  and  $X'$ .

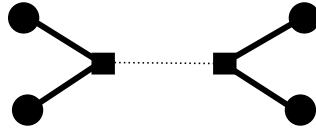


Figure 9: Situation 2: the subtrees  $T\langle X \rangle$  and  $T\langle X' \rangle$  are drawn in bold.

- Otherwise, each light part is an edge of  $T\langle S \rangle$  that contains either  $c_0$  or  $c_1$ , but not both. Therefore, there is a one-to-one mapping between the connected components of  $I_2$  and the nonempty sets among  $Real(N_T[c_0]) \setminus \{c_1\}, Real(N_T[c_1]) \setminus \{c_0\}$ . See Fig. 10 for an illustration.

◇

Overall, Claim 6.1.2 reduces the placement of internally  $S$ -constrained vertices to the construction of two stars in parallel, thereby giving  $\mathcal{O}(|S|^2)$  different possibilities.

**Processing the  $(S, X_1, X_2)$ -sandwiched vertices.** Then, we consider all the other  $S$ -constrained vertices. By Lemma 5.2 in the previous Section,  $S$  is strictly contained into

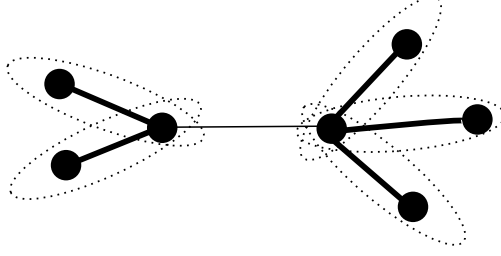


Figure 10: A case where all light parts must be edges.

exactly two maximal cliques, that we denote  $K$  and  $K'$ . In particular, if  $v \in S$  is  $(S, X_1, X_2)$ -sandwiched then, either  $X_1 = K$  or  $X_1 = K'$ . In the former case we call  $v$   $K$ -dependent, and in the latter case we call it  $K'$ -dependent.

The following Claim 6.1.3 shows that it defines two equivalence classes:

**Claim 6.1.3.** *Assume that  $G$  has a 4-Steiner root  $T$  where  $T\langle S \rangle$  is a bistar, and let  $v \in S$  be not internally  $S$ -constrained. If  $v$  is  $K$ -dependent then, either  $v \in \mathcal{C}(T\langle S \rangle) \cap \mathcal{C}(T\langle K \rangle)$ , or  $v$  is a leaf of  $T\langle S \rangle$  that is adjacent in  $T$  to the unique node in  $\mathcal{C}(T\langle K \rangle)$ .*

*In particular, a vertex  $v \in S$  cannot be both  $K$ -dependent and  $K'$ -dependent (unless it is also internally  $S$ -constrained).*

*Proof.* By the proof of Lemma 5.2,  $\mathcal{C}(T\langle S \rangle) = \{c_0, c_1\}$  where  $\mathcal{C}(T\langle K \rangle) = \{c_0\}$  and  $\mathcal{C}(T\langle K' \rangle) = \{c_1\}$ . Now, suppose by contradiction  $v$  is either  $c_1$  or a leaf adjacent to  $c_1$ .

Let  $X_2 \in \mathcal{X}(G)$  be such that  $v$  is  $(S, K, X_2)$ -sandwiched. As we have  $X_2 \cap S = \{v\}$ ,  $X_2$  is contained into at least one maximal clique  $K''$  that does not contain  $S$ . Furthermore, we can assume w.l.o.g.  $K'' \cap S = \{v\}$  since otherwise  $v$  is internally  $S$ -constrained, a contradiction. Thus from now on we will assume  $X_2 = K''$  is a maximal clique that does not contain  $S$ .

Let  $x \in (K \cap X_2) \setminus \{v\}$ . On one hand,  $x \notin K'$  (because  $x \notin S$ ) and so,  $dist_T(x, c_1) \geq 3$ . On the other hand,  $dist_T(x, c_0) \leq 2$  because  $x \in K$ . Altogether combined,  $dist_T(x, c_0) = 2$  and the unique  $xc_1$ -path in  $T$  goes by  $c_0$ . Note that it also implies that the unique  $xv$ -path in  $T$  goes by the edge  $c_0c_1$ , hence  $dist_T(x, v) \in \{3, 4\}$ .

Furthermore, let  $y \in X_2 \setminus K$ , that exists since we assume  $X_2$  to be a maximal clique. We prove as a subclaim  $y \notin K'$ . Indeed, by Theorem 2.4 there is an edge labeled by  $S$  in any clique-tree of  $G$ . This edge must be  $KK'$ , that implies there can be no edge in  $G$  between  $K \setminus K'$  and  $K' \setminus K$ . Since we have  $xy \in E(G)$ , this proves our subclaim that  $y \notin K'$ . In particular, we also have  $dist_T(y, c_1) \geq 3$ . But then, since we must also have  $dist_T(x, y) \leq 4$ , the  $yv$ -path in  $T$  also goes by the edge  $c_0c_1$  (otherwise,  $dist_T(x, y) = dist_T(x, c_1) + dist_T(c_1, y) \geq 6$ ). Since  $dist_T(y, c_0) \geq 3$  (because  $y \notin K$ ) we so obtain that  $dist_T(y, c_0) = 3$  and  $c_1 = v$ .

However, there exists  $z \in S$  a leaf adjacent to  $c_0$ . In particular, there exists a clique-intersection  $X$  that contains all of  $x, y, z, v$ . As we assume  $v$  is not internally  $S$ -constrained,  $S \subset X$ . This is a contradiction because we cannot have  $X = K$  nor  $X = K'$ .  $\diamond$

Given a fixed placement of internally  $S$ -constrained vertices, by Claim 6.1.3 we can generate all possible placements of the “sandwiched” vertices, as follows. We choose the central node

in  $\mathcal{C}(T\langle K \rangle)$  and the central node in  $\mathcal{C}(T\langle K' \rangle)$  (there are  $\mathcal{O}(|S|^2)$  possibilities). Then, all remaining  $K$ -dependent vertices, resp.  $K'$ -dependent, must be added as leaves adjacent to the unique node in  $\mathcal{C}(T\langle K \rangle)$ , resp. to the unique central node in  $\mathcal{C}(T\langle K' \rangle)$ . Overall, we have  $\mathcal{O}(|S|^2) \times \mathcal{O}(|S|^2) = \mathcal{O}(|S|^4)$  possibilities for positioning the  $S$ -constrained vertices.

**Processing the  $S$ -free vertices.** However, each such a possibility does not quite define a potential bistar for  $S$  as we also need to position the  $S$ -free vertices. By Theorem 4.1, we can always assume the  $S$ -free vertices to be leaf-nodes with all of them except maybe one adjacent to the same central node of  $T\langle S \rangle$ . In particular, given a fixed placement of the  $S$ -constrained vertices, there are  $\mathcal{O}(|S|)$  possibilities in order to place the  $S$ -free vertices (specifically, we choose among  $\mathcal{O}(|S|)$  possibilities the unique  $S$ -free vertex that is not adjacent to the same central node as the others, if any, as well as the central node to which all other  $S$ -free vertices must be adjacent).

Summarizing, we only need to add  $\mathcal{O}(|S|^5)$  different trees in  $\mathcal{T}_S$ , that takes  $\mathcal{O}(|S|^6)$ -time.  $\square$

**Remark 6.1.** We only use the fact that  $S$  is a minimal separator when we process the so-called “sandwiched vertices”. All other arguments in our proof stay valid for clique-intersections. Furthermore, as already observed for Lemma 4.2, such “sandwiched vertices” do not exist in maximal cliques. Therefore, we can also use the algorithm of Theorem 6.1 in order to generate, for any maximal clique  $K_i$  without a  $K_i$ -free vertex, all possible subtrees  $T\langle K_i \rangle$  of diameter at most 3 in any well-structured 4-Steiner root of  $G$ .

## 6.2 Case of a Leaf Node

In this section, we generalize the construction of Theorem 6.2 to the maximal cliques that can be leaves in a rooted clique-tree. For that, we use a well-known decomposition of these maximal cliques into a unique minimal separator and a set of simplicial vertices.

**Theorem 6.2.** *Given  $G = (V, E)$  and a rooted clique-tree  $T_G$  of  $G$ , let  $K_i \in \mathcal{K}(G)$  be a leaf. We can construct, in time polynomial in  $|K_i|$ , a set  $\mathcal{T}_i$  of 4-Steiner roots for  $G_i := G[K_i]$  with the following additional property: In any well-structured 4-Steiner root  $T$  of  $G$ , there exists  $T'_i \in \mathcal{T}_i$  Steiner-equivalent to  $T\langle K_i \rangle$ .*

*Proof.* Let  $K_{p(i)}$  be the father node of  $K_i$ . By Theorem 2.4,  $S_i := K_i \cap K_{p(i)}$  is a minimal separator. We compute the family  $\mathcal{T}_{S_i}$  given by Theorem 6.1. Then, in order to compute a candidate subtree  $T_i$ , to be added into  $\mathcal{T}_i$ , we consider all the subtrees  $T_{S_i} \in \mathcal{T}_{S_i}$  and we proceed as follows. We select a node in  $T_{S_i}$  that we assume to be closest to  $\mathcal{C}(T_i)$  (hence,  $\mathcal{O}(|V(T_{S_i})|) = \mathcal{O}(|S_i|)$  possibilities), and we set its distance to the center (this can only be 0, 1 or 2). In doing so, we can assume  $\mathcal{C}(T_i)$  to be added into  $T_{S_i}$ . Finally, the vertices in  $K_i \setminus S_i$  are all simplicial, and so, we can connect them to  $\mathcal{C}(T_i)$  as explained in Lemma 4.3 (one possibility up to Steiner equivalence). Note that in doing so, we may also obtain solutions  $T_i$  such that  $\text{diam}(T_i) \neq 4$ , that we will need to discard. See Fig. 11 for an illustration. Overall,  $|\mathcal{T}_i| = \mathcal{O}(|S_i| |\mathcal{T}_{S_i}|) = \mathcal{O}(|S_i|^6)$ .  $\square$



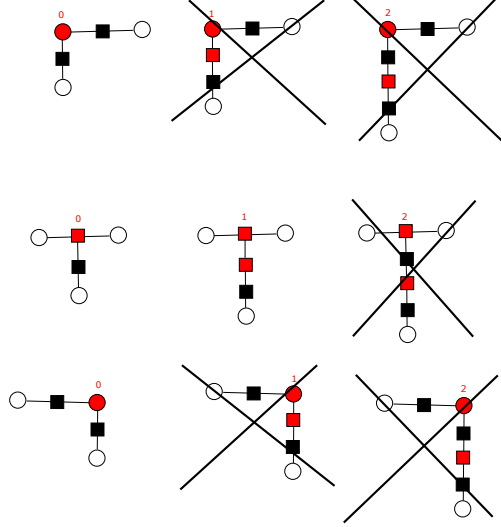


Figure 11: An execution of the algorithm of Theorem 6.2. The minimal separator has size two and induces a star. There is one simplicial vertex to add.

### 6.3 Case of an internal node

Given  $G = (V, E)$  chordal, let  $T_G$  be an arbitrary rooted clique-tree of  $G$  and let  $K_i \in \mathcal{K}(G)$  be an internal node of  $T_G$ . We want to compute a polynomial-size representation for the family  $\mathcal{F}_i$  of all possible subtrees  $T\langle K_i \rangle$  in a well-structured 4-Steiner root of  $G$ .

**Some preliminary observations.** By Lemma 4.3 there always exists a root  $T$  where all  $K_i$ -free vertices are leaves of  $T\langle K_i \rangle$  and connected to  $\mathcal{C}(T\langle K_i \rangle)$  by a path of length two whose internal node is Steiner. Thus, we can first assume for simplicity  $K_i$  does not contain any  $K_i$ -free vertex — such vertices, if they exist, will be added at the end of the construction. Furthermore, as noticed earlier (Remark 6.1) we can use the algorithmic proof of Theorem 6.1 in order to generate all the subtrees  $T\langle K_i \rangle$  of diameter at most three to be added in  $\mathcal{F}_i$ . Summarizing, as a consequence of other results in this paper we are only interested in maximal cliques  $K_i$  with no  $K_i$ -free vertex and in generating subtrees of diameter *exactly* four.

**Organization of this part.** Our main tool for this task is a careful analysis of the intersections between the minimal separators in  $K_i$  (Section 6.3.1). Unfortunately, sometimes we cannot derive from this information a polynomial bound on the number of possible subtrees. We identify the only degenerate case when this cannot be done, and show how to handle with it, in Section 6.3.2. Proposition 6.7 in Section 6.3.3 will summarize our results for this part.

#### 6.3.1 Getting more from clique-intersections

We will use the following lemma in order to prove our first result in this section:

**Lemma 6.3.** *Let  $G = (V, E)$  be strongly chordal, let  $S \in \mathcal{S}(G)$ , let  $K$  be a maximal clique containing  $S$  and let  $R, c$  be such that  $R \subset S$  and either  $c \in R$  or  $c$  is Steiner. We can compute in polynomial time a node  $c'$  with the following properties:*

- *For any well-structured 4-Steiner root  $T$  of  $G$  such that:  $T\langle S \rangle$  is a bistar,  $c \in \mathcal{C}(T\langle S \rangle) \setminus \mathcal{C}(T\langle K \rangle)$ , and  $\text{Real}(N_T[c]) = R$ , there exists a well-structured root  $T'$  with the same properties such that:*

$\mathcal{C}(T'\langle K \rangle) = \{c'\}$ , and  $dist_{T'}(u, v) \geq dist_T(u, v)$  for every  $u, v \in V$ . Moreover, either  $T \equiv_G T'$ , or  $\sum_{u, v \in V} dist_{T'}(u, v) > \sum_{u, v \in V} dist_T(u, v)$ .

In order to better understand the significance of Lemma 6.3, assume that  $T\langle S \rangle$  should be a bistar in the final solution we want to compute, and that we already identified one of its center node  $c$  and the set of real nodes  $R$  to which this node must be adjacent. What this above property says is that there is essentially one canonical way to compute the bistar given  $R$  and  $c$ . The more technical condition  $dist_{T'}(u, v) \geq dist_T(u, v)$  is simply there in order to ensure that by doing so, we cannot miss a solution of an intermediate problem we call DISTANCE-CONSTRAINED ROOT (*i.e.*, see Section 7). Finally, our condition on the potential function  $\sum_{u, v \in V} dist_{T'}(u, v)$  increasing ensures that we can repeatedly apply our “canonical completion” method for arbitrarily many minimal separators  $S$ .

*Proof.* We define a node  $c' \notin V \setminus R$  (either in  $R$  or Steiner), as follows:

- If there exists a clique-intersection  $X \subset S$  such that  $X \not\subseteq R$  and  $X \cap (R \setminus \{c\}) \neq \emptyset$  then, we pick  $c' \in (X \cap R) \setminus \{c\}$ ;
- Else, if there exists  $c' \in R \setminus c$  that is  $K$ -dependent (*i.e.*,  $c'$  is  $(S, K, X_2)$ -sandwiched for some  $X_2 \in \mathcal{X}(G)$ ) then, we output  $c'$ .
- Otherwise,  $c'$  is Steiner.

Note that as  $G$  is strongly chordal, we can compute the above  $c'$  in polynomial time by using the clique-arrangement of  $G$  (Theorem 2.5).

**Correctness.** In what follows, let  $T$  be any well-structured 4-Steiner root of  $G$  such that  $T\langle S \rangle$  is a bistar,  $c \in \mathcal{C}(T\langle S \rangle) \setminus \mathcal{C}(T\langle K \rangle)$ , and  $Real(N_T[c]) = R$ . Furthermore, let  $\mathcal{C}(T\langle S \rangle) = \{c, c_2\}$ . We divide the proof into the following claims:

**Claim 6.3.1.** *If there exists a clique-intersection  $X \subset S$  such that  $X \not\subseteq R$  and  $X \cap (R \setminus \{c\}) \neq \emptyset$  then,  $c' = c_2$ .*

*Proof.* As we have  $X \not\subseteq R$  and  $Real(N_T[c]) = R$ ,  $X$  cannot contain any leaf node of  $T\langle S \rangle$  adjacent to  $c$ . Thus, the only possible node in  $(X \cap R) \setminus \{c\}$  must be  $c_2$ .  $\diamond$

**Claim 6.3.2.** *If there exists  $c' \in R \setminus c$  that is  $K$ -dependent then,  $c' = c_2$ .*

*Proof.* Suppose for the sake of contradiction  $c' \neq c_2$ . In particular, since  $Real(N_T[c]) = R$ ,  $c'$  is a leaf of  $T\langle S \rangle$  that is adjacent to  $c$ . Let  $X_2 \in \mathcal{X}(G)$  be such that  $c'$  is  $(S, K, X_2)$ -sandwiched and let  $x \in (K \cap X_2) \setminus \{c'\}$ . As already proved for Claim 6.1.3,  $dist_T(x, c_2) = 2$  and  $dist_T(x, c) = 3$ . In particular we have  $dist_T(c', x) = 4$ . Since we have  $c', x \in X_2$ , we must have  $diam(T\langle X_2 \rangle) = 4$  and the central node of  $T\langle X_2 \rangle$  must be the central node onto the  $c'x$ -path in  $T$ , that is  $c_2$ . However, since  $K \neq X_2$  (because  $S \not\subseteq X_2$ ) the latter contradicts Lemma 3.5.  $\diamond$

By Claims 6.3.1 and 6.3.2 we may assume from now on  $c'$  to be Steiner. Assume  $c_2 \in R$ , *i.e.*,  $c_2$  is not Steiner (otherwise, we are done). We now reuse a transformation that we introduced in the proof of Theorem 4.1 (*i.e.*, Operation 1 for  $v = c_2$  and  $c$ ). Specifically, we define  $Q_{c_2}$  as the subtree of  $T$  induced by the union of  $c_2$  with all the components of  $T \setminus V(T\langle S \rangle)$  that are adjacent to  $c_2$  and do not contain any real node at a distance  $\leq 4$  from  $S \setminus \{c_2\}$  in  $T$ . We create a new tree  $T'$  by first

removing  $V(Q_{c_2}) \setminus \{c_2\}$ , then replacing  $c_2$  by the Steiner node  $c'$ , and finally adding a copy of  $Q_{c_2}$  and the edge  $cc_2$ . In doing so, we can only increase the distances in  $T'$  compared to  $T$ , and these distances strictly increase at least between  $c_2$  and all the leaves of  $T\langle S \rangle$  to which it was previously adjacent.

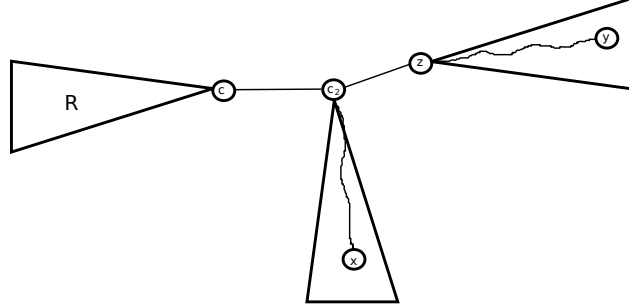


Figure 12: To the proof of Lemma 6.3.

**Claim 6.3.3.**  $T'$  keeps the property of being a well-structured 4-Steiner root of  $G$

Note that proving Claim 6.3.3 will prove the lemma.

*Proof.* This part of the proof closely follows Claim 4.3.1. First as our construction can only increase distances in the Steiner root, we can easily deduce that:

$$\forall uv \notin E, \text{dist}_{T'}(u, v) \geq \text{dist}_T(u, v) > 4.$$

However, we also need to check that conversely,  $\text{dist}_{T'}(u, v) \leq 4$  for all  $uv \in E$ . For that, we prove as an intermediate subclaim that:

$$\text{dist}_T(\text{Real}(Q_{c_2} \setminus \{c_2\}), V \setminus Q_{c_2}) > 4.$$

Indeed, suppose by contradiction there exist  $x \in \text{Real}(Q_{c_2}) \setminus \{c_2\}, y \in V \setminus V(Q_{c_2})$  such that  $\text{dist}_T(x, y) \leq 4$ . The unique  $xy$ -path in  $T$  goes by  $c_2$  (see Fig. 12 for an illustration). Furthermore by the definition of  $Q_{c_2}$  we have  $\text{dist}_T(x, S \setminus \{c_2\}) > 4$ . But then since  $c_2 \in \mathcal{C}(T\langle S \rangle)$  we have  $\text{dist}_T(c_2, S \setminus \{c_2\}) = 1$ , and so  $\text{dist}_T(x, c_2) \geq 4$ . It implies  $\text{dist}_T(x, y) > 4$ , a contradiction.

This above subclaim implies that if  $\text{dist}_{T'}(u, v) > 4$  for some  $uv \in E$  then,  $c_2 \in \{u, v\}$ . In order to prove that no such a pair  $uv$  can exist, and that  $T'$  is well-structured, we consider all the clique-intersections  $X'$  that contain  $c_2$ . There are three cases:

- Case  $X' \subset S$ . Then, either  $X' = \{c_2\}$  and we are done, or  $|X'| \geq 2$ . Furthermore in the latter subcase we have  $X' \subseteq R$  (otherwise, this would imply the existence of a  $X$  as earlier defined, that would contradict Claim 6.3.1). As a result we have  $T\langle X' \rangle = T'\langle X' \rangle$ .
- Case  $X' \supseteq S$ . We can observe that  $T\langle X' \rangle \cap Q_{c_2} = \{c_2\}$  since we proved above that we have  $\text{dist}_T(\text{Real}(Q_{c_2} \setminus \{c_2\}), V \setminus Q_{c_2}) > 4$ . In particular,  $T'\langle X' \rangle$  is obtained from  $T\langle X' \rangle$  by replacing  $c_2$  by a Steiner node (only if it were an internal node of  $T\langle X' \rangle$ ) then, making of  $c_2$  a leaf. Furthermore, since there is at least one leaf adjacent to each center node in  $T\langle S \rangle$ , we can prove as for Subclaim 4.3.1.2 that  $\text{diam}(T'\langle X' \rangle) = \text{diam}(T\langle X' \rangle)$ . As already observed in Claim 4.3.1, it implies that all properties of Theorem 4.1 are preserved provided that  $c_2$  is

not  $X'$ -free. So we only need to prove this is always the case. On one hand,  $c_2$  is not  $S$ -free because  $c_2$  is not a leaf of  $T\langle S \rangle$  and we assume  $T$  is well-structured. On the other hand,  $c_2$  is not  $X'$ -free for any  $X' \supset S$ , because  $c_2 \in S$  and  $|S| \geq 3$ .

- Otherwise, in all other cases we prove  $T'\langle X' \rangle = T\langle X' \rangle$ . To see that, first note that it may not be the case only if  $T\langle X' \rangle$  is not fully contained into  $Q_{c_2}$  (and so,  $T\langle X' \rangle \cap Q_{c_2} = \{c_2\}$ ).

We prove as a subclaim that if  $T'\langle X' \rangle \neq T\langle X' \rangle$  then,  $T\langle X' \rangle$  must intersect  $T\langle S \rangle \setminus \{c_2\}$ . Indeed, otherwise  $T\langle X' \rangle$  must intersect some subtree  $T^{(sub)}$  of  $T \setminus T\langle S \rangle$  that is not in  $Q_{c_2}$ . Let  $x \in X' \cap T^{(sub)}$ . By the definition of  $Q_{c_2}$ , there exists  $x' \in \text{Real}(T^{(sub)})$  such that  $\text{dist}_T(x', S \setminus \{c_2\}) \leq 4$  and there exists a  $xx'$ -path in  $T$  that does not go by  $c_2$ . In particular,  $x'$  must be adjacent in  $G$  to all the vertices in  $S$  (otherwise, we could derive the existence of a clique-intersection  $X$  as defined above, thereby contradicting Claim 6.3.1). Note that it implies  $\text{dist}_T(x', c_2) \leq 2$  because the unique path in  $T$  between  $x'$  and any leaf adjacent to  $c$  must go by  $c_2$ . Then,  $x' \in K$ . Furthermore since  $\text{dist}_T(x, c_2) \leq 4$  we obtain (by considering the median node of the triple  $x, x', c_2$ )  $\text{dist}_T(x, x') \leq 4$ .

Let  $K''$  be a maximal clique containing all of  $x, x', c_2$ . There are two cases.

- Assume first  $\text{dist}_T(x, S \setminus \{c_2\}) > 4$ . Then,  $S \not\subseteq K''$ , and more specifically  $K'' \cap S = \{c_2\}$ . It implies  $K'' \neq K$ , and so  $c_2$  is  $(S, K, K'')$ -sandwiched, that contradicts Claim 6.3.2.
- Otherwise,  $\text{dist}_T(x, S \setminus \{c_2\}) \leq 4$  and we may further assume  $x = x'$ . However, by the hypothesis  $S \not\subseteq X'$ ,  $X' \not\subseteq S$  resp., and so  $S \cap X' = \{c_2\}$  (otherwise,  $X = S \cap X'$  would falsify Claim 6.3.1). It implies  $v$  is  $(S, K, X')$ -sandwiched, that contradicts Claim 6.3.2.

Therefore, we proved our subclaim that  $T\langle X' \rangle$  must intersect  $T\langle S \rangle \setminus \{c_2\}$ . In fact, our proof shows more generally that  $T\langle X' \rangle$  cannot intersect any subtree of  $T \setminus T\langle S \rangle$  with a node adjacent to  $c_2$ . Then,  $T\langle X' \rangle$  must intersect a real node in  $N_T(c_2) \setminus \{c\} = S \setminus R$  since otherwise,  $T\langle R \rangle = T'\langle R \rangle$  would imply  $T\langle X' \rangle = T'\langle X' \rangle$ , a contradiction. However in this situation,  $X := S \cap X'$  satisfies  $|X| \geq 2, c_2 \in X$  and  $X \not\subseteq R$ , that contradicts Claim 6.3.1.

◇

The above case analysis ends up proving the claim, and so, the lemma. □

We will also need the following useful result which we keep using throughout most of the remaining proofs in this paper:

**Lemma 6.4.** *Given  $G = (V, E)$  and  $T$  any 4-Steiner root of  $G$ , let  $X \in \mathcal{X}(G)$  and let  $S \subset X$  be a minimal separator. If  $T\langle S \rangle$  is a non-edge-star then, there exists  $c \in N_T[\mathcal{C}(T\langle X \rangle)]$  such that  $\text{Real}(N_T[c]) = S$ .*

*Proof.* Write  $\mathcal{C}(T\langle S \rangle) = \{c\}$ . By Theorem 3.4,  $\text{Real}(N_T[c]) = S$ . Furthermore since by the hypothesis  $T\langle S \rangle$  has at least two leaves then, the unique path between at least one such a leaf and  $\mathcal{C}(T\langle X \rangle)$  must pass by  $c$ . Since  $\text{rad}(T\langle X \rangle) \leq 2$ , this implies  $\text{dist}_T(c, \mathcal{C}(T\langle X \rangle)) \leq 1$ . □

We now explain how to construct an important subfamily of  $\mathcal{F}_i$ :

**Lemma 6.5.** *Let  $K_i$  be a maximal clique of  $G = (V, E)$  with no  $K_i$ -free vertex. In  $\mathcal{O}(n|K_i|^7)$ -time, we can compute a family  $\mathcal{B}_i$  with the following special property: For any well-structured 4-Steiner root  $T$  of  $G$  where for at least one minimal separator  $S \subset K_i$ ,  $T\langle S \rangle$  is a bistar, there is a  $T'$  such that  $T'\langle S_i \rangle \equiv_G T\langle S_i \rangle$ ,  $T'\langle K_i \rangle \in \mathcal{B}_i$  and  $\text{dist}_{T'}(r, V_i \setminus S_i) \geq \text{dist}_T(r, V_i \setminus S_i)$  for every  $r \in V(T\langle S_i \rangle)$ .*

Note that we do not capture *all* well-structured roots with this above lemma, but only those maximizing certain distances' conditions.

*Proof.* Let  $\Omega(K_i) \subseteq \mathcal{S}(G)$  contain all the minimal separators in  $K_i$ . By Theorem 6.1, for any  $S \in \Omega(K_i)$  we can construct a family  $\mathcal{T}_S$  such that, in any  $T_{K_i} \in \mathcal{F}_i$ , we should have  $T_{K_i}\langle S \rangle$  is Steiner-equivalent to some tree in  $\mathcal{T}_S$ . This takes total time  $\mathcal{O}(|K_i|^6 |\Omega(K_i)|) = \mathcal{O}(n|K_i|^6)$ . Fix  $S \in \Omega(K_i)$  (there are  $\mathcal{O}(n)$  possibilities) and a bistar  $T\langle S \rangle \in \mathcal{T}_S$  (by Theorem 6.1, there are  $\mathcal{O}(|S|^5)$  possibilities, that is in  $\mathcal{O}(|K_i|^5)$ ). Note that in particular for  $S_i \subseteq S$ , this will also generate all possibilities for  $T\langle S_i \rangle$ .

Roughly we show that except in a few particular cases easy to solve, for every  $S' \in \Omega(K_i)$  there is only *one* canonical solution in  $\mathcal{T}_{S'}$  that is compatible with  $T\langle S \rangle$ . For that, assume the existence of a 4-Steiner root  $T$  of  $G$  that contains  $T\langle S \rangle$  as a subtree. We may only consider those  $S' \in \Omega(K_i)$  that are *not* contained into any other  $S'' \in \Omega(K_i)$ . Indeed if  $S' \subseteq S''$  then, trivially  $T\langle S' \rangle$  is forced by  $T\langle S'' \rangle$ . Thus from now on, we assume  $S'$  is inclusion wise maximal in  $\Omega(K_i)$ .

In what follows is a simple observation for the case  $S \cap S' = \emptyset$  (see also Fig. 13 for an illustration).

**Claim 6.5.1.** *If  $S \cap S' = \emptyset$  then,  $T\langle S' \rangle$  is a star with a Steiner central node. Moreover, the central node of  $T\langle K_i \rangle$  must be Steiner.*

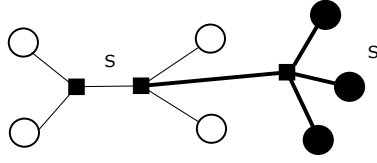


Figure 13: An example where  $S \cap S' = \emptyset$ .

*Proof.* By Lemma 5.2,  $\mathcal{C}(T\langle K_i \rangle) \subset \mathcal{C}(T\langle S \rangle)$ , and so by Theorem 3.4,  $Real(N_T[\mathcal{C}(T\langle K_i \rangle)]) \subset S$ . But then,  $T\langle K_i \rangle \setminus N_T[\mathcal{C}(T\langle K_i \rangle)]$  is a collection of isolated leaves. The latter proves either  $S' = \{v\}$  is a cut-vertex or  $T\langle S' \rangle$  is a star with a Steiner central node in  $N_T[\mathcal{C}(T\langle K_i \rangle)]$ . In the former case we so conclude that  $v$  is  $K_i$ -free by inclusion wise maximality of  $S'$  and by Lemma 4.2. Since we assume there is no  $K_i$ -free vertex, this case cannot happen. Therefore,  $T\langle S' \rangle$  is a star with a Steiner central node in  $N_T[\mathcal{C}(T\langle K_i \rangle)]$ . Finally, the center of  $T\langle K_i \rangle$  must be also Steiner (otherwise by Lemma 6.4, this vertex should be in  $S'$ ).  $\diamond$

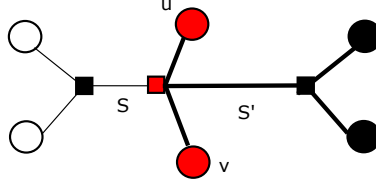
If  $S \cap S' = \emptyset$  then, by combining Claim 6.5.1 and Lemma 6.4 there is essentially one way to insert  $S'$  in  $T\langle K_i \rangle$  (*i.e.*, we construct a star  $T\langle S' \rangle$  with one Steiner central node, then we make this central node adjacent to the Steiner central node of  $T\langle K_i \rangle$ ).

For the remaining cases, we assume  $S' \cap S \neq \emptyset$  for any inclusion wise maximal  $S' \in \Omega(K_i)$ . Several cases may arise:

- Case there exist  $u, v \in S \cap S'$  nonadjacent in  $T\langle S \rangle$  (see Fig. 14 for an illustration). We prove  $T\langle S' \rangle$  is a bistar. Indeed, suppose by contradiction  $T\langle S' \rangle$  is a star. Then, since  $u, v \in S \cap S'$  are non adjacent, the center of  $T\langle S' \rangle$  must be in  $\mathcal{C}(T\langle S \rangle)$ . This implies  $S' \subset S$ , a contradiction. Therefore, we proved as claimed  $T\langle S' \rangle$  must be a bistar.

In this situation we must have  $S \cap S' = Real(N_T[\mathcal{C}(T\langle K_i \rangle)])$  (*e.g.*, see the proof of Lemma 5.2). Since there exist  $u, v \in S \cap S'$  nonadjacent in  $T\langle S \rangle$ , the central node of  $T\langle K_i \rangle$  can be uniquely

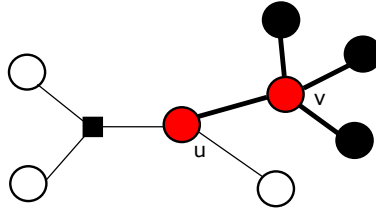
defined as the central node  $c_i \in \mathcal{C}(T\langle S \rangle)$  such that  $u, v \in N_T[c_i]$ . Finally, since  $Real(N_T[c_i])$  is fixed by  $T\langle S \rangle$ , by Lemma 6.3 this leaves at most *one* canonical possibility for the second central node in  $\mathcal{C}(T\langle S' \rangle)$ , and so, at most one possibility for  $T\langle S' \rangle$ .



Note that we always fall in this case provided  $|S \cap S'| \geq 3$ . So, we are left to study when  $|S \cap S'| \in \{1, 2\}$ .

- Case  $S \cap S' = \{u, v\}$ . We further assume  $uv \in E(T\langle S \rangle)$  (otherwise, we fall in the previous subcase). Recall that we assume  $S' \not\subseteq S$ . In particular, we must have  $S \cap S' \subseteq N_T[\mathcal{C}(T\langle K_i \rangle)]$ . W.l.o.g.,  $u \in \mathcal{C}(T\langle S \rangle)$  (or equivalently,  $u$  must be the central node of  $T\langle K_i \rangle$ ). Since  $S' \not\subseteq S$  and there is at least one leaf-node of  $T\langle S \rangle$  that is a real node adjacent to  $u$ , we so deduce that  $v$  is a leaf of  $T\langle S \rangle$  (otherwise, we would have  $S \cap S' \neq \{u, v\}$ ). Several situations force  $T\langle S' \rangle$  to be a star, for instance if:
  - $S'$  is strictly contained into another minimal separator of  $G$ ;
  - $G \setminus S'$  has at least three full components<sup>2</sup>;
  - or  $Real(N_T[u]) \neq \{u, v\}$ .

If such a situation occurs then, by Lemma 6.4,  $v$  must be the center of the star  $T\langle S' \rangle$ , thereby leaving only one possibility for  $T\langle S' \rangle$  (*i.e.*, see Fig. 15).



From now on assume that no minimal separator strictly contains  $S'$ ,  $G \setminus S'$  has exactly two full components and  $Real(N_T[u]) = \{u, v\}$ . The subtree  $T\langle S' \rangle$  is forced to be a bistar if there exists at least one  $S'' \in \Omega(K_i)$  inclusion wise maximal such that:  $S \cap S'' = \{v\}$  (otherwise,  $T\langle S'' \rangle$  should be an edge and, since we assume  $|S'| \geq 3$  this would imply  $S'' \subseteq S'$  by Lemma 6.4, a contradiction). Furthermore as explained in the previous case there is at most one canonical possibility for the bistar  $T\langle S' \rangle$ .

If no  $S''$  as above exists then,  $T\langle S' \rangle$  may be either a star or a bistar. We can bipartition all the remaining minimal separators  $S'' \in \Omega(K_i)$  that are inclusion wise maximal (including  $S'$ )

<sup>2</sup>Equivalently, given any clique-tree  $T_G$  of  $G$  we have  $|E_{S'}(T_G)| \geq 2$ .

as follows: those containing  $v$ , and those that do not. Note that in the former subcase (which includes  $S'$ ) we have  $S \cap S'' = \{u, v\}$ , whereas in the latter subcase  $S \cap S'' = \{u\}$ . Furthermore if  $S \cap S'' = \{u\}$  then,  $T\langle S'' \rangle$  must always be a star with a Steiner central node that is adjacent to  $u$  (to see that, recall that  $Real(N_T[u]) = \{u, v\}$ , and so,  $T\langle S'' \rangle$  cannot be a bistar). In particular, there is only *one* possibility for such a  $S''$ . However, the same as  $S'$ , for all other  $S''$  such that  $S \cap S'' = \{u, v\}$ ,  $T\langle S'' \rangle$  may be either a star or a bistar. The key observation here is that  $T\langle S'' \rangle$  can be a star for at most *one* such a  $S''$  (otherwise, by Lemma 6.4 there would be two non-edge stars with the same center node  $v$ , that contradicts Property P-3.4.2 of Theorem 3.4). Summarizing, since all these sets  $S'' \setminus S$  are pairwise disjoint, we are left with  $\mathcal{O}(|K_i|)$  possibilities for the unique such  $S''$  for which  $T\langle S'' \rangle$  is a star (if any); this choice fixes the corresponding subtree for all the remaining minimal separators. See Fig. 16.

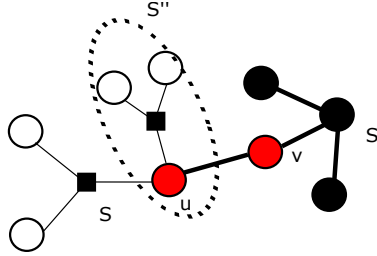


Figure 16: A case where there are two possibilities for  $T\langle S' \rangle$ .

- Case  $S \cap S' = \{v\}$ . If  $v \in \mathcal{C}(T\langle S \rangle)$  then,  $v \in \mathcal{C}(T\langle K_i \rangle)$  because we assume  $S' \not\subseteq S$ . In particular, the only possibility for  $T\langle S' \rangle$  is a star with a Steiner central node that is adjacent to  $v$  (recall that  $v$  is adjacent to at least one leaf in  $T\langle S \rangle$ , and so,  $T\langle S' \rangle$  cannot be a bistar). Assume for the remaining of the case  $v$  is a leaf of  $T\langle S \rangle$ . As in the previous case, several situations force  $T\langle S' \rangle$  to be a star, like if:
  - $S'$  is strictly contained into another minimal separator of  $G$ ;
  - $G \setminus S'$  has at least three full components;
  - or  $Real(N_T[\mathcal{C}(T\langle K_i \rangle)]) \neq \{v\}$ .

Furthermore if such a situation occurs then,  $v$  must be a center node of the star  $T\langle S' \rangle$  (possibly,  $T\langle S' \rangle$  is an edge), and so, there is only one possibility for  $T\langle S' \rangle$ .

From now on assume no minimal separator strictly contains  $S'$ ,  $G \setminus S'$  has exactly two full components and  $Real(N_T[\mathcal{C}(T\langle K_i \rangle)]) = \{v\}$ . In particular, the unique central node of  $T\langle K_i \rangle$  is some Steiner node  $\alpha_i$ . We also consider all the other minimal separators  $S'' \in \Omega(K_i)$  in the same situation as  $S'$ . As we only consider inclusion wise maximal elements  $S'' \in \Omega(K_i)$  intersecting  $S$ , we must have  $S \cap S'' = \{v\}$ . However, unlike the previous subcase, in an *arbitrary* well-structured  $T$  there may be several such  $S''$  for which  $T\langle S'' \rangle$  is a star. We now prove that up to local modifications of  $T$ , we can always assume there is at most *one* such  $S''$  for which  $T\langle S'' \rangle$  is a star. Note that by doing so, we can conclude as for the previous subcase about the number of possibilities for  $T\langle K_i \rangle$ .

Assume there exist  $S_{i_j}, S_{i_k}$  such that  $T\langle S_{i_j} \rangle, T\langle S_{i_k} \rangle$  are stars. Then,  $T\langle S_{i_j} \rangle, T\langle S_{i_k} \rangle$  must be edges with a common end  $v$  (otherwise, one should be a non-edge star and so by Lemma 6.4,

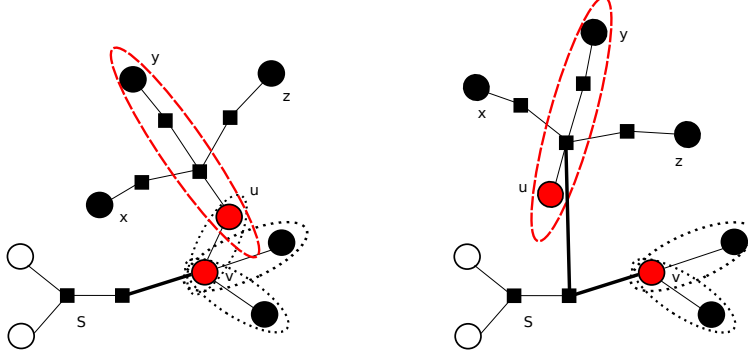


Figure 17: The transformation of an edge into a bistar.

either  $S_{i_j} \subset S_{i_k}$  or  $S_{i_k} \subset S_{i_j}$ , a contradiction). By inclusion wise maximality of  $S_{i_j}$  and  $S_{i_k}$ , there is at least one of these two separators whose intersection with  $S_i$  is either empty or reduced to  $\{v\}$ . Assume w.l.o.g. this is the case for  $S_{i_j}$  and write  $S_{i_j} = \{u, v\}$ . We first gain more insights on the structure of  $T\langle K_{i_j} \rangle$ . For that, let  $W_{i_j} := V_{i_j} \setminus S_{i_j}$ . Since  $v$  has a neighbour in  $V_i \setminus V_{i_j}$  we must have  $\text{dist}_T(v, W_{i_j}) = 4$ . This implies  $\text{diam}(T\langle K_{i_j} \rangle) = 4$ ,  $\text{dist}_T(u, W_{i_j}) = 3$ , and all other real vertices of  $T\langle K_{i_j} \rangle$  must be leaves at distance two from  $\mathcal{C}(T\langle K_{i_j} \rangle)$ . See Fig. 17 for an illustration.

We connect the unique node  $\alpha_{i_j} \in \mathcal{C}(T\langle K_{i_j} \rangle)$  (which is Steiner) to  $\alpha_i$  and then, we remove the edge  $uv$ . In doing so, we obtain a tree  $T'$  such that  $\text{Real}(T') = V$  and  $T'\langle S_{i_j} \rangle$  is a bistar. Note that in particular,  $T\langle S_i \rangle = T'\langle S_i \rangle$ , that follows from  $S_i \cap S_{i_j} \subseteq \{v\}$ . Furthermore we claim that  $T'$  keeps the property of being a 4-Steiner root of  $G$ . There are two cases:

- Suppose by contradiction  $\text{dist}_{T'}(x, y) \leq 4$  for some  $xy \notin E$ . In particular, the unique  $xy$ -path in  $T'$  must go by the edge  $\alpha_{i_j}\alpha_i$ . However since all neighbours of  $\alpha_i$  except  $v$  and all neighbours of  $\alpha_{i_j}$  except  $u$  are Steiner nodes, we obtain that  $\{u, v\} \cap \{x, y\} \neq \emptyset$ . In particular if  $x \in \{u, v\}$  then,  $y \in K_i \cup K_{i_j}$ , and so  $xy \in E(G)$ . A contradiction.
- Conversely, suppose by contradiction  $\text{dist}_{T'}(x, y) > 4$  for some  $xy \in E$ . In particular, the unique  $xy$ -path in  $T$  must go by the edge  $uv$ . Therefore, there must be a maximal clique  $K$  containing all of  $x, y, u, v$ . We have  $K \notin \{K_i, K_{i_j}\}$  since  $\text{dist}_{T'}(x, y) > 4$ . But then by maximality of  $S_{i_j}$ ,  $K \cap K_i = K \cap K_{i_j} = K_i \cap K_{i_j} = S_{i_j}$  and there are at least three full components in  $G \setminus S_{i_j}$ . A contradiction.

So, we proved as claimed that  $T'$  keeps the property of being a 4-Steiner root of  $G$ . We can prove in the same way that we have  $T\langle X \rangle = T'\langle X \rangle$  for every clique-intersection  $X \notin \{S_{i_j}, K_i, K_{i_j}\}$ , that implies  $T'$  is well-structured. We end up observing  $\text{dist}_{T'}(r, W_i) \geq \text{dist}_T(r, W_i)$  for every  $r \in T\langle S_i \rangle$  by construction, where  $W_i := V_i \setminus S_i$ .

Then, we obtain the desired property by repeating this above transformation until there is at most one  $S''$  such that  $T\langle S'' \rangle$  is a star.

Overall, given a fixed  $T\langle S \rangle$  we have at most  $\mathcal{O}(|K_i|)$  possibilities for  $T\langle K_i \rangle$ . □



### 6.3.2 A degenerate case

If there is no minimal separator  $S \subset K_i$  such that  $T\langle S \rangle$  is a bistar then, we get much less information on the structure of  $T\langle K_i \rangle$ . We identify the following as our main obstruction for bounding the number of possible subtrees:

**Definition 6.1.** Given  $G = (V, E)$  and  $T$  a 4-Steiner root of  $G$ , let  $K_i \in \mathcal{K}(G)$  and let  $S \subset K_i$  be a minimal separator of size  $|S| \geq 2$ . We call  $T\langle S \rangle$  a *thin branch* of  $T\langle K_i \rangle$  if we have:

- $T\langle S \rangle \setminus \mathcal{C}(T\langle K_i \rangle)$  is a connected component of  $T\langle K_i \rangle \setminus \mathcal{C}(T\langle K_i \rangle)$ ;
- and there is no other  $T\langle S' \rangle, S' \subset K_i$  which both intersects  $\mathcal{C}(T\langle K_i \rangle)$  and  $T\langle S \rangle \setminus \mathcal{C}(T\langle K_i \rangle)$ .

The *head* of a thin branch is the node of  $T\langle S \rangle$  that is the closest to  $\mathcal{C}(T\langle K_i \rangle)$ .

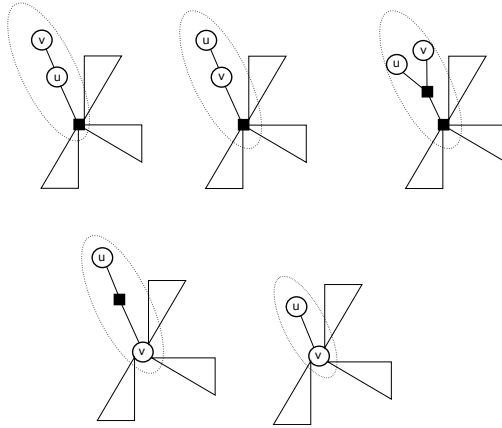


Figure 18: Examples of thin branches (represented by a dashed ellipse).

In order to understand the difficulties we met, assume on the way to construct  $T\langle K_i \rangle$  we correctly identified  $\mathcal{C}(T\langle K_i \rangle)$  and a minimal separator  $S$  for which  $T\langle S \rangle$  must be a thin branch. We can prove that  $T\langle S \rangle$  must be a star (possibly, an edge). However, without any additional information, there would be at least  $|S| \geq 2$  possibilities for  $\mathcal{C}(T\langle S \rangle)$  (e.g., see Figure 18). If there are  $p$  such minimal separators  $S^1, S^2, \dots, S^p$  for which  $T\langle S^j \rangle$  must be a thin branch then, the number of possibilities for  $T\langle K_i \rangle$  goes up to  $2^p$  at least.

Intuitively, our choice for  $T\langle S^j \rangle$  does not really matter as long as this does not violate any distance's constraints in the final solution we get. Guided by this intuition, we will sketch in Section 8 a way to process all these  $S^j$ 's – except maybe one – independently from each other. In particular, for now we do not really need to “guess” what will be exactly  $T\langle S^j \rangle$  in our final solution but just to correctly certify it has to be a thin branch. Specifically, we prove the following result:

**Lemma 6.6.** *Let  $K_i$  be a maximal clique of  $G = (V, E)$  with no  $K_i$ -free vertex. There exists a family  $\mathcal{D}_i$  that can be computed in  $\mathcal{O}(n|K_i|^4)$ -time, and such that the following hold for any well-structured 4-Steiner root  $T$  of  $G$ :*

1. *If  $\text{diam}(T\langle K_i \rangle) = 4$ , and there is no minimal separator  $S \subset K_i$  such that  $T\langle S \rangle$  is a bistar then, we have:  $(T'\langle Y_i \cup \mathcal{C}(T\langle K_i \rangle) \rangle, \mathcal{C}(T\langle K_i \rangle)) \in \mathcal{D}_i$  for some  $T' \equiv_G T$  and  $Y_i \subseteq X_i$ ;*

2. Moreover,  $S_i \subseteq Y_i$ , and for any  $v \in K_i \setminus Y_i$  there is a minimal separator  $S \subseteq (K_i \setminus Y_i) \cup \mathcal{C}(T\langle K_i \rangle)$  such that:  $v \in S$ , and  $T\langle S \rangle$  is a thin branch.

*Proof.* By the hypothesis we are left to compute the diameter-four subtrees where, for every minimal separator  $S \subset K_i$ ,  $T\langle S \rangle$  has diameter at most two. For that, we only need to consider the subset  $\mathcal{S}_i$  of all minimal separators  $S \subset K_i$  that are not strictly contained into any other minimal separator in  $K_i$ . Furthermore, we recall that by the hypothesis there is no  $K_i$ -free vertex. In particular, every  $S \in \mathcal{S}_i$  has size at least two. This implies  $T\langle S \rangle$  must be either an edge or a star. We now divide the proof into several cases:

- Case there is a  $S \in \mathcal{S}_i$  such that  $\mathcal{C}(T\langle K_i \rangle) = \mathcal{C}(T\langle S \rangle)$ . Note that in this case,  $T\langle S \rangle$  must be a non-edge star. Fix  $S \in \mathcal{S}_i$  (there are  $\mathcal{O}(n)$  possibilities) and one non-edge star  $T\langle S \rangle$  such that  $Real(T\langle S \rangle) = S$  (there are  $\mathcal{O}(|S|) = \mathcal{O}(|K_i|)$  possibilities). By Lemma 6.4,  $Real(N_T[\mathcal{C}(T\langle S \rangle)]) = S$ . Therefore, there is at most one compatible solution for any other  $S' \in \mathcal{S}_i$ , namely: if  $v \in S \cap S'$  is a leaf of  $T\langle S \rangle$  then,  $T\langle S' \rangle$  must be a star with  $v$  as a central node (possibly, an edge); otherwise,  $T\langle S' \rangle$  must be a non-edge star with a Steiner central node  $\alpha \in N_T(\mathcal{C}(T\langle S \rangle))$ . See Fig. 19 for an illustration of that case.

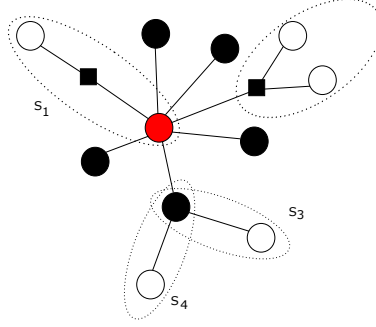


Figure 19: Case 1 of Lemma 6.6.

- Case there is no  $S \in \mathcal{S}_i$  such that  $\mathcal{C}(T\langle K_i \rangle) = \mathcal{C}(T\langle S \rangle)$ . Fix  $c_i \in \mathcal{C}(T\langle K_i \rangle)$  as being any vertex of  $K_i$  or Steiner (this gives  $\mathcal{O}(|K_i|)$  possibilities). There are several subcases:
  - For  $S \in \mathcal{S}_i$  of size  $|S| \geq 3$ , the only possibility for  $T\langle S \rangle$  is to be a non-edge star such that (by Lemma 6.4)  $Real(N_T[\mathcal{C}(T\langle S \rangle)]) = S$ , and the center of  $T\langle S \rangle$  must be adjacent to  $c_i$ . If in addition, there is a clique-intersection  $X \subset S, |X| = 2$  and  $c_i \in X$  then, the center of  $T\langle S \rangle$  must be the unique vertex in  $X \setminus \{c_i\}$ . Otherwise, we claim that  $T\langle S \rangle$  must be a thin branch. Indeed, suppose by contradiction the existence of a  $S' \subset K_i$  such that  $T\langle S' \rangle$  intersects both  $\mathcal{C}(T\langle K_i \rangle)$  and  $T\langle S \rangle \setminus \mathcal{C}(T\langle K_i \rangle)$ . As we assume  $S \not\subseteq S'$ , we should have  $T\langle S \rangle \cap T\langle S' \rangle$  that is contained into the edge between  $c_i$  and the central node of  $T\langle S \rangle$ . In particular,  $S' \not\subseteq S$  (otherwise, this would contradict the non-existence of a clique-intersection  $X$  as defined above). This implies  $T\langle S' \rangle$  should be either a bistar or a non-edge star such that  $\mathcal{C}(T\langle K_i \rangle) = \mathcal{C}(T\langle S' \rangle)$ . However, in both cases this would contradict our hypothesis that no such  $S'$  exist. Therefore, we proved as claimed that  $T\langle S \rangle$  must be a thin branch.
  - Let  $S^1, S^2, \dots, S^q$  be minimal separators of size exactly two that are pairwise intersecting into some vertex  $u \neq c_i$ . Then, their union must be a star: where the center is the unique

vertex  $u$  in  $\bigcap S^j$ , and  $Real(N_{T\langle K_i \rangle}[u]) = \{c_i\} \cup \left(\bigcup_j S^j\right)$ . In particular,  $T\langle S^j \rangle$  must be an edge for every  $j$ .

- So, the only remaining subcase is a minimal separator  $S \in \mathcal{S}_i$  such that:  $|S| = 2$ , and the intersection of  $S$  with any other minimal separator of  $\mathcal{S}_i$  is either empty or reduced to  $c_i$ . Then,  $T\langle S \rangle$  must be a thin branch.

See Fig. 20 for an illustration of these subcases.

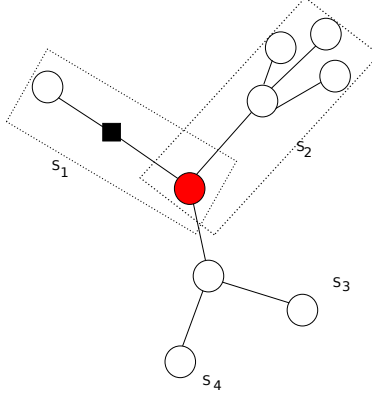


Figure 20: Case 2 of Lemma 6.6. Thin branches are identified by dotted rectangles.

Finally, according to Definition 6.1, there may be at most one  $S$  such that  $S_i \cap (S \setminus \{c_i\}) \neq \emptyset$  and  $T\langle S \rangle$  must be a thin branch. Only for this  $S$  we generate all possibilities for  $T\langle S \rangle$ , thereby generating  $\mathcal{O}(|S|)$  different pairs  $(T_{Y_i}, c_i)$  to add in the family.

□

### 6.3.3 The polynomial-time computation

Summarizing this section we get:

**Proposition 6.7.** *Let  $K_i$  be a maximal clique of  $G = (V, E)$ . In  $\mathcal{O}(n|K_i|^8)$ -time, we can compute a family  $\mathcal{F}_i$  with the following special property. For any well-structured 4-Steiner root  $T$  of  $G$ , there exists a  $T'$  and a (not necessarily maximal) clique  $Y_i \subseteq K_i$  such that  $(T'\langle Y_i \cup \mathcal{C}(T'\langle K_i \rangle) \rangle, \mathcal{C}(T'\langle K_i \rangle)) \in \mathcal{F}_i$ , and we have:*

- $S_i \subseteq Y_i$  and  $T'\langle S_i \rangle \equiv_G T\langle S_i \rangle$ ;
- $dist_{T'}(r, V_i \setminus S_i) \geq dist_T(r, V_i \setminus S_i)$  for any  $r \in V(T\langle S_i \rangle)$ ;
- For any  $v \in K_i \setminus Y_i$  there is a minimal separator  $S \subseteq (K_i \setminus Y_i) \cup \mathcal{C}(T\langle K_i \rangle)$  such that:  $v \in S$  and  $T\langle S \rangle$  is a thin branch.

Moreover,  $Y_i = K_i$  if either  $diam(T'\langle K_i \rangle) < 4$  or there exists a minimal separator  $S \subset K_i$  such that  $T'\langle S \rangle$  is a bistar.

*Proof.* If there are  $K_i$ -free vertices then, by using Lemma 4.3, there is essentially one canonical way to add these vertices at the end of the construction. For that, it suffices to fix the center of  $T\langle K_i \rangle$  which, as explained in the proof of Theorem 6.2, can only increase the total runtime by a multiplicative factor in  $\mathcal{O}(|K_i|)$ . Thus from now on we may assume that  $K_i$  has no  $K_i$ -free vertex. Furthermore, we may also assume we already computed all the diameter-three subtrees to add in  $\mathcal{F}_i$  (*i.e.*, see Remark 6.1). We explain in Lemma 6.5 how to compute the subfamily  $\mathcal{B}_i$  of all diameter-four subtrees where at least one minimal separator  $S \subset K_i$  has  $T\langle S \rangle$  being a bistar. Finally, Lemma 6.6 completes the construction of the family  $\mathcal{F}_i$ .  $\square$

## 7 Step S.3: Deciding the partial solutions to store

The next two Sections are devoted to the main loop of our algorithm proving Theorem B. Specifically, let  $T_G$  be the rooted clique-tree we defined in Section 5. We consider all the maximal cliques  $K_i$  that are *internal nodes* in  $T_G$ , starting from the twigs (*a.k.a.*, internal nodes whose all children nodes are leaves). For every such a  $K_i$  we must execute a same procedure: that corresponds to Step S.3 of the algorithm, and is the main focus of this section.

During this procedure, we activate each child node  $K_{i_j}$  of  $K_i$  sequentially, by sending a message. This activation message contains a series of constraints on the 4-Steiner roots of  $G_{i_j}$  that we want to compute. Overall, after it has been activated, we complete Step S.4 for  $K_{i_j}$  (detailed in the next Section 8) using the series of constraints it has received from its father node  $K_i$ . The output is a set  $\mathcal{T}_{i_j}$  of partial solutions whose size is polynomial in  $|S_{i_j}|$ . Finally, after we received  $\mathcal{T}_{i_j}$ , we can compute and send the activation message for the next child of  $K_i$  to be activated (if any).

As a way to formalize the constraints that we need to include in activation messages, we now introduce the following problem:

**Problem 3** (DISTANCE-CONSTRAINED ROOT).

**Input:** a graph  $G = (V, E)$  and a rooted clique-tree  $T_G$ , a maximal clique  $K_{i_j}$ , a tree  $T_{S_{i_j}}$  s.t.  $\text{Real}(T_{S_{i_j}}) = S_{i_j}$ , and a sequence  $(d_r)_{r \in V(T_{S_{i_j}})}$  of positive integers.

**Output:** Either a 4-Steiner root  $T_{i_j}$  of  $G_{i_j}$  s.t.  $T_{S_{i_j}} \equiv_G T_{i_j}\langle S_{i_j} \rangle$  and,  $\forall r \in V(T_{S_{i_j}})$ :  $\text{dist}_{T_{i_j}}(r, V_{i_j} \setminus S_{i_j}) \geq d_r$ ; Or  $\perp$  if there is no such a partial solution which can be extended to some well-structured 4-Steiner root  $T$  of  $G$ .

**Theorem 7.1.** Given  $G = (V, E)$  chordal and a rooted clique-tree  $T_G$  as in Theorem 5.4, let  $K_i$  be an internal node with children  $K_{i_1}, K_{i_2}, \dots, K_{i_p}$ . If we can solve DISTANCE-CONSTRAINED ROOT in time  $P(n, |S_{i_j}|)$  for some polynomial  $P$  then, we can compute in time  $\mathcal{O}(n|K_i|^5 P(n, |K_i|))$  a family  $\mathcal{T}_{i_1}, \mathcal{T}_{i_2}, \dots, \mathcal{T}_{i_p}$  of 4-Steiner roots for  $G_{i_1}, G_{i_2}, \dots, G_{i_p}$ , respectively, such that:

1. For any  $j \in \{1, 2, \dots, p\}$ ,  $|\mathcal{T}_{i_j}| = \mathcal{O}(|S_{i_j}|^5)$ ;
2. For any well-structured 4-Steiner root  $T$  of  $G$ , there exists a  $T'$  such that:  $T\langle K_i \rangle \equiv_G T'\langle K_i \rangle$ ,  $T'\langle V_{i_j} \rangle \in \mathcal{T}_{i_j}$  for any  $j \in \{1, 2, \dots, p\}$ , and (only if  $K_i \neq K_0$ )  $\text{dist}_{T'}(r, V_i \setminus S_i) \geq \text{dist}_T(r, V_i \setminus S_i)$  for any node  $r \in V(T\langle S_i \rangle)$ .

We postpone the proof that we can solve DISTANCE-CONSTRAINED ROOT in polynomial time to Section 8. This above result can be seen as a pre-processing phase for  $K_i$ , that is crucial in order

to bound the runtime of our algorithm by a polynomial. Note that the technical condition on the nodes in  $T\langle S_i \rangle$  is simply there to ensure that when later in the algorithm, we will need to solve DISTANCE-CONSTRAINED ROOT at  $K_i$ , we cannot miss a solution.

The remaining of this subsection is now devoted to the proof of Theorem 7.1. We will use some additional terminology that we define next:

**Definition 7.1.** Given  $G = (V, E)$ , let  $A, B, S \subset V$  satisfy  $A \cup B = V$  and  $A \cap B = S$ . Two trees  $T_A, T_B$ , where  $\text{Real}(T_A) = A$  and  $\text{Real}(T_B) = B$ , are *compatible* if  $T_A\langle S \rangle \equiv_G T_B\langle S \rangle$ . Then,  $T_A \odot T_B$  is the tree obtained from  $T_A, T_B$  by the identification of  $T_A\langle S \rangle$  with  $T_B\langle S \rangle$ .

In particular, assume  $G$  to be chordal and let  $T_G$  be a rooted clique-tree of  $G$ . For any  $K_i \in \mathcal{K}(G)$ , let  $S_i := K_i \cap K_{p(i)}$ , let  $V_i := V(G_i)$  and let  $W_i := V_i \setminus S_i$ . Given  $T, T'$  4-Steiner roots of  $G$ , we say that  $T'$  is  *$i$ -congruent* to  $T$  if  $T' \equiv_G T\langle V \setminus W_i \rangle \odot T'_i$ , for some 4-Steiner root  $T'_i$  of  $G_i$ .

Note that in particular, any two Steiner roots of  $G$  are trivially 0-congruent (*i.e.*, assuming  $S_0 = \emptyset$  by convention). Finally in what follows we also use  $d_{T_{i_j}}(r)$  as a shorthand for  $\text{dist}_{T_{i_j}}(r, W_{i_j})$ . We observe that for any  $r \in V(T\langle S_{i_j} \rangle)$  we have  $\text{dist}_{T_{i_j}}(r, W_{i_j}) \leq \text{dist}_{T_{i_j}}(r, K_{i_j} \setminus S_{i_j}) \leq 4$ .

**Outline of the proof.** We process the children nodes  $K_{i_j}$  sequentially by increasing size of the minimal separators  $S_{i_j}$ . For that, we start constructing the family  $\mathcal{T}_{S_{i_j}}$  of Theorem 6.1, and we consider the subtrees  $T_{S_{i_j}} \in \mathcal{T}_{S_{i_j}}$  sequentially. We divide the proof into several cases depending on  $|S_{i_j}|$  and on  $\text{diam}(T_{S_{i_j}})$ .

- If  $|S_{i_j}| \leq 2$  then, there can only be  $\mathcal{O}(1)$  different possibilities for the pair  $T_{S_{i_j}}, (d_r)_{r \in V(T_{S_{i_j}})}$ . We can solve DISTANCE-CONSTRAINED ROOT for all these possibilities, thereby obtaining the family  $\mathcal{T}_{i_j}$ . However, for some reasons that will become clearer in Section 8, we only keep in  $\mathcal{T}_{i_j}$  the solutions which satisfy some local optimality criteria. See Section 7.1.
- The processing of the minimal separators  $S_{i_j}$  with at least three elements is more intricate (Sections 7.2 and 7.3). For a fixed  $T_{S_{i_j}}$  we define an encoding with only  $|S_{i_j}|^{\mathcal{O}(1)}$  possibilities, that essentially summarizes at “guessing” the central nodes of  $T\langle K_i \rangle$  and  $T\langle K_{i_j} \rangle$ . Then, we show that only one solution per possible encoding needs to be stored in  $\mathcal{T}_{i_j}$ . The correctness of this part crucially depends on some additional distances’ constraints that are derived from the smaller separators contained into  $S_{i_j}$ , and on Theorem 5.4. Indeed, our approach could not work with an arbitrary rooted clique-tree.

### 7.1 Case $\text{diam}(T_{S_{i_j}}) \leq 1$ .

In this situation,  $|S_{i_j}| \leq 2$  and so, there can only be  $\mathcal{O}(1)$  possibilities for the distances’ constraints  $(d_r)_{r \in S_{i_j}}$ . We can solve DISTANCE-CONSTRAINED ROOT for all possible values, thereby obtaining the family  $\mathcal{T}_{i_j}$ . However, for reasons which will become clearer in the remaining of this Section and in the proofs of Section 8, storing all these possibilities would increase the runtime of our algorithm. We confront this issue with a local optimality criterion. Specifically:

**Claim 7.1.1.** Assume  $S_{i_j} = \{v\}$  and let  $T_{i_j}^{(M)} \in \mathcal{T}_{i_j}$  maximize  $d_{T_{i_j}^{(M)}}(v)$ . If  $T$  is a 4-Steiner root of  $G$  and  $T\langle V_{i_j} \rangle \in \mathcal{T}_{i_j}$  then,  $T\langle V \setminus W_{i_j} \rangle \odot T_{i_j}^{(M)}$  is also a 4-Steiner root of  $G$ .

◇

By Claim 7.1.1, we so conclude that if  $S_{i_j}$  is a cut-vertex then, we can keep exactly one solution in the family  $\mathcal{T}_{i_j}$ .

**Claim 7.1.2.** *Assume  $S_{i_j} = \{u, v\}$ . Let  $T$  be a 4-Steiner root of  $G$  such that  $T\langle S_{i_j} \rangle$  is an edge and  $\text{dist}_T(v, W_{i_j}) \geq \text{dist}_T(u, W_{i_j})$ . Then,  $T' := T\langle V \setminus W_{i_j} \rangle \odot T_{i_j}^{(v)}$  is also a 4-Steiner root of  $G$ , where  $T_{i_j}^{(v)} \in \mathcal{T}_{i_j}$  is, among all solutions in this set such that  $T_{i_j}\langle S_{i_j} \rangle$  is an edge and  $d_{T_{i_j}}(v)$  is maximized, one maximizing  $d_{T_{i_j}}(u)$ . Moreover,  $\text{dist}_{T'}(v, W_{i_j}) \geq \text{dist}_T(v, W_{i_j})$  and  $\text{dist}_{T'}(u, W_{i_j}) \geq \text{dist}_T(u, W_{i_j})$ .*

*Proof.* It suffices to prove  $\text{dist}_{T'}(u, W_{i_j}) \geq \text{dist}_T(u, W_{i_j})$ . We first observe  $d_{T_{i_j}^{(v)}}(v) - d_{T_{i_j}^{(v)}}(u) \leq 1$  because  $T_{i_j}^{(v)}\langle S_{i_j} \rangle$  is an edge. Therefore, there are two cases. Either  $\text{dist}_T(v, W_{i_j}) = d_{T_{i_j}^{(v)}}(v)$  was already maximized, and so we have  $d_{T_{i_j}^{(v)}}(u) \geq \text{dist}_T(u, W_{i_j})$ . Or  $d_{T_{i_j}^{(v)}}(v) \geq \text{dist}_T(v, W_{i_j}) + 1$ , and so  $d_{T_{i_j}^{(v)}}(u) \geq d_{T_{i_j}^{(v)}}(v) - 1 \geq \text{dist}_T(v, W_{i_j}) \geq \text{dist}_T(u, W_{i_j})$ . ◇

By Claim 7.1.2, if  $S_{i_j} = \{u, v\}$  and  $T_{S_{i_j}}$  is an edge then, we only need to keep two solutions, namely: among all those maximizing  $d_{T_{i_j}}(v)$  ( $d_{T_{i_j}}(u)$ , resp.) the one maximizing  $d_{T_{i_j}}(u)$  ( $d_{T_{i_j}}(v)$ , resp.).

## 7.2 Case $T_{S_{i_j}}$ is a non-edge star.

If  $|S_{i_j}| = 2$  then, as already observed in Section 7.1, there can only be  $\mathcal{O}(1)$  different possibilities for the constraints. We can solve DISTANCE-CONSTRAINED ROOT for all possible values, thereby obtaining the family  $\mathcal{T}_{i_j}$ . Thus from now on we assume  $|S_{i_j}| \geq 3$ .

Although there may be exponentially many possible sets of constraints in this case, we show that only a few of the distances' constraints we impose truly need to be considered by our algorithm. Specifically, write  $\mathcal{C}(T_{S_{i_j}}) = \{c\}$ . We consider all possible pairs of nodes  $c_i, c_{i_j}$  that are either in  $S_{i_j}$  or Steiner. For every such a fixed pair, we are interested in the existence of (well-structured) 4-Steiner roots  $T$  of  $G$  such that:  $c_i \in \mathcal{C}(T\langle K_i \rangle) \cap N_T[c]$  and  $c_{i_j} \in \mathcal{C}(T\langle K_{i_j} \rangle) \cap N_T[c]$ . In order for our algorithm to decide whether such Steiner roots exist, we will prove that we only need to consider the distances between  $\mathcal{O}(1)$  nodes in  $T_{S_{i_j}}$  and  $W_{i_j}$ . – In particular, we will prove that we only need to store  $\mathcal{O}(1)$  partial solutions in  $\mathcal{T}_{i_j}$ . – On our way to prove this result, we also use various properties of 4-Steiner powers in order to impose additional distances' constraints on the solutions in  $\mathcal{T}_{i_j}$  which we prove to be necessary in order to extend such a partial solution to all of  $G$ . This second phase is crucial in proving correctness of our approach.

Overall, by Lemma 6.4, we can always relate *any* 4-Steiner root of  $G$  with a pair  $c_i, c_{i_j}$  as defined above. Since there are  $\mathcal{O}(|S_{i_j}|)$  possibilities for every of  $T_{S_{i_j}}, c_i, c_{i_j}$ , we are left with only  $\mathcal{O}(|S_{i_j}|^3)$  different possibilities to store for stars.

Recall that  $T_G$  is a rooted clique-tree of  $G$  as stated in Theorem 5.4. Before starting our analysis, we need to derive a few properties from  $T_G$ . Indeed, our approach could not work with an arbitrary rooted clique-tree.

**Claim 7.1.3.** *If  $S_{i_j}$  strictly contains a minimal separator of  $G_{i_j}$  then, in any 4-Steiner root  $T$  of  $G$ ,  $T\langle S_{i_j} \rangle$  is a bistar.*

*Proof.* Suppose by contradiction  $T\langle S_{i_j} \rangle$  is a star. Any minimal separator of  $G_{i_j}$  that is strictly contained into  $S_{i_j}$  should have size at most two. However, Property P-5.4.2 of Theorem 5.4 ensures that all such minimal separators should have size at least three. A contradiction.  $\diamond$

By Claim 7.1.3, we may assume from now on that  $S_{i_j}$  does not strictly contain any minimal separator of  $G_{i_j}$  (otherwise,  $T_{S_{i_j}}$  cannot be a star, and we are done with this case). We are now left with two possibilities:

### 7.2.1 Subcase no minimal separator of $G_{i_j}$ contains $S_{i_j}$ .

Given any 4-Steiner root  $T_{i_j}$  of  $G_{i_j}$  where  $T_{i_j}\langle S_{i_j} \rangle \equiv_G T_{S_{i_j}}$ , we define an encoding as follows:

$$\text{short-encode}(T_{i_j}) := \langle c, c_{i_j}, d_{T_{i_j}}(c), d_{T_{i_j}}(c_{i_j}) \rangle,$$

where  $c_{i_j} \in \mathcal{C}(T_{i_j}\langle K_{i_j} \rangle) \cap N_{T_{i_j}}[c]$  is arbitrary. The relationship between short encodings and DISTANCE-CONSTRAINED ROOT is discussed at the end of the section. First we prove the following result:

**Claim 7.1.4.** *If  $\text{short-encode}(T_{i_j}) = \text{short-encode}(T'_{i_j})$  and  $T := T^{(0)} \odot T_{i_j}$  is a 4-Steiner root of  $G$  then,  $T' := T^{(0)} \odot T'_{i_j}$  is also a 4-Steiner root of  $G$ .*

*Proof.* It suffices to prove  $d_{T_{i_j}}(v) = d_{T'_{i_j}}(v)$  for every  $v \in S_{i_j}$ . For that, we need to analyze the possible intersections between  $S_{i_j}$  and the minimal separators in  $K_{i_j}$ . Recall that  $S_{i_j}$  is not a minimal separator of  $G_{i_j}$  by the hypothesis.

- Moreover, assume  $c_{i_j} \neq c$ . By Lemma 6.4 we have  $\text{Real}(N_T[c]) = S_{i_j}$ . Combined with the fact that a minimal separator of  $G_{i_j}$  can neither contain  $S_{i_j}$  nor be strictly contained into  $S_{i_j}$  (Claim 7.1.3), this implies all the paths between  $S_{i_j}$  and  $W_{i_j}$  must pass by  $c, c_{i_j}$  (see Fig. 21 for an illustration). In this situation, our partial encoding already contains all the distances' information we need.

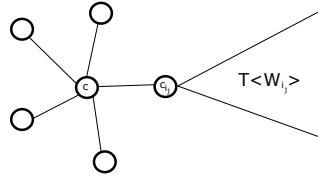


Figure 21: A schematic view of  $T\langle V_{i_j} \rangle$ .

- Otherwise,  $c_{i_j} = c$ . A simple transformation of the construction proposed in Lemma 4.3 shows that we can always assume the simplicial vertices among  $S_{i_j} \setminus \{c\}$  (in  $G_{i_j}$ ) to be leaves adjacent to  $c$  in  $T$ . Namely, we can make all these vertices leaves of  $T\langle K_{i_j} \rangle$  in such a way that they are connected to  $c$  via a path with one Steiner node. We complete this construction by contracting each such simplicial vertex with its Steiner neighbour.

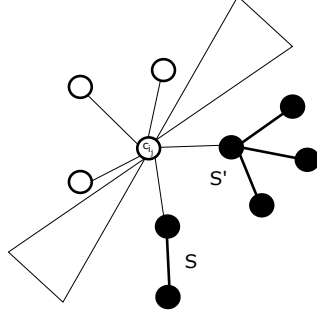


Figure 22: The case  $c = c_{i_j}$ . Two minimal separators  $S, S'$  overlapping  $S_{i_j}$  are drawn in bold.

We end up showing that *all* the vertices in  $S_{i_j} \setminus \{c\}$  that are contained into another minimal separator  $S$  of  $G_{i_j}$  are adjacent in  $T$  to some vertex in  $W_{i_j}$  (hence, their distance to this set is known implicitly and does not need to be stored in the encoding). Indeed, since  $S$  and  $S_{i_j}$  overlap, we cannot have  $T \langle S \rangle$  is a bistar (otherwise,  $S_{i_j} \subseteq S$  by Lemma 5.2). In particular: either  $T \langle S \rangle$  is an edge with exactly one end in  $S_{i_j}$ ; or  $T \langle S \rangle$  is a non edge star and, by Lemma 6.4, the unique vertex in  $(S \cap S_{i_j}) \setminus \{c\}$  is its center. See Fig. 22 for an illustration.

◇

Finally given **short – encode** $(T_{i_j})$ , we can transform such a short encoding into the constraints  $(d_r)_{r \in V(T_{i_j} \langle S_{i_j} \rangle)}$  where:

- $d_c = d_{T_{i_j}}(c)$
- If  $c_{i_j}$  is a real node that is different than  $c$  then,  $d_{c_{i_j}} = d_{T_{i_j}}(c_{i_j})$ ;
- For all other nodes  $r \in S_{i_j}$ :

$$d_r = \begin{cases} d_c + 1 & \text{if } c \neq c_{i_j} \text{ or } r \text{ is simplicial in } G_{i_j} \\ 1 & \text{otherwise.} \end{cases}$$

Note that in doing so,  $d_c \in \{2, 3\}$  and when it is defined  $d_{c_{i_j}} \in \{1, 2\}$ . Overall, there are at most  $2^2 = 4$  possibilities for a fixed pair  $T_{S_{i_j}}, c_{i_j}$ . Furthermore, this above transformation is not injective, and we can so obtain the same constraints for different short encodings (thereby further reducing the size of  $T_{i_j}$ ). The reason why this does not matter is that assuming we made a correct guess for **short – encode** $(T_{i_j})$ , we proved in Claim 7.1.4 that we have  $d_{T_{i_j}}(r) = d_r$  for any  $r \in V(T_{S_{i_j}})$ . In particular, if  $T_{i_j}$  can be extended to a 4-Steiner root of  $G$  then, so could be any partial solution  $T'_{i_j}$  that would satisfy these above constraints as we would have  $d_{T'_{i_j}}(r) \geq d_{T_{i_j}}(r)$  for any  $r \in V(T_{S_{i_j}})$ .

### 7.2.2 Subcase a minimal separator of $G_{i_j}$ contains $S_{i_j}$ .

As for the previous subcase, we start introducing a short encoding then, we explain its relationship with **DISTANCE-CONSTRAINED ROOT** at the end of this section. The novelty here is that we need to complete our encoding with additional distances' conditions, that we will also use in order to



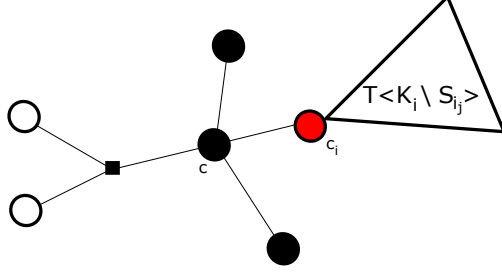


Figure 23: To the proof of Claim 7.1.6.

define our distances' constraints in the input. Specifically, let us fix a pair  $c_i, c_{i_j}$  and let us only consider the partial solutions  $T_{i_j}$  where  $T_{i_j}\langle S_{i_j}\rangle \equiv_G T_{S_{i_j}}$  and  $c_{i_j} \in \mathcal{C}(T_{i_j}\langle K_{i_j}\rangle) \cap N_{T_{i_j}}[c]$ . We set:

$$\text{short - encode - } 2(T_{i_j}) = \left[ \langle c, c_i, d_{T_{i_j}}(c), d_{T_{i_j}}(c_i) \rangle \mid d_{T_{i_j}}(v), \forall v \in S_i \cap S_{i_j} \right].$$

In order to bound the number of possible such encodings, we prove that:

**Claim 7.1.5.**  $|S_i \cap S_{i_j}| = \mathcal{O}(1)$ .

*Proof.* Recall that by Lemma 5.3,  $S_{i_j}$  must be weakly  $T_G$ -convergent. Since we assume  $|S_{i_j}| \geq 3$ , by Property P-5.4.1 of Theorem 5.4,  $S_{i_j}$  is  $T_G$ -convergent. Moreover since there is a minimal separator of  $G_{i_j}$  that contains  $S_{i_j}$ , the maximal clique incident to all edges in  $\bigcup_{S'_{i_j}, S_{i_j} \subseteq S'_{i_j}} E_{S'_{i_j}}(T_G)$  must be  $K_{i_j}$ . This implies that  $S_i := K_i \cap K_{p(i)}$  cannot contain  $S_{i_j}$ . In particular,  $|S_i \cap S_{i_j}| \leq 2$ .  $\diamond$

This new encoding above may not be informative enough in some cases. We complete it with additional *distances' conditions*. Specifically, we consider all the other minimal separators  $S_{i_k} := K_i \cap K_{i_k}$  between  $K_i$  and one of its children nodes such that  $S_{i_k} \subset S_{i_j}$ . Note that since we assume  $T_{S_{i_j}}$  to be a star, we must have  $|S_{i_k}| \leq 2$ . There are two possibilities:

- If  $S_{i_k} = \{v_{i_k}\}$  then, by Claim 7.1.1 there is only one solution left in  $\mathcal{T}_{i_k}$ . Specifically, this solution  $T_{i_k} \in \mathcal{T}_{i_k}$  maximizes  $d_{i_k} := d_{T_{i_k}}(v_{i_k})$ . We are left ensuring  $d_{T_{i_j}}(v_{i_k}) > 4 - d_{i_k}$ .
- Otherwise,  $S_{i_k} = \{u_{i_k}, v_{i_k}\}$ . Then,  $T_{i_j}\langle S_{i_k}\rangle$  must be an edge and we may assume w.l.o.g.  $c_{i_j} = u_{i_k}$ . We are left to ensure that  $d_{T_{i_j}}(v_{i_k}) \geq 2$ .

**Claim 7.1.6.** For any  $T_{i_j}$  that satisfies the above distances' conditions, one of the following properties must be true:

1.  $T_{i_j}$  can be extended to a 4-Steiner root of  $G$ ;
2. For any 4-Steiner root  $T'_{i_j}$  of  $G_{i_j}$  such that  $\text{short - encode - } 2(T_{i_j}) = \text{short - encode - } 2(T'_{i_j})$ , we cannot extend  $T'_{i_j}$  to a well-structured 4-Steiner root of  $G$ .

*Proof.* See Fig. 23 for an illustration. Assume that there is a  $T'_{i_j}$  as stated in the claim that can be extended to a well-structured 4-Steiner root  $T'$  of  $G$ . In order to prove the claim, it suffices to prove that  $T := T'\langle V \setminus W_{i_j}\rangle \odot T_{i_j}$  (or a slight modification of it) is also a 4-Steiner root of  $G$ . For that,

since  $E(G \setminus W_{i_j}) \cup E(G_{i_j}) = E(G)$  (all edges are covered), it suffices to prove  $\text{dist}_T(V \setminus V_{i_j}, W_{i_j}) > 4$ . We start observing that according to Property P-5.4.2 of Theorem 5.4, there is a minimal separator  $S'$  of  $G_{i_j}$  that strictly contains  $S_{i_j}$ . Such a minimal separator  $S'$  must induce a bistar in  $T'$  and  $T$ . Then, as  $c$  must be the center of one of the two maximal cliques containing  $S'$  (i.e., see Lemmata 5.2 and 5.3), we so obtain  $c \neq c_i$ . In particular, all the paths between  $S_{i_j}$  and  $K_i \setminus S_{i_j}$  will need to pass by  $c_i$ , and so:

$$\text{dist}_T(K_i \setminus S_{i_j}, W_{i_j}) = \text{dist}_{T'}(K_i \setminus S_{i_j}, W_{i_j}) > 4.$$

We now prove as a subclaim that we also have  $\text{dist}_T(V \setminus V_i, W_{i_j}) = \text{dist}_{T'}(V \setminus V_i, W_{i_j}) > 4$ . Indeed, in order to have  $\text{dist}_T(u, v) \leq 4$  for some  $u \notin V_i, v \in W_{i_j}$ , we need an intermediate node  $r \in T\langle S_i \rangle \cap T\langle S_{i_j} \rangle$  onto the unique  $uv$ -path in  $T$ . However, such a node  $r$  should be either  $c$  or in  $S_i \cap S_{i_j}$ . By definition of our short encoding, this would imply  $\text{dist}_{T'}(u, v) \leq 4$ , a contradiction.

We finally consider the other minimal separators  $S_{i_k} := K_i \cap K_{i_k}$  between  $K_i$  and one of its children nodes that intersect  $K_{i_j}$ . We have  $S_{i_j} \not\subseteq S_{i_k}$  (otherwise,  $S_{i_j}$  could not be  $T_G$ -convergent, thereby contradicting Property P-5.4.1 of Theorem 5.4). Thus,  $|S_{i_k} \cap S_{i_j}| \leq 2$ . Furthermore since we have  $c \neq c_i$ , Lemma 6.4 implies that either  $S_{i_k} \subset S_{i_j}$  or  $S_{i_k} \cap S_{i_j} \subseteq \{c_i, c\}$ . In the latter case:

$$\text{dist}_T(W_{i_k}, W_{i_j}) = \text{dist}_{T'}(W_{i_k}, W_{i_j}) > 4.$$

Thus from now on assume  $S_{i_k} \subset S_{i_j}$  (and so,  $|S_{i_k}| \leq 2$ ).

If  $S_{i_k}$  is a cut-vertex then, it follows from Claim 7.1.1 and the distances' conditions satisfied by  $T_{i_j}$  that we can always assume  $\text{dist}_T(W_{i_k}, W_{i_j}) > 4$ . We end up with the case  $S_{i_k} = \{u_{i_k}, v_{i_k}\}$ . Then,  $T\langle S_{i_k} \rangle$  must be an edge and we may assume w.l.o.g.  $c = u_{i_k}$ . Since we assume  $T\langle S_{i_j} \rangle$  is a non-edge star,  $c$  is adjacent to some other leaf than  $v_{i_k}$ . In other words,  $\text{dist}_T(c, V \setminus W_{i_k}) = 1$  is minimized. Then, since we have  $\text{dist}_{T'}(W_{i_k}, W_{i_j}) > 4$  we must have  $\text{dist}_{T'}(c, W_{i_k}) = 4$  and so,  $\text{dist}_{T'}(v_{i_k}, W_{i_k}) = 3$ . It follows from Claim 7.1.2 and the distances' conditions satisfied by  $T_{i_j}$  that we can always assume  $\text{dist}_T(W_{i_k}, W_{i_j}) > 4$ .  $\diamond$

Finally given **short – encode – 2**( $T_{i_j}$ ), we can transform such a short encoding into the constraints  $(d_r)_{r \in V(T_{S_{i_j}})}$  where:

- $d_c = d_{T_{i_j}}(c)$
- For any  $v \in S_{i_j} \cap S_i$ ,  $d_v = d_{T_{i_j}}(v)$ .
- If  $c_i$  is a real node that is not in  $S_i \cup \{c\}$  then,  $d_{c_i} = d_{T_{i_j}}(c_i)$ ;
- If  $v_{i_k} \in S_{i_k} \cap S_{i_j}$  has a distance-condition then,  $d_{v_{i_k}}$  is set to the largest such condition.
- Finally, for all other nodes  $r \in V(T_{S_{i_j}})$ :  $d_r = 1$  (trivial constraint).

For any fixed  $T_{S_{i_j}}$  the mapping  $\varphi : \mathbf{short – encode – 2}(T_{i_j}) \rightarrow \langle c_i, (d_r)_{r \in V(T_{S_{i_j}})} \rangle$  is injective. Moreover we proved in Claim 7.1.6 that, in any 4-Steiner root  $T$  of  $G$ , all the paths between  $T\langle S_i \rangle$  and  $W_{i_j}$  must go by  $\{c, c_i\} \cup (S_i \cap S_{i_j})$ . Therefore, our short encodings always preserve a yes-instance of DISTANCE-CONSTRAINED ROOT at  $K_i$  provided one exists.

### 7.3 Case $T_{S_{i_j}}$ is a bistar

We follow the same approach as in Section 7.2. In fact, the proof is a bit simpler in this case. For instance by Lemma 5.2, in *any* 4-Steiner root  $T$  of  $G$  such that  $T\langle S_{i_j} \rangle \equiv_G T_{S_{i_j}}$ , we must have  $\mathcal{C}(T_{S_{i_j}}) = \{c_i, c_{i_j}\}$  where  $c_i$  and  $c_{i_j}$  are the unique central nodes of  $T\langle K_i \rangle$  and  $T\langle K_{i_j} \rangle$ , respectively.

**Preliminaries.** Before choosing our short encoding, we will need to prove some new properties of the rooted clique-tree  $T_G$  (given by Theorem 5.4).

**Claim 7.1.7.** *Let  $T_{i_j}$  be a 4-Steiner root of  $G_{i_j}$  such that  $T_{i_j}\langle S_{i_j} \rangle \equiv_G T_{S_{i_j}}$  is a bistar. If  $T_{i_j}$  can be extended to a 4-Steiner root of  $G$  then, all the vertices in  $N_{T_{i_j}}(c_i) \setminus \{c_{i_j}\}$  are simplicial in  $G_{i_j}$*

*(hence, their distance to  $W_{i_j}$  is always equal to  $d_{T_{i_j}}(c_i) + 1$ ).*

*Proof.* This may not be the case only if some of these vertices are contained into a minimal separator  $S$  of  $G_{i_j}$ . Then, since  $N_{T_{i_j}}[c_i] \setminus \{c_{i_j}\}$  is a connected component of  $T_{i_j}\langle K_{i_j} \rangle \setminus \{c_{i_j}\}$ , we should have  $S \subseteq S_{i_j}$ . By Property P-5.4.2 of Theorem 5.4 we obtain that  $|S| \geq 3$ . Furthermore,  $S \subset S_{i_j}$  because, by Lemma 5.2,  $G \setminus S_{i_j}$  has exactly two full components. This implies that the only possibility for  $T_{i_j}\langle S \rangle$  is a star; moreover,  $Real(N_{T_{i_j}}[c_i]) = S$  by Lemma 5.3. However, by Property P-5.4.2 of Theorem 5.4, there should exist a minimal separator of  $G_{i_j}$  that strictly contains  $S$ . As proved in Lemma 5.3,  $S$  must be weakly  $T_G$ -convergent. Hence, all edges in  $\bigcup_{S', S \subset S'} E_{S'}(T_G)$  must be incident to  $K_{i_j}$ . This implies that in fact, the subtrees  $T_{i_j}\langle S' \rangle$ ,  $S' \supset S$  can only intersect in  $N_{T_{i_j}}[c_{i_j}]$ . So,  $Real(N_{T_{i_j}}[c_{i_j}]) = S$ , that is a contradiction since  $c_{i_j} \neq c_i$ .  $\diamond$

**Claim 7.1.8.** *Let  $T_{i_j}$  be a 4-Steiner root of  $G_{i_j}$  such that  $T_{i_j}\langle S_{i_j} \rangle \equiv_G T_{S_{i_j}}$  is a bistar, and let  $T$  be a 4-Steiner root of  $G$  extending  $T_{i_j}$ . One of the following conditions is true:*

1. *Either every real node in  $N_T(c_{i_j})$  is simplicial in  $G_{i_j}$  or is adjacent in  $T$  to a vertex of  $W_{i_j}$ .*
2. *Or  $|S_i \cap N_T[c_{i_j}]| \leq 2$ , and in the same way  $|S_{i_k} \cap N_T[c_{i_j}]| \leq 2$  for any other child  $K_{i_k}$  of  $K_i$ .*

*Proof.* Recall that according to Lemma 5.2,  $S_{i_j}$  is only contained into the maximal cliques  $K_i$  and  $K_{i_j}$ . In particular if we prove  $K \cap K' \subseteq S_{i_j}$  for some  $\{K, K'\} \neq \{K_i, K_{i_j}\}$  then, this inclusion is strict. We implicitly use this fact in what follows. Specifically if  $S_i \subseteq S_{i_j}$  (resp.,  $S_{i_k} \subseteq S_{i_j}$  for some child  $K_{i_k}$  of  $K_i$ ,  $k \neq j$ , or  $S \subseteq S_{i_j}$  for some minimal separator  $S$  of  $G_{i_j}$ ) then, we must have  $S_i \subset S_{i_j}$  (resp.,  $S_{i_k} \subset S_{i_j}$  or  $S \subset S_{i_j}$ ).

For the remaining of the proof we may assume  $|N_T[c_{i_j}]| \geq 3$  (otherwise we are done). Our proof is based on the following two observations:

1. By Property P-3.4.2 of Theorem 3.4, having  $|S_i \cap N_T[c_{i_j}]| \geq 3$  ( $|S_{i_k} \cap N_T[c_{i_j}]| \geq 3$ , resp.) would imply  $Real(N_T[c_{i_j}]) \subseteq S_i$  ( $Real(N_T[c_{i_j}]) \subseteq S_{i_k}$ , resp.). Since  $N_T[c_{i_j}] \setminus \{c_i\}$  is a connected component of  $T\langle K_i \rangle \setminus \{c_i\}$ , this would in turn imply  $S_i \subset S_{i_j}$  ( $S_{i_k} \subset S_{i_j}$ , resp.). As a result, we would have  $S_i = Real(N_T[c_{i_j}])$  (resp.,  $S_{i_k} = Real(N_T[c_{i_j}])$ ).
2. By Property P-5.4.2 of Theorem 5.4 any minimal separator of  $G_{i_j}$  that is strictly contained into  $S_{i_j}$  must have size at least 3, be  $T_G$ -convergent, and be strictly contained into a minimal separator of  $G_{i_j}$ . So, the only possibility for such a separator is also  $S = Real(N_T[c_{i_j}])$ .

If  $S_{i_j}$  contains a minimal separator  $S$  of  $G_{i_j}$  then, by the above two observations we cannot have  $S = S_i$ , neither  $S = S_{i_k}$  for any other child  $K_{i_k}$  of  $K_i$  (otherwise,  $S$  could not be  $T_G$ -convergent). In particular, we cannot have  $|S_i \cap N_T[c_{i_j}]| \geq 3$ , neither  $|S_{i_k} \cap N_T[c_{i_j}]| \geq 3$ , and so we always fall in Case 2 of the claim.

From now on we assume that  $S_{i_j}$  does not strictly contain any minimal separator of  $G_{i_j}$ . We prove as a subclaim that either  $Real(N_T[c_{i_j}]) \neq S_i$  and  $Real(N_T[c_{i_j}]) \neq S_{i_k}$  for any child  $K_{i_k}$ , or there is no separator of  $G_{i_j}$  that contains  $Real(N_T[c_{i_j}])$ . Indeed, suppose by contradiction  $Real(N_T[c_{i_j}]) = S_i$  and there exists a separator  $S'$  of  $G_{i_j}$  that contains  $Real(N_T[c_{i_j}])$ . It implies by Lemma 5.3  $S_i$  is weakly  $T_G$ -convergent but not  $T_G$ -convergent, thereby contradicting Property P-5.4.1 of Theorem 5.4. The proof for  $S_{i_k}$  is identical as the one above, thereby proving the subclaim. Finally, let us assume either  $Real(N_T[c_{i_j}]) = S_i$  or  $Real(N_T[c_{i_j}]) = S_{i_k}$  for some child  $K_{i_k}$  (otherwise, according to our first observation above, we always fall in Case 2 of the claim). Since no minimal separator of  $G_{i_j}$  can contain  $Real(N_T[c_{i_j}])$ , we can reuse the same proof as for Claim 7.1.4 (Case  $c = c_{i_j}$ ) in order to prove that we always fall in Case 1 of the claim. Specifically, we end up showing that *all* the vertices in  $N_T(c_{i_j})$  that are contained into another minimal separator  $S$  of  $G_{i_j}$  are adjacent in  $T$  to some vertex in  $W_{i_j}$ . Indeed,  $S$  and  $Real(N_T[c_{i_j}])$  must overlap because we assume  $S \not\subseteq S_{i_j}$  and  $Real(N_T[c_{i_j}]) \not\subseteq S$ . Then, we cannot have  $T\langle S \rangle$  is a bistar (otherwise,  $Real(N_T[c_{i_j}]) \subseteq S$  by Lemma 5.2). In particular: either  $T\langle S \rangle$  is an edge with exactly one end in  $N_T[c_{i_j}]$ ; or  $T\langle S \rangle$  is a non edge star and, by Lemma 6.4, the unique vertex in  $(S \cap S_{i_j}) \setminus \{c_{i_j}\}$  is its center.  $\diamond$

**The encoding.** For any 4-Steiner root  $T_{i_j}$  of  $G_{i_j}$  such that  $T_{i_j}\langle S_{i_j} \rangle \equiv_G T_{S_{i_j}}$  is a bistar, we include in our short encoding the following information:

$$\left[ c_i, d_{T_{i_j}}(c_i), d_{T_{i_j}}(c_{i_j}) \right] \text{ and } \begin{cases} \text{only if } |Real(N_{T_{i_j}}[c_{i_j}])| \leq 2) d_{T_{i_j}}(r), \forall r \in N_{T_{i_j}}[c_{i_j}] \\ \text{or (only if } |S_i \cap N_{T_{i_j}}[c_{i_j}]| \leq 2) d_{T_{i_j}}(v), \forall v \in N_{T_{i_j}}[c_{i_j}] \cap S_i \end{cases}$$

As usual, the relationship between this above encoding and DISTANCE-CONSTRAINED ROOT is made explicit at the end of the section. There are only  $\mathcal{O}(1)$  possibilities for a fixed  $T_{S_{i_j}}$ . By Theorem 6.1, we so obtain  $\mathcal{O}(|S_{i_j}|^5)$  different encodings. However, we need to complete this case with similar distances' conditions as for the star case (Section 7.2).

**Additional conditions.** Specifically, assume  $|Real(N_{T_{S_{i_j}}}[c_{i_j}])| \geq 3$  and  $|N_{T_{S_{i_j}}}[c_{i_j}] \cap S_{i_k}| \leq 2$  for any other child  $K_{i_k}$  of  $K_i$  (otherwise, no additional constraint is needed). We consider all the other minimal separators  $S_{i_k} := K_i \cap K_{i_k}$  between  $K_i$  and one of its children nodes such that  $S_{i_k} \subseteq N_{T_{S_{i_j}}}[c_{i_j}]$ . In particular,  $|S_{i_k}| \leq 2$ . There are two possibilities:

- If  $S_{i_k} = \{v_{i_k}\}$  then, by Claim 7.1.1 there is only one solution left in  $\mathcal{T}_{i_k}$ . Specifically, this solution  $T_{i_k} \in \mathcal{T}_{i_k}$  maximizes  $d_{i_k} := d_{T_{i_k}}(v_{i_k})$ . We are left ensuring  $d_{T_{i_j}}(v_{i_k}) > 4 - d_{i_k}$ .
- Otherwise,  $S_{i_k} = \{u_{i_k}, v_{i_k}\}$ . We have  $S_{i_k} \neq Real(N_{T_{i_j}}[c_{i_j}])$ . Then, the tree  $T_{S_{i_j}}\langle S_{i_k} \rangle$  must be an edge and we may assume w.l.o.g.  $c_{i_j} = u_{i_k}$ . We are left to ensure that  $d_{T_{i_j}}(v_{i_k}) \geq 2$ .

**Claim 7.1.9.** *For any  $T_{i_j}$  that satisfies all of the above distances' conditions, one of the following properties is true:*

1. Either  $T_{i_j}$  can be extended to a 4-Steiner root of  $G$ ;
2. Or for any 4-Steiner root  $T'_{i_j}$  of  $G_{i_j}$  with the same short encoding as  $T_{i_j}$ , we cannot extend  $T'_{i_j}$  to a well-structured 4-Steiner root of  $G$ .

*Proof.* Assume there is a  $T'_{i_j}$  as stated in the claim that can be extended to a well-structured 4-Steiner root  $T'$  of  $G$ . In order to prove the claim, it suffices to prove that  $T := T' \langle V \setminus W_{i_j} \rangle \odot T_{i_j}$  (or a slight modification of it) is also a 4-Steiner root of  $G$ . Equivalently, we are left proving that  $\text{dist}_T(V \setminus V_{i_j}, W_{i_j}) > 4$ . By Theorem 4.1, we have  $\text{dist}_T(v, W_{i_j}) = 2 + \text{dist}_{T_{i_j}}(c_i, W_{i_j})$  and in the same way  $\text{dist}_{T'}(v, W_{i_j}) = 2 + \text{dist}_{T'_{i_j}}(c_i, W_{i_j})$  for any simplicial vertex  $v \in K_i$ . In particular:

$$\text{dist}_T(v, W_{i_j}) = \text{dist}_{T'}(v, W_{i_j}) > 4.$$

So, we are left to consider the other minimal separators  $S_{i_k} := K_i \cap K_{i_k}$  between  $K_i$  and any other node (including its father node  $K_{p(i)}$ ). Note that  $S_{i_k}$  cannot both intersect  $N_T(c_{i_j})$  and  $N_T(c_i)$  (otherwise,  $S_{i_k} = S_{i_j}$ , thereby contradicting Lemma 5.2). Moreover by Claim 7.1.7, any real node in  $N_T(c_i) \setminus \{c_{i_j}\}$  is simplicial in  $G_{i_j}$ .

Let us first assume  $S_{i_k} \cap S_{i_j} \subseteq N_T[c_i]$ . If  $i_k \neq i$  then, we have in this situation:

$$\begin{aligned} \text{dist}_T(W_{i_j}, W_{i_k}) &= \min\{\text{dist}_T(W_{i_k}, c_i) + \text{dist}_{T_{i_j}}(c_i, W_{i_j}), \\ &\quad \text{dist}_T(W_{i_k}, c_{i_j}) + \text{dist}_{T_{i_j}}(c_{i_j}, W_{i_j})\}, \end{aligned}$$

and in the same way:

$$\begin{aligned} \text{dist}_{T'}(W_{i_j}, W_{i_k}) &= \min\{\text{dist}_{T'}(W_{i_k}, c_i) + \text{dist}_{T'_{i_j}}(c_i, W_{i_j}), \\ &\quad \text{dist}_{T'}(W_{i_k}, c_{i_j}) + \text{dist}_{T'_{i_j}}(c_{i_j}, W_{i_j})\}. \end{aligned}$$

In particular:

$$\text{dist}_T(W_{i_j}, W_{i_k}) = \text{dist}_{T'}(W_{i_j}, W_{i_k}) > 4.$$

Otherwise,  $i_k = i$ , and we also obtain that:

$$\text{dist}_T(W_{i_j}, V \setminus V_i) = \text{dist}_{T'}(W_{i_j}, V \setminus V_i) > 4.$$

As a result, we are only interested in the situation  $S_{i_k} \cap N_T(c_{i_j}) \neq \emptyset$  – that implies  $S_{i_k} \subseteq N_T[c_{i_j}]$ . We further assume  $|S_{i_k}| \leq 2$  since otherwise, we are done by Case 1 of Claim 7.1.8 and the fact that  $T_{i_j}, T'_{i_j}$  have the same short encoding. Then, there are two cases (*i.e.*, exactly the same as for the star case):

- Assume  $S_{i_k} = \{v_{i_k}\}$ . Then, as explained above (Section 7.1), we only kept in  $\mathcal{T}_{i_k}$  a partial solution  $T_{i_k}$  maximizing  $d_{i_k} := d_{T_{i_k}}(v_{i_k})$ . In this situation, it follows from the distances' conditions over  $T_{i_j}$  that we have (if  $i_k \neq i$ )  $\text{dist}_T(W_{i_k}, W_{i_j}) > 4$  or (if  $i_k = i$ )  $\text{dist}_T(V \setminus V_i, W_{i_j}) > 4$ .
- Otherwise,  $S_{i_k} = \{u_{i_k}, v_{i_k}\}$ . Recall that  $S_{i_k} \subset S_{i_j}$ . We may further assume  $|\text{Real}(N_T[c_{i_j}])| \geq 3$  and  $S_i \neq S_{i_k}$  (otherwise, the encoding already includes the distance to  $W_{i_j}$  from any node in  $T \langle S_{i_k} \rangle$ , and so we are done). Thus,  $T \langle S_{i_k} \rangle$  must be an edge and we may assume w.l.o.g.  $c_{i_j} = u_{i_k}$ . Since we assume  $|\text{Real}(N_T[c_{i_j}])| \geq 3$ ,  $c_{i_j}$  is adjacent to some other real node than  $v_{i_k}$ . In other words,  $\text{dist}_T(c_{i_j}, V \setminus W_{i_k}) = 1$  is minimized. Then, since we have  $\text{dist}_{T'}(W_{i_k}, W_{i_j}) > 4$  we must have  $\text{dist}_{T'}(c_{i_j}, W_{i_k}) = 4$  and so,  $\text{dist}_{T'}(v_{i_k}, W_{i_k}) = 3$ . It follows from Claim 7.1.2 and the distances' conditions over  $T_{i_j}$  that we can always assume  $\text{dist}_T(W_{i_k}, W_{i_j}) > 4$ .

◇

Finally, an encoding for bisters is transformed into distances' constraints as follows:

- If  $d_{T_{i_j}}(r)$  is included in the short encoding then,  $d_r := d_{T_{i_j}}(r)$ . In particular, this will be the case for  $c_i, c_{i_j}$ .
- If  $v_{i_k} \in S_{i_k} \cap S_{i_j}$  needs to satisfy some specified distance-condition then,  $d_{v_{i_k}}$  is set to the largest such a condition.
- For all other vertices  $v \in S_{i_j}$ ,  $d_v = 1$  (trivial constraint).

For any fixed bistar  $T_{S_{i_j}}$  the mapping from the encodings to the distances' constraints is bijective. Indeed, in order to prove it is the case, the only difficulty is to prove that we can correctly identify from the constraints the nodes  $c_i, c_{i_j}$ . Since we will always impose  $d_{c_{i_j}} \leq 2$  whereas  $d_{c_i} \geq 3$ , this is always possible.

## 8 Step S.4: The dynamic programming

In what follows, let  $\|G\| := \sum_{K_i \in \mathcal{K}(G)} |K_i|$ . For a chordal graph,  $\|G\| = \mathcal{O}(n + m)$  [BP93]. We can now state the core result of this paper:

**Theorem 8.1.** *Let  $G = (V, E)$  be strongly chordal, let  $T_G$  be a rooted clique-tree as in Theorem 5.4 and let  $K_i \in \mathcal{K}(G)$ . There is some polynomial  $P$  such that, after a pre-processing in time  $\mathcal{O}(n\|G_i\|^5 P(n))$ , we can solve DISTANCE-CONSTRAINED ROOT for any  $T_{S_i}, (d_r)_{r \in T_{S_i}}$  in time  $\mathcal{O}(P(n))$ .*

Theorem 8.1 proves Theorem B directly. Note that we made no effort in order to improve the running time in our analysis. A very rough analysis shows that we have  $P(n) = \mathcal{O}(n^{10})$ .

*Proof of Theorem 8.1.* If  $K_i$  is a leaf of  $T_G$  then, we construct the family given by Theorem 6.2. We only keep the trees  $T_i \in \mathcal{T}_i$  that satisfy the additional constraints we have. Thus from now on, assume  $K_i$  is an internal node with children  $K_{i_1}, K_{i_2}, \dots, K_{i_p}$ .

**Preprocessing.** Let  $\mathcal{T}_{i_1}, \mathcal{T}_{i_2}, \dots, \mathcal{T}_{i_p}$  be as in Theorem 7.1. By induction on  $T_G$ , the computation of all the  $\mathcal{T}_{i_j}$ 's requires total preprocessing time  $\sum_{j=1}^p \mathcal{O}(n\|G_{i_j}\|^5 P(n))$ , and  $\sum_{j=1}^p \mathcal{O}(|S_{i_j}|^5 P(n))$  additional time. We also need to construct the family  $\mathcal{F}_i$  of Proposition 6.7, that takes  $\mathcal{O}(n|K_i|^8)$ -time. – Recall that the elements in  $\mathcal{F}_i$  are of the form  $(T_{Y_i}, \mathcal{C}_i)$  where  $Y_i \subseteq K_i$  and  $\mathcal{C}_i$  must represent the center of  $T\langle K_i \rangle$  (missing vertices of  $K_i \setminus Y_i$  are supposed to be located in thin branches, see Lemma 6.6). – Overall, if we assume w.l.o.g. that  $P(n) = \Omega(n^3)$  then, this pre-processing phase takes total time:

$$\begin{aligned}
& \sum_{j=1}^p \mathcal{O}(n\|G_{i_j}\|^5 P(n)) + \sum_{j=1}^p \mathcal{O}(|S_{i_j}|^5 P(n)) + \mathcal{O}(n|K_i|^8) \\
&= \mathcal{O}(nP(n) \cdot \sum_{j=1}^p \|G_{i_j}\|^5) + \mathcal{O}(p|K_i|^5 P(n)) + \mathcal{O}(n|K_i|^5 P(n)) \\
&= \mathcal{O}(nP(n) \cdot (\|G_i\|^5 - |K_i|^5)) + \mathcal{O}(n|K_i|^5 P(n)) \\
&= \mathcal{O}(nP(n)\|G_i\|^5).
\end{aligned}$$

**Answering a query.** In what follows let  $T_{S_i}$  and  $(d_r)_{r \in T_{S_i}}$  be fixed. Recall that for every  $(T_{Y_i}, \mathcal{C}_i) \in \mathcal{F}_i$  we have  $S_i \subseteq Y_i$ , and so, we can check whether  $T_{S_i} \equiv_G T_{Y_i} \langle S_i \rangle$ . This takes total time  $\mathcal{O}(|S_i| |\mathcal{F}_i|) = \mathcal{O}(n |K_i|^9)$ . Then, we consider each  $(T_{Y_i}, \mathcal{C}_i) \in \mathcal{F}_i$  that passes this first test above sequentially. Simply put, we use a series of filtering rules in order to greedily find a solution to DISTANCE-CONSTRAINED ROOT, or to correctly conclude that there is none.

Assume first  $Y_i = K_i$  (no thin branch). For every  $r \in T_{S_i}$  we check whether we have:

$$\text{dist}_{T_{Y_i}}(r, K_i \setminus S_i) \geq d_r$$

(otherwise, we violate our distances' constraints). We will assume from now on it is the case. In the same way, for every  $r_{i_j} \in T_{Y_i} \langle S_{i_j} \rangle$ ,  $j \in \{1, 2, \dots, p\}$ , we only keep in  $\mathcal{T}_{i_j}$  those solutions  $T_{i_j}$  such that we have:

$$\text{dist}_{T_{Y_i}}(r, r_{i_j}) + d_{T_{i_j}}(r_{i_j}) \geq d_r.$$

Overall, since  $|\mathcal{T}_{i_j}| = \mathcal{O}(|S_{i_j}|^5) = \mathcal{O}(|K_i|^5)$ , this new verification phase takes total time  $\mathcal{O}(p |S_i| |K_i|^5) = \mathcal{O}(n |K_i|^6)$ . Furthermore in doing so, we ensure that *any* 4-Steiner root of  $G_i$  that we can obtain from  $T_{Y_i}$  and the remaining solutions in the  $\mathcal{T}_{i_j}$ 's will satisfy all our distances' constraints. Conversely, if no such a solution can be found then, we can correctly report that our distances' constraints cannot be satisfied in any well-structured 4-Steiner root of  $G$  (by Theorem 7.1).

We now introduce another filtering rule, quite similar as the one above, that we will keep using throughout the remaining of the proof. Specifically, for every  $j \in \{1, 2, \dots, p\}$  and  $r_{i_j} \in T_{Y_i} \langle S_{i_j} \rangle$ , we assign some value  $\ell^{i_j}(r_{i_j})$  that intuitively represents the distance of  $r_{i_j}$  to  $V_i \setminus V_{i_j}$ . Every time the rule is applied, we discard all solutions  $T_{i_j} \in \mathcal{T}_{i_j}$  such that  $d_{T_{i_j}}(r_{i_j}) + \ell^{i_j}(r_{i_j}) \leq 4$ . We set initially  $\ell^{i_j}(r_{i_j}) := \text{dist}_{T_{Y_i}}(r_{i_j}, K_i \setminus S_{i_j})$  and we apply the rule. Overall, updating (initializing, resp.) the values  $\ell^{i_j}$  for every  $j$  takes time  $\sum_{j=1}^p \mathcal{O}(|S_{i_j}|) = \mathcal{O}(n |K_i|)$ . Applying the rule takes time  $\sum_{j=1}^p \mathcal{O}(|S_{i_j}| |\mathcal{T}_{i_j}|) = \sum_{j=1}^p \mathcal{O}(|S_{i_j}|^6) = \mathcal{O}(n |K_i|^6)$ .

In what follows, we explain how to greedily construct a solution (if any), starting from  $T_i := T_{Y_i}$ . The procedure is divided into a constant number of phases. Every time we complete one of these phases, we need to apply this above filtering rule.

- *Phase 1: Processing the cut-vertices.* We consider all the indices  $j \in \{1, 2, \dots, p\}$  such that  $S_{i_j} = \{v\}$  is a cut-vertex. By Claim 7.1.1 there is exactly one solution left in  $\mathcal{T}_{i_j}$ . We add it to the solution, *i.e.*, we set  $T_i := T_i \odot T_{i_j}$ . Furthermore, for every  $k \in \{1, 2, \dots, p\} \setminus \{j\}$  and  $r_{i_k} \in T_{Y_i} \langle S_{i_k} \rangle$  (possibly,  $r_{i_k} = v$ ) we set  $\ell^{i_k}(r_{i_k}) := \min\{\ell^{i_k}(r_{i_k}), \text{dist}_{T_{Y_i}}(r_{i_k}, v) + d_{T_{i_j}}(v)\}$ . We end up applying the filtering rule above.
- *Phase 2: Processing the edges.* We consider all the indices  $j \in \{1, 2, \dots, p\}$  such that  $S_{i_j} = \{u, v\}$  and  $T_{Y_i} \langle S_{i_j} \rangle$  is an edge. The following claim shows that we can proceed similarly as for Phase 1 provided we know which among  $u$  or  $v$  will be closest to  $V_i \setminus V_{i_j}$ . Therefore, computing this information is the main objective of this phase.

**Claim 8.1.1.** *Assume  $S_{i_j} = \{u, v\}$ . Let  $T$  be a 4-Steiner root of  $G$  such that  $T \langle S_{i_j} \rangle$  is an edge and  $\text{dist}_T(u, V \setminus V_{i_j}) \geq \text{dist}_T(v, V \setminus V_{i_j})$ . Then,  $T' := T \langle V \setminus W_{i_j} \rangle \odot T_{i_j}^{(v)}$  is also a 4-Steiner root of  $G$ , where  $T_{i_j}^{(v)} \in \mathcal{T}_{i_j}$  is, among all solutions in this set such that  $T_{i_j}^{(v)} \langle S_{i_j} \rangle$  is an edge and  $d_{T_{i_j}^{(v)}}(v)$  is maximized, one maximizing  $d_{T_{i_j}^{(v)}}(u)$ .*

*Proof.* By maximality of  $d_{T_{i_j}^{(v)}}(v)$  the resulting  $T'$  would not be a 4-Steiner root of  $G$  only if  $\text{dist}_T(u, V \setminus V_{i_j}) + d_{T_{i_j}^{(v)}}(u) \leq 4$ . But then, since  $d_{T_{i_j}^{(v)}}(v) - d_{T_{i_j}^{(v)}}(u) \leq 1$  (because  $T \langle S_{i_j} \rangle$  is an edge), one would obtain  $\text{dist}_T(u, V \setminus V_{i_j}) = \text{dist}_T(v, V \setminus V_{i_j})$  and  $\text{dist}_T(v, V \setminus V_{i_j}) + d_{T_{i_j}^{(v)}}(v) = 5$ . In particular, we should have in the original Steiner root  $T$ :

$$\min\{\text{dist}_T(u, W_{i_j}), \text{dist}_T(v, W_{i_j})\} \geq d_{T_{i_j}^{(v)}}(v).$$

As  $T_{i_j}^{(v)}$  maximizes  $d_{T_{i_j}^{(v)}}(v)$  and, under this latter condition,  $d_{T_{i_j}^{(v)}}(u)$  is maximized, we obtain that  $d_{T_{i_j}^{(v)}}(u) \geq \text{dist}_T(u, W_{i_j}) \geq d_{T_{i_j}^{(v)}}(v)$ .  $\diamond$

By Claim 8.1.1 we are left to decide which amongst  $u$  or  $v$  will minimize its distance to  $V_i \setminus V_{i_j}$  in the final solution. If either  $u$  or  $v$  has a real neighbour in  $T_i \setminus S_{i_j}$  then, we are done. Thus from now on we assume this is not the case.

There may be several other indices  $k$  such that  $S_{i_k} = S_{i_j}$ . As an intermediate step, we explain how to merge the solutions in  $\mathcal{T}_{i_j}$  and in  $\mathcal{T}_{i_k}$  into a new set  $\mathcal{T}'_{i_j}$  when this happens. For that, we consider all the  $T_{i_j}, T_{i_k}$  sequentially. We put  $T_{i_j} \odot T_{i_k}$  into  $\mathcal{T}'_{i_j}$  if and only if we have  $\min\{d_{T_{i_j}}(v) + d_{T_{i_k}}(v), d_{T_{i_j}}(u) + d_{T_{i_k}}(u)\} > 4$ . If so then,  $d_{T_{i_j} \odot T_{i_k}}(u) = \min\{d_{T_{i_j}}(u), d_{T_{i_k}}(u)\}$ , and in the same way  $d_{T_{i_j} \odot T_{i_k}}(v) = \min\{d_{T_{i_j}}(v), d_{T_{i_k}}(v)\}$ . Overall, since there are at most two solutions stored in each of  $\mathcal{T}_{i_j}$  and  $\mathcal{T}_{i_k}$ , this takes constant-time. We end up applying Claim 7.1.2 in order to replace  $\mathcal{T}_{i_j}$  by the at most two best solutions in  $\mathcal{T}'_{i_j}$ . By repeating this above procedure, we can assume w.l.o.g. that there is no other index  $k$  such that  $S_{i_k} = S_{i_j}$ .

We may further assume that there is no index  $k$  such that  $S_{i_k} = \{u\}$  ( $S_{i_k} = \{v\}$ , resp.) for otherwise we already ensured at the last step  $d_{T_{i_k}}(u) = 4$  ( $d_{T_{i_k}}(v) = 4$ , resp.). Then, let us assume  $\text{dist}_{T_i}(u, C_i) \leq \text{dist}_{T_i}(v, C_i)$  ( $u$  is closer than  $v$  to the center of  $T_{Y_i}$ ). In most cases,  $u$  will be the closest to  $V_i \setminus V_{i_j}$ . Indeed, as we assume  $v$  has no real neighbour in  $T_i \setminus S_{i_j}$ , it is a leaf in  $T_{Y_i}$ . Therefore, a necessary condition for having  $v$  closer than  $u$  to  $V_i \setminus V_{i_j}$  is that there exists another minimal separator  $S_{i_k}$  containing  $v$ . Since  $v$  is a leaf, this implies  $S_{i_j} \subset S_{i_k}$ . In particular, as we also assume  $u$  and  $v$  have no real neighbour in  $T_i \setminus S_{i_j}$ ,  $T_{Y_i} \langle S_{i_k} \rangle$  must be a bistar. We divide our analysis in several subcases:

- Subcase  $S_{i_j} = \mathcal{C}(T_{Y_i} \langle S_{i_k} \rangle)$ . By Claim 6.1.1, there should be a heavy part in  $S_{i_k}$ , and so, one of  $u$  or  $v$  should have a real neighbour in  $T_i$ . A contradiction.
- Subcase  $C_i = \{u\}$ . By Claim 7.1.7,  $v$  is simplicial in  $G_{i_k}$ . This proves  $u$  will be closest than  $v$  to  $V_i \setminus V_{i_j}$  in this subcase.
- Otherwise, as  $u$  minimizes its distance to the center we must have  $\mathcal{C}(T_{i_k} \langle K_{i_k} \rangle) = \{u\}$  (this can only be true for at most one index  $i_k$ ). Note that  $v$  is the only leaf of  $T_{Y_i} \langle S_{i_k} \rangle$  that is adjacent to  $u$ . Therefore, by Lemma 5.2,  $v$  is the only real neighbour of  $u$  in any  $T_{i_k} \in \mathcal{T}_{i_k}$  and  $v$  must stay the only real neighbour of  $u$  in  $T_i$ . This implies that we always have  $d_{T_{i_k}}(u) = 2$ . We must ensure that the solution  $T_{i_j} \in \mathcal{T}_{i_j}$  that we will choose satisfies  $d_{T_{i_k}}(u) \geq 3$ . Conversely, among all the partial solutions in  $\mathcal{T}_{i_j}$  that satisfies this necessary condition, we can always choose the one  $T_{i_j}$  maximizing  $d_{T_{i_j}}(v)$ .



- *Phase 3: Processing the bistars.* We consider all the indices  $j \in \{1, 2, \dots, p\}$  such that  $T_{Y_i}\langle S_{i_j} \rangle$  is a bistar. Let  $\mathcal{C}_i = \{c_i\}$  and let  $\mathcal{C}(T_{Y_i}\langle S_{i_j} \rangle) = \{c_i, c_{i_j}\}$ . We keep only the solutions  $T_{i_j} \in \mathcal{T}_{i_j}$  such that  $\mathcal{C}(T_{i_j}\langle K_{i_j} \rangle) = \{c_{i_j}\}$ . Then, we have by Claim 7.1.7  $d_{T_{i_j}}(v) = d_{T_{i_j}}(c_i) + 1$  for any  $v \in N_{T_{i_j}}(c_i) \setminus \{c_{i_j}\}$ . So, we would like to pick  $T_{i_j} \in \mathcal{T}_{i_j}$  that maximizes  $d_{T_{i_j}}(c_i)$ . We prove next that we can always choose greedily such a partial solution  $T_{i_j}$ . Indeed, the only case where we cannot do that w.l.o.g. is when there exists another minimal separator  $S_{i_k}$  such that  $T_{Y_i}\langle S_{i_k} \rangle \cap N_{T_{Y_i}}[c_{i_j}] \setminus \{c_i\} \neq \emptyset$ . In fact, as by Lemma 5.2 we will always have  $d_{T_{i_j}}(c_{i_j}) = 2$  (for any choice for  $T_{i_j}$ ), we are only interested in the case when  $S_{i_k} \subseteq N_{T_{i_j}}[c_{i_j}]$ . We may further assume  $|S_{i_k}| \geq 3$  (otherwise, due to Phases 1 and 2, this was already taken into account).

However, by Claim 7.1.8 and its proof, this implies that there is only one possibility for  $d_{T_{i_j}}(r)$ , for every  $r \in T_{Y_i}\langle S_{i_j} \rangle$ . Specifically (Case 1 of the claim),  $d_{T_{i_j}}(c_{i_j}) = 2$ , and for every  $u \in N_{T_{Y_i}\langle S_{i_j} \rangle}(c_{i_j})$  either  $d_{T_{i_j}}(u) = 3$  or (if and only if  $u$  belongs to a minimal separator of  $G_{i_j}$ )  $d_{T_{i_j}}(u) = 1$ . So, in this situation, there is only one solution stored in  $\mathcal{T}_{i_j}$ , and we need to pick this one. As in the two previous phases, for every  $k \in \{1, 2, \dots, p\} \setminus \{j\}$  and  $r_{i_k} \in T_{Y_i}\langle S_{i_k} \rangle$  we update  $\ell^{i_k}(r_{i_k})$  and then, we end up applying the filtering rule above.

- *Phase 4: Processing the stars.* We finally consider all the indices  $j \in \{1, 2, \dots, p\}$  such that  $T_{Y_i}\langle S_{i_j} \rangle$  is a star. Let  $\mathcal{C}(T_{Y_i}\langle S_{i_j} \rangle) = \{c\}$ . We divide the analysis in two subphases:

- *Subphase 4.a: Processing a star when  $c \in \mathcal{C}_i$ .* As a guidance towards our next choices, we start analyzing the possibilities we still have among  $\mathcal{T}_{i_j}$ :

**Claim 8.1.2.** *The following properties are true for any  $T_{i_j} \in \mathcal{T}_{i_j}$ :*

1.  $\text{diam}(T_{i_j}\langle K_{i_j} \rangle) = 4$ ;
2. *and the unique center node  $v_j \in \mathcal{C}(T_{i_j}\langle K_{i_j} \rangle)$  is either in  $S_{i_j} \setminus \mathcal{C}_i$ , or it is a Steiner node. Moreover:*
  - (a) *if  $v_j \in S_{i_j}$  then,  $v_j$  is a leaf of  $T_{Y_i}$ ;*
  - (b) *and for every  $r \in T_{Y_i}\langle S_{i_j} \rangle$ ,  $d_{T_{i_j}}(r) = \text{dist}_{T_{i_j}}(r, v_j) + d_{T_{i_j}}(v_j)$ .*

*Proof.* We show that assuming any of these above properties does not hold, some distances' constraints would be violated w.r.t. our previous choices in the other Phases, and so, we should have discarded  $T_{i_j}$  when we applied the filtering rule. Suppose by contradiction  $\text{diam}(T_{i_j}\langle K_{i_j} \rangle) < 4$ . Then, the only possibility is  $\text{diam}(T_{i_j}\langle K_{i_j} \rangle) = 3$ , and so,  $c \in \mathcal{C}(T_{i_j}\langle K_{i_j} \rangle)$ . However, the latter would contradict Lemma 3.5 as we already assume  $c \in \mathcal{C}_i$ . Therefore,  $\text{diam}(T_{i_j}\langle K_{i_j} \rangle) = 4$ , thereby implying  $\mathcal{C}(T_{i_j}\langle K_{i_j} \rangle) = \{v_j\}$  for some  $v_j$ .

By Lemma 6.4,  $\text{Real}(N_{T_{i_j}}[c]) = S_{i_j}$  and  $c \in N_{T_{i_j}}[v_j]$ . Thus, either  $v_j$  is Steiner, or  $v_j \in S_{i_j}$ . Furthermore if  $v_j \in S_{i_j}$  then,  $v_j \in S_{i_j} \setminus \mathcal{C}_i$  (otherwise, this would contradict Lemma 3.5). As every real node adjacent to  $v_j$  in  $T_{Y_i}$  should be in  $K_{i_j}$ , and we have  $K_i \cap K_{i_j} = S_{i_j}$ , we so obtain that  $v_j \in S_{i_j} \implies v_j$  is a leaf-node of  $T_{Y_i}$ .

Finally, we prove that for every  $r \in T_{Y_i}\langle S_{i_j} \rangle$ ,  $d_{T_{i_j}}(r) = \text{dist}_{T_{i_j}}(r, v_j) + d_{T_{i_j}}(v_j)$ . Suppose for the sake of contradiction that it is not the case. We recall that by Lemma 6.4 we have  $\text{Real}(N_{T_{i_j}}[c]) = S_{i_j}$ . This implies that  $T_{Y_i}\langle S_{i_j} \rangle \setminus \{v_j\}$  is a connected component of  $T_{i_j}\langle K_{i_j} \rangle \setminus \{v_j\}$ . In particular, there must exist a minimal separator  $S$  of  $G_{i_j}$  that

intersects  $T_{Y_i}\langle S_{i_j}\rangle \setminus \{v_j\}$  (otherwise, the equality  $d_{T_{i_j}}(r) = \text{dist}_{T_{i_j}}(r, v_j) + d_{T_{i_j}}(v_j)$  would hold for any node  $r \in T_{Y_i}\langle S_{i_j}\rangle$ ). By Claim 7.1.3,  $S \not\subseteq S_{i_j}$ . So, we are only left with two subcases:

- \* Subcase  $S_{i_j} \not\subseteq S$ . Then,  $S_{i_j} \cap S \subseteq \{c, v_j\}$ . However if  $c \in S$  then, let  $C$  be the connected component of  $G_{i_j} \setminus S$  containing  $S_{i_j} \setminus S$ . We must have  $\text{dist}_{T_{i_j}}(c, V_{i_j} \setminus (S \cup C)) = 4$  (otherwise  $\text{dist}_{T_{i_j}}(S_{i_j} \setminus S, V_{i_j} \setminus (S \cup C)) \leq 4$ , a contradiction). We are done in this subcase as we always have  $1 + d_{T_{i_j}}(v_j) \leq 4$ .
- \* Subcase  $S_{i_j} \subseteq S$ . By Property P-5.4.2 of Theorem 5.4 we can assume w.l.o.g.  $S$  strictly contains  $S_{i_j}$ . However, this would imply  $T_{i_j}\langle S \rangle$  is a bistar, and so, by Lemma 5.3,  $c$  would also be in  $\mathcal{C}(T_{i_j}\langle K \rangle)$  for some maximal clique  $K$  in  $G_{i_j}$ . A contradiction.

Overall, since we reached a contradiction in both subcases, we proved as claimed that for every  $r \in T_{Y_i}\langle S_{i_j}\rangle$ ,  $d_{T_{i_j}}(r) = \text{dist}_{T_{i_j}}(r, v_j) + d_{T_{i_j}}(v_j)$ .  $\diamond$

For any  $T_{i_j} \in \mathcal{T}_{i_j}$ , let  $v_j$  be as defined in Claim 8.1.2. Let  $S_{i_k} \neq S_{i_j}$  be unprocessed (in particular,  $T_{Y_i}\langle S_{i_k}\rangle$  is a non-edge star). Since either  $v_j$  is Steiner or  $v_j$  is a leaf of  $T_{Y_i}$ , we have  $v_j \notin T_{Y_i}\langle S_{i_k}\rangle$ . Furthermore for every  $r \in T_{Y_i}\langle S_{i_j}\rangle$ ,  $d_{T_{i_j}}(r) = \text{dist}_{T_{i_j}}(r, v_j) + d_{T_{i_j}}(v_j)$ . This implies that w.r.t.  $S_{i_k}$ , any solution  $T_{i_j}$  that maximizes  $d_{T_{i_j}}(\mathcal{C}(T_{i_j}\langle K_{i_j}\rangle))$  would be a best possible choice – *i.e.*, regardless of our exact choice for  $v_j$ . However, we also need to account for the other indices  $k$  such that  $S_{i_k} = S_{i_j}$ .

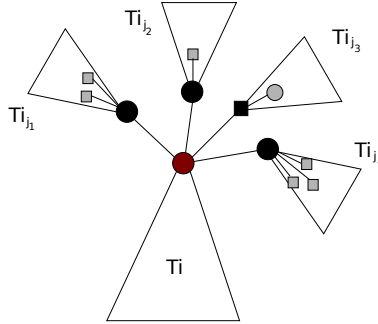


Figure 24: An illustration of Phase 4.

Let  $J = \{j' \mid S_{i_{j'}} = S_{i_j}\}$ . One should ensure that in the solutions  $T_{i_{j'}}, j' \in J$  that we will choose, the center nodes  $v_{j'}$  in  $T_{i_{j'}}\langle K_{i_{j'}}\rangle$  will be pairwise different. Furthermore, since all the  $v_{j'}$ 's are pairwise at distance two, there can be at most one  $j_{\min} \in J$  such that  $d_{T_{i_{j_{\min}}}}(v_{j_{\min}}) = 1$ . See Fig. 24 for an illustration. In order to satisfy all these constraints, while ensuring that such a  $j_{\min}$  does not exist if it is possible, we make a reduction to MAXIMUM-WEIGHT MATCHING [DPS18].

1. Specifically, let  $\text{Steiner}[J] := \{\alpha_{j'} \mid j' \in J\}$  be a set of Steiner nodes. We construct a bipartite graph  $\text{Bip}(S_{i_j})$  with respective sides  $J$  and  $(S_{i_j} \setminus \mathcal{C}_i) \cup \text{Steiner}[J]$ .
2. For every  $j' \in J$  and  $v \in S_{i_j} \setminus \mathcal{C}_i$ , there is an edge  $j'v$  if there exists a  $T_{i_{j'}} \in \mathcal{T}_{i_j}$  such that  $\mathcal{C}(T_{i_{j'}}\langle K_{i_{j'}}\rangle) = \{v\}$ . Furthermore if such a  $T_{i_{j'}}$  exists then, we choose

one maximizing  $d_{T_{i_{j'}}}(v)$  and we assign the weight  $d_{T_{i_{j'}}}(v)$  to the edge  $j'v$  (this can either be 1 or 2).

In the same way, there is an edge  $j'\alpha_{j'}$  if there exists a  $T_{i_{j'}} \in \mathcal{T}_{i_{j'}}$  such that the unique node in  $\mathcal{C}(T_{i_{j'}} \langle K_{i_{j'}} \rangle)$  is Steiner. Furthermore if such a  $T_{i_{j'}}$  exists then, we choose one maximizing  $d_{T_{i_{j'}}}(\mathcal{C}(T_{i_{j'}} \langle K_{i_{j'}} \rangle))$  and we assign the weight  $d_{T_{i_{j'}}}(\mathcal{C}(T_{i_{j'}} \langle K_{i_{j'}} \rangle))$  to the edge  $j'\alpha_{j'}$ .

3. We compute a matching in  $Bip(S_{i_j})$  of maximum total weight. This takes  $\mathcal{O}(n^{5/2})$ -time [DPS18]. By construction, such a matching should contain an edge incident to every  $j' \in J$ , and its total weight should be either  $2|J| - 1$  (if  $j_{\min}$  exists) or  $2|J|$ .

For every  $j' \in J$ , we pick a solution  $T_{i_{j'}}$  corresponding to the edge incident to  $j'$  in the matching. Then, as in all previous phases, we end up applying our filtering rule above.

- *Subphase 4.b: Processing a star when  $c \notin \mathcal{C}_i$ .* This situation can happen only if  $\text{diam}(T_{Y_i}) = 4$ . Then, the unique node  $c_i \in \mathcal{C}_i$  is a neighbour of  $c$  in  $T_i$  (by Lemma 6.4). We may further assume that, if  $\text{Real}(N_{T_i}[c_i])$  is a minimal separator  $S$  then, we already handled with  $S$  during the previous subphase. Similarly, we already handled with any minimal separator strictly contained into  $S_{i_j}$ , strictly containing  $S_{i_j}$  resp., during the previous phases. Hence, the unique path in  $T_i$  between  $T_{Y_i} \langle S_{i_j} \rangle$  and any other  $T_{Y_i} \langle S_{i_k} \rangle$  that we did not process yet goes by  $c_i$ . We are left with finding a solution  $T_{i_j} \in \mathcal{T}_{i_j}$  maximizing  $d_{T_{i_j}}(c_i)$ . However, as in the previous subphase we also need to account for the other indices  $k$  such that  $S_{i_k} = S_{i_j}$ .

Let  $J = \{j' \mid S_{i_{j'}} = S_{i_j}\}$ . We may assume  $|J| \geq 2$  since otherwise, we are done by taking any solution  $T_{i_j} \in \mathcal{T}_{i_j}$  that maximizes  $d_{T_{i_j}}(c_i)$  (i.e., as explained above). Since  $S_{i_j}$  must be weakly  $T_G$ -convergent (Lemma 5.3), and so,  $T_G$ -convergent (Property P-5.4.1 of Theorem 5.4), it implies that, for any  $j' \in J$ , there can be no minimal separator of  $G_{i_{j'}}$  that contains  $S_{i_j}$ . However, an additional difficulty compared to the previous subphase is that now the center  $c$  of the star can also be in  $\mathcal{C}(T_{i_{j'}} \langle K_{i_{j'}} \rangle)$ . So, we need to modify our approach in the previous subphase as follows:

1. We first choose the unique  $j_0 \in J$  such that  $c \in \mathcal{C}(T_{i_{j_0}} \langle K_{i_{j_0}} \rangle)$  (if any). Then, we choose a corresponding solution in  $\mathcal{T}_{i_{j_0}}$  among  $\mathcal{O}(|S_{i_{j_0}}|^5) = \mathcal{O}(|K_i|^5)$  possibilities. Overall, there are  $\mathcal{O}(n|K_i|^5)$  possibilities. We test each such a possibility sequentially (including the case where no such a  $j_0$  exists).
2. Assume for this step that we fixed a value for  $j_0$  (if we test for the case where no such a  $j_0$  exists then, we can go directly to the next step). By Claim 7.1.4, the following property holds for any  $v \in S_{i_j} \setminus \{c\}$ : either  $v$  is simplicial in  $G_{i_{j_0}}$  (and so,  $d_{T_{i_{j_0}}}(v) = 3$ ), or  $d_{T_{i_{j_0}}}(v) = 1$ . In the former case we discard all solutions  $T_{i_{j'}} \in \mathcal{T}_{i_{j'}}$ ,  $j' \in J \setminus \{j_0\}$  such that  $d_{T_{i_{j'}}}(v) < 2$  whereas in the latter case, we discard all solutions  $T_{i_{j'}} \in \mathcal{T}_{i_{j'}}$ ,  $j' \in J \setminus \{j_0\}$  such that  $d_{T_{i_{j'}}}(v) < 4$ .
3. Finally, we apply our reduction to MAXIMUM-WEIGHT MATCHING (from the previous subphase) in order to pick the solutions  $T_{i_{j'}} \in \mathcal{T}_{i_{j'}}$  for every  $j' \in J \setminus \{j_0\}$ . Indeed, we observe that for every  $j' \in J \setminus \{j_0\}$ , we will always obtain  $d_{T_{i_{j'}}}(c_i) =$

$2 + d_{T_{i_{j'}}}(C(T_{i_{j'}} \langle K_{i_{j'}} \rangle)) \in \{3, 4\}$ . Therefore, our choice for the partial solutions  $T_{i_{j'}}$  will ensure that  $d_{T_{i_{j'}}}(c_i)$  is maximized (w.r.t. our choice for  $j_0$  and  $T_{i_{j_0}}$ ).

4. Overall, among all the valid solutions computed (for any possible choice of  $j_0$  and  $T_{i_{j_0}}$ ), we keep the one maximizing  $\min_{j' \in J} d_{T_{i_{j'}}}(c_i)$ .

This last phase concludes the algorithm.

In order to complete the proof, let us finally assume  $Y_i \neq K_i$  (there are thin branches). Then,  $\mathcal{C}_i = \{c_i\}$ . We consider all the minimal separators  $S_{j_1}, S_{j_2}, \dots, S_{j_q} \subseteq (K_i \setminus Y_i) \cup \{c_i\}$  sequentially. For every  $\ell \in \{1, 2, \dots, q\}$  we must have  $T \langle S_{j_\ell} \rangle$  is a thin branch, and so, a star. We so have  $\mathcal{O}(|S_{j_\ell}|) = \mathcal{O}(|K_i|)$  possibilities. Furthermore, since according to Definition 6.1 there can be no minimal separator  $S_{i_k}$  which intersects both  $S_{j_\ell}$  and  $K_i \setminus S_{j_\ell}$ , any solution  $T_{j_\ell} \in \mathcal{T}_{j_\ell}$  that maximizes  $d_{T_{j_\ell}}(c_i)$  would be a best possible choice. This latter case resembles to the situation we met in Subphase 4.b. We can solve it by using the same tools as for this subphase. Specifically:

1. We consider each possibility for the star  $T \langle S_{j_\ell} \rangle$  sequentially;
2. Given a fixed  $T \langle S_{j_\ell} \rangle$ , every minimal separator  $S_{i_k} \subset S_{j_\ell}$  must be either a cut-vertex or induce an edge (otherwise, we can discard this possibility for  $T \langle S_{j_\ell} \rangle$ ). Then, we can process such minimal separators  $S_{i_k}$  as in Phases 1 and 2 above.
3. We end up applying the same procedure as for Subphase 4.b. in order to select the partial solutions  $T_{i_{j'}}$  such that  $S_{i_{j'}} = S_{i_j}$ . Namely, this procedure combines a brute-force enumeration with our reduction to MAXIMUM-WEIGHT MATCHING.
4. Overall, among all the valid solutions computed, we keep the one maximizing  $d(c_i)$ . Then, we can apply our filtering rule above.

Once we applied this above procedure to all the thin branches, we can reuse our previous four-phase algorithm in order to process all the other minimal separators.  $\square$

## 9 Conclusion

There are essentially two dominant approaches in order to solve  $k$ -LEAF POWER and  $k$ -STEINER POWER in the literature. The first one, and by far the most elegant, is based on structural characterization of the corresponding graph classes [BL06, BLS08]. Unfortunately such characterizations – mostly based on forbidden induced subgraphs – look challenging to derive for larger values of  $k$ . Furthermore, some recent work suggests that even a nice characterization of  $k$ -leaf powers ( $k$ -Steiner powers, resp.) by forbidden induced subgraphs might not be enough in order to obtain a polynomial-time recognition algorithm [Laf17].

The second approach consists in a clever use of dynamic programming. Although this approach is much less satisfying on the graph-theoretic side, it may be more promising than the first one. For instance, the only known algorithms so far for recognizing 5-leaf powers and 3-Steiner powers are based on this approach [CK07]. Unfortunately, standard dynamic programming techniques are challenging to apply as the value of  $k$  increases, which is probably why no improvement has been obtained for this problem for over a decade – until this paper.

We propose several new avenues for research on dynamic programming algorithms for  $k$ -leaf powers and  $k$ -Steiner powers. In particular, we hope that our structural analysis of these roots – based on a renewed interest for clique-intersections – can be helpful in order to generalize our algorithmic framework to larger values of  $k$ . Some of our side contributions, especially the design of a problem-specific clique-tree and our greedy procedures in order to select partial solutions, can also be of independent interest for future research on this topic.

## Acknowledgements

The author would like to thank the Research Institute of the University of Bucharest (ICUB) that founded the postdoctoral project of which this paper is the outcome.

## References

- [ABNT16] S. Arumugam, A. Brandstädt, T. Nishizeki, and K. Thulasiraman. *Handbook of graph theory, combinatorial optimization, and algorithms*. Chapman and Hall/CRC, 2016.
- [BFW92] H. Bodlaender, . Fellows, and T. Warnow. Two strikes against perfect phylogeny. In *ICALP*, pages 273–283. Springer, 1992.
- [BH08] A. Brandstädt and C. Hundt. Ptolemaic graphs and interval graphs are leaf powers. In *Latin American Symposium on Theoretical Informatics*, pages 479–491. Springer, 2008.
- [BHMW10] A. Brandstädt, C. Hundt, F. Mancini, and P. Wagner. Rooted directed path graphs are leaf powers. *Discrete Mathematics*, 310(4):897–910, 2010.
- [BL06] A. Brandstädt and V. Le. Structure and linear time recognition of 3-leaf powers. *Information Processing Letters*, 98(4):133–138, 2006.
- [BLR09] A. Brandstädt, V. Le, and D. Rautenbach. A forbidden induced subgraph characterization of distance-hereditary 5-leaf powers. *Discrete Mathematics*, 309(12):3843–3852, 2009.
- [BLR10] Andreas Brandstädt, V. Le, and D. Rautenbach. Exact leaf powers. *Theoretical Computer Science*, 411(31-33):2968–2977, 2010.
- [BLS08] A. Brandstädt, V. Le, and R. Sritharan. Structure and linear-time recognition of 4-leaf powers. *ACM Transactions on Algorithms (TALG)*, 5(1):11, 2008.
- [BM08] J. A. Bondy and U. S. R. Murty. *Graph theory*. Grad. Texts in Math., 2008.
- [BP93] J. Blair and B. Peyton. An introduction to chordal graphs and clique trees. In *Graph theory and sparse matrix computation*, pages 1–29. Springer, 1993.
- [BPP10] S. Bessy, C. Paul, and A. Perez. Polynomial kernels for 3-leaf power graph modification problems. *Discrete Applied Mathematics*, 158(16):1732–1744, 2010.

- [BW10] A. Brandstädt and P. Wagner. Characterising  $(k, \ell)$ -leaf powers. *Discrete Applied Mathematics*, 158(2):110–122, 2010.
- [CFM11] L. Chandran, M. Francis, and R. Mathew. Boxicity of leaf powers. *Graphs and Combinatorics*, 27(1):61–72, 2011.
- [CK07] M. Chang and M. Ko. The 3-steiner root problem. In *WG*, pages 109–120. Springer, 2007.
- [DGHN06] M. Dom, J. Guo, F. Huffner, and R. Niedermeier. Error compensation in leaf power problems. *Algorithmica*, 44(4):363–381, 2006.
- [DGHN08] M. Dom, J. Guo, F. Hüffner, and R. Niedermeier. Closest 4-leaf power is fixed-parameter tractable. *Discrete Applied Mathematics*, 156(18):3345–3361, 2008.
- [DGN05] M. Dom, J. Guo, and R. Niedermeier. Bounded degree closest k-tree power is np-complete. In *International Computing and Combinatorics Conference*, pages 757–766. Springer, 2005.
- [DPS18] R. Duan, S. Pettie, and H. Su. Scaling algorithms for weighted matching in general graphs. *ACM Transactions on Algorithms (TALG)*, 14(1):8, 2018.
- [EH18] D. Eppstein and E. Havvaei. Parameterized leaf power recognition via embedding into graph products. In *IPEC*. Springer, 2018. to appear.
- [FMR<sup>+</sup>08] M. Fellows, D. Meister, F. Rosamond, R. Sritharan, and J. Telle. Leaf powers and their properties: Using the trees. In *ISAAC*, pages 402–413. Springer, 2008.
- [GW07] F. Gurski and E. Wanke. The clique-width of tree-power and leaf-power graphs. In *WG*, pages 76–85. Springer, 2007.
- [HT10] M. Habib and T. To. Structure and recognition of 3, 4-leaf powers of galled phylogenetic networks in polynomial time. Technical Report arXiv:1012.4084, ArXiv, 2010.
- [JKL00] T. Jiang, P. Kearney, and G. Lin. Phylogenetic k-root and steiner k-root. In *ISAAC*, pages 539–551, 2000.
- [KKLY10] W. Kennedy, H. Kong, G. Lin, and G. Yan. Linear time construction of 5-phylogenetic roots for tree chordal graphs. *Journal of combinatorial optimization*, 19(1):94–106, 2010.
- [KLY06] W. Kennedy, G. Lin, and G. Yan. Strictly chordal graphs are leaf powers. *Journal of Discrete Algorithms*, 4(4):511–525, 2006.
- [Laf17] M. Lafond. On strongly chordal graphs that are not leaf powers. In *WG*, pages 386–398. Springer, 2017.
- [McK11] T. McKee. Minimal weak separators of chordal graphs. *Ars Combinatoria*, 101:321–331, 2011.

- [NR15] R. Nevries and C. Rosenke. Characterizing and computing the structure of clique intersections in strongly chordal graphs. *Discrete Applied Mathematics*, 181:221–234, 2015.
- [NR16] R. Nevries and C. Rosenke. Towards a characterization of leaf powers by clique arrangements. *Graphs and Combinatorics*, 32(5):2053–2077, 2016.
- [NRT02] N. Nishimura, P. Ragde, and D. Thilikos. On graph powers for leaf-labeled trees. *Journal of Algorithms*, 42(1):69–108, 2002.
- [PT87] R. Paige and R. Tarjan. Three partition refinement algorithms. *SIAM Journal on Computing*, 16(6):973–989, 1987.
- [Rau06] D. Rautenbach. Some remarks about leaf roots. *Discrete Mathematics*, 306(13):1456–1461, 2006.
- [Rég94] J. Régis. A filtering algorithm for constraints of difference in CSPs. In *AAAI*, volume 94, pages 362–367, 1994.
- [RTL76] D. Rose, R. Tarjan, and G. Lueker. Algorithmic aspects of vertex elimination on graphs. *SIAM Journal on computing*, 5(2):266–283, 1976.
- [Ste92] M. Steel. The complexity of reconstructing trees from qualitative characters and subtrees. *Journal of classification*, 9(1):91–116, 1992.
- [WB09] P. Wagner and A. Brandstädt. The complete inclusion structure of leaf power classes. *Theoretical Computer Science*, 410(52):5505–5514, 2009.