



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-01887266v1

HAL Id: hal-01887266

<https://hal.science/hal-01887266v1>

Submitted on 3 Oct 2018 (v1), last revised 24 Dec 2018 (v2)

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^{1,2}

¹National Institute for Research and Development in Informatics, Romania

²The Research Institute of the University of Bucharest ICUB, Romania

Abstract

The k^{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^{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^{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, there are several of these formulations that 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⁺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 bounded degeneracy [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 1. *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 also 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 2. *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 1 follows from the combination of Theorem 2 with the aforementioned reduction from k -LEAF POWER to $(k - 2)$ -STEINER POWER [BLS08]. We think that our general approach (presented next) could be generalized to larger values of k , although this would first require stronger structure theorems than the ones we use in this paper.

Overview of the techniques. In order to prove our results we use various properties of 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]. We refer to Sec. 2 for any undefined graph-theoretic terminology in this introduction. Our starting point 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 of the tree T we aim at computing. The latter extends to any k the structural results that were presented in [CK07] for $k \leq 3$. – We note by passing that clique-intersections have already been used for (partly) characterizing k -leaf powers under the names of *clique arrangement* [NR16] or “clique graph” [NRT02]. – Furthermore, only if $k \leq 4$, we prove that there always exists a “well-structured” k -Steiner root with additional properties that are also based on clique-intersections. We exploit these nice structural results in the design of a dynamic programming algorithm on a *clique-tree* [BP93] in order to prove Theorem 2.

We want to stress that although the general construction of our algorithm is quite standard, and unsurprisingly close to what has been done in previous works for the recognition of k -Steiner powers [CK07, JKL00], sometimes the devil is in the detail. Specifically, there are several difficulties arising in order to keep the running time polynomial as the value of k increases. Our proposed solutions for $k = 4$ are already quite intricate and they result in an embarrassingly long proof, that may come as a surprise.

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). We sketch in what follows the two main obstacles we met in the design of a “naive” dynamic programming algorithm for our problem:

- If G is a k -Steiner power then, so must be the induced subgraph $G[C \cup S]$, and we aim at storing the k -Steiner roots of $G[C \cup S]$ for further use. There may be exponentially many such solutions already when $G[C \cup S]$ is a clique and $k \geq 3$. Therefore, we cannot afford to store all possible solutions explicitly. However it seems at the minimum we need to keep the part of the Steiner root that 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 of the Steiner root is a subtree of diameter at most $k - 1$, and so, there may be exponentially many possibilities to store whenever $k \geq 4$.
- An additional difficulty is that, in any k -Steiner root of G , one needs to ensure that all vertices in C stay at a distance $\geq k + 1$ from all vertices in $V \setminus (S \cup C)$. So, we also need a way to retrieve, for any partial solution found for $G[C \cup S]$, the distances between vertices in C and those in S . Storing this information would further increase the number of partial roots we need to keep. 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 .

In order to derive a polynomial-time algorithm, 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 Steiner roots and clique-intersections in the graph. Furthermore, in order to bound the number of partial solutions we need to store by a polynomial we combine these stronger properties on the 4-Steiner roots with several tricks (*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).

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 high-level overview of our algorithm, as a guideline for all the other sections.

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 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 categories: “free” and “constrained”, that depend on the other clique-intersections these vertices are contained into. Our main finding is that “free” vertices are mostly responsible for the combinatorial explosion of partial solutions we should store in a naive dynamic programming algorithm for the 4-STEINER ROOT problem. We prove that there always exists a 4-Steiner root where such free vertices are leaves with very special properties, that essentially rules out one of the main difficulties we met in the design of our algorithm.

Sections 5, 6, 7 and 8 are devoted to the main steps of the algorithm. We start presenting a constructive proof of a clique-tree with quite constrained properties in Section 5. Roughly after some preprocessing, we root our clique-tree in such a way that smaller minimal separators should label the edges closer to the root. Our construction ensures that several complications that could occur by using our approach with an arbitrary clique-tree will never occur. Our technical construction is partly motivated by the results in Section 3.

It is well known that the nodes and edges of a clique-tree somewhat represent the maximal cliques and the minimal separators of the graph. In Section 6 we continue using the results in Sections 3 and 4 in order to precompute, for every node and edge in the clique-tree 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). In Section 6.3 we construct in polynomial-time a family of potential subtrees $T\langle X_i \rangle$ for all the other maximal cliques X_i – based on a careful analysis of clique-intersections in X_i and several additional tricks.

In Section 7 we introduce 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. We then explain how we can use this new problem in order to bound the number of partial solutions that we will need to store for our dynamic programming. Finally, we detail in Section 8 the resolution of our intermediate problem, 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 $E_k := \{uv \mid 0 < dist_G(u, v) \leq k\}$. Furthermore, G is a k -Steiner power if there is some tree T such that G is an

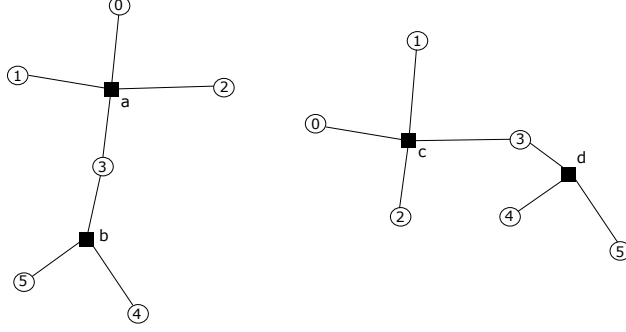


Figure 1: Two Steiner-equivalent trees. Cycles and rectangles represent real and Steiner nodes, respectively.

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 ROOT).

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

Output: A k -Steiner root of G (if any).

Problem 2 (k -LEAF POWER).

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

Question: Is G a k -leaf power?

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

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 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 in particular 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 4 ([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 5 ([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.*

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

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

Based on the above theorem, we can define $E_S(T_G) := \{XY \in E(T_G) \mid X \cap Y = S\}$. 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 X_0 as its root. Let $(X_q, X_{q-1}, \dots, X_1, X_0)$ be a postordering of T_G (obtained by depth-first search). For any $i > 0$, we define $X_{p(i)}$ as the father node of X_i . The common intersection of X_i with its father node is the minimal separator $S_i := X_i \cap X_{p(i)}$. 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 X_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 $G_0 = G$. Furthermore, we 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 some family 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

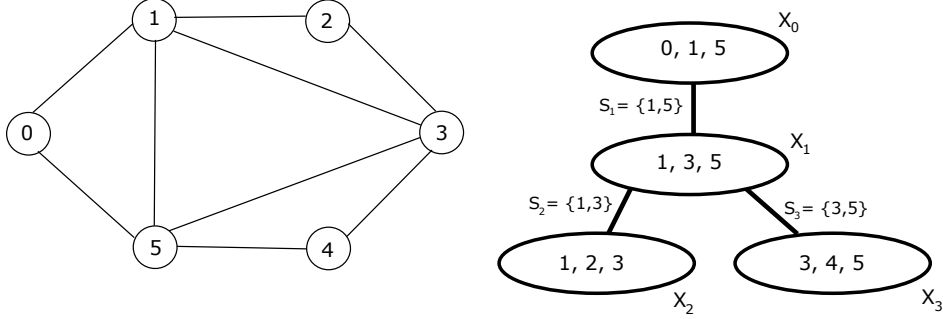


Figure 2: A chordal graph G (left) and a rooted clique-tree T_G (right).

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 Y if and only if we have $X \subseteq Y$.

Theorem 7 ([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 2. By Theorem 3, this will also imply Theorem 1. 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.

- **Initialization Step.** Given $G = (V, E)$, we check whether G is strongly chordal. If this is not the case then, by Theorem 4 G cannot be a k -Steiner power for any $k \geq 1$, and we stop. Otherwise by Theorem 7 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).
- **Step 1.** We construct a clique-tree T_G of G that we root in some $X_0 \in \mathcal{K}(G)$. This clique tree must satisfy very specific properties of which we postpone the precise statement in Section 5. Roughly, we want to ensure that a minimal separator S can occur as an edge $X_i X_{p(i)} \in E(T_G)$, between a maximal clique X_i and its father node, if and only if there is no minimal separator contained into S that appears as an edge in the subtree rooted at X_i . However, we cannot do that exactly due to some recursive complications in our algorithm. The technical motivations behind this additional structure will be further explained in Sections 7 and 8.

- **Step 2.** For every minimal separator S we compute a polynomial-size family \mathcal{T}_S of subtrees whose real nodes are exactly S (Section 6.1). The collection $(\mathcal{T}_S)_{S \in \mathcal{S}(G)}$ is constructed in such a way that assuming G has a 4-Steiner root, there must be one such a root T such that $T \langle S \rangle \in \mathcal{T}_S$ for every minimal separator S . We then proceed similarly for the maximal cliques X_i that are either leaf-nodes (Section 6.2) or internal nodes of T_G (Section 6.3). Correctness of this part follows from our structure theorem of Section 4.
- **Step 3.** Let $(X_q, X_{q-1}, \dots, X_0)$ be a post-ordering of the maximal cliques (*i.e.*, obtained by depth-first-search traversal of T_G). Here again this post-ordering is not arbitrary. Specifically, if $X_i \in \mathcal{K}(G)$ has children nodes $X_{i_1}, X_{i_2}, \dots, X_{i_p}$, where $i < i_1 < i_2 < \dots < i_p$ then, we impose that $S_{i_1}, S_{i_2}, \dots, S_{i_p}$ are ordered by decreasing size. We consider the maximal cliques $X_i \in \mathcal{K}(G)$ sequentially, from $i = q$ down to $i = 0$. If X_i is internal then, let $X_{i_1}, X_{i_2}, \dots, X_{i_p}$ be its children nodes. For every $1 \leq j \leq p$ we have if G is a 4-Steiner power then (by heredity), so is the subgraph $G_{i_j} = (V_{i_j}, E_{i_j})$ that is induced by all the maximal cliques in the subtree $T_G^{i_j}$ rooted at X_{i_j} . Our objective in the next Step will be to compute a set \mathcal{T}_{i_j} of 4-Steiner roots for G_{i_j} . As a way to avoid a combinatorial explosion of the number of partial solutions we will need to store, we sketch in Section 7 how to define – using $(\mathcal{T}_S)_{S \in \mathcal{S}(G)}$ – a polynomial-size subset of “encodings” for these solutions. By combining some local optimization rules with properties of our clique-tree T_G , we show that at most one solution per possibility for the encoding needs to be stored. In particular, we will explain how our above restrictions on the post-ordering can help us to derive additional distances’ constraints from the siblings of a node before we can process it.
- **Step 4.** We end up considering one more time the maximal cliques $X_i \in \mathcal{K}(G)$ sequentially, from $i = q$ down to $i = 0$. After Step 3 is completed, X_i received from its parent node a polynomial-size subset of constraints 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.
 - **Case X_i is a leaf-node.** After Step 2 is completed, we are given a family of all possible subtrees $T \langle X_i \rangle$. We are left verifying whether there exists a solution in this family which satisfies all of the constraints.
 - **Case X_i is an internal node.** Let $X_{i_1}, X_{i_2}, \dots, X_{i_p}$ be the children nodes of X_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 $T \langle X_i \rangle$ (computed during Step 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 stress the intriguing relationship between our approach and the implementation of the `alldifferent` constraint in constraint programming [Rég94].
- 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 Steiner roots and clique-intersections are proved in Section 3.2. 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 1 (folklore). *The following hold for any tree T :*

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

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

Lemma 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. We start the proof with the 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 vertices are exactly the center nodes (Lemma 1), we obtain as claimed that $\mathcal{C}(T_1 \cap T_2) = \mathcal{C}(T_1)$.

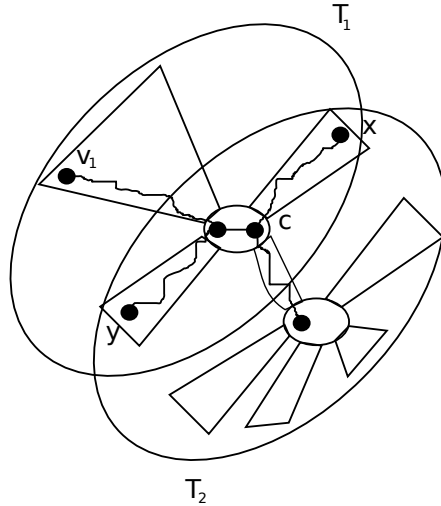


Figure 3: To the proof of Lemma 2.

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 $\text{dist}_T(s, \mathcal{C}(T_2))$ 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 vertex of $V(T_1)$ 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 (Lemma 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 the 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 $\mathcal{C}(T_1)$ (Lemma 1), 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 intersects $\mathcal{C}(T_2) \setminus \mathcal{C}(T_1)$. Then, $dist_T(z, \mathcal{C}(T_2)) = dist_T(z, \mathcal{C}(T_1))$. Furthermore by unimodality (Lemma 1) every node $v_1 \in V(T_1)$ has eccentricity $dist_T(v_1, \mathcal{C}(T_1)) + rad(T_1)$. Since z is an end in a diametral path of T_1 it maximizes $dist_T(z, \mathcal{C}(T_1))$, and so, for every $v_1 \in V(T_1)$ we have $dist_T(v_1, \mathcal{C}(T_2)) \leq dist_T(v_1, \mathcal{C}(T_1)) \leq dist_T(z, \mathcal{C}(T_1)) = dist_T(z, \mathcal{C}(T_2))$.
- Otherwise, let $c \in \mathcal{C}(T_1)$ minimize $dist_T(c, \mathcal{C}(T_2))$. 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 $dist_T(v_1, \mathcal{C}(T_2)) \leq dist_T(v_1, c) + dist_T(c, \mathcal{C}(T_2)) \leq rad(T_1) + dist_T(c, \mathcal{C}(T_2))$, 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 intersects $\mathcal{C}(T_2) \setminus \mathcal{C}(T_1)$; 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, $dist_T(z, c) = rad(T_1)$ 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, $dist_T(v_1, \mathcal{C}(T_2)) \leq dist_T(z, \mathcal{C}(T_2))$ for every $v_1 \in V(T_1)$. \square

Lemma 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 $diam(T_1 \cup T_2) = \max\{diam(T_1), diam(T_2)\}$.*

Proof. Since $\mathcal{C}(T_1) \subseteq \mathcal{C}(T_2)$ we have:

$$ecc_{T_1 \cup T_2}(v_1) \leq dist_T(v_1, \mathcal{C}(T_1)) + \max\{rad(T_1), rad(T_2)\}$$

for every $v_1 \in V(T_1)$. By the unimodality property (Lemma 1):

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

and also:

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

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

In the same way, 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. In particular since we also have $\mathcal{C}(T_1) \subseteq \mathcal{C}(T_2)$, $diam(T_1) \geq diam(T_2)$. There are two cases to consider:

- Case $diam(T_1) = diam(T_2)$. Then, $\mathcal{C}(T_1) = \mathcal{C}(T_2)$ and we can strengthen the above 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 we have $\mathcal{C}(T_1) \subseteq \mathcal{C}(T_2)$, $diam(T_1)$ is 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 8. *Given $G = (V, E)$ and T any k -Steiner root of G , the following properties hold for any clique-intersection $X \in \mathcal{X}(G)$:*

1. *We have $Real(T\langle X \rangle) = X$ and $diam(T\langle X \rangle) \leq k$;*
2. *There is no supertree $T'_X \supset T\langle X \rangle$ with $X \subset Real(T'_X)$ and $diam(T'_X) = diam(T\langle X \rangle)$;*
3. *If $X \subset X' \in \mathcal{X}(G)$ then, $diam(T\langle X \rangle) < diam(T\langle X' \rangle)$.*
4. *If $\mathcal{C}(T\langle X \rangle) \subseteq \mathcal{C}(T\langle X' \rangle)$ then, $X \cup X'$ is a clique of G .*

Proof. First assume $X \in \mathcal{K}(G)$ to be a maximal clique. Since all leaves of $T\langle X \rangle$ are in X , $diam(T\langle X \rangle) = \max_{u,v \in X} dist_T(u, v)$. By the hypothesis T is a k -Steiner root of G , and so, since X is a clique of G , $\max_{u,v \in X} dist_T(u, v) \leq k$. In particular, $diam(T\langle X \rangle) \leq k$, that implies in turn the vertices of $Real(T\langle X \rangle)$ must induce a clique of G . We can conclude that $Real(T\langle X \rangle) = X$ by maximality of X . More generally, let $X = \bigcap_{i=1}^{\ell} X_i$, for some family $X_1, X_2, \dots, X_{\ell} \in \mathcal{K}(G)$. Clearly, $T\langle X \rangle \subseteq \bigcap_{i=1}^{\ell} T\langle X_i \rangle$, and so, $X \subseteq Real(T\langle X \rangle) \subseteq \bigcap_{i=1}^{\ell} Real(T\langle X_i \rangle)$. As we proved before, $Real(T\langle X_i \rangle) = X_i$ for every $1 \leq i \leq \ell$, and so, $Real(T\langle X \rangle) \subseteq \bigcap_{i=1}^{\ell} X_i = X$. Altogether combined, we obtain that $Real(T\langle X \rangle) = X$.

Second, let $T'_X \supset T\langle X \rangle$ be such that $diam(T'_X) = diam(T\langle X \rangle)$. We claim $Real(T'_X) = X$, that will prove the second part of the theorem. Indeed, for any maximal clique X_j that contains X , we have $diam(T\langle X_j \rangle \cap T'_X) \geq diam(T\langle X \rangle) = diam(T'_X)$, and so, $diam(T\langle X_j \rangle \cup T'_X) \leq diam(T\langle X_j \rangle) \leq k$ by Lemma 2. It implies $Real(T'_X) \subseteq X_j$. Furthermore, since $X \in \mathcal{X}(G)$, it is exactly the intersection of all the maximal cliques that contains it, thereby proving the claim. In particular (Property 3), assume now $X \subset X'$. Since $T\langle X' \rangle \supset T\langle X \rangle$, we cannot have $diam(T\langle X' \rangle) = diam(T\langle X \rangle)$ (otherwise, $X' = X$ by Property 2). Therefore, $diam(T\langle X' \rangle) > diam(T\langle X \rangle)$.

Finally, assume $\mathcal{C}(T\langle X \rangle) \subseteq \mathcal{C}(T\langle X' \rangle)$. By Lemma 3 we obtain that $diam(T\langle X \rangle \cup T\langle X' \rangle) = \max\{diam(T\langle X \rangle), diam(T\langle X' \rangle)\} \leq k$. In particular, $X \cup X'$ is a clique of G . \square

Before ending this section, we slightly strenghten Property 4 of Theorem 8, as follows:

Lemma 4. *Given $G = (V, E)$ and T any $2k$ -Steiner root of G , we have $\mathcal{C}(T\langle X_i \rangle) \cap \mathcal{C}(T\langle X_j \rangle) = \emptyset$ for any two different maximal cliques $X_i, X_j \in \mathcal{K}(G)$.*

Proof. Suppose for the sake of contradiction $\mathcal{C}(T\langle X_i \rangle) \cap \mathcal{C}(T\langle X_j \rangle) \neq \emptyset$, and let $v \in \mathcal{C}(T\langle X_i \rangle) \cap \mathcal{C}(T\langle X_j \rangle)$. By Theorem 8, $\max\{diam(T\langle X_i \rangle), diam(T\langle X_j \rangle)\} \leq 2k$, and so, any vertex of $T\langle X_i \rangle \cup T\langle X_j \rangle$ is at a distance $\leq k$ from v in T . In particular, $diam(T\langle X_i \rangle \cup T\langle X_j \rangle) \leq 2k$, and so, $X_i \cup X_j$ is a clique of G . The latter contradicts the fact that X_i, X_j are maximal cliques. \square

4 Well-structured 4-Steiner roots

We refine our results in the previous Section when $k = 4$. Given $G = (V, E)$ and a rooted clique-tree T_G of G , let $X_i \in \mathcal{K}(G)$ be arbitrary. We recall that our goal will be eventually to construct, in polynomial time, a set \mathcal{T}_i of 4-Steiner roots for some subgraph G_i that can always be extended to a 4-Steiner root for G if one exists. 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 for the remaining of the paper:

Definition 2. Given $G = (V, E)$ and $X \in \mathcal{X}(G)$, a vertex $v \in X$ is called *X-free* if for any other $X' \in \mathcal{X}(G)$ we have either $v \notin X'$, $X \subseteq X'$ or $X \cap X' = \{v\}$. A vertex $v \in X$ that is not *X-free* is called *X-constrained*.

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 $T\langle X \rangle$. However, we prove that we can always force them to be *leaves* of this subtree. Specifically:

Theorem 9. *Let $G = (V, E)$ be a 4-Steiner power. There always exists a 4-Steiner root T of G where, for any clique-intersection $X \in \mathcal{X}(G)$, all the *X-free* vertices are leaves of $T\langle X \rangle$ with maximum eccentricity $\text{diam}(T\langle X \rangle)$. Moreover:*

1. *all the *X-free* vertices, except maybe one, are closest to the same central node in $\mathcal{C}(T\langle X \rangle)$;*
2. *all the internal nodes on a path between $\mathcal{C}(T\langle X \rangle)$ and a *X-free* vertex are Steiner;*
3. *and if $X \in \mathcal{K}(G)$ and it has a *X-free* vertex then, $\text{diam}(T\langle X \rangle) = 4$.*

Theorem 9 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. 5 and 6. It is crucial for the proof that in any 4-Steiner root of G all 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*.

We first prove Theorem 9 for maximal cliques (Section 4.1). Then, we prove the result in its full generality in Section 4.2.

4.1 The case of (Almost) Simplicial vertices

Let $X_i \in \mathcal{K}(G)$ be fixed. We start giving a simple characterization of X_i -free vertices in terms of simplicial vertices and cut-vertices. Then, we prove Theorem 9 in the special case when X is a maximal clique.

Lemma 5. *Given $G = (V, E)$ and $X_i \in \mathcal{K}(G)$, a vertex $v \in X_i$ is X_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 X_i that can also contain v .*

Proof. A vertex $v \in X_i$ is not contained into any other maximal clique if and only if it is simplicial (and in this case, this vertex is clearly X_i -free). From now on assume v is not simplicial. If $v \in X_i \cap X_j$ then, in any clique-tree T_G of G , the vertex v and more generally, all of $X_i \cap X_j$, is contained into all the minimal separators that label an edge of the $X_i X_j$ -path in T_G . This implies that there is always a largest clique-intersection $X \subset X_i$ containing v that is a minimal separator. Hence a non simplicial $v \in X_i$ is X_i -free if and only if it is a cut-vertex, and there is no other minimal separator in X_i that contains this vertex. \square

Lemma 6. *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 X_i with at least one X_i -free vertex:*

- $diam(T\langle X_i \rangle) = 4$;
- every X_i -free vertex v is a leaf of $T\langle X_i \rangle$ such that $dist_T(v, \mathcal{C}(T\langle X_i \rangle)) = 2$, and the internal node onto the unique $v\mathcal{C}(T\langle X_i \rangle)$ -path is Steiner.

Proof. We give an illustration of the proof in Fig. 4. 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 X_i -free for some $X_i \in \mathcal{K}(G)$. We now proceed by induction on $|S_1|$.

Assume $S_1 = \emptyset$ for the base case. While there exist $X_i \in \mathcal{K}(G)$ and $v \in X_i$ simplicial that falsify the properties of the lemma, we proceed as follows. Let $c_i \in \mathcal{C}(T\langle X_i \rangle)$ minimize $dist_T(v, c_i)$ (possibly, $v = c_i$). We first replace v by a Steiner node α . 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 α (if $c_i = v$). We connect v to c'_i via a path of length exactly $4 - \max_{u \in X_i \setminus \{v\}} dist_{T'}(c'_i, u)$ of which all internal nodes are Steiner. In doing so, we obtain a tree T'' such that $Real(T'') = V$. By construction, $\max_{u \in X_i \setminus \{v\}} dist_{T'}(c'_i, u) \leq ecc_{T\langle X_i \rangle}(c_i) \leq 2$ (since $diam(T\langle X_i \rangle) \leq 4$), hence:

$$dist_T(v, c_i) \leq ecc_{T\langle X_i \rangle}(c_i) \leq 4 - \max_{u \in X_i \setminus \{v\}} dist_{T'}(c'_i, u) = dist_{T''}(v, c'_i).$$

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, $diam(T''\langle X_i \rangle) = 4$ and the unique central node in $\mathcal{C}(T''\langle X_i \rangle)$ is onto the $v c'_i$ -path by construction. Here it is also important to observe that, since v is only contained into X_i and our transformation can only increase the distances between the real nodes, X_i 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 nodes v are contained into some clique X_i such that: $diam(T\langle X_i \rangle) = 4$, v is a leaf of $T\langle X_i \rangle$ such that $dist_T(v, \mathcal{C}(T\langle X_i \rangle)) = 2$, and the internal node onto the $v\mathcal{C}(T\langle X_i \rangle)$ -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_ℓ 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 R^i that satisfy the properties of the lemma w.r.t. G_i . Overall, by identifying all the R^i 's at v , one obtains a tree R . We claim that R satisfies the two properties of the lemma. Indeed, it follows from the characterization of Lemma 5 that for any $X_j \subseteq C_i \cup \{v\}$, the X_j -free vertices in G are still X_j -free vertices in G_i . – Note that in particular, if v is X_j -free in G then, v is simplicial in G_i . – Therefore, the claim is proved. It remains now to

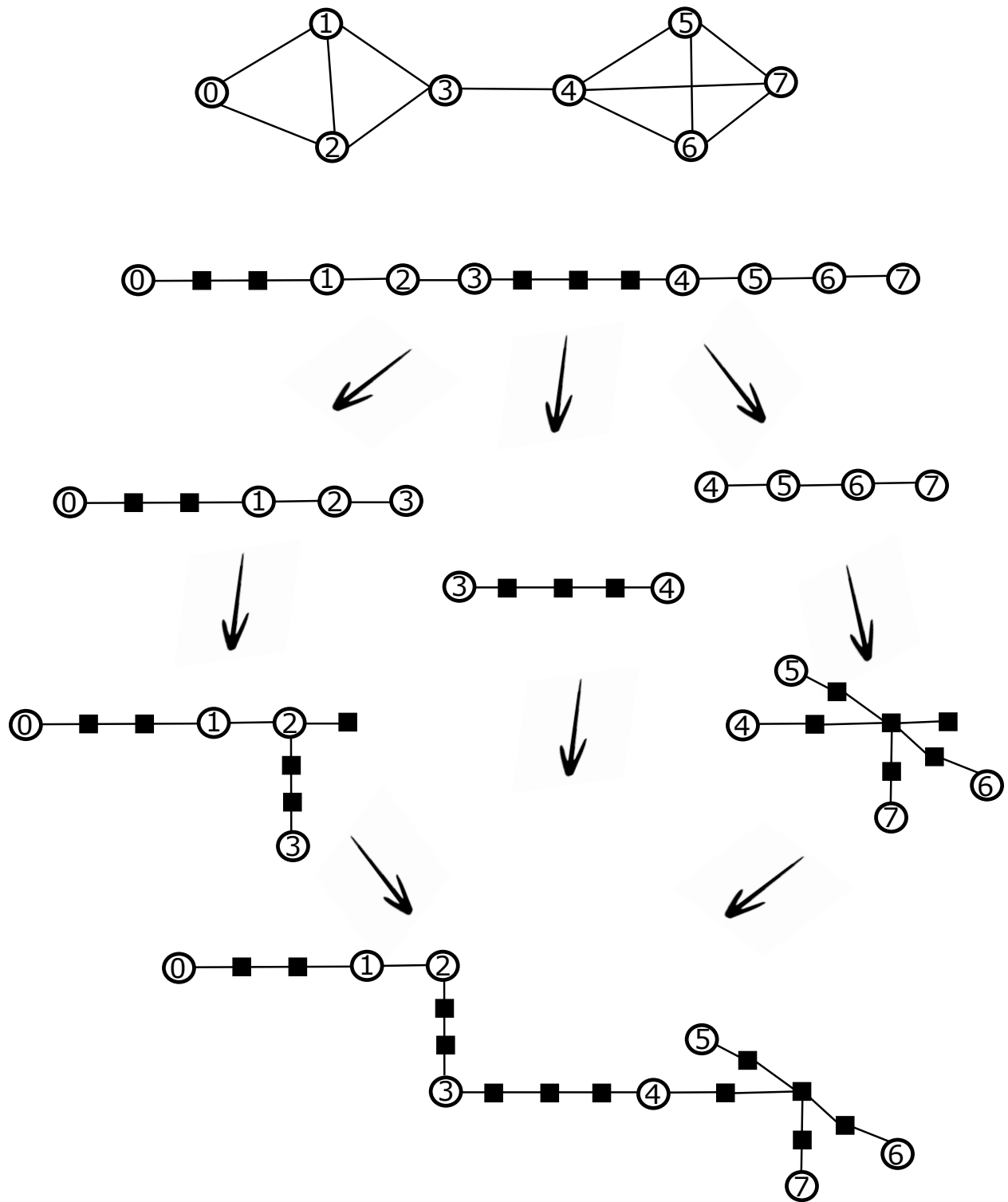


Figure 4: The transformation of Lemma 6 applied to an arbitrary 4-Steiner root.

show that R 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 $dist_R(x, v) + dist_R(v, y) \leq 4$. Our construction implies $dist_R(x, v) \geq dist_T(x, v)$ and $dist_R(y, v) \geq dist_T(y, v)$. But then, we should already have $dist_T(x, y) \leq 4$ in the original root T . Thus, since T is a 4-Steiner root of G , this case cannot happen and R is also a 4-Steiner root of the graph G . \square

4.2 The general case

We are now ready to prove Theorem 9 in its full generality:

Proof of Theorem 9. Let T be such that the result holds for maximal cliques (such a T exists by Lemma 6). 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. 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 $T \setminus T \langle X \rangle$ with one node adjacent to v . Let Q_v be the subtree of T that is induced by $V(R_v) \cup \{v\}$.

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

1. We remove R_v and we replace v by a Steiner node α_v . In doing so, we obtain an intermediate tree denoted by T_v ;
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 $T_v \langle X \setminus v \rangle$ (possibly, $c = \alpha_v$).

We refer to Fig. 5 and 6 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 1. *Assume v is X -free and 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 either $dist_T(Real(R_v), v) \geq 4$ or c is Steiner. Moreover, for any $X' \in \mathcal{X}(G) \setminus \{X\}$ and for any of the properties stated in the theorem, if this property 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 . Indeed, if there exist $x \in V(R_v)$, $y \notin V(Q_v)$ such that $dist_T(x, y) \leq 4$ then, v is onto the unique xy -path in T . In particular, there must be a node $z \in V(T \langle X \rangle) \setminus \{v\}$ such that the xy -path in T goes by the edge vz (*i.e.*, because $y \notin V(Q_v)$). But then, since $diam(T \langle X \rangle) \leq 3$ by Theorem 8 (*i.e.*, because X is strictly contained into some maximal clique), one of v or z must be in $\mathcal{C}(T \langle X \rangle)$. Thus, we can always assume w.l.o.g. y is a real node in $X \setminus \{v\}$ (possibly, $y = z$) for every node in $\mathcal{C}(T \langle X \rangle)$ is adjacent to such a vertex by the hypothesis. This implies the existence of a maximal clique containing x, y, v , and the latter cannot be X since we have $x \in V(R_v)$. However the latter contradicts that v is X -free, and so, it proves 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 .

It follows from the above result that in order for T' to be a 4-Steiner root for G , one must ensure $dist_{T'}(Real(R_v), V \setminus V(Q_v)) > 4$. Note that this is always the case if $dist_T(v, Real(R_v)) \geq 4$. 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 $dist_T(v, V \setminus V(Q_v)) \leq dist_T(v, X \setminus \{v\}) \leq 2$, and so, $dist_T(v, Real(R_v)) = 3$. Then, a necessary and sufficient condition for having $dist_{T'}(Real(R_v), V \setminus V(Q_v)) > 4$ is that c is Steiner. However, the above does not prove that T' is a 4-Steiner root of G yet, as we also need to ensure $dist_{T'}(u, v) \leq 4$ 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 $dist_T(Real(R_v), V \setminus V(Q_v)) > 4$, 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 and so, it must intersect $T \langle X \rangle \setminus \{v\}$. We then consider two different cases:

- Assume $X \subset X'$. By the above observation, $T \langle X' \rangle \cap R_v = \emptyset$. In particular, $T' \langle X' \rangle$ 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. Note that in doing so, any X' -free vertex that was a leaf in $T \langle X' \rangle$ is also a leaf of $T' \langle X' \rangle$. Furthermore, the above transformation cannot add an internal real node onto the path between such a leaf and the center nodes, that implies we cannot break Property 2 of the theorem. We cannot break Property 1 either since v cannot be X' -free, and so, we did not move any X' -free vertex during this operation. Finally, since every center node of $\mathcal{C}(T \langle X \rangle)$ is adjacent to a real node in $X \setminus \{v\}$ by the hypothesis, we have after Operation 1 $diam(T' \langle X' \rangle) = diam(T \langle X' \rangle)$ (*i.e.*, we cannot break Property 3, and we also obtain $dist_{T'}(u, v) \leq 4$ for every $u \in X'$).
- Otherwise, $X \not\subset X'$ and we prove $T' \langle X' \rangle = T \langle X' \rangle$. To see that, first note this may not be the case only if $T \langle X' \rangle$ is not fully contained into Q_v . In this situation, we also know that $T \langle X' \rangle$ must intersect $T \langle X \rangle \setminus \{v\}$. 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 β is Steiner there must exist $y \in X'$ such that the unique vy -path in T goes by β . However, this implies $diam(T \langle X' \cup N_T[\beta] \rangle) = diam(T \langle X' \rangle)$. We recall that there exists at least one leaf node $u \in Real(N_T(\beta)) \setminus \{v\}$ by the hypothesis. Thus, by Property 2 of Theorem 8 we have $u, v \in X \cap X'$, thereby contradicting that v is X -free.

The claim directly follows from this above case analysis. \diamond

The proof is now divided into two main phases.

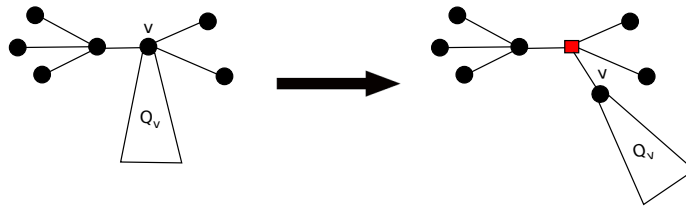


Figure 5: Forcing the X -free vertices as leaves.

Phase 1: Transformation into leaves (see Fig. 5 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 an X -free vertex $v \in X$ that is not a leaf. Note that we have $v \in \mathcal{C}(T \langle X \rangle)$. 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). Overall, by Claim 1 we can repeat the above transformation until all the X -free vertices are leaves of $T\langle X \rangle$.

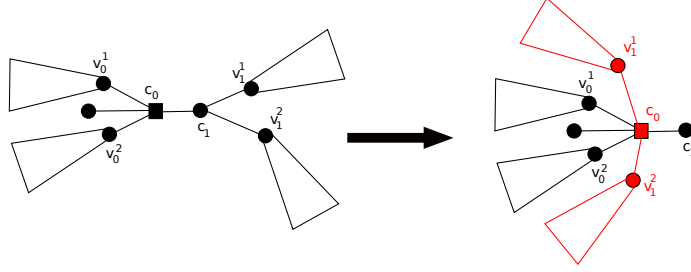


Figure 6: Grouping the X -free vertices on a same side.

Phase 2: Grouping the X -free vertices (see Fig. 6 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. 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, there would exist a maximal clique containing x_i, v_j^1, v_j^2 , and so, the two of v_j^1, v_j^2 would be X -constrained). 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, we can prove as above that v_j^i should be X -constrained). 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 $c = c_0^1$. 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 1. \square

5 Step 1: Construction of the clique-tree

The main result in this section is the construction of a very specific clique-tree (Theorem 11), of which we will use the additional properties in order to ensure that our algorithm runs in polynomial time. We present a first construction in Section 5.1. 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 10. *Given $G = (V, E)$ chordal, we can compute in polynomial time a clique-tree T_G such that, for any $S_i = X_i \cap X_{p(i)}$ and for any child X_j of $X_{p(i)}$, there is no minimal separator of G_j that is contained into S_i .*

¹As suggested by Fig. 6, we could actually apply the transformation to v_1^1 and v_1^2 simultaneously. However, we do not need this refinement for the proof.

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 X_0 . We consider all the minimal separators $S \in \mathcal{S}(G)$ by decreasing size. Let X_S be incident to an edge in $\bigcup_{S \subseteq S'} E_{S'}(T_G)$ and the closest possible to the root. We observe that X_S is the least common ancestor of all maximal cliques that are incident to an edge in $\bigcup_{S \subseteq S'} E_{S'}(T_G)$. All edges in $E_S(T_G)$ can be made incident to X_S , as follows. Assume there exists $YZ \in E_S(T_G)$ such that $X_S \notin \{Y, Z\}$. By the above observation, Y, Z are into the subtree rooted at X_S . W.l.o.g., Z is further than Y from X_S . Since $S \subseteq X_S \cap Z$, S is contained into all the maximal cliques onto the $X_S Z$ -path. In particular, we still obtain a clique-tree of T_G if we replace YZ by $X_S Z$ and in doing so, $S \subseteq X_S \cap Z \subseteq Y \cap Z = S$. Furthermore after this transformation, X_S became the new father node of the maximal clique Z 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 there exists $i > 0$ and $S_k \subseteq S_i$ a minimal separator of G_j where X_j is a child of $X_{p(i)}$ (possibly, $X_i = X_j$). Our transformation ensures $S_k \neq S_i$, i.e., $S_k \subset S_i$. Since the subtree rooted at X_j is a rooted clique-tree of G_j , there must exist some edge $X_k X_{p(k)}$ in this subtree such that $X_k \cap X_{p(k)} = S_k$. However, since we consider minimal separators by increasing size, the edge $X_i X_{p(i)}$ should already exist when we process S_k . It implies the maximal clique X_{S_k} to which we connected all edges in $E_{S_k}(T_G)$ should be an ancestor of $X_{p(i)}$, that is a contradiction. \square

We will see this “flat” clique-tree, given by Theorem 10 (or at least a technical modification of this clique-tree), considerably simplifies the analysis of our algorithm. In particular, we stress the following of its properties (that we prove on purpose under slightly weaker conditions than the ones given by Theorem 10):

Lemma 7. *Given a rooted clique-tree T_G of $G = (V, E)$, let $X_i \in \mathcal{K}(G)$ have children $X_{i_1}, X_{i_2}, \dots, X_{i_p}$, and assume that the following properties are true for any $j \in \{1, 2, \dots, p\}$:*

- *For any $j' \in \{1, 2, \dots, p\}$ there can be no minimal separator of $G_{i_{j'}}$ contained into S_{i_j} ;*
- *Moreover $S_{i_j} \not\subseteq S_i$.*

Then, all the edges in $E_{S_{i_j}}(T_G)$ are incident to X_i .

Proof. Since $S_{i_j} \not\subseteq S_i$ by the hypothesis, all edges in $E_{S_{i_j}}(T_G)$ must be in the subtree T_G^i rooted at X_i . But then, since for every j' we have that $T_G^{i_{j'}}$ is a rooted clique-tree of $G_{i_{j'}}$, there can be no edge of $E_{S_{i_j}}(T_G)$ in this subtree. This proves that all edges in $E_{S_{i_j}}(T_G)$ must be incident to X_i . \square

Corollary 1. *Let T_G be a rooted clique-tree of $G = (V, E)$ as stated in Theorem 10. For any minimal separator $S \in \mathcal{S}(G)$, the edge-set $E_S(T_G)$ is incident to a common maximal clique X_S .*

Proof. By Theorem 6, there exist $X_{i_j} X_i \in E(T_G)$ such that $X_{i_j} \cap X_i = S$ and $X_{p(i_j)} = X_i$. Hence, by Lemma 7 we can choose $X_S := X_i$. \square

5.2 Weak convergence

Unfortunately, this “flat” clique-tree that we get is not exactly what we need yet, due to some encoding issues. Specifically, consider the particular case in our dynamic programming algorithm where a minimal separator S_i induces a star in some partial solutions we found for the subgraph

G_i . In order to bound the number of partial solutions that we should store where $T\langle S_i \rangle$ is a star, we would like most vertices in S_i to be simplicial in G_i . However, this cannot be the case if we encountered a larger minimal separator $S' \supset S_i$ in the corresponding clique-subtree. See Section 7 for more details.

We now introduce the following notions:

Definition 3. 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 X_S that is incident to all edges in $\bigcup_{S \subset S'} E_{S'}(T_G)$. S is termed *T_G -convergent* if the maximal clique X_S is also incident to all edges in $E_S(T_G)$.

In order to motivate Definition 3, in what follows are two observations on the relationships between clique-trees, minimal separators and 4-Steiner roots:

Lemma 8. *Given $G = (V, E)$ and T any 4-Steiner root of G , let $X_i \in \mathcal{K}(G)$ and let $S \subset X_i$ be a clique-intersection. If T_S is a bistar then, $\mathcal{C}(T\langle X_i \rangle) \subset \mathcal{C}(T\langle S \rangle)$.*

In particular, there are exactly two maximal cliques that contain S .

Proof. We have by Theorem 8 $\text{diam}(T\langle X_i \rangle) > \text{diam}(T\langle S \rangle)$, and so, $\text{diam}(T\langle X_i \rangle) = 4$. In particular, write $\mathcal{C}(T\langle X_i \rangle) = \{c_i\}$. Every component in $T\langle X_i \rangle \setminus \{c_i\}$ has diameter at most two, thereby implying $c_i \in V(T\langle S \rangle)$. Furthermore since $\text{ecc}_{T\langle S \rangle}(c_i) \leq \text{rad}(T\langle X_i \rangle) = 2$, c_i cannot be a leaf of $T\langle S \rangle$, i.e., $c_i \in \mathcal{C}(T\langle S \rangle)$. By Lemma 4, there can be no two maximal cliques $X_i, X_j \in \mathcal{K}(G)$ such that $\mathcal{C}(T\langle X_i \rangle) = \mathcal{C}(T\langle X_j \rangle)$. Therefore, the above implies that S can only be contained in at most two maximal cliques. Finally, since S is not a maximal clique, it is contained into exactly two maximal cliques. \square

Lemma 9. *Given $G = (V, E)$ and T any 4-Steiner root of G , let $X \in \mathcal{S}(G)$ be a minimal separator. If T_X is a non-edge star then, X is weakly T_G -convergent for any clique-tree T_G of G .*

Proof. We may assume that X is strictly contained into at least one minimal separator S for otherwise there is nothing to prove. By Theorem 8, $T\langle S \rangle$ is a bistar and S must be inclusion wise maximal in $\mathcal{S}(G)$. This implies $\mathcal{C}(T\langle X \rangle) \subset \mathcal{C}(T\langle S \rangle)$. Furthermore, it follows from Lemma 8 that S must be contained into exactly two maximal cliques X_i, X_j and $\mathcal{C}(T\langle X_i \rangle) \cup \mathcal{C}(T\langle X_j \rangle) = \mathcal{C}(T\langle S \rangle)$. In particular, we may assume w.l.o.g. that $\mathcal{C}(T\langle X \rangle) = \mathcal{C}(T\langle X_i \rangle)$. But then, still by Lemma 8, any minimal separator S' that strictly contains X must be contained into X_i and exactly one other maximal clique $Y_{S'}$. By Theorem 6, the latter implies $X_i Y_{S'} \in E(T_G)$ and $X_i \cap Y_{S'} = S'$. \square

Our goal is to force weak convergence to imply convergence. Intuitively, this will make the above issue with stars, as well as several related issues with bistars, *local*, thereby allowing us to handle with them more efficiently. We end this section by analysing the cases where the flat clique-tree of Section 5.1 does not satisfy this property (see Fig. 7 for an illustration of such cases).

Lemma 10. *Given a chordal graph $G = (V, E)$ and $S \in \mathcal{S}(G)$, let T_G be a rooted clique-tree of G and let $X_S, Y_S \in \mathcal{K}(G)$ be such that: all edges in $E_S(T_G)$ are incident to X_S ; all edges in $\bigcup_{S \subset S'} E_{S'}(T_G)$ are incident to Y_S ; and no ancestor of X_S can be incident to an edge of $\bigcup_{S \subset S'} E_{S'}(T_G)$. Exactly one of the following three conditions is true:*

- $X_S = Y_S$;
- X_S is a father node of Y_S and we have $S \subseteq X_S \cap Y_S$;

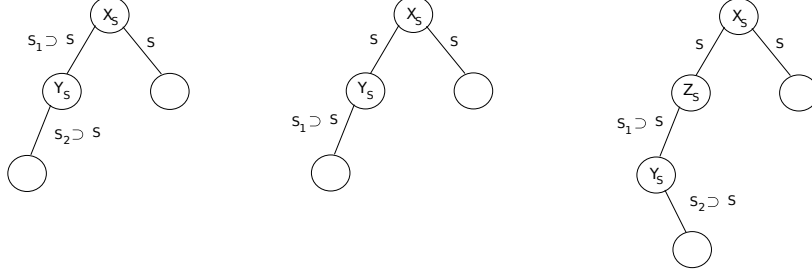


Figure 7: Examples of S weakly T_G -convergent but not T_G -convergent.

- or there exists $Z_S \in \mathcal{K}(G)$ being a child node of X_S and the father node of Y_S , and we have $X_S \cap Z_S = S \subset Z_S \cap Y_S$.

Proof. Assume $X_S \neq Y_S$. We observe that X_S must be an ancestor of Y_S (otherwise, by taking their least common ancestor we would contradict our assumption that no ancestor of X_S can be incident to an edge of $\bigcup_{S \subseteq S'} E_{S'}(T_G)$). Since $S \subseteq X_S \cap Y_S$, all edges onto the $X_S Y_S$ -path must be in $\bigcup_{S \subseteq S'} E_{S'}(T_G)$. Finally, since all edges in $E_S(T_G)$ are incident to X_S , and all edges in $\bigcup_{S \subseteq S'} E_{S'}(T_G)$ are incident to Y_S , there can be at most two edges into the $X_S Y_S$ -path. In the case when there are two edges then, exactly one such edge must be in $E_S(T_G)$ and this edge is incident to the maximal clique X_S . \square

5.3 The final construction

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

Theorem 11. *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 $S_i := X_i \cap X_{p(i)}, i > 0$:*

1. *If S_i is weakly T_G -convergent and $|S_i| \geq 3$ then, S_i is T_G -convergent;*
2. *If S_i contains a minimal separator of G_i then, $|S_i| \geq 3$ and S_i is T_G -convergent;*
3. *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 other minimal separator of G_i .*

Proof. Let T_G be the rooted clique-tree obtained from Theorem 10. We perform a dynamic programming on the internal nodes of T_G , that starts from the twigs (*a.k.a.*, the nodes whose all children nodes are leaves). Specifically, let $X_i \in \mathcal{K}(G)$ be an internal node of T_G , and assume that all the children nodes of X_i that are not leaves were already processed. Our construction ensures that the set of descendants of X_i were not modified (but the connections between these nodes may have changed). In particular, we did not modify the children nodes of X_i neither the set of their respective descendants. Let $\mathcal{S}_i := \{X_j \cap X_i \mid X_{p(j)} = X_i\}$. We consider all the minimal separators $S \in \mathcal{S}_i$ by decreasing size. Note that X_i is the maximal clique X_S : as defined by Corollary 1.

There are several cases:

- We do not modify $E_S(T_G)$ if either $|S| \leq 2$, S is T_G -convergent or S is not weakly T_G -convergent. In this situation, for any child node X_j such that $S_j = S$ we claim that S_j satisfies all three conditions of the theorem. Indeed, in order to prove this, it suffices to

prove that no minimal separator of G_j can be contained into S_j . This was holding initially by Theorem 10, and so it still holds at this step because we did not modify the set of descendants of X_j . We stress that here we need to assume that the minimal separators of \mathcal{S}_i are considered by decreasing size (otherwise, edges in $E_{S'}(T_G)$ could have been made incident to a node of T_G^j at some previous step, for some $S' \subset S$).

- Otherwise, $|S| \geq 3$, S is weakly T_G -convergent but not T_G -convergent. As we did not change the set of descendants of X_i , we can apply Lemma 10 in order to prove there exists a descendant Y_S of X_i – either a child or the child of a child – that is incident to all the edges in $\bigcup_{S' \supset S} E_{S'}(T_G)$. We use the same operation as in the proof of Theorem 10 in order to make all edges in $E_S(T_G)$ incident to Y_S . On the way, X_i becomes the father node of Y_S , and all the maximal cliques $Z \notin \{Y_S, X_i\}$ that are incident to an edge in $E_S(T_G)$ are made children of Y_S . Note that if we are in Case 3 of Lemma 10, this means we changed the orientation of $Y_S Z_S$ (*i.e.*, Z_S used to be the father node of Y_S and it is now a child of Y_S). See Fig. 8 for an illustration.

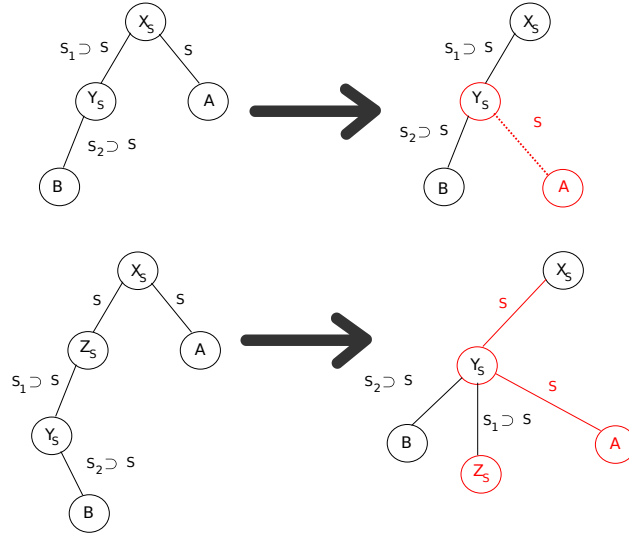


Figure 8: Local modifications of a clique-tree.

In doing so, we say that we processed S , all edges labeled by S resp.

We now need to analyse the impact of this second case above to the properties of other minimal separators. First we notice that since we assume S was not T_G -convergent, there was no minimal separator $S' \subseteq S$ T_G -convergent before we made the transformation. In particular, any minimal separator S' that was T_G -convergent before the transformation has kept this property. Conversely, if after the transformation some minimal separator S' became (not) weakly T_G -convergent then, $S' \subseteq S$. As we did not modify the descendants of X_i , this implies by Theorem 10 either $S' \in \mathcal{S}_i$ or $X_{S'}$ is an ancestor of X_i . In both cases (since we consider the minimal separators of \mathcal{S}_i by decreasing size), S' was not processed yet. Summarizing, Condition 1 of the theorem is preserved for all the minimal separators already processed, and it now also holds for S .

For the remaining two conditions, we can further restrict our study to the subtree containing Y_S . By Lemma 10 there are only two possibilities:

1. Situation # 1: Assume Y_S was a child of X_i before the transformation. Any child X_j of X_i , $X_j \neq Y_S$, such that $X_i \cap X_j = S$ is now a child of Y_S , with $X_j \cap Y_S = S$. The latter did not change the fact that S does not contain a minimal separator of G_j , nor the properties of the edges in the subtree rooted at X_j . Therefore, the properties of all edges in the subtree rooted at X_j are left unchanged, and the edge $X_j Y_S$ now satisfies Conditions 2 and 3 of the theorem (*i.e.*, because $X_j \cap Y_S = S$). We are left with $S' := X_i \cap Y_S$.

If $S' = S$ then, no minimal separator that appears in the subtree rooted at Y_S can be strictly contained into S . However, S now appears in this subtree. Since S was weakly T_G -convergent but not T_G -convergent before the transformation, there is a minimal separator in the subtree rooted at Y_S that strictly contains S . As a result, Conditions 2 and 3 of the theorem are satisfied.

From now on assume $S' \supset S$. Since S is T_G -convergent after the transformation, so is S' . Furthermore, Theorem 10 ensures that, in every $X_j \neq Y_S$ such that $X_j \cap X_i = S$ (X_j is a child of Y_S after the transformation), there can be no minimal separator of G_j contained into S' . Thus the minimal separator S is the only new one contained into S' that can appear in the subtree rooted at Y_S . In particular, there is another minimal separator $S'' \supset S$ in this subtree because S was not T_G -convergent before the transformation. As a result, Conditions 2 and 3 of the theorem keep holding for the edge $X_j Y_S$.

2. Situation # 2: Otherwise, the father node Z_S of Y_S previously was a child node of X_i , where $X_i \cap Z_S = S$ and $Y_S \cap Z_S \supset S$. The analysis here is the same as for Situation # 1 except for $S' := Y_S \cap Z_S$, as we changed the orientation of the edge $Y_S Z_S$ during the transformation. However, after the transformation we have no minimal separator containing S' in the subtree rooted at Z_S (because such a separator should also contain S , and so, should be incident to Y_S). More generally, S' is T_G -convergent because S is. We now focus on the existence of minimal separators $R \subset S'$ in this subtree, and we explain how to correct our transformation when they exist.

A central observation is that we did not change the edges incident to Z_S – as they were after we processed Z_S – until this transformation. To see that, first notice that as we have $X_i \cap Z_S = S$, Z_S was already a neighbour of X_i when we started processing X_i (recall that we consider minimal separators in \mathcal{S}_i by decreasing size). Then, if we had $Z_S \in \{Y_{S''}, Z_{S''}\}$ for some $S'' \in \mathcal{S}_i$ considered before S , the only possibility would be $Y_{S''} = Z_S$ and we fell in Situation # 1 for S'' . However, this would imply $S'' \subseteq S$ that is a contradiction. Hence, we proved our above observation that the edges incident to Z_S did not change since we started processing X_i .

Now, since when we process a node we cannot make this node incident to a new set $E_{R'}(T_G)$, there was an edge $Z_S Y$ in the initial clique-tree T_G that we obtained from Theorem 10 such that $Z_S \cap Y = S'$ and Y was a child node of Z_S (possibly, $Y = Y_S$, or Y became a child node of Y_S after some previous processing). In particular, in any subtree rooted at a child node of Z_S , Theorem 10 ensures that there was no minimal separator contained into S' . As we did not change the set of descendants of Z_S until this transformation we so obtain that, if there is some edge labeled by $R \subset S'$ in the subtree rooted at Z_S then, there was initially such an edge incident to Z_S . Such edges could have been made incident to a descendant of Z_S in

order to make R T_G -convergent but then, this should have been to Y_S . We can so conclude this did not happen and all such edges stayed incident to Z_S .

Repairing the procedure. We consider all these minimal separators $R \subset S'$ that appear in the subtree rooted at Z_S by decreasing size. For every such a R , we first verify whether R is T_G -convergent. (Note that before we start the procedure R was either T_G -convergent or not weakly T_G -convergent.)

- (a) If it is not the case then, we move all edges in $E_R(T_G)$ in order to make them incident to Y_S . Specifically, for any child B of Z_S such that $B \cap Z_S = R$, we make Y_S the new father node of B . See Fig. 9 for an illustration.
- (b) Otherwise, R is T_G -convergent, and we also verify whether $|R| \geq 3$ and R is strictly contained into at least one other minimal separator that appears in the subtree rooted at Z_S . If it is *not* the case then, we move all edges in $E_R(T_G)$ as explained above.

Overall, at the end of this phase all the edges $E_R(T_G)$, $R \subset S'$ that we could not possibly move satisfy Condition 3 of the theorem. In order to prove correctness of our repairing procedure, we must prove that we do not violate one of the conditions stated in the theorem for any minimal separator or edge that we already processed. Suppose by contradiction that we did violate one of these properties for some edge labeled by R' and already processed. There exists $R \subset S'$ such that after moving edges in $E_R(T_G)$, we violated one of the conditions stated in the theorem for this edge, and this stayed so until the end of the repairing procedure. Since we did not modify the ancestor/descendant relationships for any other node than Y_S and Z_S , the only way so that it can happen is that the move did not preserve the property for R' to be either T_G -convergent or not weakly T_G -convergent.

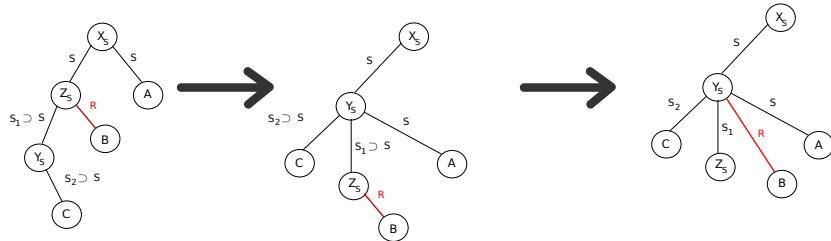


Figure 9: Repairing the procedure when $R \subset S_1$.

Clearly this above property can be modified only if $R' \subseteq R$. We can further assume $|R'| \geq 3$ since otherwise, all conditions stated in the theorem will stay true. We first prove that R' cannot become weakly T_G -convergent after the move, unless this was already the case before we started this phase. Otherwise already when we processed the node Z_S , we could have made R T_G -convergent by making all edges in $E_R(T_G)$ incident to Y_S . As a result, if R' loses the property of being not weakly T_G -convergent then, before we start the repairing procedure it was T_G -convergent (Condition 1) and so, all edges in $E_{R'}(T_G)$ must be incident to Z_S . In the same way if R' loses the property of being T_G -convergent then, as $R' \subseteq R \subset S'$, it also implies all edges in $E_{R'}(T_G)$ must be incident to Z_S . We so deduce that either $R' = R$ or we will also consider R' later during the repairing procedure. In particular if $R' \subset R$ and we

would move edges in $E_R(T_G)$ then, two possibilities could arise when we eventually consider the minimal separator R' :

- R' is weakly T_G -convergent. Since we assume that we previously moved edges in $E_R(T_G)$, R' is not T_G -convergent. Then, moving edges in $E_{R'}(T_G)$ would keep for R' its initial property of being T_G -convergent at the end of the procedure.
- R' is not weakly T_G -convergent. In particular in the subtree rooted at B' , $B' \cap Z_S = R'$, there can be no minimal separator strictly contained into R' (otherwise, by Condition 3 of the theorem such minimal separators should be T_G -convergent, which is impossible since $R' \subset S'$ and the edges labeled by S' are not in this clique-subtree). By moving edges in $E_{R'}(T_G)$, R' would stay not weakly T_G -convergent. As a result, we could move edges in $E_{R'}(T_G)$ without violating the conditions stated in the theorem for those edges.

We are left with the special case $R = R'$. Moving edges in $E_R(T_G)$ cannot change the property for R of being (not) weakly T_G -convergent. Thus, the only subcase left is when R was T_G -convergent, $|R| \geq 3$, and the move would violate this property. However, only in this subcase we do not move the edges in $E_R(T_G)$, that is a contradiction. Indeed, R must be strictly contained into at least one other minimal separator that appears in the subtree rooted at Z_S (otherwise, the move could not violate the property for R of being T_G -convergent).

□

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

We use Theorem 9 in this section in order to derive, for every clique-intersection $X \in \mathcal{K}(G) \cup \mathcal{S}(G)$, a polynomial-size family of all possible subtrees $T\langle X \rangle$ we could have in a well-structured 4-Steiner root of G . We first focus on the case of minimal separators (Section 6.1), before extending our results to the maximal cliques that are either leaf-nodes (Section 6.2) or internal nodes (Section 6.3) of the clique-tree T_G .

6.1 Case of Minimal Separators

A key intermediate Step in our algorithm is, for any minimal separator S , the polynomial-time construction of a family \mathcal{T}_S of potential subtrees $T\langle S \rangle$:

Theorem 12. *Let $\mathcal{S}(G)$ be the set of all minimal separators in $G = (V, E)$. In $\mathcal{O}(n^3m)$ -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 .*

This result will be further exploited in the next sections.

Proof. Let $S \in \mathcal{S}(G)$ be fixed. We will show how to construct \mathcal{T}_S in $\mathcal{O}(|S|^4)$ -time. Since $|\mathcal{S}(G)| = \mathcal{O}(n)$, $\max\{|S| \mid S \in \mathcal{S}(G)\} = \mathcal{O}(n)$ 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 8, 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 8 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. For ease of reasoning, fix some (unknown) well-structured 4-Steiner root T where $T\langle S \rangle$ is a bistar, and write $\mathcal{C}(T\langle S \rangle) = \{c_0, c_1\}$. 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 2. *If $\{c_0, c_1\}$ is a light part then, there is a heavy part that strictly contains it.*

Proof. Suppose by contradiction $X = \{c_0, c_1\}$ is a light part and no heavy part contains it. Let X_i be any maximal clique such that $X \subseteq X_i$ but $S \not\subseteq X_i$. Such a X_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, $X \subseteq X_i \cap S$, and so $X = X_i \cap S$ since there is no heavy part containing X . Furthermore we have $X_i \not\subseteq S$, hence there exists $j \in \{0, 1\}$ such that c_j has a neighbour in $T\langle X_i \rangle \setminus S$ (possibly, a Steiner node). Then, by Lemma 2 applied to $N_T[c_j]$ and $T\langle X_i \rangle$, $\text{diam}(N_T[c_j] \cup T\langle X_i \rangle) = \text{diam}(T\langle X_i \rangle) \leq 4$, and so, $X_i \cup \text{Real}(N_T[c_j])$ is a clique of G . By maximality of X_i , $\text{Real}(N_T[c_j]) \subseteq X_i$. However, there is at least one leaf in $\text{Real}(N_T[c_j]) \setminus X$, that implies $X \subset X_i \cap S$, a contradiction. \diamond

Then, we divide the proof in two subcases:

- 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 $X \subseteq N_T[c_1]$. By Property 2 of Theorem 8 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 bisters to put into \mathcal{T}_S , and so, this takes $\mathcal{O}(|S|^3)$ -time².
- From now on we assume that there is no heavy part. By Claim 2, $\{c_0, c_1\}$ 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 2 of Theorem 8). In this situation, we construct the *intersection graph* I_2 of the light parts in S (*i.e.*, with a vertex for each light part and an edge between every two light parts with a common intersection). We claim that assuming G has a 4-Steiner root T , I_2 has at most two connected components, that follows from 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 $\text{Real}(N_T[c_j]) = X$. Two different situations might occur:

²We will actually show in Lemma 11 this number of potential bisters 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. Situation # 1 (see Fig. 10): 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 $X' \subseteq \text{Real}(N_T[c_j]) = X$, a contradiction), this must be c_{1-j} . In particular, for any $X' \neq X$, $T\langle X' \rangle$ must be an edge between c_{1-j} and a leaf (thereby implying $X' \cap X \neq \emptyset$).

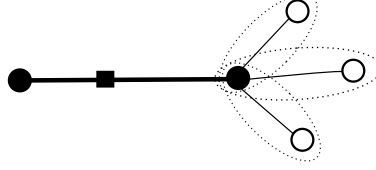


Figure 10: Situation 1: the subtree $T\langle X \rangle$ is drawn in bold. There are 3 other light parts represented by dashed ellipses.

2. Situation # 2 (see Fig. 11): 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, $\text{Real}(N_T[c_{1-j}]) = X'$, and so, there are no other light part in S than X and X' .

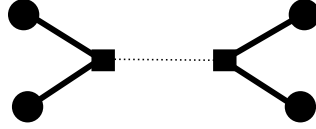


Figure 11: Situation 2: the subtrees $T\langle X \rangle$ and $T\langle X' \rangle$ are drawn in bold.

- 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 $\text{Real}(N_T[c_0]), \text{Real}(N_T[c_1])$. See Fig. 12 for an illustration.

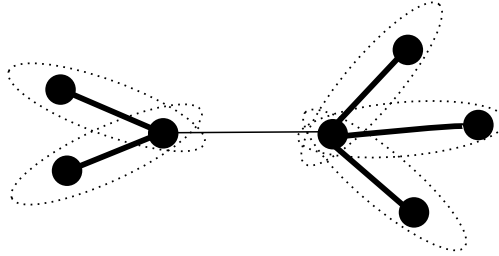


Figure 12: A case where all light parts must be edges.

The proof of this above claim also shows there are essentially two possibilities in order to position the S -constrained vertices:

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

Overall, this reduces the placement of S -constrained vertices to the construction of two stars in parallel, thereby giving $\mathcal{O}(|S|^2)$ different possibilities. 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 9, 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 $\mathcal{O}(|S|^3)$ different trees in \mathcal{T}_S , that takes $\mathcal{O}(|S|^4)$ -time. \square

Remark 1. We only use the fact that a minimal separator is a clique-intersection. In particular, we can use the algorithm of Theorem 12 in order to generate, for any maximal clique X_i without a X_i -free vertex, all possible subtrees $T\langle X_i \rangle$ of diameter at most 3 in any well-structured 4-Steiner root of G .

6.2 Step 2: Case of a Leaf Node

The main purpose of this section is to prove the following result (base case of our dynamic programming algorithm):

Theorem 13. *Given $G = (V, E)$ and a rooted clique-tree T_G of G , let $X_i \in \mathcal{K}(G)$ be a leaf. We can construct, in time polynomial in $|X_i|$, a set \mathcal{T}_i of 4-Steiner roots for $G_i := G[X_i]$ 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 $T\langle X_i \rangle$.*

Proof. Let $X_{p(i)}$ be the father node of X_i . By Theorem 6, $S_i := X_i \cap X_{p(i)}$ is a minimal separator. We compute the family \mathcal{T}_{S_i} given by Theorem 12. 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 $X_i \setminus S_i$ are all simplicial, and so, we can connect them to $\mathcal{C}(T_i)$ as explained in Lemma 6 (one possibility up to Steiner equivalence). Note that in doing so, we may also obtain solutions T_i such that $\text{diam}(T_i) > 4$ or T_i is not well-structured, that we will need to discard. See Fig. 13 for an illustration. Overall, $|\mathcal{T}_i| = \mathcal{O}(|S_i| |\mathcal{T}_{S_i}|) = \mathcal{O}(|S_i|^4) = \mathcal{O}(|X_i|^4)$. \square

In order to anticipate an intermediate problem that we will introduce in Section 7, we end up this section with the following consequence of Theorem 13:

Corollary 2. *Given $G = (V, E)$ and a rooted clique-tree T_G of G , let $X_i \in \mathcal{K}(G)$ be a leaf and let $(d_v)_{v \in V(T\langle S_i \rangle)}$ be a sequence of positive integers.*

We can construct, in time polynomial in $|X_i|$, a set \mathcal{T}_i of 4-Steiner roots for $G_i := G[X_i]$ with the following additional property: If G has a well-structured 4-Steiner root T where, for any $v \in V(T\langle S_i \rangle)$:

$$\text{dist}_T(v, X_i \setminus S_i) \geq d_v$$

then, there exists $T'_i \in \mathcal{T}_i$ Steiner-equivalent to $T\langle X_i \rangle$.

Proof. We construct the family given by Theorem 13. We only keep the trees $T_i \in \mathcal{T}_i$ that satisfy the additional distance constraints we have. \square

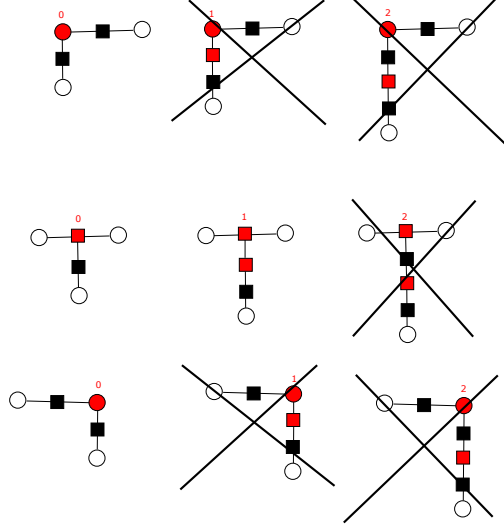


Figure 13: An execution of the algorithm of Theorem 13. The minimal separator has size two and induces a star. There is one simplicial vertex to add.

6.3 Case of an internal node

The main objective for the remaining of this section is given some internal node X_i of T_G , to compute a family \mathcal{F}_i of all possible $T\langle X_i \rangle$ in a well-structured 4-Steiner root of G . By Lemma 6 there always exists a root T where all X_i -free vertices are leaves of $T\langle X_i \rangle$ and connected to $\mathcal{C}(T\langle X_i \rangle)$ by a path of length two whose internal node is Steiner. Thus, we can first assume for simplicity X_i does not contain any X_i -free vertex (*i.e.*, such vertices will be added at the end of the construction). Furthermore, as noticed earlier (Remark 1) we can use the algorithmic proof of Theorem 12 in order to generate all the subtrees $T\langle X_i \rangle$ of diameter at most three to be added in \mathcal{F}_i . Therefore, as a consequence of other results in this paper we are only interested in generating trees of diameter *exactly* four with no X_i -free vertex.

Our main tool for this task is a careful analysis of the intersections between the minimal separators in X_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 can happen, and show how to handle with it, in Section 6.3.2. Proposition 1 in Section 6.3.3 will summarize our results for this part.

6.3.1 Getting more from clique-intersections

The following lemma will be useful in order to prove our first result in this section:

Lemma 11. *Given $G = (V, E)$ strongly chordal, let $S \in \mathcal{S}(G)$. There exists a family \mathcal{T}_S with the following two properties:*

1. *For any well-structured 4-Steiner root T of G , there exists a well-structured T' such that $T'\langle S \rangle \in \mathcal{T}_S$ and $\text{dist}_{T'}(u, v) \geq \text{dist}_T(u, v)$ for every $u, v \in V$. Moreover, either $T \equiv_G T'$, or $\sum_{u, v \in V} \text{dist}_{T'}(u, v) > \sum_{u, v} \text{dist}_T(u, v)$.*

2. For any $R \subset S$ and $c \in R$, there is at most one bistar $T\langle S \rangle \in \mathcal{T}_S$ such that $c \in \mathcal{C}(T\langle S \rangle)$ and $N[c] = R$. In the same way, there is at most one bistar $T\langle S \rangle \in \mathcal{T}_S$ such that $\alpha \in \mathcal{C}(T\langle S \rangle)$ is Steiner and $N(\alpha) = R$.

Moreover, \mathcal{T}_S has size $\mathcal{O}(|S|^3)$ and it can be computed in polynomial time.

In order to better understand the significance of Lemma 11, 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 Property 2 says is that there is essentially one canonical way to compute this bistar given R and c . The more technical Property 1 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).

Proof. The result is obtained by applying some polynomial-time post-processing to the family \mathcal{T}_S of Theorem 12. We consider all possible R, c such that, for some $T\langle S \rangle \in \mathcal{T}_S$ we have $c \in \mathcal{C}(T\langle S \rangle)$ and $Real(N[c]) = R$. W.l.o.g., for any fixed R we keep at most one such a pair such that c is a Steiner node. There are only $\mathcal{O}(|\mathcal{T}_S|)$ possibilities for a fixed S , that is in $\mathcal{O}(|S|^3)$ by Theorem 12. For each such a pair, 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 $c' \in (X \cap R) \setminus \{c\}$;
- 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 7). Furthermore, amongst all the bisters $T\langle S \rangle \in \mathcal{T}_S$ such that $c \in \mathcal{C}(T\langle S \rangle)$ and $Real(N[c]) = R$, we only keep the one such that either $\mathcal{C}(T\langle S \rangle) = \{c, c'\}$ (if $c' \in R$) or the unique node in $\mathcal{C}(T\langle S \rangle) \setminus \{c\}$ is Steiner.

In order to prove correctness of this post-processing, we fix any well-structured 4-Steiner root T of G such that $T\langle S \rangle$ is a bistar, $c \in \mathcal{C}(T\langle S \rangle)$ and $Real(N_T[c]) = R$. Let $\mathcal{C}(T\langle S \rangle) = \{c, c_2\}$. If either $c' = c_2$ or both c_2, c' are Steiner nodes then, we are done. So, we assume from now on this is not the case. In this situation, we first prove that a clique-intersection X as above cannot exist. Indeed, suppose by contradiction such a X exists. As we have $X \not\subseteq R$, 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 , which contradicts our assumption that $c_2 \neq c'$. Therefore, we proved as claimed that no such a clique-intersection X can exist. This implies $c_2 \in R$ but c' is Steiner.

We reuse a transformation we introduced in the proof of Theorem 9 (*i.e.*, Operation 1 for $v = c_2$ and c). Specifically, we define Q_{c_2} as the subtree of T that is 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 . 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 $c'c_2$. 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.

Furthermore, we claim that T' keeps the property to be a well-structured 4-Steiner root of G , that will prove the lemma. This part of the proof closely follows Claim 1. First, suppose by contradiction there exist $x \in Real(Q_{c_2}), y \in V \setminus V(Q_{c_2})$ such that $dist_T(x, y) \leq 4$ but $dist_{T'}(x, y) > 4$. The unique xy -path in T goes by c_2 (see Fig. 11 for an illustration). Furthermore, this path also goes by a node z in $T\langle S \rangle \setminus \{c_2\}$ since we have $y \notin Q_{c_2}$. We cannot have $z \in N_T[c]$ since otherwise, the distance between z and c_2 , and so, the distance between x and y , did not change

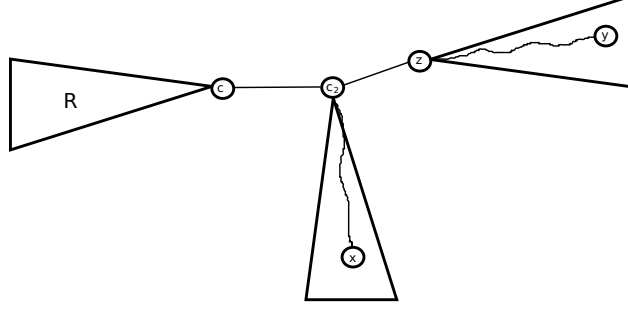


Figure 14: To the proof of Lemma 11.

in T' . Therefore, $z \notin R$ is a leaf in $T\langle S \rangle$. But then, there exists a maximal clique X_k such that $x, y, c_2, z \in X_k$, and so, $X := S \cap X_k \supseteq \{c_2, z\}$. In particular, $|X| \geq 2$ and we cannot have $X \subseteq N_T[c]$ because $X \setminus R \neq \emptyset$. Therefore, either $|X| \geq 3$ or $|X| = 2$ and $c \notin X$. In the former case, $T\langle X \rangle$ is a star with center c_2 , and so, $X \cap N_T[c] \subseteq \{c, c_2\}$. Otherwise, $X \subseteq N_T[c_2] \setminus \{c\}$ which implies $N_T[c] \cap X = \{c_2\}$. In both cases, this contradicts the non existence of a clique-intersection X as defined earlier in the proof. As a result, T' is a 4-Steiner root of G . We can prove similarly as above $\text{dist}_{T'}(\text{Real}(Q_{c_2} \setminus \{v\}), V \setminus Q_{c_2}) \geq \text{dist}_T(\text{Real}(Q_{c_2} \setminus \{v\}), V \setminus Q_{c_2}) > 4$. However, one still needs to ensure that T' is well-structured.

For that, it suffices to consider all the clique-intersections X' that contain c_2 (because otherwise, $T\langle X' \rangle = T'\langle X' \rangle$ and so we are done). 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 is a contradiction). 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 $\text{dist}_T(\text{Real}(Q_{c_2} \setminus \{v\}), V \setminus Q_{c_2}) > 4$. 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. Note that in doing so, any X' -free vertex that was a leaf in $T\langle X' \rangle$ is also a leaf of $T'\langle X' \rangle$. Furthermore, the above transformation cannot add new internal real nodes onto the path between such a leaf and the center nodes, that implies we cannot break Property 2 of Theorem 9. We can break Property 1 only if c_2 is X' -free. However, this is not possible if $S = X'$ because c_2 is not a leaf of $T\langle S \rangle$ and we assume T is well-structured. This cannot be the case if $S \subset X'$ either, because $v \in S$ and $|S| \geq 3$ (*i.e.*, because there is at least one leaf adjacent to each center node in $T\langle S \rangle$). Finally, $\text{diam}(T'\langle X' \rangle) \geq \text{diam}(T\langle X' \rangle)$ and so, we cannot break Property 3 of Theorem 9.
- Otherwise, in all other cases we prove $T'\langle X' \rangle = T\langle X' \rangle$. To see that, first note this may not be the case only if $T\langle X' \rangle$ is not fully contained into Q_{c_2} . Then, $T\langle X' \rangle$ must intersect $T\langle S \rangle \setminus \{c_2\}$. If we suppose by contradiction $T'\langle X' \rangle \neq T\langle X' \rangle$ then in particular, this implies $T\langle X' \rangle$ must contain a real node in $N_T(c_2) \setminus \{c\} = S \setminus R$. But then, $X := S \cap X'$ satisfies $|X| \geq 2$, $c_2 \in X$ and $X \not\subseteq R$, that contradicts our assumption that no such a X exists.

The above case analysis ends up proving that T' is well-structured, thereby proving the lemma. \square

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

Lemma 12. *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 8, $\text{Real}(N_T[c]) = S$. Furthermore since $T\langle S \rangle$ 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 $\text{dist}_T(c, \mathcal{C}(T\langle X \rangle)) \leq 1$. \square

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

Lemma 13. *Let X_i be a maximal clique of $G = (V, E)$ with no X_i -free vertex. In $\mathcal{O}(n|X_i|^5)$ -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 $S \subset X_i$, $T\langle S \rangle$ is a bistar, there is a T' such that $T'\langle S_i \rangle \equiv_G T\langle S_i \rangle$, $T'\langle X_i \rangle \in \mathcal{B}_i$ and $\text{dist}_{T'}(v, V_i \setminus S_i) \geq \text{dist}_T(v, V_i \setminus S_i)$ for every $v \in V(T\langle S_i \rangle)$.*

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

Proof. Let $\mathcal{S}_i \subseteq \mathcal{S}(G)$ contain all the minimal separators in X_i . By Theorem 12, for any $S \in \mathcal{S}_i$ we can construct a family \mathcal{T}_S such that, in any $T_{X_i} \in \mathcal{F}_i$, we should have $T_{X_i}\langle S \rangle$ is Steiner-equivalent to some tree in \mathcal{T}_S . This takes total time $\mathcal{O}(|X_i|^4|\mathcal{S}_i|) = \mathcal{O}(n|X_i|^4)$. Fix $S \in \mathcal{S}_i$ (there are $\mathcal{O}(n)$ possibilities) and a bistar $T\langle S \rangle \in \mathcal{T}_S$ (by Theorem 12, there are $\mathcal{O}(|S|^3)$ possibilities, that is in $\mathcal{O}(|X_i|^3)$). 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 $S' \in \mathcal{S}_i$ 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 $S' \in \mathcal{S}_i$ that are *not* contained into any other $S'' \in \mathcal{S}_i$. 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 \mathcal{S}_i .

In what follows is a simple observation for the case $S \cap S' \neq \emptyset$ (see also Fig. 15 for an illustration).

Claim 3. *If $S \cap S' = \emptyset$ then, $T\langle S' \rangle$ is a star with a Steiner central node. Moreover, the center node of $T\langle X_i \rangle$ must be Steiner.*

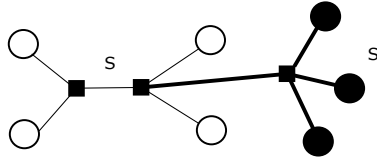


Figure 15: An example where $S \cap S' = \emptyset$.

Proof. By Lemma 8 $\text{Real}(N_T[\mathcal{C}(T\langle X_i \rangle)]) \subset S$. But then, $T\langle X_i \rangle \setminus N_T[\mathcal{C}(T\langle X_i \rangle)]$ 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 $N_T[\mathcal{C}(T\langle X_i \rangle)]$. In the former case we so conclude that v is X_i -free by inclusion wise maximality of S' and by Lemma 5. Since we assume there is no X_i -free vertex, this case cannot

happen. Therefore, $T\langle S'\rangle$ is a star with a Steiner central node in $N_T[\mathcal{C}(T\langle X_i\rangle)]$. Finally, the center of $T\langle X_i\rangle$ must be also Steiner (otherwise, this vertex should be in S'). \diamond

If $S \cap S' = \emptyset$ then, by combining Claim 3 and Lemma 12 there is essentially one way to insert S' in $T\langle X_i\rangle$ (*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 center node of $T\langle X_i\rangle$).

For the remaining cases, we assume $S' \cap S \neq \emptyset$ for any inclusion wise maximal $S' \in \mathcal{S}_i$. Several cases may arise:

- Case there exist $u, v \in S \cap S'$ nonadjacent in $T\langle S\rangle$ (see Fig. 16 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. By the proof of Lemma 8 we must have $T\langle S\rangle \cap T\langle S'\rangle = N_T[\mathcal{C}(T\langle X_i\rangle)]$. Since there exist $u, v \in S \cap S'$ nonadjacent in $T\langle S\rangle$, the central node of $T\langle X_i\rangle$ 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 the neighbourhood $N_T[c_i]$ is fixed by $T\langle S\rangle$, by Lemma 11 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 $T\langle S'\rangle$.

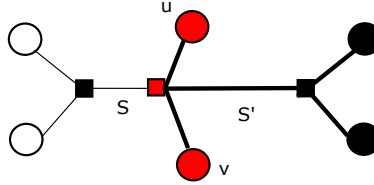


Figure 16: Two bisters intersecting.

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 $S \cap S' \subseteq N_T[\mathcal{C}(T\langle X_i\rangle)]$. W.l.o.g., $u \in \mathcal{C}(T\langle S\rangle)$ (or equivalently, u must be the central node of $T\langle X_i\rangle$) and v is a leaf of $T\langle S\rangle$. Several situations force $T\langle S'\rangle$ to be a star, for instance if:
 - S' is strictly contained into another minimal separator of G ;
 - or $Real(N_T[u]) \neq \{u, v\}$.

If such a situation occurs then, by Lemma 12 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. 17).

From now on assume no minimal separator strictly contains S' and $Real(N_T[u]) = \{u, v\}$. $T\langle S'\rangle$ is forced to be a bistar if there exists at least one $S'' \in \mathcal{S}_i$ 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 12). 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 $S'' \in \mathcal{S}_i$ that are inclusion wise maximal (including S') as

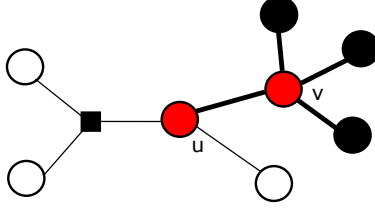


Figure 17: Star intersecting a bistar.

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 12 there would be two non-edge stars with the same center node v , that contradicts Property 2 of Theorem 8). Summarizing, since all these sets $S'' \setminus S$ are pairwise disjoint, we are left with $\mathcal{O}(|X_i|)$ 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. 18.

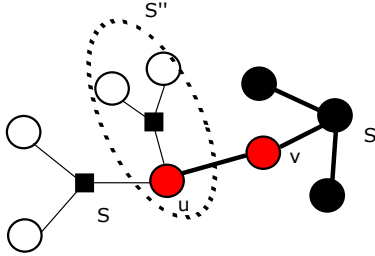


Figure 18: A case where there are two possibilities for $T\langle S' \rangle$.

- Case $S \cap S' = \{v\}$. If $v \in \mathcal{C}(T\langle S \rangle)$ then, 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 ;
 - or $Real(N_T[\mathcal{C}(T\langle X_i \rangle)]) \neq \{v\}$.

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' and $Real(N_T[\mathcal{C}(T\langle X_i \rangle)]) = \{v\}$. In particular, the unique central node of $T\langle X_i \rangle$ is some Steiner node α_i . Furthermore, as we only consider the inclusion wise maximal elements $S'' \in \mathcal{S}_i$ intersecting S , we must have $S \cap S'' = \{v\}$. 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. However, 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 $T\langle X_i \rangle$.

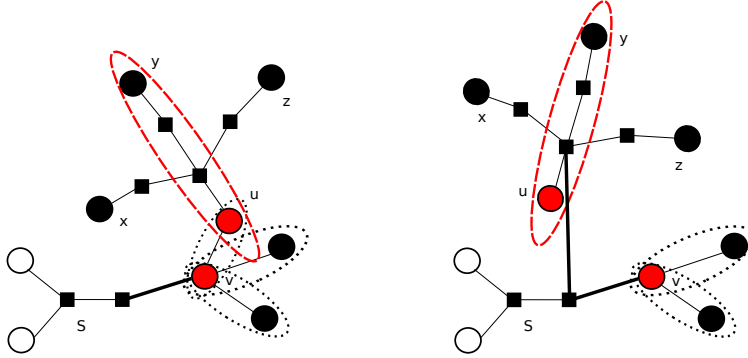


Figure 19: The transformation of an edge into a bistar.

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 12, 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 $T\langle X_{i_j} \rangle$. 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 $\text{diam}(T\langle X_{i_j} \rangle) = 4$, $\text{dist}_T(u, W_{i_j}) = 3$, and all other real vertices of $T\langle X_{i_j} \rangle$ must be leaves at distance two from $\mathcal{C}(T\langle X_{i_j} \rangle)$. See Fig. 19. We connect the unique node $\alpha_{i_j} \in \mathcal{C}(T\langle X_{i_j} \rangle)$ (which is Steiner) to α_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. Since all neighbours of α_i except v and all neighbours of α_{i_j} except u are Steiner nodes, T' keeps the property of being a 4-Steiner root of G . Furthermore, by inclusion wise maximality of S_{i_j} we have $T\langle X \rangle = T'\langle X \rangle$ for every clique-intersection $X \notin \{S_{i_j}, X_i, X_{i_j}\}$, that implies T' is well-structured. We end up observing $\text{dist}_{T'}(x, y) \geq \text{dist}_T(x, y)$ for every $x, y \in T\langle X_i \rangle$, and more generally $\text{dist}_{T'}(x, W_i) \geq \text{dist}_T(x, W_i)$ for every $x \in T\langle S_i \rangle$, 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 $\mathcal{O}(|X_i|)$ possibilities for $T\langle X_i \rangle$. \square

6.3.2 A degenerate case

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

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

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

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

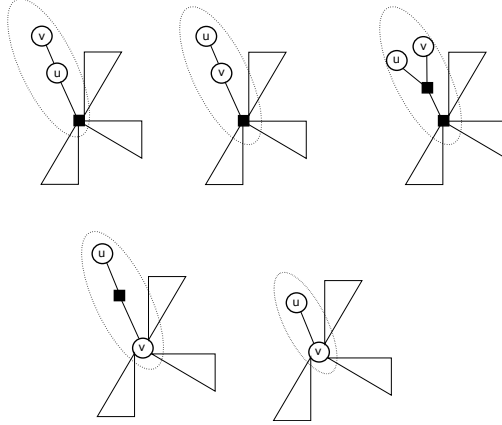


Figure 20: Examples of thin branches (represented by a dashed ellipse).

In order to understand the difficulties we met, assume on the way to construct $T\langle X_i \rangle$ we correctly identified $\mathcal{C}(T\langle X_i \rangle)$ 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 20). 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 $T\langle X_i \rangle$ 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 14. *Let X_i be a maximal clique of $G = (V, E)$ with no X_i -free vertex. There exists a family \mathcal{D}_i that can be computed in $\mathcal{O}(n|X_i|^4)$ -time and such that the following hold for any well-structured 4-Steiner root T of G :*

1. *If $\text{diam}(T\langle X_i \rangle) = 4$ and there is no minimal separator $S \subset X_i$ such that $T\langle S \rangle$ is a bistar then, we have $(T'\langle Y_i \cup \mathcal{C}(T\langle X_i \rangle) \rangle, \mathcal{C}(T\langle X_i \rangle)) \in \mathcal{D}_i$ for some $T' \equiv_G T$ and $Y_i \subseteq X_i$;*
2. *Moreover, $S_i \subseteq Y_i$, and for any $v \in X_i \setminus Y_i$ there is a minimal separator $S \subseteq (X_i \setminus Y_i) \cup \mathcal{C}(T\langle X_i \rangle)$ 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 $S \subset X_i$, $T\langle S \rangle$ has diameter at most two. For that, we only need to consider the subset \mathcal{S}_i of all minimal separators $S \subset X_i$ that are not strictly contained into any other minimal separator in X_i . Furthermore by the hypothesis there is no X_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 $\mathcal{C}(T\langle X_i \rangle) = \mathcal{C}(T\langle S \rangle)$. 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 $\mathcal{O}(|X_i|)$ possibilities). By Lemma 12, $Real(N_T[\mathcal{C}(T\langle S \rangle)]) = S$. 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 (possibly, an edge) with v as a center node; otherwise, $T\langle S' \rangle$ must be a non-edge star with a Steiner center node $\alpha \in N_T(\mathcal{C}(T\langle S \rangle))$. See Fig. 21 for an illustration of that case.

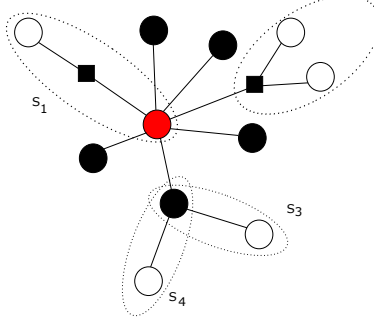


Figure 21: Case 1 of Lemma 14.

- Case there is no $S \in \mathcal{S}_i$ such that $\mathcal{C}(T\langle X_i \rangle) = \mathcal{C}(T\langle S \rangle)$. Fix $c_i \in \mathcal{C}(T\langle X_i \rangle)$ as being any vertex of X_i or Steiner (this gives $\mathcal{O}(|X_i|)$ 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 12) $Real(N_T[\mathcal{C}(T\langle S \rangle)]) = S$, 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, S is 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 $u \neq v$. Then, their union must be a star: where the center is the unique vertex u in $\bigcap S^j$, and $Real(N_{T\langle X_i \rangle}[u]) = \{v\} \cup \left(\bigcup_j S^j\right)$. 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. 22 for an illustration of these subcases.

Finally, according to Definition 4, 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 1. *Let X_i be a maximal clique of $G = (V, E)$. In $\mathcal{O}(n|X_i|^6)$ -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 $Y_i \subseteq X_i$ such that $(T'\langle Y_i \cup \mathcal{C}(T'\langle X_i \rangle) \rangle, \mathcal{C}(T'\langle X_i \rangle)) \in \mathcal{F}_i$ and we have:*

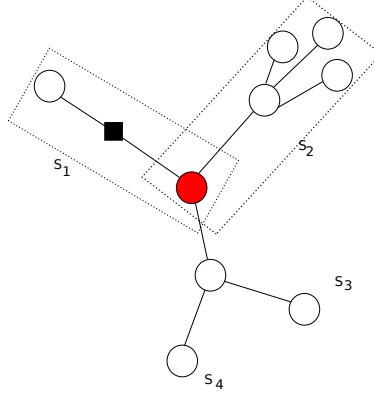


Figure 22: Case 2 of Lemma 14. Thin branches are identified by dotted rectangles.

- $S_i \subseteq Y_i$ and $T'\langle S_i \rangle \equiv_G T\langle S_i \rangle$;
- $\text{dist}_{T'}(x, V_i \setminus S_i) \geq \text{dist}_T(x, V_i \setminus S_i)$ for any $x \in V(T\langle S_i \rangle)$;
- For any $v \in X_i \setminus Y_i$ there is a minimal separator $S \subseteq (X_i \setminus Y_i) \cup \mathcal{C}(T\langle X_i \rangle)$ such that $v \in S$ and $T\langle S \rangle$ is a thin branch.

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

Proof. If there are X_i -free vertices then, by using Lemma 6, 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 $T\langle X_i \rangle$ which, as explained in the proof of Theorem 13, can only increase the total runtime by a multiplicative factor in $\mathcal{O}(|X_i|)$. Thus from now on we may assume that X_i has no X_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 1). We explain in Lemma 13 how to compute the subfamily \mathcal{B}_i of all diameter-four subtrees where at least one minimal separator $S \subset X_i$ has $T\langle S \rangle$ being a bistar. Finally, Lemma 14 completes the construction of the family \mathcal{F}_i . \square

7 Step 3: Deciding the partial solutions to store

In what follows, let X_{i_j} be a fixed child of X_i in T_G . Recall that in the next Step, we will compute a subset \mathcal{T}_{i_j} of 4-Steiner roots for G_{i_j} . During this current Step, we compute a series of “indications” to be transmitted to X_{i_j} in order to enforce the number of partial solutions that we will store in \mathcal{T}_{i_j} to stay polynomial in $|S_{i_j}|$. For that, we introduce the following problem:

Problem 3 (DISTANCE-CONSTRAINED ROOT).

Input: a graph $G = (V, E)$, a maximal clique X_{i_j} , a tree $T_{S_{i_j}} \in \mathcal{T}_{S_{i_j}}$, and a sequence $(d_v)_{v \in V(T_{S_{i_j}})}$ 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, $\forall v \in V(T_{S_{i_j}})$: $\text{dist}_{T_{i_j}}(v, V_{i_j} \setminus S_{i_j}) \geq d_v$; Or \perp if there is no such a root which can be extended to some well-structured 4-Steiner root T of G .

Theorem 14. *Given $G = (V, E)$ chordal and a rooted clique-tree T_G as in Theorem 11, let X_i be an internal node with children $X_{i_1}, X_{i_2}, \dots, X_{i_p}$. If we can solve DISTANCE-CONSTRAINED ROOT in polynomial time then, we can compute in polynomial time 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 $j \in \{1, 2, \dots, p\}$, $|\mathcal{T}_{i_j}| = \mathcal{O}(|S_{i_j}|^3)$;
2. For any well-structured 4-Steiner root T of G , there exists a T' such that $T\langle X_i \rangle \equiv_G T'\langle X_i \rangle$, $T'\langle V_{i_j} \rangle \in \mathcal{T}_{i_j}$ for any $j \in \{1, 2, \dots, p\}$, and (only if $X_i \neq X_0$) $\text{dist}_{T'}(v, V_i \setminus S_i) \geq \text{dist}_T(v, V_i \setminus S_i)$ for any node $v \in V(T\langle S_i \rangle)$.

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 X_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 X_i , we cannot miss a solution.

The remaining of this subsection is now devoted to the proof of Theorem 14. We will use some additional terminology that we define next:

Definition 5. 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 $X_i \in \mathcal{K}(G)$, let $S_i := X_i \cap X_{p(i)}$, 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 $d_{T_{i_j}}(v)$ as a shorthand for $\text{dist}_{T_{i_j}}(v, W_{i_j})$. We observe that for any $v \in V(T\langle S_{i_j} \rangle)$ we have $\text{dist}_{T_{i_j}}(v, W_{i_j}) \leq \text{dist}_{T_{i_j}}(v, X_{i_j} \setminus S_{i_j}) \leq 4$.

Outline of the proof. We process the children nodes X_{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 12, and we consider the subtrees $T\langle S_{i_j} \rangle \in \mathcal{T}_{S_{i_j}}$ sequentially. We divide the proof into several cases depending on the value of the diameter of $T\langle S_{i_j} \rangle$.

- If $|S_{i_j}| \leq 2$ then, there can only be $\mathcal{O}(1)$ different possibilities for the pair $T\langle S_{i_j} \rangle, (d_v)_{v \in V(T\langle S_{i_j} \rangle)}$. 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 treatment of the minimal separators S_{i_j} with at least three elements is more intricate (Sections 7.2 and 7.3). For a fixed $T\langle S_{i_j} \rangle$ we define an encoding with only $\mathcal{O}(|S_{i_j}|^{\mathcal{O}(1)})$ possibilities, that essentially summarizes at “guessing” the central nodes of $T\langle X_i \rangle$ and $T\langle X_{i_j} \rangle$. Then, we show that only one solution per possibility 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 small separators contained into S_{i_j} , and on Theorem 11. Indeed, our approach could not work with an arbitrary clique-tree.

7.1 Case $\text{diam}(T\langle S_{i_j} \rangle) \leq 1$.

In this situation, $|S_{i_j}| \leq 2$ and so, there can only be $\mathcal{O}(1)$ possibilities for the distances' constraints $(d_v)_{v \in S_{i_j}}$. 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 proofs of Section 8, storing all these possibilities increases the runtime of our algorithm. We confront this issue with a local optimality criterion. Specifically:

Claim 4. *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 4 if S_{i_j} is a cut-vertex then, we keep exactly one solution in \mathcal{T}_{i_j} .

Claim 5. *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 $T\langle S_{i_j} \rangle$ is an edge. Therefore, 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 5 if $S_{i_j} = \{u, v\}$ and $T\langle S_{i_j} \rangle$ 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 $T\langle S_{i_j} \rangle$ 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 $\mathcal{C}(T\langle S_{i_j} \rangle) = \{c\}$. For any fixed choices of a node $c_i \in \mathcal{C}(T\langle X_{i_j} \rangle) \cap N_T[c]$ and $c_{i_j} \in \mathcal{C}(T\langle X_{i_j} \rangle) \cap N_T[c]$ (in a final solution T), we will prove that all the choices of our algorithm can be based on $\mathcal{O}(1)$ distances. The existence of such two nodes is given by Lemma 12. Of course we do not know the center nodes c_i, c_{i_j} in advance. In what follows we propose to “guess” these nodes. Since c_i and c_{i_j} will be either in S_{i_j} or Steiner, this leaves $|S_{i_j}| + 1$ possibilities for each. Furthermore for every fixed triple $T\langle S_{i_j} \rangle, c_i, c_{i_j}$, we will show that only $\mathcal{O}(1)$ partial solutions will need to be stored. Overall, that gives only $\mathcal{O}(|S_{i_j}|^3)$ different possibilities for stars.

Finally, on our way to upper-bound the number of possibilities to store by a polynomial, 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.

Recall that T_G is a rooted clique-tree of G as stated in Theorem 11. Before starting our analysis, we need to derive a few properties from T_G . Indeed, our approach could not work with an arbitrary clique-tree.

Claim 6. S_{i_j} does not strictly contain any minimal separator of G_{i_j} (otherwise we can discard all the solutions where $T\langle S_{i_j} \rangle$ is a star).

Proof. Suppose by contradiction it is the case. Since we assume $T\langle S_{i_j} \rangle$ is a star, any such a minimal separator should have size at most two. However, Condition 3 of Theorem 11 ensures that all such minimal separators should have size at least three. A contradiction. \diamond

We are now left with two possibilities:

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

Given an arbitrary 4-Steiner root T_{i_j} of G_{i_j} , we extract the following information:

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

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. 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 X_{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 12 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 6), this implies all the paths between S_{i_j} and W_{i_j} must pass by c, c_{i_j} (see Fig. 23 for an illustration). In this situation, our partial encoding already contains all the distances' information we need.

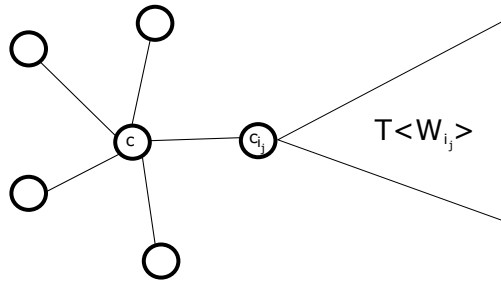


Figure 23: A schematic view of $T\langle V_{i_j} \rangle$.

- Otherwise, $c_{i_j} = c$. A simple transformation of the construction proposed in Lemma 6 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 $T\langle X_{i_j} \rangle$ in such a way that they

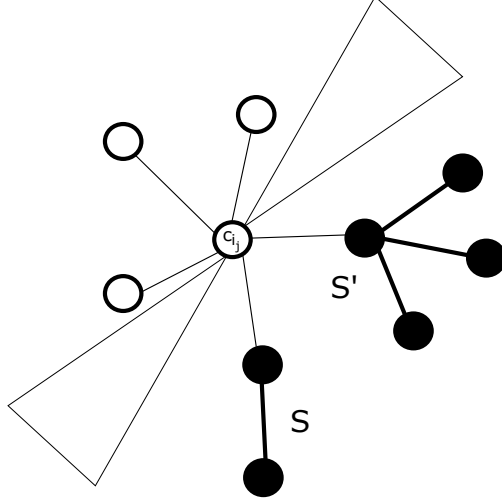


Figure 24: The case $c = c_{i_j}$. Two minimal separators S, S' overlapping S_{i_j} are drawn in bold.

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.

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 8). 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 12 the unique vertex in $(S \cap S_{i_j}) \setminus \{c\}$ is its center. See Fig. 24 for an illustration.

◇

Finally given $\text{short-encode}(T_{i_j})$, we can transform such a short encoding into the constraints $(d_v)_{v \in V(T \langle S_{i_j} \rangle)}$ 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 $v \in S_{i_j}$:

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

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 $T \langle S_{i_j} \rangle$. 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 \mathcal{T}_{i_j}). The reason why this does not matter is that assuming we made a correct guess for $\text{short-encode}(T_{i_j})$, we proved in Claim 7 that we have $d_{T_{i_j}}(v) = d_v$ for any $v \in V(T \langle S_{i_j} \rangle)$. 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 $d_{T'_{i_j}}(v) \geq d_{T_{i_j}}(v)$ for any $v \in V(T \langle S_{i_j} \rangle)$.

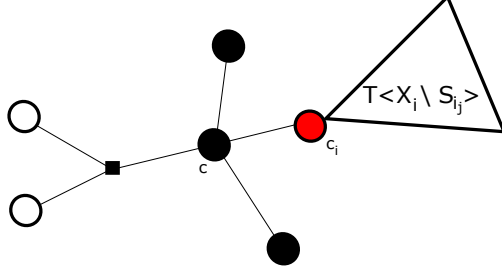


Figure 25: To the proof of Claim 8.

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 distances' conditions, that we will also use in order to define our distances' constraints. 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 $|S_i \cap S_{i_j}| = \mathcal{O}(1)$. Indeed, recall that by Lemma 9, S_{i_j} must be weakly T_G -convergent. Since we assume $|S_{i_j}| \geq 3$, by Condition 1 of Theorem 11 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 $\bigcup_{S_{i_j} \subseteq S'} ES'(T_G)$ must be X_{i_j} . This implies that $S_i := X_i \cap X_{p(i)}$ cannot contain S_{i_j} . In particular, $|S_i \cap S_{i_j}| \leq 2$.

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 $S_{i_k} := X_i \cap X_{i_k}$ between X_i and one of its children nodes such that $S_{i_k} \subset S_{i_j}$. Note that since we assume $T\langle S_{i_j} \rangle$ 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 4 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, $T\langle S_{i_k} \rangle$ 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 8. *Let T_{i_j} satisfy the above distances' conditions. Exactly one of the following conditions is true:*

1. T_{i_j} can be extended to a well-structured 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. 25 for an illustration. 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 $T := T'\langle V \setminus W_{i_j} \rangle \odot T_{i_j}$ is also a 4-Steiner root of G . For that, we start observing that by Condition 3 of Theorem 11, there is a minimal separator of G_{i_j} that strictly contains S_{i_j} . Such a

minimal separator must induce a bistar in T' and T , thereby implying $c \neq c_i$ by Lemmata 8 and 9 (possibly, $c = c_{i_j}$). In particular, all the paths between S_{i_j} and $X_i \setminus S_{i_j}$ will need to pass by c_i , and so:

$$\text{dist}_T(X_i \setminus S_{i_j}, W_{i_j}) = \text{dist}_{T'}(X_i \setminus S_{i_j}, W_{i_j}) \geq 5.$$

We finally consider the other minimal separators $S_{i_k} := X_i \cap X_{i_k}$ between X_i and one of its children nodes that intersect S_{i_j} . We have $S_{i_j} \not\subseteq S_{i_k}$ and so, $|S_{i_k} \cap S_{i_j}| \leq 2$. Furthermore since we have $c \neq c_i$, Lemma 12 implies either $S_{i_k} \subset S_{i_j}$ or $S_{i_k} \cap S_{i_j} \subseteq \{c_i, c\}$. In the latter case:

$$\text{dist}_T(W_{i_k}, W_{i_j}) = \text{dist}_{T'}(W_{i_k}, W_{i_j}) \geq 5.$$

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 4 and the distances' constraints over T_{i_j} that we can always assume $\text{dist}_T(W_{i_k}, W_{i_j}) \geq 5$. 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 $\text{dist}_{T'}(W_{i_k}, W_{i_j}) \geq 5$ 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 5 and the distances' constraints over T_{i_j} that we can always assume $\text{dist}_T(W_{i_k}, W_{i_j}) \geq 5$. \diamond

Finally given **short – encode – 2**(T_{i_j}), we can transform such a short encoding into the constraints $(d_v)_{v \in V(T\langle S_{i_j} \rangle)}$ 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 a condition.
- For all other nodes $v \in S_{i_j}$: $d_v = 1$ (trivial constraint).

For any fixed $T\langle S_{i_j} \rangle$ the mapping $\varphi : \text{short – encode – 2}(T_{i_j}) \rightarrow \langle c_i, (d_v)_{v \in V(T\langle S_{i_j} \rangle)} \rangle$ is injective. Moreover in any final solution extending to all of G , we proved in Claim 8 that we will have all the paths between $T\langle S_i \rangle$ and W_{i_j} which need to pass 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 X_i provided one exists.

7.3 Case $T\langle S_{i_j} \rangle$ 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 8, we must have $\mathcal{C}(T\langle S_{i_j} \rangle) = \{c_i, c_{i_j}\}$.

Before choosing our short encoding, we will need the properties of T_G given by Theorem 11.

Claim 9. *Let T_{i_j} be a 4-Steiner root of G_{i_j} such that $T_{i_j}\langle S_{i_j} \rangle \equiv_G T\langle S_{i_j} \rangle$ is a bistar. 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 implicitly deduced from $d_{T_{i_j}}(c_i)$).

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, $S \subseteq S_{i_j}$, and so, by Condition 3 of Theorem 11 we have $|S| \geq 3$. By Lemma 8, $S \subset S_{i_j}$. This implies by Lemma 9 the only possibility is that $T\langle S \rangle$ is a star such that $Real(N_T[c_i]) = S$. However, by Condition 3 of Theorem 11, another minimal separator of G_{i_j} must strictly contain S . By Lemma 9 this implies that in fact, $Real(N_T[c_{i_j}]) = S$, that is a contradiction since $c_{i_j} \neq c_i$. \diamond

Claim 10. *Let T_{i_j} be a 4-Steiner root of G_{i_j} such that $T_{i_j}\langle S_{i_j} \rangle \equiv_G T\langle S_{i_j} \rangle$ 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. *All the vertices in $N_T(c_{i_j})$ are either simplicial in G_{i_j} or adjacent in T to a vertex of W_{i_j} .*
2. *$|S_i \cap N_T[c_{i_j}]| \leq 2$, and in the same way $|S_{i_k} \cap N_T[c_{i_j}]| \leq 2$ for any other child X_{i_k} of X_i .*

Proof. We may assume $|N_T[c_{i_j}]| \geq 3$ (otherwise we are done). Then, $|S_i \cap N_T[c_{i_j}]| \geq 3$ ($|S_{i_k} \cap N_T[c_{i_j}]| \geq 3$, resp.) would imply $Real(N_T[c_{i_j}]) = S_i$ ($Real(N_T[c_{i_j}]) = S_{i_k}$, resp.).

Furthermore by Condition 3 of Theorem 11 any minimal separator of G_{i_j} strictly contained into S_{i_j} must have size at least 3 and be strictly contained into another minimal separator of G_{i_j} . So, the only possibility for such a separator is also $S = Real(N_T[c_{i_j}])$. In particular if such a S exists then, it is T_G -convergent (*i.e.*, Condition 3 of Theorem 11). Therefore, we cannot have $S = S_i$ ($S = S_{i_k}$, resp.), thereby implying we always fall in Case 2 of the claim.

From now on assume such a S does not exist. We prove either $Real(N_T[c_{i_j}]) \neq S_i$ and $Real(N_T[c_{i_j}]) \neq S_{i_k}$ for any child X_{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 9 S_i is weakly T_G -convergent but not T_G -convergent, thereby contradicting Condition 1 of Theorem 11. The proof for S_{i_k} is identical as the one above.

Finally, assume either $Real(N_T[c_{i_j}]) = S_i$ or $Real(N_T[c_{i_j}]) = S_{i_k}$ for some child X_{i_k} (otherwise we are done). Since no minimal separator of G_{i_j} can contain $Real(N_T[c_{i_j}])$, Case 1 of the claim follows from the same proof as for Claim 7 (Case $c = c_{i_j}$). \diamond

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

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

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 $T\langle S_{i_j} \rangle$. By Theorem 12, we so obtain $\mathcal{O}(|S_{i_j}|^3)$ 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 $|Real(N_{T_{i_j}}[c_{i_j}])| \geq 3$ and $|N_{T_{i_j}}[c_{i_j}] \cap S_{i_k}| \leq 2$ for any child X_{i_k} of X_i (otherwise, no additional constraint is needed). We consider all the other minimal separators $S_{i_k} := X_i \cap X_{i_k}$ between X_i and one of its children nodes such that $S_{i_k} \subseteq N_{T_{i_j}}[c_{i_j}]$. In particular, $|S_{i_k}| \leq 2$. There are two possibilities:

- If $S_{i_k} = \{v_{i_k}\}$ then, by Claim 4 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 $S_{i_k} \neq \text{Real}(N_{T_{i_j}}[c_{i_j}])$. Then, the tree $T\langle S_{i_k} \rangle$ 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 11. *Let T_{i_j} satisfy all of the above distances' conditions. Exactly one of the following conditions is 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} 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 $T := T'\langle V \setminus W_{i_j} \rangle \odot T_{i_j}$ is also a 4-Steiner root of G . By Theorem 9, 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 $v \in X_i$. In particular:

$$\text{dist}_T(v, W_{i_j}) = \text{dist}_{T'}(v, W_{i_j}) \geq 5.$$

So, we are left to consider the other minimal separators $S_{i_k} := X_i \cap X_{i_k}$ between X_i and any other node. 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 8). If $S_{i_k} \cap S_{i_j} \subseteq N_T[c_i]$ then, by Claim 9, we have:

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

and in the same way:

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

(*i.e.*, because all other vertices in $N_T(c_i)$ are simplicial in G_{i_j}). In particular:

$$\text{dist}_T(W_{i_j}, W_{i_k}) = \text{dist}_{T'}(W_{i_j}, W_{i_k}) \geq 5.$$

Therefore, 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 10 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' constraints over T_{i_j} that we have $\text{dist}_T(W_{i_k}, W_{i_j}) \geq 5$.

- Otherwise, $S_{i_k} = \{u_{i_k}, v_{i_k}\}$. Recall that $S_{i_k} \subset S_{i_j}$. We may further assume $|Real(N_T[c_{i_j}])| \geq 3$ and $S_i \neq S_{i_k}$ (otherwise, the encoding already includes the distance to W_{i_j} from any node in $T\langle S_{i_k} \rangle$). 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 $|Real(N_T[c_{i_j}])| \geq 3$, c_{i_j} is adjacent to some other real node than v_{i_k} . In other words, $dist_T(c_{i_j}, V \setminus W_{i_k}) = 1$ is minimized. Then, since we have $dist_{T'}(W_{i_k}, W_{i_j}) \geq 5$ we must have $dist_{T'}(c_{i_j}, W_{i_k}) = 4$ and so, $dist_{T'}(v_{i_k}, W_{i_k}) = 3$. It follows from Claim 5 and the distances' constraints over T_{i_j} that we can always assume $dist_T(W_{i_k}, W_{i_j}) \geq 5$.

◇

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

- If $d_{T_{i_j}}(v)$ is included in the encoding then, $d_v := d_{T_{i_j}}(v)$. In particular, this will be the case for c_i, c_{i_j} .
- If $v_{i_k} \in S_{i_k} \cap S_{i_j}$ has a 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 $T\langle S_{i_j} \rangle$ 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 4: The dynamic programming

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

Theorem 15. *Let $G = (V, E)$ be strongly chordal, let T_G be a rooted clique-tree as in Theorem 11 and let $X_i \in \mathcal{K}(G)$. There is some polynomial P such that, after a pre-processing in time $\mathcal{O}(n\|G_i\|^3 P(n))$, we can solve DISTANCE-CONSTRAINED ROOT for any $T\langle S_i \rangle, (d_v)_{v \in T\langle S_i \rangle}$ in time $\mathcal{O}(P(n))$.*

Theorem 15 proves Theorem 2 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 $P(n) = \mathcal{O}(n^8)$.

Proof of Theorem 15. If X_i is a leaf of T_G then, this follows from Corollary 2. Thus from now on, assume X_i is an internal node with children $X_{i_1}, X_{i_2}, \dots, X_{i_p}$.

Preprocessing. Let $\mathcal{T}_{i_1}, \mathcal{T}_{i_2}, \dots, \mathcal{T}_{i_p}$ be as in Theorem 14. By induction on T_G , the computation of all the \mathcal{T}_{i_j} 's requires total preprocessing time $\sum_{j=1}^p \mathcal{O}(n\|G_{i_j}\|^3 P(n))$, and $\sum_{j=1}^p \mathcal{O}(|S_{i_j}|^3 P(n))$ additional time. We also need to construct the family \mathcal{F}_i of Proposition 1, that takes $\mathcal{O}(n|X_i|^6)$ -time. Note that the elements in \mathcal{F}_i are of the form (T_{Y_i}, \mathcal{C}_i) where $Y_i \subseteq X_i$ and \mathcal{C}_i must represent the center of $T\langle X_i \rangle$ (missing vertices of $X_i \setminus Y_i$ are supposed to be located in thin branches, see Lemma 14). Overall, if we assume w.l.o.g. that $P(n) = \Omega(n^3)$ then, this pre-processing phase takes

total time:

$$\begin{aligned}
& \sum_{j=1}^p \mathcal{O}(n \|G_{i_j}\|^3 P(n)) + \sum_{j=1}^p \mathcal{O}(|S_{i_j}|^3 P(n)) + \mathcal{O}(n |X_i|^6) \\
&= \mathcal{O}(nP(n)) \cdot \sum_{j=1}^p \|G_{i_j}\|^3 + \mathcal{O}(p |X_i|^3 P(n)) + \mathcal{O}(n |X_i|^3 P(n)) \\
&= \mathcal{O}(nP(n)) \cdot (\|G_i\|^3 - |X_i|^3) + \mathcal{O}(n |X_i|^3 P(n)) \\
&= \mathcal{O}(nP(n) \|G_i\|^3).
\end{aligned}$$

Answering a query. In what follows let $T\langle S_i \rangle$ and $(d_v)_{v \in T\langle S_i \rangle}$ 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 $T\langle S_i \rangle \equiv_G T_{Y_i}\langle S_i \rangle$. This takes total time $\mathcal{O}(|S_i| |\mathcal{F}_i|) = \mathcal{O}(n |X_i|^7)$. 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 $Y_i = X_i$ (no thin branch). For every $v \in T\langle S_i \rangle$ we check whether we have:

$$dist_{T_{Y_i}}(v, X_i \setminus S_i) \geq d_v$$

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

$$dist_{T_{Y_i}}(v, u) + d_{T_{i_j}}(u) \geq d_v.$$

Overall, since $|\mathcal{T}_{i_j}| = \mathcal{O}(|S_{i_j}|^3) = \mathcal{O}(|X_i|^3)$, this new verification phase takes total time $\mathcal{O}(p |S_i| |X_i|^3) = \mathcal{O}(n |X_i|^4)$. 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 by Theorem 14.

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 $u \in T\langle S_{i_j} \rangle$, we assign some value $r_{i_j}(u)$ that intuitively represents the distance of u 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 $d_{T_{i_j}}(u) + r_{i_j}(u) \leq 4$. We set initially $r_{i_j}(u) := dist_{T_{Y_i}}(u, X_i \setminus S_{i_j})$ and we apply the rule. Overall, updating (initializing, resp.) the values r_{i_j} for every j takes time $\sum_{j=1}^p \mathcal{O}(|S_{i_j}|) = \mathcal{O}(n |X_i|)$. Applying the rule takes time $\sum_{j=1}^p \mathcal{O}(|S_{i_j}| |\mathcal{T}_{i_j}|) = \sum_{j=1}^p \mathcal{O}(|S_{i_j}|^4) = \mathcal{O}(n |X_i|^4)$. 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 4 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 $u \in T\langle S_{i_k} \rangle$ (possibly, $u = v$) we set $r_{i_k}(u) := \min\{r_{i_k}(u), dist_{T_i}(u, v) + d_{T_{i_j}}(v)\}$. 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 $T\langle S_{i_j} \rangle$ 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 12. *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, $T\langle V \setminus W_{i_j} \rangle \odot T_{i_j}^v$ 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 $d_{T_{i_j}^v}(v)$ 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 $\text{dist}_{T'}(u, V \setminus V_{i_j}) = \text{dist}_T(u, V \setminus V_{i_j})$ 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 $d_{T_{i_j}^v}(u)$ is maximized, $d_{T_{i_j}^v}(u) \geq \text{dist}_T(u, W_{i_j}) \geq d_{T_{i_j}^v}(v)$. \diamond

By Claim 12 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 5 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 v has no real neighbour in $T_i \setminus S_{i_j}$, $T\langle S_{i_k} \rangle$ must be a bistar. We divide our analysis in several subcases:

- Subcase $S_{i_j} = \mathcal{C}(T\langle S_{i_k} \rangle)$. By Claim 2, 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 9, 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 $\mathcal{C}(T_{i_k}\langle X_{i_k}\rangle) = \{u\}$ (this can only be true for at most one index i_k). Note that v is the only leaf of $T\langle S_{i_k}\rangle$ that is adjacent to u . Therefore, by Lemma 8, v is the only real neighbour of u in any $T_{i_k} \in \mathcal{T}_{i_k}$. 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 $d_{T_{i_j}}(u) \geq 3$. 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 $T\langle S_{i_j}\rangle$ is a bistar. Let $\mathcal{C}_i = \{c_i\}$ and let $\mathcal{C}(T\langle S_{i_j}\rangle) = \{c_i, c_{i_j}\}$. We keep only the solutions $T_{i_j} \in \mathcal{T}_{i_j}$ such that $\mathcal{C}(T_{i_j}\langle X_{i_j}\rangle) = \{c_{i_j}\}$. Then, we have by Claim 9 $d_{T_{i_j}}(v) = d_{T_{i_j}}(c_i) + 1$ for any $v \in N_{T_i}(c_i) \setminus \{c_{i_j}\}$. So, we would like to pick $T_{i_j} \in \mathcal{T}_{i_j}$ that maximizes $d_{T_{i_j}}(c_i)$. The only case where we cannot do that w.l.o.g. is when there exists a minimal separator $S_{i_k} \subseteq N_{T_i}[c_{i_j}]$. 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 10 and its proof, this implies that there is only one possibility for $d_{T_{i_j}}(v)$, for every $v \in T\langle S_{i_j}\rangle$. Specifically (Case 1 of the claim), $d_{T_{i_j}}(c_{i_j}) = 2$, and for every $u \in N_{T\langle S_{i_j}\rangle}(c_{i_j})$ 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 $u \in T\langle S_{i_k}\rangle$ we update $r_{i_k}(u)$ 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 $T\langle S_{i_j}\rangle$ is a star. Let $\mathcal{C}(T\langle S_{i_j}\rangle) = \{c\}$. 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 13. *The following properties are true for any $T_{i_j} \in \mathcal{T}_{i_j}$:*

1. $\text{diam}(T_{i_j}\langle X_{i_j}\rangle) = 4$;
2. *and the unique center node $v_j \in \mathcal{C}(T_{i_j}\langle X_{i_j}\rangle)$ is either in $S_{i_j} \setminus \mathcal{C}_i$, or it is a Steiner node. Moreover:*
 - (a) *if $v \in S_{i_j}$ then, v is a leaf of T_{Y_i} ;*
 - (b) *every vertex of $S_{i_j} \setminus \{v\}$ must be simplicial in G_{i_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. If $\text{diam}(T_{i_j}\langle X_{i_j}\rangle) < 4$ then, the only possibility is $\text{diam}(T_{i_j}\langle X_{i_j}\rangle) = 3$, and so, $c \in \mathcal{C}(T_{i_j}\langle X_{i_j}\rangle)$. However, the latter would contradict Lemma 4 as we already assume $c \in \mathcal{C}_i$. Therefore, $\text{diam}(T_{i_j}\langle X_{i_j}\rangle) = 4$, thereby implying $\mathcal{C}(T_{i_j}\langle X_{i_j}\rangle) = \{v_j\}$ for some v_j . By Lemma 12, $\text{Real}(N_{T_{i_j}}[c]) = S$. Thus, either v_j is Steiner, or $v_j \in S$. Furthermore if $v_j \in S$ then, $v_j \in S \setminus \mathcal{C}_i$ (otherwise, this would contradict Lemma 4).

Finally, we prove that every vertex of $S_{i_j} \setminus \{v\}$ must be simplicial in G_{i_j} . Indeed, we recall that by Lemma 12 we have $\text{Real}(N_{T_{i_j}}[c]) = S$. By Claim 6, no minimal separator of G_{i_j}

can be strictly contained into S_{i_j} . So, in order for having vertices of $S_{i_j} \setminus \{v\}$ that are not simplicial in G_{i_j} , there should be a minimal separator S' of G_{i_j} that contains S_{i_j} . But then, by Condition 3 of Theorem 11 we could assume at least one such S' strictly contains S_{i_j} . This would imply $T_{i_j} \langle S' \rangle$ would be a bistar, and so, by Lemma 9 c would also be in $\mathcal{C}(T_{i_j} \langle X \rangle)$ for some maximal clique X in G_{i_j} . A contradiction. \diamond

We observe that no such a v_j as above in the claim can be contained into another $S_{i_k} \neq S_{i_j}$ where $T \langle S_{i_k} \rangle$ is a star. This implies that w.r.t. every such a separator S_{i_k} , any solution T_{i_j} that maximizes $d_{T_{i_j}}(\mathcal{C}(T_{i_j} \langle X_{i_j} \rangle))$ 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 $S_{i_k} = S_{i_j}$.

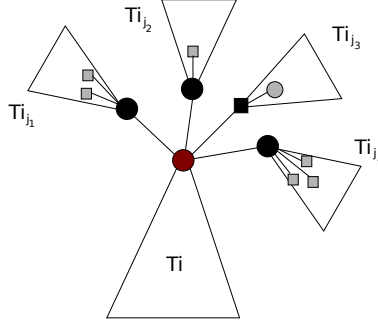


Figure 26: An illustration of Phase 4.

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 $T_{i_{j'}} \langle X_{i_{j'}} \rangle$ 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. 26 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 $Steiner[J] := \{\alpha_{j'} \mid j' \in J\}$ be a set of Steiner nodes. We construct a bipartite graph $Bip(S_{i_j})$ with respective sides J and $(S_{i_j} \setminus \mathcal{C}_i) \cup Steiner[J]$.
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 $\mathcal{C}(T_{i_{j'}} \langle X_{i_{j'}} \rangle) = \{v\}$. 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 $\mathcal{C}(T_{i_{j'}} \langle X_{i_{j'}} \rangle)$ is Steiner. Furthermore if such a $T_{i_{j'}}$ exists then, we choose one maximizing $d_{T_{i_{j'}}}(\mathcal{C}(T_{i_{j'}} \langle X_{i_{j'}} \rangle))$ and we assign the weight $d_{T_{i_{j'}}}(\mathcal{C}(T_{i_{j'}} \langle X_{i_{j'}} \rangle))$ 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 . 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 $T\langle S_{i_j} \rangle$ and any other $T\langle S_{i_k} \rangle$ 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 9), and so, T_G -convergent (Condition 1 of Theorem 11), 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 $\mathcal{C}(T_{i_{j'}}\langle X_{i_{j'}} \rangle)$. 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 $c \in \mathcal{C}(T_{i_{j_0}}\langle X_{i_{j_0}} \rangle)$ (if any). Then, we choose a corresponding solution in $\mathcal{T}_{i_{j_0}}$ among $\mathcal{O}(|S_{i_{j_0}}|^3) = \mathcal{O}(|X_i|^3)$ possibilities. Overall, there are $\mathcal{O}(n|X_i|^3)$ possibilities. We test each such a possibility sequentially.
2. By Claim 7, the following property holds for any $v \in S \setminus \{c\}$: 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 latter case, we discard all solutions $T_{i_{j'}} \in \mathcal{T}_{i_{j'}}$, $j' \in J \setminus \{j_0\}$ such that $\mathcal{C}(T_{i_{j'}}\langle X_{i_{j'}} \rangle) = \{v\}$.
3. Finally, we observe that for every $j' \in J \setminus \{j_0\}$, we will always obtain $d_{T_{i_{j'}}}(c_i) = 2 + d_{T_{i_{j'}}}(\mathcal{C}(T_{i_{j'}}\langle X_{i_{j'}} \rangle)) \in \{3, 4\}$. Since we aim at maximizing $d_{T_{i_{j'}}}(c_i)$, we can apply our reduction to MAXIMUM-WEIGHT MATCHING in order to pick the solutions $T_{i_{j'}} \in \mathcal{T}_{i_{j'}}$ for every $j' \in J \setminus \{j_0\}$.
4. Overall, among all the valid solutions computed, 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 $Y_i \neq X_i$ (there are thin branches). Then, $\mathcal{C}_i = \{c_i\}$. We consider all the minimal separators $S_{j_1}, S_{j_2}, \dots, S_{j_q} \subseteq (X_i \setminus Y_i) \cup \{c_i\}$ 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 $\mathcal{O}(|S_{j_\ell}|) = \mathcal{O}(|X_i|)$ possibilities. Furthermore, since according to Definition 4 there can be no minimal separator S_{i_k} which intersects both S_{j_ℓ} and $X_i \setminus S_{j_\ell}$, 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 (but we do not apply the filtering rules).

3. We end up applying the same procedure as for Subphase 4.b.. 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 ROOT 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.

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, ℓ) -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⁺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.