Mixed Integer Nonlinear Optimization Models for the Euclidean Steiner Tree Problem in R

Hacène Ouzia, Nelson Maculan

HAL Id: hal-02293105

https://hal.sorbonne-universite.fr/hal-02293105v2

Submitted on 23 Sep 2019

Mixed Integer Nonlinear Optimization Models for the Euclidean Steiner Tree Problem in $\mathbb{R}^d$

Hacene Ouzia*

Sorbonne Universite, CNRS, LIP-6, F-75005, Paris, France

Nelson Maculan

Universidade Federal do Rio de Janeiro, COPPE & IM, 21941-972 Rio de Janeiro, Brazil

Abstract

New mixed integer nonlinear optimization models for the Euclidean Steiner tree problem in d-space (with $d \geq 3$) will be presented in this work. Each model features a non smooth objective function but a convex set of feasible solutions. All these models are theoretically equivalent. From these models, six mixed integer linear and nonlinear relaxations will be considered. Each relaxation has the same set of feasible solutions as the model from which it is derived. Finally, preliminary computational results highlighting the main features of the presented relaxations will be discussed.

Keywords: Integer Programming, Euclidean Steiner tree problem, Steiner tree, Nonlinear optimization models, Mixed integer nonlinear optimization, Relaxation.

1. Introduction

Given n points in $\mathbb{R}^d$ (*called terminals*) the goal in the *Euclidean Steiner Tree Problem* (ESTP) is to find the $\|.\|_2$-shortest tree that spans these points using or not extra points (called *Steiner points*). The length of each edge of the tree is the Euclidean distance between its ends. A very important history of the (ESTP) can be found in [1].

The Euclidean Steiner tree problem is an NP-hard optimization problem (see [13, 12]). It has several applications, to cite a few: phylogenetics (see [16, 3, 2]) and structure and folding proteins (see [19, 17]). In this work, the dimension d will be assumed at least equal to 3 (for the case $d = 2$ see [20, 5, 4]).

As explained in [8], exact approaches to the (ESTP) can be divided into two categories. The approaches in the first category are enumeration based

---

approaches. The first enumeration based approaches are the two algorithms proposed by Gilbert and Pollak in [14] and Smith in [18]. The approaches of the second category are based on a mathematical model of the (ESTP). The first of these mathematical models is the mixed integer nonlinear optimization formulation presented in [15]. From this formulation two other formulations of the (ESTP) can be found in [9] and [10] (see [8] for more details). More recent exact approaches can be found in [11, 6, 9, 10, 7].

In this work, new mixed integer nonlinear models for the Euclidean Steiner tree problem will be presented. All these models are derived from the model given in [15], and are all theoretically equivalent. But, from a computational point of view they are different. Also, all the models feature a non smooth objective function. From these models several mixed integer linear and nonlinear relaxations will be considered. Each relaxation has a smooth objective function which lower bounds the objective function of the model from which it is derived and it has the same set of feasible solutions as the set of feasible solutions of the model from which it is derived. Solving these relaxations using an ad hoc optimization solver leads to a lower and upper bounds of an optimal solution of the Euclidean Steiner tree problem.

The paper is organized as follows. In the next section, the mixed integer nonlinear optimization models of the (ESTP) given in [15] and [10] will be recalled. In section 3, new mixed integer nonlinear optimization models for the (ESTP) will be presented. In section 4, six mixed integer linear and nonlinear relaxations of the (ESTP) will be considered. Preliminary computational results will be presented in section 5. Concluding remarks will be given in the final section.

2. Previous Models

In this section, the mixed integer nonlinear formulations of the Euclidean Steiner tree problem given in [15] and [10] will be recalled.

To simplify the presentation, some notations and definitions are necessary. Let P equal to $\{1, \ldots, n\}$ the index set the terminals $\zeta^1, \zeta^2, \ldots, \zeta^n$ and let X equal to $\{1, \ldots, n-2\}$ the index set of the Steiner points $x^1, x^2, \ldots, x^{n-2}$. The terminals and Steiner points all belong to $\mathbb{R}^d$.

A *Steiner tree* is any spanning tree connecting the n terminals $\zeta^1, \zeta^2, \ldots, \zeta^n$, having at least one Steiner point, and the coordinates of its Steiner points are all known. If the coordinates of the Steiner points of a Steiner tree are not fixed then the resulting tree is called the *topology* of the Steiner tree. A topology is called a *Steiner topology* if the degree of any Steiner point is 3 and the degree of any terminal is at most 3. In the case where the number of Steiner points in a Steiner topology equals $n-2$ then this topology is called a *full Steiner topology*.

2.1. Formulation by Maculan, Michelon and Xavier

In the formulation by Maculan, Michelon and Xavier (see [15]) the Euclidean Steiner tree problem is modeled as follows:

$$\min \sum_{p \in P, q \in X} y_{pq} \|x^q - \zeta^p\|_2 + \sum_{p,q \in X: p < q} z_{pq} \|x^q - x^p\|_2 \tag{1}$$

s.t.

$$\sum_{q \in X} y_{pq} = 1, \; p \in P, \tag{2}$$

$$\sum_{p \in P} y_{pq} + \sum_{p \in X: p < q} z_{pq} + \sum_{p \in X: p > q} z_{qp} = 3, \; q \in X, \tag{3}$$

$$\sum_{p \in X: p < q} z_{pq} = 1, \; q \in X, q > 1, \tag{4}$$

$$\sum_{p \in P} y_{pq} \leq 2, \; q \in X, \tag{5}$$

$$x^p \in \mathbb{R}^d, p \in P, \tag{6}$$

$$y_{pq} \in \{0, 1\}, p \in P, q \in X, \tag{7}$$

$$z_{pq} \in \{0, 1\}, p \in X, q \in X, p < q, \tag{8}$$

where $\|.\|_2$ is the Euclidean norm in $\mathbb{R}^d$. The sum over an empty set is equal, by convention, to zero.

In a feasible solution (y, z) satisfying the constraints (2)-(8) the variable y_{pq} equals 1 if and only if the terminal ζ^p is connected to the Steiner point x^q and the variable z_{pq} equals 1 if and only if the two Steiner points x^p and x^q are connected together. As shown in [15], to any feasible solution (y, z) corresponds a unique Steiner tree having a full topology. Any optimal solution of the optimization problem (1)-(8) is called a *minimal Steiner tree*.

Regarding the constraints (2)-(8): constraints (2) indicate that the degree of a terminal is equal to 1; constraints (3) indicate that the degree of a Steiner point is equal to 3; and constraints (4) eliminate subtours among Steiner points. Finally, constraints (5) are only used to strengthen the model and notice that they are not necessarily valid in the plane.

In the sequel it will be assumed that:

Assumption 1. *For any indexes p and q that belong to P with $p < q$:*

$$\|\zeta^p - \zeta^q\|_\infty \leq \frac{1}{d} \text{ and } \|\zeta^p\|_\infty \leq 1.$$

Notice that this assumption is not restrictive. Indeed, if

$$\max \{\|\zeta^p - \zeta^q\|_\infty : p, q \in P\} = \delta \text{ and } \max \{\|\zeta^p\|_\infty : p \in P\} = \beta,$$

where $\delta > \frac{1}{d}$ and $\beta > 1$, then the assumption will be fulfilled if the coordinates of the terminals are divided by the factor $\delta \beta d$.

2.2. Formulation by Fampa and Maculan

The formulation by Fampa and Maculan (see [9] and [10]) is derived from the formulation by Maculan, Michelon and Xavier (see [15]) as follows.

Using the following substitutions:

$$y_{pq}\|x^q - \zeta^p\|_2 = w_{pq}, p \in P, q \in X,$$
$$z_{pq}\|x^q - x^p\|_2 = v_{pq}, p, q \in X, p < q,$$

in the objective function and adding only the following valid constraints to the set of feasible solutions (2)-(8):

$$w_{pq} \geq \|x^q - \zeta^p\|_2 + y_{pq} - 1, p \in P, q \in X, \tag{9}$$
$$v_{pq} \geq \|x^q - x^p\|_2 + z_{pq} - 1, p, q \in X, p < q, \tag{10}$$
$$w_{pq} \geq 0, p \in P, q \in X, \tag{11}$$
$$v_{pq} \geq 0, p \in X, q \in X, p < q. \tag{12}$$

Indeed, for any indexes p and q that belong to P, if in an optimal solution the variable y_{pq} is equal to 0 then the corresponding constraint in (9) implies that:

$$w_{pq} \geq \|x^q - \zeta^p\|_2 - 1.$$

According to the assumption 1 and the corresponding constraint in (11) it follows that w_{pq} is non negative and since we are minimizing its value in an optimum solution is 0. Now, if y_{pq} is equal to 1 then the corresponding constraint in (9) implies that:

$$w_{pq} \geq \|x^q - \zeta^p\|_2.$$

Again, since we are minimizing then at the optimum this last constraint must be satisfied at equality. The same arguments apply to the other constraints.

Thus, one ends up with the following formulation:

$$\min \sum_{p \in P, q \in X} w_{pq} + \sum_{p \in X, q \in X: p < q} v_{pq}$$
$$s.t.$$
$$(2) - (8) \text{ and } (9) - (12).$$

In the sequel, the same technique will be applied to derive new mixed integer nonlinear optimization models for the Euclidean Steiner tree problem.

3. New Models

In this section, new models for the Euclidean Steiner tree problem will be presented. These models, as stated before, are all equivalent and derived from the model in [15]. In these models, due to the presence of the square root, only the objective function is non convex *et* non smooth. In the next section, several relaxations will be derived from these models.

Remark 2. *In what follows, the fact that the variables y and z are binary will be used frequently without mentioning it explicitly.*

3.1. First Model

Notice that the objective function (1) can be written as follows:

$$\sum_{p\in P, q\in X} \sqrt{\sum_{j=1}^{d} \left[y_{pq}\left(x_j^q - \zeta_j^p \right) \right]^2} + \sum_{p,q\in X:p<q} \sqrt{\sum_{j=1}^{d} \left[z_{pq}\left(x_j^q - x_j^p \right) \right]^2}. \tag{13}$$

Notice that we used the fact that y_{pq} and z_{pq} are equal to y_{pq}^2 and z_{pq}^2, respectively. Using the following substitutions:

$$v_{pq}^j = y_{pq}\left(x_j^q - \zeta_j^p \right), p \in P, q \in X, j \in \{1,\dots,d\}, \tag{14}$$

$$w_{pq}^j = z_{pq}\left(x_j^q - x_j^p \right), p,q \in X, p < q, j \in \{1,\dots,d\}, \tag{15}$$

the objective function (1) reads:

$$\sum_{p\in P, q\in X} \sqrt{\sum_{j=1}^{d} \left(v_{pq}^j \right)^2} + \sum_{p,q\in X:p<q} \sqrt{\sum_{j=1}^{d} \left(w_{pq}^j \right)^2}. \tag{16}$$

The following constraints can be added to strengthen the set (2)-(8):

$$-y_{pq} \le v_{pq}^j \le y_{pq}, p \in P, q \in X, j \in \{1,\dots,d\}, \tag{17}$$

$$-z_{pq} \le w_{pq}^j \le z_{pq}, p,q \in X, p < q, j \in \{1,\dots,d\}, \tag{18}$$

$$v_{pq}^j \ge y_{pq} - 1 + \left(x_j^q - \zeta_j^p \right), p \in P, q \in X, j \in \{1,\dots,d\}, \tag{19}$$

$$v_{pq}^j \le \left(x_j^q - \zeta_j^p \right) + 1 - y_{pq}, p \in P, q \in X, j \in \{1,\dots,d\}, \tag{20}$$

$$w_{pq}^j \ge z_{pq} - 1 + \left(x_j^q - x_j^p \right), p,q \in X, p < q, j \in \{1,\dots,d\}, \tag{21}$$

$$w_{pq}^j \le \left(x_j^q - x_j^p \right) + 1 - z_{pq}, p,q \in X, p < q, j \in \{1,\dots,d\}. \tag{22}$$

So, the resulting model reads:

$$\min \sum_{p\in P, q\in X} \sqrt{\sum_{j=1}^{d} \left(v_{pq}^j \right)^2} + \sum_{p,q\in X:p<q} \sqrt{\sum_{j=1}^{d} \left(w_{pq}^j \right)^2} \tag{M_1}$$

s.t.

$$(2) - (8) \text{ and } (17) - (22).$$

3.2. Second Model

The objective function (1) can be written differently using the following substitutions:

$$v_{pq} = \sum_{j=1}^{d} y_{pq}\left(x_j^q - \zeta_j^p \right)^2, p \in P, q \in X, \tag{23}$$

$$w_{pq} = \sum_{j=1}^{d} z_{pq} \left(x_j^q - x_j^p \right)^2, p, q \in X, p < q. \tag{24}$$

Then, the following constraints can be used instead to strengthen the set (2)-(8).

$$0 \leq v_{pq} \leq y_{pq}, p \in P, q \in X, \tag{25}$$

$$0 \leq w_{pq} \leq z_{pq}, p, q \in X, p < q, \tag{26}$$

$$v_{pq} \geq y_{pq} - 1 + \sum_{j=1}^{d} \left(x_j^q - \zeta_j^p \right)^2, p \in P, q \in X, \tag{27}$$

$$v_{pq} \leq \sum_{j=1}^{d} \left(x_j^q - \zeta_j^p \right)^2, p \in P, q \in X, \tag{28}$$

$$w_{pq} \geq z_{pq} - 1 + \sum_{j=1}^{d} \left(x_j^q - x_j^p \right)^2, p, q \in X, p < q, \tag{29}$$

$$w_{pq} \leq 1 - z_{pq} + \sum_{j=1}^{d} \left(x_j^q - x_j^p \right)^2, p, q \in X, p < q. \tag{30}$$

This leads to the following second model:

$$\min \sum_{p \in P, q \in X} \sqrt{v_{pq}} + \sum_{p, q \in X : p < q} \sqrt{w_{pq}} \tag{M_2}$$

s.t.

$$(2) - (8) \text{ and } (25) - (30).$$

3.3. Third Model

This third model is based on a different expression of the objective function (1). Indeed, it follows using the following substitutions:

$$u_{pq}^j = y_{pq} x_j^q, p \in P, q \in X, j \in \{1, \ldots, d\}, \tag{31}$$

that:

$$\sum_{j=1}^{d} y_{pq} \left(x_j^q - \zeta_j^p \right)^2 = \sum_{j=1}^{d} \left(u_{pq}^j - \zeta_j^p y_{pq} \right)^2. \tag{32}$$

And, using these other substitutions:

$$v_{pq}^j = z_{pq} x_j^p, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{33}$$

$$w_{pq}^j = z_{pq} x_j^q, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{34}$$

it follows that:

$$\sum_{j=1}^{d} z_{pq} \left(x_j^q - x_j^p \right)^2 = \sum_{j=1}^{d} \left(w_{pq}^j \right)^2 - \sum_{j=1}^{d} x_j^p w_{pq}^j - \sum_{j=1}^{d} x_j^q v_{pq}^j + \sum_{j=1}^{d} \left(v_{pq}^j \right)^2. \quad (35)$$

Thus, one can add the following constraints:

$$-y_{pq} \le u_{pq}^j \le y_{pq}, p \in Pq \in X, j \in \{1, \ldots, d\}, \quad (36)$$

$$u_{pq}^j \le x_j^q + 1 - y_{pq}, p \in P, q \in X, j \in \{1, \ldots, d\}, \quad (37)$$

$$u_{pq}^j \ge x_p^q - 1 + y_{pq}, p \in P, q \in X, j \in \{1, \ldots, d\}, \quad (38)$$

and

$$-z_{pq} \le v_{pq}^j \le z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \quad (39)$$

$$v_{pq}^j \le x_j^p + 1 - z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \quad (40)$$

$$v_{pq}^j \ge x_j^p - 1 + z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \quad (41)$$

$$w_{pq}^j \le z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \quad (42)$$

$$w_{pq}^j \ge -z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \quad (43)$$

$$w_{pq}^j \le x_j^q + 1 - z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \quad (44)$$

$$w_{pq}^j \ge x_j^q - 1 + z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \quad (45)$$

to strengthen the set (2)-(8). Then, one ends up with the following third model:

$$\min \sum_{p \in P, q \in X} \sqrt{\sum_{j=1}^{d} \left(u_{pq}^j - \zeta_j^p y_{pq} \right)^2} + \qquad (M_3)$$

$$\sum_{p,q \in X: p < q} \sqrt{\left(\sum_{j=1}^{d} \left(w_{pq}^j \right)^2 - \sum_{j=1}^{d} x_j^p w_{pq}^j - \sum_{j=1}^{d} x_j^q v_{pq}^j + \sum_{j=1}^{d} \left(v_{pq}^j \right)^2 \right)}$$

s.t.

$$(2) - (8) \text{ and } (36) - (45).$$

3.4. Fourth Model

Considering the same substitutions (31), (33) and (34) as in the previous model it follows that:

$$\sum_{j=1}^{d} y_{pq} \left(x_j^q - \zeta_j^p \right)^2 = \sum_{j=1}^{d} \left(u_{pq}^j - \zeta_j^p y_{pq} \right)^2, \qquad (46)$$

and

$$\sum_{j=1}^{d} z_{pq} \left(x_j^q - x_j^p \right)^2 = \sum_{j=1}^{d} \left(w_{pq}^j - v_{pq}^j \right)^2. \tag{47}$$

Now, one can add the following constraints:

$$-y_{pq} \le u_{pq}^j \le y_{pq}, \ p \in P, q \in X, j \in \{1, \ldots, d\}, \tag{48}$$

$$u_{pq}^j \le x_j^q + 1 - y_{pq}, \ p \in P, q \in X, j \in \{1, \ldots, d\}, \tag{49}$$

$$u_{pq}^j \ge x_j^q - 1 + y_{pq}, \ p \in P, q \in X, j \in \{1, \ldots, d\}, \tag{50}$$

and

$$-z_{pq} \le v_{pq}^j \le z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{51}$$

$$v_{pq}^j \le x_j^p + 1 - z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{52}$$

$$v_{pq}^j \ge x_j^p - 1 + z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{53}$$

$$w_{pq}^j \le z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{54}$$

$$w_{pq}^j \ge -z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{55}$$

$$w_{pq}^j \le x_j^q + 1 - z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{56}$$

$$w_{pq}^j \ge x_j^q - 1 + z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{57}$$

to strengthen the set (2)-(8). Then, one ends up with the following model:

$$\min \sum_{p \in P, q \in X} \sqrt{\sum_{j=1}^{d} \left(u_{pq}^j - \zeta_j^p y_{pq} \right)^2} + \sum_{p,q \in X: p < q} \sqrt{\sum_{j=1}^{d} \left(w_{pq}^j - v_{pq}^j \right)^2} \tag{M_4}$$

s.t.

$$(2) - (8) \text{ and } (48) - (57).$$

3.5. Equivalence of the Four Models

All the previous models are equivalent, that is for any optimal solution of one model, these exists an optimal solution of the other having the same objective value. To prove this fact, the main argument is that the variables y and z are kept binary in all these models.

Let us consider the equivalence of the first and fourth models. The other equivalences can be obtained using the same arguments.

Let $X = (x, y, z, u, v, w)$ be an optimal solution of the model (M_4). Consider the following point (of an appropriate dimension) $A = (a, b, c, r, s)$ obtained from X as follows:

$$a = x, b = y, c = z, \tag{58}$$

$$r_{pq}^{j} = u_{pq}^{j} - y_{pq}\zeta^{p}, p \in P, q \in X, j \in \{1, \ldots, d\}, \tag{59}$$

$$s_{pq}^{j} = v_{pq}^{j} - w_{pq}^{j}, p, q \in X, j \in \{1, \ldots, d\}. \tag{60}$$

It is clear that (b, c) satisfies the constraints (2)-(8). Also, since the variables y and z are binary then:

$$r_{pq}^{j} = y_{pq}\left(x_{j}^{q} - \zeta^{p}\right) = b_{pq}\left(a_{j}^{q} - \zeta^{p}\right), p \in P, q \in X, j \in \{1, \ldots, d\},$$

$$s_{pq}^{j} = w_{pq}^{j} - v_{pq}^{j} = c_{pq}\left(a_{j}^{q} - a_{j}^{p}\right), p, q \in X, j \in \{1, \ldots, d\},$$

which means that the vector (a, b, c, r, s) satisfies the equalities (14) and (15). Also, notice that:

$$f_4(X) = f_1(A),$$

where f_1 and f_4 are the objective functions of the models (M_1) and (M_4), respectively.

Now, assume that the point A is not an optimal solution of the model (M_1). That is, there exists a feasible points $\hat{A}$ of (M_1) such that $f_1(A) > f_1(\hat{A})$. Let $\hat{A}$ be equal to $\left(\hat{a}, \hat{b}, \hat{c}, \hat{r}, \hat{s}\right)$. The point $\hat{X}$ defined as follows:

$$\hat{x} = \hat{a}, \hat{y} = \hat{b}, \hat{z} = \hat{c},$$

$$\hat{u}_{pq}^{j} = \hat{b}_{pq}\hat{a}_{j}^{q}, p \in P, q \in X, j \in \{1, \ldots, d\},$$

$$\hat{v}_{pq}^{j} = \hat{c}_{pq}\hat{a}_{j}^{p}, p, q \in X, p < q, j \in \{1, \ldots, d\},$$

$$\hat{w}_{pq}^{j} = \hat{c}_{pq}\hat{a}_{j}^{q}, p, q \in X, p < q, j \in \{1, \ldots, d\},$$

is a feasible solution of the model (M_4) and $f_1\left(\hat{A}\right) = f_4\left(\hat{X}\right)$. Thus,

$$f_1\left(\hat{A}\right) = f_4\left(\hat{X}\right),$$
$$\geq f_4(X).$$

Consequently,

$$f_1\left(\hat{A}\right) \geq f_1(A),$$

which is a contradiction. Thus, there exists for any optimal solution of the model (M_4) an optimal solution of the model (M_1) with the same objective value. The reverse property can be shown using same arguments.

Despite the theoretical equivalence of these models, from a computational point a view they lead to a non equivalent relaxations.

4. Relaxations

In the previous section, four non convex and non smooth mixed integer models for the Euclidean Steiner tree problems were presented. In this section,

mixed integer linear and nonlinear relaxations will be derived from these models. Each relaxation has the same set of feasible solutions as the set of feasible solutions of the model from which it is derived.

Before that, recall that an optimization problem:

$$\min \left\{ f\left(x, y, z\right) : \left(x, y, z\right) \in \Omega \right\}, \tag{61}$$

is a relaxation of the problem (1)-(8) if:

1. the objective function satisfies the following condition:

$$f\left(x, y, z\right) \leq \sum_{p \in P, q \in X} y_{pq} \|x^q - \zeta^p\|_2 + \sum_{p \in X, q \in X: p < q} z_{pq} \|x^q - x^p\|_2, \tag{62}$$

2. the constraint set Ω contains the set defined by (2)-(8).

Given in the next proposition, a sufficient condition that guarantees (62) for an objective function f obtained from (1) when the square root in the expression that defining it is ignored.

Proposition 3. *If for any indexes p and q that belong to P with $p < q$ such that:*

$$\|\zeta^p - \zeta^q\|_\infty \leq \frac{1}{d},$$

then

$$\sum_{p \in P, q \in X} y_{pq} \|x^q - \zeta^p\|_2^2 + \sum_{p \in X, q \in X: p < q} z_{pq} \|x^q - x^p\|_2^2$$
$$\leq \sum_{p \in P, q \in X} y_{pq} \|x^q - \zeta^p\|_2 + \sum_{p \in X, q \in X: p < q} z_{pq} \|x^q - x^p\|_2.$$

Proof. It is sufficient to notice that, for every $p \in P$ and $q \in X$:

$$\|x^q - \zeta^p\|_2^2 \leq d \max_{p \in P, q \in X} \left\{ \|x^q - \zeta^p\|_\infty^2 \right\},$$
$$\leq d \max_{p, q \in P: p < q} \left\{ \|\zeta^q - \zeta^p\|_\infty^2 \right\} \leq 1,$$

(all the Steiner points belong to the convex hull of the terminals), for every $p \in X$ and $q \in X$ such that $p < q$:

$$\|x^p - x^q\|_2^2 \leq d \max_{p, q \in X: p < q} \left\{ \|x^p - x^q\|_\infty^2 \right\},$$
$$\leq d \max_{p, q \in P: p < q} \left\{ \|\zeta^p - \zeta^q\|_\infty^2 \right\} \leq 1,$$

and, finally, any real x that belong to the interval $[0, 1]$ is less than its square root $\sqrt{x}$. This completes the proof. ∎

4.1. First Relaxation

The first relaxation is derived from the model (1)-(8) and it is defined as follows.

$$\min \sum_{p\in P, q\in X} y_{pq} \sum_{j=1}^{d} \left(x_j^q - \zeta_j^p \right)^2 + \sum_{p\in X, q\in X: p<q} z_{pq} \sum_{j=1}^{d} \left(x_j^q - x_j^p \right)^2 \qquad (R_1)$$

$$s.t.$$
$$(2) - (8).$$

The objective function of the relaxation (R_1) is smooth but not convex.

4.2. Second Relaxation

The second relaxation is derived from the first model (M_1). It is defined as follows:

$$\min \sum_{p\in P, q\in X} \sum_{j=1}^{d} \left(v_{pq}^j \right)^2 + \sum_{p\in X, q\in X: p<q} \sum_{j=1}^{d} \left(w_{pq}^j \right)^2 \qquad (R_2)$$

$$s.t.$$
$$(2) - (8) \text{ and } (17) - (22).$$

The relaxation (R_2) has a strict convex quadratic and smooth objective function and its set of feasible solutions is a polyhedron.

4.3. Third Relaxation

The third relaxation is derived from the second model (M_2). It is defined as follows:

$$\min \sum_{p\in P, q\in X} v_{pq} + \sum_{p,q\in X: p<q} w_{pq} \qquad (R_3)$$

$$s.t.$$
$$(2) - (8) \text{ and } (25) - (30).$$

The relaxation (R_3) is a mixed integer nonlinear optimization problem.

4.4. Fourth Relaxation

The fourth relaxation is derived from the model (M_1) as follows. Let us set:

$$r_{pq} = \|v_{pq}\|_\infty, p \in P, q \in X,$$
$$s_{pq} = \|w_{pq}\|_\infty, p, q \in X, p < q.$$

where the variables v and w are defined by (14)-(15).

Noticing the following valid constraints:

$$r_{pq} \geq v_{pq}^j , p \in P, q \in X, j = 1, \ldots, d, \qquad (63)$$

$$r_{pq} \geq -v_{pq}^j, p \in P, q \in X, j = 1, \ldots, d, \tag{64}$$

$$r_{pq} \geq 0, p \in P, q \in X,, \tag{65}$$

$$s_{pq} \geq w_{pq}^j, p, q \in X, p < q, j = 1, \ldots, d, \tag{66}$$

$$s_{pq} \geq -w_{pq}^j, p, q \in X, p < q, j = 1, \ldots, d, \tag{67}$$

$$s_{pq} \geq 0, p, q \in X, p < q, \tag{68}$$

and the following fact:

$$\sum_{p \in P, q \in X} r_{pq} + \sum_{p,q \in X : p < q} s_{pq} \leq \sum_{p \in P, q \in X} y_{pq} \|x^q - \zeta^p\|_2 + \sum_{p \in X, q \in X : p < q} z_{pq} \|x^q - x^p\|_2,$$

one can consider the following relaxation:

$$\min \sum_{p \in P, q \in X} r_{pq} + \sum_{p,q \in X : p < q} s_{pq} \tag{R_4}$$

$$s.t.$$

$$(2) - (8), (17) - (22) \text{ and } (63) - (68).$$

The relaxation (R_4) is a mixed integer linear optimization problem. Finally, notice that:

$$\sum_{p \in P, q \in X} r_{pq} + \sum_{p,q \in X : p < q} s_{pq} \geq \sum_{p \in P, q \in X} \sum_{j=1}^d \left(v_{pq}^j \right)^2 + \sum_{p,q \in X : p < q} \sum_{j=1}^d \left(w_{pq}^j \right)^2,$$

which means that the lower bounds found using the relaxation (R_4) are as good as the bound found using the relaxation (R_2).

4.5. Fifth Relaxation

The fifth relaxation is derived from the third model (M_3). It is defined as follows:

$$\min \sum_{p \in P, q \in X} \sum_{j=1}^d \left(u_{pq}^j - \zeta_j^p y_{pq} \right)^2 + \tag{R_5}$$

$$\sum_{p,q \in X : p < q} \left(\sum_{j=1}^d \left(w_{pq}^j \right)^2 - \sum_{j=1}^d x_j^p w_{pq}^j - \sum_{j=1}^d x_j^q v_{pq}^j + \sum_{j=1}^d \left(v_{pq}^j \right)^2 \right)$$

$$s.t.$$

$$(2) - (8) \text{ and } (36) - (45).$$

The objective function of the relaxation (R_5) is smooth but not convex.

4.6. Sixth Relaxation

The sixth relaxation is derived from the fourth model (M_4) and it is defined as follows:

$$\min \sum_{p \in P, q \in X} \sum_{j=1}^{d} \left(u_{pq}^{j} - \zeta_{j}^{p} y_{pq} \right)^2 + \sum_{p,q \in X : p < q} \sum_{j=1}^{d} \left(w_{pq}^{j} - v_{pq}^{j} \right)^2 \qquad (R_6)$$

$$s.t.$$

$$(2) - (8) \text{ and } (48) - (57).$$

The objective function in the relaxation (R_6) is smooth and convex.

5. Preliminary Computational Results

The main goal of these preliminary computational results is to highlight the main features, from a computational point of view, of the proposed relaxations. At some points, comparisons will be made with the solutions found using Smith's exact algorithm (see [18]). The solution time of all the presented experiences was fixed to 3 hours.

5.1. The Instances

Two sets of instances are considered. The first one contains five instances related to the five platonic solids. In each instance, the coordinates of the terminals are the coordinates of the vertices of the corresponding platonic solid in $\mathbb{R}^3$. These instances are named: `Tetra`, `Octa`, `Cube`, `Icosa` and `Dodec`. Recall that: the tetrahedron has 4 vertices; the octahedron has 6 vertices; the cube has 8 vertices; the icosahedron has 12 vertices; and the dodecahedron has 20 vertices.

The second set of instances is a subset of Smith's instances representing simplicies and octahedra in $\mathbb{R}^d$. The dimension of the simplicies varies from 3 to 8 and the dimension of the octahedra varies from 4 to 6. These instances are named `NSimp-x` and `NOcta-x`, where x indicates the value of the dimension d.

All the coordinates of these instances are scaled in ordered to fulfill the assumption 1.

5.2. The Results

The following Tables 1 and 2 sum up, for each instance and every relaxation the following items: the lower bound value (denoted `lb`), the upper bound value (denoted `ub`), the gap between the upper and lower bounds (denoted `gap` and defined as being: $\frac{ub-lb}{ub} \times 100$), and the gap between the upper bound `ub` and the length of the best feasible solution found using Smith's algorithm (this gap is denoted `gap*` and defined as been $\frac{ub-Best.sol}{ub} \times 100$). For a relaxation R and an instance I: `lb` is the value of the best solution found after solving I using R (the time limit being 3 hours); `ub` is the $\|.\|_2$-length of this best solution[1].

---

[1]Remember that the value of an optimal feasible solution of any relaxation is always a lower bound of its $\|.\|_2$-length, see the proposition 3

Finally, in the column `Best sol` were reported the values of the best feasible solutions found using Smith's enumeration algorithm (see [18]).

The solver `Artelys-Knitro` (see `https://www.artelys.com/`) is used to solve the relaxations R_1 and R_5. The other relaxations are solved using the `Ilog-Cplex` (see `https://www.ibm.com/`) solver. These solvers are used via the `ampl` (see `https://ampl.com/`) modeling language. As stated before, the time limit was fixed to 3 hours. All of our instances are solved on a `8Gb Ram` and `2.7 GHz x 4 Inter-Core` processor running under `Linux Ubuntu 16.04.5 LTS 64 Bits`.

Few comments on the results reported in Tables 1 and 2 are in order. First notice that, the relative gap between the upper and lower bounds is very important! Indeed, the value of `gap` is between 83% and 95% for all the relaxations except R_4. The value of `gap` for the relaxation R_4 is between 35% and 67%. This may indicate that the relaxation R_4 may perform well. This is not the case if one compares the values of `gap*`.

	R_1				R_2				R_3				
Instance	lb	ub	gap (%)	gap* (%)	lb	ub	gap (%)	gap* (%)	lb	ub	gap (%)	gap* (%)	Best sol.
Tetra	0.138889	0.824958	83.16	1.44	0.138889	0.824958	83.16	1.44	0.138889	0.824956	83.16	1.44	0.8130525127
Octa	0.107071	0.969439	88.96	1.39	0.107071	0.969439	88.96	1.39	0.107069	0.969435	88.96	1.39	0.9560044889
Cube	0.11358	1.202142	90.55	0.81	0.113469	1.204798	90.58	1.02	0.113468	1.204795	90.58	1.02	1.1924500991
Icosa	0.214883	2.009197	89.31	19.09	0.13004	1.633776	92.04	0.50	0.131442	1.642929	92.00	1.05	1.6256392494
Dodeca	0.322825	3.238799	90.03	31.73	0.146715	2.305906	93.64	4.12	0.192682	2.592707	92.57	14.72	2.2110017813
NOcta-4	0.07568	0.974522	92.23	2.39	0.075348	0.972001	92.25	2.14	0.075344	0.972004	92.25	2.14	0.9512411857
NOcta-5	0.072482	1.027447	92.95	7.62	0.05863	0.978012	94.01	2.95	0.05863	0.978011	94.01	2.95	0.9491744441
NOcta-6	0.052298	1.020536	94.88	7.07	0.04819	0.984459	95.10	3.66	0.048428	0.987845	95.10	3.99	0.9484338012
NSimp-3	0.078125	0.618718	87.37	1.44	0.078125	0.618718	87.37	1.44	0.078125	0.61871	87.37	1.44	0.6097893868
NSimp-4	0.060952	0.64171	90.50	2.29	0.060952	0.64171	90.50	2.29	0.060952	0.641716	90.50	2.29	0.6269985606
NSimp-5	0.05	0.656676	92.39	2.98	0.05	0.656676	92.39	2.98	0.049995	0.656683	92.39	2.98	0.6371368899
NSimp-6	0.043132	0.664619	93.51	3.09	0.042375	0.667347	93.65	3.49	0.042355	0.66732	93.65	3.48	0.6440799752
NSimp-7	0.036762	0.675356	94.56	3.96	0.036762	0.675356	94.56	3.96	0.036712	0.675357	94.56	3.96	0.6486057997
NSimp-8	0.032459	0.681588	95.24	4.31	0.032459	0.681588	95.24	4.31	0.032407	0.681696	95.25	4.32	0.6522373981

Table 1: Lower and upper bounds obtained using the relaxations R_1, R_2 and R_3.

	R_4				R_5				R_6				
Instance	lb	ub	gap (%)	gap* (%)	lb	ub	gap (%)	gap* (%)	lb	ub	gap (%)	gap* (%)	Best sol.
Tetra	0.58317	0.909546	35.88	10.61	0.138889	0.824958	83.16	1.44	0.138889	0.824958	83.16	1.44	0.8130525127
Octa	0.666667	1.101728	39.49	13.23	0.107071	0.969439	88.96	1.39	0.107071	0.969439	88.96	1.39	0.9560044889
Cube	0.7698	1.333333	42.26	10.57	0.115929	1.213361	90.45	1.72	0.113469	1.204795	90.58	1.02	1.1924500991
Icosa	1.107124	1.727303	35.90	5.89	0.254982	2.116967	87.96	23.21	0.13004	1.633776	92.04	0.50	1.6256392494
Dodeca	1.559627	2.521275	38.14	12.31	na	na	na	na	0.155806	2.35431	93.38	6.09	2.2110017813
NOcta-4	0.625	1.25	50.00	23.90	0.075348	0.972001	92.25	2.14	0.075348	0.972001	92.25	2.14	0.9512411857
NOcta-5	0.6	1.341641	55.28	29.25	na	na	na	na	0.05863	0.978012	94.01	2.95	0.9491744441
NOcta-6	0.583333	1.342596	56.55	29.36	na	na	na	na	0.04819	0.984459	95.10	3.66	0.9484338012
NSimp-3	0.353553	0.707107	50.00	13.76	0.078125	0.618718	87.37	1.44	0.078125	0.618718	87.37	1.44	0.6097893868
NSimp-4	0.353553	0.790569	55.28	20.69	0.060952	0.64171	90.50	2.29	0.060952	0.64171	90.50	2.29	0.6269985606
NSimp-5	0.353553	0.866025	59.18	26.43	0.05	0.656676	92.39	2.98	0.05	0.656676	92.39	2.98	0.6371368899
NSimp-6	0.353553	0.935414	62.20	31.14	0.042375	0.667347	93.65	3.49	0.042375	0.667347	93.65	3.44	0.6440799752
NSimp-7	0.353553	1	64.64	35.14	0.036762	0.675356	94.56	3.96	0.036762	0.675356	94.56	3.96	0.6486057997
NSimp-8	0.353553	1.06066	66.67	38.51	na	na	na	na	0.032459	0.681588	95.24	4.31	0.6522373981

Table 2: Lower and upper bounds obtained using the relaxations R_4, R_5 and R_6.

Considering all instances, except `Icosa` and `Dodec`, the value of `gap*` of all the relaxations except the relaxation R_4 is at most 4% while the minimum value of `gap*` for R_4 is 10.5%. So, using the relaxation R_4 is not a good choice if one

is interested in finding only a feasible solution to the Euclidean Steiner tree problem.

The relaxation R_5 is the most difficult to solve since several instances were not solved. Recall that this relaxation consists of optimizing a non convex objective function over a polyhedron.

The two instances Icosa and Dodec are the most difficult to solve, especially Dodec. The best upper bound for these two instances are found using the relaxation R_2, a gap* equal to 0.5% for Icosa and around 4% for Dodec. Also, the feasible solution found using the relaxation R_6 has a gap* equal to 0.5%.

Regarding Smith's instances, the value of gap* is at most 5% for all relaxations except R_4 for which the minimum gap* is equal to 13.76%. The value of gap* increases slightly (1% in average) with the dimension d. Because of the size of the relaxation (number of variables and constraints) which increases with the dimension d.

Finally, depicted below are the feasible solutions of the platonic instances found using relaxations R_2 and R_6. As indicated in each picture, the left image is a feasible solution found using R_2 and the right image is the feasible solution found using R_6. Notice that, for each considered instance, the topologies are different.
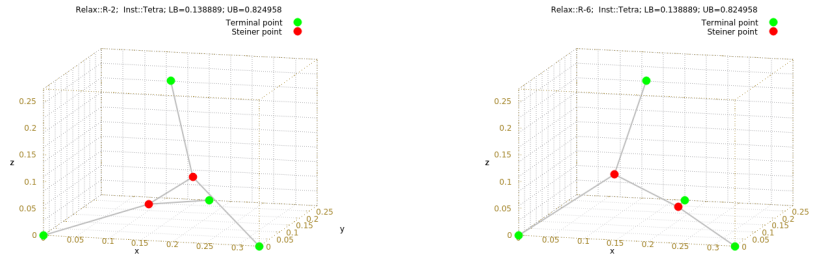

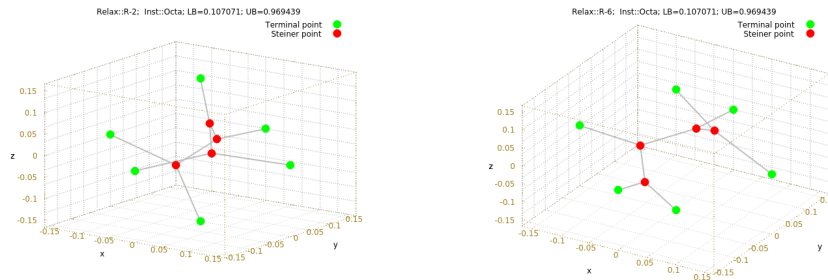Figure 1: Tetrahedron instance.

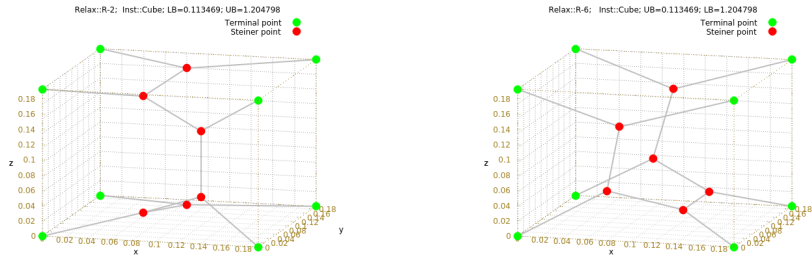
Figure 2: Octahedron instance.
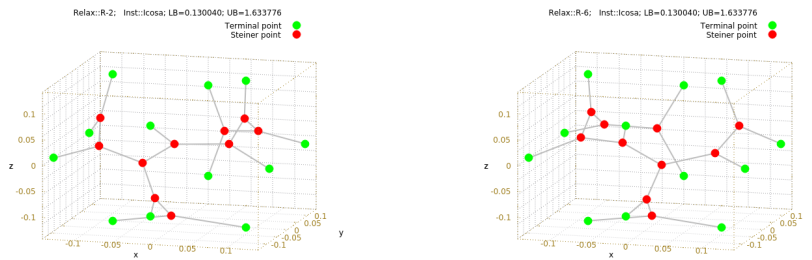
Figure 3: Cube instance.

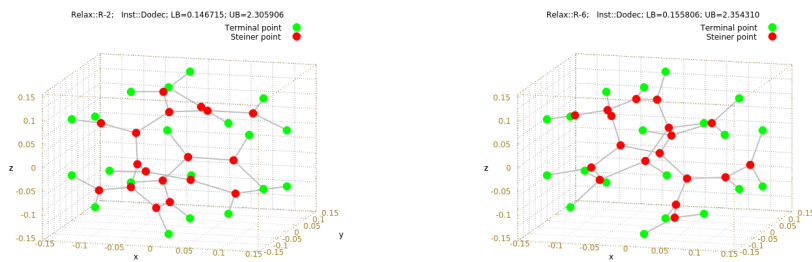
Figure 4: Icosahedron instance.

Figure 5: Dodecahedron instance.

6. Conclusion

In this work, new models for the Euclidean Steiner tree problem were presented. These models are shown to be theoretically equivalent. Six linear versus nonlinear mixed integer relaxations were derived from these models: R_1 is a nonlinear and non convex mixed integer relaxation; R_2 and R_6 are a convex quadratic

16

mixed integer relaxations; R_3 and R_4 are mixed integer linear relaxations; and, finally, R_5 is a non convex quadratic relaxation. Solving these relaxations permits, for a given instance, the computation of the lower and upper bounds of the length of the minimum Steiner tree. The upper bound found with any relaxation corresponds to the length of a feasible Steiner tree. The preliminary computational results show that even if the values of the lower bounds are very poor, the values of the upper bounds are rather tight. So, one can for instance investigate the efficiency of an ad hoc algorithm to solve, exactly or approximately (generating *good* feasible solutions), both relaxations R_2 and R_6.

Acknowledgements

References

[1] M. Brazil, R.L. Graham, D.A. Thomas, and M. Zachariasen. On the history of the euclidean steiner tree problem. *Archive for History of Exact Sciences*, 68:327–354, 2014.

[2] M. Brazil, D. A. Thomas, B. K. Nielson, P. Winter, C. Wulff-Nilsen, and M. Zachariasen. A novel approach to phylogenetic trees: *d*-dimensional geoemtric steiner trees. *Networks*, 53(2):104–111, 2009.

[3] L. L. Cavalli-Sforza and A. W. F. Edwards. Phylogenetic analysis: Models and estimation procedures. *Evolution*, 21:550–570, 1967.

[4] Warm D. M., P. Winter, and M. Zachariasen. Exact algorithms for plan steiner tree problems: A computational study. Technical report, 1998.

[5] Warm D. M. and M. Zachariasen. Large euclidean steiner minimum trees in an hour. In *ISMP*, 1997.

[6] C. D'Ambrosio, M.H.C Fampa, J. Lee, and S. Vigerske. Formulation of the euclidean steiner tree problem in n-space: missing proofs. *Optimization Letters, accepted*.

[7] M. Fampa and K. Anstreicher. An improved algorithm for computing steiner minimal trees in eucidean *d*-space. *Discrete Optimization*, (5):530–540, 2008.

[8] M.H.C Fampa, J. Lee, and N. Maculan. An overview of exact algorithms for the euclidean steiner tree problem in *n*-space. *International Transactions in Operational Research (ITOR)*, 23(4):861–874, 2016.

[9] M.H.C Fampa and N. Maculan. A new relaxation in conic form for the euclidean steiner problem in $\mathbb{R}^n$. *RAIRO - Operations Research*, 35(4):283–394, 2001.

[10] M.H.C Fampa and N. Maculan. Using a conic formulation for finding steiner minimal trees. *Numerical Algorithms*, 35(4):315–330, 2004.

[11] R. Fonseca, M. Brazil, P. Winter, and M. Zachariasen. Faster exact algorithms for computing steiner trees in higher dimensional euclidean spaces. *Presented at the* 11*th DIMACS Implementation Challenge Workshop, http://dimacs11.zib.de/workshop/FonsecaBrazilWinterZachariasen.pdf*, 2014.

[12] M. R. Garey, R. L. Graham, and D. S. Johnson. Complexity of computing steiner minimal trees. *SIAM J. Appl. Math.*, 32(4):835–859, June 1977.

[13] M. R. Garey and D. S. Johnson. *Computers and Intratability: A Guide to the Theory of NP-Completeness*. W.H. Freeman and Company, San Francisco, CA, USA, 1979.

[14] E. N. Gilbert and H. O. Pollak. Steiner minimal trees. *SIAM Journal on Applied Mathematics*, 16(1):1–29, 1968.

[15] N. Maculan, Ph. Michelon, and A.E. Xavier. The euclidean steiner tree problem in $\mathbb{R}^n$: a mathematical programming formulation. *Annals of Operations Research*, 96:209–220, 2000.

[16] F. Montenegro, J.R.A Torreão, and N. Maculan. Microcanonical optimization for the euclidean steiner problem in $\mathbb{R}^n$ with application to phylogenetic inference. *Physical Review E*, 68:056702–1– 056702–5, 2003.

[17] J. M. Smith, Y. Jang, and M. K. Kim. Steiner minimal trees, twist angles, and the protein folding problem. *Proteins: Structures, Functions, and Bioinformatics*, 66(4):889–902, 2007.

[18] W. D. Smith. How to find steiner minimal trees in euclidean *d*-space. *Algorithmica*, 7(2-3):137–177, 1992.

[19] C. Stanton and J. M. Smith. Steiner trees and 3-d macromolecular conformation. *Informs journal on computing*, 16:470–485, 2004.

[20] M. Zachariasen. *Large Euclidean Steiner Minimum Trees in an Hour*. PhD thesis, University of Copenhagen, 1998.