

# On the star forest polytope for 4-cactus graphs

Nguyen Thanh Loan<sup>1</sup>, Nguyen Viet Hung<sup>1</sup> Minh Hieu Nguyen<sup>2,3</sup>, and  
Thi Viet Thanh Vu<sup>4</sup>

<sup>1</sup> INP Clermont Auvergne, Univ Clermont Auvergne, Mines Saint-Etienne, CNRS,  
UMR 6158 LIMOS, 1 Rue de la Chebarde, 63178 Aubiere Cedex, France  
`thanh_loan.nguyen@doctorant.uca.fr`

Corresponding author: `viet_hung.nguyen@uca.fr`

<sup>2</sup> Unité de Mathématiques Appliquées, ENSTA, Institut Polytechnique de Paris,  
91120 Palaiseau, France

<sup>3</sup> CEDRIC-Cnam, 292 rue Saint Martin, F-75141 Paris Cedex 03, France  
`minh-hieu.nguyen@ensta-paris.com`

<sup>4</sup> Hanoi Open University, Viet Nam  
`thanhvtv82@gmail.com`

**Abstract.** This paper investigates the polyhedral structure of the Maximum Weight Star Forest Problem (MWSFP) in an undirected weighted graph  $G = (V, E)$ , where each edge has a non-negative weight. A star in  $G$  is either an isolated node or a connected subgraph in which all edges share a common endpoint, and a star forest is a collection of disjoint stars. The objective of the MWSFP is to find a star forest with the maximum total edge weight. This problem is NP-hard in general but can be solved in polynomial time when  $G$  is a cactus graph [19].

In this paper, we provide a complete polyhedral description of the star forest polytope  $SFP(G)$  when  $G$  is a 4-cactus graph, a subclass of cactus graphs where each cycle has at most four edges. More precisely, we introduce a new class of facet-defining inequalities, called  $M$ -cactus inequalities, which hold for any graph. We then show that when  $G$  is a 4-cactus graph, the  $M$ -cactus inequalities, together with the non-negativity inequalities, completely describe  $SFP(G)$ .

**Keywords:** Graph theory · Spanning star forest · Polyhedral combinatorics · Cactus graph

## 1 Introduction

The Maximum Weight Star Forest Problem (MWSFP) arises in graph theory and optimization, where the objective is to find a collection of disjoint stars in an undirected graph  $G = (V, E)$  that maximizes the total weight of the included edges. A *star* in  $G$  is defined as either an isolated node or a subgraph where all edges share a common endpoint known as the *center*. A node adjacent to the center is called a *leaf*. A *star forest* refers to a collection of disjoint stars in  $G$ . The weight of a star forest is the sum of the weights of its edges. Given a weight vector  $c \in \mathbb{R}_+^{|E|}$ , the MWSFP seeks to find a star forest in the graph  $G$  with the

maximum possible weight, i.e.,  $\max \sum_{e \in F} c_e$  subject to  $F \in \mathcal{F}(G)$ , where  $\mathcal{F}(G)$  is the set of star forests in  $G$ . Introduced by Nguyen et al. [18], the problem is NP-hard since it can be reduced to the minimum dominating set problem in the unweighted case. The MWSFP has several applications. In computational biology, it is used for aligning genomic sequences with high duplication, improving methods such as the threaded blockset aligner program [16, 18], and for comparing phylogenetic trees [9]. In the automobile industry, a directed version of the problem helps optimize configuration management in inclusion relationships, where the objective is to find a maximum-weight spanning subgraph in which each connected component forms an outward star [1].

Previous research on the MWSFP has primarily focused on two directions. The first involves developing approximation algorithms for the unweighted version and its generalizations. Nguyen et al. [18] introduced a 0.6-approximation algorithm, later improved to 0.803 by Athanassopoulos et al. [3] using its connection to the complementary set cover problem. Stronger inapproximability results and APX-hardness proofs have also been established for node-weighted and edge-weighted versions [18, 11, 14, 10]. The second direction focuses on efficient algorithms for special graph classes in the weighted version. Nguyen et al. [18] developed a linear time algorithm for solving the MWSFP on trees, leading to a  $\frac{1}{2}$ -approximation algorithm for general graphs. More recently, Nguyen extended this approach to cactus graphs by introducing a linear time algorithm for this class [19].

Despite significant progress in algorithmic studies of the MWSFP, its polyhedral structure remains largely unexplored. Given a star forest  $F$  in a graph  $G$ , let  $x_F$  be its incidence vector in  $\mathbb{R}^{|E|}$ . The star forest polytope of  $G$ , denoted  $SFP(G)$ , is the convex hull of the incidence vectors of all star forests in  $G$ . To the best of our knowledge, the only existing work on  $SFP(G)$  is by Aider et al. [2], who provided a complete description of this polytope when  $G$  is a tree or a cycle. However, extending these results to more general graph classes remains an open problem. Motivated by this issue, this paper focuses on the polyhedral study of the MWSFP by introducing a new class of facet-defining inequalities for the star forest polytope in cactus graphs, which generalizes both trees and cycles. Formally, a cactus graph is a graph where each edge belongs to at most one cycle. These graphs are relevant in applications such as telecommunication network design in sparsely populated areas [17] and material handling systems using automated guided vehicles [15]. Using an extended formulation inspired by the work of Baiou and Barahona on the Uncapacitated Facility Location Problem [8], we provide a complete description of the star forest polytope for 4-cactus graphs, a subclass of cactus graphs in which each cycle has at most four edges. Furthermore, we demonstrate that the facet-defining inequalities for  $SFP(G)$  when  $G$  is a 4-cactus can be extended to facet-defining inequalities for  $SFP(G)$  in the case of an arbitrary graph.

The paper is structured as follows. In Section 2, we introduce a class of facet-defining inequalities, called the  $M$ -cactus inequalities for  $SFP(G)$  in the general case of  $G$ . Section 3 presents how the  $M$ -cactus inequalities combined

with non-trivial inequalities, provide a complete polyhedral description when  $G$  is a 4-cactus graph. Finally, in Section 4, we summarize our work and discuss the directions for future research.

In the rest of this section, we introduce some useful definitions and notations used throughout this paper. Given an undirected graph  $G = (V, E)$ , a node  $u$  is a *pendant* node if it has degree 1, and a *non-pendant* node otherwise. A *2-matching*  $S$  in  $G$  is a subgraph where each component is either an edge or two edges sharing a common endpoint. For a subgraph  $M$  of  $G$  and an edge  $e \in G \setminus M$ , we denote  $M + e$  as the graph obtained by adding the edge  $e$  to  $M$ , and  $M - e$  as the graph obtained by removing the edge  $e$  from  $M$ . For  $x \in \mathbb{R}^{|E|}$  and  $F \subseteq E$ , let  $x(F) = \sum_{e \in F} x(e)$ . A cycle  $C = \{u_1, u_2, \dots, u_p\}$  in a graph of length  $p \geq 3$  is a sequence of distinct vertices such that each pair  $u_i, u_{i+1}$  for  $i = 1, \dots, p-1$  and  $(u_p, u_1)$  forms an edge in the graph. A 3-cycle (triangle) is a cycle of length 3, and a 4-cycle is a cycle of length 4. We denote by  $\mathcal{C}_3$  and  $\mathcal{C}_4$  the sets of all 3-cycles and 4-cycles in  $G$ , respectively.

## 2 Facet-defining inequalities for $SFP(G)$ : $M$ -cactus inequalities

This section discusses the general case where  $G$  is an arbitrary connected undirected graph. Aider et al. [2] introduced  $M$ -tree inequalities and proved that, together with non-negativity inequalities, they provide a complete description of  $SFP(G)$  when  $G$  is a tree. Inspired by this result, we introduce a more general class of inequalities, called  $M$ -cactus inequalities, to extend this characterization beyond trees. Furthermore, we show that these inequalities define facets of  $SFP(G)$  for general graphs. To begin, we define the concept of a  $M$ -cactus.

**Definition 1.** (*M-cactus*) A  $M$ -cactus is a cactus graph where

- Every node in a 3-cycle or 4-cycle is adjacent to at most one pendant node.
- Every remaining non-pendant node is adjacent to exactly one pendant node.

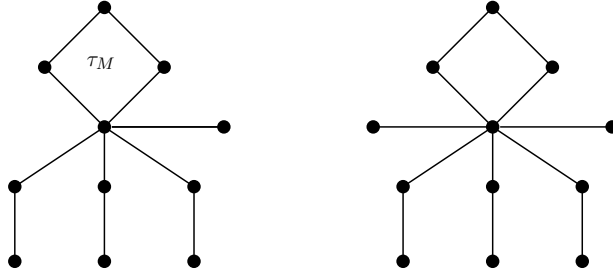


Fig. 1: A  $M$ -cactus  $\tau_M$  (left) and a non  $M$ -cactus (right).

Observe that, a  $M$ -tree is a special case of a  $M$ -cactus that does not contain any cycles. For a  $M$ -cactus  $\tau_M$ , we define an  $M$ -2-matching  $M$  as a cactus graph whose node set is exactly that of  $\tau_M$ , including all its pendant nodes. Figure 1 illustrates a  $M$ -cactus  $\tau_M$  that is not a  $M$ -tree, as well as an example of a non- $M$ -cactus. Figure 2 illustrates the corresponding  $M$ -2-matching  $M$  of  $\tau_M$ .

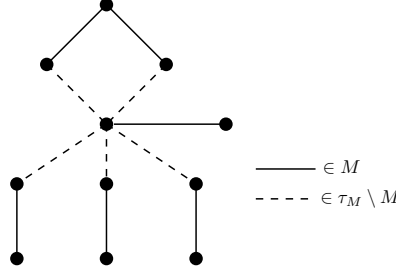


Fig. 2: The  $M$ -2-matching  $M$  corresponds to  $\tau_M$ .

**Definition 2.** (*M-cactus inequality*) Given a  $M$ -cactus  $\tau_M$ , let  $M$  be the corresponding  $M$ -2-matching. We define the  $M$ -cactus inequalities as follows:

$$x(\tau_M) \leq |M|. \quad (1)$$

Before proving the validity of the  $M$ -cactus inequality, we first recall an important property of  $SFP(G)$  that is given by Aider et al. [2].

**Lemma 1.** [2] For any graph  $G$ ,  $SFP(G)$  is a full dimensional polytope.

**Theorem 1.** The  $M$ -cactus inequalities are valid for  $SFP(G)$ .

*Proof.* Assume that  $\tau_M$  is a  $M$ -cactus and  $M$  is a  $M$ -2-matching of  $\tau_M$ . To prove validity, we show that

$$|F \cap \tau_M| \leq |M|,$$

where  $F$  is any star forest of  $G$ . If  $F$  only contains the edges in  $M$ , i.e.,  $M \subseteq F$ , then  $|F \cap \tau_M| \leq |M|$ . If  $(\tau_M \setminus M) \subseteq F$ , then, by the definition of a  $M$ -cactus graph we obtain

$$|F \cap \tau_M| \leq |\tau_M \setminus M| < |M|.$$

The remaining case is where  $F$  contains edges in both  $M$  and  $\tau_M \setminus M$ . If  $F \cap \tau_M$  is a 2-matching, then  $|F \cap \tau_M| \leq |M|$  because  $M$  is a maximum 2-matching in  $\tau_M$ . Therefore, we now examine the case where  $F \cap \tau_M$  contains a star  $S$  with center  $u$ , such that  $S$  includes at least two edges from  $F$ : one edge  $uu_1$  in  $M$  and another edge  $uu_2$  in  $\tau_M \setminus M$ . There exists an edge  $u_2u_3 \in M$ , where  $u_3$  may coincide with  $u_1$ , because  $M$  covers all vertices of  $\tau_M$ . Since  $uu_1$  and  $uu_2$  already belong to  $F$ , it follows that  $u_2u_3 \notin F$ , that is,  $u_2u_3 \in M \setminus F$ . By the definition of a  $M$ -cactus,  $u_3$  must either have degree 1 in  $\tau_M$  or belong to the

cycle  $\{u_1, u, u_2, u_3\} \in \mathcal{C}_4$ , with  $u_3u_1 \notin F$ . This ensures that for each edge in  $F \cap (\tau_M \setminus M)$ , there is a corresponding edge in  $M \setminus F$ . Therefore,

$$|F \cap (\tau_M \setminus M)| \leq |M \setminus F|.$$

Combining these, we get

$$|(F \cap (\tau_M \setminus M)) \cup (F \cap M)| \leq |(M \setminus F) \cup (F \cap M)| = |M|.$$

Since  $M \subseteq \tau_M$ , we get that  $|(F \cap (\tau_M \setminus M)) \cup (F \cap M)| = |F \cap \tau_M|$ . As a result, we have  $|F \cap \tau_M| \leq |M|$ . Thus, the  $M$ -cactus inequalities are valid.  $\square$

**Theorem 2.** *The  $M$ -cactus inequalities define facets of  $SFP(G)$ .*

*Proof.* Suppose there exists a facet-defining inequality

$$\alpha^t x \leq \beta \tag{2}$$

for  $SFP(G)$  such that all star forests satisfying (1) at equality also satisfy (2) at equality. As stated in Lemma (1), since  $SFP(G)$  is full dimensional, it suffices to demonstrate that Inequality (2) must be a positive multiple of the  $M$ -cactus inequality (1).

Let  $\tau$  be any  $M$ -cactus and  $M$  be the  $M$ -2-matching associated with  $\tau$ . Let us examine the corresponding  $M$ -cactus inequality

$$x(\tau) \leq |M|. \tag{3}$$

Observe that  $M$  is a star forest satisfying (1) at equality. Thus,  $M$  also satisfies (2) at equality. This implies that

$$\alpha(M) = \beta. \tag{4}$$

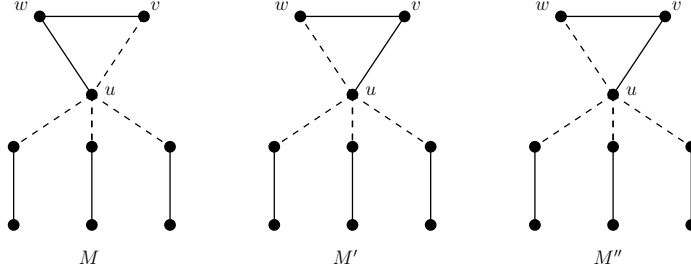
• We first show that  $\alpha(e) = \alpha(e')$  for all  $e, e' \in \tau$ . Indeed, let  $uv$  be an edge in  $\tau \setminus M$ . We consider the following cases.

- *Case 1.*  $uv$  belongs to a 3-cycle  $\{u, v, w\} \in \mathcal{C}_3$  with  $uw, vw \in M$ . We consider the star forests  $M' = M - uw + uv$  and  $M'' = M - vw + uv$  (see figure 3). Observe that  $M, M', M''$  are star forests that satisfy (3) at equality. Thus, we obtain

$$\alpha(M) = \alpha(M - uw + uv) = \alpha(M - vw + uv).$$

It follows that

$$\alpha(uv) = \alpha(uw) = \alpha(vw).$$

Fig. 3: An example of the star forests  $M, M', M''$ .

- *Case 2.*  $uv$  belongs to a 4-cycle  $\{u, v, v', u'\} \in \mathcal{C}_4$ . We consider two subcases:
  - *Subcase 2.1.* If  $uu'$  and  $vv'$  are in  $M$ , then  $M, M - uu' - vv' + uv + v'u', M - uu' + uv$ , and  $M - vv' + uv$  are the star forests satisfying (3) at equality, which implies that

$$\alpha(uv) = \alpha(uu') = \alpha(vv') = \alpha(u'v').$$

- *Subcase 2.2.* If  $uu'$  and  $u'v'$  are in  $M$ , then using a similar argument, considering the star forests  $M, M - uu' - u'v' + uv + vv', M - uu' + uv$ , and  $M - u'v' + uv$ , we conclude that

$$\alpha(uv) = \alpha(uu') = \alpha(u'v') = \alpha(vv').$$

- *Case 3.*  $uv$  does not belong to any cycle  $C \in \{\mathcal{C}_3, \mathcal{C}_4\}$ . Let  $uu'$  and  $vv'$  be edges in  $M$  that are incident to  $u$  and  $v$ , respectively. Consider the star forests  $M' = M - uu' + uv$  and  $M'' = M - vv' + uv$ . Since both  $M'$  and  $M''$  satisfy (3) at equality, it follows

$$\alpha(uv) = \alpha(uu') = \alpha(vv').$$

Thus, we have established that all edges in  $\tau$  have the same coefficient, i.e.,  $\alpha_{e'} = \alpha_e$  for all  $e', e \in \tau$ .

- We now show that  $\alpha(e) = 0$  for all  $e \notin \tau$ . Considering an edge  $uv$  belongs to  $\tau$ , there are two possible cases.

- *Case 1.* Neither  $u$  nor  $v$  belongs to any cycle  $C \in \{\mathcal{C}_3, \mathcal{C}_4\}$ . Since  $M$  and  $M + uv$  are both star forests satisfying (3) at equality, it follows that

$$\alpha(M) = \alpha(M + uv) = \beta.$$

This implies that  $\alpha(uv) = 0$ .

- *Case 2.* Exactly one endpoint of the edge  $uv$  belongs to a cycle  $C \in \{\mathcal{C}_3, \mathcal{C}_4\}$ . Suppose that this endpoint is  $u \in C$ . If  $u \notin M$ , then  $M + uv$  remains a star forest satisfying (3), and we obtain  $\alpha(uv) = 0$ . Therefore, in the rest of this case, we only consider the scenario where  $u \in M$ . Since  $C \in \{\mathcal{C}_3, \mathcal{C}_4\}$ ,

it follows that there are exactly two edges of  $C$ , denoted  $e_1$  and  $e_2$  belong to  $M$ . If three edges  $uv, e_1, e_2$  share a common endpoint  $u$ , then  $M + uv$  is also a star forest satisfying (3) at equality, and we also obtain  $\alpha(uv) = 0$ . Otherwise, without loss of generality, we assume that the edge sharing the common endpoint  $u$  with  $uv$  and  $e_1$  is  $e_3 \in C \setminus M$ . In this case  $M + uv + e_3 - e_2$  remains a star forest satisfying (3) at equality. Note that  $\alpha(e_2) = \alpha(e_3)$  since  $e_2, e_3$  are two edges in  $\tau$ . Consequently, we obtain  $\alpha(uv) = 0$ .

From the above arguments, we conclude that  $\alpha(e) = \alpha(e')$  for all  $e, e' \in \tau$  and  $\alpha(e) = 0$  for all  $e \notin \tau$ . Therefore, we obtain  $\alpha^t x \leq \beta$  as a positive multiple of the  $M$ -cactus inequality, which completes the proof.  $\square$

### 3 Complete description of $SFP(G)$ in 4-cactus

In this section, we focus on a special class of cactus graphs called 4-cactus graphs and provide a characterization of  $SFP(G)$  for 4-cactus graphs. We begin by formally defining this concept.

**Definition 3.** A 4-cactus graph is a cactus graph in which every cycle has length at most 4, meaning that every cycle is either a 3-cycle or a 4-cycle.

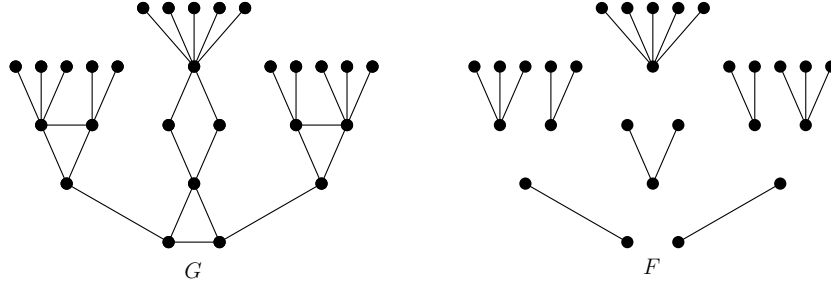


Fig. 4: A 4-cactus graph  $G$  (left) and a star forest  $F$  of  $G$  (right).

For the remainder of this section, we consider  $G$  as a 4-cactus graph. Consequently, any  $M$ -cactus graph in  $G$  is also a 4-cactus graph. With this setting, we now describe the size of maximal star forests in  $M$ -cactus in the following lemma.

**Lemma 2.** Every maximal star forest with respect to a  $M$ -cactus  $\tau_M$  has cardinality  $|M|$ , where  $M$  is the  $M$ -2-matching in  $\tau_M$ .

*Proof.* Let  $H$  be any maximal star forest with respect to  $\tau_M$ . By the validity of  $M$ -cactus inequalities (1), we obtain  $|H| \leq |M|$ . It suffices to show that  $|H| \geq |M|$ . We demonstrate this by proving that each edge in  $M$  corresponds to one edge in  $H$ . Since  $M$  is a 2-matching, each of its components is either (1) two edges sharing a common endpoint or (2) an independent edge.

*Type 1. The component in  $M$  that is an independent edge.* Consider an independent edge  $v_1v_2 \in M$  where  $v_1$  is a pendant node in  $\tau_M$  and  $v_2$  is a non-pendant node. We examine two possibilities:

- If  $v_1$  is a node in  $H$ , then  $v_1v_2$  must also belong to  $H$ . Thus, it corresponds to itself.
- If  $v_1$  is not a node in  $H$ , then by maximality of  $H$ ,  $v_2$  must be a leaf in some star of  $H$ . Otherwise, we could add  $v_1v_2$  to  $H$ , contradicting its maximality. In this case,  $v_1v_2$  corresponds to the edge incident to  $v_2$  in  $H$ .

*Type 2. The component in  $M$  that are two edges share a common endpoint.* Suppose that two edges  $v_1v_2, v_1v_3$  form a component in  $M$ . By definition of  $\tau_M$ , the nodes  $v_1, v_2, v_3$  must belong to the same cycle  $C \in \{\mathcal{C}_3, \mathcal{C}_4\}$  of  $\tau_M$ . We assume that  $C$  is a 3-cycle (similar arguments apply to the case of 4-cycle). We consider the following possible cases.

- If  $C$  has two edges in  $H$ , then these two edges correspond to  $v_1v_2$  and  $v_1v_3$ .
- If  $C$  has exactly one edge in  $H$ . Assume, without loss of generality, that the edge is  $v_2v_3 \in H$  and that  $v_2$  is a center of a star in  $H$ . This implies that the edges  $v_1v_2, v_2v_3$  are not in  $H$ . Since  $H$  is a maximal star forest,  $v_1$  should be a node in  $H$  (otherwise, adding the edge  $v_1v_2$  to  $H$  would contradict its maximality). Moreover,  $v_1$  must be a leaf (otherwise, we could also add the edge  $v_1v_2$  to  $H$  and  $H$  would remain a star forest). In this case, the two edges  $v_1v_2$  and  $v_1v_3$  in  $M$  correspond to the edge  $v_2v_3$  and the edge  $e'$ , that is incident to  $v_1$  in  $H$  (see Figure 5).

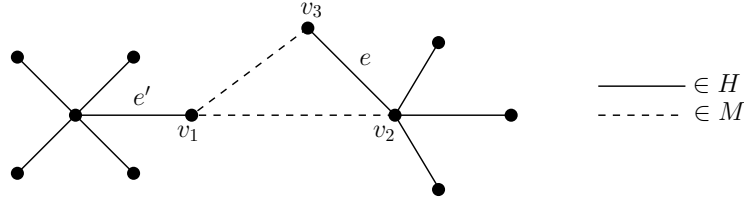


Fig. 5: The edges  $e, e'$  correspond to two edges  $\{v_1v_2, v_1v_3\}$  in  $M$  when  $C$  has exactly one edge in  $H$ .

- If  $C$  has no edges in  $H$ , then by the maximality of  $H$ , at least two nodes of  $C$  must be the nodes in  $H$  since otherwise, we could add the edge connecting those two nodes to  $H$  and  $H$  remains a star forest. We now assume, without loss of generality, that  $v_2$  and  $v_3$  are the nodes in  $H$ . Then  $v_2$  and  $v_3$  should be two leaves belonging to two distinct stars in  $H$ . Indeed, if there is at least one such node—assume  $v_2$  is a center or an isolated node in  $H$ , then we could add  $v_2v_1$  to  $H$  and  $H$  would still be a star forest. This implies that the two edges  $v_1v_2$  and  $v_1v_3$  correspond to the edges  $e, e'$ , which are incident to  $v_2$  and  $v_3$  in  $H$ , respectively (see Figure 6).

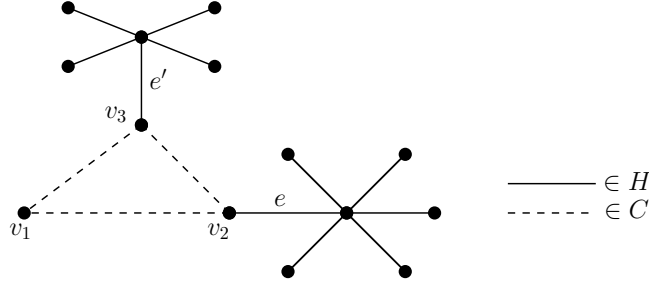


Fig. 6: The edges  $e, e'$  correspond to two edges  $\{v_1v_2, v_1v_3\}$  in  $M$  when  $C$  has no edges in  $H$ .

Conclusion,  $|H| \geq |M|$  and  $|H| \leq |M|$ , we then deduce that  $|H| = |M|$ .  $\square$

To characterize the star forest polytope for 4-cactus graphs, we recall an important concept: *supported graph*. The support graph corresponding to the inequalities  $a^t x \leq b$  is a subgraph of  $G$  containing all the edges with nonzero coefficients in  $a^t x \leq b$ . We now revisit the relationship between facet-defining inequalities and their associated support graphs, as shown in [2]. This relationship holds for any graph and directly applies to 4-cactus graphs.

**Lemma 3.** [2] *Let  $a^t x \leq b$  be a facet-defining inequality for  $SFP(G)$ , and let  $G_a = (V_a, E_a)$  be its corresponding support graph. Then, every tight star forest associated with  $a^T x \leq b$  is maximal with respect to  $E_a$ .*

We now present the main result of this paper in the following theorem.

**Theorem 3.** *Let  $G$  be a 4-cactus graph. Then,  $SFP(G)$  is completely defined by the non-negativity inequalities and the  $M$ -cactus inequalities (1).*

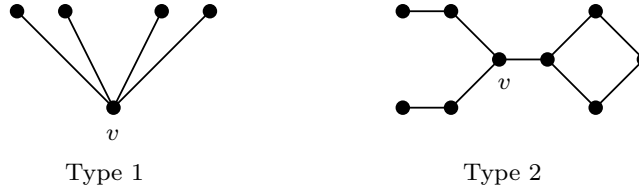
*Proof.* Assume that the inequality  $a^t x \leq b$  defines a facet of  $SFP(G)$  but is neither a non-negativity inequality nor a  $M$ -cactus inequality. Let  $G_a$  be the support graph associated with  $a^t x \leq b$ . We assume without loss of generality that  $G_a$  is a 4-cactus.  $G_a$  is thus a subgraph of  $G$ .

We now consider the case where  $G_a$  is a  $M$ -cactus graph. Consider any tight star forest  $S$  with respect to  $a^t x \leq b$ . By Lemma 3,  $S$  is a maximal subgraph of  $G_a$  satisfying the given inequality. Furthermore, Lemma 2 implies that  $S$  is also tight with respect to the  $M$ -cactus inequalities. This contradicts our assumption that  $a^t x \leq b$  is a facet-defining inequality that is distinct from a  $M$ -cactus inequality. Hence,  $G_a$  cannot be a  $M$ -cactus graph.

Therefore,  $G_a$  must contain at least one non-pendant node  $v$  belonging to one of the following two categories (see Figure 7):

- *Type 1.*  $v$  is adjacent to at least two pendant nodes in  $G_a$ .
- *Type 2.*  $v$  does not belong to any cycle and is not connected to any pendant nodes in  $G_a$ .

We now analyze the two possible cases.

Fig. 7: An illustration of the two types of nodes  $v$ .

• *Case 1. Every non-pendant node in  $G_a$  is of Type 1.* This means that each such node is adjacent to at least two pendant nodes. In this situation, we modify  $G_a$  by keeping only one pendant node for each non-pendant node of Type 1, resulting in a  $M$ -cactus  $\tau_M$ . Let  $M_\tau$  be the  $M$ -2-matching associated with  $\tau_M$ . Then, we have the corresponding  $M$ -cactus inequality  $x(\tau_M) \leq |M_\tau|$ .

Furthermore, observe that for any non-pendant node  $w$  of Type 1, any star forest  $S$  satisfying  $a^t x \leq b$  at equality must be a maximal star forest in  $\tau_M$  and contain either all or none of the pendant nodes connected to  $w$ . Let  $F$  be any tight star forest for  $a^t x \leq b$ . Then,  $F \cap \tau_M$  is a maximal star forest in  $\tau_M$ . By Lemma 2, every maximal star forest in  $\tau_M$  satisfies the  $M$ -cactus inequality at equality. This implies that any star forest satisfying  $a^t x \leq b$  at equality must also satisfy the corresponding  $M$ -cactus inequality at equality, leading to a contradiction.

• *Case 2.  $G_a$  contains at least one non-pendant node of Type 2.* We begin by proving the following claim.

*Remark 1.* There exists a non-pendant node  $p$  of Type 2 such that all the other non-pendant nodes of Type 2 belong to the same connected component obtained by removing  $p$  from  $G_a$ .

*Proof.* We prove this claim constructively. We begin by selecting any non-pendant node  $p_0$  of Type 2 and setting  $i = 0$ . We define the iterative process as follows. At iteration  $i$ , let  $G_1^i, \dots, G_{k_i}^i$  be the subgraphs obtained by removing  $p_i$  from  $G_a$ , and assume without loss of generality that  $G_1^i$  is the 4-cactus containing  $p_0$ . Two cases arise:

- If all remaining non-pendant nodes of Type 2 are in a single 4-cactus  $G_{t_i}$ , then  $p_i$  satisfies the claim, and we stop.
- Otherwise, if Type 2 nodes exist in at least two 4-cactus graphs, assume  $G_{k_i}^i$  is one such component. Choose  $p_{i+1}$  as an arbitrary non-pendant Type 2 node in  $G_{k_i}^i$  and set  $i \leftarrow i + 1$ . This set of nodes forms the kernel.

Since each iteration adds a distinct node to the kernel, the process must terminate in at most  $|V(G_a)|$  iterations, resulting in the required non-pendant node  $p$  of Type 2.  $\square$

Let  $p$  be a non-pendant node of Type 2 that satisfies Remark 1 and let  $G_p$  be the connected component containing the other non-pendant Type 2 nodes,

obtained by removing  $p$  from  $G_a$ . Observe that  $p$  connects to  $G_p$  through at most two edges, and the subgraph  $S$  of  $G_a$  induced by the edges in  $G_a \setminus G_p$  forms a 4-cactus (see Figure 8). It follows that  $S$  is either a  $M$ -cactus or a 4-cactus containing only Type 1 nodes. As demonstrated in Case 1,  $S$  includes a  $M$ -cactus  $\tau_M$ , which includes all non-pendant nodes of  $S$ .

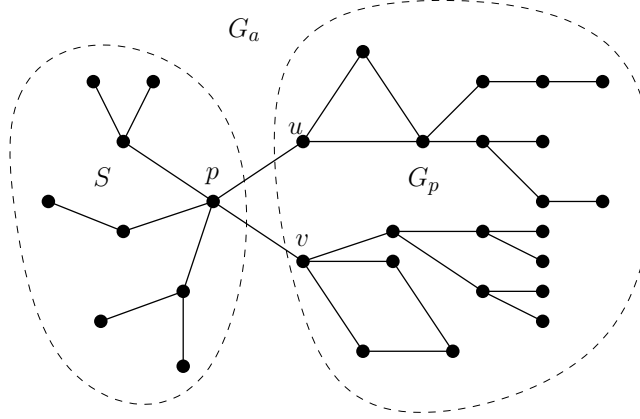


Fig. 8: The two subgraphs  $G_p$  and  $S$  of  $G_a$ .

Let  $u$  and  $v$  be neighbors of  $p$  in  $G_p$ . Note that  $u$  and  $v$  may coincide; however, the subsequent arguments remain unchanged. Thus, the edges  $pu$  and  $p v$  link  $S$  to  $G_p$ . Observe that any maximal star forest in  $G_a$ , whether it contains the edges  $pu$  and  $p v$  or not, must include a maximal star forest in  $S$ . By Case 1, this star forest contains a maximal star forest in  $\tau_M$ . By Lemma 3, any star forest  $F$  satisfying  $a^t x \leq b$  at equality is maximal in  $G_a$ , implying that  $F \cap \tau_M$  is a maximal star forest in  $\tau_M$ . By Lemma 2,  $F \cap \tau_M$  satisfies the  $M$ -cactus inequality associated with  $\tau_M$  at equality. This contradicts the assumption that  $a^t x \leq b$  is facet-defining, which completes the proof.  $\square$

## 4 Conclusions

In this paper, we provide a complete description of the star forest polytope for 4-cactus graphs. Our work extends previous polyhedral descriptions by introducing a new class of facet-defining inequalities. Furthermore, we demonstrate that the  $M$ -cactus inequalities define facets in the general case. Beyond 4-cactus graphs, our results provide useful insights that could aid in developing efficient approximation algorithms for the Maximum Weight Star Forest Problem in general graphs. Future research could leverage this polyhedral framework to improve combinatorial optimization techniques in broader graph structures.

## References

1. A. Agra, D. Cardoso, O. Cerdeira, and E. Rocha. A spanning star forest model for the diversity problem in the automobile industry. In: *ECCO XVIII, Minsk* (2005).
2. M. Aïder, L. Aoudia, M. Baiou, A. R. Mahjoub, and V. H. Nguyen. On the star forest polytope for trees and cycles. *RAIRO - Operations Research* **53** (2019).
3. S. Athanassopoulos, I. Caragiannis, C. Kaklamanis, and M. Kyropoulou. An improved approximation bound for spanning star forest and color saving. In: *Mathematical Foundations of Computer Science (MFCS)*, (2009) 90-101.
4. P. Avella, A. Sassano, and I. Vasilev. Computational study of large-scale p-median problems. *Mathematical Programming* (2007) 89-114.
5. M. Baiou and F. Barahona. On the integrality of some facility location polytopes. *SIAM Journal on Discrete Mathematics* (2009).
6. M. Baiou and F. Barahona. Simple extended formulation for the dominating set polytope via facility location. In: *Tech. Rep. RC25325, IBM Research* (2012).
7. M. Baiou and F. Barahona. Algorithms for minimum weighted dominating sets in cycles and cacti. In: *Tech. Rep. RC25488, IBM Research* (2014).
8. M. Baiou and F. Barahona. The dominating set polytope via facility location. In: *Combinatorial Optimization. ISCO 2014. In Vol. 8596 of Lecture Notes in Computer Sciences*, (2014) 38-49.
9. V. Berry, S. Guillemot, F. Nicolas, and C. Paul. On the approximation of computing evolutionary trees. In: *Proc. of the Eleventh Annual International Computing and Combinatorics Conference. Springer, Berlin, Heidelberg*, (2005) 115-123.
10. D. Chakrabarty and G. Goel. On the approximability of budgeted allocations and improved lower bounds for submodular welfare maximization and GAP. In: *FOCS* (2008) 687-696.
11. N. Chen, R. Engelberg, C. T. Nguyen, P. Raghavendra, A. Rudra, and G. Singh. Improved approximation algorithms for the spanning star forest problem. In: *APPROX/RANDOM In Vol. 4627 of Lecture Notes in Computer Science book series* (2007) 44-58.
12. S. Ferneyhough, R. Haas, D. Hannson, and G. MacGillivray. Star forests, dominating sets, and Ramsey-type problems. *Discrete Mathematics* **245** (2002) 255-262.
13. S. I. Hakimi, J. Mitchem, and E. Schmeichel. Star arboricity of graphs. *Discrete Mathematics* (1996) 93-98.
14. J. He and H. Liang. On variants of the spanning star forest problem. In: *FAW-AAIM* (2011) 70-81.
15. O. Kariv and S. L. Hakimi. An algorithmic approach to network location problems, Part 1: The p-center. *SIAM Journal on Applied Mathematics* (1979) 513-537.
16. W. J. Kent, C. Riemer, L. Elnitski, A. F. Smit, K. M. Roskin, R. Baertsch, K. Rosenbloom, H. Clawson, E. D. Green, D. Haussler, and W. Miller. Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Research* (2004) 708-715.
17. W. L. G. Koontz. Economic evaluation of loop feeder relief alternatives. *Bell System Technical Journal* (1980) 277-281.
18. C. T. Nguyen, J. Shen, M. Hou, L. Sheng, W. Miller, and L. Zhang. Approximating the spanning star forest problem and its applications to genomic sequence alignment. *SIAM Journal on Computing* (2008) 946-962.
19. V. H. Nguyen. The maximum weight spanning star forest problem on cactus graphs. In: *Discrete Mathematics, Algorithms and Applications*, World Scientific Publishing (2015).