Properties, Extensions and Application of Piecewise Linearization for Euclidean Norm Optimization in R 2

Aloïs Duguet, Christian Artigues, Laurent Houssin, Sandra Ulrich Ngueveu

Properties, Extensions and Application of Piecewise Linearization for Euclidean Norm Optimization in $\mathbb{R}^2$

Aloïs Duguet[1], Christian Artigues[1], Laurent Houssin[2], Sandra Ulrich Ngueveu[1]

[1]: LAAS-CNRS, Université de Toulouse, CNRS, INP, Toulouse, France

[2]: ISAE-SUPAERO, Université de Toulouse, Toulouse, France

Abstract

This work considers non-convex mixed integer nonlinear programming where nonlinearity comes from the presence of the two-dimensional euclidean norm in the objective or the constraints. We build from the euclidean norm piecewise linearization proposed by Camino et al. [2019], that allows to solve such non-convex problems via mixed-integer linear programming with an arbitrary approximation guarantee. Theoretical results are established that prove that this linearization is able to satisfy any given approximation level with the minimum number of pieces. An extension of the piecewise linearization approach is proposed. It shares the same theoretical properties for elliptic constraints and/or objective. An application shows the practical appeal of the elliptic linearization on a non-convex beam layout mixed optimization problem coming from an industrial application.

1 Introduction

This work deals with non-convex mixed-integer nonlinear programming (MINLP) problems involving euclidean norms either in the objective, the constraints or both. Typical examples are minimization or maximization of distances between objects, as well as convex proximity or separation constraints. Recent applications involve service infrastructure placement in 5G networks Santoyo-González and Cervelló-Pastor [2018], relay node deployment in wireless networks Zhou et al. [2018] and beam layout optimization in telecommunication satellites Camino et al. [2014, 2019]. We use the latter problem as a case-study.

MINLP problems are very challenging. Exact methods are generally based on Branch-and-Bound Smith and Pantelides [1999] to ensure global optimality, using different relaxations Adams and Sherali [1986], Nowak [2005], for example convex relaxation Liberti [2004] or convex envelopes Tardella [2007]. To exhibit these relaxations, numerous works deal with problem reformulation, which allows to simplify the computation of the relaxation or to strengthen it Liberti et al. [2009b], Liberti [2004], Sherali and Liberti [2009], Smith and Pantelides [1999]. The reformulation-linearization technique (RLT) Sherali and Adams [1999], Sherali and Liberti [2009] is a widely used reformulation method. Applying this method requires two steps: first, reformulate the problem so as to add valid nonlinear constraints; second, linearize the model, replacing variable products by new continuous variables. Generally, MINLP problems are considered harder to solve than Mixed-Integer Linear Programming (MILP) ones, leading to the use of reformulation methods Geißler et al. [2012], Liberti et al. [2009a]. Indeed, MILP solvers are often able to tackle industrial-sized problems Borghetti et al. [2008], Camponogara et al. [2011].

Linearizing a MINLP problem consists in replacing all its nonlinear functions with piecewise linear ones. Theoretical advantages of MINLP linearizations are discussed in Geißler et al. [2012]. In some cases, there are methods to obtain guarantees on the linear approximation, which motivate linearization. However, linearization also has drawbacks like the fastly increasing number of variables and constraints needed to represent a piecewise linear function depending on the dimension of the input Hughes and Anderson [1996], Smith [2000]. Another drawback is the control of the approximation error Geißler et al. [2012].

Despite those disadvantages, linearization is a widely used method to solve MINLP problems and several linearizing methods have been published. Some of them are valid for any dimensions Geißler et al. [2012], Zhang and Wang [2008] and others for two dimensions or more D'Ambrosio et al. [2010], Misener et al. [2009], Rovatti et al. [2014], Silva and Camponogara [2014].

A particularly relevant topic in linearization is the modeling of a piecewise linear function. Good properties of formulations have been identified, such as the locally ideal property, which corresponds to when each vertex of the linear relaxation is integral Keha et al. [2004], Padberg [2000]. A comparison of formulations which focuses on this property is available in Sridhar et al. [2013]. Also, a formulation described in Vielma and Nemhauser [2011] uses only a logarithmic number of binary variables and constraints in the number of pieces.

Eventually, as MINLP linearizations only yield an approximation of the original problem, being able to quantify the approximation error is a major advantage. There are at least two general ways of controlling the approximation when the linearization is only in the objective function. The first is a trial and error procedure: linearize the MINLP problem, check the quality of the candidate solution found, and if it is not good enough try to linearize with more pieces. Some linearization schemes only refine the linearization around the solution candidate Burlacu et al. [2020]. Despite being easy to implement, this linearization scheme as a disadvantage: the number of necessary iterations is not known in advance. This is why the second way of controlling the linearization approximation is to enforce before optimization that the biggest approximation error will be smaller than $\delta > 0$. In this case, the optimal value of the MILP is by construction at a distance of no more than δ from the optimal value of the

MINLP. Examples are described in Dunham [1986], Geißler et al. [2012], Rosen and Pardalos [1986]. Piecewise linear bounding of the nonlinear functions is one of those. Such boundings are used in Geißler et al. [2012], Ngueveu [2019], Rebennack and Kallrath [2015a], Rebennack and Krasko [2020]. The first way of controlling the approximation error may introduce at the end less linear pieces in the linearization, leading to smaller MILP, while the latter ensures the satisfaction of the approximation error bound in a single MILP solving iteration. In this work, the second way is preferred for two main reasons. First, the euclidean norm has nice properties allowing to control a priori the approximation error. Second, when linearizing constraints, ensuring their satisfaction at the first MILP run is an advantage.

As the number of pieces used in a piecewise linear function increases, the MILP problem associated requires an increasing number of binary variables and constraints, which means that the problem is expected to take more time to be solved. Thus, using as few pieces as possible to achieve a satisfying approximation error is of particular interest.

This article focuses on using the lowest number of pieces to satisfy a given approximation level rather than minimizing the approximation error with a given number of pieces. To the best of our knowledge, articles in the literature are few to choose the same point of view Ngueveu [2019], Rebennack and Kallrath [2015b], Rebennack and Krasko [2020], Rosen and Pardalos [1986]. The goal of this article is to develop such a linearizing method for the euclidean norm, building on the linearization approach proposed by Camino et al. [2019].

This article has two main contributions. First, we prove that given an approximation level, the linearizing method described in Camino et al. [2019] uses the minimal number of pieces. Second, we extend this bounding method to other functions. In particular we address the linearization of a norm that has an application in the telecommunication satellite domain. It evaluates the gain obtained by covering user areas with elliptic beams instead of circular ones. The article is organized as follows: Section 2 describes the euclidean norm linearization of Camino et al. [2019], and proves that given an approximation error threshold, it creates a piecewise linear bounding with the minimal number of pieces. In Section 3, it is shown that the linearizing method of Camino et al. [2019] can be used to linearize the euclidean norm of $\mathbb{R}^2$ in the objective function, and that it can be adapted to linearize norm with level set that are ellipses. In Section 4, the interest of linearizing elliptic constraints is demonstrated on the beam layout satellite telecommunication problem. Eventually, conclusions are drawn in Section 5.

2 Euclidean Norm Linearization

In this section, the linearization scheme of the euclidean norm of $\mathbb{R}^2$ of Camino et al. [2019] is described, then a definition of the approximation error used throughout this article is given, and finally it is proved that the linearization of Camino et al. [2019] can be used to create a piecewise linear bounding with

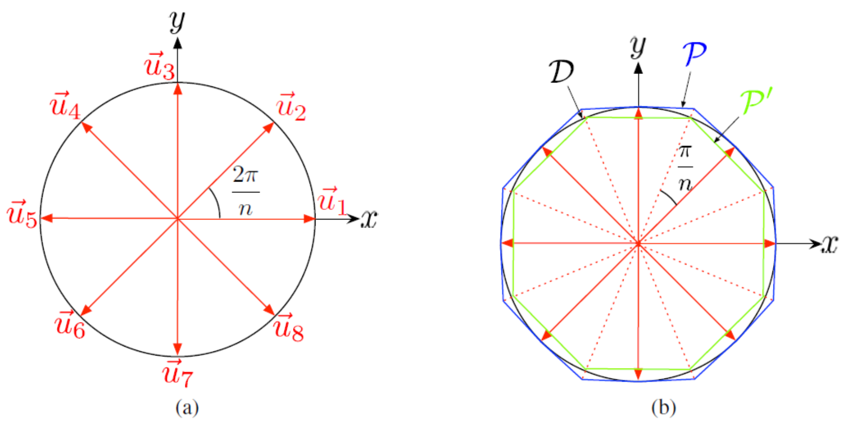

Figure 1: Euclidean norm linearization for 8 directions Camino et al. [2019]

the minimal number of pieces that satisfies a given approximation error. For the remainder of this article, the term *euclidean norm linearization* denote the linearization scheme of the euclidean norm of $\mathbb{R}^2$ of Camino et al. [2019].

2.1 Method Description

The euclidean norm linearization describes how to linearize a constraint with a single euclidean norm of $\mathbb{R}^2$ term, for both inequality directions. However, it cannot be applied to a sum of euclidean norms. The linearization is based on the evaluation of multiple scalar products between the vector appearing in the euclidean norm and unit vectors regularly spaced out. Let $||\cdot||_2$ be the euclidean norm of $\mathbb{R}^n$. Let us consider the constraint

$$||x||_2 \leq \Delta \quad \forall x \in \mathbb{R}^2 \tag{1}$$

for $\Delta \in \mathbb{R}^+$. This constraint is satisfied if x is in the centered disk of radius Δ. Set $\mathbf{u_i} = (\cos \frac{2i\pi}{p}, \sin \frac{2i\pi}{p})$ for $i = 1, ..., p$ with p the number of scalar products used for the linearization. The directions $\mathbf{u_i}$ for $i = 1, ..., 8$ are depicted in Figure 1(a). The linearization consists in replacing constraint (1) by the constraints

$$x.\mathbf{u_i} \leq \Delta \quad \forall i = 1, ..., p, \tag{2}$$

with . being the usual scalar product. Throughout this article, the word "polygon" denotes a solid plane region inside a closed polygonal chain without intersection. A polygonal chain is a union of segments where one segment intersects at its endpoints with the previous and the next segment of the chain. In addition, a polygonal chain is closed if its startpoint and endpoint are the same. If constraint (2) is satisfied, x is in the blue polygon of Figure 1(b), and according to ([Camino et al., 2019, Proposition 1]) the following upper bound on $||x||_2$

holds:

$$||x||_2 \leq \frac{\Delta}{\cos\frac{\pi}{p}}. \tag{3}$$

In addition, if one of the constraints of (2) is not satisfied, x is out of the blue polygon, and ([Camino et al., 2019, Proposition 2]) gives that

$$||x||_2 \geq \Delta. \tag{4}$$

In this example, (2) is a relaxation of (1), thus some x satisfying (2) might be infeasible for (1), but a constraint stronger than (1) can be obtained by linearizing it with

$$x.\mathbf{u_i} \leq \Delta \cos\frac{\pi}{P} \quad i = 1, ..., p, \tag{5}$$

which has the green polygon of Figure 1(b) for feasible set.

The linearization of constraint (4) differs from that of (1) because it is non-convex. Thus the constraints linearizing (4) are

$$x.\mathbf{u_i} \geq \Delta - M_i(1 - b_i), \quad i = 1, ..., p, \tag{6}$$

$$\sum_{i=1}^{p} b_i = 1, \tag{7}$$

$$b_i \in \{0, 1\}, \quad i = 1, ..., p, \tag{8}$$

which use the big-M method, but the principle stays the same as constraints (5). Here, M_i must be a valid upper bound of $\Delta - x.\mathbf{u_i}$, which is easily obtained if x is bounded. Remark that the polygons created by the linearization of (1) into both (2) and (5) are convex regular polygons with p sides, for all $p \geq 3$, because they are both convex, equiangulars and equilaterals. Note that in the remainder of this work, regular polygon is used for convex regular polygon.

Finally remark that, intuitively, the blue polygon in Figure 1 is the "smallest" regular polygon of 8 sides containing the black disk because each segment middle point touches the circle forming the disk. Conversely, the green polygon is the "biggest" inside the black disk because each vertex touches the circle forming the disk. In the following, we focus on the approximation of the frontier of constraints of type $||x||_2 \leq \Delta$ or $||x||_2 \geq \Delta$. The blue polygon is denoted an *outer approximation* of the disk while the green polygon is denoted an *inner approximation* of the disk. Those two terms are properly defined in the following subsection for the need of this work.

2.2 Approximation error

An approximation error tailored for the euclidean norm linearization is introduced in Definition 2.2. It measures the highest proportion of euclidean norm of two points inside the difference of a disk $B(x_0, \Delta)$ and a polygon. If the

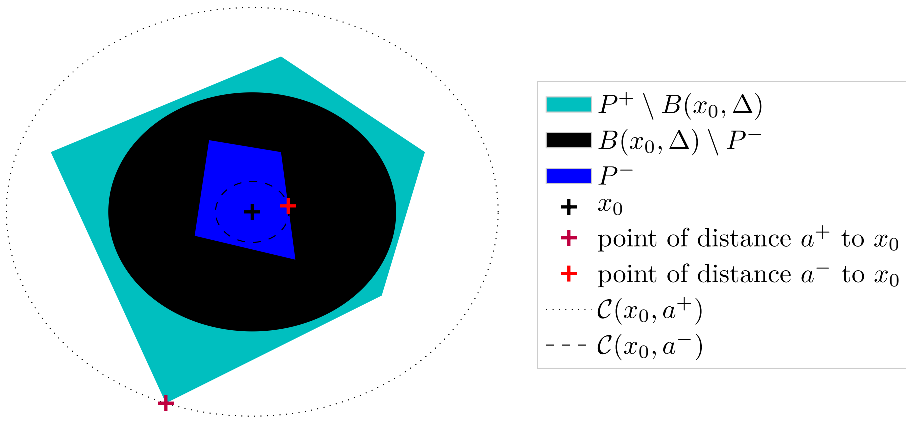

Figure 2: Illustration of Definition 2.2

polygon contains the disk, it is denoted an outer approximation; if the polygon is included in the disk, it is denoted an inner approximation.

Let $n \in \mathbb{N}^*$ and $x_0 \in \mathbb{R}^2$. Let $B(x_0, \Delta) = \{x \in \mathbb{R}^2 \ : \ ||x - x_0||_2 \leq \Delta\}$ be the disk of radius Δ and center x_0, with $\Delta \in \mathbb{R}^+$. Let P^+ be a polygon and $a^+ \in \mathbb{R}^+$ be such that $B(x_0, \Delta) \subset P^+$ and $a^+ = \sup\{||x - x_0||_2 \ : \ x \in P^+ \backslash B(x_0, \Delta)\}$. Finally, let P^- be a polygon and $a^- \in \mathbb{R}^+$ be such that $P^- \subset B(x_0, \Delta)$ and $a^- = \inf\{||x - x_0||_2 \ : \ x \in B(x_0, \Delta) \backslash P^-\}$. We say that:

- P^+ is an outer approximation of $B(x_0, \Delta)$ with error $\epsilon := \frac{a^+}{\Delta} - 1$,
- P^- is an inner approximation of $B(x_0, \Delta)$ with error $\epsilon := \frac{\Delta}{a^-} - 1$,
- P^+ and P^- are a bounding of $B(x_0, \Delta)$ with error $\epsilon := \max\{\frac{a^+}{\Delta}, \frac{\Delta}{a^-}\} - 1$.

The -1 term in the formula of error ϵ is added so that an approximation with error $\epsilon = 0$ is exact. A similar definition can be made for constraints $||x - x_0||_2 \geq \Delta$. Figure 2 illustrates the definition. The region P^+ is in cyan, black and blue while the region P^- is in blue, and the region $B(x_0, \Delta)$ is in black and blue. The cyan region is $P^+ \setminus B(x_0, \Delta)$, the region where the approximation error of the outer approximation occurs. The point of maximal distance from x_0 is the purple one. The norm of the purple point minus x_0 is the radius of the dotted circle. Similarly, the black region is $B(x_0, \Delta) \setminus P^-$, the region where the approximation error of the inner approximation occurs. The point of minimal distance from x_0 is the red one. The norm of the red point minus x_0 is the radius of the dashed circle.

In this article, the outer approximation polygons P^+ (resp. inner approximation polygons P^-) that interest us are those defined with points $x \in \mathbb{R}^2$ satisfying constraints of type (5) (resp. constraints of type (2)) or constraints (5) with reversed inequality (resp. constraints (2) with reversed inequality).

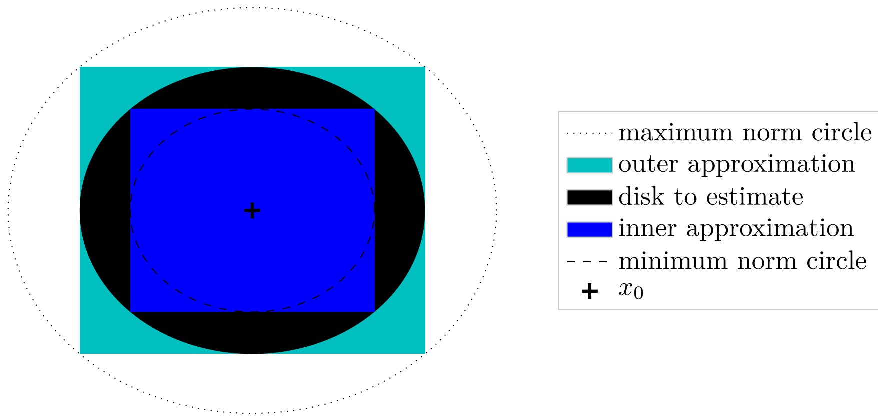

Figure 3: Bounding of a disk by (2) and (5) for $p = 4$ scalar products

It is illustrated in Figure 3 where the black and blue disk is bounded by the cyan and blue squares induced by the linearization of (2) and (5) for $p = 4$. The maximal norm of the outer approximation square (resp. the minimal norm of the inner approximation square) is depicted by the dotted circle (resp. the dashed circle). Together, the inner and outer approximation of the disk are a bounding of it. Thus, according to Definition 2.2, the approximation error is the maximum between the ratio of radius of dotted and black circles minus one and the ratio of radius of black and dashed circles minus one.

A calculation shows that the approximation error associated to (2) and (5) are both $\frac{1}{\cos \frac{\pi}{p}} - 1$ because the intervals in which approximation errors occur are $[\Delta, \frac{\Delta}{\cos \frac{\pi}{p}}]$ and $[\Delta \cos \frac{\pi}{p}, \Delta]$ respectively. Remark that the approximation error tends to 0 when the number of scalar products p tends to infinity, thus the euclidean norm linearization is an approximation that can be as good as needed by choosing p big enough.

2.3 Uniform Inner Approximation

In this subsection, we show that for a fixed number of sides p, regular polygons with all vertices on the circle are the only polygons with minimal error to approximate a disk. To simplify the notation, we introduce a useful notation for the approximation error of a disk by a polygon. Let $\mathcal{P}$ be a polygon inside a disk $\mathcal{D}$ centered at $(0, 0)$. $\epsilon(\mathcal{P})$ denote the approximation error of the disk $\mathcal{D}$ by the polygon $\mathcal{P}$ according to Definition 2.2.

Let P be a polygon inside the unit disk $\mathcal{D}$, recall that $\epsilon(P)$ is -1 plus the inverse of the minimal norm of a point of $\mathcal{D} \backslash P$.

Let $\mathcal{D}$ be the unit disk. Let $\mathcal{P} \subset \mathcal{D}$ be a simple nonconvex polygon, where simple means that no edges of $\mathcal{P}$ intersect except in the extreme points. Let $\mathcal{P}' = conv(\mathcal{P})$ be the convex envelope of $\mathcal{P}$. Then, we have that $\mathcal{P}' \subset \mathcal{D}$, $\mathcal{P}'$ has strictly less sides than $\mathcal{P}$ and $\epsilon(\mathcal{P}) \geq \epsilon(\mathcal{P}')$.

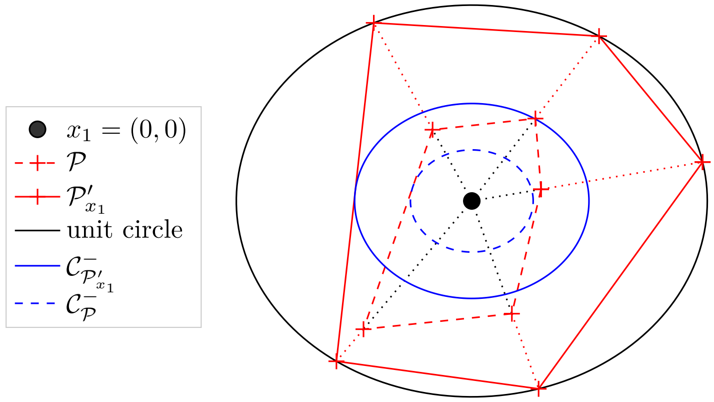

Figure 4: Inflation of $\mathcal{P}$ by $x_1 = (0, 0)$

Proof. First, $\mathcal{P} \subset \mathcal{D}$ implies that $\mathcal{P}' \subset \mathcal{D}$ because $\mathcal{D}$ is convex. Second, a polygon is convex if each angle of this polygon is smaller than π. Now as $\mathcal{P}$ is not convex, its convex envelope $\mathcal{P}'$ has strictly less sides because at least one vertex of $\mathcal{P}$ has an angle strictly greater than π and thus will not be a vertex of $\mathcal{P}'$. Third, $\mathcal{P} \subset \mathcal{P}' = conv(\mathcal{P})$ implies that the set $\mathcal{D} \backslash \mathcal{P}'$ over which is calculated the minimum a^- of Definition 2.2 is smaller than $\mathcal{D} \backslash \mathcal{P}$, thus leading to $\epsilon(\mathcal{P}) \geq \epsilon(conv(\mathcal{P}))$. □

According to Lemma 2.3, a nonconvex polygon with p sides cannot induce a smaller approximation error than its convex envelope when approximating a disk. Thus there is a convex polygon with p sides that has the minimal error to approximate the disk. Let $\mathcal{A}_p^-$ be the set of convex polygons of p sides lying inside $\mathcal{D}$, the unit disk centered at $(0, 0)$. The minus symbol refers to the inner approximation case. We consider $p \geq 3$ to avoid undefined polygons (with $p \leq 1$) and lines (with $p = 2$). The unit disk is centered at $(0, 0)$ because the number of sides of a polygon is invariant by homothetic transformation so it is always possible to return to this case from a disk centered at $x_0 \in \mathbb{R}^2$. The goal is to show that a regular polygon with p sides and all vertices on the circle is an optimal solution of the following optimization problem.

$$(\mathcal{P}_p^-) \quad \begin{cases} \min & \epsilon(\mathcal{P}) \\ s.c. & \mathcal{P} \in \mathcal{A}_p^- \end{cases} \tag{9}$$

A definition of a specific transformation of a polygon is introduced as a first argument to show that this regular polygon is a solution of (9). [inflation] Let $\mathcal{P} \in \mathcal{A}_p^-$ and let $x \in Int(\mathcal{P}) = \{x \in \mathcal{P} \setminus Fr(\mathcal{P})\}$ the interior of $\mathcal{P}$, where $Fr(\mathcal{P})$ is the frontier of $\mathcal{P}$. Let *inflation of* $\mathcal{P}$ *by* x denote the polygon $\mathcal{P}'_x$ with vertices $v \in V$ where each v is obtained by intersecting the unit disk with the ray of initial point x and direction $\overrightarrow{xw}$, with w a vertex of $\mathcal{P}$. Figures (4) and (5) illustrate this definition. They show the same dashed red polygon $\mathcal{P}$ inflated by two different x. The red dashed polygon is inflated by $x_1 = (0, 0)$ in the

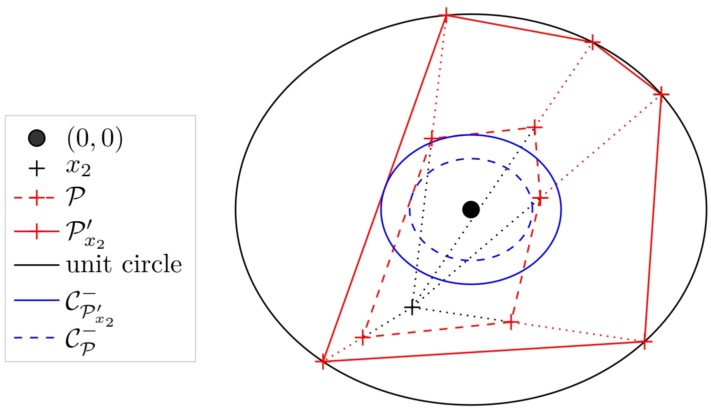

Figure 5: Inflation of $\mathcal{P}$ by $x_2 = (-0.25, -0.5)$

first figure, and in the second by $x_2 = (-0.25, -0.5)$. The inflated polygons $\mathcal{P}'_{x_1}$ and $\mathcal{P}'_{x_2}$ are in red, with dotted black and red lines showing the movement of the vertices. Finally, the blue dashed circles and the blue circles show the minimal norm of the polygons $\mathcal{P}$ and $\mathcal{P}'_x$ respectively. Their radius are the a^- of Definition 2.2.

Let $p \geq 3$ and let $\mathcal{P}'_x$ be the inflation of $\mathcal{P} \in \mathcal{A}^-_p$ by $x \in Int(\mathcal{P})$. Then we have:

1. $\mathcal{P}'_x$ is a polygon
2. $Int(\mathcal{P}) \subset Int(\mathcal{P}'_x)$
3. The vertices of $\mathcal{P}'_x$ are on the circle
4. $\mathcal{P}'_x$ is convex
5. $\mathcal{P}'_x$ has the same number of sides as $\mathcal{P}$
6. $\epsilon(\mathcal{P}'_x) \leq \epsilon(\mathcal{P})$

Proof. 1. By definition, $\mathcal{P}'_x$ is a polygon if its frontier is a closed polygonal chain without intersections. As the vertices of $\mathcal{P}'_x$ come from vertices of $\mathcal{P}$, by adding the straight line segments induced by the straight line segments of $\mathcal{P}$, $\mathcal{P}'_x$ must be in a closed polygonal chain. To prove that there are no intersections in this polygonal chain, remark that the p rays used to build the inflation are intersecting in x. Also, an intersection of the polygonal chain of $\mathcal{P}'_x$ would mean that two of those rays intersect in another point meaning that the two rays overlap completely, which is impossible because $\mathcal{P}$ is convex and $x \in Int(\mathcal{P})$.
2. Remark that $\mathcal{P}$ can be decomposed into p triangles in the following way: each triangle has three vertices: x and two vertices of $\mathcal{P}$ connected by an edge. Decompose $\mathcal{P}'_x$ in the same way. Each triangle T of $\mathcal{P}$ is included in one triangle T' of $\mathcal{P}'_x$: the one for which the two vertices of T moved to the two vertices of T'

with the inflation. The inclusion stands because those two vertices have been moved along the sides of triangle T in the opposite direction of the third vertex x.
3. By definition of inflation.
4. A polygon with all vertices on the unit circle and with no self intersection is convex because the angle of each vertex with its two neighbours is strictly less than π.
5. By definition, $\mathcal{P}'_x$ has no more than p vertices. Suppose by contradiction that $\mathcal{P}'_x$ has at least one vertex less than $\mathcal{P}$. Thus, there exists a vertex v' of $\mathcal{P}'_x$ that has been obtained by two different vertices v_1 and v_2 of $\mathcal{P}$. It means that those two vertices lie in the same ray with initial point x. Thus, v_1 is a convex combination of x and v_2, which contradicts with it being an extreme point because $\mathcal{P}$ is convex.
6. By point 2, it is known that $\mathcal{D}\backslash Int(\mathcal{P}') \subset \mathcal{D}\backslash Int(\mathcal{P})$, thus $\min\{||x||_2 \ : \ x \in \mathcal{D}\backslash Int(\mathcal{P}')\} \geq \min\{||x||_2 \ : \ x \in \mathcal{D}\backslash Int(\mathcal{P})\}$, leading to point 6. $\square$

Points 3, 4, 5 and 6 of Proposition 2.3 show that there is a polygon with each vertex on the circle that is solution of (9). Thus, the focus is on the analysis of such polygons. The approximation error is calculated with the point of minimal norm of $\mathcal{D}\backslash Int(\mathcal{P})$. The following proposition exhibits the minimal norm point among the points in $\mathcal{D}\backslash Int(\mathcal{P})$.

Let $\mathcal{P} \in \mathcal{A}_p^-$ be a polygon with vertices on the circle. If $(0,0) \in Int(\mathcal{P})$, then the minimal norm point $x \in \mathcal{D}\backslash Int(\mathcal{P})$ is D the middle point of the longest side of $\mathcal{P}$.

Proof. It is obvious that the minimal norm point is on a side of $\mathcal{P}$ because $(0,0) \in Int(\mathcal{P})$.

Each side of polygon $\mathcal{P}$ has endpoints of norm one, and a computation shows that the minimal norm point of a side with this property is its middle point M. Moreover, by considering the triangle decomposition of a polygon like in proof of point 2 of Proposition 2.3, the norm of M can be expressed with the angle α of the triangle at point $(0,0)$. A computation shows $||M||_2 = \cos\frac{\alpha}{2} =: g(\alpha)$. Moreover, as $(0,0) \in Int(\mathcal{P})$, the angle α is in $]0,\pi[$.

Finally, we have $g'(\alpha) < 0$ on $]0,\pi[$ so the norm of the middle point of a side of $\mathcal{P}$ decreases when the angle α increases, which is equivalent to when the length of the side increases, leading to the result. $\square$

The hypothesis made for Proposition 2.3 is not a limiting one because if the origin $(0,0)$ is not in the interior of $\mathcal{P}$, the approximation error of $\mathcal{P}$ is infinite. Proposition 2.3 shows that the longest side of a polygon with vertices on the circle determines the approximation error of the polygon.

The only solutions of (9) are regular polygons of p sides with all vertices on the circle and their objective value is $\frac{1}{\cos\frac{\pi}{p}} - 1$.

Proof. Let $(\alpha_1, ..., \alpha_p)$ be the angles at $(0,0)$ associated to each of the p triangles of $\mathcal{P}$ as in the proof of Proposition 2.3. A solution of (9) is a polygon with all vertices on the circle that is also solution of the optimization problem

$$\min_{\mathcal{P} \in \mathcal{A}_p^-} \max_{i=1,\ldots,p} \alpha_i, \tag{10}$$

which minimizes the maximum angle α_i of polygon $\mathcal{P}$. The only solution is $\alpha_i = \frac{2\pi}{p}$, $i = 1, ..., p$. Thus the p triangles making the polygon have the same angle at $(0,0)$ and thus they are all similar. Going back to the polygon, it means that every sides have same length. As the polygon is equilateral and inscriptible in a circle, it is a regular polygon with p sides. Finally, the optimal value is $\frac{1}{\cos \frac{\pi}{p}} - 1$ by calculating the norm of the middle point of any side of the regular polygon of p sides with vertices on the circle. □

2.4 Minimality of the Number of Scalar Product Used in the Linearization Scheme

This subsection is dedicated to the proof of Theorem 2.4, which states that the euclidean norm linearization needs the minimum number of scalar products to satisfy a given approximation error threshold ϵ_0 as defined in Definition 2.2. The linearization creates a piecewise linear bounding for a given approximation error threshold that uses the minimal number of pieces. More precisely, we have the following result.

Let $\epsilon_0 > 0$ be the approximation error threshold to be satisfied for the approximation of a disk $\mathcal{D}$ by a polygon. Let $p \in \mathbb{N}$ be such that

$$p = \min\{k \in \mathbb{N}^* : \frac{1}{\cos \frac{\pi}{k}} - 1 \leq \epsilon_0, k > 2\},$$

then a polygon contained in (resp. containing) the disk $\mathcal{D}$ satisfying the approximation error threshold ϵ_0 has at least p sides, and the inner approximation (resp. outer approximation) of the disk with p scalar products given by the euclidean norm linearization produces a polygon with p sides satisfying the approximation error threshold.

The proof of Theorem 2.4 treats separately the outer approximation and the inner approximation. The inner approximation case is first proved, then a justification that the proof scheme is similar for the outer approximation case is given.

A corollary of Proposition 2.3 is needed before the proof of Theorem 2.4 can be stated. The optimal value of (9) decreases with p.

Theorem 2.4. Let $\epsilon_0 > 0$ be the approximation error threshold. Let $\mathcal{D}$ be the unit disk centered at $(0,0)$. Let $p = \min\{k \in \mathbb{N}^* : \frac{1}{\cos \frac{\pi}{k}} - 1 \leq \epsilon_0, k > 2\}$. Then, Proposition 2.3 and Corollary 2.4 give that a polygon satisfying the approximation error threshold ϵ_0 has at least p sides and a regular polygon with p sides satisfies the approximation error threshold ϵ_0. As the inner approximation of a disk with p scalar products yields a regular polygon with p sides, it satisfies the approximation error threshold.

If $\mathcal{D}$ is not the unit disk, an homothetic transformation can transform $\mathcal{D}$ into the unit disk. The same transformation applied to the polygon would change neither its number of sides nor the approximation error of the disk, so the result is valid for any disk.

Finally, the outer approximation case can be treated in the same manner. Indeed, Lemma 2.3 is easily adapted to this case, with a stronger property on the approximation error: $\epsilon(\mathcal{P}) = \epsilon(conv(\mathcal{P}))$. Proposition 2.3 is proved in a similar way, with two substitutions:

- replace the inflation of Definition 2.3 by a *deflation*: the sides of the polygon are moved in the direction of the center of the disk. Indeed, it keep them parallel to their initial position until they are tangent to the circle,
- replace the angles α_i associated to each of the p triangles at point $(0, 0)$ in the proof of Theorem 2.3 by angles directly on the vertices of the deflated polygon.

Finally, the equivalent to Corollary 2.4 as well as the final proof are easily adaptable. □

3 Extensions of the Euclidean Norm Linearization

In this section, two extensions of the euclidean norm linearization are discussed. Moreover, the result of Theorem 2.4 is transposable to the extensions.

The first is the linearization of the euclidean norm of $\mathbb{R}^2$ directly in the objective function. An optimal polyhedron in terms of number of pieces for a given error ϵ is constructed. This result is based on the absolutely homogeneous property of a norm and Theorem 2.4. One application to that linearization is that an arbitrary nonconvex twice differentiable function f can be decomposed in a Difference of Convex function (DC) $f(x) = g(x) - \rho||x||_2$, with $\rho > 0$ sufficiently large [Muts, 2021, Section 3.2]). Thus, it is possible to compute piecewise linear under- and overestimators of f by replacing $||x||_2$ with a piecewise linear function. However, the major drawback of this decomposition is the weakness of the estimators Muts and Nowak [2020]. Another interest is to avoid having to deal with a nondifferentiable function.

The second extension concerns constraints. The linearization is valid with a class of norms of elliptic level set. This result comes from the fact that an ellipse is a linear deformation of a circle, so that an optimal polygon overestimating a disk with a given error can be deformed into an optimal polygon overestimating an ellipse. Indeed, it suffices to apply the linear function that transforms the circle into the ellipse. An application to the beam layout problem is shown in Section 4.

3.1 Linearization in the Objective Function

The euclidean norm linearization can be used to linearize the euclidean norm of $\mathbb{R}^2$ in a constraint, but also in the objective function as shown in this subsection. After introducing the parametric equation of a positive cone, a construction of the optimal piecewise linear bounding of the euclidean norm of $\mathbb{R}^2$ for a given error is shown.

The surface $\mathcal{C}$ generated by the euclidean norm of $\mathbb{R}^2$ is defined as

$$\begin{aligned}\mathcal{C} &= \{(x_1, x_2, x_3) : x_3 = ||(x_1, x_2)||_2, x_1 \in \mathbb{R}, x_2 \in \mathbb{R}\} \\ &= \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1^2 + x_2^2 - x_3^2 = 0 \text{ and } x_3 \geq 0\},\end{aligned} \tag{11}$$

which is a positive cone. Before introducing the bounding of the positive cone, recall the definition of equally spaced unit vectors $\mathbf{u_i}$ for the linearization, with an integer $p > 2$:

$$\mathbf{u_i} = (\cos\frac{2i\pi}{p}, \sin\frac{2i\pi}{p}) \text{ for } i = 1, ..., p, \tag{12}$$

and let $\mathcal{C}_p^+$ and $\mathcal{C}_p^-$ be defined as

$$\mathcal{C}_p^+ = \{(x_1, x_2, f_p^+(x_1, x_2)) : x = (x_1, x_2) \in \mathbb{R}^2\} \text{ and} \tag{13}$$

$$\mathcal{C}_p^- = \{(x_1, x_2, f_p^-(x_1, x_2)) : x = (x_1, x_2) \in \mathbb{R}^2\} \tag{14}$$

with $f_p^+(x_1, x_2)$ and $f_p^-(x_1, x_2)$ defined as follows

$$f_p^+(x_1, x_2) = \max_{i=1,\ldots,p} \frac{(x_1, x_2)^T.\mathbf{u_i}}{\cos\frac{\pi}{p}}, \tag{15}$$

$$f_p^-(x_1, x_2) = \max_{i=1,\ldots,p} (x_1, x_2)^T.\mathbf{u_i}. \tag{16}$$

The surface induced by the inner approximation of p scalar products is denoted $\mathcal{C}_p^+$ because it is an *overestimation* of $\mathcal{C}$ in the sense that for a given $(x_1, x_2) \in \mathbb{R}^2$, x_3 such that $(x_1, x_2, x_3) \in \mathcal{C}$ and x_3^+ such that $(x_1, x_2, x_3^+) \in \mathcal{C}_p^+$, then we have $x_3 \leq x_3^+$.

Similarly, the surface induced by the outer approximation of p scalar products is an *underestimation* of $\mathcal{C}$ and is denoted $\mathcal{C}_p^-$. The terms underestimation, overestimation and bounding for functions are defined more precisely below.

Let $f : \mathbb{R}^2 \to \mathbb{R}^+$ be a function, $F_0 := \{x \in \mathbb{R}^2 : f(x) = 0\}$ be the level set of f for value 0 and f^-, $f^+ : \mathbb{R}^2 \to \mathbb{R}^+$ be functions with level set F_0 of value 0. Function f^- with property

$$\exists\epsilon > 0,\ \forall x \in \mathbb{R}^2 \setminus F_0,\ \frac{f(x) - f^-(x)}{f^-(x)} \in [0, \epsilon] \tag{17}$$

underestimates function f with approximation error ϵ. Function f^+ with property

$$\exists\epsilon > 0,\ \forall x \in \mathbb{R}^2 \setminus F_0,\ \frac{f^+(x) - f(x)}{f(x)} \in [0, \epsilon] \tag{18}$$

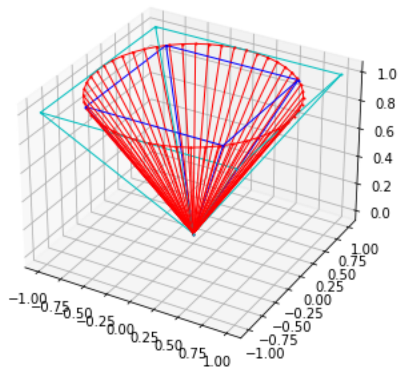

Figure 6: Linear bounding of the positive cone seen from above for $p = 4$ scalar products

overestimates function f with approximation error ϵ. In addition, functions f^- and f^+ with properties (17) and (18) respectively form a *bounding* of function f with approximation error ϵ. Finally, if f^- and f^+ are piecewise linear functions, they form a *linear bounding* of f.

Figure 6 shows a 3D representation of $\mathcal{C}$ in red, $\mathcal{C}_p^-$ in cyan and $\mathcal{C}_p^+$ in blue with $p = 4$.

Let $p \in \mathbb{N}$, $p > 2$. The euclidean norm $||x||_2$ is linearly bounded by f_p^- and f_p^+ with approximation error $\frac{1}{\cos\frac{\pi}{p}} - 1$.

Proof. The proof that f is overestimated by f_p^+ with approximation error $\frac{1}{\cos\frac{\pi}{p}} - 1$ is given here. The inner approximation case is similar. Proving property

$$\frac{f_p^+(x)}{||x||_2} \in [1, \frac{1}{\cos\frac{\pi}{p}}], \quad \forall x \in \mathbb{R}^2 \tag{19}$$

is sufficient since it is equivalent to property (18) for $f^+(x) = f_p^+(x)$, $f(x) = ||x||_2$ and $\epsilon = \frac{1}{\cos\frac{\pi}{p}} - 1$. Remark that $\mathcal{C}_p^+$ is stable by a $\frac{2\pi}{p}$ angle rotation around axis x_3 because the regular polygon with p sides used to construct $\mathcal{C}_p^+$ also has this property. Thus proving property (19) for $x \in X_p = \{(\alpha\cos\beta, \alpha\sin\beta) : \alpha \geq 0, \beta \in [-\frac{\pi}{p}, \frac{\pi}{p}]\}$ is sufficient because $Z_p^+ := \{(x_1, x_2, f_p^+(x_1, x_2)) : (x_1, x_2) \in X_p\}$ is the set of points on one face of the outer approximation $\mathcal{C}_p^+$. Moreover, equation

$$Z_p^+ = \{(\alpha_+\cos\frac{\pi}{p}, \beta_+, \alpha_+) : \alpha_+ \geq 0, \beta_+ \in [-\alpha_+\sin\frac{\pi}{p}, \alpha_+\sin\frac{\pi}{p}]\} \tag{20}$$

shows a parametrization of the face Z_p^+. Vectors of Z_p^+ can all be written uniquely as $\alpha_+ u + \beta_+ v$ with $\alpha_+ > 0$ and $\beta_+ \in [-\alpha_+\sin\frac{\pi}{p}, \alpha_+\sin\frac{\pi}{p}]$. Let

$x \in X_p$. We will show that $1 \leq \frac{f_p^+(x)}{||x||_2} \leq \frac{1}{\cos\frac{\pi}{p}}$. Let $\alpha \geq 0$ and $\beta \in [-\frac{\pi}{p}, \frac{\pi}{p}]$ so that $x = (\alpha\cos\beta, \alpha\sin\beta)$. It is known that $f(x) = \alpha$. To find the point $(\alpha_+\cos\frac{\pi}{p}, \beta_+, \alpha_+)$ of Z_p^+, it suffices to solve the linear system

$$\begin{cases} \alpha\cos\beta = \alpha_+\cos\frac{\pi}{p} \\ \alpha\sin\beta = \beta_+ \end{cases} \tag{21}$$

which has solution

$$\begin{cases} \alpha_+ = \alpha\frac{\cos\beta}{\cos\frac{\pi}{p}} \\ \beta_+ = \alpha\sin\beta. \end{cases} \tag{22}$$

Thus, we have:

$$\frac{f^+(x)}{||x||_2} = \frac{\cos\beta}{\cos\frac{\pi}{p}} \in [1, \frac{1}{\cos\frac{\pi}{p}}] \quad \forall\beta \in [-\frac{\pi}{p}, \frac{\pi}{p}] \tag{23}$$

and the upper bound of the interval $[1, \frac{1}{\cos\frac{\pi}{p}}]$ is reached for $\beta = 0$. Thus, $\epsilon = \frac{1}{\cos\frac{\pi}{p}}$ is the approximation error. □

Let the required approximation error be $\epsilon \in [\frac{1}{\cos\frac{\pi}{p}}, \frac{1}{\cos\frac{\pi}{p-1}}[$, with $p \in \mathbb{N}$, $p > 2$. Then the linear bounding of the euclidean norm by f_p^- and f_p^+ uses the minimum number of pieces.

Proof. The number of sides of the polygon induced by the euclidean norm linearization for a constraint is optimal. Moreover, the linear bounding of the euclidean norm uses the same number of pieces for the same approximation error. As this linear bounding is constructed upon the polygon, it gives a lower bound of the number of pieces for the linear bounding, a lower bound that is reached. □

Thus the euclidean norm linearization can be used to derive a linear bounding of the euclidean norm of $\mathbb{R}^2$ which uses the minimum number of pieces.

3.2 Extension to Elliptic Constraints

As mentionned previously, it is possible to adapt the euclidean norm linearization for norms with ellipses as level sets in constraints. Such constraints are denoted *elliptic constraints*. In this section, a gathering of useful definitions on ellipses is followed by the construction of the linear bounding of elliptic constraints using the minimal number of pieces.

Recall that an ellipse $\mathcal{E}$ with center (u, v), width a, height b (i.e. with no angle between the horizontal axis and the major axis) is a subset of $\mathbb{R}^2$ such that

$$\frac{(x-u)^2}{a^2} + \frac{(y-v)^2}{b^2} = 1 \quad \forall(x, y) \in \mathbb{R}^2. \tag{24}$$

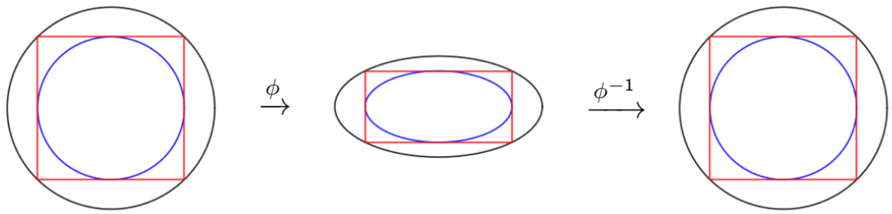

Figure 7: Linear transformation between a circle and an ellipse

By adding an angle θ, width and height are replaced by semi-major axis a and semi-minor axis b and we have

$$\frac{((x-u)\cos\theta-(y-v)\sin\theta)^2}{a^2}+\frac{((x-u)\sin\theta+(y-v)\cos\theta)^2}{b^2}=1 \quad \forall(x,y)\in\mathbb{R}^2. \tag{25}$$

In this subsection, the previous results on the approximation error of the euclidean norm linearization are extended to norms with ellipses as level sets. Such a norm is denoted $||\cdot||_{a,b,\theta}$, and is defined as

$$||x||^2_{a,b,\theta}=\frac{((x_1\cos\theta-x_2\sin\theta)^2}{a^2}+\frac{((x_1\sin\theta+x_2\cos\theta)}{b^2} \quad \forall x=(x_1,x_2)\in\mathbb{R}^2,\ a,b\in\mathbb{R}^+_*. \tag{26}$$

The vocabulary used for ellipses is defined below:

- the shape of an ellipse is denoted (a,b) where a and b are the semi-major axis and semi-minor axis respectively (which means $a\geq b>0$),
- the angle of an ellipse θ refers to the angle between the horizontal axis and the major axis, with $\theta\in[-\frac{\pi}{2},\frac{\pi}{2}]$.

Remark that any non-degenerate ellipse can be obtained by a linear transformation of a circle: a circle with a radius R is a specific ellipse with $a=b=R$ and $\theta=0$. This linear transformation is a change of basis to an orthogonal basis followed by a translation. Without the translation, ϕ is the linear transformation from a circle of radius 1 centered at the origin to an ellipse of shape (a,b), angle θ and centered at the origin:

$$\phi_{a,b,\theta}:\mathbb{R}^2\to\mathbb{R}^2 \tag{27}$$
$$(x,y)\mapsto(ax\cos\theta-by\sin\theta,ax\sin\theta+by\cos\theta). \tag{28}$$

This transformation is linear, its inverse transformation $\phi^{-1}_{a,b,\theta}$ is also linear (see Figure 7) and finally it preserves the norm ratio between two colinear vectors. Thus, the regular polygon with p sides obtained by the euclidean norm linearization is, after transformation by $\phi_{a,b,\theta}$, a polygon that uses the minimum number of sides to satisfy a given approximation error for the ellipse of shape (a, b) and angle θ.

This analysis shows that constraints involving elliptic norms can also be linearized by the euclidean norm linearization, after a suitable change of basis (as well as the linearization of elliptic norms in the objective function). A procedure to linearize a constraint $||x||_{a,b,\theta} \leq \Delta$ is thus to replace it by:

$$||\phi^{-1}_{a,b,\theta}(x)||_{a,b,\theta}.\mathbf{u_i} \leq d\Delta \quad \forall i = 1, ..., p, \tag{29}$$

where d is equal to $\frac{1}{\cos\frac{\pi}{p}}$ for an inner approximation and 1 for an outer approximation. A downside of this linearization is that a, b and θ have to be constants of the model. Otherwise, nonlinear terms are introduced. this means that to linearize a constraint "x is in an ellipse" in the case where the shape of the ellipse and angle are not fixed, a discretization of the possible ellipses parameters is needed. However, this yields many constraints: with n tuples (a, b, θ) and p scalar products, it is necessary to use np big-M constraints, instead of p for the circle, to achieve the same approximation error. Thus, the model is not suited for fine-grain discretization of possible ellipses. However, the next section presents an application for which the proposed linearization is useful.

4 Application to the Beam Layout Problem

The interest of the extension of the euclidean norm linearization is demonstrated on a mixed continous/discrete optimization problem with elliptic constraints. The considered problem is derived from a beam layout problem arising in satellite telecommunications. The original problem has been tackled in Camino et al. [2014, 2016, 2019] and does not involve elliptic constraints. We refer to Camino et al. [2019] for a state-of-the-art review on this problem. Informally, a satellite equipped with a multibeam antenna has to cover a set of end users inside a predefined area on earth by means of a set of movable beams (see Figure 8). In the original problem the projected surface of the beam is circular. Hence, a user is covered if its coordinates are inside one of the beam disks, i.e. if a proximity constraint between the user and the beam center is satisfied. There are also separation constraints between the beam centers depending on discrete beam/reflector assignment constraints as explained in details below. However, in modern multibeam antennas systems the beam shape need not be circular as reconfigurable antennas allow to obtain different beam shapes including elliptic ones Rao et al. [2006]. Adding the possibility of elliptic beam shape in addition to circular ones obviously increases the covering power of the system by a better potential adaptation to the area containing the end users. However, the complexity of proximity and separation constraint modeling is increased, due to the extra constraints required for linearization of ellipses as explained above. The aim of the experimental study carried out in this section is to determine if this complexity increase is compensated by the gain in the objective function

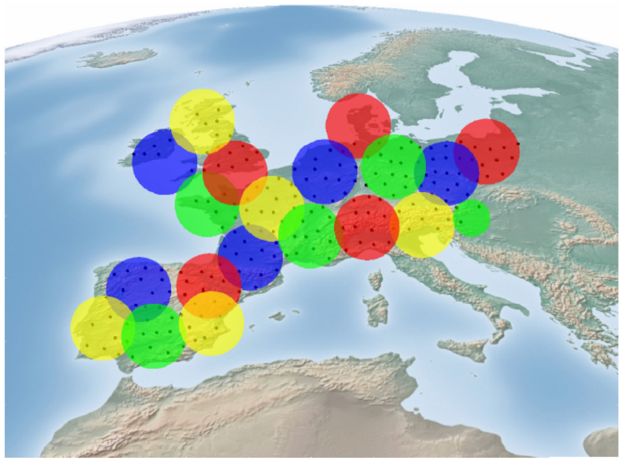

Figure 8: Example of beam layout solution with only circles as beam shape Camino et al. [2019]

for a limited amount of CPU time compared to the model allowing only circular beams.

The first subsection explains the modeling of the beam layout problem, while the second provides the numerical results.

4.1 Problem Definition and Formulation

A multibeam satellite is a telecommunication satellite that uses relatively narrow beams to provide a service to end-users on earth. It has different reflector antennas and each beam is associated to a reflector antenna.

In the considered beam layout problem, a multibeam satellite provides services via beams $b \in B$, each produced by a reflector $r \in R$. On earth, end-users are user stations $s \in S$ of coordinates (X_s, Y_s). They are characterised by their traffic demand $T_s \in \mathbb{R}^+$. They can be covered by a beam of the multibeam satellite. A station s is considered covered by the satellite if one beam b covers the station and transmits a service to it. Each beam covers a portion of the earth in the shape of an ellipse of predefined parameters, with W the radius if this shape is a circle. The goal is to maximise the traffic covered while satisfying the maximum capacity of covering of each beam and some separation constraints coming from technology constraints: each beam of the satellite is associated to a reflector and two beams issued from the same reflector cannot be too close to each other as explained in Camino et al. [2014]. It forces such two beams, enlarged by a factor κ, to not intersect. Beams used in Camino et al. [2019] cover a circular region on earth, but, as mentioned above, it is technically possible to use beams of elliptic shapes. Each beam can be an ellipse of shape $(a_k, \frac{1}{a_k})$, $a_k \geq 1$ with $k \in K$ given, and angle $\theta \in \Theta$ given. This choice of shape ensures that every ellipse covers the same surface on earth, thus no shape of ellipse has an advantaged over another.

A model of the beam layout problem is

$$\max \sum_{b\in B, s\in S} T_s \alpha_{s,b} \tag{30}$$

$$\text{s.t.} \sum_{b\in B} \alpha_{s,b} \le 1 \qquad \forall s \in S, \tag{31}$$

$$\sum_{r\in R} \beta_{b,r} \le 1 \qquad \forall b \in B, \tag{32}$$

$$\sum_{s\in S} \alpha_{s,b} \le |S| \sum_{r\in R} \beta_{b,r} \qquad \forall b \in B, \tag{33}$$

$$\sum_{s\in S} \alpha_{s,b} \ge \sum_{r\in R} \beta_{b,r} \qquad \forall b \in B, \tag{34}$$

$$\sum_{s\in S} T_s \alpha_{s,b} \le \gamma \qquad \forall b \in B \tag{35}$$

$$\sum_{k\in K, \theta\in\Theta} \delta_{b,k,\theta} = 1 \qquad \forall b \in B, \tag{36}$$

$$||(x_b, y_b) - (X_s, Y_s)||_{a_k, \frac{1}{a_k}, \theta} \le W + (2 - \alpha_{s,b} - \delta_{b,k,\theta}) M_s$$

$$\forall s \in S, \forall b \in B, \forall k \in K, \forall \theta \in \Theta, \tag{37}$$

$$\gamma_{b,b'} \ge \textstyle\sum_{r'\in R} \beta_{b,r'} + \sum_{r'\in R} \beta_{b',r'} + \beta_{b,r} + \beta_{b',r} - 3$$

$$\forall b, b' \in B, b' > b, \forall r \in R, \tag{38}$$

$$||(x_b, y_b) - (x_{b'}, y_{b'})||_{a_k, \frac{1}{a_k}, \theta} \ge 2\kappa W - N(1 - \gamma_{b,b'})$$

$$\forall b, b' \in B, b' > b, \tag{39}$$

$$x_b, y_b \ge 0 \qquad \forall b \in B, \tag{40}$$

$$\alpha_{s,b} \in \{0,1\} \qquad \forall s \in S, \forall b \in B, \tag{41}$$

$$\beta_{b,r} \in \{0,1\} \qquad \forall b \in B, \forall r \in R, \tag{42}$$

$$\gamma_{b,b'} \in \{0,1\} \qquad \forall b, b' \in B, b' > b, \tag{43}$$

$$\delta_{b,k,\theta} \in \{0,1\} \qquad \forall b \in B, \forall k \in K, \forall \theta \in \Theta. \tag{44}$$

The objective (30) is the maximisation of the covered traffic, with $\alpha_{s,b}$ a binary variable equal to 1 if user station s is covered by beam b. Constraint (31) forces a station to be covered by at most one beam. Constraint (32) associates at most one reflector antenna r to each beam b due to the binary variable $\beta_{b,r}$. Constraints (33) and (34) force an unused beam to be associated to no stations and a used beam to be associated to at least one station. Constraint (35) ensures that each beam cannot cover more than a traffic of γ, the capacity of a beam. Constraint (36) associates a shape $(a_k, \frac{1}{a_k})$ and an angle $\theta \in \Theta$ to a beam b

through the use of the binary variable $\delta_{b,k,\theta}$. Constraint (37) checks that each station s is in the beam b of center (x_b, y_b) to which it is affected. M_s denote the big-M constant associated and $||\cdot||_{a_k,\frac{1}{a_k},\theta}$ the norm defined in (26) with level set of value 1 the ellipse of shape $(a_k, \frac{1}{a_k})$ and angle θ. Constraint (38) enforces that the binary variable $\gamma_{b,b'}$ is equal to 1 if beams b and b' come from the same reflector antenna. Constraint (39) ensures that two beams associated to the same reflector antenna b and b' are not too close. Finally, (40)-(44) define variable types.

The details of the linearization into an MILP model are discussed here. The nonlinearity comes from constraints (37) and (39). As explained in (29), the linearization procedure is to apply a suitable linear transformation and then apply the euclidean norm linearization. Moreover, constraint (38) needs to be adapted to the linearization. Thus, the only constraints that change in the linearization are (37), (38) and (39).

Let p be the number of scalar products used to linearize, according to an error approximation ϵ wanted. Let $\phi^{-1}_{a_k,\frac{1}{a_k},\theta}$ be the linear transformation defined in (29). The linearization of (37) with the euclidean norm linearization is

$$\forall s \in S, \forall b \in B, \forall k \in K, \forall \theta \in \Theta, \forall i = 1, .., p$$

$$\phi_{a_k,\frac{1}{a_k},\theta}\left(\begin{pmatrix} x_b - X_s \\ y_b - Y_s \end{pmatrix}\right) \cdot \begin{pmatrix} \cos\frac{2\pi i}{p} \\ \sin\frac{2\pi i}{p} \end{pmatrix} \leq W \cos\frac{\pi}{P} + (2 - \alpha_{s,b} - \delta_{b,k,\theta})M_s, \tag{45}$$

where i denotes the index of vector u_i used for the scalar product. The right part contains $\cos\frac{\pi}{p}$ to ensure feasibility, which corresponds to the inner approximation of the disk.

Constraint (39) is not as easy to linearize. Indeed, this is not a convex constraint: it is about ensuring that two ellipses do not intersect. In the case of ensuring that two disks do not intersect (created by two beams of circle shape), it suffices to evaluate the distance between the two centers with the euclidean norm linearization and then ensure with big-M constraints that this distance is higher than what is necessary. For two ellipses, this linearization scheme does not work. The convexity of an ellipse allows to model the intersection with linear separation in a similar way as in Kallrath and Rebennack [2014].

Let A,$B \subset \mathbb{R}^2$. The sets A and B are linearly separated if $\exists v \in \mathbb{R}^2$, $\exists b \in \mathbb{R}$: $\forall x \in A$, $\forall y \in B : x.v \leq b \leq y.v$. The linear separation is illustrated in Figure 9: the two octogons are linearly separated because a line d can be drawn between them, which corresponds to the hyperplane $d = \{u \in \mathbb{R}^2 : u.v = b\}$. The ellipses approximations are convex polygons so it suffices to test the linear separation with the vertices of those polygons. The last problem for the linearization is that making scalar products between variable vectors is a quadratic operation. To adress this issue, the linear separation is tested only for a finite number N_{dir} of hyperplanes of normal vectors $v_j = (\cos\frac{2j\pi}{p}, \sin\frac{2j\pi}{p})$ for $j = 1, ..., N_{dir}$. This means that a stronger version of the linear separation is ensured in the model,

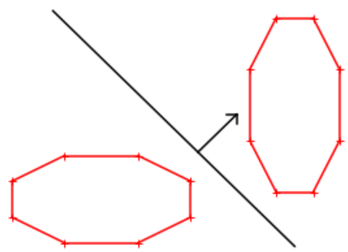

Figure 9: Linear separation of two ellipses outer approximations by the black line

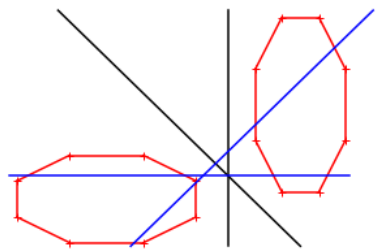

Figure 10: The linear separation is working with the two black lines and is not working with the two blue ones

but with a reasonably high value for N_{dir}, it is probably sufficiently good in a comparison with the others approximations of the model.

Constraint (38) becomes constraints

$$\sum_{j=1}^{N_{dir}} \gamma_{b,b',j} \geq \sum_{r \in R} \beta_{b,r} + \sum_{r \in R} \beta_{b',r} + \beta_{b,r} + \beta_{b',r} - 3 \tag{46}$$

$$\forall r \in R, \forall b, b' \in B \text{ such that } b' > b \tag{47}$$

where the new binary variables $\gamma_{b,b',j}$ replace $\gamma_{b,b'}$ because for the linearization, selecting which line j is associated to v_j for $j = 1, ..., N_{dir}$ ensures the linear separation. Finally the linearization of (39) becomes

$$\begin{pmatrix} S_i^{x,b} \\ S_i^{y,b} \end{pmatrix} \cdot \begin{pmatrix} \cos \frac{2\pi k}{N_{dir}} \\ \sin \frac{2\pi k}{N_{dir}} \end{pmatrix} \leq w_{b,b'} + N(1 - \gamma_{b,b',k}) \tag{48}$$

$$\begin{pmatrix} S_i^{x,b'} \\ S_i^{y,b'} \end{pmatrix} \cdot \begin{pmatrix} \cos \frac{2\pi k}{N_{dir}} \\ \sin \frac{2\pi k}{N_{dir}} \end{pmatrix} \geq w_{b,b'} + N(1 - \gamma_{b,b',k}) \tag{49}$$

$$\forall b, b' \in B \text{ such that } b' \neq b, \forall i = 1, ..., P, \forall k = 1, ..., N_{dir} \tag{50}$$

where the coordinates of the vertex i of the outer approximation of beam b are $(S_i^{x,b}, S_i^{y,b})$. The outer approximation of the disk is used to ensure feasibility. The value N is used as big-M constant of the constraint.

4.2 Computational Experiments

The goal of this section is to evaluate the potential of the extension of the euclidean norm linearization to ellipses with the beam layout problem. The description of instances [1] as well as the software and hardware used for the numerical experiments are described in §4.2.1. Then, a comparison between the MILP models resulting from the linearizations described previously and two MINLP models is made to show the linearization interest. Finally, the euclidean norm linearization (CL for Circle Linearization) and its extension to ellipses (EL for Ellipse Linearization) are compared on the beam layout problem.

4.2.1 Description of instances and experimental set-up

A set of 100 instances is considered, corresponding to different sizes, different number of scalar products p used in the linearizations and different densities. There are five different sizes: the number of stations $|S|$ varies in $\{20, 40, 60, 80, 100\}$ with associated number of beams $|B| = \min\{\frac{|S|}{8}, 10\}$ which is $|B| \in \{3, 5, 7, 10\}$. As for p, it is taken from the set $\{4, 8, 12, 16, 20\}$. Concerning densities, two different values $d \in \{30, 70\}$ are used. Thus there are 10 different instances for each couple (size, density).

For each instance, the stations positions are drawn randomly from a set of 157 positions from a real instance and there are three reflector antennas so $R = \{1, 2, 3\}$. The ellipses allowed in the *EL* models are $\phi_{a,\frac{1}{a},\theta}$ with $a \in A = \{\frac{1}{2}, 1, 2\}$ and $\theta \in \Theta = \{\frac{2\pi}{3}, \frac{4\pi}{3}, 1\}$. Each ellipse allowed is visible on Figure 11. Moreover, the capacity γ is fixed to 500 in all instances. It is a major limitation for the objective value as the naive upper bound induced by the capacity is pretty close to the objective value in many instances.

Examples of solutions are plotted in Figure 12 and 13. Black points are stations not covered, colored points are covered stations, and the color indicates the reflector antenna which possesses the beam covering the station. The main advantage of using ellipses beam shapes is that there are much more possibilities to cover stations with the same number of beams compared to only circles beam shapes.

Due to the size of instances, the parameter N_{dir} has a really low impact on both time and objective value, so for each instance it is set to 10.

For *CL* and *EL* models, each instance is solved on one core of a Xeon E5-2695 v3 @ 2.30GHz CPU with a RAM limit of 3.5 Go. The solver used is CPLEX 12.9 with the global thread count parameter set to 1. As for the resolution of the MINLP model, the NEOS server Czyzyk et al. [1998], Dolan [2001], Gropp and Moré [1997] with the BARON 21.1.13 solver were used, with a RAM of 3 Go, a single thread, an absolute gap tolerance of 10^{-6} and a relative gap tolerance of 10^{-4} to match up with CPLEX. NEOS server CPUs used in our experiments are different to the one used for *CL* and *EL* models, but as we will see, the results make it clear that a difference in CPU is not sufficient to make up for the worse performance of the MINLP models. The two MINLP

[1] All data and models are available upon request to the corresponding author, A. Duguet.

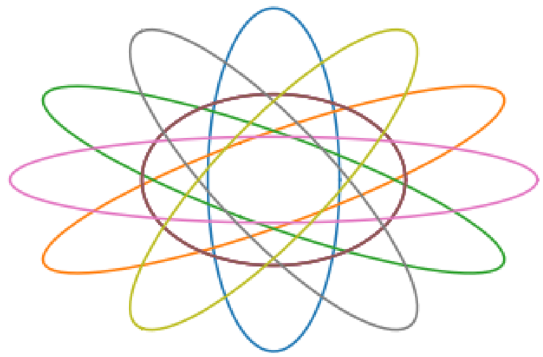

Figure 11: Beam shapes allowed for the instances solved with the *EL* linearization

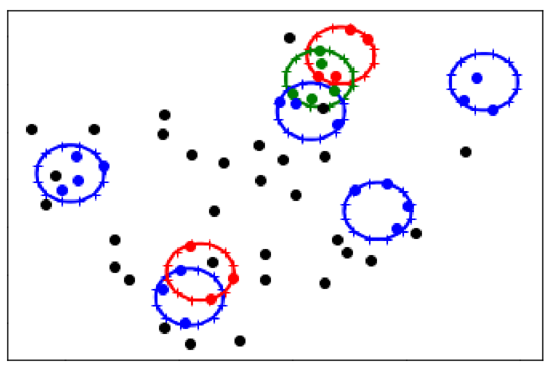

Figure 12: Example of a feasible solution with the circle model

models tested are denoted BARON-C for the beam layout problem where beam shapes are only circles, and BARON-E for beam shapes with ellipses. They are created according to Equations (30)-(44), with the obvious simplifications for BARON-C because only circle beam shapes are allowed.

In Table 1, the number of binary variables, continuous variables, linear constraints as well as nonlinear constraints for all models are displayed. In addition, the number of binary variables, continuous variables and constraints are shown for some MILP models with p the number of scalar products used for the linearization. It can be seen that the difference of constraints between MINLP models and MILP models, as well as between models allowing circle beam shapes and ellipse beam shapes is important. It can be explained by the constraints introduced during the linearization of constraints (37) into (45) and (39) into (46)-(47). Comparatively, the number of binary variables appear less fluctuating than the number of constraints across models.

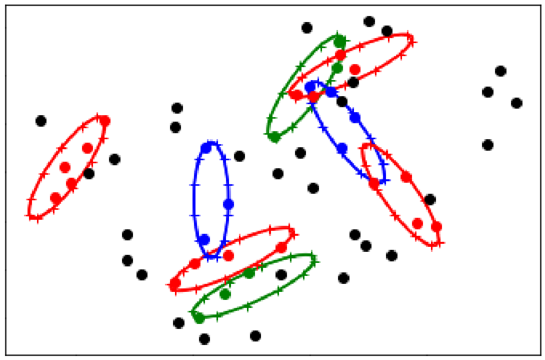

Figure 13: Example of a feasible solution with the ellipse model

	BARON-C				BARON-E			
$\|S\|$	#bin	#var	#Lcons	#NLcons	#bin	#var	#Lcons	#NLcons
20	78	6	41	63	105	6	44	567
40	240	10	90	210	285	10	95	1890
60	568	16	176	508	640	16	184	4572
80	930	20	255	845	1020	20	265	7605
100	1130	20	275	1045	1220	20	285	9405

	CL - $p = 4$			CL - $p = 20$			EL - $p = 4$			EL - $p = 20$		
$\|S\|$	#bin	#var	#cons	#bin	#var	#cons	#bin	#var	#cons	#bin	#var	#cons
20	81	6	293	129	6	1301	126	15	2444	126	15	12044
40	255	10	930	415	10	4290	360	35	8095	360	35	40095
60	616	16	2208	1064	16	10336	856	80	19704	856	80	97784
80	1010	20	3635	1730	20	17155	1370	120	32665	1370	120	162265
100	1210	20	4455	1930	20	21175	1570	120	39885	1570	120	198285

Table 1: Number of variables, binary variables and constraints for different models

4.2.2 Comparison with a MINLP Solver

All results displayed in Tables 2, 3 and 4 are limited in time by 3600 seconds. The results for instances with densities 30 and 70 are aggregated, as well as the 10 instances with same parameter $|S|$, which means that instances are grouped in batches of 20.

Table 2 compares instance by instance MILP models to MINLP models in terms of objective value. There are 10 MILP models: the use of circles (*CL*) or ellipses (*EL*) associated to the number of scalar product $p \in \{4, 8, 12, 16, 20\}$ of the linearization. There are two MINLP models, which are BARON-C and BARON-E. Each cell of the table shows two numbers. The first one is the number of times (out of the 20 instances) the objective value of the MILP model is greater or equal to the MINLP model, while the second one is the number of times the objective value of the MINLP model is greater or equal to the MILP model. The bigger of those two numbers is in bold to help visualize which of the MILP model or the MINLP model has the greatest number of best objective value. The *EL* models have a greater number of best objective value for each batch. As for the *CL* models, for 7 batches out of 50 they have a greater number of best objective value, namely 6 for $|S| = 20$ and 1 for $|S| = 40$. Thus the MINLP models with BARON are not competitive with the MILP models with CPLEX in terms of best objective values, and this difference is accentuated when the size of the instances increases.

In Table 3, the number of instances solved to optimality for each model is displayed. It shows that BARON-E is unable to solve any instance to optimality. Moreover, the *EL* models and BARON-C only solve instances with the smallest number of stations to optimality while *CL* models can solve to optimality some instances with 40 stations. It means that the MILP models have a greater potential to find an optimal solution than the MINLP models.

Finally, Table 4 shows the number of instances with no integer solution with strictly positive objective value (non trivial solution). For most instances with at least 60 stations, BARON-C cannot find non trivial integer solutions, while BARON-E only finds positive objective value for the smallest size of instances. It proves that the MINLP model is too hard to be solved within one hour for both BARON-C and BARON-E. In comparison, *CL* and *EL* models find a positive objective value for most instances no matter the size. Thus the MILP models are the most reliable to get a nontrivial solution to the beam layout problem.

To conclude, our comparison between MINLP models using the BARON solver and MILP models with CPLEX shows that a MILP approximation of the beam layout is a viable option for its resolution because the exact MINLP models are too big to be handled comparatively to our MILP models.

4.2.3 Comparison between the Ellipse Linearization and the Circle Linearization

Each one of the 100 instances has been solved using the two different linearizations, *CL* *CL* and *EL*, for three different maximum computation times which

		$\|S\| = 20$		$\|S\| = 40$		$\|S\| = 60$	
model	p	BARON-C	BARON-E	BARON-C	BARON-E	BARON-C	BARON-E
CL	4	0 / **20**	8 / **12**	7 / **13**	**20** / 0	**19** / 1	**20** / 0
	8	4 / **20**	**12** / 8	**14** / 6	**20** / 0	**19** / 3	**20** / 3
	12	12 / **20**	**15** / 6	**17** / 4	**20** / 0	**20** / 3	**20** / 3
	16	15 / **20**	**15** / 6	**19** / 3	**20** / 0	**19** / 5	**20** / 5
	20	15 / **20**	**15** / 5	**20** / 2	**20** / 0	**19** / 3	**20** / 3
EL	4	**15** / 5	**20** / 0	**13** / 7	**20** / 0	**20** / 0	**20** / 0
	8	**18** / 2	**20** / 0	**19** / 1	**20** / 0	**20** / 0	**20** / 0
	12	**18** / 2	**20** / 0	**20** / 0	**20** / 0	**20** / 0	**20** / 0
	16	**19** / 1	**20** / 0	**20** / 0	**20** / 0	**20** / 0	**20** / 0
	20	**19** / 1	**20** / 0	**20** / 0	**20** / 0	**20** / 0	**20** / 0

		$\|S\| = 80$		$\|S\| = 100$	
model	p	BARON-C	BARON-E	BARON-C	BARON-E
CL	4	**20** / 0	**20** / 0	**20** / 1	**20** / 1
	8	**19** / 1	**20** / 1	**19** / 3	**20** / 3
	12	**19** / 3	**20** / 3	**18** / 3	**20** / 3
	16	**20** / 0	**20** / 0	**20** / 3	**20** / 3
	20	**20** / 3	**20** / 3	**19** / 2	**20** / 2
EL	4	**20** / 0	**20** / 0	**20** / 0	**20** / 0
	8	**20** / 0	**20** / 0	**20** / 0	**20** / 0
	12	**20** / 0	**20** / 0	**20** / 0	**20** / 0
	16	**20** / 0	**20** / 0	**20** / 0	**20** / 0
	20	**20** / 0	**20** / 0	**20** / 0	**20** / 0

Table 2: Comparison between linearization and BARON models: number of instances with a greater objective value

$\|S\|$	BARON-C	BARON-E	CL					EL				
			4	8	12	16	20	4	8	12	16	20
20	20	0	20	20	20	20	20	20	20	17	13	16
40	0	0	20	20	18	15	12	0	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0	0	0
80	0	0	0	0	0	0	0	0	0	0	0	0
100	0	0	0	0	0	0	0	0	0	0	0	0

Table 3: Number of instances solved to optimality for the different models

$\lvert S \rvert$	BARON-C	BARON-E	CL					EL				
			4	8	12	16	20	4	8	12	16	20
20	0	3	0	0	0	0	0	0	0	0	0	0
40	0	20	0	0	0	0	0	0	0	0	0	0
60	17	20	0	3	3	5	3	0	0	0	0	0
80	17	20	0	1	3	0	3	0	0	0	0	0
100	14	20	1	3	3	3	2	0	0	0	0	0

Table 4: Number of instances with no strictly positive objective value for the different models

are 600, 1800 and 3600 seconds, as well as for each parameter p. Looking specifically to the instances solved to optimality of *CL* and *EL* in Table 3, there are few instances solved to optimality due to the size of the optimization problems tackled, and *EL* produces bigger optimization problems than *CL* so less instances solved with *EL* are solved to optimality than for *CL*. There are only 17 % of all instances solved to optimality with *EL*, and all those instances have 20 stations. As for instances solved with *CL*, 37 % are solved to optimality, with all 20 stations instances, and 85 % of 40 stations instances. None of the others are solved to optimality. Thus, the following comparison should only be taken as a practical comparison and not as a comparison of optimal values since most of the instances are not solved to optimality.

Results of the resolution of the instances of the MILP models for low and high densities respectively are shown in Figures 14 and 15. Each point represents a relative gap of a mean over 10 instances from a naive upper bound, in percentage. The gap is calculated with formula $100\frac{M-\gamma|B|}{M}$ with M the mean objective value and $\gamma|B|$ the naive upper bound derived from the maximum capacity $|S|$ a beam can cover. This naive upper bound is plotted in the dashed black line. The points linked represent values obtained from instances with the same model, the same time limit and the same number of stations $|S|$, thus there are five different series of plots, representing the 5 different numbers of stations used.

There are some visible facts on Figures 14 and 15. For a high number of stations, *EL* really benefit from extra computation time whereas *CL* does not need more than 600 seconds. So, when the size of the instance does not exceed 40 stations and when sufficient CPU time is allowed, *EL* outperforms *CL*. Another noticeable trend is that high density instances are far closer to the naive upper bound (the majority of the solutions are between 0 and 5 % away from it), while low density instances are seldom at less than 5 % of difference. As before, the combinatorics of the models is the key to understand that point: higher density leads to more possible covered stations by a single beam. It is also visible that the objective value obtained with *CL* are lower than for *EL* for small number of stations, and higher for a high number. It can be explained by the size of *EL* optimization problems that become unmanageable faster than for *CL* when the number of stations increases. Indeed, the number of binary variables and

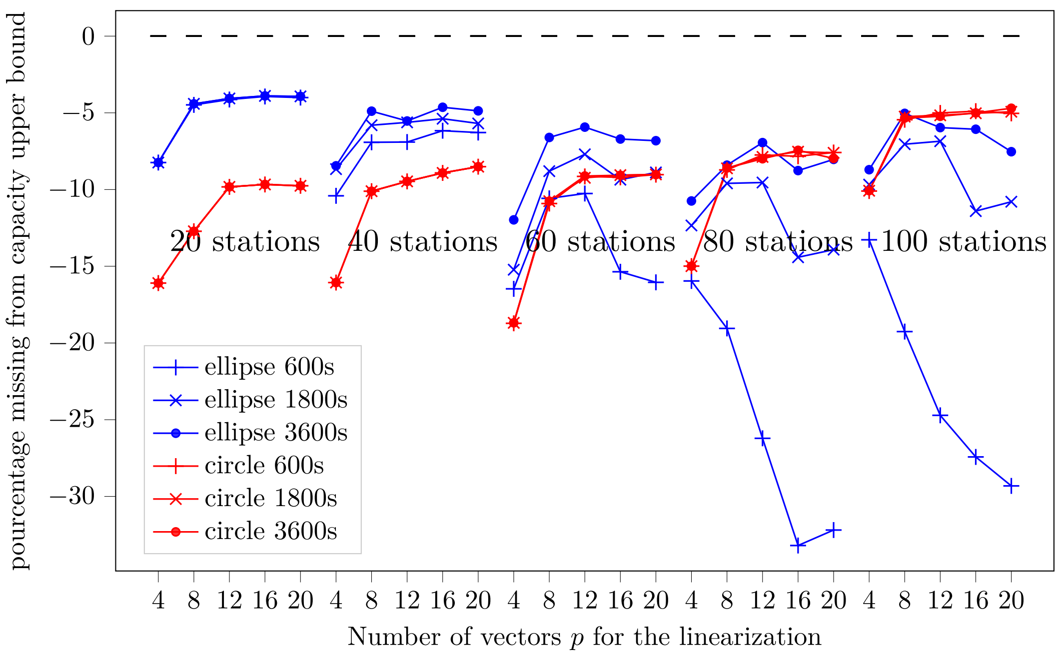

Figure 14: Results for low density instances

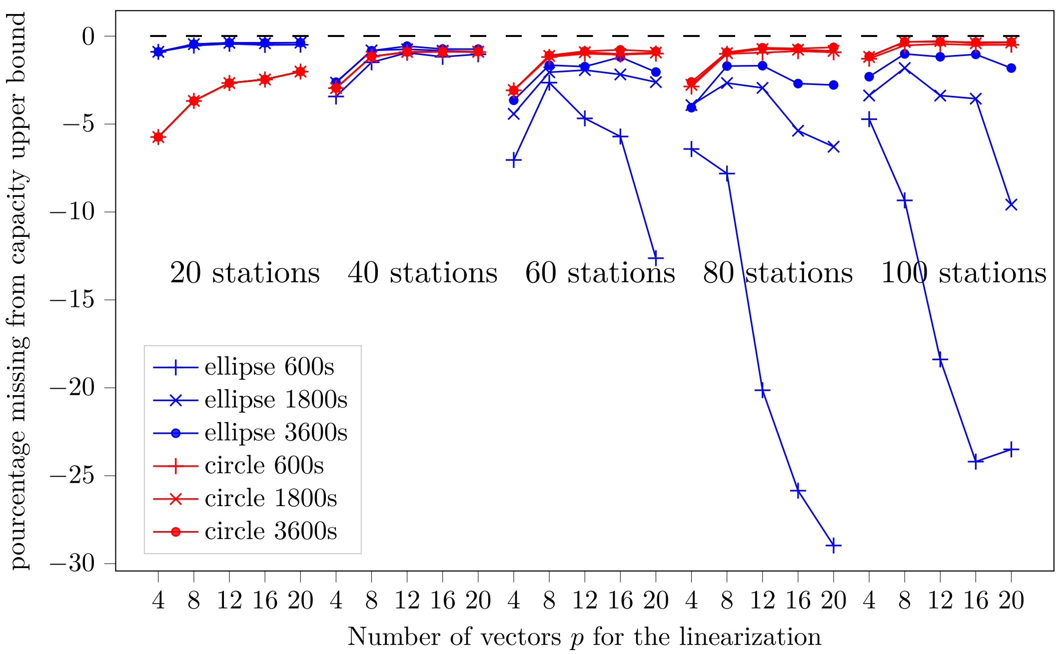

Figure 15: Results for high density instances

d \ time (s)	600	1800	3600
30	40	60	80
70	20	40	40

Table 5: Maximal number of stations for which *EL* has a better mean objective function than CL

model \ p	4	8	12	16	20
EL	4	5	12	7	2
CL	0	1	4	11*	16*

Table 6: Number of best means depending on the number of scalar product P used

constraints increases faster in the *EL* models than in the *CL* model. It is due to the different orientations and elongations allowed for ellipses.

Table 5 shows the maximum number of stations for which *EL* gives better mean results than *CL*, depending on the density and the time limit. For high density, *EL* are the best for 40 stations and less, while it is better than *CL* with up to 80 stations for a low density and a time limit of 3600 seconds. This can be interpreted by considering that the elliptic beams are more efficient to capture users spread in less dense areas, while circular beams are sufficient for highly dense areas. This would also suggest an adaptative selection of elliptic or circular beams depending on the density of the subareas to be covered.

Table 6 gives indications on the impact of the number of scalar product p for the linearization. The number of times a certain value of p gives better mean results than the other values of p depending on the model is shown. Character $*$ expresses that the results for two values of 16 and 20 scalar products are the same. It is shown that the increase in size of *EL* optimization problems makes it less efficient for a too high number of scalar products p.

Overall this application to the beam layout problem shows that, under the above-described alloted CPU time and problem size limit conditions, the proposed extension of the euclidean norm linearization to elliptic constraints is manageable and can allow a significant gain in practical applications, even with respect to a direct MINLP model of the problem.

5 Conclusion

This work on the linearization of non-convex MINLP proposed theoretical and practical results based on the euclidean norm linearization. A linearization of the euclidean norm based on the scalar product between regularly spaced unit vectors and the vector that needs to be measured has been developed. From a theoretical point of view, the guarantee that given any approximation error the linearization method uses the minimum number of pieces has been proven, which means that the model uses polygons with the least number of sides to approximate a circle with a modulable and measurable approximation of nonlinear constraints. Moreover, it has been shown that linearizing elliptic constraints while keeping the previously mentioned guarantee is achievable. Thus the linearization method can be applied to a wider range of problems. As for the practical point of view, the elliptic constraint linearization increases the model complexity with respect to the euclidean norm linearization, which makes this linearization useful under CPU time and problem size limit conditions. This work can be continued in many topics. From the theoretical side, one may focus on MINLP linearization and approximation guarantees, for example for piecewise linearization of more general functions of two variables, with the euclidean norm of $\mathbb{R}^2$ being a case study. From the practical point of view, heuristics and especially matheuristics could be derived from the proposed linearization scheme to handle more efficiently larger problems.

References

W. P. Adams and H. D. Sherali. A tight linearization and an algorithm for zero-one quadratic programming problems. *Management Science*, 32(10):1274–1290, 1986. doi: 10.1287/mnsc.32.10.1274.

A. Borghetti, C. D'Ambrosio, A. Lodi, and S. Martello. An MILP approach for short-term hydro scheduling and unit commitment with head-dependent reservoir. *IEEE Transactions on Power Systems*, 23(3):1115–1124, 2008. doi: 10.1109/TPWRS.2008.926704.

R. Burlacu, B. Geißler, and L. Schewe. Solving mixed-integer nonlinear programmes using adaptively refined mixed-integer linear programmes. *Optimization Methods and Software*, 35(1):37–64, 2020. doi: 10.1080/10556788.2018.1556661.

J.-T. Camino, S. Mourgues, C. Artigues, and L. Houssin. A greedy approach combined with graph coloring for non-uniform beam layouts under antenna constraints in multibeam satellite systems. In *2014 7th Advanced Satellite Multimedia Systems Conference and the 13th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, pages 374–381, 2014. doi: 10.1109/ASMS-SPSC.2014.6934570.

J.-T. Camino, C. Artigues, L. Houssin, and S. Mourgues. Mixed-integer linear programming for multibeam satellite systems design: Application to the beam layout optimization. In *2016 Annual IEEE Systems Conference (SysCon)*, pages 1–6, 2016. doi: 10.1109/SYSCON.2016.7490613.

J.-T. Camino, C. Artigues, L. Houssin, and S. Mourgues. Linearisation of euclidean norm dependent inequalities applied to multibeam satellites design. *Computational Optimization and Applications*, 2019. doi: 10.1007/s10589-019-00083-z.

E. Camponogara, M. P. de Castro, A. Plucenio, and D. J. Pagano. Compressor scheduling in oil fields. *Optimization and Engineering*, 12:153–174, 2011. doi: 10.1007/s11081-009-9093-3.

J. Czyzyk, M. P. Mesnier, and J. J. Moré. The neos server. *IEEE Journal on Computational Science and Engineering*, 5(3):68 – 75, 1998.

C. D'Ambrosio, A. Lodi, and S. Martello. Piecewise linear approximation of functions of two variables in MILP models. *Operations Research Letters*, 38(1):39–46, 2010. doi: 10.1016/j.orl.2009.09.005.

E. D. Dolan. The neos server 4.0 administrative guide. Technical Memorandum ANL/MCS-TM-250, Mathematics and Computer Science Division, Argonne National Laboratory, 2001.

J. G. Dunham. Optimum uniform piecewise linear approximation of planar curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(1):67–75, 1986. doi: 10.1109/TPAMI.1986.4767753.

B. Geißler, A. Martin, A. Morsi, and L. Schewe. Using piecewise linear functions for solving MINLPs. In J. Lee and S. Leyffer, editors, *Mixed Integer Nonlinear Programming*, pages 287–314, New York, NY, 2012. Springer New York. ISBN 978-1-4614-1927-3.

W. Gropp and J. J. Moré. Optimization environments and the neos server. In M. D. Buhman and A. Iserles, editors, *Approximation Theory and Optimization*, pages 167 – 182. Cambridge University Press, 1997.

R. B. Hughes and M. R. Anderson. Simplexity of the cube. *Discrete Mathematics*, 158(1):99–150, 1996. doi: 10.1016/0012-365X(95)00075-8.

J. Kallrath and S. Rebennack. Cutting ellipses from area-minimizing rectangles. *Journal of Global Optimization*, 59:405–437, 2014. doi: 10.1007/s10898-013-0125-3.

A. B. Keha, I. R. de Farias, and G. L. Nemhauser. Models for representing piecewise linear cost functions. *Operations Research Letters*, 32(1):44–48, 2004. doi: 10.1016/S0167-6377(03)00059-2.

L. Liberti, S. Cafieri, and F. Tarissan. Reformulations in mathematical programming: A computational approach. In A. Abraham, A.-E. Hassanien, P. Siarry, and A. Engelbrecht, editors, *Foundations of Computational Intelligence Volume 3: Global Optimization*, pages 153–234, Berlin, Heidelberg, 2009a. Springer Berlin Heidelberg. ISBN 978-3-642-01085-9. doi: 10.1007/978-3-642-01085-9_7.

L. Liberti, N. Maculan, and Y. Zhang. Optimal configuration of gamma ray machine radiosurgery units: The sphere covering subproblem. *Optimization Letters*, 3:109–121, 2009b. doi: 10.1007/s11590-008-0095-4.

L. S. Liberti. *Reformulation and Convex Relaxation Techniques for Global Optimization.* PhD thesis, Imperial College London, 2004.

R. Misener, C. E. Gounaris, and C. A. Floudas. Global optimization of gas lifting operations: A comparative study of piecewise linear formulations. *Industrial & Engineering Chemistry Research*, 48(13):6098–6104, 2009. doi: 10.1021/ie8012117.

P. Muts. *Decomposition methods for mixed-integer nonlinear programming.* PhD thesis, 07 2021.

P. Muts and I. Nowak. Towards multi-tree methods for large-scale global optimization. In H. A. Le Thi, H. M. Le, and T. Pham Dinh, editors, *Optimization of Complex Systems: Theory, Models, Algorithms and Applications*, pages 498–506, Cham, 2020. Springer International Publishing. ISBN 978-3-030-21803-4. doi: 10.1007/978-3-030-21803-4_50.

S. U. Ngueveu. Piecewise linear bounding of univariate nonlinear functions and resulting mixed integer linear programming-based solution methods. *European Journal of Operational Research*, 275(3):1058–1071, 2019. doi: 10.1016/j.ejor.2018.11.021.

I. Nowak. *Relaxation and Decomposition Methods for Mixed Integer Nonlinear Programming*. Birkhäuser, 2005. doi: 10.1007/3-7643-7374-1.

M. Padberg. Approximating separable nonlinear functions via mixed zero-one programs. *Operations Research Letters*, 27(1):1–5, 2000. doi: 10.1016/S0167-6377(00)00028-6.

S. Rao, M. Tang, and C.-C. Hsu. Multiple beam antenna technology for satellite communications payloads. *ACES Journal*, 21(3):1054–4887, 2006. doi: 10.2514/6.2007-3179.

S. Rebennack and J. Kallrath. Continuous piecewise linear delta-approximations for bivariate and multivariate functions. *J Optim Theory Appl*, 167:102–117, 2015a. doi: 10.1007/s10957-014-0688-2.

S. Rebennack and J. Kallrath. Continuous piecewise linear delta-approximations for univariate functions: Computing minimal breakpoint systems. *J Optim Theory Appl*, 167:617–643, 2015b. doi: 10.1007/s10957-014-0687-3.

S. Rebennack and V. Krasko. Piecewise linear function fitting via mixed-integer linear programming. *INFORMS Journal on Computing*, 32(2):507–530, 2020. doi: 10.1287/ijoc.2019.0890.

J. Rosen and P. Pardalos. Global minimization of large-scale constrained concave quadratic problems by separable programming. *Mathematical Programming*, 34:163–174, 1986. doi: 10.1007/BF01580581.

R. Rovatti, C. D'Ambrosio, A. Lodi, and S. Martello. Optimistic MILP modeling of non-linear optimization problems. *European Journal of Operational Research*, 239(3):32–45, 2014. doi: 10.1016/j.ejor.2014.03.020.

A. Santoyo-González and C. Cervelló-Pastor. Latency-aware cost optimization of the service infrastructure placement in 5g networks. *Journal of Network and Computer Applications*, 114:29–37, 2018. doi: 10.1016/j.jnca.2018.04.007.

H. D. Sherali and W. P. Adams. *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*. Kluwer Academic Publishers, Dodrecht, 1999.

H. D. Sherali and L. Liberti. Reformulation-linearization technique for global optimization. In C. A. Floudas and P. M. Pardalos, editors, *Encyclopedia of Optimization*, pages 3263–3268, Boston, MA, 2009. Springer US. ISBN 978-0-387-74759-0. doi: 10.1007/978-0-387-74759-0_559.

T. L. Silva and E. Camponogara. A computational analysis of multidimensional piecewise-linear models with applications to oil production optimization. *European Journal of Operational Research*, 232(3):630–642, 2014. doi: 10.1016/j.ejor.2013.07.040.

E. Smith and C. Pantelides. A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of nonconvex MINLPs. *Computers & Chemical Engineering*, 23(4):457–478, 1999. doi: 10.1016/S0098-1354(98)00286-5.

W. D. Smith. A lower bound for the simplexity of then-cube via hyperbolic volumes. *European Journal of Combinatorics*, 21(1):131–137, 2000. doi: 10.1006/eujc.1999.0327.

S. Sridhar, J. Linderoth, and J. Luedtke. Locally ideal formulations for piecewise linear functions with indicator variables. *Operations Research Letters*, 41(6):627–632, 2013. doi: 10.1016/j.orl.2013.08.010.

F. Tardella. Existence and sum decomposition of vertex polyhedral convex envelopes. *Optimization Letters*, 2:363–375, 2007. doi: 10.1007/s11590-007-0065-2.

J. P. Vielma and G. L. Nemhauser. Modeling disjunctive constraints with a logarithmic number of binary variables and constraints. *Mathematical Programming Series A*, 128:49–72, 2011. doi: 10.1007/s10107-009-0295-4.

H. Zhang and S. Wang. Linearly constrained global optimization via piecewise-linear approximation. *Journal of Computational and Applied Mathematics*, 214(1):111–120, 2008. doi: 10.1016/j.cam.2007.02.006.

C. Zhou, A. Mazumder, A. Das, K. Basu, N. Matin-Moghaddam, S. Mehrani, and A. Sen. Relay node placement under budget constraint. In *Proceedings of the 19th International Conference on Distributed Computing and Networking*, pages 1–11, 2018. doi: 10.1145/3154273.3154302.