BEAM ANGLE OPTIMIZATION IN IMRT USING PATTERN SEARCH METHODS: INITIAL MESH-SIZE CONSIDERATIONS

Humberto Rocha[1], Joana Matos Dias[1,2], Brígida Ferreira[3,4] and Maria do Carmo Lopes[3,4]

[1]*INESCC, Rua Antero de Quental 199, 3000-033 Coimbra, Portugal*

[2]*FEUC, Av. Dias da Silva 165, 3004–512 Coimbra, Portugal*

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

[4]*I3N, Departamento de Física, Universidade de Aveiro, Aveiro, Portugal*

Keywords: Beam angle optimization, Derivative-free optimization, Radiation therapy.

Abstract: In radiotherapy treatments, the selection of appropriate radiation incidence directions is decisive for the quality of the treatment, both for appropriate tumor coverage and for enhance better organs sparing. However, the beam angle optimization (BAO) problem is still an open problem and, in clinical practice, beam directions continue to be manually selected by the treatment planner in a time-consuming trial and error iterative process. The goal of BAO is to improve the quality of the radiation incidence directions used and, at the same time, release the treatment planner for other tasks. The objective of this paper is to discuss the benefits of using pattern search methods in the optimization of the BAO problem. Pattern search methods are derivative-free optimization methods that require few function value evaluations to progress and converge and have the ability to avoid local entrapment. These two characteristics gathered together make pattern search methods suited to address the BAO problem. Considerations about the initial mesh-size importance and other strategies for a better coverage and exploration of the BAO problem search space will be debated.

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. 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, the treatment planning is usually divided into three smaller problems which can be solved sequentially: beam angle optimization (BAO) problem, fluence map optimization (FMO) problem, and leaf sequencing problem. 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 (Das and Marks, 1997). 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 mixed integer programming approaches (Lee et al., 2006), neighborhood search approaches (Aleman et al., 2008), hybrid multiobjective evolutionary optimization approaches (Schreibmann et al., 2004), and gradient search approaches (Craft, 2007). The BAO problem is quite difficult since it is a highly non-convex optimization problem with many local minima (Craft, 2007). Therefore, methods that avoid being easily trapped in local minima should be used. Pattern search methods are suited to address the BAO problem since they have the ability to avoid local entrapment. Here, we will discuss the benefits of using pattern search methods in the optimization of the BAO problem. Considerations about the initial mesh-size importance and other strategies for a better coverage and exploration of the BAO problem search space will be debated.

Rocha H., Matos Dias J., Ferreira B. and do Carmo Lopes M..
BEAM ANGLE OPTIMIZATION IN IMRT USING PATTERN SEARCH METHODS: INITIAL MESH-SIZE CONSIDERATIONS.
DOI: 10.5220/0003716803550360
In *Proceedings of the 1st International Conference on Operations Research and Enterprise Systems* (ICORES-2012), pages 355-360
ISBN: 978-989-8425-97-3

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. Most of the previous BAO studies are based on a variety of scoring methods or approximations of 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. Here, for modelling the BAO problem, we will use the optimal solution value of the FMO problem as measure of the quality of a given beam angle set (Aleman et al., 2008; Craft, 2007). 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 k to be the fixed number of (coplanar) beam directions, i.e., k 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). Here we will consider all continuous $[0^\circ, 360^\circ]$ gantry angles 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:

$$\begin{array}{ll} \min & f(\theta_1,\ldots,\theta_k) \\ s.t. & \theta_1,\ldots,\theta_k \in \mathbb{R}^k. \end{array} \qquad (1)$$

Here, the objective $f(\theta_1,\ldots,\theta_k)$ that measures the quality of the set of beam directions $\theta_1,\ldots,\theta_k$ 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 (Romeijn et al., 2003), mixed integer linear models (Lee et al., 2006), nonlinear models (Aleman et al., 2008), and multiobjective models (Craft et al., 2006).

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 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 (V) and the number of columns equals the number of beamlets (N) from all beam directions considered. Therefore, using matrix format, we can say that the total dose received by the voxel i is given by $\sum_{j=1}^{N} D_{ij} w_j$, 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 (Aleman et al., 2008). 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:

$$\begin{array}{ll} \min_w & \sum_{i=1}^{V} F_i \left(\sum_{j=1}^{N} D_{ij} w_j \right) \\ s.t. & w_j \geq 0,\ j=1,\ldots,N, \end{array} \qquad (2)$$

with F_i defined as asymmetric quadratic penalty functions (Romeijn et al., 2003):

$$F_i\left(\sum_{j=1}^{N} D_{ij} w_j\right) = \frac{1}{v_S}\left[\underline{\lambda}_i \left(T_i - \sum_{j=1}^{N} D_{ij} w_j\right)_+^2 + \overline{\lambda}_i \left(\sum_{j=1}^{N} D_{ij} w_j - T_i\right)_+^2\right],$$

where T_i is the desired dose for voxel i, $\underline{\lambda}_i$ and $\overline{\lambda}_i$ 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 (Aleman et al., 2008), 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 PATTERN SEARCH METHODS

Pattern search methods are directional direct search methods that belong to a broader class of derivative-free optimization methods. Pattern search methods are iterative methods generating a sequence of iterates $\{x_k\}$ using positive bases (or positive spanning sets) and moving in the 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 $\mathbb{R}^n$, but no proper set does.

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 v_i in a positive basis (or positive spanning set) that is a descent direction (Davis, 1954), i.e., there is an $\alpha > 0$ such that $f(x_k + \alpha v_i) < f(x_k)$. This is the core of directional direct search methods and in particular of pattern search methods.

Pattern search methods are iterative methods generating a sequence of non-increasing iterates $\{x_k\}$. Given the current iterate x_k, at each iteration k, the next point x_{k+1} is chosen from a finite number of candidates on a given mesh M_k (defined using the vectors forming a positive spanning set) aiming to provide a decrease on the objective function: $f(x_{k+1}) < f(x_k)$. Pattern search methods consider two steps at every iteration. The first step consists of a finite search on the mesh, 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 is unsuccessful a second step, called the poll step, is performed around the current iterate with the goal of decreasing the objective function. The poll step follows stricter rules and appeals to the concepts of positive bases. The poll step attempts to perform a local search in a mesh neighborhood that, for a sufficiently small mesh-size parameter Δ_k, is guaranteed to provide a function reduction, unless the current iterate is at a stationary point (Alberto et al., 2004). So, if the poll step also fails, the mesh-size parameter Δ_k must be decreased. 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 (Moré and Wild, 2009) for comparison of several derivative-free optimization algorithms is $\Delta_0 = \max\{1, \|x_0\|_\infty\}$. 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. However, this choice of initial mesh-size parameter might not be adequate as we will illustrate along this paper.

For driving the resolution of the BAO problem, we will use the last version of SID-PSM (Custódio and Vicente, 2007; Custódio et al., 2010) which is a MATLAB implementation of the pattern search methods that incorporate 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. The default initial mesh-size parameter of SID-PSM is also $\Delta_0 = \max\{1, \|x_0\|_\infty\}$. The benefits of using this particular implementation of pattern search methods in the optimization of the BAO problem and the influence of the initial mesh-size parameter choice on the quality of the solution obtained are illustrated using a clinical example of head and neck case that is presented next.

4 NUMERICAL TESTS AND DISCUSSION

A clinical example of a retrospective treated case of head and neck tumor at the Portuguese Institute of

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

Structure	Mean dose	Max dose	Prescribed dose
Spinal cord	–	45 Gy	–
Brainstem	–	54 Gy	–
Left parotid	26 Gy	–	–
Right parotid	26 Gy	–	–
PTV left	–	–	59.4 Gy
PTV right	–	–	50.4 Gy
Body	–	70 Gy	–

Oncology of Coimbra is used to verify the benefits and issues of using pattern search methods in the optimization of the BAO problem. The patients' CT set and delineated structures were exported via Dicom RT to a freeware computational environment for radiotherapy research. 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 case in study it was separated in two parts: PTV left and PTV right. The prescribed doses for all the structures considered in the optimization are presented in Table 1.

Our tests were performed on a 2.66Ghz Intel Core Duo PC with 3 GB RAM. In order to facilitate convenient access, visualization and analysis of patient treatment planning data, the computational tools developed within MATLAB and CERR (Deasy et al., 2003) (computational environment for radiotherapy research) were used as the main software platform to embody our optimization research and provide the necessary dosimetry data to perform optimization in IMRT. The dose was computed using CERR's pencil beam algorithm (QIB). 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. In order to initialize the algorithm we need to choose an initial point x_0, a positive spanning set, and an initial mesh-size parameter $\Delta_0 > 0$. Typically, in head and neck cancer cases, patients are treated with 5 to 9 equispaced beams in a coplanar arrangement. Here, we will consider the equispaced 5 beam configuration with angles 0, 72, 144, 216 and 288, and with 0 collimator angle. Since our goal is to improve the typically used treatment plans, this is a good starting point and thus we will consider x_0 as the previous 5-beam equispaced angle configuration. 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. The spanning set used was the positive spanning set ($[e\ -e\ I\ -I]$, with I being the identity matrix and $e = [1\ \ 1]^T$). Each of these directions corresponds to, respectively, the rotation of all incidence directions clockwise, 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 default initial mesh-size parameter, as mentioned before, is $\Delta_0 = \max\{1, \|x_0\|_\infty\}$. For the considered initial point this would give an initial mesh-size of $\Delta_0 = 288$. For this "cyclic" problem such initial mesh-size is too large implying in practice that huge rotation of angles would occur. Moreover, convergence would take too long leading to an excessive number of function evaluations. Obtaining the optimal solution for a beam angle set is time costly and even if only a beam angle is changed in that set, a complete dose computation is required in order to compute and obtain the corresponding optimal FMO solution. Therefore, few function value evaluations should be used to tackle the BAO problem within a clinically acceptable time frame.

Note that if the initial mesh-size parameter is a power of 2, ($\Delta_0 = 2^p, p \in \mathbb{N}$), and the initial point is a vector of integers, using the default mesh update, i.e., to half the mesh parameter at unsuccessful iterations and to keep it at successful ones, all iterates will be a vector of integers until the mesh parameter size becomes inferior to 1. This possibility is rather interesting for our BAO problem at hand and, for initial mesh-size parameter, we tested ($\Delta_0 = 2^p, p = 1, 2, \ldots$). For the initial point selected, the distance between two consecutive beam angle directions is 72, thus, the maximum Δ_0 we considered was 64. The history of the beam angle optimization process using SID-PSM, for each initial mesh-size parameter considered, is presented in Figure 1. By simple inspection we can verify that only mesh-size parameters 32 and 64 originate a sequence of iterates that are reasonably well distributed by amplitude in $\mathbb{R}^2$, while mesh-size parameters inferior to 32 fail to cover in amplitude all the search space. At first sight, larger mesh-size parameters, which obtain iterates better distributed by amplitude in $\mathbb{R}^2$, should be desirable for an improved global search in conjunction with the search step. One of the main ad-

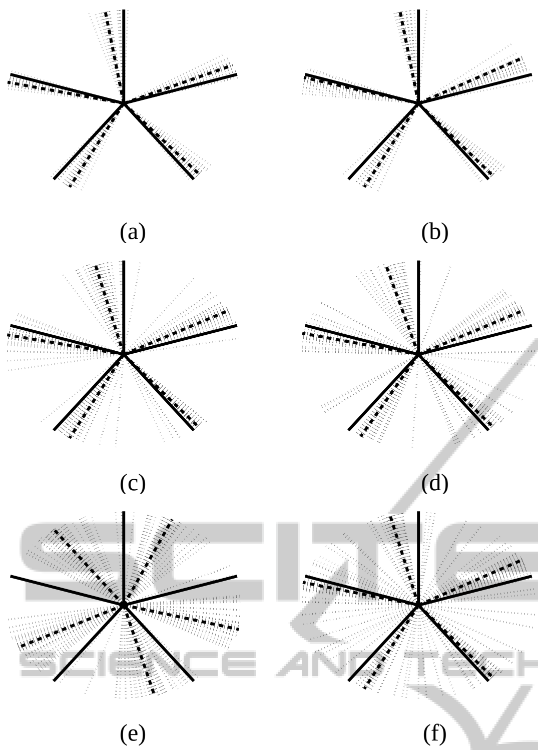

Figure 1: History of the beam angle optimization process using SID-PSM for mesh-size parameters 2 to 64, 1(a) to 1(f) respectively. Initial angle configuration, optimal angle configuration and intermediate angle configurations are displayed with solid, dashed and dotted lines, respectively.

vantages 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 previously used and tested strategies/heuristics that successfully address the BAO problem and enhance for a global search by influencing the quality of the local minimizer or stationary point found by the method. In the last version of SID-PSM, the search step computes a single trial point using minimum Frobenius norm quadratic models to be minimized within a trust region, which enhanced a significant improvement of direct search for black-box non-smooth functions (Custódio et al., 2010) similar to the BAO problem at hand. The size of the trust region is coupled to radius of the sample set. Thus, for an effective global search, the sample points should span all the search space. 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 have the advantage of a better coverage of

Table 2: Results obtained using $x_0 = (0, 72, 144, 216, 288)$ for different initial mesh-size parameters.

Δ_0	f init.	f opt.	f evals.	time (s)
2	90.21	79.70	51	3162
4	90.21	79.60	65	4183
8	90.21	79.26	113	7159
16	90.21	79.36	115	7216
32	90.21	83.16	137	8645
64	90.21	79.26	139	8649

the search space but may cause the algorithm to jump over lower local minima than the obtained one. That was the case for $\Delta_0 = 32$ which originated the worst result. The results obtained for the different initial mesh-size parameters are presented in Table 2. The quality of the treatment plan obtained is directly proportional to the correspondent final objective function value. For this initial point the treatment plans obtained for all initial mesh-size considered except $\Delta_0 = 32$ are equivalent. This means that larger initial mesh parameters do not lead to better local minima despite the improved search space coverage.

An alternative popular approach to keep small mesh-size parameters and still have a good coverage of the search space is to use a multi-start approach. 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 better span of $\mathbb{R}^2$ in amplitude, that is only obtained by overlapping all the iterates which might be fallacious for this particular problem. In future work we aim to use a single starting point, a small initial mesh-size parameter, and obtain a good span in amplitude of $\mathbb{R}^2$ by incorporating an additional global strategy in the search step such as response surface approach or radial basis functions interpolation.

Let us now illustrate the benefits of using a treatment plan with the best optimal angle configuration obtained by SID-PSM (5 *PSM*) compared with the usual treatment plan with equispaced beam directions (5 *equi*). 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. An ideal DVH for the tumor would present 100% volume for all dose values ranging from zero to the prescribed dose value and then drop immediately to zero, indicating that the whole target volume is treated exactly as prescribed. Ideally, the curves for the organs at risk would instead drop immediately to zero, meaning that no volume receives radiation. Another metric usually used for plan evaluation is the volume of PTV that receives 95% of the

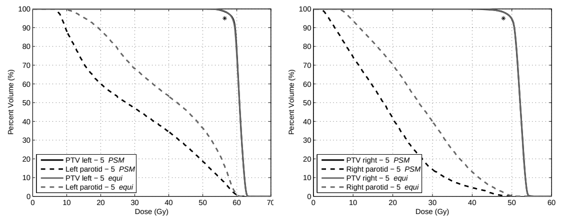

Figure 2: Cumulative dose volume histogram comparing the treatment plans 5 *PSM* and 5 *equi*.

prescribed dose. Typically, 95% of the PTV volume is required. DVH results are displayed in Figure 2. Since parotids are the most difficult organs to spare, for clarity, the DVH only includes the targets and the parotids. The asterisk indicates 95% of PTV volume versus 95% of the prescribed dose. By observing Figure 2 we confirm that both treatment plans fulfill the goal of having 95% of the prescribed dose for 95% of the volume for both PTV right and PTV left. Focusing in parotid sparing we can observe that a better parotid sparing can be obtained using the beam angle solution obtained by SID-PSM.

5 CONCLUSIONS

The BAO problem is a continuous global highly non-convex optimization problem known to be extremely challenging and yet to be solved satisfactorily. Pattern search methods framework is a suitable approach for the resolution of the non-convex BAO problem due to their structure, organized around two phases at every iteration. The poll step, where convergence to a local minima is assured, and the search step, where flexibility is conferred to the method since any strategy can be applied. We have shown that a beam angle set can be locally improved in a continuous manner using pattern search methods. The initial mesh-size parameter importance and other strategies for a better coverage and exploration of the BAO problem search space were tested and debated. In future work, pattern search methods improvement will be tested with the incorporation of additional global strategies in the search step such as response surface approaches or radial basis functions interpolation. We have to highlight the low number of function evaluations required by SID-PSM to obtain locally optimal solutions. The efficiency on the number of function value computation is of the utmost importance, particularly when the BAO problem is modeled using the optimal values of the FMO problem. Thus, the global strategies to be incorporated must comply with this requirement.

ACKNOWLEDGEMENTS

This work has been partially supported by FCT under 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

Alberto, P., Nogueira, F., Rocha, H. and Vicente, L. N. (2004). Pattern search methods for user-provided points: Application to molecular geometry problems. *SIAM J. Optim.*, 14, 1216–1236.

Aleman, D. M., Kumar, A., Ahuja, R. K., Romeijn, H. E. and Dempsey, J. F. (2008). Neighborhood search approaches to beam orientation optimization in intensity modulated radiation therapy treatment planning. *J. Global Optim.*, 42, 587–607.

Craft, D., Halabi, T., Shih, H. and Bortfeld, T. (2006). Approximating convex Pareto surfaces in multiobjective radiotherapy planning. *Med. Phys.*, 33, 3399–3407.

Craft, D. (2007). Local beam angle optimization with linear programming and gradient search. *Phys. Med. Biol.*, 52, 127–135.

Custódio, A. L. and Vicente, L. N. (2007). Using sampling and simplex derivatives in pattern search methods. *SIAM J. Optim.*, 18, 537–555.

Custódio, A. L., Rocha, H. and Vicente, L. N. (2010). Incorporating minimum Frobenius norm models in direct search. *Comput. Optim. Appl.*, 46, 265–278.

Das, S. K. and Marks, L. B. (1997). 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.

Davis, C. (1954). Theory of positive linear dependence. *Am. J. Math.*, 76, 733–746.

Deasy, J. O., Blanco, A. I. and Clark, V. H. (2003). CERR: A Computational Environment for Radiotherapy Research. *Med. Phys.*, 30, 979–985.

Lee, E. K., Fox, T. and Crocker, I. (2006). Simultaneous beam geometry and intensity map optimization in intensity-modulated radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 64, 301–320.

Moré, J. and Wild, S. (2009). Benchmarking Derivative-Free Optimization Algorithms, *SIAM J. Optim.*, 20, 172–191.

Romeijn, H. E., Ahuja, R. K., Dempsey, J. F., Kumar, A. and Li, J. (2003). A novel linear programming approach to fluence map optimization for intensity modulated radiation therapy treatment planing. *Phys. Med. Biol.*, 48, 3521–3542.

Schreibmann, E., Lahanas, M., Xing, L. and Baltas, D. (2004). Multiobjective evolutionary optimization of the number of beams, their orientations and weights for intensity-modulated radiation therapy. *Phys. Med. Biol.*, 49, 747–770.