A NEW PENALTY/STOCHASTIC APPROACH TO AN APPLICATION OF THE COVERING PROBLEM: THE GAMMA KNIFE TREATMENT

Marilis Bahr Karam Venceslau[1,2], Helder Manoel Venceslau[1,3,*],
Renan Vicente Pinto[1,4], Gustavo Dias[5] and Nelson Maculan[1]

Abstract. The covering problem of a three dimensional body using different radii spheres is considered. The motivating application – the treatment planning of Gamma Knife radiosurgery – is briefly discussed. We approach the problem only by the geometric covering point of view, that is, given a set of spheres and a body, the objective is to cover the body using the smallest possible number of spheres, regardless of the dosage issue. In order to solve this mathematical programming problem, we consider an approach based on the application of penalty and stochastic local search techniques. Finally, some illustrative results and comparisons are presented.

Mathematics Subject Classification. 90C90, 52C17, 90C26, 90C27.

Received January 2, 2018. Accepted February 19, 2020.

1. Introduction

Covering regions of the Euclidean space using regular geometric figures are classic problems in the literature. Their theoretical aspects are studied in the areas of discrete geometry and combinatorial geometry [24] as well as mathematical optimization [25, 27, 28].

Discrete mathematics and optimization techniques applied to medicine emerged as a new and important research area [5]. In the last years a large number of medical procedures have benefited from the use of optimization techniques. Among them, radiotherapy is unquestionably the most popular area in the medical treatment planning where optimization techniques are used [12, 19].

This article deals with the covering problem of three dimensional bodies using different radii spheres. Our formulation is motivated by the treatment planning of Gamma Knife radiosurgery [1, 6, 7, 22]. It is one of the most advanced means for treatment of brain tumors, vascular malformations and functional disorders. A Gamma

[1] COPPE – PESC, Federal University of Rio de Janeiro, Cidade Universitária, Centro de Tecnologia, Bloco H, 21941-972 Rio de Janeiro, RJ, Brazil.
[2] Colégio Pedro II, Campo de São Cristóvão, 177, São Cristóvão, 20921-903 Rio de Janeiro, RJ, Brazil.
[3] CEFET/RJ – Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Av. Maracanã, 229, Maracanã, 20271-110 Rio de Janeiro, RJ, Brazil.
[4] Universidade Federal Rural do Rio de Janeiro, BR-465, Km 7, 23897-000 Seropédica, RJ, Brazil.
[5] LIX, École Polytechnique, 91128 Palaiseau, France.
*Corresponding author: `heldermv@cos.ufrj.br`, `helder.venceslau@cefet-rj.br`

Knife unit accurately delivers a suitable dose of ionizing radiation, called shot, to the target tumor region. In good approximation, these shots can be modeled as spheres of different sizes. Multiple shots can be used to cover the tumor, while avoiding an excessive dose to the surrounding healthy tissue. The intrinsic characteristics of this problem enable its formulation as a covering problem in the three-dimensional Euclidean space $\mathbb{R}^3$.

Most of the existing literature about the use of optimization techniques and mathematical models to automate the treatment focus on obtaining complete solutions for the planning process, including the dosage in addition to the positions of the shots. Jitprapaikulsarn [9] contains a detailed summary of some of the existing approaches. On the other hand, Liberti *et al.* [11], Nascimento *et al.* [17], Soutou *et al.* [23] and Pinto [20] approached the problem only by the geometric point of view. In other words, they aimed to solve just the covering problem of a three dimensional body using different radii spheres, without dealing with the dosage issue in their formulations.

Pinto [20] formulated the problem as a non-convex Mixed Integer Nonlinear Programming (non-convex MINLP) with quadratic constraints and a linear objective function. Discretization techniques were used to obtain a linear model.

Our proposal is to provide a different method to solve the non-convex formulation proposed in [20]. After reformulating the model, we adopt an alternative approach in order to search for its solutions. More specifically, global optimization heuristics which mix stochastic searches and penalty techniques are used. Although we cannot ensure optimality of the solutions, our method selects the best solution from a set of prominent candidates.

The remainder of this paper is organized as follows. Section 2 introduces the covering problem and its basic mathematical formulation. A description of our new proposed formulation and solving method is presented at the end of this section. In Section 3, a smooth version of the spheres constraints and a heuristic for the "best" distance between unequal spheres are developed. Section 4 develops the security region concept, included in our formulation of the covering problem. Section 5 defines some important measures of coverage quality in our context. Section 6 deals with the combinatorial aspects of the problem. Section 7 presents illustrative computational results. Finally, Section 8 summarizes and presents the conclusions.

2. Problem definition

There is a myriad of variations of the covering problem, depending on the main object to be covered and the objects used to build the covering. This article firstly discusses a general version of the covering problem that uses different radii spheres to cover a solid body. It can be seen as an initial stage for the development of our proposed formulation, which will include additional constraints.

2.1. Mathematical preliminaries

The cardinality of a set A will be denoted by $|A|$ and the symmetric difference between two arbitrary sets A and B will be denoted by $A \, \Delta \, B = (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$.

Given an integer $n > 0$, let us define $L(n) = \{1, \ldots, n\}$ and denote its power set (the set of all its subsets) by $\mathcal{P}(L(n))$. Let Ω denote an arbitrary subset of $\mathcal{P}(L(n))$, such that $\varnothing \in \Omega$. This set will be described as $\Omega = \{\Omega_k \mid k \in \{0, \ldots, K\}\}$, such that $K + 1 = |\Omega|$ is the number of subsets of $L(n)$ in Ω and we will always fix the assignment $\Omega_0 = \varnothing$.

A (solid) sphere (or closed ball) in three dimensional space will be denoted by $S(x, r)$, where $x \in \mathbb{R}^3$ is its center location and r is its radius. An indexed location will be denoted by x^i, with i in a predefined index set. Whenever the location and radius of a sphere are not the focus of an analysis, they will be suppressed.

The volume (Lebesgue measure in $\mathbb{R}^3$) of a solid T will be denoted by V_T. The distances considered in the presented formulations will always use the Euclidean norm.

2.2. The optimal full covering problem

We start with the covering problem version presented in [11], which can be mathematically described as:

(P): Given a compact set $T \subset \mathbb{R}^3$, an integer $n > 0$, a finite set $R \subset \mathbb{R}_+$ of radii, a function $\rho\colon L(n) \to R$ and a subset $\Omega \subseteq \mathcal{P}(L(n))$, find a set of spheres $B = \{S(x^i, \rho(i)) \mid i \in \Omega_k\}$, $k \in \{0, \ldots, K\}$, of minimum cardinality covering each point of T.

The subset Ω is chosen intentionally to eliminate elements of the full power set $\mathcal{P}(L(n))$ which *via* ρ generate equivalent spheres selections. This formulation implicitly defines k and x^i as decision variables and the number of spheres, which must be minimized, as the objective function. A particular configuration of selected spheres and their locations will be referred as a *covering structure* and generically denoted as B.

This problem can be formulated as a mathematical programming problem, which searches for the appropriate subset of a pool of n spheres, with respective radii associated *via* function ρ, and the appropriate places for their centers, such that the union of these spheres covers the set T.

Problem (P) will be referred to just as *the optimal covering problem*, and a feasible solution will be a *full covering* of T. In practice, however, insisting on obtaining full coverings may come into conflict with other specific constraints of the problem at hand. Thus, we also define a *partial covering* as a covering structure that partially covers T. For example, an 80% partial covering of T means a covering structure that covers 80% of the volume of T. Without additional remarks, in this text a covering will always be assumed to be a partial covering.

2.3. Motivation for a new partial covering problem for the Gamma Knife radiosurgery application

The intrinsic characteristics of the Gamma Knife radiosurgery treatment planning problem enables its formulation as a covering problem using definition (P). Modeling shots as spheres, covering structures for Gamma Knife should have the following special attributes:

- No sphere is totally overlapped by the other spheres;
- Each sphere (normally a part of it) contributes to cover the tumor;
- The total number of spheres is small to prevent use of high doses of radiation;
- No sphere covers a significant part of healthy tissue.

The last item is a constraint not included in the original definition of problem (P). One of its most important consequences is to make it extremely difficult to obtain full coverings that do not put healthy tissue in danger of receiving excessive doses of radiation. In other words, for the Gamma Knife treatment planning, 90% partial coverings can be considered totally acceptable when compared to full coverings that may submit healthy tissue to high doses of radiation [12].

A convenient way to enforce healthy tissue protection is to define an enlarged version of the tumor and keep all spheres inside this region, which will be denominated the *security region* associated to a tumor. Note that this is a constraint commonly found in packing formulations. As the spheres in a covering are allowed to overlap each other, we are actually facing a sort of *relaxed packing* inside the security region.

Containment in a security region will be an important constraint in the forthcoming formulations. A detailed analysis of this concept and the associated constraint will be presented in Section 4.

We will now develop some assessment tools to measure the quality of (partial) coverings. They will be used to justify the modeling techniques that will be employed soon.

Let us fix a certain covering structure B, which means to fix a selection of spheres Ω_k and the location of their centers. For the purpose of this analysis, let us renumber the set of selected spheres and denote it by $S = \{S_1, \ldots, S_m\}$, where $m = |\Omega_k|$. Free indexes in the expressions below will take values from $L(m)$, which means we will simply write $B = \bigcup S_i$.

For the considered covering structure B, let $\hat{S}_i = S_i \setminus (\bigcup_{j \neq i} S_j)$. If $\hat{S}_i \neq \varnothing$, $\forall i \in L(m)$, then we will say that this covering is a *nonredundant* covering. In a nonredundant covering each sphere contributes to the volume V_B.
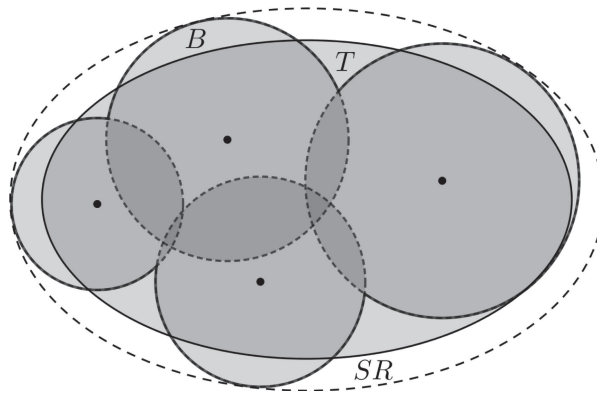
FIGURE 1. Solid T partially covered by a (supposed) admissible covering structure B. Label SR denotes the boundary of the security region.

Additionally, if $V_{T \cap \hat{S}_i} > 0$, $\forall i \in L(m)$, then we will say that this covering is an *effective* covering (clearly effective implies nonredundant). In this context, "effective" means that each sphere cannot be discarded because it effectively contributes to cover T.

An arbitrary covering structure B does not need to be effective (neither nonredundant), which means that some spheres may have to be discarded to obtain an effective covering. It is important to notice that the discard order matters, affecting the resulting final effective covering. A simple example consists of a covering using spheres S_1 and S_2 with $T \subset (S_1 \cap S_2)$. Either S_1 or S_2 can be discarded in order to obtain two different effective coverings. It is clear that some kind of ordering is necessary to organize the selection of spheres by an optimization process in order to minimize $m = |\Omega_k|$ using effective coverings. This will be discussed in Section 6.

Covering structures that are contained in the security region, abbreviated to SR, will be denominated *SR-compliant* coverings. Security regions are the subject of Section 4.

Partial coverings may not integrally cover a set T, which is not desirable. On the other hand, it is considered acceptable (but not desirable for applications like Gamma Knife) that points outside T are covered. Therefore, a covering index I_Δ measuring the covering fitness can be defined as any monotonic function of the volume of the symmetric difference $V_{B \triangle T}$. A covering whose index I_Δ attains a certain predefined threshold, which prescribes a low symmetric difference volume, will be denominated an *attaining I_Δ* covering. Section 5 introduces different covering indexes, including the *Paddick index*, which is widely used and convenient in our context.

Finally, for some applications the total volume of the intersection of two (or more) spheres must be kept below some predefined threshold (for example, to prevent excessive doses of radiation at some regions of the tumor). Thus, we also define $I_\cap$ indexes taking into account this volume, and partial coverings satisfying this constraint will be denominated *attaining $I_\cap$* coverings. Section 5 introduces the $I_\cap$ indexes I_{B2}, I_{B3} and I_{B4}.

We define a *pre-admissible covering* as an SR-compliant effective covering, and an *admissible covering* as a pre-admissible covering that is also an attaining I_Δ and $I_\cap$ covering. Figure 1 illustrates an admissible partial covering.

Specific actual applications of coverings like Gamma Knife look for admissible coverings. It is now clear that covering optimality in our context has to do with obtaining the "best admissible covering", whose precise meaning will be explained at the end of Section 2.4.

2.4. Mathematical formulation of the new partial covering problem

Instead of solving the classical covering problem (P), which is not adequate for applications like Gamma Knife, the approach taken by [20] is to develop an alternative formulation based on desirable properties of

partial coverings, hereafter named $(P1)$, which will be the foundation of our proposed formulation in the next section.

Considering a solid T and a set $S = \{S_1, \ldots, S_n\}$ of n spheres, possibly with different radii r_i, the formulation $(P1)$ is defined using decision variables x and y:

- $x \in T^n$;
- $y \in \{0,1\}^n$.

where:

- $x = \{x^1, \ldots, x^n\}$ is a vector containing the centers of the spheres;
- $y = \{y_1, \ldots, y_n\}$ specifies which spheres of the set S are present in the solution (an alternative way to express a subset Ω_k of S).

In order to present the formulation $(P1)$, let us firstly define the parameters

$$\alpha_{ij}, \quad 1 \le i < j \le n,$$

such that

$$0 < \alpha_{ij} < \min\{r_i, r_j\},$$

which are used to derive the minimum allowed distances

$$d_{ij} = r_i + r_j - \alpha_{ij},$$

between spheres S_i and S_j, as depicted in Figure 2, and the real numbers

$$c_i, \quad 1 \le i \le n,$$

which are used to select what spheres are employed in a specific covering.

Let us present a detailed description of these parameters and their roles. The parameters d_{ij}, derived from the parameters α_{ij}, are used, in the final analysis, to define the maximum allowed intersection of two spheres. This is accomplished by requiring that

$$d(S_i, S_j) = ||x^i - x^j|| \ge d_{ij} = r_i + r_j - \alpha_{ij}.$$

For a specific sphere S_i, the upper bound on parameter α_{ij} implies that $r_i - \alpha_{ij} > 0$. As a consequence, $d(S_i, S_j) - r_j > 0$, for $1 \le j \le n$, $j \ne i$. If we define $\gamma_i = \min_{j \ne i}\{d(S_i, S_j) - r_j\}$, then $\gamma_i > 0$. This last expression means that a hypothetical sphere $\bar{S}$ with center x^i and radius less than γ_i can be drawn such that $\bar{S} \bigcap (\bigcup_{j \ne i} S_j) = \varnothing$, which immediately implies that $\hat{S}_i = S_i \setminus (\bigcup_{j \ne i} S_j) \ne \varnothing$. In other words, we get a nonredundant covering.

Additionally, if $x^i \in T$ then $T \bigcap \bar{S} \ne \varnothing$ and, as a consequence, we have $V_{T \bigcap \bar{S}} > 0$, which implies that $V_{T \bigcap \hat{S}_i} > 0$ (because $\bar{S} \subset \hat{S}_i$). In other words, we get an effective covering.

The parameters c_i are "weights" associated to the spheres S_i, which increase in direct proportion with r_i^3. They are used in the objective function to favor the choice of bigger spheres, which tends to reduce the number of necessary spheres used to cover T, in accordance with the optimization direction of the classical covering problem (P), as discussed in [20].

Finally, we have the mathematical programming formulation $(P1)$:

$$(P1): \max \sum_{i=1}^{n} c_i\, y_i \tag{2.1}$$

$$\text{s. t. } ||x^i - x^j||^2 \ge d_{ij}^2\, (y_i + y_j - 1), \quad 1 \le i < j \le n, \tag{2.2}$$

$$x \in T^n, \tag{2.3}$$

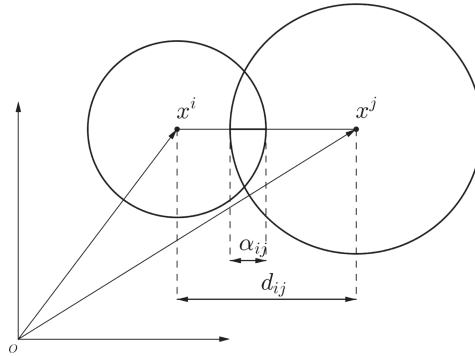$$y \in \{0,1\}^n. \tag{2.4}$$

FIGURE 2. Parameters describing the intersection of two spheres.

This formulation can be analyzed using the definitions of the previous section to conclude that it is designed to produce effective partial coverings. Constraints (2.2) prevent excessive overlap of the selected spheres, producing nonredundant coverings. Constraints (2.3) force the centers of the spheres to be inside T, which together with constraints (2.2), produce effective coverings. Constraints (2.4) define which spheres are selected from the set of available spheres. The objective function (2.1) includes the different weights given to the different spheres, based on their radii, which provides a mechanism to favor the bigger ones. Problem $(P1)$ is clearly a non-convex MINLP [2,8].

Formulation $(P1)$ produces effective partial coverings with high coverage of the solid T. On the other hand, it does not produce SR-compliant coverings[6]. It also does not consider any kind of fitness or intersection indexes. But formulation $(P1)$ is sufficiently flexible to be the base for an improved formulation, which will be presented in the next paragraphs.

In this work, the adopted approach to solve $(P1)$ involves the following steps:

- Organize the selected spheres into nodes of an underlying *sphere allocation tree* (each node is composed of a fixed selection Ω_k of spheres);
- Penalize violations of the "spheres overlap" constraints (2.2);
- Penalize violations of the "spheres" centers in solid" constraints (2.3).

The way we integrate the three points above into our method uses the following unconstrained mathematical programming (MP) formulation, valid for the kth node of the sphere allocation tree:

$$(PN) : \max C_k - \sum_{\substack{i,j \in \Omega_k \\ i<j}} p_u \left(\frac{\|x^i - x^j\|}{d_{ij}} \right) - \sum_{i \in \Omega_k} q_\mu(d^i, x^i) \tag{2.5}$$

where

- $C_k = \sum_{i \in \Omega_k} c_i$, because $y_i = 1$ for $i \in \Omega_k$ and $y_i = 0$ for $i \notin \Omega_k$;
- Function p_u penalizes violations of the "spheres overlap" constraints (Sect. 3);
- Function q_μ penalizes violations of the "spheres centers in solid" constraints (Sect. 4, which defines the parameter d^i).

---

[6]In Pinto [20] the author subsequently includes his version of security region to deal with Gamma Knife treatment.

This approach results in some interesting consequences:

- The formulation (PN) is not linked to the sphere selection combinatorial aspect of the original problem. It is used just to reduce penalization of the constraints violations in order to obtain effective partial coverings.
- Function q_μ can be designed to incorporate the security region constraint in order to obtain pre-admissible partial coverings.
- Quality criteria based on I_Δ and $I_\cap$, which are not present in the objective function of $(P1)$, can now be used to deal only with admissible partial coverings (Sect. 5).

It is important to note that penalization methods are adequate for this problem because small deviations of the locations of the spheres' centers still produces viable solutions in the context of partial coverings. The overall strategy to solve our modified version of problem $(P1)$ can then be stated as:

- Find a reasonable way to organize the spheres into nodes of a sphere allocation tree, where each node is composed of a fixed selection of spheres (Sect. 6);
- Heuristically determine a solution of (PN) satisfying the security region constraint, for each node of the sphere allocation tree (Sects. 3 and 4);
- Measure the quality of the covering, keeping track of admissible covering nodes, and iterate until it is no more possible to get admissible covering nodes (Sects. 5 and 6).

This strategy will finally choose the best covering from a set of admissible coverings. From now on this will be the meaning of "best (admissible) covering" in our context.

3. Penalty for violation of spheres overlap constraints

3.1. Definition of the penalty function p_u

In order to define p_u, let us firstly define the penalty function p_v, which is based on a left shifted version of the Van der Walls potential [10]:

$$p_v(t) = E \left(1 - \frac{1}{\left((1 - b_k)\, t + b_k \right)^4} \right)^k,$$
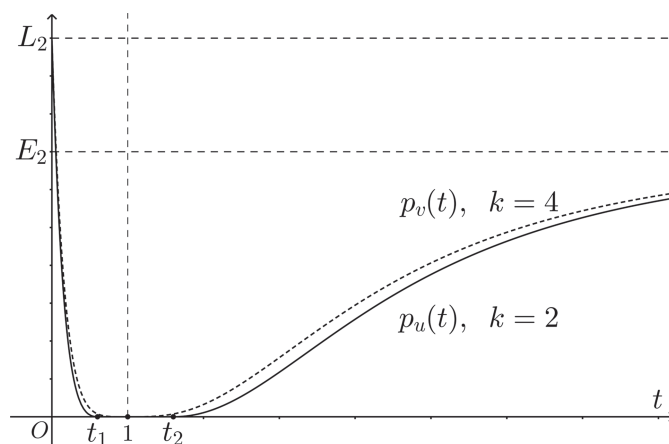
where

$$b_k = \frac{1}{\sqrt[4]{\sqrt[k]{L/E} + 1}},$$

and $k \geq 2$ is an even number. The parameter L is the penalty value for $t = 0$ and the parameter E is the asymptotic penalty value when $t \to \infty$. The minimum value zero is attained for $t = 1$ $(p_v(1) = 0)$, which can then be used to penalize violations of the equality $\|x^i - x^j\| = d_{ij}$ through the inclusion of a penalty term in the objective function.

The function p_v, as defined above, has the following characteristics that make it very adequate as a penalty function for the spheres overlap:

- As $p_v(0) = L$, solvers will not have to deal with not defined values of the penalty function. This is why a left shift has been applied to the original Van der Walls potential: it is not defined for $t = 0$ (goes to $+\infty$ as $t \to 0$);
- The asymptotic value E implies that spheres that are far apart from each other will not present too much mutual influence when compared to spheres that are close to each other. But they will tend to bunch together (if there are no other spheres in between them to prevent it) due to the small decrease on the distance associated penalty values;
- The width and depth of the basin of low penalty values $(p_v(t) \leq 0.5E)$ can be varied adjusting the parameters L, E and k.

FIGURE 3. Graph of functions $p_v(t)$ and $p_u(t)$ for $\delta = 0.3$.

The function p_v has only one minimum at $t = 1$. But the study of optimal coverings of $\mathbb{R}^3$ using equal radii spheres shows that, different from the case of covering of $\mathbb{R}^2$ using equal radii circles, d_{ij} cannot be taken as only one fixed value (more details in Sect. 3.2).

This means that it is convenient to define a null penalty interval that would allow a certain degree of freedom for the placement of the spheres. This can be achieved starting with function p_v, shifting it down by δ, eliminating its negative values and squaring the result to obtain a differentiable penalty function p_u:

$$p_u(t) = \left[\max(0, p_v(t) - \delta) \right]^2,$$

where the p_v parameters $L = \sqrt{L_2} + \delta$ and $E = \sqrt{E_2} + \delta$ are obtained from the definition of the equivalent parameters $L_2 = p_u(0) = (L - \delta)^2$ and $E_2 = \lim_{t \to \infty} p_u(t) = (E - \delta)^2$ of p_u.

It is worth noting that for $\delta = 0$, p_u defined using p_v with parameters k, $L = \sqrt{L_2}$ and $E = \sqrt{E_2}$ is exactly the same as p_v with parameters $2k$, $L = L_2$ and $E = E_2$.

Figure 3 depicts a function p_u built from a function p_v with $k = 2$, $L = \sqrt{L_2} + \delta$ and $E = \sqrt{E_2} + \delta$ which, for small values of δ, is very close to a function p_v with $k = 4$, $L = L_2$ and $E = E_2$. Setting $k = 2$ is enough for our purposes and will be adopted hereafter.

Points $0 < t_1 < 1$ and $t_2 > 1$ are the ends of the null penalty interval. They can be obtained solving $p_v(t) = \delta$:

$$t_1 = \frac{\frac{1}{\sqrt[4]{1 + \sqrt{\delta/E}}} - b_2}{1 - b_2}, \quad t_2 = \frac{\frac{1}{\sqrt[4]{1 - \sqrt{\delta/E}}} - b_2}{1 - b_2}.$$

Substituting b_2 by its definition, calculating the quotient of t_2 and t_1 and simplifying the final expression we get

$$\frac{t_2}{t_1} = \frac{\sqrt[4]{\frac{1 + \sqrt{L/E}}{1 - \sqrt{\delta/E}}} - 1}{\sqrt[4]{\frac{1 + \sqrt{L/E}}{1 + \sqrt{\delta/E}}} - 1}.$$

Considering that the distances between the centers x^i and x^j of the spheres S_i and S_j, $1 \le i < j \le n$, will be allowed to vary, with null penalty, between positive lower and upper bounds

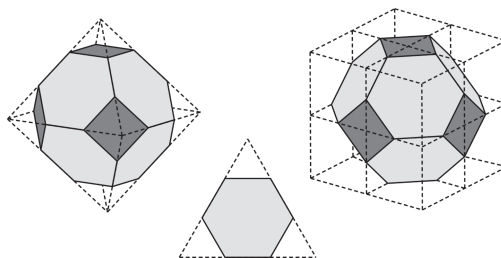$$m_{ij} \le \left\| x^i - x^j \right\| \le M_{ij},$$

FIGURE 4. Truncated octahedron.

it is necessary to determine a certain δ^* such that

$$\frac{t_2}{t_1} = \frac{M_{ij}}{m_{ij}} \geq 1.$$

A straightforward analysis shows that for $\delta = 0$ we have $t_1 = t_2 = 1$ and for $\delta \to \infty$ we have $t_1 \to 0$ and $t_2 \to \infty$. This means that it is always possible to find a number δ^* such that the fraction t_2/t_1 assumes any desirable value greater than or equal to one. The numeric value of δ^* can be determined using Newton's method, for example.

Proportionality then enables the determination of the parameter d_{ij} to be used in function p_u:

$$d_{ij} = \frac{m_{ij}}{t_1} = \frac{M_{ij}}{t_2}.$$

3.2. Values of α_{ij}

An important question has now to be answered: How much two spheres should be allowed to overlap in order to obtain a satisfactory covering by spheres?

3.2.1. Values of α_{ij} for same radii spheres

For the case of spheres with identical radii the optimum space covering arrangement of spheres is the body-centred cubic (bcc) lattice where each sphere circumscribes a truncated octrahedron (see [3], for example). The plane containing the intersection of two spheres also contains one face of their respective truncated octahedron, and the value of α (Fig. 2) depends on how each octrahedron "kisses" the other: using square faces or hexagonal faces.

Due to this fact, there are two different values of α for spheres with centers in a bcc lattice, one for truncated octahedra square faces and another for truncated octahedra hexagonal faces, respectively:

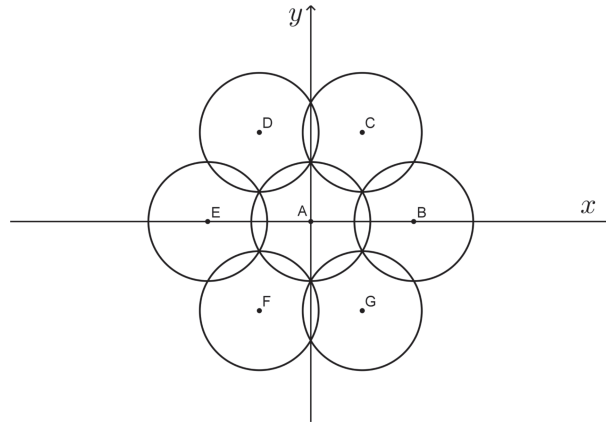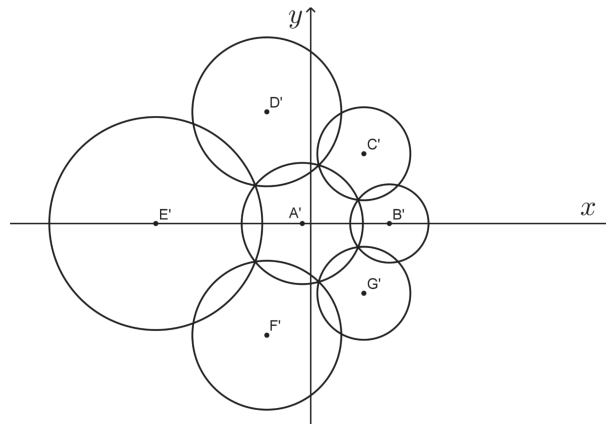$$\alpha_m \cong 0,2111R \quad \text{and} \quad \alpha_M \cong 0,4508R,$$

where R is the common radius of the spheres. The truncated octahedron, inscribed in an octahedron and in a cube, is presented in Figure 4.

3.2.2. Values of α_{ij} for different radii spheres

For the case of different sphere radii, we decided to use a heuristic approach: apply a conformal mapping (which preserves angles) to a regular arrangement of spheres. A basic conformal mapping for $\mathbb{R}^2$, here represented by a complex function, is

$$f(z) = \frac{z}{cz + 1},$$

which is the identity for parameter $c = 0$. The effect of applying this conformal map to an hexagonal covering of $\mathbb{R}^2$ by circles can be realized by comparing Figures 5 and 6.
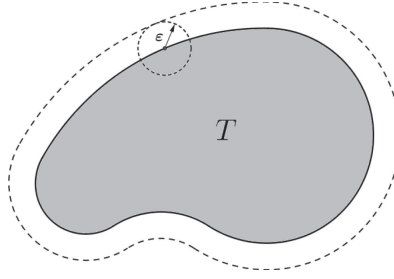
FIGURE 5. Seven circles from an hexagonal covering of $\mathbb{R}^2$ by circles.

FIGURE 6. The same seven circles, after applying a basic conformal map f with $c = 0.06$. Due to the mapping distortion, the new centers are not obtained as images of the original centers.

A similar reasoning can be applied for the covering of $\mathbb{R}^3$ by spheres with centers pertaining to the bcc lattice[7]. As the conformal mappings preserve the angles, the analysis of the distances between the spheres can be performed on their intersections with a plane that contains their centers. The angles before the transformation can be easily determined using basic geometry and the fact that the radii are the same. After the transformation, the angles are preserved and the distances can then be determined using the law of cosines.

Based on this process, it is possible to calculate the values of d_{ij} (associated to the values of α_{ij}) for the bcc lattice:

- $m_{ij}^2 = r_i^2 + r_j^2 + \frac{2}{5}r_i r_j$, associated to α_M.
- $M_{ij}^2 = r_i^2 + r_j^2 + \frac{6}{5}r_i r_j$, associated to α_m.

---

[7]Smooth conformal mappings on a domain of $\mathbb{R}^3$ ($\mathbb{R}^n$ in general, for $n > 2$) are Möbius transformations. This is the celebrated Liouville's theorem of conformal geometry.

FIGURE 7. Security region of a solid T (ε is the tolerance).

As presented in the previous section, the determination of t_1 and t_2 for function p_u implies finding δ^* such that:

$$\frac{t_2}{t_1} = \frac{M_{ij}}{m_{ij}} = \sqrt{\frac{r_i^2 + r_j^2 + \frac{6}{5}r_i r_j}{r_i^2 + r_j^2 + \frac{2}{5}r_i r_j}} = \sqrt{1 + \frac{4}{5\left(\frac{r_i}{r_j} + \frac{r_j}{r_i}\right) + 2}}.$$

But the radii of the spheres (previously chosen) are known, which means that:

- the quotients t_2/t_1 are known;
- the values of δ^* can be pre-calculated (*via* Newton's method);
- the parameters t_1 and t_2 can be calculated;
- finally, the parameters d_{ij} of p_u can be calculated.

4. PENALTY FOR VIOLATION OF SECURITY REGION AND SPHERES' CENTERS CONSTRAINTS
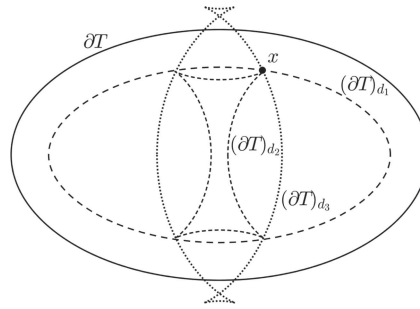
4.1. Definition of the security region

We now turn our attention to the security region. For practical applications like Gamma Knife treatment planning (as mentioned in Sect. 2.3) it is important to avoid that a large volume of the covering spheres be located outside the target volume T. In our model that is accomplished maintaining all spheres inside the security region. Figure 7 depicts a plane section of a solid T and its security region with tolerance parameter ε.

In order to simplify our mathematical[8] development, we will assume that $T \subset \mathbb{R}^3$ is homeomorphic to the unitary ball $B^3 = \{(u, v, w) \in \mathbb{R}^3 \mid u^2 + v^2 + w^2 \leq 1\}$. In other words, T is compact (closed and bounded), connected, with uniform dimension equal to three and its boundary ∂T is homeomorphic to the sphere $S^2 = \partial B^3$. It will also be assumed that the boundary ∂T is either a differentiable (smooth) surface or is composed of a finite number of differentiable surfaces.

By convention, the unitary vector $\hat{n}(x)$, normal to the surface ∂T at x, will point outside T. This normal vector can always be defined for points on the differentiable pieces of ∂T. To any point x located on the junction of two or more differentiable surfaces (like the edges or vertices of a polyhedron's surface) we associate not only one vector, but a set of unitary vectors, hereafter denominated $N(x)$, obtained from the normalization of the conic combinations of the normal vectors of each surface at the junction point x. These normal vectors can be obtained calculating the limit $\lim_{y \to x} \hat{n}(y)$ using points y close do x on each differentiable surface of the junction. This definition can be extended to all points $x \in \partial T$ if we specify that $N(x) = \{\hat{n}(x)\}$ for x located on a differentiable piece of ∂T.

As will be shown in the next paragraphs, keeping a sphere inside the security region is equivalent to keeping its center inside an internal *parallel body*. The parallel body, on the other hand, will be defined in terms of parallel surfaces.

---

[8]In this paragraph we temporarily use the traditional topological notation.

FIGURE 8. $x \in T$ pertains to three distinct parallel surfaces.

The parallel surfaces to the boundary surface ∂T of a solid T will be represented by $(\partial T)_d$ and defined as

$$(\partial T)_d = \{y \in \mathbb{R}^3 \mid y = x + d \cdot N(x),\ x \in \partial T\},$$

where, for each $x \in \partial T$, $x + d \cdot N(x)$ is the subset composed of the vectors obtained from the vector sum of x and each d scaled vector $\hat{n} \in N(x)$.

It is important to notice that $|d|$ can be interpreted as distance, but depending on the sign of d we have:

- $d > 0$: Parallel surfaces $(\partial T)_d$ are in general external to T. It is clear that for T convex the surfaces $(\partial T)_d$ are always external and never self-intersect.
- $d \leq 0$: Parallel surfaces $(\partial T)_d$ can be internal or external to T (even for T convex) but are always internal for small values of $|d|$.

To simplify the analysis, from this point on we will deal only with convex bodies T, although it will not be required that the parallel surfaces $(\partial T)_d$ are differentiable.

We define the internal ($d \leq 0$) parallel body at a (signed) distance d as the set of points inside T that are at a distance greater than or equal to $|d|$ from ∂T. To unambiguously define the internal parallel body it would be necessary that each point inside T was located over only one parallel surface. But the internal surfaces often present self-intersections [13, 14] as d decreases starting from zero, which can be verified in a section of T, as shown in Figure 8.

To overcome this difficulty, let us define the following auxiliary function ($d \leq 0$):

$$D(x) = \arg \max_{x \in (\partial T)_d} d,\ \forall x \in \mathbb{R}^3.$$

This function helps to eliminate parallel surface ambiguity whenever x pertains to more than one surface $(\partial T)_d$. It is straightforward to verify that it measures the (signed) distance from $x \in T$ to ∂T, given by $D(x) = -\mathcal{D}(x, \partial T)$, where $\mathcal{D}$ represents the distance function from a point to a geometric object.

The main properties of $D(x)$ are:

- If $x \in \partial T$ then $D(x) = 0$;
- $D(x)$ is continuous (because distance functions are continuous);
- $D(x)$ attains the minimum value inside T (by the extreme value theorem);
- If $T \subset \mathbb{R}^3$ is homeomorphic to the unitary ball and convex then $D(x)$ is a convex function.

A proof of the last property can be found in [26]. It makes it possible to define the body parallel to T at (signed) distance d as:

$$T_d = \{x \in \mathbb{R}^3 \mid D(x) \leq d\}.$$

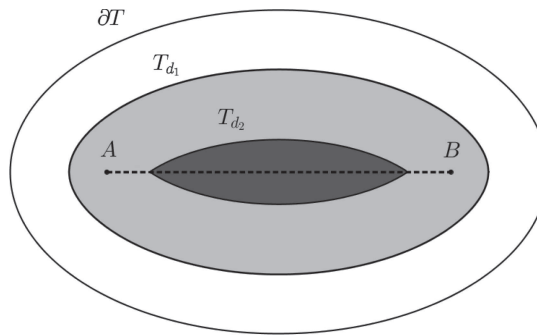Some examples of parallel bodies are presented in Figures 9 and 10.

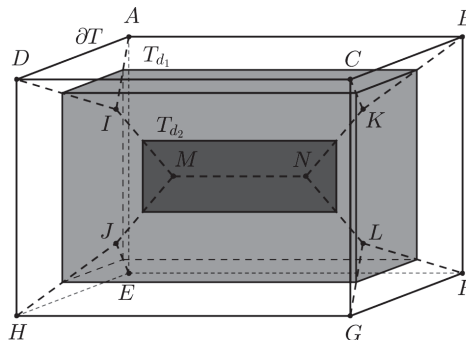FIGURE 9. Prolate ellipsoid section: ∂T regular, but $D(x)$ is non-differentiable in segment $\overline{AB}$.

FIGURE 10. Parallelepiped section: ∂T non regular and $D(x)$ is non-differentiable in segments $\overline{MN}$, $\overline{AI}$, $\overline{DI}$, $\overline{IM}$, $\overline{HJ}$, $\overline{EJ}$, $\overline{JM}$, $\overline{CK}$, $\overline{BK}$, $\overline{KN}$, $\overline{GL}$, $\overline{FL}$ e $\overline{LN}$.

Since T convex implies $D(x)$ convex, the parallel body T_d, as a sublevel set of a convex function, is consequently a convex set. The Motzkin theorem [16], from Convex Analysis, ensures then that the distance function $\mathcal{D}(x, T_d)$ is always differentiable for $x \in \mathbb{R}^3 \backslash T_d$. Using $\mathcal{D}(x, T_d)$ in our definitions is therefore a simple way to work with functions that are always differentiable, which simplifies implementations using non-linear solvers. This comes into play when we consider the following equivalence:

$$S_i \subset T_\varepsilon \iff x^i \in T_{\varepsilon - r_i}, \ \ i \in L(n).$$

Figure 11 shows geometrically the meaning of this equivalence, which is very intuitive. A formal proof is presented in [26].

The second part of the equivalence can also be expressed as:

$$D(x^i) \leq \varepsilon - r_i, \ i \in L(n). \tag{4.1}$$

4.2. Definition of the penalty function q_μ

If now we define the function

$$q(d, x) = \begin{cases} 0, & D(x) \leq d \ (\text{in other words, } x \in T_d) \\ +\infty, & D(x) > d \ (\text{in other words, } x \notin T_d) \end{cases},$$
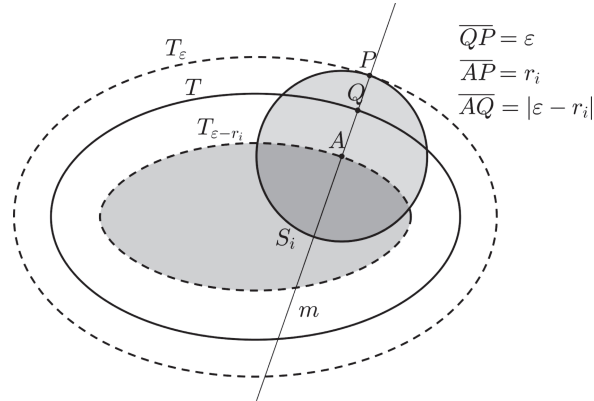
FIGURE 11. Solids T_ε, T and $T_{\varepsilon-r_i}$. Sphere S_i is tangent to ∂T_ε with common normal line m at P.

the constraints (4.1) can also be expressed as:

$$q(\varepsilon - r_i, x^i) = q(d^i, x^i) \leq 0, \ i \in N,$$

where $d^i = \varepsilon - r_i$. It is worth to notice that defining $d^i = \min(0, \varepsilon - r_i)$ automatically includes the constraint $x \in T$.

Despite its useful theoretical value, function q cannot be implemented in practice. But a smooth version of it can be created to be used as a penalty function q_μ, such that $q_\mu \to q$ when $\mu \to \infty$. The most natural smooth version of q is:

$$q_\mu (d, x) = \mu \mathcal{D}(x, T_d)^2.$$

It is clear that the function q_μ is built to be differentiable for all $x \in \mathbb{R}^3$, even for $x \in \partial T_d$, a side effect of squaring the generally non-differentiable function $\mathcal{D}(x, T_d)$.

Assuming the definition $d^i = \min(0, \varepsilon - r_i)$, the function q_μ can then be used to simultaneously penalize violations of the "security region" and "spheres' centers in solid" constraints in the problem formulation (PN) for each node of the sphere allocation tree (defined later).

4.3. q_μ for specific solids

The function q_μ depends directly on the parallel body T_d, which in general can be difficult to describe. The simplest case to be considered is a (solid) sphere T centered at the origin and with radius R. It is straightforward to verify that for this solid we have

$$q_\mu (d, x) = \mu[\max(0, ||x|| - (R - d))]^2.$$

Other solids T are considered in Section 7, which demanded the design of specific functions q_μ for each one of them.

5. COVERING INDEXES

Measuring the quality of a covering is an important issue in Gamma Knife radiosurgery and some conformity indexes have already been proposed. In this simplified approach we suppose that the radiation energy density is constant over the balls that represent the shots. It means that the aspects affecting the quality of a covering are the following ones:

(1) Is each point of the solid T covered?

(2) Are all covered points inside the solid T?

(3) Is every covered point covered by only one ball of the covering?

As already mentioned before, the aspect #1 is a prerequisite in a traditional mathematical covering problem. This is not our case because a full covering of T can severely impact the aspect #2, which means that healthy tissue is being irradiated. Finally, the aspect #3 means that part of the tissue (maybe healthy) is receiving an extra dose of radiation, which is undesirable from a medical point of view.

Instead of considering isolated points, it is better to consider volumes in order to measure compliance with the above mentioned aspects. Let us firstly define some basic volumes:

- V_T: Volume of solid T;
- V_B: Volume of the covering structure $B = \bigcup_{i \in \Omega_k} S_i$;
- $V_{T \cup B}$: Volume of the union $T \cup B$;
- $V_{T \cap B}$: Volume of the intersection $T \cap B$;
- $V_{T \Delta B}$: Volume of the symmetric difference $T \Delta B$;
- V_{IBi}: Volume of all the intersections of "i" or more spheres, for $i \in \{2, 3, 4\}$.

Using the basic volumes, we are now able to build a set of conformity indexes:

- The covering index $I_{\text{COV}} = \dfrac{V_{T \cap B}}{V_T}$: This index measures what fraction of the solid T is covered by the set of spheres B;

- The efficiency index $I_{E1} = \dfrac{V_{T \cap B}}{V_B}$: This index measures what fraction of the set of spheres B is used to cover the solid T;

- The Intersection index $I_{Bi} = \dfrac{V_{IBi}}{V_B}$, for $i \in \{2, 3, 4\}$: This index measures what fraction of the set of spheres B is composed of the intersection of "i" or more spheres;

The conformity indexes most mentioned in literature and widely used in radiosurgery centers are:

- PITV index: The best known and most used index, before the appearance of the Paddick index. It is described by Shaw *et al.* [21] and defined by

$$I_{\text{PITV}} = \frac{V_{T \cap B}}{V_T} = I_{\text{COV}}.$$

- Paddick index: Better than the Shaw index because it also takes into consideration the relative position of the volumes V_T and V_B. It is described by Paddick in [18] and defined by

$$I_{\text{PDK}} = \frac{V_{T \cap B}^2}{V_T V_B} = \frac{V_{T \cap B}}{V_T} \frac{V_{T \cap B}}{V_B} = I_{\text{COV}} I_{E1} = \frac{(V_B + V_T - V_{T \Delta B})^2}{4 V_T V_B}.$$

Paddick calls I_{COV} the "target coverage" and I_{E1} the "plan selectivity".

As $\varnothing \subseteq T \Delta B \subseteq T \cup B$, it follows that $0 \leq V_{T \Delta B} \leq V_{T \cup B}$. If $V_{T \Delta B} = V_{T \cup B}$ we have $I_{\text{PDK}} = 0$ and if $V_{T \Delta B} = 0$ we have $I_{\text{PDK}} = 1$. It is clear that I_{PDK} is a monotonic function of $V_{T \Delta B}$, thus an I_Δ index. By definition we always have $I_{\text{PDK}} \leq I_{\text{COV}}$ and $I_{\text{PDK}} \leq I_{E1}$. Thus I_{PDK} is a lower bound for I_{COV} and I_{E1}, which means that improving (increasing) I_{PDK} automatically improves I_{COV} and I_{E1}.

Considering their definitions, it is also clear that I_{B2}, I_{B3} and I_{B4} are $I_\cap$ indexes. We obviously always have $I_{B4} \leq I_{B3} \leq I_{B2}$. Thus I_{B2} is an upper bound for the other intersection indexes and improving (decreasing) I_{B2} automatically improves I_{B3} and I_{B4}.

Figure 12 depicts the typical behaviour of the main indexes as functions of the number N_S of equal spheres in a covering structure:
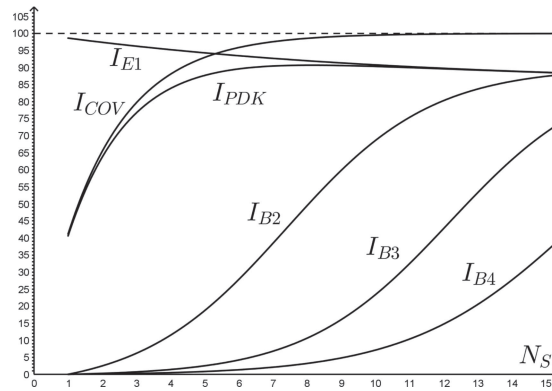
FIGURE 12. General behaviour of the conformity indexes.

As the indexes I_{PDK} and I_{B2} are bounds to all the other mentioned indexes, they were chosen as the main quality indexes. Based on preliminary results of our experiments, it was clearly difficult to obtain values of I_{PDK} superior than 75% and insisting on increasing the number of spheres made the intersection index I_{B2} to increase, reaching values above 35%. Both indexes are then used to together measure the quality of a covering, by the prescription of a minimum value for I_{PDK} and a maximum value for I_{B2}.

6. SPHERE ALLOCATION TREE

The previous sections deal with the problem of covering a solid T using a fixed selection of different radii spheres, which was formulated as the (PN) problem. It has already been mentioned that solving (PN) results in pre-admissible partial coverings, which are viable solutions for practical problems like the Gamma Knife planning process. The Paddick and Intersection indexes were presented as a way to measure the quality of the covering solutions associated to each selection of spheres.

Now, assuming we have a heuristic method for solving the problem (PN) associated to each selection of spheres, there is a final step to be accomplished: find a way to choose the best selection of spheres. The most convenient way to do this is to firstly define a structure to organize the selections of spheres and then define a search strategy to get nodes from this structure. The adopted structure is denominated a sphere allocation tree and each node of this tree represents a particular selection of spheres. The root of the tree is the empty set of spheres and the new node choice, taken at each node during the search procedure, is associated to the conformity indexes.

A point that deserves attention is the fact that the tree nodes are not readily available: they must be created. This prompts for the definition of a node generation procedure.

The following subsections will detail the sphere allocation tree, the search strategy and the node generation procedure.

6.1. Sphere allocation tree structure

Given a set $\Omega \subseteq \mathcal{P}(L(n))$, the underlying sphere allocation tree organizes its elements (Ω_k nodes representing spheres selections for $k \in \{0, 1, \ldots, K\}$) in order to make a search. For example, suppose we have a pool of spheres such that:

- F is the total number of different radii. For each radius we say we have a different "type" of sphere;
- G is the total number of spheres of each radius.

Of course $n = FG$ and we could simply define $\Omega = \mathcal{P}(L(n))$. In this case we would have the number of nonempty subsets of spheres K calculated as

$$K = 2^{FG} - 1 = 2^n - 1.$$

But it is clear that organizing the spheres by radius, such that for each radius the first spheres are always selected before the last ones, eliminates symmetries and reduces the size of the set of nonempty subsets of spheres to

$$K = (G + 1)^F - 1.$$

Both expressions for K are the same for $G = 1$, but the second is always less than the first for $G > 1$ (since $2^G > G + 1$ for $G > 1$). This second definition of Ω is obviously better and will be adopted hereafter.

In order to identify each node Ω_k, let us define what we will call the signature of the kth node, for $k \in \{0, 1, \ldots, K\}$:

$$\eta_k = (l_1, l_2, \ldots, l_F),$$

where l_i is the number of spheres for each different radius, with $0 \leq l_i \leq G$, $i = 1, \ldots, F$. In other words, $\eta_k \in \{0, 1, \ldots, G\}^F$. A sphere allocation tree can then be created and organized by signatures.

6.2. Search strategy

We want to search and locate good nodes within the tree, organized by signatures, using the covering indexes to guide the selection of the subsequent node. The following search strategy was designed, due to the influence of the Gamma Knife application, which aims to obtain a good tumor covering while minimizing the total number of shots:

- Firstly select the bigger spheres;
- Increase the number or spheres of the current selected size;
- If the quality criteria, based on the covering indexes, rejects the current set of spheres, we backtrack to the previous selection of spheres (by removing the last added sphere) and start using spheres of immediately smaller size;
- This process is repeated until it is no more possible to increase the number of spheres used in the covering.

The *sphere allocation diagram*, depicting the search strategy, is shown in Figure 13.

Some details of the search decision at each node are listed below:

- Quality criteria: the biggest Paddick index I_{PDK} associated to an acceptable intersection index I_{B2}. In practice, the Paddick index must be bigger than the parameter $IPdkMin$ and the intersection index I_{B2} must be smaller than the parameter $IB2Max$ to get a valid node;
- The first valid node becomes the incumbent node;
- The I_{PDK} index of a new valid node is compared with the I_{PDK} index of the incumbent node. If it is greater, the new node is set to be the incumbent node.

6.3. Node generation

The search strategy considers that all tree nodes are available when taking a decision to jump from the current node to another node. As the nodes are heuristic solutions of (PN) problems associated to a fixed selection of spheres, there must be a way to generate a new node, starting from the current node. We introduce now the stochastic component of our method. The new node is generated by inserting a new sphere into the current set of covering spheres, at a spatially uniform random location outside them. The optimization process then runs for this new (PN) problem and we get the new node.

Considering the full search space, which includes the selection of spheres and their locations, this strategy can be seen as a stochastic local search whose candidate generator procedure is the process describe above. Since new nodes are heuristic solutions, in practice this node generation process is iterated a limited number of times to choose the best valid node, based on its I_{PDK} index.
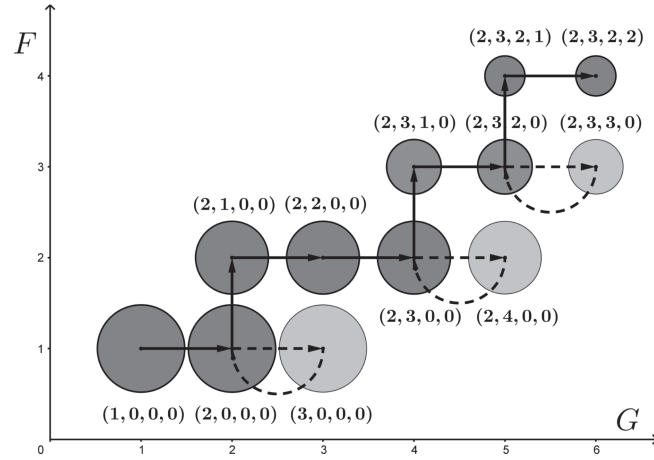
FIGURE 13. Sphere allocation diagram, detailing the search strategy in the set S, for $F = 4$ and $G = 6$.

7. COMPUTATIONAL IMPLEMENTATION AND RESULTS

7.1. Algorithm pseudocode

The proposed method was implemented and tested. For the sake of clarity, a simplified pseudocode of the method is listed in Algorithm 1. It describes how the various components of the method, as presented in the previous sections, are put together. Some parameters and variables used in the pseudocode are explained below:

- Parameter $IPdkMin$: Minimum acceptable value for the Paddick index;
- Parameter $IB2Max$: Maximum acceptable value for the intersection index I_{B2};
- Variable $CNode$: Represents the current node with a specific selection of spheres and spheres' centers configuration. It is used as an index for the computed conformity indexes, for example, as in $IPdk(CNode)$. An optimized $CNode$ is a current node after solving (PN), which heuristically chooses the best place for the centers of the spheres;
- Parameter $NumIterEachNode$: The number of times a $CNode$ is optimized;
- Variable $BestCNode$: Stores the best $CNode$ found after $NumIterEachNode$ iterations;
- Variable $IncumbentNode$: The so far best $CNode$ that satisfies $IPdk(CNode) \geq IPdkMin$ and $IB2(CNode) \leq IB2Max$, representing an admissible covering structure.

Some of the details that were omitted in the pseudocode listing include, for example, timestamping calls and events logging, which are important in a practical implementation.

7.2. Test instances

The test instances were created based on the Gamma Knife application. The set of available spheres is presented in Table 1.

The selected solids for the experiments are all convex, but the surface ∂T of some solids is non differentiable. They are listed in Table 2.

The specific parallel body T_d associated to each solid was calculated and implemented in order to define the function $\mathcal{D}(x, T_d)$ used to create the function q_μ.

7.3. Setup for the computational experiments

We list below the main characteristics of the computational environment used to run the experiments:

Algorithm 1. Find a heuristic covering of the solid T.

Input: A solid T to be covered and an ordered pool of balls
Output: A heuristic covering with $IPdk \geq IPdkMin$ and $IB2 \leq IB2Max$
  Create a T based grid for quasi-Monte Carlo calculation [15] of conformity indexes
  $IncumbentNode \leftarrow null$
  Build p_u for each sphere and q_μ for the solid T
  Get first available ball and create $CNode$
  loop
    Locate last included ball at a random outside position
    Optimize $CNode$ heuristically solving the associated PN
    $BestCNode \leftarrow CNode$
    for $i = 1$ **to** $NumIterEachNode$ **do**
      Locate last included ball at a random outside position
      Optimize $CNode$ heuristically solving the associated PN
      if $IPdk(CNode) > IPdk(BestCNode)$ **and** $IB2(CNode) \leq IB2Max$ **then**
        $BestCNode \leftarrow CNode$
      end if
    end for
    $CNode \leftarrow BestCNode$
    if $IB2(CNode) > IB2Max$ **then**
      Backtrack and select next ball type
    else if $IPdk(CNode) \geq IPdkMin$ **then**
      if $IncumbentNode = null$ **or** $IPdk(CNode) > IPdk(IncumbentNode)$ **then**
        $IncumbentNode \leftarrow CNode$
      end if
    end if
    Get Next Available ball and create $CNode$
    if it was not possible to get another ball from the pool **then**
      break
    end if
  end loop
  if $IncumbentNode \neq null$ **then**
    print $IncumbentNode$
  end if

TABLE 1. Pool of available spheres for Gamma Knife application.

Sphere type	Radius	Available
Type a	9	8
Type b	7	8
Type c	4	8
Type d	2	10

TABLE 2. Characteristics of the selected solids.

Solid	Length	Width	Height	∂T
Parallelepiped	15	9	9	Non diff.
Cube	12	12	12	Non diff.
Sphere	12	12	12	Diff.
Prolate ellipsoid	15	9	9	Diff.
Oblate ellipsoid	9	15	15	Diff.

TABLE 3. Results of the computational experiments.

Solid	Signature	Parameters		Covering indexes				
		IPm	$IB2M$	I_{COV}	I_{E1}	I_{B2}	I_{B3}	I_{PDK}
Parallelepiped	$(3,1,5,0)$	70%	34%	80%	91%	33%	6%	73%
	$(2,6,0,0)$	70%	40%	85%	89%	40%	14%	76%
Cube	$(4,4,1,0)$	70%	34%	91%	91%	26%	6%	83%
Sphere	$(2,2,8,0)$	70%	34%	91%	89%	34%	13%	82%
Prolate ellipsoid	$(2,2,2,0)$	70%	34%	99%	77%	31%	10%	76%
Oblate ellipsoid	$(3,2,4,0)$	70%	34%	95%	87%	34%	6%	83%

- Non linear solver: CONOPT v 3.14 [4];
- Parameter ε was arbitrarily set to 1.5, an adequate fraction of the linear dimensions of the selected solids;
- Too big spheres are immediately rejected, if they don't fit in T_ε;
- A 10.000 points grid was set for Quasi-Monte Carlo volume computations [15];
- Three iterations were employed to generate each new node, taking the best one;
- Valid node: $IPdkMin = 70\%$ and $IB2Max = 34\%$ (or 40% for some hard cases).

Some comments must be made regarding the last item. Firstly, increasing $IPdkMin$ too much (setting it to 75%, for example), reduces drastically the success rate of the algorithm: it finishes without finding valid nodes. This is caused by the fact that I_{PDK} is an I_Δ index, harder to be satisfied than a simpler covering index like I_{PITV}.

Secondly, reducing $IB2Max$ too much (setting it to 30%, for example), again reduces drastically the success rate of the algorithm. The value 34%, for example, was adequate for solids with differentiable ∂T surfaces, but we had to allow bigger values (like 40%, for example) for some cube and parallelepiped instances, apparently because these solids present non differentiable ∂T surfaces.

7.4. Results

A small sample of the results of the computational experiments are presented in Table 3, where the parameters $IPdkMin$ and $IB2Max$ are abbreviated to IPm and $IB2M$.

The running times were always low, in general considerably less than $300\,\text{s}$. This was caused in part by the small size of the chosen instances of the (actual) problem to be solved, but also by the "bigger spheres first" characteristic of the proposed search strategy.

It is interesting to visually assess the quality of a covering. For this purpose, and based on the results presented in Table 3, Figure 14 presents the covering of a cube, Figure 15 presents the covering of a prolate ellipsoid and Figure 16 presents the covering of an oblate ellipsoid.

7.5. Comparisons with other works

Albeit it is not possible to make direct comparisons between the presented method and other methods mentioned in the literature, mainly due to the specific combination of indexes I_{PDK} and I_{B2} used to measure covering quality, this section presents some trade-off comparisons.

In [20] alternative methods are presented to cover a parallelepiped T with dimensions $14 \times 12 \times 10$ and security region with $\varepsilon = 1$ using spheres with radii in the same set $\{9, 7, 4, 2\}$. The best methods, "XPRESS SLP", "Heur IPOPT" and "Graph B&C", were chosen to make the comparisons, as presented in Table 4. "XPRESS SLP" is an algorithm that employs sucessive linear programming (SLP) techniques, "Heur IPOPT" is a VNS-based heuristic that uses IPOPT as its main subsolver for nonlinear programming and "Graph B&C" encodes a branch and cut algorithm based on a graph representation of the solutions. The first column lists the compared aspects, where $|S|$ indicates the number of spheres used in the covering structure and t (sec) indicates the running time
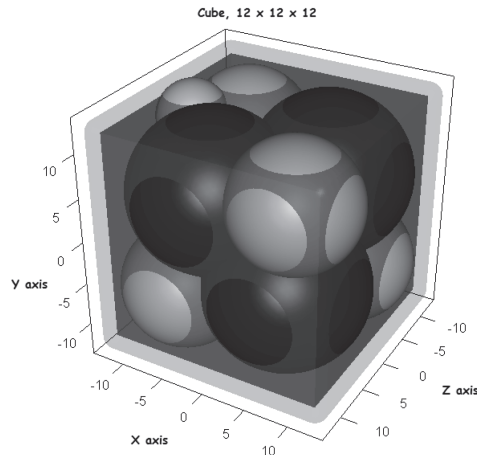
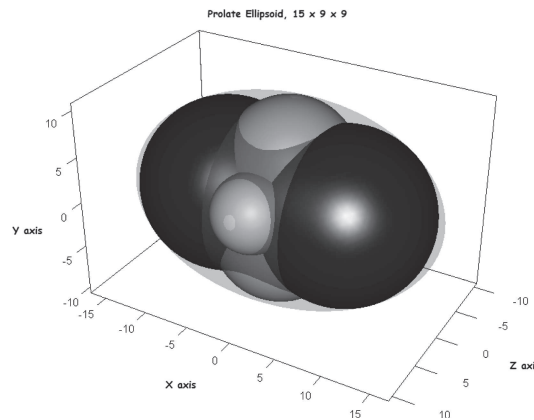FIGURE 14. Covering a cube with unequal spheres.

FIGURE 15. Covering a prolate ellipsoid with unequal spheres.

in seconds. Indexes I_{PDK} and I_{B2} were calculated using their spheres' centers positions. Due to its stochastic nature, two results obtained *via* Algorithm 1, "Sol. 1" and "Sol. 2", are also presented.

Despite not being specifically designed to maximize the covering index, both solutions obtained with the method developed in this work present better I_{COV} values. Better Paddick values were expected and ultimately obtained. Also, it is remarkable that the number of spheres used is much smaller than the number of spheres used by the other methods, which is certainly a good property in the context of Gamma Knife treatment planning. The drawbacks are a worse I_{B2} index and a higher running time than the first two methods (although perfectly acceptable for the intended application). Increasing $IB2Max$ to 42% allowed us to obtain "Sol. 2", which presents even better I_{COV} and I_{PDK} indexes, at the cost of even worse values for the index I_{B2} and the running time.

Still using the same radii set $\{9, 7, 4, 2\}$, Pinto [20] includes comparisons with the instances presented in Liberti *et al.* [11]. The instance bt4 is a sphere T with radius 10 and security region defined with $\varepsilon = 6$. It was
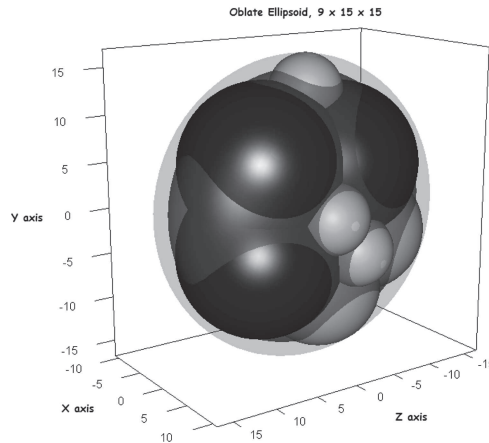
FIGURE 16. Covering an oblate ellipsoid with unequal spheres.

TABLE 4. Comparison of results for parallelepiped solid using different methods.

Parallelepiped T $\varepsilon = 1$	XPRESS SLP	Heur IPOPT	Graph B&C	Algorithm 1 Sol. 1	Sol. 2
$|S|$	35	29	74	19	20
t (sec)	5	15	747	46	51
I_{COV}	89.15	91.41	92.62	94.13	95.13
I_{PDK}	74.58	76.81	74.25	78.81	78.91
I_{B2}	15.83	15.33	24.30	33.75	40.94

TABLE 5. Comparison of results for instance `bt4`.

Sphere T $\varepsilon = 6$	Liberti *et al.* P_2'' (CPLEX)	Pinto B&C	Algorithm 1
$|S|$	3	6	1
t (sec)	58608	1	52
I_{COV}	90.7	98.1	73.5
I_{PDK}	–	41.5	73.5
I_{B2}	–	30.8	0.0

chosen to make the comparison, as presented in Table 5, because of the peculiar behavior of Algorithm 1 when dealing with this instance[9].

At first glance it may be surprising to have a covering structure using only one sphere, as obtained by using Algorithm 1. But considering that this algorithm looks for the valid node with the best I_{PDK}, it will reject some valid nodes, even with better I_{COV} values (which can be obtained during the search process), in favor of valid nodes with better I_{PDK} values. For the instance bt4, the high value of ε implies that increasing I_{COV} means including new spheres inside the security region, but with a large portion of their volumes outside the solid T,

---

[9]It was not possible to compute I_{PDK} and I_{B2} for the result presented in reference [11] because it did not publish the locations of the spheres' centers.

which reduces I_{PDK}. This instance is a good example of the difference between a conventional optimal covering (best I_{COV}) and the specific optimal covering considered in this work (best I_{PDK}).

8. Conclusion and future research

The results presented in this article demonstrate that it is possible to obtain good solutions for practical covering problems like the Gamma Knife treatment planning employing the presented method.

The use of the Paddick index to select valid nodes gives to the method an early compliance with the expected good coverings, and better nodes are just promoted to be the incumbent. On the other hand, the first intersection index provides a way to avoid accumulating radiation exposure beyond acceptable values on certain parts of a tumor (or healthy tissue on the security region).

It is the aim of future research to:

- Analyze whether it is worthy or not to expand the search scope, currently constrained to the order imposed by the sizes of the available spheres;
- Explore alternative penalty curves for the distance between two spheres;
- Explore the possibility of creating constraints that directly restrict the values of the intersection indexes.
- Deal with non-convex solids, starting with simple geometric configurations.
- Develop general constraints to deal with complex solids.
- Analyze the possibility of proposing a single index that takes into consideration the values of the intersection indexes.

References

[1] J.D. Bourland and Q.J. Wu, Morphology-guided radiosurgery treatment planning and optimization for multiple isocenters. *Med. Phys.* **26** (1999) 2151–2160.

[2] S. Burer, A.N. Letchford, Non-convex mixed-integer nonlinear programming: A survey. *Surv. Oper. Res. Manage. Sci.* **17** (2012) 97–106.

[3] J.H. Conway and N.J.A. Sloane, Sphere Packings, Lattices and Groups, 3rd edition. Springer-Verlag, New York, NY (1999).

[4] A. Drud, CONOPT: A GRG code for large sparse dynamic nonlinear optimization problems. *Math. Program.* **31** (1985) 153–191.

[5] D.Z. Du, P. Pardalos and J. Wang, Vol. 55 of *Discrete Mathematical Problems with Medical Applications*. American Mathematical Society, Providence, RI (2000).

[6] M. Ferris and D. Shepard, Optimization of Gamma Knife Radiosurgery. In: Vol. 55 of *Discrete Mathematical Problems with Medical Applications*, edited by D.Z. Du, P. Pardalos, J. Wang. DIMACS Series in Discrete Mathematics and Theoretical Computer Science, American Mathematical Society, Providence, RI (2000) 27–44.

[7] M. Ferris, J. Lim and D. Shepard, An optimization approach for the radiosurgery treatment planning. *SIAM J. Optim.* **13** (2003) 921–937.

[8] C.A. Floudas, Nonlinear and Mixed-integer Optimization: Fundamentals and Applications. Oxford University Press, Oxford (1995).

[9] S. Jitprapaikulsarn, An optimization-based treatment planner for gamma knife radiosurgery, Ph.D. thesis, Case Western Reserve University, Cleveland, OH (2005).

[10] J.E. Jones, On the Determination of Molecular Fields. II. From the Equation of State of a Gas. In: Vol. 106 of *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. The Royal Society of London, London (1924) 463–477.

[11] L. Liberti, N. Maculan and Y. Zhang, Optimal configuration of gamma ray machine radiosurgery units: The sphere covering subproblem. *Optim. Lett.* **3** (2009) 109–121.

[12] J. Lim, Optimization in radiation treatment planning, Ph.D. thesis, University of Wisconsin, Madison, WI (2002).

[13] T. Maekawa, Self-intersections of offsets of quadratic surfaces: Part I, explicit surfaces. *Eng. Comput.* **14** (1998) 1–13.

[14] T. Maekawa, Self-intersections of offsets of quadratic surfaces: Part II, implicit surfaces. *Eng. Comput.* **14** (1998) 14–22.

[15] W.J. Morokoff and R.E. Caflisch, Quasi-monte carlo integration. *J. Comput. Phys.* **122** (1995) 218–230.

[16] T. Motzkin, Sur quelques propriétés caractéristiques des ensembles convexes. *Atti Acad. Naz. Lincei. Rend. VI* **21** (1935) 562–567.

[17] R.Q. Nascimento, A.F.U.S. Macambira, L.F. Cabral, R.V. Pinto, The discrete ellipsoid covering problem: A discrete geometric programming approach. *Discret. Appl. Math.* **164** (2014) 276–285.

[18] I.I. Paddick, A simple scoring ratio to index the conformity of radiosurgical treatment plans. *J. Neurosurg.* **93** (2000) 219–222.

[19] P. Pardalos and H. Edwin, Vol. 26 of *Handbook of Optimization in Medicine*, Series Springer Optimization and its Applications. Springer US, New York, NY (2009).

[20] R.V. Pinto, O problema de recobrimento de sólidos por esferas de diâmetros diferentes. Tese de Doutorado, COPPE/UFRJ, Rio de Janeiro (2015).

[21] E. Shaw, R. Kline, M. Gillin, L. Souhami, A. Hirschfeld, R. Dinapoli, *et al.*, Radiation therapy oncology group: Radiosurgery quality assurance guidelines. *Int. J. Radiat. Oncol. Biol. Phys.* **27** (1993) 1231–1239.

[22] D.M. Shepard, M.C. Ferris, R. Ove and L. Ma, Inverse treatment planning for gamma knife radiosurgery. *Med. Phys.* **27** (2000) 2748–2756.

[23] A. Soutou and Y. Dai, Global optimization approach to unequal sphere packing problems in 3D. *J. Optim. Theory Appl.* **114** (2002) 671–694.

[24] K. Bezdek, Classical Topics in Discrete Geometry. Springer US, New York, NY (2010).

[25] C. Uhler and S.J. Wright, Packing ellipsoids with overlap. *Soc. Ind. Appl. Math.* **55** (2013) 4.

[26] M.B.K. Venceslau, O problema de recobrimento mínimo de um corpo em três dimensões por esferas de diferentes raios. Tese de Doutorado, COPPE/UFRJ, Rio de Janeiro (2015).

[27] H.M. Venceslau, D.C. Lubke and A.E. Xavier, Optimal covering of solid bodies by spheres via the hyperbolic smoothing technique. *Optim. Meth. Softw.* **30** (2014) 391–403.

[28] A.E. Xavier and A.A.F.D. Oliveira, Optimal covering of plane domains by circles via hyperbolic smoothing. *J. Glob. Optim.* **31** (2005) 493–504.