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 has a non smooth objective function but a convex set of feasible solutions. All these models are theoretically equivalent in the sens that for any optimal solution of one model there exists an optimal solution of the other model having the same objective value. 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, computational results showing the main features of the presented relaxations will be discussed.

Keywords: 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

*Corresponding author

Email addresses: `hacene.ouzia@sorbonne.universite.fr` (Hacene Ouzia), `maculan@cos.ufrj.br` (Nelson Maculan)

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 [9, 10]). It has several applications, to cite a few: phylogenetics (see [13, 3, 2]) and structure and folding proteins (see [16, 14]). In this paper, the dimension d will be assumed at least equal to 3 (for the case $d = 2$ see [17, 5, 4]).

As explained in [6], 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 [11] and Smith in [15]. 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 [12]. From this formulation two other formulations of the (ESTP) can be found in [7] and [8] (see [6] for more details).

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 [12]. And they are all theoretically equivalent in the sens that for any optimal solution of one model there exists an optimal solution of the other model having the same objective value. But, from a computational point of view they are different because some of them are convex and others not. Also, they all have 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 solution as the set of feasible of the model from which it is derived. Solving these relaxations using a dedicated optimization solver leads to a lower and upper bounds of an optimal solution to 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 [12] and [8] 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 [12] and [8] will be recalled.

To simplify the presentation, some notations and definitions are necessary. Let P equal to $\{1, \ldots, n\}$ be the index set the terminals $\zeta^1, \zeta^2, \ldots, \zeta^n$ and let X equal to $\{1, \ldots, n-2\}$ be 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 [12]) 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.

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

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

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

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

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

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

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

where $\|.\|_2$ is the Euclidean norm in $\mathbb{R}^d$.

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 [12] 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): the the constraints (2) indicate that the degree of a terminal is equal to 1; the constraints (3) indicate that the degree of a Steiner point is equal to 3; and the constraints (4) eliminate subtours among Steiner points. Finally, the 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 belonging 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 multiplied by a factor $\frac{1}{\delta\beta d}$.

2.2. Formulation by Fampa and Maculan

The formulation by Fampa and Maculan (see [7] and [8]) is derived from the formulation by Maculan, Michelon and Xavier (see [12]) 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)-(7):

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

Indeed, for any indexes p and q belonging 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 the optimum solution is 0. Now, if y_{pq} is equal to 1 then the corresponding constraint in (9) implies that:

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

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 said before, are all equivalent and derived from the model in [12]. 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:

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

Using the following substitutions:

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

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

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):

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

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

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

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

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

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

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:

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

Thus, adding the following strengthen constraints:

$$-y_{pq} \leq u_{pq}^{j} \leq y_{pq}, p \in P q \in X, j \in \{1, \ldots, d\}, \tag{36}$$

$$u_{pq}^{j} \leq x_{j}^{q} + 1 - y_{pq}, p \in P, q \in X, j \in \{1, \ldots, d\}, \tag{37}$$

$$u_{pq}^{j} \geq x_{p}^{q} - 1 + y_{pq}, p \in P, q \in X, j \in \{1, \ldots, d\}, \tag{38}$$

and

$$-z_{pq} \leq v_{pq}^{j} \leq z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{39}$$

$$v_{pq}^{j} \leq x_{j}^{p} + 1 - z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{40}$$

$$v_{pq}^{j} \geq x_{j}^{p} - 1 + z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{41}$$

$$w_{pq}^{j} \leq z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{42}$$

$$w_{pq}^{j} \geq -z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{43}$$

$$w_{pq}^{j} \leq x_{j}^{q} + 1 - z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{44}$$

$$w_{pq}^j \geq x_j^q - 1 + z_{pq}, p, q \in X, p < q, j \in \{1, \ldots, d\}, \tag{45}$$

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

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

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

$s.t.$

$(2) - (8)$ and $(36) - (45)$.

3.4. Fourth Model

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

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

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, adding the following strengthen constraints:

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

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

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

and

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

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

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

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

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

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

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

to the set (2)-(8) 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) 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 there exists an optimal solution of the other having the same objective value. To prove it, the main argument is the fact 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 (Ω_1) and (Ω_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 $f_1(A) > f_1\left(\hat{A}\right)$. 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:

$$\begin{aligned}
\hat{x} &= \hat{a}, \hat{y} = \hat{b}, \hat{z} = \hat{c}, \\
\hat{u}^j_{pq} &= \hat{b}_{pq}\hat{a}^q_j, p \in P, q \in X, j \in \{1, \ldots, d\}, \\
\hat{v}^j_{pq} &= \hat{c}_{pq}\hat{a}^p_j, p, q \in X, p < q, j \in \{1, \ldots, d\}, \\
\hat{w}^j_{pq} &= \hat{c}_{pq}\hat{a}^q_j, p, q \in X, p < q, j \in \{1, \ldots, d\},
\end{aligned}$$

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

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

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 they are not equivalent from a computational point a view. Indeed, solving a non convex optimization problem is more challenging than solving a convex one.

4. Relaxations

In the previous section, four non convex and non smooth mixed integer models for the Euclidean Steiner tree problems was 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 of feasible solutions (2)-(8).

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

Proposition 3. *If for any indexes p and q belonging 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 belonging to the interval $[0,1]$ is less than its square root $\sqrt{x}$. This completes the proof. ■

4.1. First Relaxation

This 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 not convex but it is smooth.

4.2. Second Relaxation

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

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

$s.t.$

$(2) - (8)$ 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 linear optimization problem.

4.4. Fourth Relaxation

This fourth relaxation is derived from (1)-(8). Let us set:

$$r_{pq} = \|x^q - \zeta^p\|_\infty, p \in P, q \in X,$$
$$s_{pq} = \|x^q - x^p\|_\infty, p, q \in X, p < q.$$

According to the assumption 1, notice that:

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

Thus, one can consider the following relaxation:

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

$s.t.$

$$(2) - (8).$$

The relaxation (R_4) is a mixed integer linear optimization problem.

4.5. Fifth Relaxation

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

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

$s.t.$

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

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) 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 results is to compare the quality of the lower and upper bounds computed using the proposed relaxation. For each instance, the comparison being with the value of the best known Steiner tree.

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 are Smith instances representing the simplicies and octahedra in $\mathbb{R}^d$. For the simplicies instances the dimension d belongs to $\{3, \ldots, 8\}$ and for the octahedra the dimension d belongs to $\{4, 5, 6\}$. These instances are named `NSimp-x`, where `x` indicates the dimension d.

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

5.2. The Results

In the following Tables 1 and 2 are summed up, for each instance and for each relaxation its lower bound value (denoted `lb`) and its upper bound (denoted `ub`) value. 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]. Finally, in the column `Best sol.` is reported the value of the best solution found using Smith's enumeration algorithm (the time limit being fixed to 3 hours too.).

	R_1		R_2		R_3		
`Instance`	`lb`	`ub`	`lb`	`ub`	`lb`	`ub`	`Best sol.`
`Tetra`	0.138889	0.824958	0.138889	0.824958	0.138889	0.824956	0.8130525127
`Octa`	0.107071	0.969439	0.107071	0.969439	0.107069	0.969435	0.9560044889
`Cube`	0.11358	1.202142	0.113469	1.204798	0.113468	1.204795	1.1924500991
`Icosa`	0.214883	2.009197	0.13004	1.633776	0.131442	1.642929	1.6256392494
`Dodec`	0.322825	3.238779	0.146715	2.305906	0.192682	2.592707	2.2110017813
`NOcta-4`	0.07568	0.974522	0.075348	0.972001	0.075344	0.972004	0.9512411857
`NOcta-5`	0.072482	1.027447	0.05863	0.978012	0.05863	0.978011	0.9491744441
`NOcta-6`	0.052298	1.020536	0.04819	0.984459	0.048428	0.987845	0.9484338012
`NSimp-3`	0.078125	0.618718	0.078125	0.618718	0.078125	0.61871	0.6097893868
`NSimp-4`	0.060952	0.64171	0.060952	0.64171	0.060952	0.641716	0.6269985606
`NSimp-5`	0.05	0.656676	0.05	0.656676	0.049995	0.656683	0.6371368899
`NSimp-6`	0.043132	0.664619	0.042375	0.667347	0.042355	0.66732	0.6440799752
`NSimp-7`	0.036762	0.675356	0.036762	0.675356	0.036712	0.675357	0.6486057997
`NSimp-8`	0.032459	0.681588	0.032459	0.681588	0.032407	0.681696	0.6522373981

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

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

[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

the `ampl` (see `https://ampl.com/`) modeling language. The time limit was fixed to 3 hours.

	R_4		R_5		R_6		
Instance	lb	ub	lb	ub	lb	ub	Best sol.
Tetra	0.58317	0.909546	0.138889	0.824958	0.138889	0.824958	0.8130525127
Octa	0.666667	1.101728	0.107071	0.969439	0.107071	0.969439	0.9560044889
Cube	0.7698	1.333333	0.115929	1.213361	0.113469	1.204798	1.1924500991
Icosa	1.107124	1.727303	0.254982	2.116967	0.13004	1.633776	1.6256392494
Dodec	1.559627	2.521275	na	na	0.155806	2.35431	2.2110017813
NOcta-4	0.625	1.25	0.075348	0.972001	0.075348	0.972001	0.9512411857
NOcta-5	0.6	1.341641	na	na	0.05863	0.978012	0.9491744441
NOcta-6	0.583333	1.342596	na	na	0.04819	0.984459	0.9484338012
NSimp-3	0.353553	0.707107	0.078125	0.618718	0.078125	0.618718	0.6097893868
NSimp-4	0.353553	0.790569	0.060952	0.64171	0.060952	0.64171	0.6269985606
NSimp-5	0.353553	0.866025	0.05	0.656676	0.05	0.656676	0.6371368899
NSimp-6	0.353553	0.935414	0.042375	0.667347	0.042375	0.667347	0.6440799752
NSimp-7	0.353553	1	0.036762	0.675356	0.036762	0.675356	0.6486057997
NSimp-8	0.353553	1.06066	na	na	0.032459	0.681588	0.6522373981

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

All 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`.

Below few comments on the results reported in Tables 1 and 2. First notice that, the relative gap between the upper and lower bounds (which is defined as been: $\frac{up-lb}{ub}$ and denoted in the sequel by `GAP`) 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 indicates that the relaxation R_4 may performs well. This is not the case if one compares the relative gap between the value of the best known solution and the value of the upper bound found by the relaxation (this gap is defined as been $\frac{ub-Best.sol}{ub}$ and denoted in the sequel by `GAP*`.)

Considering all instances except `Icosa` and `Dodec`, the value of `GAP*` for all 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. Notice that the relaxation R_5 is the most difficult to solve. Recall that this relaxation consists in 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 a around 4% for `Dodec`. Also, the feasible solution returned by 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 the size of the relaxation (number of variables and constraints) increases with the dimension d.

Finally, below are depicted the feasible solutions of the platonic instances found by the two 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 of the solution found are different.

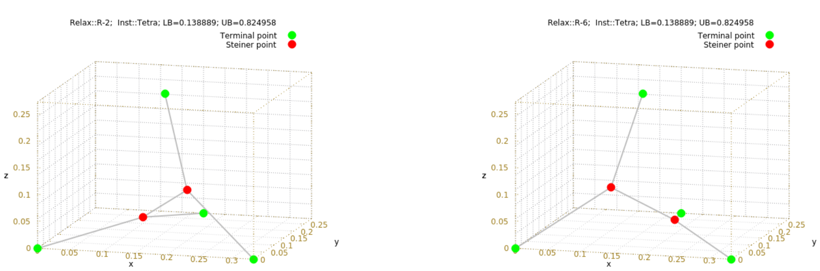

Figure 1: Tetrahedron instance.

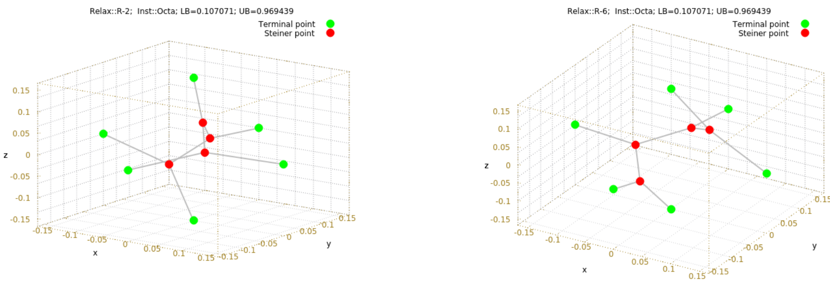

Figure 2: Octahedron instance.

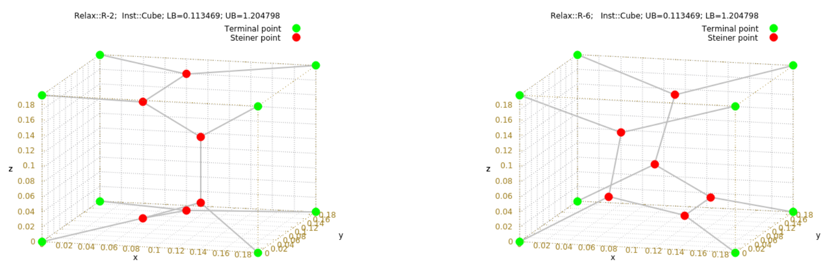

Figure 3: Cube instance.

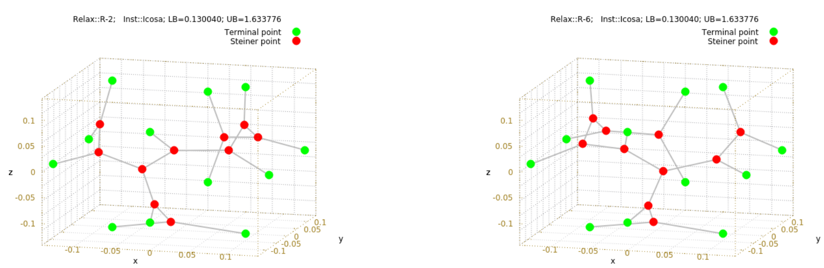

Figure 4: Icosahedron instance.

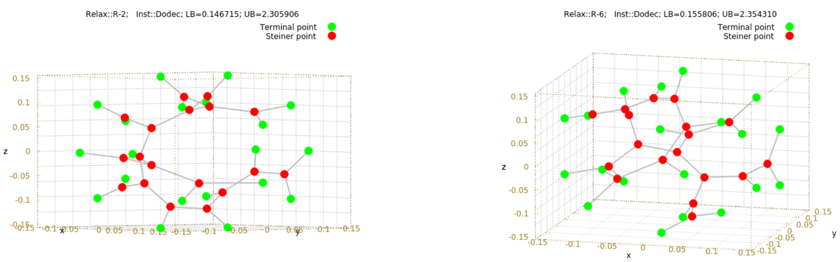

Figure 5: Dodecahedron instance.

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/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 mixed integer relaxations; R_3 and R_4 are mixed integer linear relaxations; and R_5 is a non convex quadratic relaxation. Solving these relaxations permits, for a given instance, the computation of lower and upper bounds to the the length of the minimum Steiner tree. The preliminary computational results show that the values of the lower bounds are very poor. But the values of the upper bounds are rather tight.

Acknowledgements

This work was partially supported by the National Council for Scientific and Technological Development - CNPq, under grant 302578/2014-5, and by CAPES-MEC.

References

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

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

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

[4] D. M., W., Winter, P., & Zachariasen, M. (1998). *Exact Algorithms for Plan Steiner tree Problems: A Computational Study*. Technical Report.

[5] D. M., W., & Zachariasen, M. (1997). Large euclidean steiner minimum trees in an hour. In *ISMP*.

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

[7] Fampa, M., & Maculan, N. (2001). A new relaxation in conic form for the euclidean steiner problem in r^n. *RAIRO - Operations Research*, *35*, 283–394.

[8] Fampa, M., & Maculan, N. (2004). Using a conic formulation for finding steiner minimal trees. *Numerical Algorithms*, *35*, 315–330.

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

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

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

[12] Maculan, N., Michelon, P., & Xavier, A. (2000). The euclidean steiner tree problem in r^n : a mathematical programming formulation. *Annals of Operations Research*, *96*, 209–220.

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

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

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

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

[17] Zachariasen, M. (1998). *Large Euclidean Steiner Minimum Trees in an Hour*. Ph.D. thesis University of Copenhagen.