

Instituto de Engenharia de Sistemas e Computadores de Coimbra
Institute of Systems Engineering and Computers
INESC - Coimbra

Humberto Rocha Joana Matos Dias
Brígida da Costa Ferreira Maria do Carmo Lopes

**Beam angle optimization in intensity-modulated radiation therapy
using radial basis functions within the pattern search methods framework**

No. 1

2012

ISSN: 1645-2631

Instituto de Engenharia de Sistemas e Computadores de Coimbra
INESC - Coimbra
Rua Antero de Quental, 199; 3000-033 Coimbra; Portugal
www.inescc.pt



Financiamento participado pelo Fundo Social Europeu e por Fundos Nacionais do MCTES



Beam angle optimization in IMRT using RBFs within the pattern search methods framework

H. Rocha ^{*} J. M. Dias ^{*,†} B.C. Ferreira ^{§,‡} M.C. Lopes ^{§,‡}

January 3, 2012

Abstract

In radiation therapy treatment planning, the selection of appropriate radiation incidence directions is decisive for the quality of the treatment plan, both for appropriate tumor coverage and for enhance better organs sparing. However, the beam angle optimization (BAO) problem is yet to be solved in a satisfactory way. The objective of this research report is to discuss the benefits of using radial basis functions within the pattern search methods framework in the optimization of the highly non-convex BAO problem. The pattern search methods framework is composed by a search step and a poll step at each iteration. The poll step performs a local search in a mesh neighborhood and assures convergence to a local minimizer or stationary point. The search step provides the flexibility for a global search since it allows searches away from the neighborhood of the current iterate. Radial basis functions are used and tested in this step both to influence the quality of the local minimizer found by the method and to obtain a better coverage of the search space in amplitude. A set of clinical examples of head-and-neck cases is used to discuss the benefits of using this approach in the optimization of the BAO problem.

Key words. Radiotherapy, IMRT, Beam Angle Optimization, Direct Search, Radial Basis Functions.

^{*}*INESC-Coimbra, Coimbra, Portugal.*

[†]*Faculdade de Economia, Universidade de Coimbra, Coimbra, Portugal.*

[‡]*I3N, Departamento de Física, Universidade de Aveiro, Aveiro, Portugal.*

[§]*Serviço de Física Médica, IPOC-FG, EPE, Coimbra, Portugal.*

1 Introduction

The purpose of radiation therapy is to deliver a dose of radiation to the tumor volume to sterilize all cancer cells minimizing the collateral effects on the surrounding healthy organs and tissues. Typically, radiation is generated by a linear accelerator mounted on a gantry that can rotate along a central axis and is delivered with the patient immobilized on a couch that can rotate (see Fig. 1). The rotation of the couch combined with the rotation of the gantry allows radiation from almost any angle around the tumor. Many authors consider non-coplanar angles [1, 2, 3, 4, 5, 6]. However, despite the fact that almost every angle is possible for radiation delivery, the use of coplanar angles is predominant. This is a way to simplify an already complex problem, and the angles considered lay in the plane of the rotation of the gantry around the patient. In clinical practice, most of the time, the number of beam angles is assumed to be defined a priori by the treatment planner and the beam directions are still manually selected by the treatment planner that relies mostly on his experience, despite the evidence presented in the literature that appropriate radiation beam incidence directions can lead to a plan's quality improvement [5, 7, 8].



Figure 1: Linear accelerator rotating through different angles [9].

An important type of radiation therapy is intensity modulated radiation therapy (IMRT), where the radiation beam is modulated by a multileaf collimator. Multileaf collimators enable the transformation of the beam into a grid of smaller beamlets of independent intensities. A common way to solve the inverse planning in IMRT optimization problems is

to use a beamlet-based approach leading to a large-scale programming problem. Due to the complexity of the whole optimization problem, many times the treatment planning is divided into three smaller problems which can be solved sequentially: BAO problem, fluence map optimization (FMO) problem, and leaf sequencing problem. Most of the efforts in the IMRT optimization community have been devoted at optimizing beamlet intensities [10], once the beam angles have been selected by the treatment planner. Comparatively fewer research effort has been directed to the optimization of beam angles [11]. Here we will focus our attention in the BAO problem, using coplanar angles, and will assume that the number of beam angles is defined a priori by the treatment planner.

Many attempts to address the BAO problem can be found in the literature including simulated annealing [3, 12, 13], genetic algorithms [11, 14, 15, 16], particle swarm optimization [17] or other heuristics incorporating a priori knowledge of the problem. Although those global heuristics can avoid local optima theoretically, globally optimal or even clinically better solutions can not be obtained without a large number of objective function evaluations. The concept of beam's eye view has been a popular approach to address the BAO problem as well [3, 18, 19]. The concept is similar to a bird's eye view, where the object being viewed is the tumor as seen from a beam. The bigger the area of the tumor and the smaller the area of the surrounding organs is seen by the beam, the better candidate the beam is to be used in the treatment plan. Other approaches include the projection of the surrounding organs into the tumor. Pugachev and Xing [20] present a computer assisted selection of coplanar angles based on a variation of the beam's eye view concept. Ehrgott et al. [11] discuss a mathematical framework that unifies the approaches found in literature. Aleman et al. [2] propose a response surface approach and include non-coplanar angles in beam orientation optimization. Lim and Cao [21] propose an approach that consists of two sequential phases: branch-and-prune and local neighborhood search. Lee et al. [22] suggests a mixed integer programming (MIP) approach for simultaneously determining an optimal intensity map and optimal beam angles for IMRT delivery. Schreibmann et al. [23] propose a hybrid multiobjective evolutionary optimization algorithm for IMRT inverse planning and apply it to the optimization of the number of incident beams, their orientations and intensity profiles. Other approaches include maximal geometric separation of treatment beams [5] or gradient searches [10].

Here, similarly to [1, 2, 5, 10, 15, 23, 24, 25], we will use the optimal solution of the FMO problem to drive our BAO problem. Most of the previous BAO studies are based on a variety of scoring methods or approximations to the FMO to gauge the quality of the beam angle set. When the BAO problem is not based on the optimal FMO solutions, the resulting beam angle set has no guarantee of optimality and has questionable reliability since it has been extensively reported that optimal beam angles for IMRT are often non-intuitive [26]. The BAO problem is quite difficult since it is a highly non-convex optimization problem with many local minima [10, 27]. Therefore, methods that avoid being easily trapped in local minima should be used. Obtaining the optimal solution for a beam angle set is time costly and even if only one beam angle is changed in that set, a complete dose computation is required in order to compute and obtain the corresponding optimal FMO solution. To minimize this time issue, methods that require few function value evaluations should be used to tackle the BAO problem. The pattern search methods framework is suited to address the BAO problem since it requires few function value evaluations to converge and have the ability to avoid local entrapment. Here, we will discuss the benefits of using radial basis functions within the pattern search methods framework in the optimization of the highly non-convex BAO problem. The pattern search methods framework is composed by a search step and a poll step at each iteration. The poll step perform a local search in a mesh neighborhood and assures convergence to a local minimizer or stationary point. The search step provides the flexibility for a global search since it allows searches away from the neighborhood of the current iterate. Radial basis functions are used and tested in this step both to influence the quality of the local minimizer found by the method and also to obtain a better coverage of the search space in amplitude. A set of clinical examples of head-and-neck cases is used to discuss the benefits of using this approach in the optimization of the BAO problem. The research report is organized as follows. In the next section we describe the BAO problem formulation and the coupled FMO problem formulation. Section 3 briefly presents the pattern search methods framework used. Radial basis functions interpolation and its use within the search step of the pattern search methods framework is presented in section 4. Clinical examples of head-and-neck cases used in the computational tests are presented in section 5. Section 6 presents the obtained results. In the last section we have the conclusions and future work.

2 Beam angle optimization problem

In order to model the BAO problem as a mathematical programming problem, a quantitative measure to compare the quality of different sets of beam angles is required. For the reasons presented in section 1, our approach for modeling the BAO problem uses the optimal solution value of the FMO problem as the measure of the quality of a given beam angle set. Thus, we will present the formulation of the BAO problem followed by the formulation of the FMO problem we used.

2.1 BAO model

Let us consider n to be the fixed number of (coplanar) beam directions, i.e., n beam angles are chosen on a circle around the CT-slice of the body that contains the isocenter (usually the center of mass of the tumor). Typically, the BAO problem is formulated as a combinatorial optimization problem in which a specified number of beam angles is to be selected among a beam angle candidate pool. The continuous $[0^\circ, 360^\circ]$ gantry angles are generally discretized into equally spaced directions with a given angle increment, such as 5 or 10 degrees. We could think of all possible combinations of n beam angles as an exhaustive global search method. However, this requires an enormous amount of time to calculate and compare all dose distributions for all possible angle combinations. For example if we choose $n = 5$ angles out of 72 candidate beam angles $\{0, 5, \dots, 355\}$, there are $C_5^{72} = 13,991,544$ combinations. By decreasing the number of candidate beam angles to 36, $\{0, 10, \dots, 350\}$, the number of different combinations is still $C_5^{36} = 376,992$, requiring an enormous amount of time to compare all the resulting plans regardless the measure considered. Therefore, an exhaustive search of a large-scale combinatorial problem is considered to be too slow and inappropriate for a clinical setting. Many heuristics and meta-heuristics have been presented as an attempt to reduce the number of combinations to compare. However, most require a prohibitive number of function evaluations when the measure considered is the optimal value of the FMO problem.

We will consider a different approach for the formulation of the BAO problem. All continuous $[0^\circ, 360^\circ]$ gantry angles will be considered instead of a discretized sample. Since the angle -5° is equivalent to the angle 355° and the angle 365° is the same as the angle 5° ,

we can avoid a bounded formulation. A basic formulation for the BAO problem is obtained by selecting an objective function such that the best set of beam angles is obtained for the function's minimum:

$$\begin{aligned} \min \quad & f(\theta_1, \dots, \theta_n) \\ \text{s.t.} \quad & (\theta_1, \dots, \theta_n) \in \mathbb{R}^n. \end{aligned} \tag{1}$$

Here, the objective $f(\theta_1, \dots, \theta_n)$ that measures the quality of the set of beam directions $\theta_1, \dots, \theta_n$ is the optimal value of the FMO problem for each fixed set of beam directions. Such functions have numerous local optima, which increases the difficulty of obtaining a good global solution. Thus, the choice of the solution method becomes a critical aspect for obtaining a good solution. Our formulation was mainly motivated by the ability of using a class of solution methods that we consider to be suited to successfully address the BAO problem: pattern search methods. The FMO model used is presented next.

2.2 FMO model

For a given beam angle set, an optimal IMRT plan is obtained by solving the FMO problem - the problem of determining the optimal beamlet weights for the fixed beam angles. Many mathematical optimization models and algorithms have been proposed for the FMO problem, including linear models [28], mixed integer linear models [29], nonlinear models [30], and multiobjective models [31].

Radiation dose distribution deposited in the patient, measured in Gray (Gy), needs to be assessed accurately in order to solve the FMO problem, i.e., to determine optimal fluence maps. Each structure's volume is discretized in voxels (small volume elements) and the dose is computed for each voxel using the superposition principle, i.e., considering the contribution of each beamlet. Typically, a dose matrix D is constructed from the collection of all beamlet weights, by indexing the rows of D to each voxel and the columns to each beamlet, i.e., the number of rows of matrix D equals the number of voxels (N_v) and the number of columns equals the number of beamlets (N_b) from all beam directions considered. Therefore, using matrix format, we can say that the total dose received by the voxel i is given by $\sum_{j=1}^{N_b} D_{ij}w_j$, with w_j the weight of beamlet j . Usually, the total number of voxels considered reaches the tens of thousands, thus the row dimension of the dose matrix is

of that magnitude. The size of D originates large-scale problems being one of the main reasons for the difficulty of solving the FMO problem.

Here, we will use a convex penalty function voxel-based nonlinear model [24]. In this model, each voxel is penalized according to the square difference of the amount of dose received by the voxel and the amount of dose desired/allowed for the voxel. This formulation yields a quadratic programming problem with only linear non-negativity constraints on the fluence values [28]:

$$\min_w \sum_{i=1}^{N_v} \frac{1}{v_S} \left[\underline{\lambda}_i \left(T_i - \sum_{j=1}^{N_b} D_{ij} w_j \right)_+^2 + \bar{\lambda}_i \left(\sum_{j=1}^{N_b} D_{ij} w_j - T_i \right)_+^2 \right]$$

$$s.t. \quad w_j \geq 0, \quad j = 1, \dots, N_b,$$

where T_i is the desired dose for voxel i , $\underline{\lambda}_i$ and $\bar{\lambda}_i$ are the penalty weights of underdose and overdose of voxel i , and $(\cdot)_+ = \max\{0, \cdot\}$. Although this formulation allows unique weights for each voxel, similarly to the implementation in [24], weights are assigned by structure only so that every voxel in a given structure has the weight assigned to that structure divided by the number of voxels of the structure (v_S). This nonlinear formulation implies that a very small amount of underdose or overdose may be accepted in clinical decision making, but larger deviations from the desired/allowed doses are decreasingly tolerated [24].

The optimal set of beam angles depends on the penalty weight values of the objective function selected. For example, for a head-and-neck cancer case, a higher penalty weight on the parotid objective results in a set of beams which will enhance better parotid sparing. For multi-objective IMRT optimization [31], each Pareto optimal treatment plan will have a distinct set of optimal beam angles. For traditional trial-and-error parameter tuning for IMRT planning, it is not clear how BAO should be incorporated into the planning process [10]. The penalty weights were manually selected to obtain acceptable treatment plans.

The FMO model is used as a black-box function. Other models used before for BAO include convex penalty function structure-based approaches [25] and a variety of linear approaches [10, 32]. It is beyond the scope of this study to discuss if this formulation of the FMO problem is preferable to others. The conclusions drawn regarding BAO coupled with this nonlinear model are valid also if different FMO formulations are considered.

3 Pattern search methods

Pattern search methods are directional direct search methods that belong to a broader class of derivative-free optimization methods (see [33] for a detailed overview of derivative-free optimization methods), such that iterate progression is solely based on a finite number of function evaluations in each iteration, without explicit or implicit use of derivatives. Since we are interested in resolution methods for the beam angle problem formulation presented in Eq. (1), we will summarily describe pattern search methods for unconstrained optimization problems of the form

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbb{R}^n, \end{aligned}$$

where the decision vector $\mathbf{x} = (x_1, \dots, x_n)$ is used as input into the black-box function f .

Pattern search methods are iterative methods generating a sequence of non-increasing iterates $\{\mathbf{x}^k\}$ using positive bases (or positive spanning sets) and moving towards a direction that would produce a function decrease. A positive basis for \mathbb{R}^n can be defined as a set of nonzero vectors of \mathbb{R}^n whose positive combinations span \mathbb{R}^n (positive spanning set), but no proper set does. A positive spanning set contains at least one positive basis. It can be shown that a positive basis for \mathbb{R}^n contains at least $n + 1$ vectors and cannot contain more than $2n$ [34]. Positive basis with $n + 1$ and $2n$ elements are referred to as minimal and maximal positive basis, respectively. Commonly used minimal and maximal positive basis are $[I - e]$, with I being the identity matrix of dimension n and $e = [1 \ \dots \ 1]^\top$, and $[I - I]$, respectively (see Figure 2).

One of the main features of positive bases (or positive spanning sets), that is the motivation for directional direct search methods, is that, unless the current iterate is at a stationary point, there is always a vector \mathbf{v}^i in a positive basis (or positive spanning set) that is a descent direction [34], i.e., there is an $\alpha > 0$ such that $f(\mathbf{x}^k + \alpha\mathbf{v}^i) < f(\mathbf{x}^k)$. This is the core of directional direct search methods and in particular of pattern search methods. The notions and motivations for the use of positive bases, its properties and examples can be found in [34, 35].

The pattern search methods framework [36] is the class of the most used and implemented directional direct search methods. Pattern search methods framework as described in [36] or later in [35] is presented next. Let us denote by \mathbf{V} the $n \times p$ matrix whose

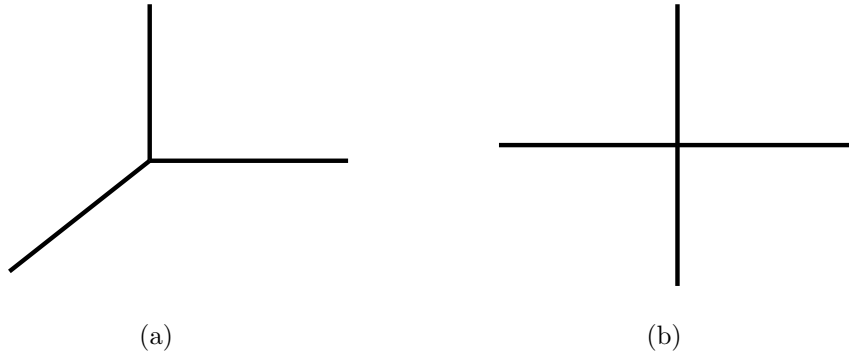


Figure 2: Examples of minimal (a) and maximal (b) positive bases in \mathbb{R}^2 .

columns correspond to the p ($\geq n + 1$) vectors forming a positive spanning set. Given the current iterate \mathbf{x}^k , at each iteration k , the next point \mathbf{x}^{k+1} , aiming to provide a decrease on the objective function, is chosen from a finite number of candidates on a given mesh $M_k = \{\mathbf{x}^k + \alpha_k \mathbf{V}\mathbf{z} : \mathbf{z} \in \mathbb{Z}_+^p\}$, where α_k is the mesh-size (or step-size) parameter and \mathbb{Z}_+ is the set of nonnegative integers.

Pattern search methods are organized around two steps at every iteration. The first step consists of a finite search on the mesh, free of rules, with the goal of finding a new iterate that decreases the value of the objective function at the current iterate. This step, called the search step, has the flexibility to use any strategy, method or heuristic, or take advantage of a priori knowledge of the problem at hand, as long as it searches only a finite number of points in the mesh. The search step provides the flexibility for a global search since it allows searches away from the neighborhood of the current iterate, and influences the quality of the local minimizer or stationary point found by the method.

If the search step fails to produce a decrease in the objective function, a second step, called the poll step, is performed around the current iterate. The poll step follows stricter rules and, using the concepts of positive bases, attempts to perform a local search in a mesh neighborhood around \mathbf{x}^k , $\mathcal{N}(\mathbf{x}^k) = \{\mathbf{x}^k + \alpha_k \mathbf{v} : \text{for all } \mathbf{v} \in P_k\} \subset M_k$, where P_k is a positive basis chosen from the finite positive spanning set \mathbf{V} . For a sufficiently small mesh-size parameter α_k , the poll step is guaranteed to provide a function reduction, unless the current iterate is at a stationary point [35]. So, if the poll step also fails to produce a function reduction, the mesh-size parameter α_k must be decreased. On the other hand, if both the search and poll steps fail to obtain an improved value for the objective function,

the mesh-size parameter is increased or held constant. We are able to describe now the pattern search methods framework.

Algorithm 3.1 (Pattern search methods framework)

- 0. Initialization** Set $k = 0$. Choose $\mathbf{x}^0 \in \mathbb{R}^n$, $\alpha_0 > 0$, and a positive spanning set \mathbf{V} .
Choose a rational number $\tau > 1$ and an integer number $m_{max} \geq 1$.
- 1. Search step** Evaluate f at a finite number of points in M_k with the goal of decreasing the objective function value at \mathbf{x}^k . If $\mathbf{x}^{k+1} \in M_k$ is found satisfying $f(\mathbf{x}^{k+1}) < f(\mathbf{x}^k)$, go to step 4 and expand M_k . Both search step and iteration are declared successful. Otherwise, go to step 2 and search step is declared unsuccessful.
- 2. Poll step** This step is only performed if the search step is unsuccessful. If $f(\mathbf{x}^k) \leq f(\mathbf{x})$ for every \mathbf{x} in the mesh neighborhood $\mathcal{N}(\mathbf{x}^k)$, go to step 3 and shrink M_k . Both poll step and iteration are declared unsuccessful. Otherwise, choose a point $\mathbf{x}^{k+1} \in \mathcal{N}(\mathbf{x}^k)$ such that $f(\mathbf{x}^{k+1}) < f(\mathbf{x}^k)$, go to step 4 and expand M_k . Both poll step and iteration are declared successful.
- 3. Mesh reduction** Let $\alpha_{k+1} = \tau^{m_k^-} \alpha_k$, with $-m_{max} \leq m_k^- \leq -1$. Set $k = k + 1$ and return to step 1 for a new iteration.
- 4. Mesh expansion** Let $\alpha_{k+1} = \tau^{m_k^+} \alpha_k$, with $0 \leq m_k^+ \leq m_{max}$. Set $k = k + 1$ and return to step 1 for a new iteration.

The most common choice for the mesh-size parameter update is to half the mesh-size parameter at unsuccessful iterations and to keep it or double it at successful ones. The purpose of the mesh-size parameter is twofold: to bound the size of the minimization step and also to control the local area where the function is sampled around the current iterate. Most derivative-free methods couple the mesh-size (or step-size) with the size of the sample set (or the search space). The initial mesh-size parameter value defined in [37] for comparison of several derivative-free optimization algorithms is $\alpha_0 = \max\{1, \|\mathbf{x}^0\|_\infty\}$. This choice of initial mesh-size parameter is commonly used as default in many derivative-free algorithms such as implementations of the Nelder-Mead method [38]. However, this choice of initial mesh-size parameter might not be adequate. Note that, if the initial mesh

parameter is a power of 2, ($\alpha_0 = 2^l, l \in \mathbb{N}$), and the initial point is a vector of integers, using this common mesh update, all iterates will be a vector of integers until the mesh-size parameter becomes inferior to 1. This possibility is rather interesting for the BAO problem.

Typically, the stopping criteria of the pattern search methods is based either on convergence criteria related with the mesh-size parameter or on the maximum number of function value evaluations allowed. Pattern search methods share the following convergence result, provided that each column vector \mathbf{v}_i of \mathbf{V} is given by $\mathbf{G}\bar{\mathbf{z}}_i$, where $\mathbf{G} \in \mathbb{R}^{n \times n}$ is a nonsingular generating matrix and $\bar{\mathbf{z}}_i$ is an integer vector in \mathbb{Z}^n , [35]:

Theorem 3.1 *Suppose that the level set $L(\mathbf{x}^0) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leq f(\mathbf{x}^0)\}$ is compact and that f is continuously differentiable in an open set containing $L(\mathbf{x}^0)$. Then*

$$\liminf_{k \rightarrow +\infty} \|\nabla f(\mathbf{x}^k)\| = 0,$$

and there exists at least one limit point \mathbf{x}^* such that $\nabla f(\mathbf{x}^*) = 0$.

Furthermore, if $\lim_{k \rightarrow +\infty} \alpha_k = 0$, $\|\mathbf{x}^{k+1} - \mathbf{x}^k\| \leq C\alpha_k$ for some constant $C > 0$ independent of the iteration counter k , and $\mathbf{x}^{k+1} = \operatorname{argmin}_{\mathbf{x} \in \mathcal{N}(\mathbf{x}^k)} f(\mathbf{x})$ in the poll step, then

$$\lim_{k \rightarrow +\infty} \|\nabla f(\mathbf{x}^k)\| = 0,$$

and every limit point \mathbf{x}^* satisfies $\nabla f(\mathbf{x}^*) = 0$.

The results of Theorem 3.1 concern the ability of pattern search methods to converge globally, i.e., from arbitrary points, to local minimizers candidates. We recall that, despite the nonexistence of any supporting theory, due to their blindness caused by the nonuse of derivatives, and also by the flexibility of the search step to incorporate global search procedures while the poll step continues to assure convergence to local minima, numerical evidence about the capability of pattern search methods to compute global minimizers has been reported – see, e.g., [35, 39].

To address the BAO problem, efficiency on the number of function value computations is of the utmost importance. Therefore, the number of trial points in the search step should be minimalist, and guided by some physical or biological meaning. On the other hand, when the search step fails to obtain a decrease on the function value, polling should also be oriented in order to further reduce the number of function value evaluations (at least for

successful iterations). Recently, the efficiency of pattern search methods improved significantly by reordering the poll directions according to descent indicators built from simplex gradients [40]. One of the main advantages of this pattern search methods framework is the flexibility provided by the search step, where any strategy can be applied as long as only a finite number of points is tested. This allows the insertion of strategies/heuristics that enhance for a global search by influencing the quality of the local minimizer or stationary point found by the method. The use of minimum Frobenius norm quadratic models, to be minimized within a trust region, to compute a single trial point in the search step, enhanced a significant improvement of direct search for black-box non-smooth functions [39] similar to the BAO problem at hand. The size of the trust region is coupled to the radius of the sample set. Thus, for an effective global search, the sample points should span all the search space. That could be achieved by using larger initial step-size parameters. However, since the BAO problem has many local minima and the number of sample points is scarce, the polynomial interpolation or regression models (usually quadratic models) used within the trust region struggle to find the best local minima. Therefore, starting with larger mesh-size parameters has the advantage of a better coverage of the search space but may cause the algorithm to jump over lower local minima than the obtained one [41].

An alternative and popular approach to keep small mesh-size parameters and still have a good coverage of all the search space is to use a multi-start approach. However, the multi-start approach has the disadvantage of increasing the total number of function evaluations and with that the overall computational time. Moreover, despite the obtained good span of \mathbb{R}^2 in amplitude, that is only obtained by overlapping all the iterates which might be fallacious for this particular problem. Here, we will adopt a different strategy. We will consider a single starting point, a small initial mesh-size parameter, and try to obtain a good span in amplitude of \mathbb{R}^2 by incorporating an additional global strategy in the search step: radial basis functions interpolation. Other global strategies like particle swarm methods can be included in the search step to enhance a global search – see [42]. However, the number of function evaluations required for this type of strategies is prohibitive for obtaining an answer in a clinically acceptable time frame. We describe next the radial basis functions interpolation strategy we used in the search step for improving the coverage of the search space while keeping a low number of function value evaluations.

4 Radial basis function interpolation

For numerical approximation of multivariate functions, radial basis functions (RBFs) can provide excellent interpolants. For any finite data set in any Euclidean space, one can construct an interpolation of the data by using RBFs, even if the data points are unevenly and sporadically distributed in a high dimensional Euclidean space. There is a wide range of applications where RBF interpolation methods can be successfully applied, such as aeronautics, meteorology and medical imaging (see [43, 44, 45], [46], and [47]). However, RBF interpolant trends between and beyond the data points depend on the RBF used and may exhibit undesirable trends using some RBFs while the trends may be desirable using other RBFs. Numerical choice of the most adequate RBF for the problem at hand should be done instead of an usual a priori choice [48].

Next, we will formulate RBF interpolation problems, discuss the solvability of RBF interpolation problems, and describe the strategy used to take advantage of RBF interpolants in the search step of the pattern search method framework applied to the BAO problem.

4.1 RBF interpolation problems

Let $f(\mathbf{x})$ be the true response to a given input vector \mathbf{x} (of n components) such that the value of f is only known at a set of N input vectors $\mathbf{x} = \mathbf{x}^1, \dots, \mathbf{x}^N$, i.e., only $f(\mathbf{x}^k)$ ($k = 1, \dots, N$) are known. An interpolation model $g(\mathbf{x}) = \sum_{j=1}^N \alpha_j \varphi_j(\mathbf{x})$ is used as an approximation of $f(\mathbf{x})$, where α_j are the coefficients to be determined by interpolation conditions $g(\mathbf{x}^k) = f(\mathbf{x}^k)$ or $\sum_{j=1}^N \alpha_j \varphi_j(\mathbf{x}^k) = f(\mathbf{x}^k)$ ($k = 1, \dots, N$), and $\varphi_1, \dots, \varphi_N$ are the basis functions depending on the choice of interpolation methods.

Interpolation functions generated from a RBF $\varphi(t)$ can be represented in the following form:

$$g(\mathbf{x}) = \sum_{j=1}^N \alpha_j \varphi(\|\mathbf{x} - \mathbf{x}^j\|), \quad (2)$$

where $\|\mathbf{x} - \mathbf{x}^j\|$ denotes the parameterized distance between \mathbf{x} and \mathbf{x}^j defined as

$$\|\mathbf{x} - \mathbf{x}^j\| = \sqrt{\sum_{i=1}^n |\theta_i| (x_i - x_i^j)^2}, \quad (3)$$

and $\theta_1, \dots, \theta_n$ are scalars (see [43]).

A standard data normalization approach is to scale each component x_i by an estimation of its standard deviation σ_i calculated from the data:

$$\sigma_i = \sqrt{\frac{\sum_{j=1}^N (x_i^j - \text{ave}(x_i))^2}{N-1}}, \quad \text{with} \quad \text{ave}(x_i) = \frac{1}{N} \sum_{j=1}^N x_i^j.$$

Scaling each data attribute by its estimated standard deviation also helps the initial formulation of the approximation problem. The scalars $\theta_1, \dots, \theta_n$ in Eq. (3) are the model tuning parameters that will be determined by a cross-validation method for the best prediction model of the given data. Mathematically, one could rewrite $\|\mathbf{x} - \mathbf{x}^j\|$ as

$$\|\mathbf{x} - \mathbf{x}^j\| = \sqrt{\sum_{i=1}^n |\bar{\theta}_i| (x_i - x_i^j)^2}, \quad \text{with} \quad \bar{\theta}_i = \frac{\theta_i}{\sigma_i^2}. \quad (4)$$

In practice, starting without any scaling (i.e., $\bar{\theta}_i = 1$ in Eq. (4)) may lead to ill-conditioning of the interpolation problem. The coefficient matrix of the linear equations $\sum_{j=1}^N \alpha_j \varphi_j(\mathbf{x}^k) = f(\mathbf{x}^k)$ ($k = 1, \dots, N$) is called the interpolation matrix. The condition number of the unscaled interpolation matrix is usually very large compared to the scaled interpolation matrix condition number. The purpose of using two sets of scaling parameters in Eq. (2) is to allow a non-dimensional initial choice of $\bar{\theta}_i = 1$. Note that, for simplicity, most of the time, $\bar{\theta}_i = 1$ is considered, which can lead to bias results.

The most popular examples of RBF [49, 50, 51] are cubic spline $\varphi(t) = t^3$, thin plate spline $\varphi(t) = t^2 \ln t$, multiquadric $\varphi(t) = \sqrt{1+t^2}$, and Gaussian $\varphi(t) = \exp(-t^2)$ (see Fig. 3). These RBFs can be used to model cubic, almost quadratic, and linear growth rates, as well as exponential decay, of the response for trend predictions.

For fixed parameters $\bar{\theta}_i$, the coefficients $\alpha_1, \dots, \alpha_N$ in Eq. (2) can be calculated by solving the following linear system of interpolation equations:

$$\sum_{j=1}^N \alpha_j \varphi(\|\mathbf{x}^k - \mathbf{x}^j\|) = f(\mathbf{x}^k), \quad \text{for } k = 1, \dots, N. \quad (5)$$

One can rewrite Eq. (5) in matrix form as

$$\Phi \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{pmatrix} = \begin{pmatrix} f(\mathbf{x}^1) \\ f(\mathbf{x}^2) \\ \vdots \\ f(\mathbf{x}^N) \end{pmatrix}, \quad (6)$$

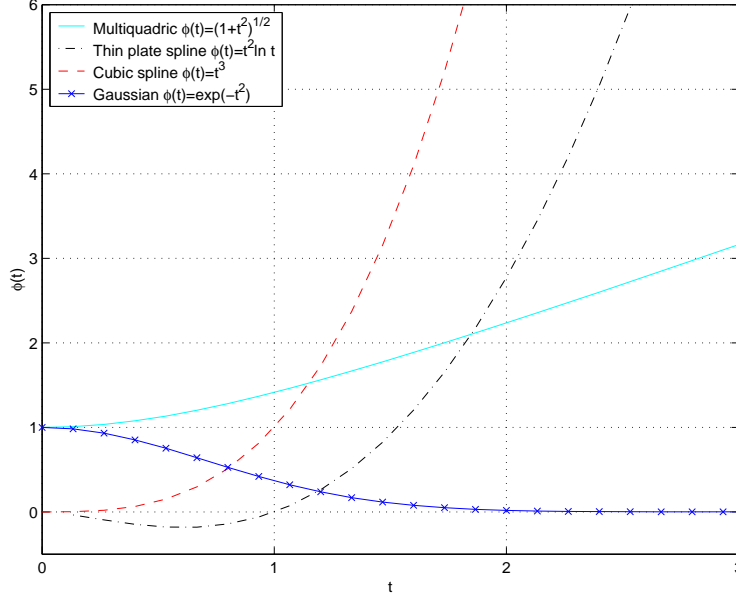


Figure 3: Graphs of commonly used radial basis functions.

where Φ is the interpolation matrix defined as

$$\Phi = \begin{pmatrix} \varphi(\|\mathbf{x}^1 - \mathbf{x}^1\|) & \varphi(\|\mathbf{x}^1 - \mathbf{x}^2\|) & \dots & \varphi(\|\mathbf{x}^1 - \mathbf{x}^N\|) \\ \varphi(\|\mathbf{x}^2 - \mathbf{x}^1\|) & \varphi(\|\mathbf{x}^2 - \mathbf{x}^2\|) & \dots & \varphi(\|\mathbf{x}^2 - \mathbf{x}^N\|) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi(\|\mathbf{x}^N - \mathbf{x}^1\|) & \varphi(\|\mathbf{x}^N - \mathbf{x}^2\|) & \dots & \varphi(\|\mathbf{x}^N - \mathbf{x}^N\|) \end{pmatrix}.$$

A unique interpolant is guaranteed for multiquadric and Gaussian RBFs, (i.e., Φ is a nonsingular matrix) even if the input vectors \mathbf{x}^j are few and poorly distributed, provided only that the input vectors are all different when $N > 1$. However, for cubic and thin plate spline RBFs, Φ might be singular. See [49] for an example of singular Φ when $\varphi(t) = t^3$. If $\varphi(t) = t^2 \ln t$, we can easily find an example where the interpolation matrix Φ is singular for a nontrivial set of distinct points $\mathbf{x}^1, \dots, \mathbf{x}^N$. For example, let $\mathbf{x}^2, \dots, \mathbf{x}^N$ be any distinct points on the sphere centered at \mathbf{x}^1 with radius 1. For this set of points, the first row and column of Φ consist of zeros, which implies the singularity of Φ .

An easy way to avoid this problem on the cubic and thin plate spline RBF interpolants is to add low-degree polynomials to interpolation functions in Eq. (2) and formulate an interpolation problem with constraints. That is, let $p(\mathbf{x}) = \sum_{j=1}^M \beta_j p_j(\mathbf{x})$, where p_1, \dots, p_M form a basis of algebraic polynomials in \mathbb{R}^n with degree at most m . Then interpolation

functions are of the following form:

$$g(\mathbf{x}) = p(\mathbf{x}) + \sum_{j=1}^N \alpha_j \varphi(\|\mathbf{x} - \mathbf{x}^j\|). \quad (7)$$

The M extra degrees of freedom in $g(\mathbf{x})$ can be eliminated by forcing the following M constraints:

$$\sum_{j=1}^N \alpha_j p_k(\mathbf{x}^j) = 0 \quad \text{for } k = 1, \dots, M, \quad (8)$$

which has the following matrix form:

$$\mathbf{P} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{pmatrix} = 0,$$

where

$$\mathbf{P} = \begin{pmatrix} p_1(\mathbf{x}^1) & \dots & p_1(\mathbf{x}^N) \\ \vdots & \ddots & \vdots \\ p_M(\mathbf{x}^1) & \dots & p_M(\mathbf{x}^N) \end{pmatrix}.$$

The interpolation equations using $g(\mathbf{x})$ in Eq. (7) become

$$\sum_{j=1}^N \alpha_j \varphi(\|\mathbf{x}^k - \mathbf{x}^j\|) + \sum_{j=1}^M \beta_j p_j(\mathbf{x}^k) = f(\mathbf{x}^k) \quad \text{for } k = 1, \dots, N. \quad (9)$$

Combining Eqs. (8) and (9), we obtain the following equation for the constrained RBF interpolation in matrix form:

$$\begin{pmatrix} \Phi & \mathbf{P}^T \\ \mathbf{P} & 0 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_N \\ \beta_1 \\ \vdots \\ \beta_M \end{pmatrix} = \begin{pmatrix} f(\mathbf{x}^1) \\ \vdots \\ f(\mathbf{x}^N) \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (10)$$

The key results on solvability of RBF interpolations related to the four RBFs shown in Fig. 3 are the following:

1. As seen before, Eq. (6) is always solvable if $\varphi(t) = \sqrt{1+t^2}$ or $\varphi(t) = \exp(-t^2)$;
2. For quadratic polynomials ($m = 2$), Eq. (10) is solvable if $\varphi(t) = t^3$ or $\varphi(t) = t^2 \ln t$, provided that the input vectors $\mathbf{x}^1, \dots, \mathbf{x}^N$ do not fall into the zero set of the polynomial.

The proofs of second statements are based on mathematical concepts called (conditionally) positive definiteness and their Schoenberg-Micchelli characterizations [52, 53] as described by Schaback and Wendland [54].

The constructed interpolant $g(\mathbf{x})$ in Eq. (2) depends on “subjective” choice of $\varphi(t)$, and model parameters $\theta_1, \dots, \theta_n$. While one can try all the possible choices of $\varphi(t)$ in search of a desirable interpolant, there are infinitely many choices for $\theta_1, \dots, \theta_n$. One could, however, use cross-validation to choose an optimal value of $\theta_1, \dots, \theta_n$ that yield an interpolant $g(\mathbf{x})$ with the most accurate trend prediction.

4.2 Model parameter tuning by cross-validation

RBF interpolation models use the parameterized distance:

$$\|\mathbf{x} - \mathbf{x}^j\| = \sqrt{\sum_{i=1}^n |\theta_i| \left(\frac{x_i - x_i^j}{\sigma_i} \right)^2}, \quad (11)$$

where $\theta_1, \dots, \theta_n$ are scalars. Mathematically, one could pick any fixed set of $\theta_1, \dots, \theta_n$ and construct the interpolation function for the given data. However, two different sets of $\theta_1, \dots, \theta_n$ will lead to two interpolation models that behave very differently between the input vectors $\mathbf{x}^1, \dots, \mathbf{x}^N$. Model parameter tuning for RBF interpolation aims at finding a set of parameters $\theta_1, \dots, \theta_n$ that results in the best prediction model of the unknown response based on the available data. Cross-validation (CV) [55, 56, 57, 58, 59, 60] and maximum likelihood estimation [61, 62, 63] are two statistical methods for tuning the model parameters $\theta_1, \dots, \theta_n$ for best prediction models. CV can be used for general model parameter tuning, while maximum likelihood estimation can only be applied for density function parameter estimation.

Other metrics that are not based on fitting errors must be used to determine which basis function $\varphi(t)$ and what scaling parameters θ_i are most appropriate to model the response function $f(\mathbf{x})$, because RBF interpolation method yields a fitting function $g(\mathbf{x})$ whose value

at \mathbf{x}^k is exactly $f(\mathbf{x}^k)$ for $k = 1, \dots, N$. One can always try to obtain values of $f(\mathbf{x})$ at some additional data points $\mathbf{x}^{N+1}, \dots, \mathbf{x}^{\bar{N}}$ and use the prediction errors $|g(\mathbf{x}^k) - f(\mathbf{x}^k)|$ for $k = N + 1, \dots, \bar{N}$ to assess the prediction accuracy of $g(\mathbf{x})$, but this technique is often impractical and always expensive. The prediction accuracy can be used as a criterion for choosing the best basis function $\varphi(t)$ and parameters θ_i .

Without additional sample points, CV [57, 59] was proposed to find $\varphi(t)$ and θ_i that lead to an approximate response model $g(\mathbf{x})$ with optimal prediction capability and proved to be effective [58, 60]. The leave-one-out CV procedure is usually used in model parameter tuning for RBF interpolation (see [58, 64]).

Algorithm 4.1 (Leave-one-out cross-validation for RBF Interpolation)

1. Fix a set of parameters $\theta_1, \dots, \theta_n$.
2. For $j = 1, \dots, N$, construct the RBF interpolant $g_{-j}(\mathbf{x})$ of the data points $(\mathbf{x}^k, f(\mathbf{x}^k))$ for $1 \leq k \leq N, k \neq j$.
3. Use the following CV root mean square error as the prediction error:

$$E^{CV}(\theta_1, \dots, \theta_n) = \sqrt{\frac{1}{N} \sum_{j=1}^N (g_{-j}(\mathbf{x}^j) - f(\mathbf{x}^j))^2}. \quad (12)$$

Remark. One could also use other forms of CV errors such as the CV average absolute error: $\frac{1}{N} \sum_{j=1}^N |g_{-j}(\mathbf{x}^j) - f(\mathbf{x}^j)|$.

The goal of model parameter tuning by CV is to find $\theta_1, \dots, \theta_n$ that minimize the CV error $E^{CV}(\theta_1, \dots, \theta_n)$ so that the interpolation model has the highest prediction accuracy when CV error is the measure. It is worth pointing out that, typically, it is difficult to minimize $E^{CV}(\theta_1, \dots, \theta_n)$ numerically because $E^{CV}(\theta_1, \dots, \theta_n)$ is a highly nonlinear and non-convex function. One could make the model parameter tuning much easier by assuming $\theta_1 = \dots = \theta_n$, which reduces the problem to unconstrained minimization of a univariate function (see [58]). This approach has the obvious benefit of dealing with a unidimensional optimization problem but the disadvantage of not using all different θ_i . Different θ_i allow the model parameter tuning to scale each variable x_i based on its significance in modeling the variance in the response, thus, have the benefit of implicit variable screening built in the model parameter tuning.

4.3 RBF interpolation within the pattern search methods framework tailored for the BAO problem

The common approach of incorporating interpolation models within pattern search methods framework consists of forming an interpolation model and finding its minimum in the search step. For example, in [39], the search step computes a single trial point using minimum Frobenius norm quadratic models to be minimized within a trust region. In this research report we are only interested in discussing the advantages of using a RBF model in the search step of the pattern search method applied to the BAO problem. Therefore, the strategy sketched here is tailored for the BAO problem, it might only be generalized for similar problems, and does not include the formal minimization of the RBF model. When RBF models are built, n RBF trial points are computed using the following strategy.

Algorithm 4.2 (RBF trial points)

Initialization: Build a RBF model.

For each beam direction ($i = 1, \dots, n$)

1. Evaluate the RBF model every degree between the previous beam direction and the next one.
2. Find the minimum of those values that correspond to a beam direction that was not evaluated yet and, is at least 4 degrees away from a previously evaluated one, for the beam direction at stake.
3. Take as RBF trial point the current iterate updating the beam direction corresponding to the minimum found in step 2.

Our main goal for using a RBF model in the search step of the pattern search methods framework is to properly explore the search space in amplitude without a random criteria. Therefore, each beam direction is tested every degree between the previous beam direction and the next one as stated in step 1. A proper minimization is unnecessary since we are interested in integer beam angle directions. Step 2 forces the algorithm to consider only directions in regions not yet explored which is the main goal of the RBF models here (directions that are less than 4 degrees apart are considered to be clinically equivalent).

We choose to implement the use of RBF interpolation in the search step taking advantage of the availability of an existing pattern search methods framework implementation used successfully by us to tackle the BAO problem [41, 65]. We used the last version of SID-PSM [39, 40] which is a MATLAB [66] implementation of the pattern search methods that incorporates improvements for the search step, with the use of minimum Frobenius norm quadratic models to be minimized within a trust region, and improvements for the poll step, where efficiency on the number of function value computations improved significantly by reordering the poll directions according to descent indicators. Similarly to the minimum Frobenius norm quadratic models, the RBF models are not built in the search step until the number of points previously evaluated is greater than $n + 1$. Therefore, the search step is skipped while the number of points is not greater than $n + 1$. The SID-PSM algorithm incorporating RBFs applied to the BAO problem is the following.

Algorithm 4.3 (Pattern search methods framework with RBFs in search step)

- 0. Initialization** Set $k = 0$. Choose $\mathbf{x}^0 \in \mathbb{R}^n$, $\alpha_0 > 0$, and a positive spanning set \mathbf{V} .
- 1. Search step** If the number of evaluated points is not superior to $n + 1$ skip the search step. Otherwise, compute the RBFs trial points followed by the minimum Frobenius norm quadratic trial point until one of them decreases the objective function value. If none achieve that goal, go to step 2 and the search step is declared unsuccessful. Otherwise, go to step 4 and both the search step and iteration are declared successful.
- 2. Poll step** This step is only performed if the search step is unsuccessful. If $f(\mathbf{x}^k) \leq f(\mathbf{x})$ for every \mathbf{x} in the mesh neighborhood $\mathcal{N}(\mathbf{x}^k)$, go to step 3 and shrink M_k . Both poll step and iteration are declared unsuccessful. Otherwise, choose a point $\mathbf{x}^{k+1} \in \mathcal{N}(\mathbf{x}^k)$ such that $f(\mathbf{x}^{k+1}) < f(\mathbf{x}^k)$ and go to step 4. Both poll step and iteration are declared successful.
- 3. Mesh reduction** Let $\alpha_{k+1} = \frac{1}{2}\alpha_k$. Set $k = k + 1$ and return to step 1 for a new iteration.
- 4. Mesh expansion** Let $\alpha_{k+1} = \alpha_k$. Set $k = k + 1$ and return to step 1 for a new iteration.

The inclusion of RBF trial points in the search step has three distinct purposes. Improve the coverage of the search space in amplitude, influence the quality of the local minimum obtained by performing searches away from the neighborhood of the current iterate and last but not least to improve the efficacy of the minimum Frobenius norm quadratic models. This last goal is a consequence of the first two, trying to obtain data points that are less poorly distributed in a high dimensional space and span all the search space in amplitude. The minimum Frobenius norm quadratic models are minimized within a trust region whose size is coupled to the radius of the sample set. Thus, for an effective global search, the sample points should span all the search space.

The benefits of using RBFs in the search step of the pattern search methods framework in the optimization of the BAO problem are illustrated using a set of clinical examples of head-and-neck cases that are presented next.

5 Head-and-neck clinical examples

Four clinical examples of retrospective treated cases of head-and-neck tumors at the Portuguese Institute of Oncology of Coimbra are used to test the pattern search methods framework proposed. The selected clinical examples were signalized at IPOC as complex cases where proper target coverage and organ sparing, in particular parotid sparing, proved to be difficult to obtain with the typical 7-beam equispaced coplanar treatment plans. The patients' CT sets and delineated structures were exported via Dicom RT to a freeware computational environment for radiotherapy research (see Figure 4). Since the head-and-neck region is a complex area where, e.g., the parotid glands are usually in close proximity to or even overlapping with the target volume, careful selection of the radiation incidence directions can be determinant to obtain a satisfying treatment plan.

The spinal cord and the brainstem are some of the most critical organs at risk (OARs) in the head-and-neck tumor cases. These are serial organs, i.e., organs such that if only one subunit is damaged, the whole organ functionality is compromised. Therefore, if the tolerance dose is exceeded, it may result in functional damage to the whole organ. Thus, it is extremely important not to exceed the tolerance dose prescribed for these type of organs. Other than the spinal cord and the brainstem, the parotid glands are also important OARs.

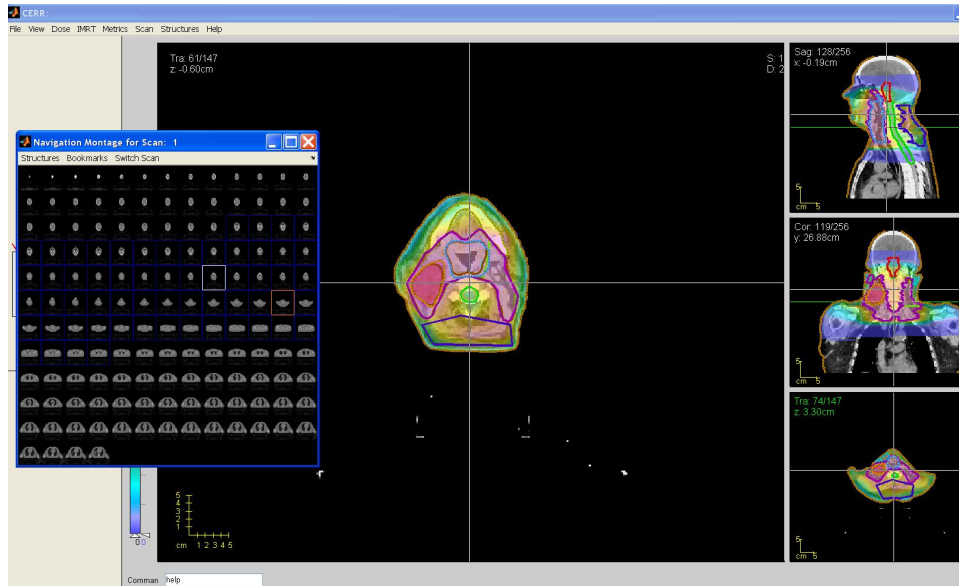


Figure 4: Structures considered in the IMRT optimization visualized in CERR.

The parotid gland is the largest of the three salivary glands. A common complication due to parotid glands irradiation is xerostomia (the medical term for dry mouth due to lack of saliva). This decreases the quality of life of patients undergoing radiation therapy of head-and-neck, causing difficulties to swallow. The parotids are parallel organs, i.e., if a small volume of the organ is damaged, the rest of the organ functionality may not be affected. Their tolerance dose depends strongly on the fraction of the volume irradiated. Hence, if only a small fraction of the organ is irradiated the tolerance dose is much higher than if a larger fraction is irradiated. Thus, for these parallel structures, the organ mean dose is generally used instead of the maximum dose as an objective for inverse planning optimization.

In general, the head-and-neck region is a complex area to treat with radiotherapy due to the large number of sensitive organs in this region (e.g., eyes, mandible, larynx, oral cavity, etc.). For simplicity, in this study, the OARs used for treatment optimization were limited to the spinal cord, the brainstem and the parotid glands.

The tumor to be treated plus some safety margins is called planning target volume (PTV). For the head-and-neck cases in study it was separated in two parts with different prescribed doses: PTV1 and PTV2. The prescription dose for the target volumes and tolerance doses for the OARs considered in the optimization are presented in Table 1.

Structure	Mean dose	Max dose	Prescribed dose
Spinal cord	–	45 Gy	–
Brainstem	–	54 Gy	–
Left parotid	26 Gy	–	–
Right parotid	26 Gy	–	–
PTV1	–	–	70.0 Gy
PTV2	–	–	59.4 Gy
Body	–	80 Gy	–

Table 1: Prescribed doses for all the structures considered for IMRT optimization.

The parotid glands are in close proximity to or even overlapping with the PTV which helps explaining the difficulty of parotid sparing. Adequate beam directions can help on the overall optimization process and in particular in parotid sparing.

6 Results

The RBFs within the pattern search methods framework approach were tested using a set of four clinical examples of retrospective treated cases of head-and-neck tumors at the Portuguese Institute of Oncology of Coimbra (IPOC). A typical head-and-neck treatment plan consists of radiation delivered from 5 to 9 equally spaced coplanar orientations around the patient. Treatment plans with 7 equispaced coplanar beams were used at IPOC and are commonly used in practice to treat head-and-neck cases [2]. Therefore, treatment plans of 7 coplanar orientations were obtained using our BAO algorithms, denoted *SID-PSM*, *PSM-RBF1*, *PSM-RBF2*, *PSM-RBF3*, *PSM-RBF4*, wether the algorithm used was the pattern search framework alone or coupled with multiquadric, thin plate spline, cubic spline or gaussian RBFs, respectively. These treatment plans were compared with the typical 7-beam equispaced coplanar treatment plans denoted *equi*.

In order to facilitate convenient access, visualization and analysis of patient treatment planning data, as well as dosimetric data input for treatment plan optimization research, the computational tools developed within MATLAB and CERR – computational environment

for radiotherapy research [67] are used widely for IMRT treatment planning research. The ORART – operations research applications in radiation therapy [68] collaborative working group developed a series of software routines that allow access to influence matrices, which provide the necessary dosimetry data to perform optimization in IMRT. CERR was elected as the main software platform to embody our optimization research.

Our tests were performed on a 2.66Ghz Intel Core Duo PC with 3 GB RAM. We used CERR 3.2.2 version and MATLAB 7.4.0 (R2007a). The dose was computed using CERR’s pencil beam algorithm (QIB). An automatized procedure for dose computation for each given beam angle set was developed, instead of the traditional dose computation available from IMRTP module accessible from CERR’s menubar. This automatization of the dose computation was essential for integration in our BAO algorithm. To address the convex nonlinear formulation of the FMO problem we used a trust-region-reflective algorithm (*fmincon*) of MATLAB 7.4.0 (R2007a) Optimization Toolbox.

The last version of SID-PSM was used as our pattern search methods framework. The spanning set used was the positive spanning set ($[e \ -e \ I \ -I]$, with I being the identity matrix of dimension k and $e = [1 \ \dots \ 1]^T$). Each of these directions correspond to, respectively, the rotation of all incidence directions clockwise, to the rotation of all incidence directions counter-clockwise, the rotation of each individual incidence direction clockwise, and the rotation of each individual incidence direction counter-clockwise. The initial size of the mesh parameter was set to $\alpha_0 = 4$. Since the initial points were integer vectors, all iterates will have integer values as long as the mesh parameter does not become less than 1. Therefore, the stopping criteria adopted was the mesh parameter becoming less than 1. Since we want to improve the quality of the typical equispaced treatment plans, the 7 equispaced coplanar beam angles were considered as the initial point for the 7-beam angle optimization process. The choice of this initial point and the non-increasing property of the sequence of iterates generated by SID-PSM imply that each successful iteration correspond to an effective improvement with respect to the usual equispaced beam configuration.

One of the main advantages of this pattern search framework is the flexibility provided by the search step, where any strategy can be applied as long as only a finite number of points is tested. This allows the insertion of our RBFs approach to enhance a global search by influencing the quality of the local minimizer mainly by covering the search space in a

more appropriate fashion. The main goal of the present work is to verify the contribution of RBFs in the improvement of our BAO approach using pattern search methods. Simultaneously we aim to verify which RBF performs better for the problem at hand. The CV error of an interpolation model can be a useful and objective tool to decide which model is better. We used the MATLAB code *fminsearch*, an implementation of the Nelder-Mead [69] multidimensional search algorithm, to minimize the CV error $E^{CV}(\theta_1, \dots, \theta_n)$ in Eq. (12) and to find the best model parameters $\theta_1, \dots, \theta_n$. The results of BAO optimization for the four retrospective treated cases of head-and-neck tumors using SID-PSM and the four different RBFs approaches used within SID-PSM are presented in Table 2.

The most important outcome of the results shown in Table 2 is that regardless the RBF considered, the use of RBFs within the pattern search methods framework approach corresponds always to an improvement of the solution in terms of objective function value. The best results in terms of objective function value were obtained by the multiquadric and thin plate splines for two tumor cases each. The CV prediction errors are also used as a tool to help on the decision of the best approximation. We can verify that, in fact, the best result for each case corresponds to the smallest CV error. The results are presented in terms of number of function evaluations instead of overall computational time since for different dose engines, beamlet optimization methods or even other objective function strategies, the overall computational time may have a totally different magnitude. Dose computation using QIB consumed most of the overall computational time. In average it took 2 hours to run the BAO optimization using the SID-PSM. Our objective is to emphasize the few number of function evaluations required by pattern search methods even when using RBFs within the search step. In Figure 5, the objective function value decrease versus number of function evaluations required is presented to compare the performances of *SID-PSM* and the two best RBFs approaches, *PSM-RBF1* and *PSM-RBF2*. By inspection of Figure 5 and analysis of Table 2 we concluded that, in average, *PSM-RBF2* leads to better predictions in terms of CV errors and better improvements with respect to *SID-PSM* in terms of optimal objective function value.

The history of the 7-beam angle optimization process using *SID-PSM* and *PSM-RBF2*, in terms of beam directions tested, for each case, is presented in Figure 6. By simple inspection we can verify that the sequence of iterates are better distributed by amplitude

Case	algorithm	Finit	Fopt	% decrease	Fevals	mean E^{CV}
1	<i>SID-PSM</i>	175.9	168.4	4.3	74	–
	<i>PSM-RBF1</i>	175.9	164.5	6.5	216	1.1158e+005
	<i>PSM-RBF2</i>	175.9	162.7	7.5	198	1.0274e+004
	<i>PSM-RBF3</i>	175.9	167.4	4.8	106	5.7068e+005
	<i>PSM-RBF4</i>	175.9	166.0	5.6	197	1.6131e+005
2	<i>SID-PSM</i>	216.3	197.9	8.5	79	–
	<i>PSM-RBF1</i>	216.3	193.8	10.4	176	1.3458e+004
	<i>PSM-RBF2</i>	216.3	195.4	9.7	136	1.7541e+004
	<i>PSM-RBF3</i>	216.3	195.6	9.6	129	1.6213e+004
	<i>PSM-RBF4</i>	216.3	198.9	8.0	123	4.3663e+005
3	<i>SID-PSM</i>	487.8	459.8	5.7	107	–
	<i>PSM-RBF1</i>	487.8	450.6	7.6	203	3.4954e+005
	<i>PSM-RBF2</i>	487.8	451.3	7.5	186	7.1020e+005
	<i>PSM-RBF3</i>	487.8	451.2	7.5	201	7.2608e+005
	<i>PSM-RBF4</i>	487.8	452.8	7.2	166	6.3432e+011
4	<i>SID-PSM</i>	257.6	248.5	3.5	69	–
	<i>PSM-RBF1</i>	257.6	244.8	5.0	102	1.2585e+005
	<i>PSM-RBF2</i>	257.6	243.7	5.4	172	1.7162e+004
	<i>PSM-RBF3</i>	257.6	244.0	5.3	161	1.7194e+005
	<i>PSM-RBF4</i>	257.6	244.2	5.2	116	1.2882e+006

Table 2: Results of the beam angle optimization process using *SID-PSM*, *PSM-RBF1*, *PSM-RBF2*, *PSM-RBF3* and *PSM-RBF4*.

in \mathbb{R}^2 when using *PSM-RBF2*, with an appropriate coverage in amplitude of all the search space. Note that in a previous work [41] we conclude that increasing the step parameter of the pattern search method framework lead to better search space coverage in amplitude but with no correspondent benefits in terms of optimum function value reduction. Moreover, in previous works [41, 65] comparisons between plans obtained after BAO optimization using *SID-PSM* and typical equidistant treatment plans were done and was demonstrated

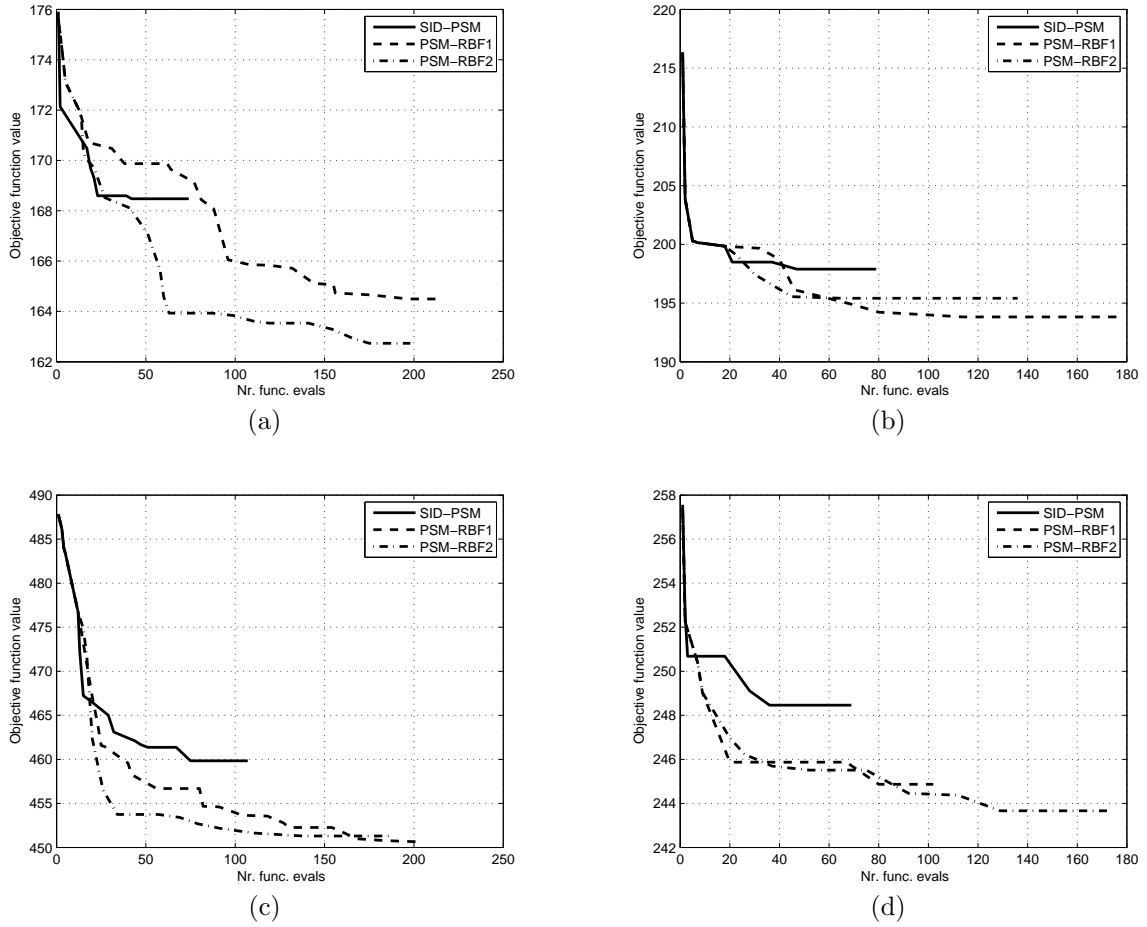


Figure 5: History of the 7-beam angle optimization process using *SID-PSM*, *PSM-RBF1* and *PSM-RBF2* for cases 1 to 4, 5(a) to 5(d) respectively.

that the quality of the plan can be improved with our BAO approach using pattern search methods. Therefore, the main goal of the present comparisons is to verify the contribution of RBFs in the improvement of our BAO approach using pattern search methods.

Despite the improvement in FMO value, the quality of the results can be perceived considering a variety of metrics. Typically, results are judged by their cumulative dose-volume histogram (DVH). The DVH displays the fraction of a structure's volume that receives at least a given dose. Another metric usually used for plan evaluation is the volume of PTV that receives 95% of the prescribed dose. Typically, 95% of the PTV volume is required. The occurrence of coldspots, less than 93% of PTV volume receives the prescribed dose, and the existence of hotspots, the percentage of the PTV volume that receives more than 110% of the prescribed dose, are other measures usually used to evaluate

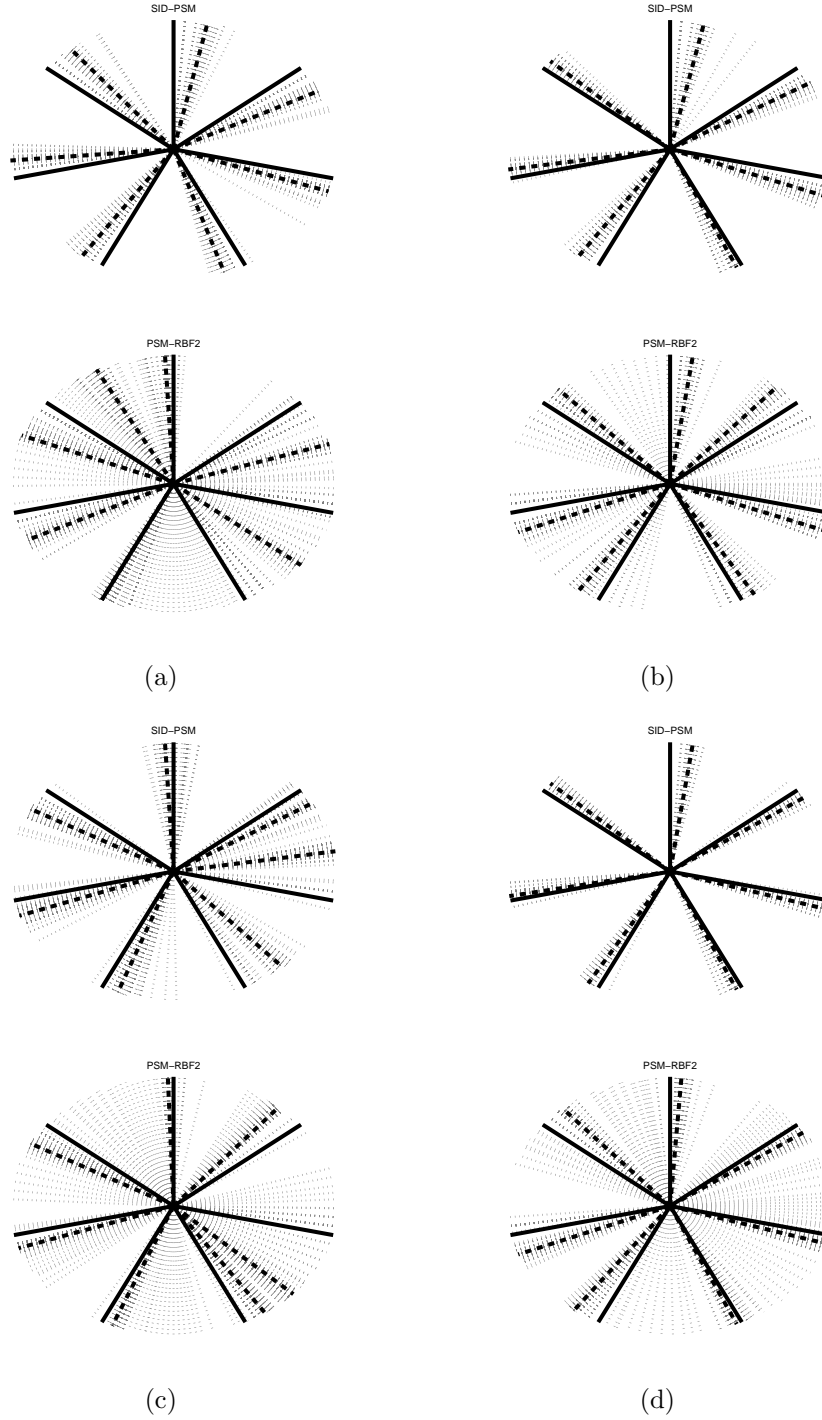


Figure 6: History of the 7-beam angle optimization process using *SID-PSM* and *PSM-RBF2* for cases 1 to 4, 6(a) to 6(d) respectively. Initial angle configuration, optimal angle configuration and intermediate angle configurations are displayed with solid, dashed and dotted lines, respectively.

Case	Target coverage	<i>PSM-RBF2</i>	<i>SID-PSM</i>	<i>equi</i>
1	PTV1 at 95 % volume	67.225 Gy	67.175 Gy	67.075 Gy
	PTV1 % > 93% of Rx (%)	99.650	99.625	99.425
	PTV1 % > 110% of Rx (%)	0.000	0.000	0.000
	PTV2 at 95 % volume	57.975 Gy	57.925 Gy	57.825 Gy
	PTV2 % > 93% of Rx (%)	97.844	97.619	97.410
	PTV2 % > 110% of Rx (%)	15.691	15.855	16.064
2	PTV1 at 95 % volume	66.675 Gy	66.475 Gy	66.675 Gy
	PTV1 % > 93% of Rx (%)	99.125	99.125	99.125
	PTV1 % > 110% of Rx (%)	0.000	0.000	0.000
	PTV2 at 95 % volume	57.375 Gy	57.425 Gy	56.975 Gy
	PTV2 % > 93% of Rx (%)	97.427	97.514	96.875
	PTV2 % > 110% of Rx (%)	4.387	4.321	4.321
3	PTV1 at 95 % volume	66.725 Gy	66.675 Gy	66.625 Gy
	PTV1 % > 93% of Rx (%)	98.687	98.687	98.415
	PTV1 % > 110% of Rx (%)	0.000	0.000	0.000
	PTV2 at 95 % volume	56.775 Gy	56.675 Gy	56.825 Gy
	PTV2 % > 93% of Rx (%)	96.685	96.595	96.682
	PTV2 % > 110% of Rx (%)	17.357	17.639	17.597
4	PTV1 at 95 % volume	67.425 Gy	67.475 Gy	67.425 Gy
	PTV1 % > 93% of Rx (%)	99.697	99.778	99.677
	PTV1 % > 110% of Rx (%)	0.000	0.000	0.000
	PTV2 at 95 % volume	57.575 Gy	57.725 Gy	57.475 Gy
	PTV2 % > 93% of Rx (%)	98.095	98.302	98.034
	PTV2 % > 110% of Rx (%)	11.417	11.421	11.785

Table 3: Target coverage obtained by treatment plans.

the target coverage. Mean and/or maximum doses of OARs are usually displayed to verify organ sparing.

The results regarding targets coverage are presented in Table 3. We can verify that both

Case	OAR	Mean Dose (Gy)			Max Dose (Gy)		
		<i>PSM-RBF2</i>	<i>SID-PSM</i>	<i>equi</i>	<i>PSM-RBF2</i>	<i>SID-PSM</i>	<i>equi</i>
1	Spinal cord	–	–	–	38.587	38.886	38.687
	Brainstem	–	–	–	50.502	49.079	53.790
	Left parotid	24.338	25.203	27.023	–	–	–
	Right parotid	24.311	24.743	26.153	–	–	–
2	Spinal cord	–	–	–	44.490	44.537	44.827
	Brainstem	–	–	–	52.764	52.537	52.336
	Left parotid	25.947	26.190	27.287	–	–	–
	Right parotid	25.388	26.099	27.659	–	–	–
3	Spinal cord	–	–	–	41.177	41.244	44.810
	Brainstem	–	–	–	48.248	48.988	48.840
	Left parotid	27.733	28.527	28.388	–	–	–
	Right parotid	29.452	30.035	30.536	–	–	–
4	Spinal cord	–	–	–	40.630	40.524	40.610
	Brainstem	–	–	–	47.290	48.079	47.458
	Left parotid	26.783	26.770	27.170	–	–	–
	Right parotid	26.205	26.934	27.030	–	–	–

Table 4: OARs sparing obtained by treatment plans.

SID-PSM and *PSM-RBF2* treatment plans consistently obtained slightly better target coverage numbers compared to *equi* treatment plans. Moreover, target coverage numbers are favorable to *PSM-RBF2* treatment plans compared to *SID-PSM* treatment plans. Organ sparing results are shown in Table 4. All the treatment plans fulfill the maximum dose requirements for the spinal cord and the brainstem. However, as expected, the main differences reside in parotid sparing. The *equi* treatment plans could never enhance parotid sparing while *SID-PSM* treatment plans fulfill the parotid’s mean dose requirements once and *PSM-RBF2* treatment plans for half of the cases. Furthermore, in average, *SID-PSM* treatment plans reduced the parotid’s mean dose irradiation in 0.8 Gy compared to the *equi* treatment plans and *PSM-RBF2* treatment plans reduced the parotid’s mean dose

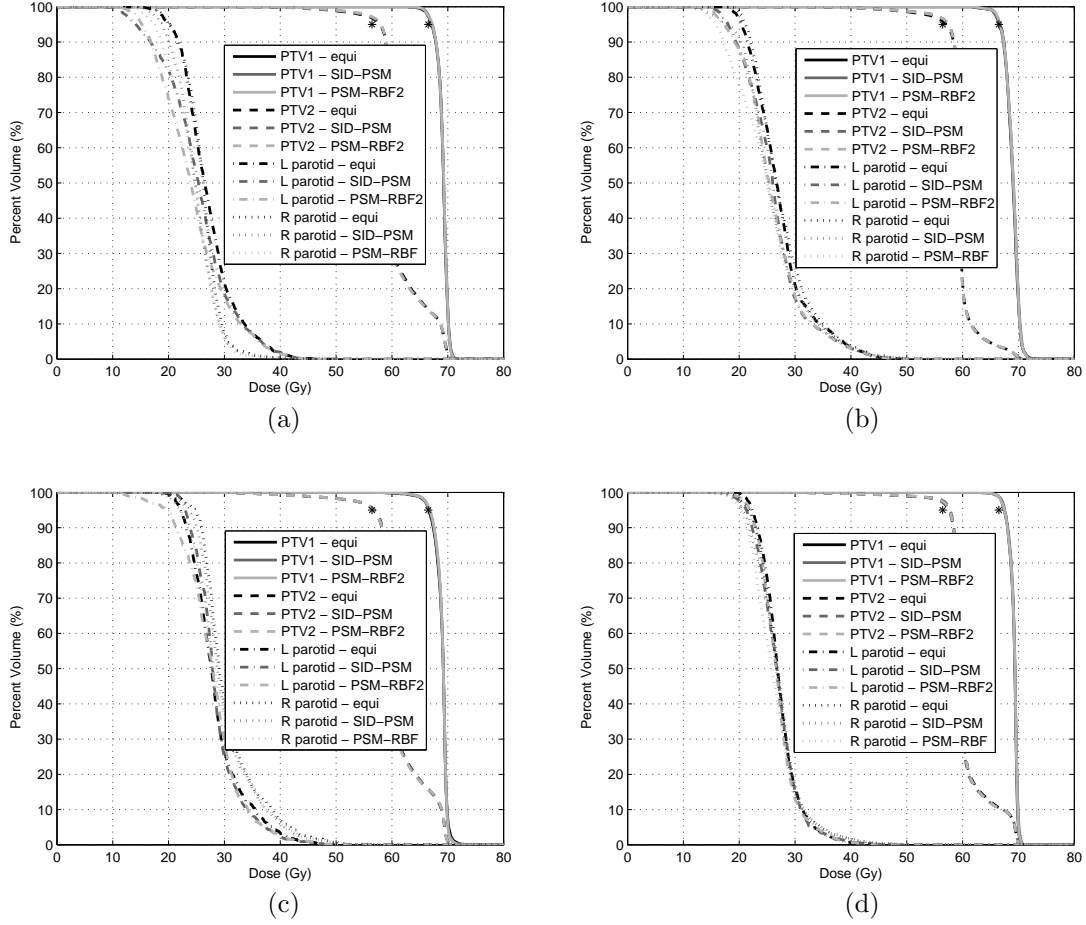


Figure 7: Cumulative dose volume histogram comparing the results obtained by *SID-PSM*, *PSM-RBF2* and *equi* for cases 1 to 4, 7(a) to 7(d) respectively.

irradiation in 1.4 Gy compared to the *equi* treatment plans. The differences between *SID-PSM* treatment plans and *PSM-RBF2* treatment plans, concerning parotid sparing, show a clear advantage for the *PSM-RBF2* treatment plans. DVH results illustrating the numbers presented in Tables 3 and 4 are displayed in Figure 7. Since parotids are the most difficult organs to spare, as shown in Table 4, for clarity, the DVHs only include the targets and the parotids. The asterisks indicate 95% of PTV volumes versus 95% of the prescribed doses. The results displayed in Figure 7 confirm the benefits of using the optimized beam directions, in particular using the directions obtained and used in *PSM-RBF2* treatment plan.

7 Conclusions

The BAO problem is a continuous global highly non-convex optimization problem known to be extremely challenging and yet to be solved satisfactorily. A new approach for the resolution of the BAO problem, using RBFs within a pattern search methods framework, was proposed and tested using a set of clinical head-and-neck cases. Pattern search methods are suited for the BAO problem since they require few function value evaluations and, similarly to other derivative-free optimization methods, have the ability to avoid local entrapment. The pattern search methods approach seems to be similar to neighborhood search approaches in which the neighborhood is constructed using the pattern search method. However, local neighborhood search approaches are only similar to the poll step of the pattern search methods framework. The existence of a search step with the flexibility to use any strategy, method or heuristic, or take advantage of a priori knowledge of the problem at hand, is an advantage that was explored successfully in this work. We have shown that a beam angle set can be locally improved in a continuous manner using pattern search methods. Moreover, it was shown that the incorporation of RBFs in the search step leads to an improvement of the local solution obtained. For numerical approximation of multivariate functions, RBFs can provide excellent interpolants, even if the data points available are unevenly and sporadically distributed. For the retrospective tumor cases tested, our RBFs tailored approach showed a positive influence on the quality of the local minimizer found and a clearly better coverage of the whole search space in amplitude. The improvement of the local solutions in terms of objective function value corresponded, for the head-and-neck cases tested, to high quality treatment plans with better target coverage and with improved organ sparing, in particular better parotid sparing. Moreover, we have to highlight the low number of function evaluations required to obtain locally optimal solutions, which is a major advantage compared to other global heuristics. This advantage should be even more relevant when considering non-coplanar directions since the number of possible directions to consider increase significantly. The efficiency on the number of function value computations is of the utmost importance, particularly when the BAO problem is modeled using the optimal values of the FMO problem.

Acknowledgements

The work of H. Rocha was supported by the European social fund and Portuguese funds from MCTES. This work has been partially supported by FCT under project grant PEst-C/EEI/UI0308/2011.

References

- [1] V. V. Mišić, D. M. Aleman, M. B. Sharpe, Neighborhood search approaches to non-coplanar beam orientation optimization for total marrow irradiation using IMRT, *Eur. J. Oper. Res.* 205 (2010) 522–527.
- [2] D. M. Aleman, H. E. Romeijn, J. F. Dempsey, A response surface approach to beam orientation optimization in intensity modulated radiation therapy treatment planning, *INFORMS J. Comput.: Computat. Biol. Med. Appl.* 21 (2009) 62–76.
- [3] H. M. Lu, H. M. Kooy, Z. H. Leber, R. J. Ledoux, Optimized beam planning for linear accelerator-based stereotactic radiosurgery, *Int. J. Radiat. Oncol. Biol. Phys.* 39 (1997) 1183–1189.
- [4] G. Meedt, M. Alber, F. Nüsslin, Non-coplanar beam direction optimization for intensity-modulated radiotherapy, *Phys. Med. Biol.* 48 (2003) 2999–3019.
- [5] S. K. Das, L. B. Marks, Selection of coplanar or non coplanar beams using three-dimensional optimization based on maximum beam separation and minimized nontarget irradiation, *Int. J. Radiat. Oncol. Biol. Phys.* 38 (1997) 643–655.
- [6] X. Wang, X. Zhang, L. Dong, H. Liu, M. Gillin, A. Ahamad, K. Ang, R. Mohan, Effectiveness of noncoplanar IMRT planning using a parallelized multiresolution beam angle optimization method for paranasal sinus carcinoma, *Int. J. Radiat. Oncol. Biol. Phys.* 63 (2005) 594–601.
- [7] H. H. Liu, M. Jauregui, X. Zhang, X. Wang, L. Dongand, R. Mohan, Beam angle optimization and reduction for intensity-modulated radiation therapy of non-small-cell lung cancers, *Int. J. Radiat. Oncol. Biol. Phys.* 65 (2006) 561–572.

- [8] C. G. Rowbottom, S. Webb, M. Oldham, Improvements in prostate radiotherapy from the customization of beam directions, *Med. Phys.* 25 (1998) 1171–1179.
- [9] RadiologyInfo. <http://www.radiologyinfo.org/en/info.cfm?pg=linac>.
- [10] D. Craft, Local beam angle optimization with linear programming and gradient search, *Phys. Med. Biol.* 52 (2007) 127–135.
- [11] M. Ehrgott, A. Holder, J. Reese, Beam selection in radiotherapy design, *Linear Algebra Appl.* 428 (2008) 1272–1312.
- [12] D. Djajaputra, Q. Wu, Y. Wu, R. Mohan, Algorithm and performance of a clinical IMRT beam-angle optimization system, *Phys. Med. Biol.* 48 (2003) 3191–3212.
- [13] T. Bortfeld, W. Schlegel, Optimization of beam orientations in radiation therapy: some theoretical considerations, *Phys. Med. Biol.* 38 (1993) 291–304.
- [14] G. A. Ezzell, Genetic and geometric optimization of three-dimensional radiation therapy treatment planning, *Med. Phys.* 23 (1996) 293–305.
- [15] O. C. Haas, K. J. Burnham, J. Mills, Optimization of beam orientation in radiotherapy using planar geometry, *Phys. Med. Biol.* 43 (1998) 2179–2193.
- [16] Y. Li, J. Yao, D. Yao, Automatic beam angle selection in IMRT planning using genetic algorithm, *Phys. Med. Biol.* 49 (2004) 1915–1932.
- [17] Y. Li, D. Yao, J. Yao, W. Chen, A particle swarm optimization algorithm for beam angle selection in intensity modulated radiotherapy planning, *Phys. Med. Biol.* 50 (2005) 3491–3514.
- [18] M. Goitein, M. Abrams, D. Rowell, H. Pollari, J. Wiles, Multidimensional treatment planning: II. Beams eye-view, back projection, and projection through CT sections, *Int. J. Radiat. Oncol. Biol. Phys.* 9 (1983) 789–797.
- [19] A. Pugachev, L. Xing, Pseudo beam’s-eye-view as applied to beam orientation selection in intensity-modulated radiation therapy, *Int. J. Radiat. Oncol. Biol. Phys.* 51 (2001) 1361–1370.

- [20] A. Pugachev, L. Xing, Computer-assisted selection of coplanar beam orientations in intensity-modulated radiation therapy, *Phys. Med. Biol.* 46 (2001) 2467–2476.
- [21] G. J. Lim, W. Cao, A two-phase method for selecting IMRT treatment beam angles: Branch-and-Prune and local neighborhood search, *Eur. J. Oper. Res.* 217 (2012) 609–618.
- [22] E. K. Lee, T. Fox, I. Crocker, Simultaneous beam geometry and intensity map optimization in intensity-modulated radiation therapy, *Int. J. Radiat. Oncol. Biol. Phys.* 64 (2006) 301–320.
- [23] E. Schreibmann, M. Lahanas, L. Xing, D. Baltas, Multiobjective evolutionary optimization of the number of beams, their orientations and weights for intensity-modulated radiation therapy, *Phys. Med. Biol.* 49 (2004) 747–770.
- [24] D. M. Aleman, A. Kumar, R. K. Ahuja, H. E. Romeijn, J. F. Dempsey, Neighborhood search approaches to beam orientation optimization in intensity modulated radiation therapy treatment planning, *J. Global Optim.* 42 (2008) 587–607.
- [25] G. J. Lim, J. Choi, R. Mohan, Iterative solution methods for beam angle and fluence map optimization in intensity modulated radiation therapy planning, *OR Spectrum* 30 (2008) 289–309.
- [26] J. Stein, R. Mohan, X. H. Wang, T. Bortfeld, Q. Wu, K. Preiser, C. C. Ling, W. Schlegel, Number and orientation of beams in intensity-modulated radiation treatments, *Med. Phys.* 24 (1997) 149–160.
- [27] S. Soderstrom, A. Brahme, Optimization of the dose delivery in a few field techniques using radiobiological objective functions, *Med. Phys.* 20 (1993) 1201–1210.
- [28] H. E. Romeijn, R. K. Ahuja, J. F. Dempsey, A. Kumar, J. Li, A novel linear programming approach to fluence map optimization for intensity modulated radiation therapy treatment planing, *Phys. Med. Biol.* 48 (2003) 3521–3542.
- [29] E. K. Lee, T. Fox, I. Crocker, Integer programming applied to intensity-modulated radiation therapy treatment planning, *Ann. Oper. Res.* 119 (2003) 165–181.

- [30] K. Cheong, T. Suh, H. Romeijn, J. Li, J. Dempsey, Fast Nonlinear Optimization with Simple Bounds for IMRT Planning, *Med. Phys.* 32 (2005) 1975–1975.
- [31] D. Craft, T. Halabi, H. Shih, T. Bortfeld, Approximating convex Pareto surfaces in multiobjective radiotherapy planning, *Med. Phys.* 33 (2006) 3399–3407.
- [32] G. J. Lim, M. C. Ferris, S. J. Wright, D. M. Shepard, M. A. Earl, An optimization framework for conformal radiation treatment planning, *INFORMS J. Comput.* 19 (2007) 366–380.
- [33] A. R. Conn, K. Scheinberg, L. N. Vicente, *Introduction to Derivative-Free Optimization*, SIAM, Philadelphia, 2009.
- [34] C. Davis, Theory of positive linear dependence, *Am. J. Math.* 76 (1954) 733–746.
- [35] P. Alberto, F. Nogueira, H. Rocha, L. N. Vicente, Pattern search methods for user-provided points: Application to molecular geometry problems, *SIAM J. Optim.* 14 (2004) 1216–1236.
- [36] C. Audet, J. E. Dennis Jr., Analysis of generalized pattern search methods, *SIAM J. Optim.* 13 (2003) 889–903.
- [37] J. Moré, S. Wild, Benchmarking Derivative-Free Optimization Algorithms, *SIAM J. Optim.* 20 (2009) 172–191.
- [38] N. J. Higham, *The Matrix Computation Toolbox*. www.ma.man.ac.uk/higham/mctoolbox.
- [39] A. L. Custódio, H. Rocha, L. N. Vicente, Incorporating minimum Frobenius norm models in direct search, *Comput. Optim. Appl.* 46 (2010) 265–278.
- [40] A. L. Custódio, L. N. Vicente, Using sampling and simplex derivatives in pattern search methods, *SIAM J. Optim.* 18 (2007) 537–555.
- [41] H. Rocha, J. M. Dias, B. C. Ferreira, M. C. Lopes, Beam angle optimization using pattern search methods: initial mesh-size considerations, *Proc. of the 1st International Conference on Operations Research and Enterprise Systems*, Vilamoura, Portugal, February 4 - 6, 2012.

- [42] A. I. F. Vaz, L. N. Vicente, A particle swarm pattern search method for bound constrained global optimization, *J. Global Optim.* 39 (2007) 197–219.
- [43] H. Rocha, Principal Component Regression For Construction Of Wing Weight Estimation Models, Ph.D. Thesis, Department of Mathematics and Statistics, Old Dominion University, Norfolk, Virginia, USA, 2005.
- [44] H. Rocha, W. Li, A. Hahn, Principal Component Regression Using Radial Basis Functions Interpolation, *Wavelets and Splines: Athens 2005*, Nashboro PRes, Brentwood, (402–415) 2006.
- [45] H. Rocha, W. Li, A. Hahn, Principal Component Regression for Fitting Wing Weight Data of Subsonic Transports, *J. Aircraft* 43 (2006) 1925–1936.
- [46] M. Buhmann, *Radial Basis Functions: Theory and Implementations*, Cambridge University Press, Cambridge, UK, 2003.
- [47] J. Carr, W. Fright, R. Beatson, Surface interpolation with radial basis functions for medical imaging, *IEEE T. Med. Imaging* 16 (1997) 96–107.
- [48] H. Rocha, On the selection of the most adequate radial basis function, *Appl. Math. Model.* 33 (2009) 1573–1583.
- [49] M. Powell, Radial Basis Function Methods for Interpolation to Functions of Many Variables, *HERMIS: Int. J. Computer Maths & Appl.* 3 (2002) 1–23.
- [50] M. Powell, Recent Research at Cambridge on Radial Basis Functions, *New Developments in Approximation Theory*, Internat. Ser. Numer. Math., Birkhauser, Basel, (215–232) 1999.
- [51] C. Baxter, The Interpolation Theory of Radial Basis Functions, Ph.D. Thesis, Department of Applied Mathematics & Theoretical Physics (DAMTP), Cambridge University, UK, 2002.
- [52] I. Schoenberg, Metric Spaces and Completely Monotone Functions, *Ann. Math.* 39 (1938) 811–841.

- [53] C. Micchelli, Interpolation of Scattered Data: Distance Matrices and Conditionally Positive Definite Functions, *Constr. Approx.* 2 (1986) 11–22.
- [54] R. Schaback, H. Wendland, Characterization and Construction of Radial Basis Functions, In: *Multivariate Approximation and Applications*, N. Dyn, D. Leviatan, D. Levin and A. Pinkus (Eds), Cambridge University Press, UK, (1–24) 2001.
- [55] S. Rippa, An algorithm for selecting a good value for the parameter c in radial basis function interpolation, *Adv. Comput. Math.* 11 (1999) 193–210.
- [56] G. E. Fasshauer, J. G. Zhang, On choosing “optimal” shape parameters for RBF approximation, *Numer. Algor.* 45 (2007) 345–368.
- [57] J. Tu, D. R. Jones, Variable Screening in metamodel design by cross-validated moving least squares method, *Proceedings of the 44th AIAA*, Norfolk, Virginia, USA, 2003.
- [58] J. Tu, Cross-validated Multivariate Metamodeling Methods for Physics-based Computer Simulations, *Proceedings of the IMAC-XXI*, Kissimmee, Florida, USA, 2003.
- [59] M. Stone, Cross-Validatory Choice and Assessment of Statistical Predictions, *J. R. Stat. Soc.* 36 (1974) 111–147.
- [60] B. Efron, R. J. Tibshirani, *An introduction to the Bootstrap*, Chapman & Hall, London, UK, 1993.
- [61] J. Fan, R. Li, Variable selection via nonconcave penalized likelihood and its oracle properties, *J. Amer. Stat. Assoc.* 96 (2001) 1348–1360.
- [62] R. Li, A. Sudjianto, Analysis of computer experiments using penalized likelihood Gaussian Kriging model, *Proceedings of 2003 American Society of Mechanical Engineers (ASME) International Design Automation Conference, DETC2003/DAC-48758*, Chicago, Illinois, USA, 2003.
- [63] J. Fan, I. Gijbels, *Local Polynomial Modeling and Its Applications*, Chapman & Hall, New York, USA, 1996.
- [64] G. Wahba, *Spline Models for Observational Data*, SIAM, Philadelphia, 1990.

- [65] H. Rocha, J. M. Dias, B. C. Ferreira, M. C. Lopes, Direct search applied to beam angle optimization in radiotherapy design, Inesc Research Reports, nr. 6, 2010.
- [66] MATLAB, The MathWorks Inc., <http://www.mathworks.com>.
- [67] J. O. Deasy, A. I. Blanco, V. H. Clark, CERR: A Computational Environment for Radiotherapy Research, *Med. Phys.* 30 (2003) 979–985.
- [68] J. O. Deasy, E. K. Lee, T. Bortfeld, M. Langer, K. Zakarian, J. Alaly, Y. Zhang, H. Liu, R. Mohan, R. Ahuja, A. Pollack, J. Purdy, R. Rardin, A collaboratory for radiation therapy planning optimization research, *Ann. Oper. Res.* 148 (2006) 55–63.
- [69] J. A. Nelder, R. Mead, A simplex method for function minimization, *Comput. J.* 7 (1965) 308–313.