Available online at www.sciencedirect.com

ScienceDirect

Linear Algebra and its Applications 428 (2008) 1365–1387

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Dose-volume-based IMRT fluence optimization: A fast least-squares approach with differentiability

Yin Zhang *,[1], Michael Merritt [2]

Department of Computational and Applied Mathematics, Rice University, Houston, TX 77005, USA

Received 24 August 2006; accepted 19 September 2007
Available online 26 November 2007
Submitted by J. Galvin

Abstract

In intensity-modulated radiation therapy (IMRT) for cancer treatment, the most commonly used metric for treatment prescriptions and evaluations is the so-called dose-volume constraint (DVC). These DVCs induce much needed flexibility but also non-convexity into the fluence optimization problem, which is an important step in the IMRT treatment planning. Currently, the models of choice for fluence optimization in clinical practice are weighted least-squares models. When DVCs are directly incorporated into the objective functions of least-squares models, these objective functions become not only non-convex but also non-differentiable. This non-differentiability is a problem when software packages designed for minimizing smooth functions are routinely applied to these non-smooth models in commercial IMRT planning systems. In this paper, we formulate and study a new least-squares model that allows a monotone and differentiable objective function. We devise a greedy approach for approximately solving the resulting optimization problem. We report numerical results on several clinical cases showing that, compared to a widely used existing model, the new approach is capable of generating clinically relevant plans at a much faster speed. This improvement can be more than one-order of magnitude for some large-scale problems.

Keywords: IMRT treatment planning; Fluence optimization; Weighted least-squares model; Dose-volume constraint; Differentiability

* Corresponding author.
E-mail address: yzhang@caam.rice.edu (Y. Zhang).
[1] This author's work was supported in part by NSF Grants DMS-0442065 and DMS-0405831.
[2] This author's work was supported in part by NSF Grant DMS-0240058.

0024-3795/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2007.09.037

1. Introduction

This section presents a very brief introduction to a particular problem, the fluence optimization problem, that often occurs in intensity-modulated radiation therapy (IMRT). For more comprehensive information on this topic, the reader is referred to articles collected in two recent books [16,13]. Survey papers on optimization models and methods in this area include [18,4], and a historical perspective can be found in [2].

1.1. IMRT fluence optimization

The most common form of IMRT uses mounted linear accelerator with a treatment head that can rotate around the patient. Usually, fixed treatment beam directions are used for dose delivery. Beam modulation is achieved through the use of a multileaf collimator (MLC) attached to the head of the treatment unit. The gantry mounted linear accelerator and the multi-leaf collimator (MLC) are shown below (see Fig. 1). In order to precisely target the tumors while minimizing exposure of the neighboring healthy structures, the MLC shapes radiation beam through a sequences of movements of its metal leaves.

There are three major parts of the optimization process that can arise during IMRT planning. The first is the *beam-angle optimization* problem which determines the "optimal" number of gantry positions and their angles. This part of the process is often formulated as a combinatorial optimization problem. The second part of the process is called intensity map *optimization*, and is aimed at finding the "optimal" intensity distribution for each beam angle. The third part of the process finds the MLC *leaf sequencing* needed to efficiently deliver the desired intensity pattern. In this paper, however, we will exclusively concentrate on the second problem.

Specifically, we need to determine the X-ray intensity pattern at each point (x, y) on the plane of the MLC aperture for all gantry angles. These beam intensity profiles are represented by two-dimensional, non-negative functions $I_a(x, y)$ for $a = 1, 2, \ldots, k$, where k is the number of gantry angles. See Fig. 2 where five beam angles are used, each with its own beam intensity profile. The purpose of fluence optimization is to find the functions $I_a(x, y)$ such that the tumor (or target) structures receive the prescribed doses and the healthy structures receive as little dose as possible. These goals are fundamentally conflicting and have ill-defined evaluation criteria, making fluence optimization a difficult modeling problem.

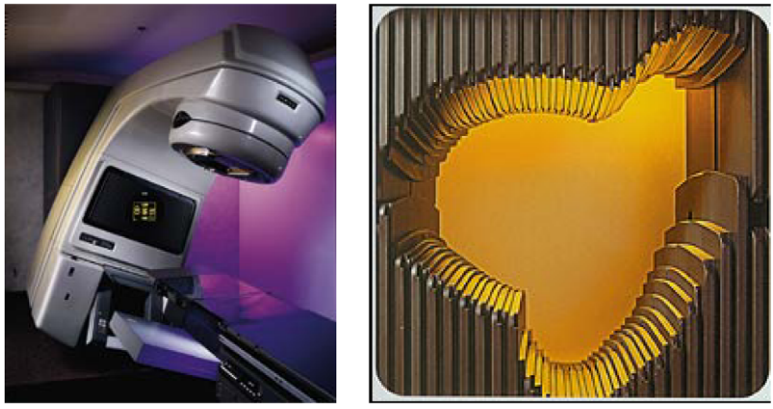

Fig. 1. Linear accelerator (left) and multi-leaf collimator (right).

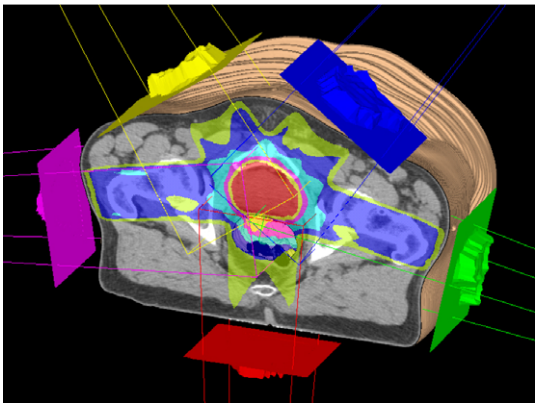

Fig. 2. A five-beam treatment scheme.

In practice, each fluence map function $I_a(x, y)$ is discretized at a rectangular grid on the plane of the MLC aperture and approximated by a set of discrete values $\{I_a(x_i, y_j)\}$. The actual number of these small rectangular elements, or "beamlets," will vary from case to case. For notational convenience, let us collect the unknown beamlet intensity values $\{I_a(x_i, y_j)\}$, $a = 1, 2, \ldots, k$, for all beam angles into a single vector x of n ordered elements, where n is the total number of beamlets for all beam angles. Hence, *the task of fluence optimization is to find a beamlet intensity vector x that produces a dose distribution as close as possible to a prescribed dose distribution*. In the IMRT literature, this process is called inverse planning.

1.2. Dose calculation and dose-volume constraints

In practice, one also needs to discretize the "region of treatment" – the three-dimensional volume of the patient's anatomy containing the target structures and any nearby critical structures that might be adversely affected by the radiation, as well as any parts of the body that the individual treatment beams might enter or exit through. This volume is discretized into small three-dimensional rectangular elements known as "voxels." During the treatment, each voxel will absorb a dose of radiation. We denote the dose values absorbed by the voxels by a vector $d \in \mathbb{R}^m$, where m is the total number of voxels in the region of treatment.

The standard IMRT model for calculating dose absorbed at the ith voxel in the region of treatment is the linear model $d = Ax$, or

$$d_i = \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, 2, \ldots, m, \tag{1}$$

where a_{ij} represents the amount of dose absorbed by the ith voxel per unit intensity emission from the jth beamlet. The collection of values a_{ij} for all the voxels and beamlets forms a matrix $A \in \mathbb{R}^{m\times n}$, known as the (dose) "influence matrix" (or kernel matrix).

The linear dose model $d = Ax$ can be considered as a first-order approximation. Although difficult to solve, radiation absorption as a function of radiation intensities can be modeled with Boltzmann transport equations [9]. Different approximation methods have been proposed for computing the matrix A. Monte Carlo sampling techniques are, for example, among the more popular methods because of their accuracy, but they are also very slow Many faster, but less accurate, dose

calculation engines exist. Currently, dose calculation is still considered an important research area. While acknowledging its importance, we will assume in this paper that a constant influence matrix A is provided to us *a priori*, and we will use it throughout our optimization process. Without loss of generality, we also assume A has no zero rows or columns. This means, respectively, that all voxels receive some non-zero amount of radiation and every beamlet influences at least one voxel's dose. These conditions can be easily met by pre-processing. Typically, $m \gg n$ with m on the order of 10^5 or larger and n of 10^3 up to 10^4. Note the entries a_{ij} are necessarily non-negative. Depending on how much scattered radiation is included, the influence matrix A can be very sparse or fairly dense.

It is generally impossible to avoid giving some dose to nearby critical structures during radiation treatment. This occurs even through multiple beam angles are used in an attempt to focus radiation on the targets while sparing critical structures. The radiation oncologists utilizes dose-volume constraints (DVCs) to prescribe and control the dose at these critical structures. A typical DVC specifies a certain percentage of the volume in a structure that is allowed to be overdosed. For example, a prescription for a left-sided lung cancer case may contain the dose-volume constraint for the healthy right lung that reads

"*no more than 30% volume of the right lung should receive 20 Gy or higher*,"

where "Gy" is the shorthand for "Gray" – the international unit for radiation dose absorption. In addition, the radiation oncologists may specify additional constraints on the same organ, such as "*no more than 40% volume of the right lung should receive 10 Gy or higher*." These DVCs control the amount of radiation reaching the right lung during the treatment. Although other metrics have been proposed, the metrics based on dose-volume constraints have become the *de facto* standard for to prescribing the radiation dose in clinical practice.

DVCs are useful in describing a desired dose distribution, but they also introduce a high degree of complexity to the planning process. In the above example, the exact voxels comprising the 30% of the right lung volume allowed to absorb more than 20 Gy is not specified. This results in a combinatorial component to the problem (once it is discretized). Mathematically, finding an exact global optimum for such a problem can be extremely difficult.

1.3. A perspective on least-squares models

Several classes of models have been proposed and studied for fluence optimization; however, this paper concentrates on a particular class – weighted least-squares models. In particular, it introduces dose-volume-based least-squares models and points out the main advantage and disadvantage of these models.

Least-squares models were the first practical models used in inverse planning [19,3]. Today they continue to be the models of choice in clinical practice, implemented in most commercial IMRT planning systems. The main advantage of these models is their "speed"; that is, they can be approximately solved relatively quickly.

It is perhaps widely-agreed that there are two major sources of difficulties in developing models for fluence optimization:

1. Multi-objectiveness: The problem is inherently multi-objective due to the presence of conflicting goals for multiple structures. The conflicts are not only between the typical two classes of structures-targets and critical organs, but also between different structures in the same class. Despite years of research, multi-objective optimization remains a very difficult problem.

2. Vagueness of clinical objectives [8]: There is not a single accepted method for ranking different treatment plans. This is partly because of the above multi-objectiveness, and partly because of the lack of quantitative metrics to measure biological responses of irradiated tissues.

Due to these fundamental difficulties, the IMRT planning process has been human-dependent and experience-driven, relying heavily on repeated trial-and-error and close interaction between treatment planners and oncologists. More often than not, the experiences of the radiation oncologists plays a decisive role in accepting or rejecting a given treatment plan.

Under these conditions, a fast turn-around time for trial-and-error becomes a great advantage. Weighted least-squares models have a speed advantage over other models such as integer programming (for example, see [10,17]). Integer programming models are more mathematically rigorous but more time consuming to solve. However, weighted least-squares models are generally capable of producing clinically relevant treatment plans (even though they have less control over constraint satisfaction than models imposing explicit constraints). These attributes are responsible for the popularity of least-squares models in clinical practice.

For the weighted least-squares model, the difficulties of the multi-objectiveness and the vagueness of clinical objectives are addressed by adjusting a set of importance weights in a trial-and-error process. In this approach, a positive weight is attached to each anatomical structure, representing the relative priority of fitting the calculated dose to the prescription for that structure. If there are N structures each consisting of a set of voxels, the objective function in a weighted least-squares model takes the form

$$f(x) = \sum_{j=1}^{N} w_j f_j(d(x)), \tag{2}$$

where $d(x) = Ax$ is the calculated dose vector corresponding to a given influence matrix A and a beamlet intensity vector x. The parameter w_j is the weight for the jth structure, and $f_j(d)$ is a piecewise quadratic function that penalizes the deviation of the calculated dose $d(x)$ relative to the prescription dose for the jth structure. For any fixed set of weights, $f(x)$ can be approximately minimized by one of the existing algorithms for continuous optimization subject to non-negativity of the beamlet intensities.

Each penalty function $f_j(d(x))$ in (2) is a sum of quadratic penalty terms, one for each voxel. If the prescription at the ith voxel is either

$$d_i(x) = b_i \quad \text{or} \quad d_i(x) \leqslant b_i,$$

where $d_i(x)$ is the calculated dose value at the ith voxel and b_i is the prescription dose, then the corresponding penalty terms are either

$$(d_i(x) - b_i)^2 \quad \text{or} \quad \max(0, d_i(x) - b_i)^2.$$

In either case, this term is convex and differentiable. Thus, without DVCs the function $f(x)$ in (2) is convex and differentiable, and the resulting optimization problem is theoretically simple.

However, the addition of DVCs adds significant complication. Consider the DVC that *no more than 10% of the volume should receive a dose 20 Gy or higher*. In this case, the standard approach is to let go the 10% voxels with the highest calculated dose values and penalize the rest of the voxels, if any, whose dose values exceed 20 Gy.

The improved flexibility that occurs when DVCs are added to the optimization process comes at a price in that the resulting penalty function $f(x)$ becomes non-convex and non-differentiable. To illustrate this point, consider the example of a 2-voxel structure with the DVC that *no more*

than 50% *of the volume* (*i.e.* 1 *voxel*) *should receive a dose of* 5 *Gy or higher*. In this case, the corresponding penalty function is

$$f(d_1, d_2) = \begin{cases} \max(0, d_1 - 5)^2, & d_1 < d_2, \\ \max(0, d_2 - 5)^2, & d_1 \geqslant d_2, \end{cases}$$

where either none or one of the two voxels is penalized, but never both. A plot of this function will clearly show that it is indeed non-convex and non-differentiable along the line $d_1 = d_2$ (starting from the point (5, 5)).

Given the combinatorial nature of DVCs, the non-convexity is perhaps unavoidable. However, it is not clear if it is possible to avoid the non-differentiability. It is problematic that optimization algorithms designed for differentiable functions are being routinely used in daily clinical practice to minimize non-differentiable functions. Experiments discussed here indicate that this non-differentiability may be responsible for many failed trials in the trial-and-error process of planning.

1.4. Notation and organization

Let n be the total number of relevant beamlets and m the total number of voxels in the treatment region. We define the n-dimensional non-negative orthant as $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x \geqslant 0\}$, and similarly for $\mathbb{R}^{m\times n}_+$. Let $A \in \mathbb{R}^{m\times n}_+$ be a given influence matrix. Let m_t be the number of target voxels. We collect the m_t rows of A corresponding to the target voxels to form a sub-matrix $A_t \in \mathbb{R}^{m_t\times n}$ of A. Thus, for any beamlet intensity vector $x \in \mathbb{R}^n_+$, the vector $A_t x$ gives the dose values in the target voxels. Accordingly, the prescribed target dose values are collected in a vector $b_t \in \mathbb{R}^{m_t}_+$. Recall that vector inequalities are always treated component-wise.

Let $\mathscr{D}_v \subset \mathbb{R}^m_+$ be the set of dose vectors that satisfy all the dose-volume constraints for a given problem. We note that there may be many targets and critical organs with (possibly multi-level) DVCs, but vectors in $\mathscr{D}_v$ satisfy all such requirements.

This paper is organized as follows. Section 2 describes a new least-squares model, which was first introduced by the authors in [20]. A greedy algorithm is developed and studied for this model in Section 3. Numerical results on eight clinical cased are presented in Section 4.

2. A least-squares model with differentiability

Since the fundamental difficulties of multi-objectiveness and vagueness of clinical objectives will persist at least for the foreseeable future, the dominance of weighted least-squares models in clinical practice will most likely continue as well. The purpose of this paper is to carry out an in-depth theoretical and numerical study of a new least-squares model that allows a differentiable objective function and faster numerical optimization.

Fast fluence optimization is also essential for resolving the problem of beam-angle optimization, which is still considered unsolved [2]. Beam-angle optimization needs to repeatedly use this fast fluence optimization as a necessary subroutine.

2.1. Geometric considerations

To support the approach presented here, the geometry of the dose-volume constraint set in the IMRT fluence optimization problem is discussed. The goal is to provide sufficient dose to the targets while satisfying the dose-volume constraints (DVCs) as closely as possible.

Suppose there is a vector $u \in \mathscr{D}_v$ that satisfies the dose-volume constraints. Then finding beamlet intensities $x \in \mathbb{R}^n_+$ whose resulting doses satisfy the dose-volume constraints requires that $Ax \leqslant u$. (For this reason, this paper loosely refers to the u values as "bounds" throughout.) Moreover, it is desirable to meet the target prescription $A_t x = b_t$. Equivalently, it is desirable to find a beamlet intensity vector $x \in \mathbb{R}^n_+$ that satisfies

$$A_t x = b_t, \quad Ax + s = u,$$

for some auxiliary (slack) variable $s \in \mathbb{R}^m_+$.

It is possible to define the *prescription set*

$$\mathscr{H} = \left\{ \begin{bmatrix} b_t \\ u \end{bmatrix} : u \in \mathscr{D}_v \right\} \subset \mathbb{R}^{m_t+m}_+ \tag{3}$$

to contain doses that meet the fixed prescription b_t and satisfies the DVCs. Since b_t is fixed in the definition (3), $\mathscr{H}$ consists of $\mathscr{D}_v$ embedded in the higher dimensional subspace. Furthermore, it is possible to define the (augmented) *physical set*

$$\mathscr{K} = \left\{ \begin{bmatrix} A_t x \\ Ax + s \end{bmatrix} : x, s \geqslant 0 \right\} \subset \mathbb{R}^{m_t+m}_+ \tag{4}$$

to contain the doses that can be realized physically (up to dose calculation and delivery errors). The word "augmented" refers to the addition of the slack variable to the last m components.

It should be clear that $\mathscr{K}$ is a closed convex cone and $\mathscr{H}$ is non-convex since the set $\mathscr{D}_v$, defined by DVCs of a combinatorial nature, is a non-convex set.

2.2. Problem formulation

In order to treat the targets while trying to achieve the dose limits for other structures, it is desirable to find $x \in \mathbb{R}^n_+$ and $s \in \mathbb{R}^m_+$ such that both $A_t x = b_t$ and $Ax + s = u$ hold. However, in trying to solve the IMRT fluence problem, it is not uncommon to find that there is no physically-achievable dose that both satisfies the DVCs and meets the prescription. That is, generally speaking $\operatorname{dist}(\mathscr{H}, \mathscr{K}) > 0$, or equivalently $\mathscr{H} \cap \mathscr{K} = \emptyset$. Thus, it is necessary to determine $d_H \in \mathscr{H}$ and $d_K \in \mathscr{K}$ such that

$$\operatorname{dist}(\mathscr{H}, \mathscr{K}) = \|d_H - d_K\| = \min_{u \in \mathscr{D}_v} \min_{x,s \geqslant 0} \left\| \begin{bmatrix} b_t \\ u \end{bmatrix} - \begin{bmatrix} A_t x \\ Ax + s \end{bmatrix} \right\|, \tag{5}$$

where the distance is in the Euclidean norm. However, it is more convenient to replace the norm in the right-hand side of (5) with the quadratic function

$$q(x, s, u) = \frac{1}{2}\|A_t x - b_t\|^2 + \frac{1}{2}\|Ax + s - u\|^2. \tag{6}$$

Thus, the objective function over the set of all $u \in \mathscr{D}_v$ can written as

$$f(u) = \min_{(x,s) \geqslant 0} q(x, s, u). \tag{7}$$

Namely, $f(u)$ is itself the optimal value of a linear least-squares problem with a non-negativity constraint. Using this notation, it is possible to restate the problem (5) with the equivalent formulation

$$\min_{u \in \mathscr{D}_v} f(u). \tag{8}$$

Solving this problem finds a $u^* \in \mathscr{D}_v$ and associated beamlet intensities $x^* \in \mathbb{R}^n_+$ such that the "deliverable" dose distribution Ax^* is as close as possible to being feasible with respect to the dose-volume constraints and to meeting the prescribed dose b_t in the targets.

It should be emphasized that for any given prescription (b_t, u^0), where b_t is a prescribed target dose and $u^0 \in \mathscr{D}_v$, the model

$$\min_{x \geqslant 0} q(x, s, u^0)$$

is nothing but a regular least-squares model without dose-volume constraints. Therefore, starting from $u = u^0$ and monotonically decrease $f(u)$ gives a better solution (in terms of improved target conformality while having a similar degree of DVC compliance) than the regular least-squares solution. This clearly illustrates the benefit of using dose-volume-based models.

Since it is desirable to allow weightings in the least-squares formulation, in place of $q(x, s, u)$ it is possible to write the equation

$$q_W(x, s, u) := \frac{1}{2}\|W_t(A_t x - b_t)\|^2 + \frac{1}{2}\|W(Ax + s - u)\|^2, \tag{9}$$

where W_t and W are diagonal weighting matrices of appropriate sizes that hold the importance weights for each structure. The weights can, for instance, incorporate scaling factors such as the number of voxels for each structure.

Without loss of generality, it can be assumed that the weighting matrices have already been absorbed into the quantities A, A_t, b_t and $\mathscr{D}_v$, i.e., $A \leftarrow WA$, $A_t \leftarrow W_t A_t$, $b_t \leftarrow W_t b_t$, …, etc. For this reason, it is not necessary to explicitly mention weighting in the formulation that follows.

2.3. Optimality of the subproblem

Our objective f is itself the minimum value of another optimization problem, making it somewhat complicated. This section will examines the properties of f on the domain $\mathscr{D}_v$ for the purpose of constructing an algorithm for solving the overall problem (8). The following assumption is made in all our theoretical results that follow.

Assumption 1. The matrix A_t is full column rank.

Experience shows that this assumption almost always holds. The explanation for this is that the number of target voxels far exceeds that of beamlets, hence A_t has far more rows than columns. The theoretical work presented here depends on uniqueness properties that follow from this assumption. In particular, the work reported here uses this assumption to prove that the subproblem solutions are unique and continuously differentiable functions of u.

Evaluating $f(u)$ requires solving a subproblem parameterized by $u \in \mathscr{D}_v$ which is denoted as $\mathscr{Q}(u)$:

$$\min_{(x,s) \geqslant 0} q(x, s, u). \tag{10}$$

Subproblem $\mathscr{Q}(u)$ is a convex bound-constrained quadratic program that needs to be solved repeatedly. The following examines the subproblem in more detail.

Proposition 1. *Under Assumption 1, the subproblem objective function $q(\cdot, \cdot, u)$ is a strongly convex quadratic function for every $u \in \mathbb{R}^m$. Hence, f is well-defined and*

$$f(u) = q(x(u), s(u), u), \tag{11}$$

where $(x(u), s(u))$ is the unique solution pair of $\mathscr{Q}(u)$.

Assumption 1 implies this strong convexity property since the matrix $A_t^{\mathrm{T}} A_t$ is the Schur complement of the Hessian of $q(x, s, u)$ with respect to (x, s). Thus, it follows from well-known facts of convex optimization that $\mathcal{Q}(u)$ has a unique solution pair $(x(u), s(u))$ for any u. Proposition 1 guarantees the existence and uniqueness of $x(u)$ and $s(u)$ for every u, i.e. they are well-defined functions of u themselves.

It is obvious that the Karush–Kuhn–Tucker (KKT) conditions for problem $\mathcal{Q}(u)$ are necessary and sufficient for optimality. Note that for two vectors v and w of the same dimension, the component-wise minimum $\min(v, w) = 0$ is equivalent to $v \circ w = 0$ and $v, w \geqslant 0$. Using this notation, the KKT conditions for $\mathcal{Q}(u)$ can be written as

$$\min(x, \nabla_x q(x, s, u)) = 0, \tag{12a}$$

$$\min(s, \nabla_s q(x, s, u)) = 0, \tag{12b}$$

where

$$\nabla_x q(x, s, u) = A_t^{\mathrm{T}}(A_t x - b_t) + A^{\mathrm{T}}(Ax + s - u), \tag{13a}$$

$$\nabla_s q(x, s, u) = Ax + s - u. \tag{13b}$$

Therefore, $(x(u), s(u))$ is the solution to $\mathcal{Q}(u)$ if and only if $x(u)$ and $s(u)$ satisfy the KKT conditions (12a) and (12b), respectively. In addition, strict complementarity holds for $\mathcal{Q}(u)$ at (x, s) if

$$x + \nabla_x q(x, s, u) > 0, \tag{14a}$$

$$s + \nabla_s q(x, s, u) > 0. \tag{14b}$$

Lemma 1. *For any u,*

$$s(u) = \max(0, u - Ax(u)), \tag{15a}$$

$$\nabla_s q(x(u), s(u), u) = \max(Ax(u) - u, 0), \tag{15b}$$

where the maximum is taken component-wise.

Proof. Substituting (13b) into (12b) gives $\min(s(u), Ax(u) + s(u) - u) = 0$.

If $[Ax(u)]_i < u_i$, then necessarily $[s(u)]_i > [Ax(u) + s(u) - u]_i \geqslant 0$. Complementarity then implies that in fact $[Ax(u) + s(u) - u]_i = 0$ and thus $[s(u)]_i = [u - Ax(u)]_i$. Otherwise, $[Ax(u)]_i \geqslant u_i$ implies $[Ax(u) + s(u) - u]_i \geqslant [s(u)]_i$, meaning that $[s(u)]_i = 0$. This proves (15a). To obtain (15b), it is necessary to substitute (15a) into (13b). □

Definition 1. Let the "sensitive" index set $S \equiv S(x, u) := \{i : (Ax)_i > u_i\}$ and the matrix $E(S) = \sum_{i \in S} e_i e_i^{\mathrm{T}}$, where e_i is the ith column of the identity matrix.

That is, $E(S)$ is the diagonal matrix with diagonal elements $[E(S)]_{ii} = 1$ for all $i \in S$ and zero otherwise. Using Lemma 1 and this definition of $E(S)$ eliminates s and simplifies the KKT conditions for $\mathcal{Q}(u)$.

Proposition 2. *The KKT conditions for $\mathcal{Q}(u)$ implies*

$$\min(x, A_t^{\mathrm{T}}(A_t x - b_t) + A^{\mathrm{T}} E(S)(Ax - u)) = 0. \tag{16}$$

We observe that while all the target voxels are involved in the above condition, only the "sensitive" healthy voxels show up. One can infer from this observation that the only healthy

voxels ultimately involved in determining $x(u)$ are those with $[Ax(u)]_i > u_i$. As shown in the next section, these voxels are included in the set that has sensitivity to the objective function f. This is a fact that will have significant algorithmic implications later.

Finally, it is necessary to state an explicit expression for the solution $x(u)$ of $\mathcal{Q}(u)$. Note that $s(u)$ can then be computed from (15a). It is first important to partition the indices for $x(u)$ as follows:

$$P = \{i : x(u)_i > 0\}, \quad O = \{i : x(u)_i = 0\}. \tag{17}$$

For a vector v, let v_P denotes the sub-vector of v consisting of the components with indices in P (similarly for v_O). For a matrix M, M_{OP} is the sub-matrix of M with row indices from O and column indices from P (similarly for M_{PP}). The next result follows directly from Proposition 2.

Proposition 3. *Let the index set S^* and the matrix $E(S^*)$ be defined as in Definition 1 for $x = x(u)$. Then the non-zero elements of $x(u)$ are given by*

$$x(u)_P = [A_t^{\mathrm{T}} A_t + A^{\mathrm{T}} E(S^*) A]_{PP}^{-1} (A_t^{\mathrm{T}} b_t + A^{\mathrm{T}} E(S^*) u)_P > 0, \tag{18}$$

and at the same time satisfy

$$[A_t^{\mathrm{T}} A_t + A^{\mathrm{T}} E(S^*) A]_{OP} x(u)_P \geqslant (A_t^{\mathrm{T}} b_t + A^{\mathrm{T}} E(S^*) u)_O. \tag{19}$$

Note that the matrix $[A_t^{\mathrm{T}} A_t + A^{\mathrm{T}} E(S^*) A]_{PP}$ is positive definite under Assumption 1.

2.4. Sensitivity of the objective function

In order to solve the problem stated in Eq. (8), it is important to show how f changes as u is adjusted. Assumption 1 is applied for all the following results.

Theorem 1. *The function $f(u)$ defined in (11) is monotone and non-increasing as u increases; i.e.,*

$$f(u + d) \leqslant f(u) \quad \forall d \in \mathbb{R}^m_+.$$

Moreover, the equality holds if and only if

$$x(u + d) = x(u), \quad s(u + d) = s(u) + d.$$

Proof. Let $d \in \mathbb{R}^m_+$. Then

$$\begin{aligned} f(u + d) &= \frac{1}{2}\|A_t x(u + d) - b_t\|^2 + \frac{1}{2}\|Ax(u + d) + s(u + d) - (u + d)\|^2 \\ &\leqslant \frac{1}{2}\|A_t x(u) - b_t\|^2 + \frac{1}{2}\|Ax(u) + (s(u) + d) - (u + d)\|^2 \\ &= f(u), \end{aligned}$$

since the subproblem $\mathcal{Q}(u + d)$ has the unique minimizer $(x(u + d), s(u + d))$ and the point $(x(u), s(u) + d)$ is feasible with respect to $\mathcal{Q}(u + d)$. The second statement also follows immediately. □

Note that it is always possible to decrease $f(u)$ by increasing u as long as $x(u)$ does not stay the same.

Definition 2. Let the sensitivity of f to increases in u_i be

$$\sigma(u)_i = \limsup_{t\to 0^+} \frac{f(u+te_i) - f(u)}{t}.$$

The limit superior is used in the sensitivity definition because the limit may not exist. Observe that when the limit does exist, $\sigma(u)_i$ is the one-sided partial derivative that gives information about how $f(u)$ changes locally as u_i is increased. Furthermore, whenever f is Gâteaux differentiable, $\sigma(u) = \nabla f(u)$. A mild condition will be provided later in this report that guarantees f being Fréchet differentiable. However, it is possible to examine a straightforward yet important consequence of the definition of $\sigma(u)$.

Theorem 2. *The sensitivity vector satisfies*

$$\sigma(u)_i \begin{cases} \leqslant -[Ax(u) - u]_i, & [Ax(u)]_i > u_i, \\ = 0, & \textit{otherwise}. \end{cases} \tag{20}$$

Proof. First consider any index i such that $[Ax(u)]_i \leqslant u_i$. It can be easily verified that for any $t > 0$ the pair $(x(u), s(u) + te_i)$ satisfies the KKT conditions for $Q(u + te_i)$, so that

$$x(u + te_i) = x(u), \quad s(u + te_i) = s(u) + te_i.$$

Moreover, for any $t > 0$, $f(u + te_i) \equiv f(u)$ and $\sigma(u)_i = 0$.

Now consider any index i such that $Ax(u)]_i > u_i$. Since $f(u + te_i)$ is the optimal value of $\mathcal{Q}(u + te_i)$,

$$\begin{aligned} f(u + te_i) &= q(x(u + te_i), s(u + te_i), u + te_i) \\ &\leqslant q(x(u), s(u), u + te_i) \\ &= f(u) - [Ax(u) - u]_i t + \frac{1}{2}t^2. \end{aligned}$$

Therefore,

$$\frac{f(u + te_i) - f(u)}{t} \leqslant -[Ax(u) - u]_i + \frac{1}{2}t.$$

Letting t go to zero gives the inequalities in (20). □

Consequently, $f(u)$ is sensitive to increases in u_i if and only if $[Ax(u)]_i > u_i$. This characterization of sensitivity is extremely simple and yet completely natural. Also, it agrees with the conclusion following Proposition 2 concerning the healthy voxels' influence on $x(u)$. With no more information about the solutions $x(u)$ and $s(u)$, this is all that can be said about the sensitivity $\sigma(u)$. It is necessary to introduce a condition on the subproblem solutions that will be used for the rest of this report.

Assumption 2. The solution of $\mathcal{Q}(u)$ is strictly complementary.

Under this condition, it is possible to we will examine the differentiability of $f(u)$, beginning with the well-defined functions $x(u)$ and $s(u)$.

Lemma 2. *The functions $x(u)$ and $s(u)$ are continuously differentiable in a neighborhood of points where Assumption 2 holds.*

Proof. Let "$\circ$" denote the component multiplication for vectors and so that is possible to write the complementarity equations in the KKT conditions (12a) as

$$K(x,s,u) := \begin{pmatrix} x \circ \nabla_x q(x,s,u) \\ s \circ \nabla_s q(x,s,u) \end{pmatrix} = 0,$$

which is clearly satisfied at $(\hat{x},\hat{s},\hat{u}) = (x(\hat{u}), s(\hat{u}), \hat{u})$. The invertibility of the Jacobian of K with respect to (x,s) at $(\hat{x},\hat{s},\hat{u})$ will now be verified. The Jacobian $K'_{(x,s)}(x,s,u)$ is

$$\begin{bmatrix} \mathrm{Diag}(\nabla_x q(x,s,u)) + \mathrm{Diag}(x)\nabla_x^2 q(x,s,u) & \mathrm{Diag}(x)A^{\mathrm{T}} \\ \mathrm{Diag}(s)A & \mathrm{Diag}(s + \nabla_s q(x,s,u)) \end{bmatrix},$$

where $\mathrm{Diag}(v)$ is the diagonal matrix with the vector v on the diagonal. Suppose that the equation $K'_{(x,s)}(\hat{x},\hat{s},\hat{u})z = 0$ holds with $z^{\mathrm{T}} = [z_1^{\mathrm{T}} z_2^{\mathrm{T}}]$. Then

$$\mathrm{Diag}(\nabla_x \hat{q})z_1 + \mathrm{Diag}(\hat{x})\nabla_x^2 \hat{q} z_1 + \mathrm{Diag}(\hat{x})A^{\mathrm{T}} z_2 = 0, \tag{21}$$

$$\mathrm{Diag}(\hat{s})Az_1 + \mathrm{Diag}(\hat{s} + \nabla_s \hat{q})z_2 = 0, \tag{22}$$

where "hats" over q denote a gradient or Hessian evaluated at $(\hat{x},\hat{s},\hat{u})$. The assumption of strict complementarity implies the (2, 2) block of $K'_{(x,s)}(\hat{x},\hat{s},\hat{u})$ is positive definite. Thus, (22) can be solved for z_2:

$$z_2 = -\mathrm{Diag}(\hat{s} + \nabla_s \hat{q})^{-1}\mathrm{Diag}(\hat{s})Az_1 = -[I - E(\widehat{S})]Az_1,$$

where $E(\widehat{S})$ is defined as in Definition 1 with $\widehat{S} = S(\hat{x},\hat{u})$. Let

$$D_1 = \mathrm{Diag}(\hat{x} + \nabla_x \hat{q})^{-1}\mathrm{Diag}(\nabla_x \hat{q}),$$
$$D_2 = \mathrm{Diag}(\hat{x} + \nabla_x \hat{q})^{-1}\mathrm{Diag}(\hat{x}),$$
$$M = A_t^{\mathrm{T}} A_t + A^{\mathrm{T}} E(\widehat{S})A.$$

Moreover, partition $\{1,\ldots,n\}$ into $P = \{i : \nabla_x \hat{q} = 0\}$ and $O = \{i : \hat{x} = 0\}$. Substituting z_2 into (21), collecting terms, and pre-multiplying by $\mathrm{Diag}(\hat{x} + \nabla_x \hat{q})^{-1}$ gives

$$(D_1 + D_2 M)z_1 = \sum_{i \in O} e_i e_i^{\mathrm{T}} z_1 + \sum_{i \in P} e_i e_i^{\mathrm{T}} M z_1 = 0. \tag{23}$$

Note that D_1 and D_2 could be simplified by using strict complementarity in x. The equation in (23) can be further simplify to $[z_1]_O = 0$ and $M_{PP}[z_1]_P = 0$. Since M is positive definite under Assumption 1, M_{PP} is positive definite and thus $z = 0$.

Having established the non-singularity of $K'_{(x,s)}(x(\hat{u}), s(\hat{u}), \hat{u})$, the result follows from the Implicit Function Theorem. □

By adding the strict complementarity condition, it can be concluded that f is differentiable and the inequalities in (20) become equalities, providing a closed form expression for the gradient, or sensitivity, of f.

Theorem 3. *If Assumption 2 holds at u, then f is differentiable at u with*

$$\sigma(u) = \nabla f(u) = -\max(Ax(u) - u, 0) \leqslant 0. \tag{24}$$

Proof. Observe from (11) that it is possible to write $f(u) = q(x(u), s(u), u)$. That is, f is q composed with the functions $x(\cdot)$, $s(\cdot)$, and the identity mapping. From Lemma 2, it can be shown that $x(\cdot)$ and $s(\cdot)$ are differentiable at u and clearly the quadratic function q is differentiable. Thus, the well-known chain rule theorem implies that f is differentiable at u so that $\sigma \equiv \nabla f$.

Let $\mathbf{x} \equiv x(\cdot)$ and $\mathbf{s} \equiv s(\cdot)$. Using this notation to distinguish arbitrary variables x and s from the functions $x(u)$ and $s(u)$, the chain rule is applied to (11) to obtain the jth component of the gradient

$$\sigma(u)_j = \nabla f(u)_j = \sum_{i=1}^{n} \frac{\partial q}{\partial x_i}\frac{\partial \mathbf{x}_i}{\partial u_j} + \sum_{i=1}^{m} \frac{\partial q}{\partial s_i}\frac{\partial \mathbf{s}_i}{\partial u_j} + \frac{\partial q}{\partial u_j},$$

where the partial derivatives of q are evaluated at $(x(u), s(u), u)$ and the partial derivatives of $\mathbf{x}$ and $\mathbf{s}$ are evaluated at u. Each term is zero in the two summations. It suffices to consider the kth term in the first summation, $\frac{\partial q}{\partial x_k}\frac{\partial \mathbf{x}_k}{\partial u_j}$. From the KKT conditions in (12a), complementarity gives $\frac{\partial q}{\partial x_k}\mathbf{x}_k = 0$, where $\frac{\partial q}{\partial x_k}$ and $\mathbf{x}_k$ are evaluated at $(x(u), s(u), u)$ and u, respectively. If $\mathbf{x}_k > 0$, then $\frac{\partial q}{\partial x_k} = 0$; so the product $\frac{\partial q}{\partial x_k}\frac{\partial \mathbf{x}_k}{\partial u_j} = 0$. On the other hand, if $\mathbf{x}_k = 0$, then differentiating both sides of (12a) with respect to u_j gives,

$$0 = \frac{\partial}{\partial u_j}\left(\frac{\partial q}{\partial x_k}\mathbf{x}_k\right) = \frac{\partial^2 q}{\partial u_j \partial x_k}\mathbf{x}_k + \frac{\partial q}{\partial x_k}\frac{\partial \mathbf{x}_k}{\partial u_j} = \frac{\partial q}{\partial x_k}\frac{\partial \mathbf{x}_k}{\partial u_j}.$$

Similarly, each term in the second summation must be zero. Substituting (15a) for $s(u)$, we obtain the formula

$$\nabla f(u) = \nabla_u q(x(u), s(u), u) = -(Ax(u) + s(u) - u) = -\max(Ax(u) - u, 0). \qquad \square$$

Strict complementarity is sufficient, but not necessary, for the desirable property of differentiability. Although degenerate behavior in x is difficult to predict, Lemma 1 provides a simple characterization of strict complementarity in s. This means that (14b) is violated in component i exactly when $[Ax(u)]_i = u_i$. Given the least-squares formulation, this degenerate case seems unlikely in practice.

It is important to emphasize that under any circumstance (even without Assumption 1) the vector $-\max(Ax(u) - u, 0)$ can always be computed once a solution $x(u)$ to the convex optimization problem $\mathcal{Q}(u)$ is obtained. In the worse case, this gives a conservative estimate for the sensitivity $\sigma(u)$ (see Theorem 2), and actually is the gradient $\nabla f(u)$ whenever it exists. For convenience, in the remainder of this report, the quantity $-\max(Ax(u) - u, 0)$ will be called the sensitivity.

3. A greedy algorithm

Summarizing from above, it is possible to write least-squares formulation (8):

$$\min_{u \in \mathcal{D}_v} f(u) := q(x(u), s(u), u),$$

where $\mathcal{D}_v$ is the set of dose distributions that satisfy all the DVCs for a given problem, and $(x(u), s(u))$ solves the subproblem $\mathcal{Q}(u)$,

$$(x(u), s(u)) = \arg\min_{(x,s)\geqslant 0} q(x, s, u).$$

As has been mentioned, the set $\mathcal{D}_v$ is non-convex. Therefore, global minimization of f in $\mathcal{D}_v$ is generally intractable. A realistic goal is to find a good local minimum in a reasonable amount

of time. To this end, consider a simple algorithm framework that exploits the monotonicity of the function f established in Theorem 1. In this framework, it is possible to decrease f along a sequence of increasing dose bounds $\{u^k \in \mathscr{D}_v\}$ so that

$$u^0 \leqslant u^1 \leqslant u^2 \leqslant \cdots \Longrightarrow f(u^0) \geqslant f(u^1) \geqslant f(u^2) \geqslant \cdots$$

This general framework can have different "relaxation schemes" to generate increasing dose bound sequences $\{u^k\} \subset \mathscr{D}_v$, resulting in different approximate solutions to (8).

3.1. A sensitivity-driven greedy algorithm

In particular, the experimental work reported here uses a relaxation scheme based on the sensitivity of f. The resulting algorithm is called a *Sensitivity-Driven Greedy* (*SDG*) algorithm.

Algorithm 1 (*Sensitivity-Driven Greedy* (*SDG*) *Algorithm*).

– **Inputs:** Initial dose bound $u^0 \in \mathscr{D}_v$.

– **Output:** Beamlet intensities $x(u^k)$.

for $k = 0, 1, 2, \ldots$

1. Solve $\mathscr{Q}(u^k)$ for $x(u^k)$.
2. Compute $\nabla f(u^k) = -\max(Ax(u^k) - u^k, 0)$.
3. If stopping criteria are met, output $x(u^k)$ and stop.
4. Set $u^{k+1} = \mathrm{Proj}_{\mathscr{D}_v^k}(u^k - \nabla f(u^k))$ where $\mathscr{D}_v^k = \{u : u \geqslant u^k\} \cap \mathscr{D}_v$.

end

Note that in Step 4 the vector inside the projection operator is

$$u^k - \nabla f(u^k) \equiv \max(u^k, Ax(u^k)) \geqslant u^k.$$

Hence, a dose bound u_i^k is replaced by the calculated dose value $[Ax(u^k)]_i$ whenever the latter is greater. The resulting larger vector is then projected onto the set $\mathscr{D}_v^k$ to obtain the next dose bound u^{k+1}. Since the set $\mathscr{D}_v$ is non-convex, so is $\mathscr{D}_v^k$. Thus, the projection operation in Step 4 calls for an explanation, which is provided later in this report.

The bulk of the computation in this framework is to solve the subproblem $\mathscr{Q}(u^k)$ in Step 1 at each iteration which is a convex quadratic program known as a non-negative least-squares (NNLS) problem. Given their relative large sizes in IMRT applications, a fast algorithm for solving these NNLS problems is of primary importance. In the implementation discussed here, an interior-point gradient algorithm that was originally designed to strike a balance between reasonable accuracy and efficiency in this application [15] is used. For a comparison of this solver's performance versus some other leading algorithms, see [14].

For given stopping criteria, the output of SDG algorithm depends solely on the choice of the initial $u^0 \in \mathscr{D}_v$. The choice for u^0 used in the present work is the prescribed dose bounds at the lowest level DVCs. For example, in the case of 2-level DVCs for the right lung, which are "*no more than 30% volume of the right lung should receive a dose of 20 Gy or higher*" and "*no more than 40% volume of the right lung should receive 10 Gy or higher*," $u^0 = 10$ Gy would be set for all voxels in the right lung.

It is worth noting that although initial guesses for the beamlet intensities are needed to start solving the subproblem $\mathscr{Q}(u^k)$ at each iteration, they have no theoretical bearing on the final outcome of the algorithm.

3.2. Projection operations

Despite $\mathscr{D}_v$ being non-convex, projection onto it is straightforward. For example, suppose $\mathscr{D}_v \subset \mathbb{R}^{10}_+$ describes the dose-volume constraint: at least 70% of the voxels must have dose values no more than 5 Gy (i.e. only 3 of the 10 components can have values greater than 5). Then

$$\mathrm{Proj}_{\mathscr{D}_v}[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]^{\mathrm{T}} = [1, 2, 3, 4, 5, 5, 5, 8, 9, 10]^{\mathrm{T}}.$$

Here the projection sets the 7 smallest components equal to the minimum of their value and 5. Clearly, this is the closest point in $\mathscr{D}_v$ as it affects the least change on the original point in $\mathbb{R}^{10}_+$. Applying the projection only requires a sorting operation and can be done quickly. However, to define the projection uniquely it is necessary to employ an *a priori* tie-break rule. To break a tie, it is possible we may choose to give priority to the component with the higher indices (by not setting it to a smaller value), or to assign priority based on some available information such as relative distance to the tumor, tissue density, individual voxel weights, etc. In numerical computation, however, ties almost never occur and the effect of a tie-break rule is inconsequential.

The projection onto $\mathscr{D}^k_v$ is just as simple except for some additional bookkeeping. Continuing with the above example, let

$$\mathscr{D}^k_v := \left\{u \in \mathbb{R}^{10}_+ : u \geqslant u^k := [1, 2, 3, 4, 5, 6, 6, 5, 5, 5]^{\mathrm{T}}\right\} \cap \mathscr{D}_v.$$

Note that $u^k \in \mathscr{D}_v$ with $u^k_6 = u^k_7 = 6 > 5$. Then

$$\mathrm{Proj}_{\mathscr{D}^k_v}[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]^{\mathrm{T}} = [1, 2, 3, 4, 5, 6, 7, 5, 5, 10]^{\mathrm{T}}.$$

Again the projection sets the 7 smallest components equal to the minimum of their value and 5, but excludes the two (the 6th and 7th) corresponding to those in u^k whose values are greater than 5 to ensure that the resulting vector is component-wise greater than or equal to u^k. Should u^k already have 3 components greater than 5, then the projection would return u^k unchanged since, being in $\mathscr{D}_v$, the remaining seven components of u^k must be less than or equal to 5.

3.3. Convergence of the SDG algorithm

Recall that solving the subproblem $\mathscr{Q}(u)$ means projecting a point in the prescription set $\mathscr{H}$ onto the physical set $\mathscr{K}$ (see the definitions (3) and (4)). Under this projection $\mathrm{Proj}_{\mathscr{K}}(\cdot)$, it happens that the image of u is $Ax(u) + s(u) \equiv u - \nabla f(u)$ (which follows from (15a)). Therefore, in view of the Step 4 of the SDG algorithm, it is possible to write

$$\begin{pmatrix} b_t \\ u^{k+1} \end{pmatrix} = \mathrm{Proj}_{\mathscr{H}^k}\left(\mathrm{Proj}_{\mathscr{K}}\begin{pmatrix} b_t \\ u^k \end{pmatrix}\right), \quad \mathscr{H}^k := \{b_t\} \times \mathscr{D}^k_v. \tag{25}$$

The set $\mathscr{D}_v$, thus the augmented set $\mathscr{H}$ defined in (3), is a union of a large (but nevertheless finite) number of convex "branches," each containing a local minimum – a point closest to the physical set $\mathscr{K}$ defined in (4). The monotonicity of the iterates $\{u^k\}$ guarantees that, without stopping, they will eventually enter and stay in a fixed branch. Afterwards, the SDG algorithm reduces to the alternating projection algorithm between two convex sets (a fixed branch of $\mathscr{H}$ and $\mathscr{K}$), as is indicated by (25). Since the convergence of the alternating projection is well known (see, for example [1]), the following convergence result is obtained.

Theorem 4. *Under Assumption* 1, *the Sensitivity-Driven Greedy Algorithm, without stopping, generates an infinite sequence* $\{u^k\}$ *that converges to a local minimum of* f *in* $\mathscr{D}_v$.

However, since most of the local minima of f in $\mathscr{D}_v$ have little or no clinical relevance, this convergence property has limited practical value. Interested readers are referred to [14] for a detailed proof. The clinical relevance of the SDG algorithm ultimately lies in the effectiveness of the relaxation scheme to select a "good" branch (or a "good" set of voxels to sacrifice). This effectiveness can only be verified through experiments.

4. Numerical comparison

This section compares the model developed here with a dose-volume-based least-squares model implemented in the popular commercial planning system Pinnacle^{3}® RayOptimizer [11,12]. Without access to the commercial system, we have implemented a model that we call a Pinnacle-like (PL) formulation, which is our best attempt to duplicate what is implemented in Pinnacle^{3} RayOptimizer based on publicly available information.

4.1. A pinnacle-like approach

The implemented Pinnacle-like (PL) model takes the same general form of weighted least-squares models

$$\min_{x\geqslant 0} p(x) := \sum_{k=1}^{N} w_k p_k(d(x)), \tag{26}$$

where N is the number of planning structures, $d(x) = Ax$ and each structure k has its own weight w_k (importance factor) and penalty function p_k. We will denote the set of voxel indices in the kth structure by V_k with cardinality $|V_k|$. At the individual structure level, there are two forms of penalty functions, one for critical (healthy) structures and one for target (tumor) structures.

Assume that critical structure k has the dose-volume constraints "*no more than η_j^k% of the voxels can have doses above b_j^k*" for $j = 1, 2, \ldots, c_k$, where c_k is the number of DVCs for critical structure k. Then the penalty function for this critical structure is of the form

$$p_k(d) = \frac{1}{|V_k|} \sum_{j=1}^{c_k} \sum_{i\in V_k} H(d_i - b_j^k) H(d[\eta_j^k] - d_i) \left(\frac{d_i - b_j^k}{b_j^k} \right)^2, \tag{27}$$

where $d[\eta_j^k]$ is the current dose value received at the η_j^k dose-volume level (i.e. at $d = d(x)$, η_j^k% of voxels receive a dose above $d[\eta_j^k]$) and H is the heaviside function $H(y) = \max(y, 0)$. This function only penalizes voxels with dose values between d_j^k and $d[\eta_j^k]$.

Assume that target structure k has a prescribed dose b^k and a dose upper bound $b_{\max}^k$. Then the penalty function for this target structure is of the form

$$p_k(d) = \frac{1}{|V_k|} \sum_{i\in V_k} [H(b^k - d_i) + H(d_i - b_{\max}^k)] \left(\frac{d_i - \delta^k}{\delta^k} \right)^2, \tag{28}$$

where $\delta^k = (b^k + b_{\max}^k)/2$ is our target fit value. Note that at any voxel only one of the above two heaviside functions can be non-zero. It is unclear in the open literature what is the exact form of the target penalty functions used in RayOptimizer, so (28) may not be precisely what is implemented inside that software, but is in the same spirit as its critical structure objective definitions. Namely, penalization by relative deviation from δ^k occurs whenever dose is below the prescribed dose d^k

or above the maximum dose $b^k_{\max}$. We have chosen this particular value of δ^k because we feel it should lead to good target dose distributions and make this formulation more comparable to our own (this form of δ is the same value which we use in our SDG implementation for b_t). With this exception, we have attempted to stay as close as possible to the Pinnacle^3 formulation as described in [11,12].

It is well known that the dose-volume-based penalty functions of the form (27) are both non-convex and non-differentiable. The non-convexity makes starting point selection an issue because different starting points can lead to different solutions. The non-differentiability can be a potential source of bad numerical behavior for optimization algorithms designed for smooth functions such as the Fortran solver NPSOL [7] employed by Pinnacle$^{3®}$ RayOptimizer [11].

4.2. Setup of experiments

Our numerical comparison has been carried out primarily under the Matlab®7 environment. While the SDG algorithm is implemented entirely in Matlab, the PL implementation uses, just like the commercial system Pinnacle$^{3®}$ RayOptimizer [11,12], the Fortran 77 package NPSOL® [7] as its optimization engine. We believe that the use of a strong Fortran solver should give PL a considerable advantage on execution speed.

The NPSOL documentation [7] indicates that it stops whenever the relative change in the objective function drops below a given tolerance. In all of our tests, we used the same stopping tolerance of 10^{-2} in both SDG and PL algorithms. All the numerical results have been produced in Matlab 7 on a Linux workstation with a 3.8 GHz Intel Xeon processor and 8Gb of memory.

The software environment used for this project is called CEER (Computational Environment for Radiotherapy Research) and is available for free download [6,5]. This is a Matlab implementation of the vital features we need for fluence optimization, namely clinical data interface, dose calculation, and visualization. In our experiments, we used the QIB dose calculation engine native to CERR to generate an influence matrix for each test case using nine equally-spaced beams.

Our numerical experiments have been conducted on eight clinical cases: 6 relatively large cases (2 esophagus and 4 lung cases) and 2 small ones (a head-and-neck and a prostate case). By the size of a case we refer to the number of voxels involved in the region of treatment. More details on the test data have been included in Appendix A.

For each test case and each model, we settled at a computed solution corresponding to a set of tuned weights that we considered to be the best seen after an extensive trial-and-error process. The procedure for selecting weights always started with tuning weights for our model first, then we used the best weights found for our model as the "initial guesses" for the PL model. In some cases, the final choices are the same for both, or very similar. In others, they differ greatly.

4.3. Summary of results

Due to the aforementioned multi-objectiveness and the vagueness of clinical objectives, comparing the quality of IMRT treatment plans is a job for qualified medical specialists. However, we can still make assessments based on dose-volume constraint compliance.

In two test cases, the lung B case and the prostate case, the SDG solutions were clearly better than those of PL (more information on these two cases is given in Appendix A). On

Table 1
Run time comparison (in CPU seconds)

Case name	PL time	SDG time	PL/SDG ratio
Esophagus A	1510	244	6.2
Esophagus B	1643	287	5.7
Lung A	1536	169	9.1
Lung B	112*	118	0.95*
Lung C	954	90	10.6
Lung D	2237	178	12.6
Headneck	80	93	0.86
Prostate	65	31	2.1

the other hand, in the two esophagus cases the SDG solutions violated the dose-volume constraints more significantly than the PL did. For the rest of the four cases, the SDG solutions appeared either comparable or marginally better than the PL ones. Overall, we can only conclude that in terms of DVC compliance, the two methods have generated clinically relevant plans of competitive qualities over the eight tested cases, though none consistently outperformed the other.

The reason for the lesser performance of SDG in the two esophagus cases appeared to be that the relaxation scheme in use was a bit too greedy. There is certainly room for further research on relaxation strategies in our approach.

Another important aspect of the comparison is the speed of planning. In Table 1, we list run times in CPU seconds for each case. Note that these are planning times for the best importance factors found, not the total time spent during the trial-and-error process. Timing information on the head-and-neck and the prostate cases in Table 1 is less meaningful because these are two very small cases with fewer voxels. On the six larger (esophagus and lung) cases, the average improvement in treatment planning time of SDG over PL is greater than 7.5. We consider this to be a significant improvement, especially in view of the fact that PL used the commercial Fortran 77 package NPSOL as its optimization engine, while SDG is programmed entirely in Matlab.

The SDG time improvements are even more promising if we take into account the fact that during our trial-and-error process of importance factor selection, the SDG times were much more stable than those of PL. For example, on Lung B, one choice of weights for PL led to a solution that bore a striking resemblance in quality to the final one we have reported, but for which PL took 17 minutes, not 2 minutes as reported in the table (thus the asterisk). Similar behavior of wildly varying times was observed for Esophagus B and Lung D as well. In this context, the run time for the best weights found is really not as indicative as the total elapsed time spent on tuning to find a good plan. Our numerical experiments suggest that our algorithm be not only faster but also more reliable than PL given the trial-and-error environment.

In summary, at this point the primary advantages of the SDG algorithm over PL appear to be speed and stability. SDG demonstrated substantial speedups over PL in almost all cases despite the fact that PL used a well-established commercial Fortran package as its optimization solver. The reported times did not include the many hours (in some cases) spent in tuning weights, nor did they indicate the erratic behavior of PL's running times. We submit that the time spent tuning PL weights with PL runs was greatly reduced by starting with the tuned SDG weights

since it is much faster per run than PL. Based on this fact alone, we believe that SDG can be useful in improving computation time for the trial-and-error process in IMRT treatment planning.

5. Conclusions

From a theoretical perspective, the use of PL-type models leaves much to be desired. As presented in this paper, the issue of non-differentiability is problematic because optimization software designed for differentiable functions, such as NPSOL, is routinely applied to non-differentiable PL-type models. The frequency with which abnormal terminations of NPSOL are encountered hints at a volatility to PL-type models that should be at least partly attributable to non-differentiability. In contrast, the objective function in our formulation is continuously differentiable under a mild condition. (In fact, it is well defined in any circumstance with or without differentiability.)

We claim that from a practical perspective, the inherent non-convexity and nonlinearity present in dose-volume-based models is more pronounced in the PL approach than in the approach presented here. Several of our PL experiments gave unacceptable results due to declared "convergence" by NPSOL to some spurious (possibly stationary) points, and it was necessary to restart from different initial guesses. In practice, this amounts to another level of "tuning" that must be done with PL to find suitable starting beamlet intensities. The SDG method, on the other hand, only requires solving well-behaved convex quadratic subproblems at every iteration and is therefore insensitive to the initial guesses for beamlet intensities.

The speed and stability advantages exhibited by the approach developed in this paper could be attributed to two desirable properties: (a) the model has a monotone objective function, and (b) the algorithm only requires solving convex quadratic programs. This combination enables the algorithm to alleviate the non-convexity problem inherent in dose-volume-based fluence optimization.

In conclusion, the new dose-volume-based, least-squares approach presented here has demonstrated a promising potential as a practical tool for IMRT treatment planning. It strikes a critical balance between the computational tractability needed in this application and the theoretical rigor lacking from some existing dose-volume-based least-squares models.

Acknowledgments

The authors would like to thank Dr. Joseph Deasy at Washington University in St. Louis for his generous help with the CERR environment, and Dr. Thomas Guerrero at the M.D. Anderson Cancer Center for providing data and help with clinical test cases.

Appendix A. Details on data and results

Appendix A.1. Clinical data description

The data set consists of eight clinical cases: 2 esophagus, 4 lung, a single head-and-neck, and one prostate. The esophagus and lung cases are courtesy of M.D. Anderson Cancer Center Thoracic Oncology Department, while the head–neck and prostate cases are distributed with CERR [5]. For the esophagus and lung cases, all the planning requirements have been

specified by the responsible radiation oncologist. The prescriptions for the head–neck and prostate cases, on the other hand, have been chosen in accordance with treatment guidelines laid out by physicians at Memorial Sloan-Kettering Cancer Center [13]. For the sake of space, we only present detailed prescription information and computational results for two cases: the lung B case and the prostate case. More details are contained in the second author's thesis [14].

For the lung B case, an upper bound of 45 Gy is put on the spinal cord. The heart has the DVC that no more than 40% of the volume can receive doses greater than 40 Gy, or in short "volume($\leqslant$40%) > 40 Gy". The esophagus has the DVC: "volume($\leqslant$50%) > 50 Gy." Moreover, the total lung has the DVCs: "volume($\leqslant$45%) > 10 Gy" and "volume($\leqslant$35%) > 20 Gy." Finally, at least 95% of the target volume must receive at least 70 Gy with an upper bound of 75 Gy.

For the prostate case, the planning target volume (PTV) has overlaps with both the bladder and the rectum. The bladder excluding the PTV volume has the DVC that "volume($\leqslant$53%) > 47 Gy". The rectum excluding the PTV volume has the DVCs that "volume($\leqslant$53%) > 47 Gy" and "volume($\leqslant$30%) > 75.6 Gy". In addition, all voxels are limited by an upper bound of 100 Gy to prevent hot spots. At least 95% of the target must receive at least 77 Gy with an upper bound of 90 Gy.

Appendix A.2. Results on lung B and prostate cases

In IMRT, the DVC compliance is visualized by dose-volume histograms (DVHs), where the x-axis represents dose values and y-axis represent accumulated volume percentage. In the DVH, each planning structure has a corresponding curve. For example, the point (30, 50) on the curve for esophagus means that 50% of esophagus voxels have dose values 30 Gy or higher. The ideal curve for a target structure is a step function dropping from 100 to zero at the prescribed dose

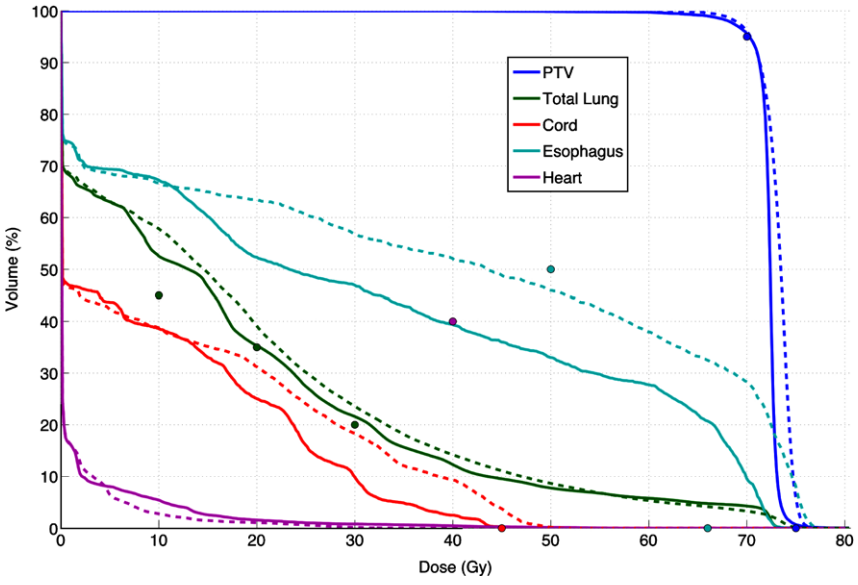

Fig. 3. Lung B – DVH (SDG solid/PL dashed).

value. For a healthy structure, the lower the curve is, the better. Another common tool for plan evaluation is contours of dose values called isodose lines.

The DVHs for the Lung B case are given in Fig. 3. For the spinal cord and esophagus, the DVH curves for SDG (solid lines) are much lower than those for PL (dashed lines), respectively. The corresponding isodose lines are shown in Fig. 4.

Structure overlap occurs in the prostate case. For the optimization process, the voxels shared by the rectum and the bladder with the planning target volume have been removed from these two organs. This approach allows an emphasis on critical structure sparing over target dose coverage. The DVHs in Fig. 5 suggest a much better SDG solution compared to the PL solution. Notice that the target dose coverage is much more homogeneous with SDG. The corresponding isodose lines are shown in Fig. 6.

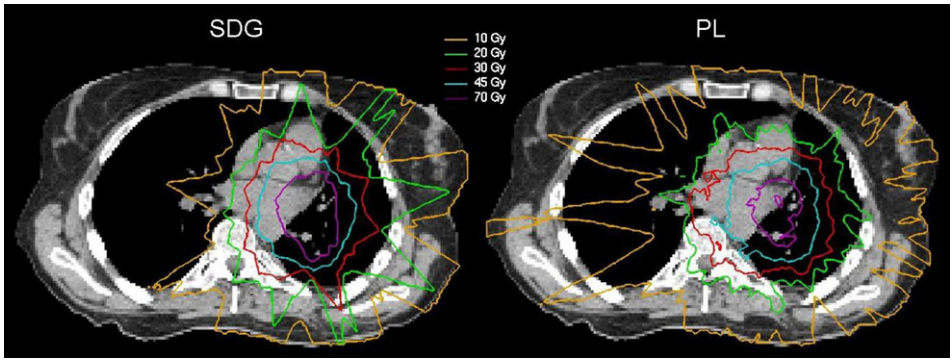

Fig. 4. Lung B – dose distributions.

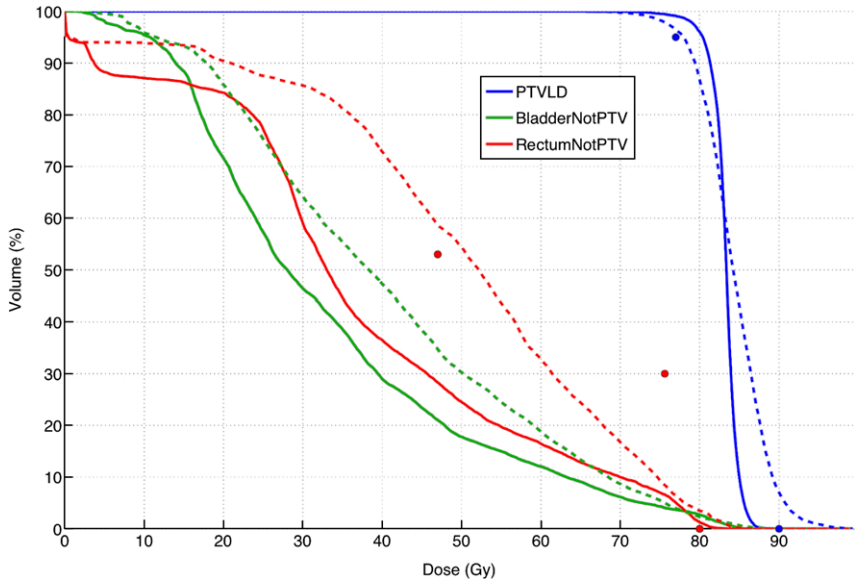

Fig. 5. Prostate – DVH (SDG solid/PL dashed).

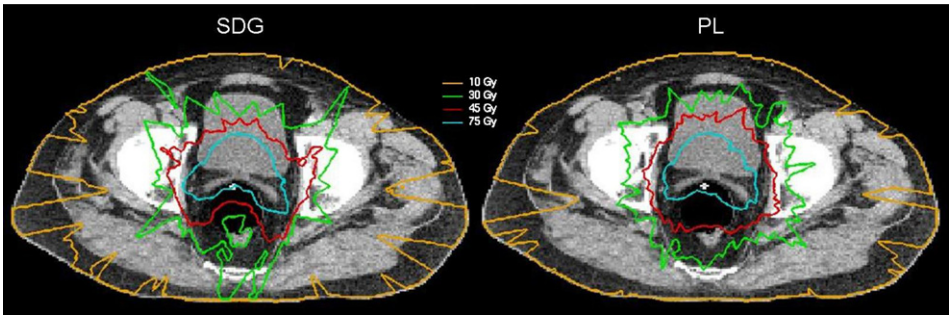

Fig. 6. Prostate – dose distributions.

References

[1] H. Bauschke, J. Borwein, On the convergence of von Neumann's alternating projection algorithm for two sets, Set-Valued Anal. 1 (1993) 185–212.
[2] T. Bortfeld, IMRT: a review and preview, Phys. Med. Biol. 51 (2006) R363–R379.
[3] T. Bortfeld, J. Bürkelbach, R. Boesecke, W. Schlegel, Methods of image reconstruction from projections applied to conformation radiotherap, Phys. Med. Biol. 35 (1990) 1423–1434.
[4] Y. Censor, Mathematical optimization for the inverse problem of intensity modulated radiation therapy, in: J.R. Palta, T.R. Mackie (Eds.), Intensity-Modulated Radiation Therapy: The State of The Art, Medical Physics Publishing, Madison, WI, USA, 2003, pp. 25–49.
[5] J. Deasy, Computational environment for radiotherapy research, Washington University in St. Louis, School of Medicine, 2006. Website: <http://radium.wustl.edu/CERR/>.
[6] J. Deasy, A. Blanco, V. Clark, CEER: a computational environment for radiotherapy research, Med Phys. 30 (2004) 979–985.
[7] P. Gill, W. Murray, M. Saunders, M. Wright, User's guide for NPSOL 5.0: a Fortran package for nonlinear programming, Stanford University, California, System Optimization Lab, NA 98-2, 1998.
[8] M. Langer, E. Lee, J. Deasy, R. Rardin, J. Deye, Operations research applied to radiotherapy, an NCI-NSF-sponsored workshop February 7–9, 2002, Int. J. Radiat. Oncol. Biol. Phys. 57 (2003) 762–768.
[9] E.W. Larsen, Tutorial: The nature of transport calculations used in radiation oncology, Transport Theory and Statistical Phys. 26 (1997) 739–764.
[10] E. Lee, T. Fox, I. Crocker, Integer programming applied to intensity-modulated radiation treatment planning optimization, Ann. Oper. Res. 119 (2003) 165–181.
[11] J. Löf, H. Rehbinder, Inverse planning optimization with RayOptimizer in Pinnacle^3, RaySearch Laboratories AB, Stockholm, Sweden, 2002.
[12] J. Löf, H. Rehbinder, T. McNutt, S. Johnson, Pinnical$^{3®}$ While Paper: P^{3}IMRT Inverse planning optimization, Philips Medical Systems, Global Information Center, I.B.R.S./C.C.R.I. Numro 11088, 5600 VC Eindhoven, Pays-Bas, The Netherlands.
[13] Memorial Sloan-Kettering Cancer Center, A Practical Guide to Intensity-Modulated Radiation Therapy, Medical Physics Publishing Corp., Madison, WI, 2003.
[14] M. Merritt, A sensitivity-driven greedy approach to fluence map optimization in intensity-modulated radiation therapy, Ph.D. thesis, Department of Computational and Applied Mathematics, Rice University, 2006.
[15] M. Merritt, Y. Zhang, An interior-point gradient method for large-scale totally nonnegative least squares problems, J. Optim. Theory Appl. 126 (2005) 191–202.
[16] J.R. Palta, T.R. Mackie (Eds.), Intensity-Modulated Radiation Therapy: The State of the Art, American Association of Physicists in Medicine, Medical Physics Monograph, vol. 29, Medical Physics Publishing, Madison, WI, 2003.
[17] F. Preciado-Walters, R. Rardin, M. Langer, V. Thai, A coupled column generation, mixed integer approach to optimal planning of intensity modulated radiation therapy for cancer, Math. Program. 101 (2004) 319–338.

[18] D. Shepard, M. Ferris, G. Olivera, T. Mackie, Optimizing the delivery of radiation therapy to cancer patients, SIAM Rev. 41 (1999) 721–744.
[19] S. Webb, Optimisation of conformal radiotherapy dose distribution by simulated annealing, Phys. Med. Biol. 34 (1989) 1349–1370.
[20] Y. Zhang, M. Merritt, A geometric approach to fluence map optimization in IMRT cancer treatment planning, in: W. Hager, S. Huang, P. Pardalos, O. Prokopyev (Eds.), Multiscale Optimization Methods and Applications, Springer, 2005, pp. 205–228.