

Selection of intensity modulated radiation therapy treatment beam directions using radial basis functions within a pattern search methods framework

H. Rocha · J. M. Dias · B. C. Ferreira ·
M. C. Lopes

Received: date / Accepted: date

Abstract The selection of appropriate radiation incidence directions in radiation therapy treatment planning is important for the quality of the treatment plan, both for appropriate tumor coverage and for better organ sparing. The objective of this paper is to discuss the benefits of using radial basis functions within a pattern search methods framework in the optimization of the highly non-convex beam angle optimization (BAO) problem. Pattern search methods are derivative-free optimization methods that require few function value evaluations to converge and have the ability to avoid local entrapment. These two characteristics gathered together make pattern search methods suited to address the 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

H. Rocha
INESCC, Rua Antero de Quental 199,
3000-033 Coimbra, Portugal
E-mail: hrocha@mat.uc.pt

J. M. Dias
FEUC, Av. Dias da Silva 165,
3004-512 Coimbra, Portugal
E-mail: joana@fe.uc.pt

B. C. Ferreira
I3N, Campus Universitário de Santiago,
3810-193 Aveiro, Portugal
E-mail: brigida@ua.pt

M. C. Lopes
IPOC-FG, EPE, Av. Bissaya Barreto 98,
3000-075 Coimbra, Portugal
E-mail: mclopes@ipocoimbra.min-saude.pt

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 retrospective treated cases of head-and-neck tumors at the Portuguese Institute of Oncology of Coimbra is used to discuss the benefits of using this approach in the optimization of the BAO problem.

Keywords Intensity modulated radiation therapy · Beam angle optimization · Pattern search methods · Radial basis functions

1 Introduction

The purpose of radiation therapy is to deliver a dose of radiation to the tumor volume to sterilize all cancer cells while 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. 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 (e.g., [5, 13, 29, 31, 32]). However, despite the fact that almost every angle is possible for radiation delivery, the use of coplanar angles, i.e., angles that lay in the plane of rotation of the gantry, 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, relying mostly on experience, despite the evidence presented in the literature that appropriate radiation beam incidence directions can lead to a plan's quality improvement [13, 28, 46].

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 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: beam angle optimization (BAO) problem, fluence map optimization (FMO) problem, and leaf sequencing problem. Most of the efforts in the IMRT optimization community have been devoted to optimizing beamlet intensities [2, 7, 9, 10, 22, 36, 43, 44, 45, 50], once the beam angles have been selected by the treatment planner. Comparatively, less research effort has been directed to the optimization of beam angles [18]. 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 [6, 17, 29], genetic algorithms [18, 20, 24] or

particle swarm optimization [23]. Although those global heuristics can theoretically avoid local optima, 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 [19,29,37]. 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 [38] present a computer assisted selection of coplanar angles based on a variation of the beam's eye view concept. Ehr Gott et al. [18] discuss a mathematical framework that unifies the approaches found in literature. Aleman et al. [5] propose a response surface approach and include non-coplanar angles in beam orientation optimization. Lim and Cao [25] propose an approach that consists of two sequential phases: branch-and-prune and local neighborhood search. Schreiber et al. [48] propose a hybrid multiobjective evolutionary optimization algorithm for IMRT treatment 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 [13], set cover [21] or gradient searches [9].

Here, similarly to [3,5,9,13,20,26,32,48], we will use the optimal solution of the FMO problem to drive our BAO problem. Many 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 [51]. The BAO problem is quite difficult since it is a highly non-convex optimization problem with many local minima [9,49]. 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. Recent literature addressed this time issue with warm starts [4]. Alternatively, 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 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 also

to obtain a better coverage of the search space in amplitude. A set of retrospective treated cases of head-and-neck tumors at the Portuguese Institute of Oncology of Coimbra is used to discuss the benefits of this approach in the optimization of the BAO problem. We will compare the performance of our approach with a gradient based local search approach and both with the typical equispaced coplanar treatment plans. The paper 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 for 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. 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:

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

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 (e.g., [43, 44]), mixed integer linear models (e.g., [22, 36]), nonlinear models (e.g., [7, 50]), and multiobjective models (e.g., [10, 45]). Recently, the use of a convex FMO formulation and interior point methods greatly speed up computational times, yielding optimal treatment plans in seconds, making it a viable option for clinical applications [2].

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 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 which originates large-scale problems.

Here, we will use a convex penalty function voxel-based nonlinear model (e.g., [3, 5, 32]). 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 [44]:

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

$$\text{s.t. } w_j \geq 0, j = 1, \dots, N_b,$$

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 [3], 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 [3].

The FMO model is used as a black-box function. Other models used before for BAO include convex penalty function structure-based approaches [26] and a variety of linear approaches [9,27]. 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 [8] 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

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

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 $2n$ [14]. Positive bases with $n + 1$ and $2n$ elements are referred to as minimal and maximal positive basis, respectively. Commonly used minimal and maximal positive bases are $[I - e]$, with I being the identity matrix of dimension n and $e = [1 \dots 1]^\top$, and $[I - I]$, respectively.

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 [14], 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 [1, 14].

The pattern search methods framework [1] is the class of the most used and implemented directional direct search methods. Pattern search methods framework as described in [1] is presented next. Let us denote by \mathbf{V} the $n \times p$ matrix whose 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 [1]. 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. The most common choice for the mesh-size parameter update

is to halve the mesh-size parameter at unsuccessful iterations and to keep it or double it at successful ones. We are able to describe now the pattern search methods framework.

Algorithm 1 Pattern search methods framework

Initialization:

- Set $k = 0$.
- Choose $\mathbf{x}^0 \in \mathbb{R}^n$, $\alpha_0 > 0$, and a positive spanning set \mathbf{V} .

Iteration:

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 $\alpha_{k+1} = \frac{1}{2} \times \alpha_k$. Set $k \leftarrow k + 1$ and return to step 1 for a new iteration.
 4. Mesh expansion: let $\alpha_{k+1} = \alpha_k$ (or $\alpha_{k+1} = 2 \times \alpha_k$). Set $k \leftarrow k + 1$ and return to step 1 for a new iteration.
-

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 called mesh or pattern generator and $\bar{\mathbf{z}}_i$ is an integer vector in \mathbb{Z}^n . This assumption is crucial for global convergence, ensuring that the mesh has only a finite number of points in a compact set [1]:

Theorem 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 $\mathbf{x}^{k+1} = \arg \min_{\mathbf{x} \in \mathcal{N}(\mathbf{x}^k)} f(\mathbf{x})$ 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 1 (the proof can be found in [53]) 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., [1, 11].

The purpose of the mesh-size parameter, beyond its use as stopping criteria, 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 [33] for comparison of several derivative-free optimization algorithms is $\alpha_0 = \max\{1, \|\mathbf{x}^0\|_\infty\}$. However, this commonly used choice of initial mesh-size parameter might not be adequate for our problem at hand. Note that, if the initial mesh parameter is a power of 2, ($\alpha_0 = 2^\ell, \ell \in \mathbb{N}$), 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.

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 [12]. 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 the 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 [11] 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 may cause the algorithm to jump over lower local minima than the obtained one, leading to worse or

similar results obtained when starting with smaller mesh-size parameters and at the cost of more function value evaluations [41].

An alternative and popular approach to keep small mesh-size parameters and still have a good coverage of the whole 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 consequently the overall computational time. Moreover, the obtained good span of \mathbb{R}^2 in amplitude is only obtained by overlapping all the iterates giving the illusion that unusual beam angle configurations were tested while in fact only local searches around the initial beam angle configurations were performed. Here, we will adopt a different strategy. We will consider a single starting point, a small initial mesh-size parameter, and will 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 [56]. However, the number of function evaluations required for these types 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.

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 (e.g., aeronautics [42]). 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 selecting RBFs a priori without a valid criteria [40].

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

$g(\mathbf{x})$ 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 α_j are the coefficients to be determined by interpolation conditions, $g(\mathbf{x}^k) = f(\mathbf{x}^k)$ ($k = 1, \dots, N$), $\|\mathbf{x} - \mathbf{x}^j\|$ 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 [40]).

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 [35] 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. 1). 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:

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

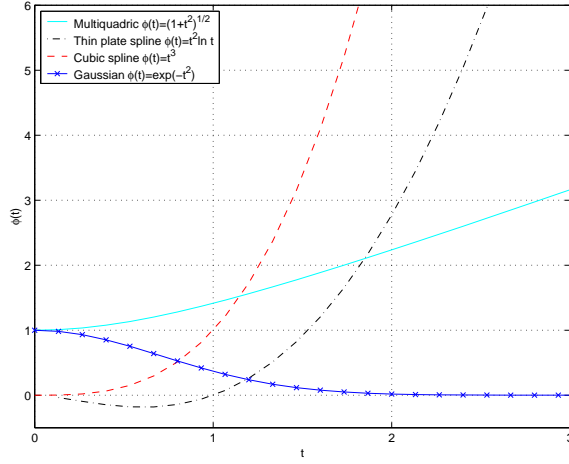


Fig. 1 Graphs of commonly used radial basis functions.

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

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

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 [35]. 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:

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

The key results on solvability of RBF interpolations related to the four RBFs shown in Fig. 1 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 [47].

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 values of $\theta_1, \dots, \theta_n$ that yield an interpolant $g(\mathbf{x})$ with the most accurate trend prediction possible.

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.

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, cross-validation (CV) [52,55] was proposed to find $\varphi(t)$ and θ_i that lead to an approximate response model $g(\mathbf{x})$ with good prediction capability and proved to be effective [54]. The leave-one-out CV procedure is usually used in model parameter tuning for RBF interpolation (see [54], e.g.) and has the following three steps:

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:

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

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. Using 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 Incorporation of RBF interpolation in the pattern search methods framework tailored for the BAO problem

The most common approach for incorporating interpolation models in pattern search methods consists of forming an interpolation model and finding its minimum in the search step. For example, in Custódio et al. [11], the search step computes a single trial point using minimum Frobenius norm quadratic models to be minimized within a trust region. Here, 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 addressing the BAO problem, it might only be generalized for similar problems, and does not include the formal minimization of the RBF model. RBF trial points are computed using the following strategy:

Algorithm 2 RBF trial points

Initialization: Build an RBF model.

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

1. Evaluate the RBF model for 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 m 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 influence the quality of the local minimizer found and simultaneously to properly explore the search space in amplitude without a random criteria. Therefore, each beam direction is tested for every degree between the previous beam direction and the next one as stated in step 1. Step 2 forces the algorithm to consider RBF model directions in regions not yet explored. In our implementation we considered $m = 4$.

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]. We used the last version of SID-PSM [11,12] which is a MATLAB [30] implementation of the pattern search methods (PSM) 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 the RBF interpolation models applied to the BAO problem is the following.

Algorithm 3 PSM framework using RBFs within search step

Initialization:

- Set $k = 0$.
- Choose $\mathbf{x}^0 \in \mathbb{R}^n$, $\alpha_0 > 0$, and a positive spanning set \mathbf{V} .

Iteration:

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 achieves 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 $\alpha_{k+1} = \frac{1}{2} \times \alpha_k$. Set $k \leftarrow k + 1$ and return to step 1 for a new iteration.
 4. Mesh expansion: let $\alpha_{k+1} = \alpha_k$ (or $\alpha_{k+1} = 2 \times \alpha_k$). Set $k \leftarrow k + 1$ and return to step 1 for a new iteration.
-

The total number of points computed in the search step is $n + 1$ (n RBF trial points plus the minimum Frobenius norm quadratic trial point). More conservative strategies could be adopted, e.g., only considering the best of the RBF trial points. The inclusion of RBF trial points in the search step has three distinct purposes: to improve the coverage of the search space in amplitude, to 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 proved to be difficult to obtain. The patients' CT sets and delineated structures were exported via Dicom RT to a freeware computational environment for radiotherapy research (see Fig. 2). 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 salivary glands are also important OARs. The parotid gland is the largest of the three salivary glands. The submandibular gland is the second largest salivary gland and are also important and difficult to spare in most head-and-neck cases. At IPOC both the submandibular glands and the third largest salivary gland, the sublingual gland, are included on a single structure named oral cavity. A common complication due to salivary 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 salivary glands 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 the treatment 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, etc.). For simplicity, in this study, the OARs used for treatment optimization were limited to the spinal cord, the brainstem, the parotid glands and the oral cavity.

The tumor to be treated plus some safety margins is called the 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, the tolerance doses for the OARs and the sizes of all structures considered in the optimization are presented in Table 1. The dose grid resolution was $3.5 \text{ mm} \times 3.5 \text{ mm} \times 3 \text{ mm}$.

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.

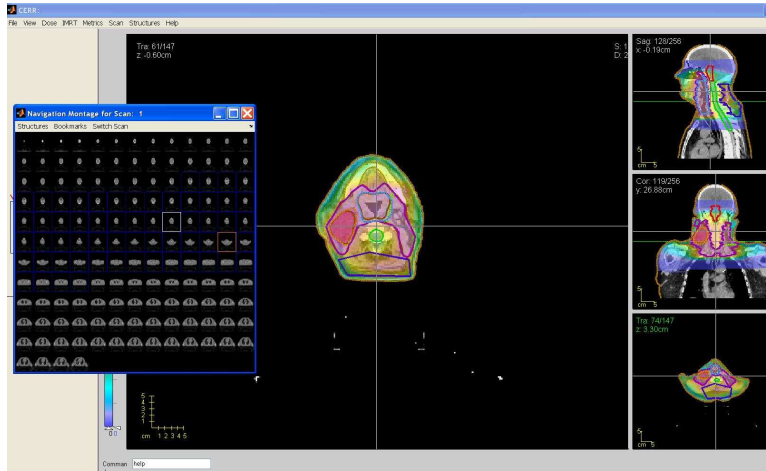


Fig. 2 Illustration of structures visualized in CERR [15].

Table 1 Sizes and prescribed doses for all the structures considered for IMRT optimization.

Structure	Mean dose	Max dose	Prescribed dose	Number of voxels			
				case 1	case 2	case 3	case 4
Spinal cord	–	45 Gy	–	1382	3567	3265	1424
Brainstem	–	54 Gy	–	1715	2072	2087	1569
Left parotid	26 Gy	–	–	1576	1536	2538	676
Right parotid	26 Gy	–	–	1390	1807	2367	684
Oral cavity	30 Gy	–	–	4646	4511	6768	2853
PTV1	–	–	70.0 Gy	4001	1485	16860	4417
PTV2	–	–	59.4 Gy	31119	43649	69748	28721
Body	–	80 Gy	–	1790592	1413138	1608589	664886

6 Results

The RBFs incorporation in the pattern search methods framework was 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 five to nine equally spaced coplanar orientations around the patient. Treatment plans with seven equispaced coplanar beams were used at IPOC and are commonly used in practice to treat head-and-neck cases [5]. Treatment plans that require fewer beams than in current practice may reduce the length of the treatment time, which is an important clinical aspect in IMRT. Moreover, using fewer beams increases the importance of each beam and BAO becomes more critical. Therefore, both seven-beam and five-beam treatment plans were obtained using our BAO algorithms, denoted *PSM*, *PSM-RBF1*, *PSM-RBF2*, *PSM-RBF3*, *PSM-RBF4*, whether 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 seven-beam and five-beam equispaced coplanar treatment plans denoted *equi7* and *equi5*, respectively.

Gradient based search algorithms are used to tackle the BAO problem by many treatment planning systems including the one used at IPOC. A reason for the popular use of gradient based search algorithms is that they converge very quickly to the closest local minimum. However, for a highly non-convex optimization problem with many local minima, such as the BAO problem, that is more a disadvantage than an advantage. As an attempt to contour this issue, many local searches are usually performed using a multi-start strategy. However, such strategy might not be effective due to the so called ‘curse of dimensionality’. Since the search space has five or seven dimensions, the number of beams we aim to optimize, the search space is too large for an efficient multi-start strategy (at least when trying to use few function value evaluations). We will compare the performance of our BAO algorithm not only with the typical equispaced coplanar treatment plans but also with a gradient based local search approach, similar to the one implemented by Craft [9], denoted *LS*.

Despite the fact that the purpose of this work is not to compare the results obtained by our approach with the results obtained by many other approaches previously used, we also tested different versions of beam’s-eye-view, a technique widely used and spread in the literature. The results were not as good as the presented by our approach. Nevertheless, the motivation for testing beam’s-eye-view among many other strategies and methods present in the literature includes the fact that it seems a good strategy to incorporate beam’s-eye-view in the search step in future works, similarly to the inclusion of RBFs in this work.

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 [15] are used widely for IMRT treatment planning research. The ORART – operations research applications in radiation therapy [16] 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.

For the gradient based local search approach we used the unconstrained multivariable minimization function (*fminunc*) of MATLAB 7.4.0 (R2007a) Optimization Toolbox. The starting beam angle sets include the equispaced coplanar beam angle set and random starting beam angle sets produced by generating n random spacings from a lognormal distribution of mean $360/n$ and standard deviation 25, and the starting angle is a random number uniform between 0 and 360 [9].

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 n and $e = [1 \ \dots \ 1]^T$). Each of these directions corresponds 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 mesh-size parameter was set to $\alpha_0 = 4$ when RBFs were used in the search step. Larger powers of 2 were tested without improvements in the quality of the results and at a cost of more function value evaluations. The initial mesh-size parameter for PSM was set to $\alpha_0 = 32$ to obtain similar numbers of function value evaluations and similar coverage of the search space in amplitude. Since the initial points were integer vectors, all iterates will have integer values as long as the mesh parameter does not become less than one. Therefore, the stopping criteria adopted was the mesh parameter becoming less than one. Since we want to improve the quality of the typical equispaced treatment plans, the equispaced coplanar beam angles were considered as the initial points for the beam angle optimization processes. The choice of these initial points 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.

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 would like to verify which particular RBF performs best 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 [34] 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 the seven and five-beam angle optimization process for the four retrospective treated cases of head-and-neck tumors are presented in Tables 2 and 3, respectively.

The most important outcome of the results shown in Tables 2 and 3 is that regardless the RBF considered, the use of RBFs within the pattern search methods framework approach outperforms the other approaches in terms of objective function value. Moreover, in average they require less function value evaluations and consequently less computational time to obtain a better solution. Another important conclusion is that all the pattern search methods framework approaches obtained clearly better objective function value than the gradient based local search approach. Moreover, the gain in terms of ob-

Table 2 Results of the seven-beam angle optimization process using *LS*, *PSM*, *PSM-RBF1*, *PSM-RBF2*, *PSM-RBF3* and *PSM-RBF4*. The initial function value (Finit), the best obtained function value (Fopt), the percentage of decrease in the objective function value, the number of function evaluations (Fevals) and the time in hours are displayed for each case.

Case	algorithm	Finit	Fopt	% decrease	Fevals	time (h)
1	<i>LS</i>	186.6	184.1	1.3	336	6.6
	<i>PSM</i>	193.2	179.5	7.1	201	4.2
	<i>PSM-RBF1</i>	193.2	178.1	7.8	178	3.7
	<i>PSM-RBF2</i>	193.2	178.2	7.8	146	2.9
	<i>PSM-RBF3</i>	193.2	177.6	8.1	182	3.8
	<i>PSM-RBF4</i>	193.2	177.8	8.0	172	3.6
2	<i>LS</i>	228.6	221.7	3.0	272	6.3
	<i>PSM</i>	242.7	213.9	11.9	145	3.6
	<i>PSM-RBF1</i>	242.7	212.3	12.5	126	3.2
	<i>PSM-RBF2</i>	242.7	211.9	12.7	184	4.3
	<i>PSM-RBF3</i>	242.7	212.2	12.6	192	4.6
	<i>PSM-RBF4</i>	242.7	212.7	12.4	130	3.2
3	<i>LS</i>	232.1	227.2	2.1	272	11.1
	<i>PSM</i>	238.6	224.5	5.9	172	7.8
	<i>PSM-RBF1</i>	238.6	222.2	6.9	91	3.8
	<i>PSM-RBF2</i>	238.6	221.3	7.3	167	7.1
	<i>PSM-RBF3</i>	238.6	222.8	6.6	138	5.9
	<i>PSM-RBF4</i>	238.6	222.3	6.8	203	8.9
4	<i>LS</i>	200.5	195.4	2.5	305	6.8
	<i>PSM</i>	202.2	191.4	5.3	140	3.4
	<i>PSM-RBF1</i>	202.2	190.1	6.0	171	3.9
	<i>PSM-RBF2</i>	202.2	189.5	6.3	163	3.8
	<i>PSM-RBF3</i>	202.2	189.6	6.2	168	3.9
	<i>PSM-RBF4</i>	202.2	191.5	5.3	197	4.8

jective function value of the gradient based local search approach is mostly due to the random choice of a better starting point rather than due to the gradient based local search performance. In terms of deciding which RBF suits better the problem at hand no conclusion can be drawn. For different cases, different RBF obtained slightly better results. The CV error of an interpolation model can also be used as an objective tool to decide which model is better. However, those values were also inconclusive and therefore are not included.

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). For the pattern search methods framework approaches, the best results in terms of objective function value, in average, were obtained by *PSM-RBF2*. Since the BAO influence is better perceived when using less beam angles, because greater improvements with respect to the equispaced beam angle configuration are obtained, DVH comparisons will be made for five-beam angle treatment plans obtained by *PSM-RBF2*, *LS* and *equi5*. 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,

Table 3 Results of the five-beam angle optimization process using *LS*, *PSM*, *PSM-RBF1*, *PSM-RBF2*, *PSM-RBF3* and *PSM-RBF4*. The initial function value (Finit), the best obtained function value (Fopt), the percentage of decrease in the objective function value, the number of function evaluations (Fevals) and the time in hours are displayed for each case.

Case	algorithm	Finit	Fopt	% decrease	Fevals	time (h)
1	<i>LS</i>	194.3	192.2	1.1	216	3.1
	<i>PSM</i>	202.5	185.6	8.3	152	2.4
	<i>PSM-RBF1</i>	202.5	185.5	8.4	95	1.4
	<i>PSM-RBF2</i>	202.5	183.9	9.2	139	2.1
	<i>PSM-RBF3</i>	202.5	184.0	9.1	129	1.9
	<i>PSM-RBF4</i>	202.5	184.9	8.7	171	2.8
2	<i>LS</i>	253.8	245.3	3.3	222	3.8
	<i>PSM</i>	267.1	233.8	12.5	129	2.4
	<i>PSM-RBF1</i>	267.1	232.6	12.9	89	1.6
	<i>PSM-RBF2</i>	267.1	231.9	13.2	91	1.7
	<i>PSM-RBF3</i>	267.1	233.6	12.5	109	1.9
	<i>PSM-RBF4</i>	267.1	233.7	12.5	194	3.8
3	<i>LS</i>	243.6	236.4	3.0	174	4.9
	<i>PSM</i>	243.6	221.4	9.1	123	3.8
	<i>PSM-RBF1</i>	243.6	218.1	10.5	164	5.1
	<i>PSM-RBF2</i>	243.6	220.8	9.4	105	3.2
	<i>PSM-RBF3</i>	243.6	218.8	10.2	104	3.0
	<i>PSM-RBF4</i>	243.6	224.5	7.8	67	1.9
4	<i>LS</i>	231.2	228.3	1.3	282	4.3
	<i>PSM</i>	240.2	217.0	9.7	112	1.9
	<i>PSM-RBF1</i>	240.2	213.7	11.0	138	2.3
	<i>PSM-RBF2</i>	240.2	213.8	11.0	105	1.7
	<i>PSM-RBF3</i>	240.2	215.0	10.5	90	1.5
	<i>PSM-RBF4</i>	240.2	216.6	9.8	179	3.3

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 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 4. We can verify that *PSM-RBF2* treatment plans consistently obtained slightly better target coverage numbers compared to *equi5* treatment plans. Moreover, target coverage numbers are favorable to *PSM-RBF2* treatment plans compared to *LS* treatment plans. Organ sparing results are shown in Table 5. All the treatment plans fulfill the maximum dose requirements for the spinal cord and the brainstem. However, as expected, the main differences reside in salivary glands sparing. The *equi5* treatment plans could never enhance parotid sparing and fail to spare the oral cavity for the first and fourth cases. While *LS* treatment plans do slightly improve the parotid's mean dose and oral cavity mean dose, *PSM-RBF2* treatment plans obtained significant improvements and manage to fulfill the mean dose requirements for most of the cases. The differences between *equi5* treatment plans and *PSM-RBF2* treatment plans, concerning salivary glands sparing, and also target coverage, show a clear advantage for the

Table 4 Target coverage obtained by five-beam treatment plans.

Case	Target coverage	<i>PSM-RBF2</i>	<i>LS</i>	<i>equi5</i>
1	PTV1 at 95 % volume	67.6 Gy	67.6 Gy	67.3 Gy
	PTV1 % > 93% of Rx (%)	99.8	99.7	99.6
	PTV1 % > 110% of Rx (%)	0.0	0.0	0.0
	PTV2 at 95 % volume	57.6 Gy	57.4 Gy	57.3 Gy
	PTV2 % > 93% of Rx (%)	97.5	96.9	97.1
	PTV2 % > 110% of Rx (%)	17.3	17.5	17.1
2	PTV1 at 95 % volume	66.6 Gy	66.6 Gy	66.6 Gy
	PTV1 % > 93% of Rx (%)	98.8	98.7	98.8
	PTV1 % > 110% of Rx (%)	0.0	0.0	0.0
	PTV2 at 95 % volume	56.2 Gy	56.2 Gy	55.7 Gy
	PTV2 % > 93% of Rx (%)	96.3	96.2	95.6
	PTV2 % > 110% of Rx (%)	5.6	6.3	6.5
3	PTV1 at 95 % volume	66.6 Gy	66.6 Gy	66.4 Gy
	PTV1 % > 93% of Rx (%)	98.3	98.2	97.9
	PTV1 % > 110% of Rx (%)	0.0	0.0	0.0
	PTV2 at 95 % volume	55.6 Gy	54.5 Gy	53.8 Gy
	PTV2 % > 93% of Rx (%)	95.4	94.3	93.8
	PTV2 % > 110% of Rx (%)	18.3	18.8	19.0
4	PTV1 at 95 % volume	66.4 Gy	66.3 Gy	66.2 Gy
	PTV1 % > 93% of Rx (%)	98.0	97.8	97.7
	PTV1 % > 110% of Rx (%)	0.0	0.0	0.0
	PTV2 at 95 % volume	56.4 Gy	55.9 Gy	55.6 Gy
	PTV2 % > 93% of Rx (%)	96.2	95.6	95.4
	PTV2 % > 110% of Rx (%)	26.6	26.3	26.5

PSM-RBF2 treatment plans. DVH results for the first and fourth cases, where both parotid and oral cavity sparing were not obtained by *equi5* treatment plans, are displayed in Fig. 3 to illustrate the numbers presented in Tables 4 and 5. For clarity, the DVHs were split in PTV1 and PTV2 and the remaining structures distributed as an attempt to better visualize the results. The asterisks indicate 95% of PTV volumes versus 95% of the prescribed doses. The results displayed in Fig. 3 confirm the benefits of using the optimized beam directions obtained and used in *PSM-RBF2* treatment plan. The benefits are clear with respect both to the equispaced solution and also to the gradient based local search solution.

In order to perceive the differences between the best five and seven-beam treatment plans, DVH results for the treatment plans obtained by *PSM-RBF2* for the second case using five and seven beams are displayed in Fig. 4. There is a slight advantage of the seven-beam treatment plan both in terms of target coverage and organ sparing. Nevertheless, the best five-beam treatment plan obtained competitive results compared to the best seven-beam treatment plan results.

Table 5 OARs sparing obtained by five-beam treatment plans.

Case	OAR	Mean Dose (Gy)			Max Dose (Gy)		
		<i>PSM-RBF2</i>	<i>LS</i>	<i>equi5</i>	<i>PSM-RBF2</i>	<i>LS</i>	<i>equi5</i>
1	Spinal cord	–	–	–	39.1	40.0	39.8
	Brainstem	–	–	–	51.5	52.8	52.8
	Left parotid	25.3	26.9	27.7	–	–	–
	Right parotid	21.6	24.4	27.1	–	–	–
	Oral cavity	30.2	32.1	33.7	–	–	–
2	Spinal cord	–	–	–	44.9	45.5	44.8
	Brainstem	–	–	–	53.7	54.6	53.8
	Left parotid	26.0	28.9	28.5	–	–	–
	Right parotid	25.5	28.8	28.7	–	–	–
	Oral cavity	29.3	30.3	29.9	–	–	–
3	Spinal cord	–	–	–	44.2	44.0	44.4
	Brainstem	–	–	–	47.3	49.8	48.5
	Left parotid	26.6	25.5	27.7	–	–	–
	Right parotid	28.6	29.6	29.5	–	–	–
	Oral cavity	27.8	28.9	29.4	–	–	–
4	Spinal cord	–	–	–	40.7	41.1	41.2
	Brainstem	–	–	–	48.5	49.4	49.3
	Left parotid	24.5	28.1	28.6	–	–	–
	Right parotid	23.9	26.4	27.0	–	–	–
	Oral cavity	31.1	35.2	35.7	–	–	–

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. The improvement of the local solutions in terms of objective function value

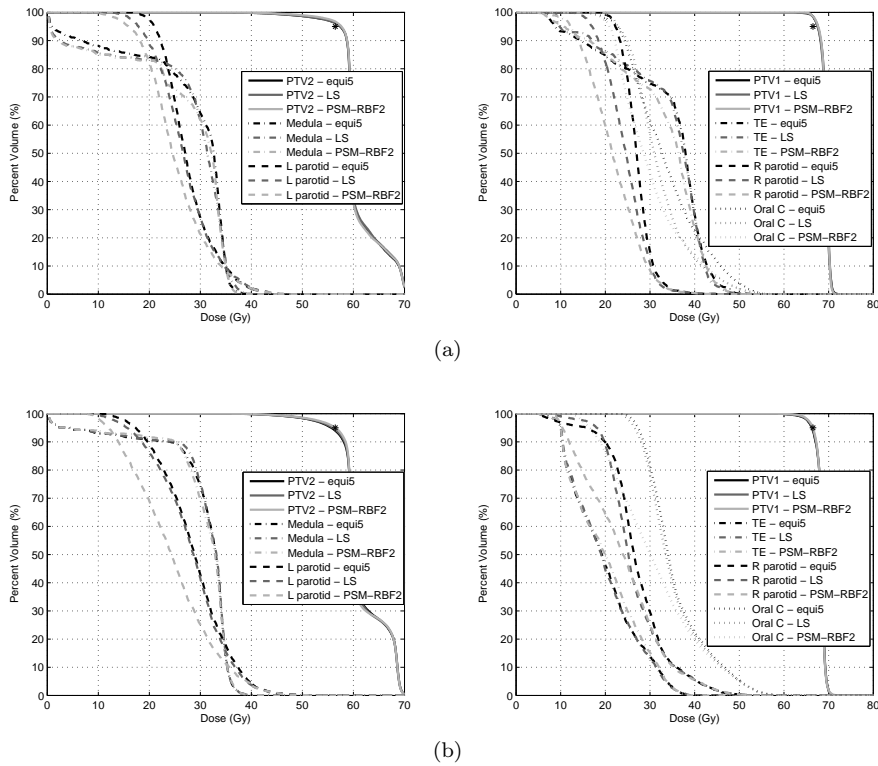


Fig. 3 Cumulative dose volume histogram comparing the results obtained by *equi5*, *LS* and *PSM-RBF2* for the first and fourth cases, 3(a) and 3(b) respectively.

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 increases 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

This work was supported by FEDER funds through the COMPETE program and Portuguese funds through FCT under project grant PTDC/EIA-CCO/121450/2010. This work has also been partially supported by FCT under

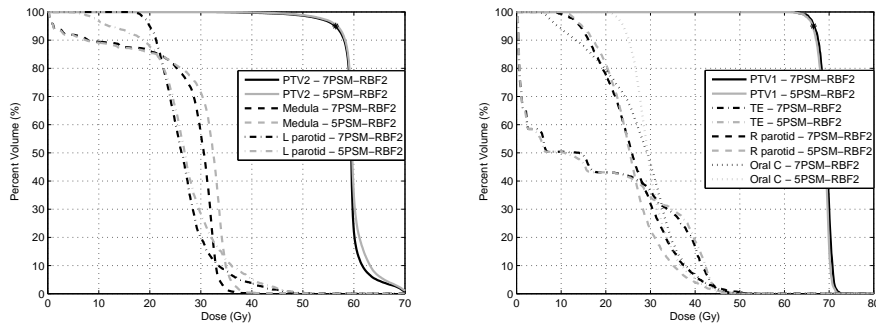


Fig. 4 Cumulative dose volume histogram comparing the results obtained by *PSM-RBF2* for the second case using 7 and 5 beams, *7PSM-RBF2* and *5PSM-RBF2* respectively.

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

References

1. Alberto, P., Nogueira, F., Rocha, H., Vicente, L.N.: Pattern search methods for user-provided points: Application to molecular geometry problems. *SIAM J. Optim.* **14**, 1216–1236 (2004)
2. Aleman, D.M., Glaser, D., Romeijn, H.E., Dempsey, J.F.: A primal-dual interior point algorithm for fluence map optimization in intensity modulated radiation therapy treatment planning. *Phys. Med. Biol.* **55**, 5467–5482 (2010)
3. Aleman, D.M., Kumar, A., Ahuja, R.K., Romeijn, H.E., Dempsey, J.F.: Neighborhood search approaches to beam orientation optimization in intensity modulated radiation therapy treatment planning. *J. Global Optim.* **42**, 587–607 (2008)
4. Aleman, D.M., Mišić, V.V., Sharpe, M.B.: Computational enhancements to fluence map optimization for total marrow irradiation using IMRT. Forthcoming in *Comput. Oper. Res.* DOI: 10.1016/j.cor.2011.05.028.
5. Aleman, D.M., Romeijn, H.E., Dempsey, J.F.: A response surface approach to beam orientation optimization in intensity modulated radiation therapy treatment planning. *INFORMS J. Comput.: Computat. Biol. Med. Appl.* **21**, 62–76 (2009)
6. Bortfeld, T., Schlegel, W.: Optimization of beam orientations in radiation therapy: some theoretical considerations. *Phys. Med. Biol.* **38**, 291–304 (1993)
7. Cheong, K., Suh, T., Romeijn, H., Li, J., Dempsey, J.: Fast Nonlinear Optimization with Simple Bounds for IMRT Planning. *Med. Phys.* **32**, 1975–1975 (2005)
8. Conn, A.R., Scheinberg, K., Vicente, L.N.: *Introduction to Derivative-Free Optimization*. SIAM, Philadelphia (2009)
9. Craft, D.: Local beam angle optimization with linear programming and gradient search. *Phys. Med. Biol.* **52**, 127–135 (2007)
10. Craft, D., Halabi, T., Shih, H., Bortfeld, T.: Approximating convex Pareto surfaces in multiobjective radiotherapy planning. *Med. Phys.* **33**, 3399–3407 (2006)
11. Custódio, A.L., Rocha, H., Vicente, L.N.: Incorporating minimum Frobenius norm models in direct search. *Comput. Optim. Appl.* **46**, 265–278 (2010)
12. Custódio, A.L., Vicente, L.N.: Using sampling and simplex derivatives in pattern search methods. *SIAM J. Optim.* **18**, 537–555 (2007)
13. Das, S.K., Marks, L.B.: 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**, 643–655 (1997)

14. Davis, C.: Theory of positive linear dependence. *Am. J. Math.* **76**, 733–746 (1954)
15. Deasy, J.O., Blanco, A.I., Clark, V.H.: CERR: A Computational Environment for Radiotherapy Research. *Med. Phys.* **30**, 979–985 (2003)
16. Deasy, J.O., Lee, E.K., Bortfeld, T., Langer, M., Zakarian, K., Alaly, J., Zhang, Y., Liu, H., Mohan, R., Ahuja, R., Pollack, A., Purdy, J., Rardin, R.: A collaboratory for radiation therapy planning optimization research. *Ann. Oper. Res.* **148**, 55–63 (2006)
17. Djajaputra, D., Wu, Q., Wu, Y., Mohan, R.: Algorithm and performance of a clinical IMRT beam-angle optimization system. *Phys. Med. Biol.* **48**, 3191–3212 (2003)
18. Ehr Gott, M., Holder, A., Reese, J.: Beam selection in radiotherapy design. *Linear Algebra Appl.* **428**, 1272–1312 (2008)
19. Goitein, M., Abrams, M., Rowell, D., Pollari, H., Wiles, J.: Multidimensional treatment planning: II. Beams eye-view, back projection, and projection through CT sections. *Int. J. Radiat. Oncol. Biol. Phys.* **9**, 789–797 (1983)
20. Haas, O.C., Burnham, K.J., Mills, J.: Optimization of beam orientation in radiotherapy using planar geometry. *Phys. Med. Biol.* **43**, 2179–2193 (1998)
21. Lee, C-H, Aleman, D.M., Sharpe, M.B.: A set cover approach to fast beam orientation optimization in intensity modulated radiation therapy for total marrow irradiation. *Phys. Med. Biol.* **56**, 5679–5695 (2011)
22. Lee, E.K., Fox, T., Crocker, I.: Integer programming applied to intensity-modulated radiation therapy treatment planning. *Ann. Oper. Res.* **119**, 165–181 (2003)
23. Li, Y., Yao, D., Yao, J., Chen, W.: A particle swarm optimization algorithm for beam angle selection in intensity modulated radiotherapy planning. *Phys. Med. Biol.* **50**, 3491–3514 (2005)
24. Li, Y., Yao, J., Yao, D.: Automatic beam angle selection in IMRT planning using genetic algorithm. *Phys. Med. Biol.* **49**, 1915–1932 (2004)
25. Lim, G.J., Cao, W.: A two-phase method for selecting IMRT treatment beam angles: Branch-and-Prune and local neighborhood search. *Eur. J. Oper. Res.* **217**, 609–618 (2012)
26. Lim, G.J., Choi, J., Mohan, R.: Iterative solution methods for beam angle and fluence map optimization in intensity modulated radiation therapy planning. *OR Spectrum* **30**, 289–309 (2008)
27. Lim, G.J., Ferris, M.C., Wright, S.J., Shepard, D.M., Earl, M.A.: An optimization framework for conformal radiation treatment planning. *INFORMS J. Comput.* **19**, 366–380 (2007)
28. Liu, H.H., Jauregui, M., Zhang, X., Wang, X., Dongand, L., Mohan, R.: Beam angle optimization and reduction for intensity-modulated radiation therapy of non-small-cell lung cancers. *Int. J. Radiat. Oncol. Biol. Phys.* **65**, 561–572 (2006)
29. Lu, H.M., Kooy, H.M., Leber, Z.H., Ledoux, R.J.: Optimized beam planning for linear accelerator-based stereotactic radiosurgery. *Int. J. Radiat. Oncol. Biol. Phys.* **39**, 1183–1189 (1997)
30. MATLAB, The MathWorks Inc.. <http://www.mathworks.com>.
31. Meedt, G., Alber, M., Nüsslin, F.: Non-coplanar beam direction optimization for intensity-modulated radiotherapy. *Phys. Med. Biol.* **48**, 2999–3019 (2003)
32. Mišić, V.V., Aleman, D.M., Sharpe, M.B.: Neighborhood search approaches to non-coplanar beam orientation optimization for total marrow irradiation using IMRT. *Eur. J. Oper. Res.* **205**, 522–527 (2010)
33. Moré, J., Wild, S.: Benchmarking Derivative-Free Optimization Algorithms. *SIAM J. Optim.* **20**, 172–191 (2009)
34. Nelder, J.A., Mead, R.: A simplex method for function minimization. *Comput. J.* **7**, 308–313 (1965)
35. Powell, M.: Radial Basis Function Methods for Interpolation to Functions of Many Variables. *HERMIS: Int. J. Computer Maths & Appl.* **3**, 1–23 (2002)
36. Preciado-Walters, F., Langer, M.P., Rardin, R.L., Thai, V.: Column generation for IMRT cancer therapy optimization with implementable segments. *Ann. Oper. Res.* **148**, 65–79 (2006)
37. Pugachev, A., Xing, L.: Pseudo beam’s-eye-view as applied to beam orientation selection in intensity-modulated radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.* **51**, 1361–1370 (2001)

38. Pugachev, A., Xing, L.: Computer-assisted selection of coplanar beam orientations in intensity-modulated radiation therapy. *Phys. Med. Biol.* **46**, 2467–2476 (2001)
39. Rippa, S.: An algorithm for selecting a good value for the parameter c in radial basis function interpolation. *Adv. Comput. Math.* **11**, 193–210 (1999)
40. Rocha, H.: On the selection of the most adequate radial basis function. *Appl. Math. Model.* **33**, 1573–1583 (2009)
41. Rocha, H., Dias, J.M., Ferreira, B.C., Lopes, M.C.: Beam angle optimization using pattern search methods: initial mesh-size considerations. *Proceedings of the 1st International Conference on Operations Research and Enterprise Systems* (2012)
42. Rocha, H., Li, W., Hahn, A.: Principal Component Regression for Fitting Wing Weight Data of Subsonic Transports. *J. Aircraft* **43**, 1925–1936 (2006)
43. Romeijn, H.E., Ahuja, R.K., Dempsey, J.F., Kumar, A.: A column generation approach to radiation therapy treatment planning using aperture modulation. *SIAM J. Optim.* **15**, 838–862 (2005)
44. Romeijn, H.E., Ahuja, R.K., Dempsey, J.F., Kumar, A., Li, J.: A novel linear programming approach to fluence map optimization for intensity modulated radiation therapy treatment planing. *Phys. Med. Biol.* **48**, 3521–3542 (2003)
45. Romeijn, H.E., Dempsey, J.F., Li, J.: A unifying framework for multi-criteria fluence map optimization models. *Phys. Med. Biol.* **49**, 1991–2013 (2004)
46. Rowbottom, C.G., Webb, S., Oldham, M.: Improvements in prostate radiotherapy from the customization of beam directions. *Med. Phys.* **25**, 1171–1179 (1998)
47. Schaback, R., Wendland, H.: *Characterization and Construction of Radial Basis Functions. Multivariate Approximation and Applications*, Cambridge University Press, 1–24 (2001)
48. Schreibmann, E., Lahanas, M., Xing, L., Baltas, D.: Multiobjective evolutionary optimization of the number of beams, their orientations and weights for intensity-modulated radiation therapy. *Phys. Med. Biol.* **49**, 747–770 (2004)
49. Soderstrom, S., Brahme, A.: Optimization of the dose delivery in a few field techniques using radiobiological objective functions. *Med. Phys.* **20**, 1201–1210 (1993)
50. Spirou, S., Chui, C.-S.: A gradient inverse planning algorithm with dose-volume constraints. *Med. Phys.* **25**, 321–333 (1998)
51. Stein, J., Mohan, R., Wang, X.H., Bortfeld, T., Wu, Q., Preiser, K., Ling, C.C., Schlegel, W.: Number and orientation of beams in intensity-modulated radiation treatments. *Med. Phys.* **24**, 149–160 (1997)
52. Stone, M.: Cross-Validatory Choice and Assessment of Statistical Predictions. *J. R. Stat. Soc.* **36**, 111–147 (1974)
53. Torczon, V.: On the convergence of pattern search algorithms. *SIAM J. Optim.* **7**, 1–25 (1997)
54. Tu, J.: Cross-validated Multivariate Metamodeling Methods for Physics-based Computer Simulations. *Proceedings of the IMAC-XXI* (2003)
55. Tu, J., Jones, D.R.: Variable Screening in metamodel design by cross-validated moving least squares method. *Proceedings of the 44th AIAA* (2003)
56. Vaz, A.I.F., Vicente, L.N.: A particle swarm pattern search method for bound constrained global optimization. *J. Global Optim.* **39**, 197–219 (2007)