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

Yin Zhang* and Michael Merritt†

Technical Report TR06-11
Department of Computational and Applied Mathematics
Rice University, Houston, Texas, 77005, U.S.A.

August 24, 2006

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 makes it problematic that 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, with speedups above one-order of magnitude for some large-scale problems.

Key words: IMRT treatment planning, Fluence optimization, Weighted least-squares model, Dose-volume constraint, Differentiability.

*This author's work was supported in part by NSF Grants DMS-0442065 and DMS-0405831.
†This author's work was supported in part by NSF Grant DMS-0240058.

Contents

1 Introduction

We will only give a very brief introduction to a particular problem, the fluence optimization problem, in intensity-modulated radiation therapy (IMRT). For more comprehensive information on this topic, we refer interested readers 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

In the usual practice of IMRT, a linear accelerator rotates on a gantry around the patient, emitting "modulated" beams of X-rays from a number of pre-fixed angles, where modulation mechanism is achieved by a multi-leaf collimator (MLC) attached to the head of the linear accelerator (see Figure 1). The MLC shapes the pattern of outgoing radiation beam, through a sequence of movements of its metal leaves, in order to precisely target the tumors while minimizing exposure of the neighboring healthy structures.

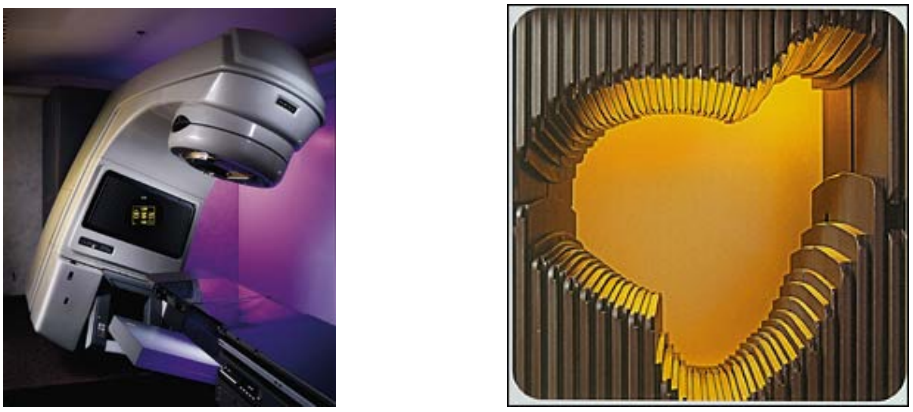

Figure 1: Linear Accelerator (left) and Multi-leaf Collimator (right)

Three major optimization problems arise in the IMRT planning process. The first is the *beam-angle optimization* problem which is to determine the "optimal" number and values of gantry angles, which is often formulated as a combinatorial optimization problem. The second is called *fluence (map) optimization* problem which is to find a set of "optimal" intensity profiles corresponding to the given set of beam angles. The third is the *leaf sequencing* problem which is to determine an "optimal" sequence of MLC leaf movements that delivers the intensity profile for each beam angle. In this paper, however, we will exclusively concentrate on the second problem – the fluence optimization problem.

Specifically, we need to determine how intense the X-ray beam should be at each point (x, y) on the MLC aperture surface for all gantry angles. These beam intensity profiles, or fluence maps, are represented by two-dimensional, nonnegative functions $I_a(x, y)$ for $a = 1, 2, \cdots, k$, where k is the number of gantry angles in use. See Figure 2 where five beam angles are in use, 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 as possible. These goals are fundamentally conflicting and have ill-defined evaluation criteria, making fluence optimization a difficult problem from a modeling viewpoint.

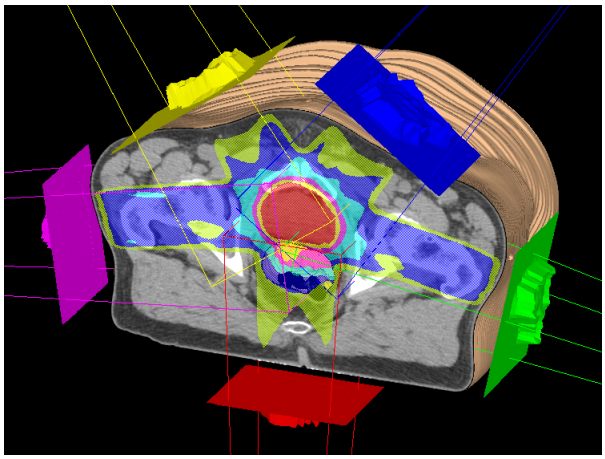

Figure 2: A five-beam treatment scheme

In practice, each fluence map function $I_a(x, y)$ is discretized at a rectangular grid on the MLC surface 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 $\{I_a(x_i, y_j)\}$, $a = 1, 2, \cdots, k$, 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, which naturally leads to the issue of how a dose distribution is calculated or prescribed.

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. 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 i-th voxel in the region of treatment is the linear model $d = Ax$, or

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

where a_{ij} represents the amount of dose absorbed by the i-th voxel per unit intensity emission from the j-th 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. Radiation absorption as a function of radiation intensities can be modeled with Boltzmann transport equations [9], which are difficult to solve. 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 and expensive. Many other less expensive, and less accurate, dose calculation engines exist. Currently, dose calculation is still considered an important research area in its own right. 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 nonzero amount of radiation and every beamlet influences at least one voxel's dose. These conditions can be easily met by pre-processing, if necessary. 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 nonnegative. Depending on how much scattering is included, the influence matrix A can be very sparse or fairly dense.

Since it is generally impossible to avoid collateral damage in radiation treatment despite using multiple angles in an attempt to focus radiation on the targets, oncologists utilize so-called dose-volume constraints (DVCs) to prescribe and evaluate treatment plans. 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 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, oncologists may specify another level of constraint on the same organ, such as "*no more than 40% volume of the right lung should receive 10 Gy or higher*." These DVCs try to control the amount of collateral damage to the

right lung during the treatment so that it can still function adequately and eventually recover after the treatment. Although other metrics have been proposed, the metrics based on dose-volume constraints has become the *de facto* standard metric to prescribe and evaluate radiation treatment plans in clinical practice.

Clearly, DVCs provide much-needed flexibility for the escalation of tumor doses. On the other hand, they also introduce a high degree of complexity to the planning process. In the above example, for instance, which 30% of the right lung volume should be allowed to absorb more than 20 Gy? This brings 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, in this paper we are only concerned with a particular class – weighted least-squares models. In particular, we will briefly introduce dose-volume based least-squares models and and point out what we see as 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 on the market. In our view, the main advantage of these models is their "speed"; that is, they can be approximately solved relatively quickly.

It is perhaps widely-agreeable 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 tumors and critical organs, but also between different structures of the same kind. As is well known, despite years of research, multi-objective optimization remains one of the most difficult problems.

2. Vagueness of clinical objectives [8]: In many cases, there is hardly consensus among radiation oncologists on a set of definitive criteria that can be used to rank the quality of 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 tissues to irradiation (which may vary from one individual patient to another).

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

Under such an environment, plus additional resource considerations, a fast turn-around time for trial-and-error becomes a great advantage. This speed advantage is enjoyed by weighted least-squares models more than other models such as integer programming models (for example, see [10, 17]), which are more mathematically rigorous but more expensive to solve. In the meantime, 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 lead to the popularity of least-squares models in clinical practice.

In a 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 prescribed one for that structure. Suppose there are N structures each consisting of a set of voxels. Then 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, w_j is the weight for the j-th structure, and $f_j(d)$ is a piecewise quadratic function that penalizes the misfit of the calculated dose $d(x)$ to the prescribed one for the j-th structure. For any fixed set of weights, $f(x)$ can be approximately minimized by one of the existing algorithms for continuous optimization subject to nonnegativity 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 i-th voxel is either

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

where $d_i(x)$ is the calculated dose value at the i-th voxel and b_i is the prescribed one, then the corresponding penalty term is, respectively, 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. Therefore, without DVCs the function $f(x)$ in (2) is convex and differentiable, and the resulting optimization problem is simple at least from a theoretical point of view.

However, the addition of DVCs makes the matter more complicated. 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 penalize only those voxels where the calculated dose values, evaluated at the current beamlet intensity x, are higher than 20 Gy but not high enough to be ranked in the top 10%. In other words, voxels with the highest 10% dose values are not penalized.

The benefit of more flexibility provided by DVCs comes at a price. That is, the resulting penalty function $f(x)$ becomes non-convex and non-differentiable. To illustrate this point, consider a hypothetic 2-voxel structure with the DVC that *no more than 50% of the volume (i.e. 1 voxel) should receive a dose 5 Gy or higher*. In this case, the corresponding penalty function is

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

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, is it possible to avoid the non-differentiability? It is troublesome that optimization algorithms designed for differentiable functions are being routinely used in daily clinical practice to minimize non-differentiable functions. Our own experiments 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 $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n : x \geq 0\}$, 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 submatrix $A_t \in \mathbb{R}^{m_t \times n}$ 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 $\mathcal{D}_v \subset \mathbb{R}^m_+$ 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 $\mathcal{D}_v$ satisfy all such requirements.

This paper is organized as follows. In Section 2, we describe and study 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. Finally, we conclude the paper in Section 5.

2 A Least-Squares Model with Differentiability

Since the fundamental difficulties of the 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 improve this state of affairs. Specifically, we carry out an in-depth theoretical and numerical study on a new least-squares model that allows a differentiable objective function and faster numerical optimization.

We mention that fast fluence optimization is also essential for beam-angle optimization, which is still considered unsolved [2], since the latter needs to use the former as a necessary subroutine to be solved repeatedly.

2.1 Geometric Considerations

To motivate our approach, we discuss the geometry of the dose-volume constraint set in the IMRT fluence optimization problem. Fundamentally, our goal is to sufficiently dose the targets while satisfying the dose-volume constraints (DVCs) as closely as possible.

Suppose we have a vector $u \in \mathcal{D}_v$ satisfying the dose-volume constraints. Then to find beamlet intensities $x \in \mathbb{R}^n_+$ whose resulting doses satisfy the dose-volume constraints, we must have $Ax \leq u$. (For this reason, we will loosely refer to the u values as "bounds" throughout.) Moreover, we wish to meet the target prescription $A_t x = b_t$. Equivalently, we would like 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 $s \in \mathbb{R}^m_+$.

We define the *prescription set*

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

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

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

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 $\mathcal{K}$ is a closed convex cone and $\mathcal{H}$ is non-convex since the set $\mathcal{D}_v$, defined by DVCs of a combinatorial nature, is a non-convex set.

2.2 Problem Formulation

To treat the targets and yet make the most of our sacrifice in the critical structures, we would ideally like to find $x \in \mathbb{R}_+^n$ and $s \in \mathbb{R}_+^m$ such that both $A_t x = b_t$ and $Ax + s = u$, but most likely this is impossible. The reality of the IMRT fluence problem is that there may be no physically-achievable dose that both satisfies the DVCs and meets the prescription. That is, generally speaking $\text{dist}(\mathcal{H}, \mathcal{K}) > 0$, or equivalently $\mathcal{H} \cap \mathcal{K} = \emptyset$. Thus, we are motivated to determine $d_H \in \mathcal{H}$ and $d_K \in \mathcal{K}$ such that

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

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, we define our objective function over the set of all $u \in \mathcal{D}_v$ to be

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

Namely, $f(u)$ is itself the optimal value of a linear least-squares problem with a nonnegativity constraint. Using this notation, we re-pose the problem (5) with the equivalent formulation

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

If we can solve this problem, then we will have found a $u^* \in \mathcal{D}_v$ 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 $u^0 \in \mathcal{D}_v$, the model

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

is nothing but a regular least-squares model without dose-volume constraints. Therefore, if we start from $u = u^0$ and monotonically decrease $f(u)$, then we should obtain a better solution (in terms of stronger target fit 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 practically desirable to allow weightings in our least-squares formulation, in place of $q(x, s, u)$ we can easily consider

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

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, we will always assume that the weighting matrices have already been absorbed into the quantities A, A_t, b_t and $\mathcal{D}_v$, i.e., $A \leftarrow WA$, $A_t \leftarrow W_t A_t$, $b_t \leftarrow W_t b_t$, ..., etc. For this reason, from now on we will not explicitly mention weighting unless necessary.

2.3 Optimality of the Subproblem

Our objective f is itself the minimum value of another optimization problem, making it somewhat complicated. In this section, we examine the properties of f on the domain $\mathcal{D}_v$ with a mind toward constructing an algorithm for solving our overall problem (8). We make the following assumption in all our theoretical results.

Assumption 1. *The matrix A_t is full column rank.*

In our experience, this assumption almost always holds and certainly does in all of our experiments. This is so partly because usually the number of target voxels far exceeds that of beamlets, hence A_t has far more rows than columns. Our theoretical work depends on uniqueness properties that follow from this assumption. In particular, we will use this assumption to prove the subproblem solutions are unique and continuously differentiable functions of u.

Evaluating $f(u)$ requires solving a subproblem parameterized by $u \in \mathcal{D}_v$ which we will denote $\mathcal{Q}(u)$:

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

Subproblem $\mathcal{Q}(u)$ is a convex bound-constrained quadratic program that we will need to solve repeatedly, so let us examine it 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 $\mathcal{Q}(u)$.

Assumption 1 implies this strong convexity property since the matrix $A_t^T A_t$ 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 $v, w \geq 0$. Using this notation, we can write the KKT conditions for $\mathcal{Q}(u)$ as

$$\begin{aligned} \min\left(x, \nabla_x q(x,s,u)\right) &= 0, & \text{(12a)} \\ \min\left(s, \nabla_s q(x,s,u)\right) &= 0, & \text{(12b)} \end{aligned}$$

where

$$\begin{aligned} \nabla_x q(x,s,u) &= A_t^T(A_t x - b_t) + A^T(Ax + s - u), & \text{(13a)} \\ \nabla_s q(x,s,u) &= Ax + s - u. & \text{(13b)} \end{aligned}$$

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, we say that strict complementarity holds for $\mathcal{Q}(u)$ at (x,s) if

$$\begin{aligned} x + \nabla_x q(x,s,u) &> 0, & \text{(14a)} \\ s + \nabla_s q(x,s,u) &> 0. & \text{(14b)} \end{aligned}$$

Lemma 1. *For any u,*

$$\begin{aligned} s(u) &= \max(0, u - Ax(u)), && \text{(15a)} \\ \nabla_s q(x(u), s(u), u) &= \max(Ax(u) - u, 0), && \text{(15b)} \end{aligned}$$

where the maximum is taken component-wise.

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

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

Definition 1. *Let the "sensitive" index set $S \equiv S(x, u) := \{i : (Ax)_i > u_i\}$ and the matrix $E(S) = \sum_{i \in S} e_i e_i^T$, where e_i is the i-th 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 $E(S)$, we can eliminate s and simplify the KKT conditions for $\mathcal{Q}(u)$.

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

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

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 we will see in the next section, these voxels are included in the set that the objective function f is sensitive to, a fact that will have significant algorithmic implications.

Finally, we give an explicit expression for the solution $x(u)$ of $\mathcal{Q}(u)$. Note that $s(u)$ can then be computed from (15a). We first partition the indices for $x(u)$:

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

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 nonzero elements of $x(u)$ are given by*

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

and at the same time satisfy

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

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

2.4 Sensitivity of the Objective Function

To solve the problem (8) we are interested in how f changes as we adjust u. Again, we have made Assumption 1 for all the following results.

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

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

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

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

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 we can always decease $f(u)$ 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 our 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 us information about how $f(u)$ changes locally as we increase u_i. Furthermore, whenever f is Gâteaux differentiable, $\sigma(u) = \nabla f(u)$. We will provide a mild condition under which f is indeed Fréchet differentiable. However, before we restrict ourselves, we examine a straightforward yet important consequence of the definition of $\sigma(u)$.

Theorem 2. *The sensitivity vector satisfies*

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

Proof. First consider any index i such that $[Ax(u)]_i \leq u_i$. 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 $t > 0$, $f(u + te_i) \equiv f(u)$. Hence, we have $\sigma(u)_i = 0$.

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

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

Therefore,

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

Let t go to zero and we obtain 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 our 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 we can say about the sensitivity $\sigma(u)$. So, we introduce a condition on the subproblem solutions that we will assume from now on.

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

Under this condition 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 let us write the complementarity equations in the KKT conditions (12) 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 $(\hat{x}, \hat{s}, \hat{u}) = (x(\hat{u}), s(\hat{u}), \hat{u})$. We will verify the invertibility of the Jacobian of K with respect to (x, s) at $(\hat{x}, \hat{s}, \hat{u})$. The Jacobian $K'_{(x,s)}(x, s, u)$ is

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

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

$$\begin{aligned} \text{Diag}(\nabla_x \hat{q})z_1 + \text{Diag}(\hat{x})\nabla_x^2 \hat{q} z_1 + \text{Diag}(\hat{x})A^T z_2 &= 0 \quad (21) \\ \text{Diag}(\hat{s})Az_1 + \text{Diag}(\hat{s} + \nabla_s \hat{q})z_2 &= 0, \quad (22) \end{aligned}$$

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, we can solve (22) for z_2:

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

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

$$\begin{aligned} D_1 &= \text{Diag}(\hat{x} + \nabla_x \hat{q})^{-1}\text{Diag}(\nabla_x \hat{q}) \\ D_2 &= \text{Diag}(\hat{x} + \nabla_x \hat{q})^{-1}\text{Diag}(\hat{x}) \\ M &= A_t^T A_t + A^T E(\hat{S})A. \end{aligned}$$

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 $\text{Diag}(\hat{x} + \nabla_x \hat{q})^{-1}$ we get

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

Note that D_1 and D_2 could be simplified by using strict complementarity in x. We can further simplify (23) 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 nonsingularity 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, we can conclude 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*

$$\sigma(u) = \nabla f(u) = -\max(Ax(u) - u, 0) \leq 0. \quad (24)$$

Proof. Observe from (11) that we can 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, we know 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 $s(u)$, we apply the chain rule to (11) to obtain the j-th 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. We claim that in the two summations, each term is zero. It suffices to consider the k-th 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,

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

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

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

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. We see that (14b) is violated in component i exactly when $[Ax(u)]_i = u_i$. Given our least-squares formulation, this degenerate case seems unlikely in practice.

We emphasize that under any circumstance (even without Assumption 1) the vector $-\max(Ax(u) - u, 0)$ can always be computed once we obtain a solution $x(u)$ to the convex optimization problem $\mathcal{Q}(u)$ which, in the worse case, 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, from now on we will always call $-\max(Ax(u) - u, 0)$ the sensitivity.

3 A Greedy Algorithm

Recall our 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 $\mathcal{Q}(u)$,

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

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 and find it quickly. To this end, we consider a simple algorithm framework that exploits the monotonicity of the function f established in Theorem 1. In this framework, we decrease f along a sequence of increasing dose bounds $\{u^k \in \mathcal{D}_v\}$ so that

$$u^0 \leq u^1 \leq u^2 \leq \cdots \implies f(u^0) \geq f(u^1) \geq f(u^2) \geq \cdots .$$

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

3.1 A Sensitivity-Driven Greedy Algorithm

In particular, we have experimented with a relaxation scheme based on the sensitivity of f. We will call the resulting algorithm a *Sensitivity-Driven Greedy (SDG)* algorithm.

Algorithm 1 (Sensitivity-Driven Greedy (SDG) Algorithm).
— **Inputs:** *Initial dose bound* $u^0 \in \mathcal{D}_v$.
— **Output:** *Beamlet intensities* $x(u^k)$.

for $k = 0, 1, 2, \ldots$
1. *Solve* $\mathcal{Q}(u^k)$ *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* $u^{k+1} = \text{Proj}_{\mathcal{D}_v^k}\left(u^k - \nabla f(u^k)\right)$ *where* $\mathcal{D}_v^k = \{u : u \geq u^k\} \cap \mathcal{D}_v$.

end

We note that in Step 4 the vector inside the projection operator is

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

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 $\mathcal{D}_v^k$ to obtain the next dose bound u^{k+1}. Since the set $\mathcal{D}_v$ is non-convex, so is $\mathcal{D}_v^k$. Thus, the projection operation in Step 4 calls for an explanation, which we will provide later.

The bulk of the computation in this framework is to solve the subproblem $\mathcal{Q}(u^k)$ 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. On the other hand, due to errors in leaf-sequencing, measurement, imaging, dose calculation, patient motion, etc., high accuracy solutions are not necessary. In our implementation, we use an interior-point gradient algorithm that was originally designed to strike a balance between reasonable accuracy and efficiency in this application [15]. 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 $u^0 \in \mathcal{D}_v$. Our choice for u^0 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*," we would set $u^0 = 10$ Gy 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 $\mathcal{Q}(u^k)$ at each iteration, they have no bearing whatsoever to the final outcome of the algorithm (at least in theory).

3.2 Projection Operations

Despite $\mathcal{D}_v$ being non-convex, projection onto it happens to be trivial. For example, suppose $\mathcal{D}_v \subset \mathbb{R}_+^{10}$ 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

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

Here the projection sets the 7 smallest components equal to the minimum of their value and 5. Clearly, this is the closest point in $\mathcal{D}_v$ 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 we do need an *a priori* tie-break rule. To break a tie, we may choose to give priority to the component with the higher indices (by not setting it to a smaller value), or we may 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 $\mathcal{D}_v^k$ is just as simple except for some additional bookkeeping. We continue with the above example. Let

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

We note that $u^k \in \mathcal{D}_v$ with $u_6^k = u_7^k = 6 > 5$. Then

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

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 $\mathcal{D}_v$, 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 $\mathcal{Q}(u)$ means projecting a point in the prescription set $\mathcal{H}$ onto the physical set $\mathcal{K}$ (see the definitions (3) and (4)). Under this projection $\text{Proj}_{\mathcal{K}}(\cdot)$, 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, we can write

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

The set $\mathcal{D}_v$, thus the augmented set $\mathcal{H}$ 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 $\mathcal{K}$ 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 $\mathcal{H}$ and $\mathcal{K}$), as is indicated by (25). Since the convergence of the alternating projection is well known (see, for example [1]), we have the following convergence result.

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 $\mathcal{D}_v$.*

However, since most of the local minima of f in $\mathcal{D}_v$ have little or no clinical relevance, this convergence property has limited practical value. We refer interested readers 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

We compare our model 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 known to us.

4.1 A Pinnacle-Like Approach

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

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

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 k-th 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 $b^k_{\max}$. Then the penalty function for this target structure is of the form

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

where $\delta^k = (b^k + b^k_{\max})/2$ is our target fit value. Note that at any voxel only one of the above two heaviside functions can be nonzero. 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.

We have used the freely-available Computational Environment for Radiotherapy Research (CERR) from Joe Deasy's group at Washington University in St. Louis

[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 the Appendix.

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 the Appendix). On 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 speedup of SDG over PL is greater

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

than 7.5. We consider this to be a significant speedup, 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 speedups 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 is 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 already be useful in speeding up the trial-and-error process in IMRT treatment planning.

5 Final Remarks

Although the wide-spread use of the PL-type models in clinical practice may be justified by their good track record, from a theoretical perspective, these models leave much to be desired. In our view, the issue of non-differentiability is the most disturbing because optimization software designed for differentiable functions, such as NPSOL, is routinely applied to non-differentiable PL-type models in today's clinical practice. Indeed, the frequency with which we encountered abnormal terminations of NPSOL 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 ours. Several of our PL experiments were frustrated due to declared "convergence" by NPSOL to some spurious (possibly stationary) points and we had 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.

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

In conclusion, our new dose-volume-based, least-squares approach 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.

Acknowledgment

Many thanks to Joe Deasy at Washington University in St. Louis for his generous help with the CERR environment, and to Thomas Guerrero at the M.D. Anderson Cancer Center for providing data and help with clinical test cases.

References

[1] H. Bauschke and J. Borwein. On the Convergence of von Neumann's Alternating Projection Algorithm for Two Sets," Set-Valued Analysis, 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 and 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 and T.R. Mackie (Eds.), Intensity-Modulated Radiation Therapy: The State of The Art. Medical Physics Publishing, Madison, Wisconsin, 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 and V. Clark. CEER: a computational environment for radiotherapy research. Med Phys., 30 (2004) 979-985.

[7] P. Gill, W. Murray, M. Saunders and 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 and 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, and I. Crocker. Integer Programming Applied to Intensity-Modulated Radiation Treatment Planning Optimization. Annals of Operations Research, 119 (2003) 165-181.

[11] J. Löf and H. Rehbinder. Inverse planning optimization with RayOptimizer in Pinnacle^3. RaySearch Laboratories AB, Stockholm, Sweden. 2002.

[12] J. Löf, H. Rehbinder, T. McNutt and 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. Numìro 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, Wisconsin, 2003.

[14] M. Merritt. A sensitivity-driven greedy approach to fluence map optimization in intensity-modulated radiation therapy. PhD thesis, Department of Computational and Applied Mathematics. Rice University. 2006.

[15] M. Merritt and Y. Zhang. An Interior-Point Gradient Method for Large-Scale Totally Nonnegative Least Squares Problems. Journal of Optimization Theory and Applications, 126 (2005) 191-202.

[16] J. R. Palta and T. R. Mackie (Eds). Intensity-Modulated Radiation Therapy: The State of the Art. American Association of Physicists in Medicine, Medical Physics Monograph No. 29, Medical Physics Publishing, Madison, Wisconsin, 2003.

[17] F. Preciado-Walters, R. Rardin, M. Langer and V. Thai. A coupled column generation, mixed integer approach to optimal planning of intensity modulated radiation therapy for cancer. Mathematical Programming, 101 (2004) 319-338.

[18] D. Shepard, M. Ferris, G. Olivera, and T. Mackie. Optimizing the delivery of radiation therapy to cancer patients, SIAM Review, 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 and M. Merritt. A Geometric Approach to Fluence Map Optimization in IMRT Cancer Treatment Planning. In: Multiscale optimization methods and applications, pp.205–228. W. Hager, S. Huang, P. Pardalos, O. Prokopyev (Eds.) Springer, 2005.

A Details on Data and Results

A.1 Clinical Data Description

The data set consists of eight clinical cases: 2 esophagus cases, 4 lung cases, a head and neck case and a prostate case. 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 an actual physician. 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($\leq 40\%$) $>$ 40 Gy". The esophagus has the DVC: "volume($\leq 50\%$) $>$ 50 Gy." Moreover, the total lung has the DVCs: "volume($\leq 45\%$) $>$ 10 Gy" and "volume($\leq 35\%$) $>$ 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($\leq 53\%$) $>$ 47 Gy". The rectum excluding the PTV volume has the DVCs that "volume($\leq 53\%$) $>$ 47 Gy" and "volume($\leq 30\%$) $>$ 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.

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 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 Figure 3, where PTV stands for Planned Target Volume. For the spinal cord and esophagus, the DVH curves for

SDG (solid lines) are much lower than those for PL (dashed lines), respectively. The isodose lines shown in Figure 4 seem to a give a consistent assessment.

Structure overlap occurs in the prostate case. We have chosen to remove the voxels shared by the rectum and bladder with the planned target volume from these two organs since we expected inevitable damage to those tissues. The DVHs in Figure 5 suggest a much better SDG solution on all fronts than that of PL. Also, the target dose is much more homogeneous with SDG. Again, the isodose lines in Figure 6 seem to confirm this assessment. Comparing the lower part of the two 45 Gy isodose lines, we see that the SDG line is conspicuously curved in to avoid the critical organ below.

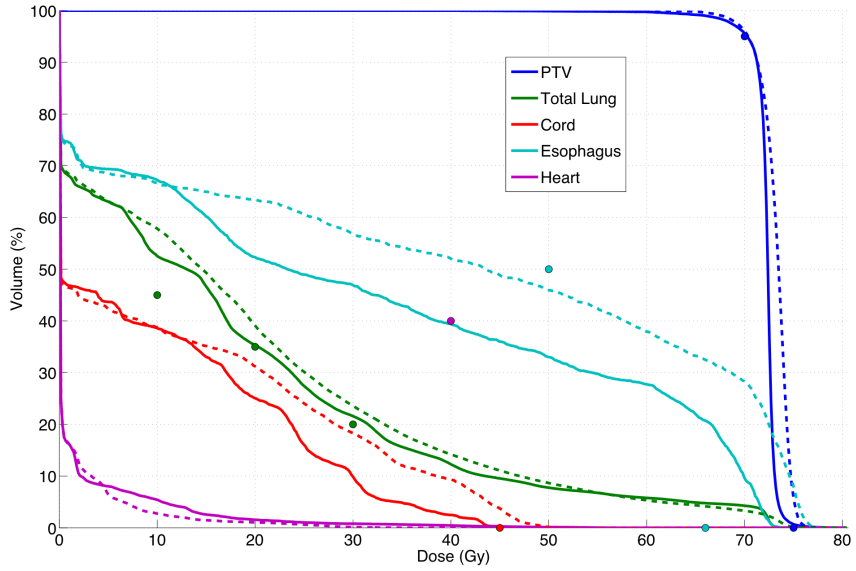

Figure 3: Lung B - DVH (SDG solid/PL dashed).

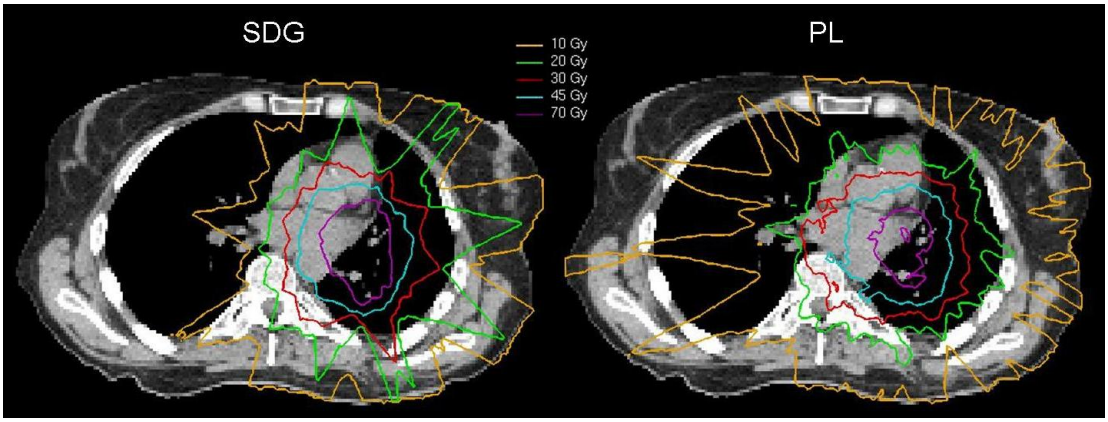

Figure 4: Lung B - Dose Distributions.

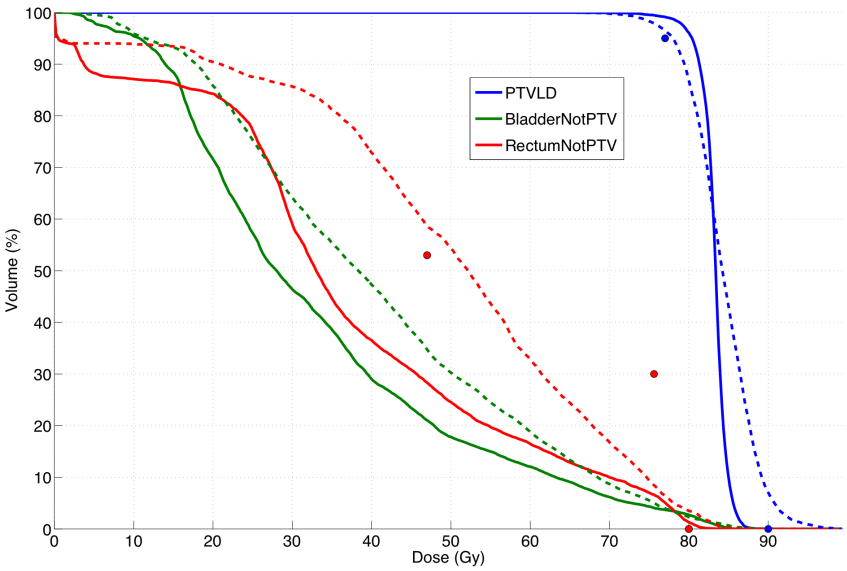

Figure 5: Prostate - DVH (SDG solid/PL dashed).

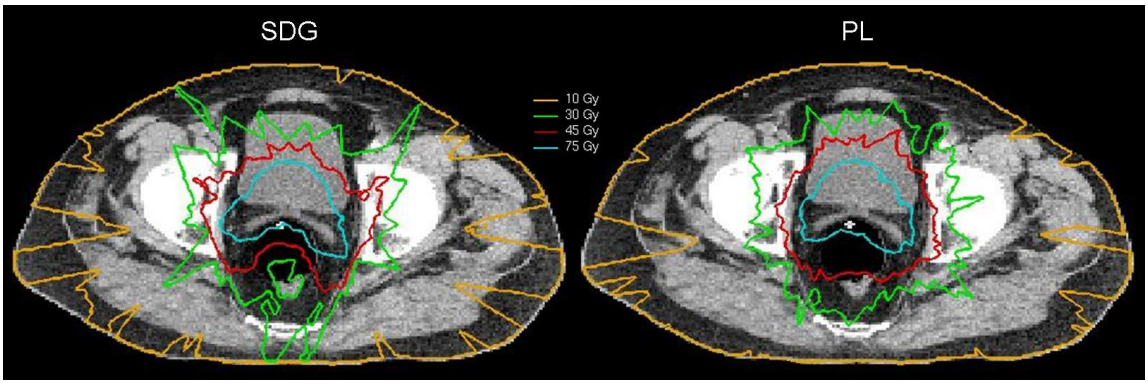

Figure 6: Prostate - Dose Distributions.