

Accounting for the tongue-and-groove effect using a robust direct aperture optimization approach

Ehsan Salari^{a)}

Department of Industrial and Systems Engineering, University of Florida, Gainesville, Florida 32611-6595

Chunhua Men

Department of Radiation Oncology, University of California San Diego, La Jolla, California 92037-0843

H. Edwin Romeijn

Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan 48109-2117

(Received 2 August 2010; revised 4 January 2011; accepted for publication 6 January 2011; published 14 February 2011)

Purpose: Traditionally, the tongue-and-groove effect due to the multileaf collimator architecture in intensity-modulated radiation therapy (IMRT) has typically been deferred to the leaf sequencing stage. The authors propose a new direct aperture optimization method for IMRT treatment planning that explicitly incorporates dose calculation inaccuracies due to the tongue-and-groove effect into the treatment plan optimization stage.

Methods: The authors avoid having to accurately estimate the dosimetric effects of the tongue-and-groove architecture by using lower and upper bounds on the dose distribution delivered to the patient. They then develop a model that yields a treatment plan that is robust with respect to the corresponding dose calculation inaccuracies.

Results: Tests on a set of ten clinical head-and-neck cancer cases demonstrate the effectiveness of the new method in developing robust treatment plans with tight dose distributions in targets and critical structures. This is contrasted with the very loose bounds on the dose distribution that are obtained by solving a traditional treatment plan optimization model that ignores tongue-and-groove effects in the treatment planning stage.

Conclusions: A robust direct aperture optimization approach is proposed to account for the dosimetric inaccuracies caused by the tongue-and-groove effect. The experiments validate the ability of the proposed approach in designing robust treatment plans regardless of the exact consequences of the tongue-and-groove architecture. © 2011 American Association of Physicists in Medicine. [DOI: [10.1118/1.3547722](https://doi.org/10.1118/1.3547722)]

Key words: IMRT treatment planning, aperture modulation, tongue-and-groove effect

I. INTRODUCTION

Intensity-modulated radiation therapy (IMRT) treatment planning is concerned with the design of a treatment plan for individual cancer patients. Such a treatment plan consists of a collection of apertures to be formed by a multileaf collimator (MLC) system along with associated intensities. Leaves in most commercial MLCs have a tongue-and-groove design that helps reduce interleaf leakage. However, the exposed leaf stepped sides (tongues) may undesirably block or scatter part of the radiation, resulting in underdosing of targets. It has been clinically indicated that the tongue-and-groove effect may cause underdosing as large as 10%–25% (see Refs. 1–8). On the other hand, accurately estimating the dosimetric effects of the tongue-and-groove architecture is a difficult task, especially with the commonly used beamlet-based dose models. The goal of this paper is to develop, implement, and test a *robust* method that takes dosimetric inaccuracies with respect to the MLC architecture into account explicitly. We will achieve this by employing upper and lower bounds on the dose distribution delivered to the patient and tailoring the optimization model to find a treat-

ment plan that is clinically attractive with respect to these bounds, i.e., one that is of high-quality *regardless* of the exact effect of the MLC tongues. By testing our approach on ten clinical cases of head-and-neck cancer, we show that our approach is successful in the sense that tight dose distribution bounds can be achieved even under coarse and easily obtainable bounds on the dosimetric effects. In contrast, a treatment plan optimization approach that does not take the dosimetric inaccuracies into account yields dose distribution bounds that are loose, making it hard to accurately assess the treatment plan quality.

IMRT treatment plan optimization is traditionally performed in two sequential stages: (i) Fluence map optimization (FMO) and (ii) leaf sequencing (LS). In particular, each orientation of the accelerator head defines a rectangular beam, each of which is conceptually discretized into a set of beamlets. The FMO problem then involves determining the optimal intensities for all beamlets. Given the optimal fluence map, the LS problem decomposes the fluence map for each beam into a manageable set of deliverable apertures. Both the FMO and the LS problems have been extensively

studied in literature; for modeling and solution approaches to FMO, we refer to the review papers by Shepard *et al.*⁹ and Romeijn and Dempsey.¹⁰ More specifically, Lee *et al.*^{11,12} studied mixed integer programming approaches; Hamacher and Küfer¹³ and Küfer *et al.*¹⁴ proposed a multicriteria approach to the problem; and Romeijn *et al.*^{15,16} developed convex programming models. Formally, the LS problem is to determine a set of deliverable apertures for delivering a fluence map that is optimal with respect to, typically, beam-on time, number of apertures used, or total treatment time. If the objective is to minimize beam-on time and any row-convex aperture is deliverable, the LS problem is efficiently solvable (see, e.g., Refs. 17–20). In addition, Baatar *et al.*,²¹ Boland *et al.*,²² Kamath *et al.*,²³ Dai and Hu,²⁴ and Siochi²⁰ studied the problem under additional MLC hardware constraints. In contrast, as Baatar *et al.*²¹ showed, the problem of decomposing a fluence map into the minimum number of apertures is NP-hard. This has led to the development of a large number of heuristics for solving this problem, such as in Refs. 20, 21, 25, and 26. Additionally, Engel,²⁷ Kalinowski,^{28,29} and Lim and Choi³⁰ developed heuristics to minimize the number of apertures while constraining the total beam-on time to be minimal. Finally, Taşkın *et al.*³¹ proposed an integer programming approach to minimize the total treatment time. The tongue-and-groove effect has been widely addressed in literature in the LS stage. In particular, LS algorithms are divided into two categories: Dynamic and static delivery; in dynamic delivery, the radiation is on when the MLC leaves are in motion, whereas in static delivery, also called the step-and-shoot, radiation is off when the leaves are in motion and it is turned on once the leaves are repositioned. Van Santfoort and Heijmen³² and Webb *et al.*³³ proposed LS algorithms that reduce the tongue-and-groove effect for dynamic delivery, and Que *et al.*³⁴ and Kamath *et al.*³⁵ incorporated the tongue-and-groove effect in the LS problem for the static delivery. Finally, Kamath *et al.*³⁶ compared different LS algorithms for static delivery with respect to the tongue-and-groove effect as well as the total treatment time.

A major issue with the traditional two-stage method is that the dose delivered to a patient depends not only on the fluence maps but also on the actual shape of the apertures used. To address this issue, an integrated approach to the FMO and LS problems, usually referred to as aperture modulation or direct aperture optimization (DAO), has been proposed. DAO explicitly solves for aperture shapes and intensities rather than beamlet intensities (see, e.g., Refs. 37–41). In contrast to the traditional method, DAO explicitly incorporates the shape of the apertures while optimizing for the aperture intensities. In general, treatment plans obtained with DAO use much fewer apertures (in fact, this is considered to be one of the attractive side effects of DAO since it reduces treatment times and beam-on times). Therefore, one could argue that the tongue-and-groove error is more significant when using DAO since much fewer apertures are used. However, the smaller number of apertures also implies that DAO fluence maps tend to be much smoother than the ones obtained with traditional two-stage approaches, and smoother fluence maps can be expected to suffer less from the most

severe inaccuracies caused by the tongue-and-groove effect, namely, those close to the isocenter (rather than the boundary). In general, it is difficult to assess *a priori* which of these effects is larger. As Earl *et al.*⁴² say, “If the tongue-and-groove effect is not accounted for in the planning, an underestimation of the absolute dose is observed for DAO IMRT plans. The magnitude of the underestimation is dependent upon the aperture shapes.” This suggests that the treatment plan optimization model should make the corresponding determination and find a fluence map and a corresponding collection of apertures that yield a high-quality treatment plan *despite* the tongue-and-groove architecture of the MLC. In this paper, we address this by developing, implementing, and testing a robust DAO model that explicitly takes into account the tongue-and-groove effect and the corresponding inaccuracies of the dose calculation.

II. METHODS

II.A. A robust DAO model

We are assuming that a patient is irradiated from several predetermined beam directions. Each of these beams is discretized into a grid of beamlets, and the collection of all beamlets is denoted by N . Furthermore, K represents the set of all deliverable apertures. Each aperture consists of a number of exposed beamlets, so we denote the set of beamlets in aperture $k \in K$ by $A_k \subseteq N$. In order to evaluate the dose distribution delivered to the patient, the patient geometry is discretized into a set V of the so-called voxels. The dose delivered to voxel $j \in V$ by aperture $k \in K$ at unit intensity is the so-called *aperture dose deposition coefficient* denoted by D_{kj} . We associate a decision variable z_j with each voxel $j \in V$ representing the dose received by voxel $j \in V$. Moreover, we let decision variable y_k represent the intensity of aperture $k \in K$. Voxel doses can then be expressed as a linear function of the aperture intensities through the dose deposition coefficients as follows:

$$z_j = \sum_{k \in K} D_{kj} y_k, \quad j \in V.$$

II.A.1. Bounds on the aperture dose deposition coefficients

Typically, the aperture dose deposition coefficients are expressed implicitly in terms of the beamlet dose deposition coefficients D_{ij} , the dose delivered to voxel $j \in V$ by beamlet $i \in N$ at unit intensity,

$$D_{kj} = \sum_{i \in A_k} D_{ij}, \quad k \in K, \quad j \in V. \quad (1)$$

This expression only provides an approximation to the aperture dose deposition coefficients (see, e.g., Refs. 41 and 43–45). In this paper, we will focus on the fact that Eq. (1) ignores the presence of the tongue-and-groove architecture of the leaves. In particular, beamlets on the boundary of aperture $k \in K$ that expose a tongue are partially blocked; Fig. 1 illustrates this issue. This will cause inaccuracies in evaluat-

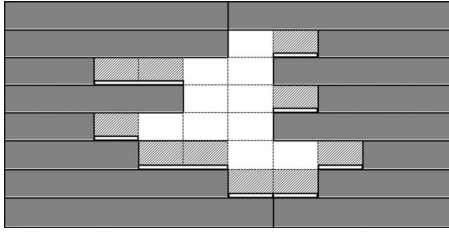


FIG. 1. Shaded beamlets on the boundary of the aperture are partially blocked by the exposed leaf tongues.

ing the dose distribution z_j , $j \in V$ when using the aperture dose deposition coefficients in Eq. (1). In fact, the right-hand side of Eq. (1) is an *upper bound* on \mathcal{D}_{kj} since it effectively assumes that no leaf tongue is present; in the remainder, we will denote this upper bound by

$$\bar{\mathcal{D}}_{kj} = \sum_{i \in A_k} D_{ij}.$$

Thus, the realized dose distribution will be smaller than the one that is planned based on the approximation in Eq. (1), potentially causing underdosing of the target voxels.

To be able to measure the effect of the leaf tongue, let D_{ij}^δ denote the dose delivered to voxel $j \in V$ by beamlet $i \in N$ at unit intensity from which a strip of width δ is removed. (Clearly, $D_{ij}^0 = D_{ij}$ and $D_{ij}^\delta < D_{ij}^{\delta'}$ if $\delta > \delta'$.) We then define

$$\underline{\mathcal{D}}_{kj}^\delta \equiv \sum_{i \in A_k \setminus \partial_k} D_{ij} + \sum_{i \in \partial_k} D_{ij}^\delta, \quad j \in V, \quad k \in K,$$

where $\partial_k \subseteq A_k$ is the set of beamlets in aperture k that expose a tongue. Now if the leaf tongue has width δ , then

$$\mathcal{D}_{kj} = \underline{\mathcal{D}}_{kj}^\delta, \quad k \in K, \quad j \in V.$$

However, since the width of the leaf tongue is typically on the order of 0.5 mm (see, e.g., Ref. 46 or Ref. 47), it is difficult for commonly used approximate dose models to accurately account for its effect on the beamlet dose deposition coefficients. Our approach will therefore be to obtain a lower bound for these coefficients by overestimating the width of the tongue to allow for an accurate estimation of $\underline{\mathcal{D}}_{kj}^\delta$. Clearly, when δ is chosen to be larger than the width of the tongue, $\underline{\mathcal{D}}_{kj}^\delta$ provides a *lower bound* on \mathcal{D}_{kj} since it effectively assumes that the leaf tongue is larger (and therefore blocks more dose) than the physical leaf tongue. For convenience, we will, in the remainder, denote the matrices of aperture dose deposition coefficients as well as their lower and upper bounds by \mathcal{D} , $\underline{\mathcal{D}}^\delta$, and $\bar{\mathcal{D}}$.

II.A.2. Treatment plan evaluation criteria

Now suppose a collection of treatment plan evaluation criteria, say L , has been identified to measure the treatment plan quality and are expressed as functions of the dose distribution: $G_\ell: \mathbb{R}^{|V|} \rightarrow \mathbb{R}$ for $\ell \in L$. Without loss of generality, we assume that smaller values are preferred to larger values. For convenience, we also assume that all criteria are convex. This is the case for many criteria proposed in literature, such

as voxel-based penalties, equivalent uniform dose, tail means, or conditional value-at-risk, etc. In addition, it has been shown that in a multicriteria framework where the relative weights of the different criteria are varied, many others can equivalently be replaced by convex ones; e.g., tumor control probability (TCP) and normal tissue complication probability (NTCP) (see Ref. 48). Moreover, we assume that the set L can be partitioned into two subsets $L = \underline{L} \cup \bar{L}$; the criteria that deal with the effects of underdosing correspond to $\ell \in \underline{L}$, while the criteria that deal with overdosing correspond to $\ell \in \bar{L}$. This is usually easy to do, especially if each of the criteria is a function of the dose distribution in a specific structure only. Finally, we make the mild assumption that each of the criteria is monotone in each voxel dose. In particular, if z, z' are two dose distributions such that $z \geq z'$, then $G_\ell(z) \leq G_\ell(z')$ for all $\ell \in \underline{L}$ and $G_\ell(z) \geq G_\ell(z')$ for all $\ell \in \bar{L}$. In other words, uniformly increasing the dose distribution cannot deteriorate an underdosing criterion or improve an overdosing criterion.

II.A.3. Robust DAO

We next propose to use a robust DAO approach to treatment planning, i.e., an approach that identifies a treatment plan that is of high-quality for all values of the aperture dose deposition coefficients within the bounds derived in Sec. II A 1 (see, e.g., Refs. 49 and 50). This is an optimization problem formulated in terms of all deliverable apertures and their associated intensities,

$$\text{minimize} \quad \max_{\underline{\mathcal{D}}^\delta \leq \mathcal{D} \leq \bar{\mathcal{D}}} \sum_{\ell \in L} \gamma_\ell G_\ell \left(\sum_{k \in K} \mathcal{D}_{k,y_k} \right),$$

subject to

$$y_k \geq 0, \quad k \in K,$$

where the values γ_ℓ ($\ell \in L$) are the (non-negative) weights associated with the criteria and $\mathcal{D}_k = (\mathcal{D}_{kj}; j \in V)$ is the vector of aperture dose deposition coefficients for aperture $k \in K$. The objective function therefore measures the worst-case weighted sum of all criteria over all aperture dose deposition coefficients within the specified bounds. [In principle, upper bounds on any or all of the convex criteria functions G_ℓ ($\ell \in L$) could be accommodated as well; however, for ease of exposition, we will not explicitly incorporate this model extension.]

The objective function of (R) has a mathematically inconvenient and cumbersome form. In the remainder of this section, we will therefore propose two reformulations of this model as tractable convex optimization problems. The first is a reformulation that uses the mathematical properties of the treatment plan evaluation criteria to derive a conservative bound on the objective function. The second one is equivalent to (R), but applies only if the treatment plan evaluation criteria are convex voxel-based penalty functions. Both reformulations rely on new decision variables \underline{z}_j and \bar{z}_j that represent a lower and an upper bound on the dose received by voxel $j \in V$, respectively,

$$\underline{z}_j = \sum_{k \in K} \underline{D}_{kj}^\delta y_k, \quad j \in V, \quad (2)$$

$$\bar{z}_j = \sum_{k \in K} \bar{D}_{kj} y_k, \quad j \in V. \quad (3)$$

In other words, Eqs. (2) and (3) provide lower and upper bounds on the dose distribution.

II.A.3.a. Reformulation 1. The first reformulation is based on the following intuitive and appealing argument. First, recall that \underline{z} and \bar{z} are the lower and upper bounds on the delivered dose distribution, respectively. Then suppose we let the objective function measure the treatment plan quality conservatively: Any treatment plan evaluation criterion concerned with underdosing is evaluated at the lower bound of the delivered dose distribution, while any treatment plan evaluation criterion concerned with overdosing is evaluated at the upper bound of the delivered dose distribution. This then leads to the following DAO model:

$$\text{minimize } \sum_{\ell \in L} \gamma_\ell G_\ell(\underline{z}) + \sum_{\ell \in \bar{L}} \gamma_\ell G_\ell(\bar{z}),$$

subject to (P^c)

$$\underline{z}_j = \sum_{k \in K} \underline{D}_{kj}^\delta y_k, \quad j \in V, \quad (4)$$

$$\bar{z}_j = \sum_{k \in K} \bar{D}_{kj} y_k, \quad j \in V, \quad (5)$$

$$y_k \geq 0, \quad k \in K. \quad (6)$$

In Appendix A 1, we show that this reformulation is indeed, if anything, more conservative than (R), i.e., it is robust against larger deviations in the dose deposition coefficients. Moreover, this optimization problem is only marginally larger than a traditional DAO model that does not account for the tongue-and-groove effect.

II.A.3.b. Reformulation 2. Now suppose that the treatment plan evaluation criteria are all voxel-based penalty functions. This means that the objective function of (R), for fixed \mathcal{D} , can be written as

$$\sum_{j \in V} F_j \left(\sum_{k \in K} \mathcal{D}_{kj} y_k \right),$$

where the functions $F_j: \mathbb{R} \rightarrow \mathbb{R}$ are convex ($j \in V$). Note that we have implicitly incorporated the criterion weights γ_ℓ into the voxel-based penalty functions. Typically, but not necessarily, the function F_j will depend only on the structure containing the voxel $j \in V$.

Now let us evaluate, for each voxel, the penalty function at both the upper and the lower bounds of the voxel dose and associate the maximum of these two penalties with this voxel. This yields the following DAO model:

$$\text{minimize } \sum_{j \in V} \max\{F_j(\underline{z}_j), F_j(\bar{z}_j)\},$$

subject to (P)

$$\underline{z}_j = \sum_{k \in K} \underline{D}_{kj}^\delta y_k, \quad j \in V,$$

$$\bar{z}_j = \sum_{k \in K} \bar{D}_{kj} y_k, \quad j \in V,$$

$$y_k \geq 0, \quad k \in K.$$

In Appendix A 2, we show that (P) is equivalent to (R) for voxel-based penalty objectives. Moreover, this optimization problem is of comparable size to (P^c).

II.B. Column generation algorithm

The total number of potential apertures that needs to be included in (P) (i.e., the cardinality of the set K) is very large. For example, for an MLC in which all possible leaf settings are allowable, a beamlet grid of 20×20 for each of five beams yields a total of about 2×10^{46} deliverable apertures. Since this means that it is intractable to solve problems (P) [and (P^c)] directly, we employ a column generation approach; in this iterative approach, we start by choosing a limited set of apertures, denoted by \hat{K} . We then, at each iteration, solve a restricted version of (P) using only the apertures in \hat{K} . Given the corresponding solution, we solve an optimization subproblem that either (i) identifies one or more promising apertures that improve the current solution when added to \hat{K} or (ii) concludes that no such aperture exists and therefore the current solution is optimal. (This problem is often referred to as the pricing problem.) In case (i), we add the identified apertures to \hat{K} and repeat the procedure. Intuitively, the pricing problem identifies those apertures for which the improvement in the objective function per unit intensity is largest, and therefore, once added to the current set of apertures, will significantly improve the treatment plan quality.

In Appendix B, we derive the mathematical form of the pricing subproblem and develop an efficient algorithm for solving this problem to optimality in case the only deliverability constraints of the MLC are the so-called row-convexity constraints. Depending on the manufacturer of the MLC, apertures can be subject to other deliverability constraints as well, and the algorithm described in Appendix B 2 can easily be extended to account for (i) interdigitation constraints and (ii) connectedness constraints.

III. RESULTS

III.A. Patient cases

We used a set of ten head-and-neck cases to study our model. For all cases, we designed plans using five equispaced ^{60}Co beams around the patient. The nominal size of each beam is $40 \times 40 \text{ cm}^2$. The beams are discretized into beamlets with size of $1 \times 1 \text{ cm}^2$, yielding on the order of 1600 beamlets. However, we reduced the number of beamlets considered in the model by defining a mask for each beam eliminating any beamlets that only have a negligible contribution to target coverage. We generated data for a

TABLE I. Problem dimensions.

Case	No. of structures	No. of beamlets	No. of voxels	
			Model	Full
1	14	813	17 108	85 017
2	13	1282	23 998	104 298
3	8	1320	36 288	189 234
4	11	1471	38 609	195 113
5	12	935	15 916	86 255
6	13	692	13 783	58 636
7	10	1044	21 241	102 262
8	10	1005	18 609	84 369
9	10	822	14 520	71 873
10	12	1721	40 198	148 294

voxel grid with size of $4 \times 4 \times 4$ mm³ for all targets and critical structures. These data were used for the evaluation of all treatment plans; however, in the optimization model, we used a coarser resolution of $8 \times 8 \times 8$ mm³ for unspecified tissue (full resolution was used for all targets and other critical structures). Table I shows the problem dimensions for the ten cases.

Each case contains two planning target volumes, PTV1 and PTV2, with prescription doses of 73.8 and 54 Gy, respectively. To determine the clinical quality of the treatment plans obtained by our method, we employ the following clinical dose-volume histogram (DVH) criteria (from the treatment planning protocol used in the Department of Radiation Oncology at the University of Florida):

- PTV1
 - At least 99% should receive at least 93% of the prescribed dose ($0.93 \times 73.8 = 68.6$ Gy).
 - At least 95% should receive at least the prescribed dose (73.8 Gy).
 - No more than 10% should receive more than 110% of the prescribed dose ($1.1 \times 73.8 = 81.2$ Gy).
 - No more than 1% should receive more than 120% of the prescribed dose ($1.2 \times 73.8 = 88.6$ Gy).
- PTV2
 - At least 99% should receive at least 93% of the prescribed dose ($0.93 \times 54 = 50.2$ Gy).
 - At least 95% should receive at least the prescribed dose (54 Gy).
- Salivary glands (left and right parotid glands—LPG/RPG, left and right submandibular glands—LSG/RSG)
 - No more than 50% of each gland should receive more than 30 Gy.
- Other structures
 - Spinal cord (SC) should receive no more than 45 Gy.
 - Brainstem (BS) should receive no more than 54 Gy.
 - Unspecified tissue (UT) should receive no more than 60 Gy.

DVH constraints on additional critical structures, such as, for example, optic nerves and chiasm, were always easily satisfied in our experiments, so we have omitted them from our results and analysis.

III.B. Treatment plan evaluation criteria

For our proof of concept, we used treatment plan evaluation criteria that consist of one-sided quadratic voxel-based penalty functions. In particular, let T denote the set of targets and S the set of *all* structures (including targets). Moreover, let $V_s \subseteq V$ denote the set of voxels in the structure $s \in S$. (For convenience, we will, in the optimization model, assume that each voxel is assigned to a single structure only; the discussion below can, without any problems, be generalized to situations where this is not the case.) We then associate a function penalizing overdosing with the voxels in all structures and a function penalizing underdosing with the voxels in all targets. More specifically, we let (with a slight abuse of notation) $\underline{L} = T \times \{-\}$ and $\bar{L} = S \times \{+\}$ and define

$$G_{(s,-)}(\underline{z}) = \sum_{j \in V_s} \underline{F}_s(z_j), \quad s \in T,$$

$$G_{(s,+)}(\bar{z}) = \sum_{j \in V_s} \bar{F}_s(\bar{z}_j), \quad s \in S,$$

where

$$\underline{F}_s(z_j) = \frac{1}{|V_s|} (\max\{0, \underline{T}_s - z_j\})^2, \quad j \in V_s, \quad s \in T, \quad (7)$$

$$\bar{F}_s(\bar{z}_j) = \frac{1}{|V_s|} (\max\{0, \bar{z}_j - \bar{T}_s\})^2, \quad j \in V_s, \quad s \in S, \quad (8)$$

where Eq. (7) quadratically penalizes underdosing below the underdosing threshold \underline{T}_s in target $s \in T$, while Eq. (8) quadratically penalizes overdosing above the overdosing threshold \bar{T}_s in the structure $s \in S$. In terms of the notation of Sec. II A 3 b, we have the following voxel-based penalty functions:

$$F_j(z_j) = \gamma_{(s,-)} \underline{F}_s(z_j) + \gamma_{(s,+)} \bar{F}_s(z_j), \quad j \in V_s; \quad s \in T,$$

$$F_j(z_j) = \gamma_{(s,+)} \bar{F}_s(z_j), \quad j \in V_s; \quad s \in S \setminus T.$$

We can then write this using a smaller number of treatment plan evaluation criteria by defining

$$G_s(\underline{z}, \bar{z}) = \sum_{j \in V_s} \max\{F_j(\underline{z}_j), F_j(\bar{z}_j)\}, \quad s \in S.$$

III.C. Tongue-and-groove effect bounds

To account for the inaccuracies caused by the tongue-and-groove effect, we provide a lower and an upper bound on the realized dose distribution according to Eqs. (2) and (3). For that purpose, we calculated the dose deposition coefficients for the actual beamlets as well as for the reduced beamlets from which a strip of width δ was removed. In commercial

TABLE II. Lower and upper bounds for target DVH criteria obtained by the new and traditional models for $\delta=0.5$ mm (in % volume).

Case	New model						Traditional model					
	PTV1			PTV2			PTV1			PTV2		
	68.6 Gy	73.8 Gy	81.2 Gy	88.6 Gy	50.2 Gy	54 Gy	68.6 Gy	73.8 Gy	81.2 Gy	88.6 Gy	50.2 Gy	54 Gy
1	(100,100)	(95,98)	(0,2)	(0,0)	(99,99)	(95,96)	(100,100)	(79,96)	(0,1)	(0,0)	(98,99)	(92,95)
5	(100,100)	(100,100)	(0,0)	(0,0)	(98,99)	(96,97)	(100,100)	(92,100)	(0,0)	(0,0)	(98,99)	(94,96)
6	(100,100)	(96,99)	(0,0)	(0,0)	(99,99)	(95,97)	(100,100)	(84,97)	(0,0)	(0,0)	(99,99)	(93,96)

MLCs, the tongue-and-groove offset of the leaf can be as small as 0.5 mm (see Ref. 47). We therefore generated values of D_{ij}^δ for $\delta \in \{0, 0.5, 1, 3, 5\}$ (where δ is measured in mm) and solved the DAO model for each of these values. Since it is hard to accurately quantify the effect of a very small value of δ , we do not advocate using $\delta=0.5$ in a clinical setting. However, we used this value to assess the importance of even small deviations from the “ideal” (i.e., without leaf tongue) dose deposition coefficients. Next, we compared the quality of the treatment plans for the coarser bounds obtained using the larger values $\delta \in \{1, 3, 5\}$ to the treatment plan obtained with the traditional model ($\delta=0$).

III.D. Implementation and results

The optimization problems of the form (P) were solved from our UFORT treatment planning system using our custom primal-dual interior point algorithm (see Ref. 51). We manually tuned the model parameters (i.e., the underdosing and overdosing thresholds as well as the weights associated with the evaluation criteria) based on the two cases. We then used this set of parameters to solve (different variants of) the problem for all ten patient cases. All experiments were per-

formed in MATLAB 2009B on a 2.33 GHz Intel Core 2 Duo computer with 2 GB of RAM using Windows operating system.

As mentioned earlier in this section, we use our column generation algorithm to solve the instances of (P), and at each iteration of the algorithm, we solve the pricing problem to determine if there exists any promising aperture that can improve the treatment plan. As the algorithm progresses and the number of apertures explicitly incorporated in the model increases, the marginal benefit of additional apertures becomes clinically insignificant. Therefore, we terminate the algorithm by monitoring the clinical DVH criteria described above. More specifically, we stop the algorithm if, in the last five iterations, the range of observed DVH criterion value spans less than Δ , where $\Delta=0.2\%$ for targets and $\Delta=2\%$ for critical structures. On the average, it takes around 3–5 min for the traditional DAO method and 7–9 min for the robust DAO method per clinical case to converge.

Table II compares the DVH criteria associated with target coverage obtained by the robust and traditional DAO methods using $\delta=0.5$ for three clinical cases. For all cases and DVH criteria, we show both the lower and the upper bounds

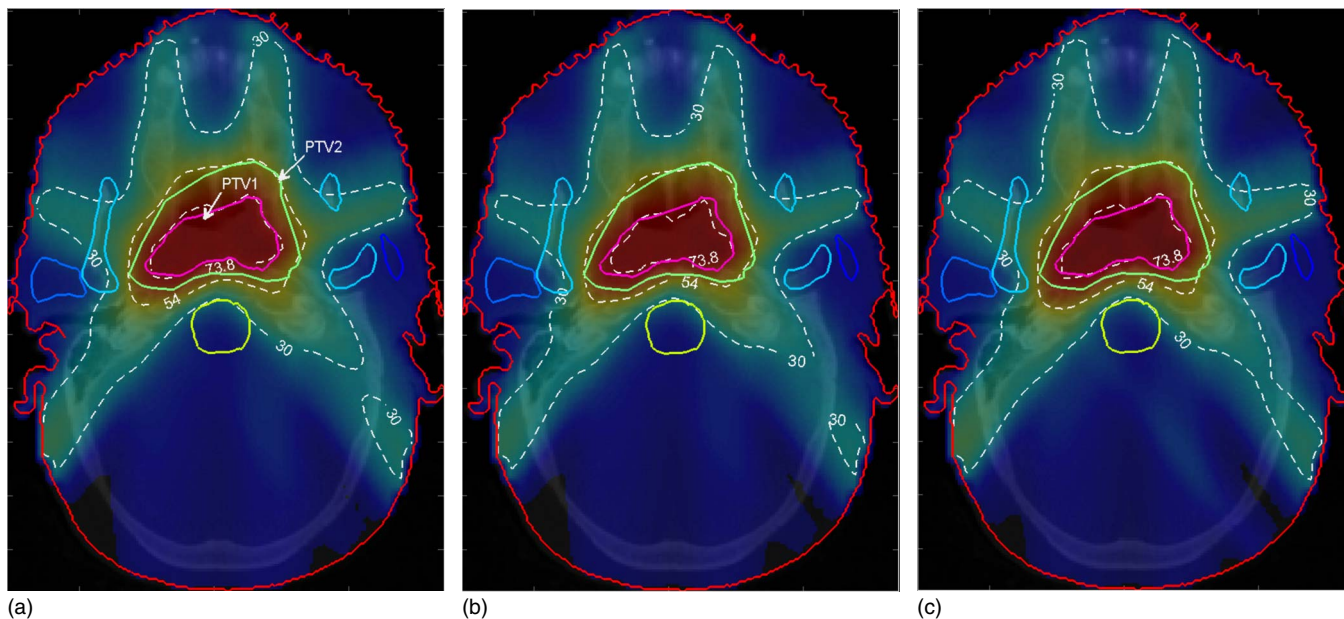


FIG. 2. Isodose curves (dashed lines) for 73.8, 54, and 30 Gy on a typical CT slice corresponding to the optimal treatment plans obtained for case 1 using the traditional and robust models with $\delta=0.5$. (a) and (b) correspond to the upper and lower bounds obtained using the traditional model, and (c) corresponds to the lower bound obtained using the robust model.

TABLE III. Lower and upper bounds for target DVH criteria obtained by the new and traditional models for $\delta=5$ mm (in % volume).

Case	New model						Traditional model					
	PTV1			PTV2			PTV1			PTV2		
	68.6 Gy	73.8 Gy	81.2 Gy	88.6 Gy	50.2 Gy	54 Gy	68.6 Gy	73.8 Gy	81.2 Gy	88.6 Gy	50.2 Gy	54 Gy
1	(100,100)	(91,98)	(0,7)	(0,0)	(97,100)	(92,98)	(0,100)	(0,96)	(0,1)	(0,0)	(39,99)	(22,95)
2	(100,100)	(93,100)	(0,11)	(0,0)	(100,100)	(100,100)	(1,100)	(0,97)	(0,3)	(0,0)	(95,100)	(85,100)
3	(100,100)	(96,99)	(0,1)	(0,0)	(99,100)	(96,98)	(2,100)	(0,99)	(0,0)	(0,0)	(50,100)	(36,97)
4	(100,100)	(100,100)	(0,0)	(0,0)	(98,99)	(95,97)	(0,100)	(0,100)	(0,0)	(0,0)	(31,99)	(15,96)
5	(100,100)	(97,100)	(0,0)	(0,0)	(98,99)	(94,97)	(0,100)	(0,100)	(0,0)	(0,0)	(39,99)	(21,96)
6	(100,100)	(93,97)	(0,0)	(0,0)	(99,99)	(94,97)	(0,100)	(0,97)	(0,0)	(0,0)	(41,99)	(23,96)
7	(100,100)	(97,99)	(0,0)	(0,0)	(99,99)	(96,98)	(0,100)	(0,99)	(0,0)	(0,0)	(53,99)	(32,97)
8	(100,100)	(98,100)	(0,0)	(0,0)	(98,99)	(94,98)	(0,100)	(0,100)	(0,0)	(0,0)	(46,99)	(30,96)
9	(100,100)	(94,100)	(0,2)	(0,0)	(98,99)	(94,97)	(0,100)	(0,99)	(0,0)	(0,0)	(48,98)	(28,95)
10	(100,100)	(100,100)	(0,0)	(0,0)	(100,100)	(100,100)	(25,100)	(0,100)	(0,0)	(0,0)	(92,100)	(30,100)

on the percent volume of a structure that receives at least the specified dose. However, due to the realistic size of the tongue width, the lower bounds are most representative of the realized delivered dose to the patient. The results therefore clearly show that treatment plans obtained by the traditional DAO model exhibit significant underdosing of PTV1 and also, to a somewhat lesser extent, of PTV2. In particular, the actual percent volume of PTV1 receiving at least the prescribed dose (73.8 Gy) in the traditional DAO model (that ignores the tongue-and-groove effect) varies from 79% to 92%, while the corresponding values in the robust DAO model are 95%–100%. This highlights the potential risk of underdosing the target when ignoring the tongue-and-groove effect during the treatment planning phase. In particular, Fig. 2 illustrates the isodose curves on a typical CT slice for clinical case 1. More specifically, isodose lines corresponding to the lower and upper bounds on the dose distribution obtained by the traditional and robust models are separately shown in the figure. For the traditional model, the lower bound on the dose distribution yields several cold spots in PTV1. However, the lower bound obtained by the robust model has significantly fewer cold spots and yields a better target coverage.

In the remainder of this section, we will focus on the robust DAO model with more conservative bounds on the dose distribution obtained with larger δ values. Tables III–VIII compare the results of our robust DAO method that accounts for the tongue-and-groove effect with results of the traditional DAO model on ten clinical cases for $\delta=1, 3$, and 5.

With respect to target coverage, the robust DAO model provides tight DVH bounds compared to the traditional DAO model. In contrast to the traditional DAO model, the robust DAO model is capable of maintaining acceptable DVH lower bounds on target coverage in the majority of the patient cases even for large values of δ . It is important to note that the width of the MLC leaf tongues is (much) smaller than the values of δ that we used in our experiments. This means that the lower bounds calculated for both models can be expected to be quite loose, especially for $\delta=3$ or 5. This means that even in the few cases where the target coverage lower bounds fall slightly short of what is desired, the actual coverage should be clinically acceptable. This argument, of course, also applies to the traditional DAO model. However, for this model, the lower bounds are so far away from what is clinically acceptable that we cannot conclude that the

TABLE IV. Lower and upper bounds for critical-structure DVH criteria obtained by the new and traditional models for $\delta=5$ mm (in % volume).

Case	New model							Traditional model						
	LPG	RPG	LSG	RSG	BS	SC	UT	LPG	RPG	LSG	RSG	BS	SC	UT
	30 Gy	30 Gy	30 Gy	30 Gy	45 Gy	54 Gy	60 Gy	30 Gy	30 Gy	30 Gy	30 Gy	45 Gy	54 Gy	60 Gy
1	(25,28)	(21,21)	(48,59)	(72,78)	(0,0)	(0,0)	(0,0)	(9,21)	(10,18)	(14,37)	(49,66)	(0,0)	(0,0)	(0,0)
2	(76,77)	(94,94)	(100,100)	(100,100)	(0,2)	(0,0)	(1,1)	(67,75)	(88,91)	(100,100)	(100,100)	(1,3)	(0,1)	(0,1)
3	(21,21)	(18,18)	N/A	N/A	N/A	(0,0)	(0,0)	(17,21)	(15,19)	N/A	N/A	N/A	(0,0)	(0,0)
4	(10,10)	(2,2)	(60,61)	(19,22)	(0,0)	(0,0)	(0,0)	(5,9)	(0,2)	(49,57)	(9,19)	(0,0)	(0,0)	(0,0)
5	(39,45)	(0,0)	(47,56)	(48,54)	(0,0)	(0,0)	(0,0)	(28,43)	(0,0)	(26,47)	(22,43)	(0,0)	(0,0)	(0,0)
6	(28,30)	(41,43)	N/A	N/A	(6,6)	(0,0)	(0,0)	(19,27)	(22,37)	N/A	N/A	(0,5)	(0,0)	(0,0)
7	(0,0)	(49,49)	(32,38)	(100,100)	(0,0)	(0,0)	(0,0)	(0,1)	(38,48)	(4,25)	(100,100)	(0,0)	(0,0)	(0,0)
8	(29,30)	(7,8)	(100,100)	(61,69)	(0,0)	(0,0)	(0,0)	(18,32)	(0,4)	(96,100)	(41,59)	(0,0)	(0,0)	(0,0)
9	(2,2)	(45,47)	(43,48)	(100,100)	(0,0)	(0,0)	(0,0)	(0,3)	(26,41)	(19,41)	(83,95)	(0,0)	(0,0)	(0,0)
10	(7,7)	(47,48)	(27,32)	(100,100)	(0,0)	(0,1)	(0,0)	(25,27)	(44,50)	(27,36)	(97,100)	(0,0)	(0,0)	(0,0)

TABLE V. Lower and upper bounds for target DVH criteria obtained by the new and traditional models for $\delta=3$ mm (in % volume).

Case	New model						Traditional model					
	PTV1			PTV2			PTV1			PTV2		
	68.6 Gy	73.8 Gy	81.2 Gy	88.6 Gy	50.2 Gy	54 Gy	68.6 Gy	73.8 Gy	81.2 Gy	88.6 Gy	50.2 Gy	54 Gy
1	(100,100)	(92,98)	(0,4)	(0,0)	(98,100)	(93,97)	(1,100)	(0,96)	(0,1)	(0,0)	(76,99)	(57,95)
2	(100,100)	(95,100)	(0,8)	(0,0)	(100,100)	(100,100)	(27,100)	(1,97)	(0,3)	(0,0)	(100,100)	(98,100)
3	(100,100)	(97,100)	(0,0)	(0,0)	(100,100)	(97,99)	(56,100)	(1,99)	(0,0)	(0,0)	(84,100)	(56,97)
4	(100,100)	(100,100)	(0,0)	(0,0)	(98,99)	(95,97)	(17,100)	(0,100)	(0,0)	(0,0)	(80,99)	(41,96)
5	(100,100)	(99,100)	(0,0)	(0,0)	(98,99)	(94,97)	(24,100)	(0,100)	(0,0)	(0,0)	(78,99)	(51,96)
6	(100,100)	(94,99)	(0,0)	(0,0)	(99,100)	(95,97)	(3,100)	(0,97)	(0,0)	(0,0)	(79,99)	(54,96)
7	(100,100)	(96,100)	(0,0)	(0,0)	(99,99)	(96,98)	(10,100)	(0,99)	(0,0)	(0,0)	(84,99)	(64,97)
8	(100,100)	(97,100)	(0,0)	(0,0)	(98,99)	(94,98)	(13,100)	(0,100)	(0,0)	(0,0)	(79,99)	(57,96)
9	(100,100)	(95,100)	(0,2)	(0,0)	(98,99)	(94,97)	(5,100)	(0,99)	(0,0)	(0,0)	(79,98)	(61,95)
10	(100,100)	(100,100)	(0,0)	(0,0)	(100,100)	(100,100)	(95,100)	(9,100)	(0,0)	(0,0)	(100,100)	(100,100)

treatment plans obtained without taking tongue-and-groove effects into account are clinically acceptable. In particular, employing the robust DAO model can significantly improve the uncertainties in PTV1 and PTV2 coverage due to the tongue-and-groove effect. For example, with the robust DAO model and $\delta=3$, the lower bound on PTV1 target coverage at the prescription dose of 73.8 Gy is at least 95% in eight out

of ten cases, and 92% and 94% in the other two cases, respectively. For the traditional model, the corresponding lower bounds are below 10% in all cases.

With respect to structure sparing, for larger values of δ the traditional DAO model provides smaller upper bounds on the corresponding DVH criteria than the robust DAO model. This can be explained by the fact that to ensure a desirable

TABLE VI. Lower and upper bounds for critical-structure DVH criteria obtained by the new and traditional models for $\delta=3$ mm (in % volume).

Case	New model							Traditional model						
	LPG	RPG	LSG	RSG	BS	SC	UT	LPG	RPG	LSG	RSG	BS	SC	UT
	30 Gy	30 Gy	30 Gy	30 Gy	45 Gy	54 Gy	60 Gy	30 Gy	30 Gy	30 Gy	30 Gy	45 Gy	54 Gy	60 Gy
1	(24,25)	(19,20)	(40,51)	(68,73)	(0,0)	(0,0)	(0,0)	(15,21)	(14,18)	(24,37)	(58,66)	(0,0)	(0,0)	(0,0)
2	(75,75)	(92,93)	(100,100)	(100,100)	(1,3)	(1,1)	(1,1)	(70,75)	(90,91)	(100,100)	(100,100)	(0,3)	(0,1)	(0,1)
3	(21,21)	(18,18)	N/A	N/A	N/A	(0,0)	(0,0)	(19,21)	(17,19)	N/A	N/A	N/A	(0,0)	(0,0)
4	(10,11)	(1,2)	(59,60)	(19,20)	(0,0)	(0,0)	(0,0)	(8,9)	(0,2)	(53,57)	(11,19)	(0,0)	(0,0)	(0,0)
5	(42,45)	(0,0)	(48,56)	(44,48)	(0,0)	(0,0)	(0,0)	(35,43)	(0,0)	(34,47)	(30,43)	(0,0)	(0,0)	(0,0)
6	(32,33)	(41,44)	N/A	N/A	(4,5)	(0,0)	(0,0)	(21,27)	(28,37)	N/A	N/A	(1,5)	(0,0)	(0,0)
7	(0,1)	(48,49)	(33,38)	(100,100)	(0,0)	(0,0)	(0,0)	(0,1)	(43,48)	(11,25)	(100,100)	(0,0)	(0,0)	(0,0)
8	(29,30)	(7,8)	(100,100)	(63,69)	(0,0)	(0,0)	(0,0)	(23,32)	(2,4)	(98,100)	(49,59)	(0,0)	(0,0)	(0,0)
9	(2,2)	(42,45)	(40,44)	(100,100)	(0,0)	(0,0)	(0,0)	(0,3)	(33,41)	(28,41)	(90,95)	(0,0)	(0,0)	(0,0)
10	(9,11)	(47,49)	(30,33)	(99,100)	(0,0)	(0,0)	(0,0)	(25,27)	(46,50)	(30,36)	(98,100)	(0,0)	(0,0)	(0,0)

TABLE VII. Lower and upper bounds for target DVH criteria obtained by the new and traditional models for $\delta=1$ mm (in % volume).

Case	New model						Traditional model					
	PTV1			PTV2			PTV1			PTV2		
	68.6 Gy	73.8 Gy	81.2 Gy	88.6 Gy	50.2 Gy	54 Gy	68.6 Gy	73.8 Gy	81.2 Gy	88.6 Gy	50.2 Gy	54 Gy
1	(100,100)	(94,98)	(0,3)	(0,0)	(99,99)	(94,97)	(99,100)	(30,96)	(0,1)	(0,0)	(96,99)	(88,95)
2	(100,100)	(96,99)	(0,6)	(0,0)	(100,100)	(100,100)	(100,100)	(56,97)	(0,3)	(0,0)	(100,100)	(100,100)
3	(100,100)	(97,100)	(0,0)	(0,0)	(100,100)	(97,99)	(100,100)	(67,99)	(0,0)	(0,0)	(99,100)	(91,97)
4	(100,100)	(100,100)	(0,0)	(0,0)	(98,98)	(95,97)	(100,100)	(50,100)	(0,0)	(0,0)	(98,99)	(90,96)
5	(100,100)	(99,100)	(0,0)	(0,0)	(98,99)	(95,97)	(100,100)	(50,100)	(0,0)	(0,0)	(97,99)	(89,96)
6	(100,100)	(95,99)	(0,0)	(0,0)	(99,100)	(95,97)	(100,100)	(35,97)	(0,0)	(0,0)	(98,99)	(89,96)
7	(100,100)	(98,100)	(0,0)	(0,0)	(99,99)	(96,98)	(100,100)	(36,99)	(0,0)	(0,0)	(98,99)	(91,97)
8	(100,100)	(99,100)	(0,0)	(0,0)	(99,99)	(96,97)	(100,100)	(44,100)	(0,0)	(0,0)	(98,99)	(89,96)
9	(100,100)	(98,100)	(0,0)	(0,0)	(98,99)	(95,97)	(100,100)	(37,99)	(0,0)	(0,0)	(96,98)	(89,95)
10	(100,100)	(100,100)	(0,0)	(0,0)	(100,100)	(100,100)	(100,100)	(91,100)	(0,0)	(0,0)	(100,100)	(100,100)

TABLE VIII. Lower and upper bounds for critical-structure DVH criteria obtained by the new and traditional models for $\delta=1$ mm (in % volume).

Case	New model							Traditional model						
	LPG	RPG	LSG	RSG	BS	SC	UT	LPG	RPG	LSG	RSG	BS	SC	UT
	30 Gy	30 Gy	30 Gy	30 Gy	45 Gy	54 Gy	60 Gy	30 Gy	30 Gy	30 Gy	30 Gy	45 Gy	54 Gy	60 Gy
1	(22,23)	(19,20)	(38,41)	(66,66)	(0,0)	(0,0)	(0,0)	(19,21)	(16,18)	(31,37)	(65,66)	(0,0)	(0,0)	(0,0)
2	(74,74)	(91,91)	(100,100)	(100,100)	(2,2)	(1,5)	(1,1)	(74,75)	(91,91)	(100,100)	(100,100)	(1,3)	(0,1)	(0,1)
3	(21,21)	(19,19)	N/A	N/A	N/A	(0,0)	(0,0)	(20,21)	(18,19)	N/A	N/A	N/A	(0,0)	(0,0)
4	(9,10)	(1,1)	(58,58)	(18,18)	(0,0)	(0,0)	(0,0)	(9,9)	(1,2)	(56,57)	(17,19)	(0,0)	(0,0)	(0,0)
5	(42,43)	(0,0)	(46,49)	(39,41)	(0,0)	(0,0)	(0,0)	(40,43)	(0,0)	(41,47)	(38,43)	(0,0)	(0,0)	(0,0)
6	(28,29)	(38,41)	N/A	N/A	(5,7)	(0,0)	(0,0)	(24,27)	(34,37)	N/A	N/A	(1,5)	(0,0)	(0,0)
7	(0,0)	(48,48)	(24,26)	(100,100)	(0,0)	(0,0)	(0,0)	(0,1)	(47,48)	(21,25)	(100,100)	(0,0)	(0,0)	(0,0)
8	(29,30)	(3,4)	(100,100)	(59,60)	(0,0)	(0,0)	(0,0)	(29,32)	(3,4)	(100,100)	(56,59)	(0,0)	(0,0)	(0,0)
9	(3,5)	(41,43)	(38,41)	(97,97)	(0,0)	(0,0)	(0,0)	(2,3)	(39,41)	(39,41)	(95,95)	(0,0)	(0,0)	(0,0)
10	(13,13)	(48,48)	(30,30)	(97,97)	(0,0)	(0,0)	(0,0)	(26,27)	(49,50)	(35,36)	(100,100)	(0,0)	(0,0)	(0,0)

target coverage for *all* aperture dose deposition coefficients in the range $\underline{D}^\delta \leq D \leq \overline{D}$ (i.e., a desirable lower bound on target coverage), the robust DAO model is forced to select apertures that provide more dose to critical structures. However, please note that in the majority of the cases, the robust DAO model satisfies the DVH criterion for a given structure whenever the traditional DAO model satisfies this criterion. Not surprisingly, the difference between the bounds obtained by the robust and traditional models decreases with the value of δ .

Figures 3 and 4 illustrate the DVHs of the optimal treatment plans obtained by the robust and traditional DAO models for case 5 with $\delta=3$ mm. For a given structure, the solid and dashed lines represent the upper and lower bounds on DVH values, respectively. The upper bounds correspond to an idealized dose distribution that ignores the presence of the

leaf tongues, while the lower bounds indicate how much the actual delivered dose distribution may deviate from the optimized one.

Finally, we also evaluated the ability of the more conservative model (P^c) in providing high-quality robust treatment plans that account for the tongue-and-groove effect. As noted before, when using treatment plan evaluation criteria that are voxel-based penalty functions (as described in Sec. III B), we do not need to use the more conservative model, but can derive an exact robust model. However, it is still interesting to compare the two approaches since the more conservative one should be used if other criteria are desired. These experiments showed that the differences between the conservative and the exact robust models in terms of DVH bounds are negligible. We illustrate the results of the two models for $\delta=3$ in Tables IX and X.

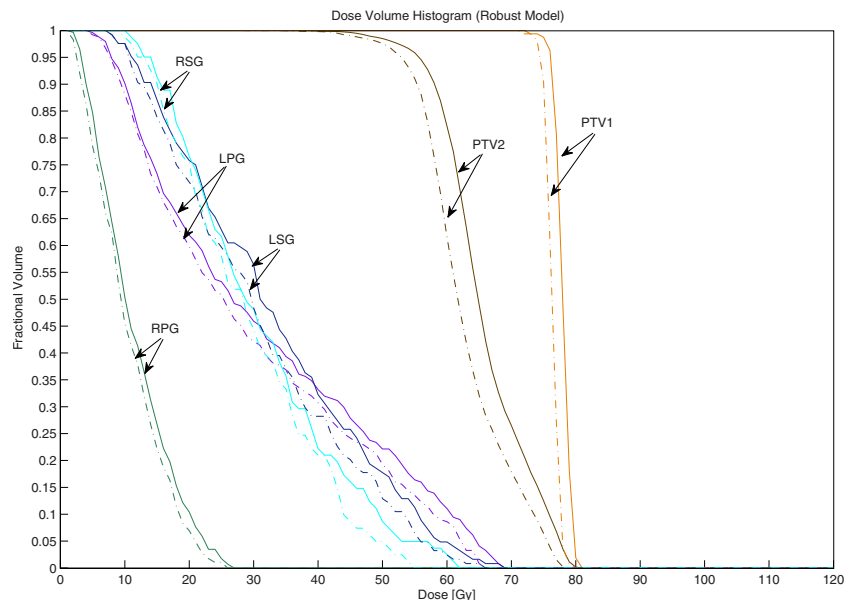


FIG. 3. Lower (dashed) and upper (solid) bounds on the DVHs of the optimal treatment plan obtained by the robust model for case 5 using $\delta=3$ mm.

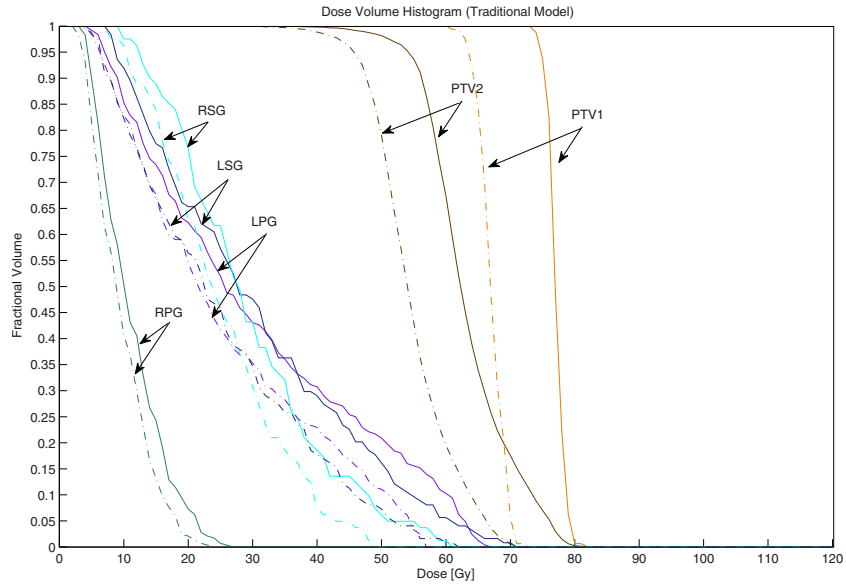


FIG. 4. Lower (dashed) and upper (solid) bounds on the DVHs of the optimal treatment plan obtained by the traditional model for case 5 using $\delta=3$ mm.

IV. CONCLUSION

In this paper, we developed a robust direct aperture optimization method for IMRT treatment planning that explicitly

accounts for the dosimetric inaccuracies caused by the tongue-and-groove architecture of the MLC leaves. Our approach does not rely on being able to accurately compute these dosimetric effects. Instead, we employ dose calcula-

TABLE IX. Lower and upper bounds for target DVH criteria obtained by the conservative and exact robust models for $\delta=3$ mm (in % volume).

Case	Conservative model						Robust model					
	PTV1			PTV2			PTV1			PTV2		
	68.6 Gy	73.8 Gy	81.2 Gy	88.6 Gy	50.2 Gy	54 Gy	68.6 Gy	73.8 Gy	81.2 Gy	88.6 Gy	50.2 Gy	54 Gy
1	(100,100)	(92,98)	(0,4)	(0,0)	(98,99)	(93,97)	(100,100)	(92,98)	(0,4)	(0,0)	(98,100)	(93,97)
2	(100,100)	(95,99)	(0,7)	(0,0)	(100,100)	(100,100)	(100,100)	(95,100)	(0,8)	(0,0)	(100,100)	(100,100)
3	(100,100)	(97,100)	(0,0)	(0,0)	(100,100)	(97,99)	(100,100)	(97,100)	(0,0)	(0,0)	(100,100)	(97,99)
4	(100,100)	(100,100)	(0,0)	(0,0)	(98,99)	(95,97)	(100,100)	(100,100)	(0,0)	(0,0)	(98,99)	(95,97)
5	(100,100)	(95,100)	(0,0)	(0,0)	(98,99)	(94,97)	(100,100)	(99,100)	(0,0)	(0,0)	(98,99)	(94,97)
6	(100,100)	(95,98)	(0,0)	(0,0)	(99,99)	(95,98)	(100,100)	(94,99)	(0,0)	(0,0)	(99,100)	(95,97)
7	(100,100)	(96,100)	(0,0)	(0,0)	(99,99)	(96,98)	(100,100)	(96,100)	(0,0)	(0,0)	(99,99)	(96,98)
8	(100,100)	(97,100)	(0,0)	(0,0)	(98,99)	(97,98)	(100,100)	(97,100)	(0,0)	(0,0)	(98,99)	(94,98)
9	(100,100)	(95,100)	(0,2)	(0,0)	(98,99)	(94,97)	(100,100)	(95,100)	(0,2)	(0,0)	(98,99)	(94,97)
10	(100,100)	(100,100)	(0,0)	(0,0)	(100,100)	(100,100)	(100,100)	(100,100)	(0,0)	(0,0)	(100,100)	(100,100)

TABLE X. Lower and upper bounds for critical-structure DVH criteria obtained by the conservative and exact robust models for $\delta=3$ mm (in % volume).

Case	Conservative model							Robust model						
	LPG	RPG	LSG	RSG	BS	SC	UT	LPG	RPG	LSG	RSG	BS	SC	UT
	30 Gy	30 Gy	30 Gy	30 Gy	45 Gy	54 Gy	60 Gy	30 Gy	30 Gy	30 Gy	30 Gy	45 Gy	54 Gy	60 Gy
1	(23,25)	(20,21)	(40,51)	(68,70)	(0,0)	(0,0)	(0,0)	(24,25)	(19,20)	(40,51)	(68,73)	(0,0)	(0,0)	(0,0)
2	(73,74)	(92,93)	(100,100)	(100,100)	(2,4)	(1,1)	(1,1)	(75,75)	(92,93)	(100,100)	(100,100)	(1,3)	(1,1)	(1,1)
3	(21,21)	(18,19)	N/A	N/A	N/A	(0,0)	(0,0)	(21,21)	(18,18)	N/A	N/A	N/A	(0,0)	(0,0)
4	(9,10)	(2,2)	(60,60)	(18,18)	(0,0)	(0,0)	(0,0)	(10,11)	(1,2)	(59,60)	(19,20)	(0,0)	(0,0)	(0,0)
5	(43,46)	(0,0)	(48,53)	(46,50)	(0,0)	(0,0)	(0,0)	(42,45)	(0,0)	(48,56)	(44,48)	(0,0)	(0,0)	(0,0)
6	(28,31)	(43,45)	N/A	N/A	(4,4)	(0,0)	(0,0)	(32,33)	(41,44)	N/A	N/A	(4,5)	(0,0)	(0,0)
7	(0,1)	(48,49)	(28,35)	(100,100)	(0,0)	(0,0)	(0,0)	(0,1)	(48,49)	(33,38)	(100,100)	(0,0)	(0,0)	(0,0)
8	(29,30)	(8,9)	(100,100)	(60,66)	(0,0)	(0,0)	(0,0)	(29,30)	(7,8)	(100,100)	(63,69)	(0,0)	(0,0)	(0,0)
9	(3,5)	(43,45)	(40,45)	(100,100)	(0,0)	(0,0)	(0,0)	(2,2)	(42,45)	(40,44)	(100,100)	(0,0)	(0,0)	(0,0)
10	(10,11)	(47,48)	(37,37)	(98,100)	(0,0)	(0,0)	(0,0)	(9,11)	(47,49)	(30,33)	(99,100)	(0,0)	(0,0)	(0,0)

tions that overestimate the dimensions of the leaf tongues so that commonly used dose models can be applied. Our model then aims to find a robust treatment plan that can obtain treatment plans of high clinical quality regardless of the exact consequences of the tongue-and-groove architecture. Due to the computational difficulty of the resulting robust optimization model, we propose a reformulation that provides a conservative bound on the robust model as well as an exact and tractable reformulation for the case of voxel-based penalty criteria. We investigated the performance of the robust and traditional DAO models in the presence of the tongue-and-groove effect on ten clinical head-and-neck cancer cases. The experiments validate the ability of the proposed approach in designing robust treatment plans. Although the suggested robust optimization approach is presented to account for the tongue-and-groove effect, in principle, it could be applied to any source of dosimetric inaccuracies for which a lower and an upper bound on the beamlet dose deposition coefficients can be provided. Future research can extend the research in this paper by generalizing this approach to account for dosimetric inaccuracies caused by employing approximate dose calculation methods.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under Grant No. DMI-0457394/CMMI-0852727.

APPENDIX A: REFORMULATIONS

1. Reformulation 1

It is easy to see that the objective function satisfies

$$\begin{aligned} & \max_{\underline{\mathcal{D}}^{\delta} \leq \mathcal{D} \leq \bar{\mathcal{D}}} \sum_{\ell \in L} \gamma_{\ell} G_{\ell} \left(\sum_{k \in K} \mathcal{D}_{k,j} y_k \right) \\ & \leq \sum_{\ell \in L} \gamma_{\ell} \max_{\underline{\mathcal{D}}^{\delta} \leq \mathcal{D} \leq \bar{\mathcal{D}}} G_{\ell} \left(\sum_{k \in K} \mathcal{D}_{k,j} y_k \right). \end{aligned} \quad (\text{A1})$$

Since for each $\ell \in \underline{L}$ the function G_{ℓ} is nonincreasing, we have

$$\max_{\underline{\mathcal{D}}^{\delta} \leq \mathcal{D} \leq \bar{\mathcal{D}}} G_{\ell} \left(\sum_{k \in K} \mathcal{D}_{k,j} y_k \right) = G_{\ell} \left(\sum_{k \in K} \underline{\mathcal{D}}_{k,j}^{\delta} y_k \right),$$

while for each $\ell \in \bar{L}$ the function G_{ℓ} is nondecreasing, we have

$$\max_{\underline{\mathcal{D}}^{\delta} \leq \mathcal{D} \leq \bar{\mathcal{D}}} G_{\ell} \left(\sum_{k \in K} \mathcal{D}_{k,j} y_k \right) = G_{\ell} \left(\sum_{k \in K} \bar{\mathcal{D}}_{k,j} y_k \right).$$

This implies that the objective function of (P^c) is an upper bound on that of (R).

2. Reformulation 2

In contrast to Eq. (A1), we can now exactly reformulate the objective function as follows:

$$\begin{aligned} & \max_{\underline{\mathcal{D}}^{\delta} \leq \mathcal{D} \leq \bar{\mathcal{D}}} \sum_{j \in V} F_j \left(\sum_{k \in K} \mathcal{D}_{k,j} y_k \right) \\ & = \sum_{j \in V} \max_{\underline{\mathcal{D}}_j^{\delta} \leq \mathcal{D}_j \leq \bar{\mathcal{D}}_j} F_j \left(\sum_{k \in K} \mathcal{D}_{k,j} y_k \right), \end{aligned}$$

where $\mathcal{D}_j = (\mathcal{D}_{k,j}; k \in K)$ is the vector of aperture dose deposition coefficients corresponding to voxel $j \in V$ (and similarly for $\underline{\mathcal{D}}_j^{\delta}$ and $\bar{\mathcal{D}}_j$). Then for each $j \in V$, we have, due to the convexity of F_j and the non-negativity of y_k ($k \in K$),

$$\begin{aligned} & \max_{\underline{\mathcal{D}}_j^{\delta} \leq \mathcal{D}_j \leq \bar{\mathcal{D}}_j} F_j \left(\sum_{k \in K} \mathcal{D}_{k,j} y_k \right) \\ & = \max \left\{ F_j \left(\sum_{k \in K} \underline{\mathcal{D}}_{k,j}^{\delta} y_k \right), F_j \left(\sum_{k \in K} \bar{\mathcal{D}}_{k,j} y_k \right) \right\}. \end{aligned}$$

This implies that the objective function of (P) is equal to that of (R).

APPENDIX B: PRICING PROBLEM

1. Formulating the pricing problem

In this section, we formulate the pricing problem of the column generation algorithm described in Sec. II B and develop a solution approach to solve it. Let π_j and $\bar{\pi}_j$ ($j \in V$) and ρ_k ($k \in K$) be the dual multipliers associated with constraints (4)–(6). In order to be able to accommodate both the two reformulations in Sec. II A 3, we will actually study a slightly more general model, which, with a slight abuse of notation, allows for the treatment plan evaluation criteria to be a function of both \underline{z} and \bar{z} : $G_{\ell}(\underline{z}, \bar{z})$. Since the objective function is convex and the constraints are linear, the Karush–Kuhn–Tucker (KKT) conditions (see, e.g., Ref. 52) are necessary and sufficient conditions for optimality of (P). Assuming, for convenience, that the objective function is differentiable, they can be written as follows:

$$\underline{z}_j = \sum_{k \in K} \underline{\mathcal{D}}_{k,j}^{\delta} y_k, \quad j \in V,$$

$$\bar{z}_j = \sum_{k \in K} \bar{\mathcal{D}}_{k,j} y_k, \quad j \in V,$$

$$\pi_j = \sum_{\ell \in L} \gamma_{\ell} \frac{dG_{\ell}(\underline{z}, \bar{z})}{dz_j}, \quad j \in V,$$

$$\bar{\pi}_j = \sum_{\ell \in L} \gamma_{\ell} \frac{dG_{\ell}(\underline{z}, \bar{z})}{d\bar{z}_j}, \quad j \in V,$$

$$\rho_k = \sum_{j \in V} \underline{\mathcal{D}}_{k,j}^{\delta} \pi_j + \sum_{j \in V} \bar{\mathcal{D}}_{k,j} \bar{\pi}_j, \quad k \in K,$$

$$y_k \rho_k = 0, \quad k \in K,$$

$$y_k, \rho_k \geq 0, \quad k \in K.$$

(If the treatment plan evaluation criteria are convex but not everywhere differentiable, we can use the generalized KKT

conditions derived by Hiriart-Urruty.⁵³ The analysis in the remainder of this section remains essentially unchanged.) Any solution of the system above can be characterized by a vector of aperture intensities $y \geq 0$; this vector then determines \underline{z} , \bar{z} , $\underline{\pi}$, $\bar{\pi}$, and ρ . Now let $(\hat{y}; \hat{\pi}, \hat{\pi}, \hat{\rho})$ be an optimal pair of primal and dual solutions to a restricted version of (P) in which only apertures in the set $\hat{K} \subset K$ are considered. In other words, $\hat{y}_k = 0$ for $k \in K \setminus \hat{K}$. This solution is optimal to (P) if and only if $\hat{\rho}_k \geq 0$ for all $k \in K$ since in this case the set of KKT conditions for (P) will be fully satisfied. In order to check whether or not these constraints are satisfied, we formulate the so-called *pricing problem* as follows:

$$\text{minimize}_{k \in K} \hat{\rho}_k = \sum_{j \in V} \underline{D}_{kj}^{\delta} \hat{\pi}_j + \sum_{j \in V} \bar{D}_{kj} \hat{\pi}_j.$$

Furthermore, we have

$$\begin{aligned} \sum_{j \in V} \underline{D}_{kj}^{\delta} \hat{\pi}_j + \sum_{j \in V} \bar{D}_{kj} \hat{\pi}_j &= \sum_{j \in V} \left(\sum_{i \in A_k \setminus \partial_k} D_{ij} + \sum_{i \in \partial_k} D_{ij}^{\delta} \right) \hat{\pi}_j \\ &\quad + \sum_{j \in V} \left(\sum_{i \in A_k} D_{ij} \right) \hat{\pi}_j \\ &= \sum_{i \in A_k \setminus \partial_k} \left(\sum_{j \in V} D_{ij} \hat{\pi}_j + \sum_{j \in V} D_{ij}^{\delta} \hat{\pi}_j \right) \\ &\quad + \sum_{i \in \partial_k} \left(\sum_{j \in V} D_{ij}^{\delta} \hat{\pi}_j + \sum_{j \in V} D_{ij} \hat{\pi}_j \right). \end{aligned}$$

We can interpret the objective function of the pricing problem as follows: As we increase the intensity of a beamlet $i \in \partial_k$, $\sum_{j \in V} D_{ij}^{\delta} \hat{\pi}_j$ represents the per unit change to the objective function value due to underdosing effects, while $\sum_{j \in V} D_{ij} \hat{\pi}_j$ represents the per unit change in the objective function value due to overdosing effects. Similarly, for a beamlet $i \in A_k \setminus \partial_k$, $\sum_{j \in V} D_{ij} \hat{\pi}_j$ and $\sum_{j \in V} D_{ij}^{\delta} \hat{\pi}_j$ represent the analogous per unit changes to the objective function value, respectively.

2. Solving the pricing problem

In our pricing problem, the cost of aperture $k \in K$, $\hat{\rho}_k$, clearly depends on the beamlets that are exposed in the aperture (i.e., the set A_k). First, note that the pricing problem decomposes by beam direction; i.e., we can solve the pricing problem for each individual beam direction. Without incorporating the tongue-and-groove effect, Romeijn *et al.*⁴⁰ showed that the pricing problem decomposes by beamlet row, i.e., we can find the optimal solution to the pricing problem by independently finding an optimal pair of leaf settings for each beamlet row. However, when the tongue-and-groove effect is taken into account, the leaf settings in adjacent beamlet rows are important since these determine the set ∂_k . We therefore formulate and solve the pricing problem for a given beam using a modification of the network flow model that was developed in Ref. 40 for solving the pricing problem under interdigation constraints. In particular, consider a fixed beam with a beamlet grid of dimensions $m \times n$ (i.e., m rows and n columns). We then create a network

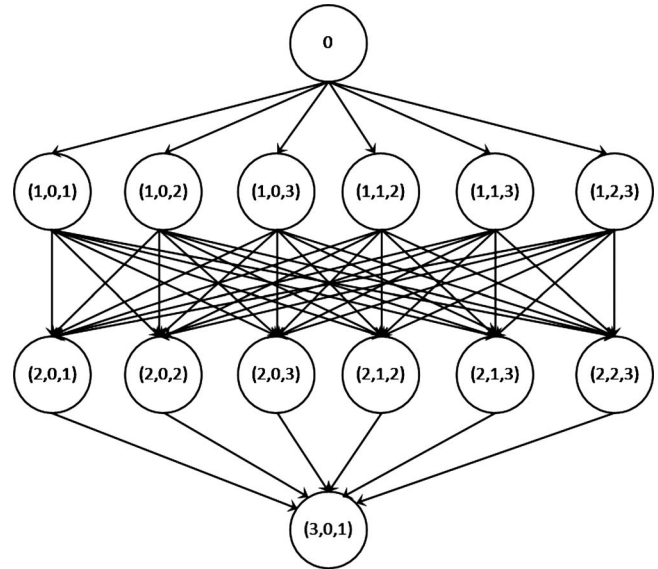


FIG. 5. The network model for the pricing problem under row-convexity constraints.

where each node corresponds to a potential leaf setting in a particular beamlet row; i.e., a typical node is characterized as (r, c_1, c_2) , where r indicates the beamlet row and (c_1, c_2) represent the rightmost beamlet blocked by the left leaf and the leftmost beamlet blocked by the right leaf in row r ($r = 1, \dots, m$; $c_1 = 0, \dots, n$; and $c_2 = 1, \dots, n+1$ with $c_1 < c_2$). (If $c_1 = 0$, the left leaf blocks no beamlets, and if $c_2 = n+1$, the right leaf blocks no beamlets.) In addition, we define source node, say 0, and sink node, say $(m+1, 0, 1)$, representing the top (“beamlet row 0”) and bottom (“beamlet row $m+1$ ”) of the aperture. We then define arcs from all nodes in beamlet row r to all nodes in a beamlet row $r+1$ ($r = 0, \dots, m$). Figure 5 illustrates the structure of this network for a small case of $m=2$ rows and $n=2$ columns.

Now note that there is a one-to-one correspondence between paths from the source node 0 to the sink node $(m+1, 0, 1)$ and deliverable apertures. Without loss of generality, assume that the tongue between rows r and $r+1$ are part of the latter and hence partially block the former, i.e., row r . Next, we assign a cost to each arc, which is defined as the cost of the exposed beamlets corresponding to the origin node of the arc. In particular, consider nodes (r, c_1, c_2) and $(r+1, c'_1, c'_2)$. Then c_1+1, \dots, c_2-1 are the exposed beamlets at row r and $\{c_1+1, \dots, c'_1\} \cup \{c'_2, \dots, c_2-1\}$ are the ones among those that expose a tongue due to the leaf setting in row $r+1$. [Note that if $c_1 \geq c'_1$, the former set is empty (i.e., the left leaf of row $r+1$ does not expose a tongue), and similarly, if $c_2 \leq c'_2$, the latter set is empty (i.e., the right leaf of row $r+1$ does not expose a tongue) (see Fig. 6).] Representing the beamlet in row r and column c by (r, c) , the cost associated with the arc from node (r, c_1, c_2) to node $(r+1, c'_1, c'_2)$ is equal to

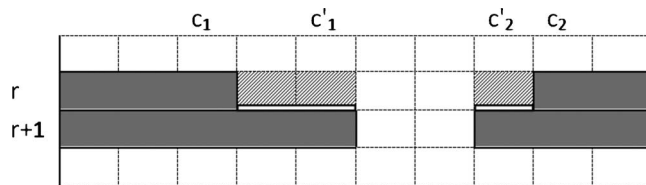


FIG. 6. Leaf settings at two adjacent rows.

$$\begin{aligned} & \sum_{c=c_1+1}^{c'_1} \left(\sum_{j \in V} D_{(r,c)j}^{\delta} \hat{\pi}_j + \sum_{j \in V} D_{(r,c)j} \hat{\pi}_j \right) \\ & + \sum_{c=c'_2}^{c_2-1} \left(\sum_{j \in V} D_{(r,c)j}^{\delta} \hat{\pi}_j + \sum_{j \in V} D_{(r,c)j} \hat{\pi}_j \right) \\ & + \sum_{c=\max\{c_1, c'_1\}+1}^{\min\{c_2, c'_2\}-1} \left(\sum_{j \in V} D_{(r,c)j}^{\delta} \hat{\pi}_j + \sum_{j \in V} D_{(r,c)j} \hat{\pi}_j \right) \end{aligned}$$

(for $r = 1, \dots, m$). Finally, we assign a cost of zero to the arcs from the source node 0 to the nodes corresponding to beamlet row 1. The length of any path from the source node to the sink node is then equal to the objective function value in the pricing problem of the corresponding aperture. Therefore, the optimal solution to the pricing problem can be found by solving a shortest path problem in the network described above. Since the network is acyclic, this problem can be solved in an amount of time that is proportional to the number of arcs in the network. Since it is easy to see that the total number of nodes in the network is $O(mn^2)$ and the total number of arcs in this network is $O(mn^4)$, the problem can be solved in $O(mn^4)$ time (see, e.g., Ref. 54).

Finally, we note that under interdigitation constraints (i.e., the left leaf of a row cannot overlap with the right leaf of an adjacent row) and connectedness constraints (i.e., the rows in which at least one beamlet is exposed are consecutive, relevant if the left and right leaves cannot entirely block a beamlet row and backup jaws are required), the pricing problem can be solved in a very similar way (and with the same running time) using slight modifications of the algorithms for these cases presented in Ref. 40.

^{a)}Electronic mail: esalari@ufl.edu

¹C. S. Chui, T. LoSasso, and S. Spirou, "Dose calculation for photon beams with intensity modulation generated by dynamic jaw or multileaf collimations," *Med. Phys.* **21**, 1237–1244 (1994).

²J. Deng, T. Pawlicki, Y. Chen, J. Li, S. B. Jiang, and C. M. Ma, "The MLC tongue-and-groove effect on IMRT dose distributions," *Phys. Med. Biol.* **46**(4), 1039–1060 (2001).

³J. M. Galvin, A. R. Smith, and B. Lally, "Characterization of a multileaf collimator system," *Int. J. Radiat. Oncol., Biol., Phys.* **25**, 181–192 (1993).

⁴R. Mohan, "Field shaping for three-dimensional conformal radiation therapy and multileaf collimation," *Semin. Radiat. Oncol.* **5**, 86–99 (1995).

⁵S. Luan, C. Wang, D. Z. Chen, X. S. Hu, S. A. Naqvi, X. Wu, and C. X. Yu, "An improved MLC segmentation algorithm and software for step-and-shoot IMRT delivery without tongue-and-groove error," *Med. Phys.* **33**, 1199–1212 (2006).

⁶R. A. C. Siochi, "Optimized removal of the tongue-and-groove underdose via constrained partial synchronization and variable depth recursion,"

Phys. Med. Biol. **54**, 1369–1381 (2009).

⁷J. R. Sykes and P. C. Williams, "An experimental investigation of the tongue and groove effect for the Philips multileaf collimator," *Phys. Med. Biol.* **43**, 3157–3165 (1998).

⁸X. Wang, S. Spirou, T. LoSasso, J. Stein, C. S. Chui, and R. Mohan, "Dosimetric verification of intensity-modulated fields," *Med. Phys.* **23**, 317–327 (1996).

⁹D. M. Shepard, M. C. Ferris, G. H. Olivera, and T. Rockwell Mackie, "Optimizing the delivery of radiation therapy to cancer patients," *SIAM Rev.* **41**, 721–744 (1999).

¹⁰H. E. Romeijn and J. F. Dempsey, "Intensity modulated radiation therapy treatment plan optimization," *TOP* **16**(2), 215–243 (2008).

¹¹E. K. Lee, T. Fox, and I. Crocker, "Optimization of radiosurgery treatment planning via mixed integer programming," *Med. Phys.* **27**, 995–1004 (2000).

¹²E. K. Lee, T. Fox, and I. Crocker, "Integer programming applied to intensity-modulated radiation therapy treatment planning," *Ann. Operat. Res.* **119**(1–4), 165–181 (2003).

¹³H. W. Hamacher and K. H. Küfer, "Inverse radiation therapy planning—A multiple objective optimization approach," *Discrete Appl. Math.* **118**(1–2), 145–161 (2002).

¹⁴K. H. Küfer, A. Scherrer, M. Monz, F. Alonso, H. Trinkaus, T. Bortfeld, and C. Thieke, "Intensity-modulated radiotherapy—A large scale multicriteria programming problem," *OR Spectrum* **25**(2), 223–249 (2003).

¹⁵H. E. Romeijn, R. K. Ahuja, J. F. Dempsey, and A. Kumar, "A novel linear programming approach to fluence map optimization in intensity modulated radiation therapy treatment planning," *Phys. Med. Biol.* **48**(21), 3521–3542 (2003).

¹⁶H. E. Romeijn, R. K. Ahuja, J. F. Dempsey, and A. Kumar, "A new linear programming approach to radiation therapy treatment planning problems," *Oper. Res.* **54**(2), 201–216 (2006).

¹⁷R. K. Ahuja and H. W. Hamacher, "A network flow algorithm to minimize beam-on time for unconstrained multileaf collimator problems in cancer radiation therapy," *Networks* **45**(1), 36–41 (2005).

¹⁸T. R. Bortfeld, D. L. Kahler, T. J. Waldron, and A. L. Boyer, "X-ray field compensation with multileaf collimators," *Int. J. Radiat. Oncol., Biol., Phys.* **28**(3), 723–730 (1994).

¹⁹S. Kamath, S. Sahni, J. Li, J. Palta, and S. Ranka, "Leaf sequencing algorithms for segmented multileaf collimation," *Phys. Med. Biol.* **48**(3), 307–324 (2003).

²⁰R. A. C. Siochi, "Minimizing static intensity modulation delivery time using an intensity solid paradigm," *Int. J. Radiat. Oncol., Biol., Phys.* **43**(3), 671–680 (1999).

²¹D. Baatar, H. W. Hamacher, M. Ehrgott, and G. J. Woeginger, "Decomposition of integer matrices and multileaf collimator sequencing," *Discrete Appl. Math.* **152**(1–3), 6–34 (2005).

²²N. Boland, H. W. Hamacher, and F. Lenzen, "Minimizing beam-on time in cancer radiation treatment using multileaf collimators," *Networks* **43**(4), 226–240 (2004).

²³S. Kamath, S. Sahni, J. Palta, and S. Ranka, "Algorithms for optimal sequencing of dynamic multileaf collimators," *Phys. Med. Biol.* **49**, 33–54 (2004).

²⁴J. R. Dai and Y. M. Hu, "Intensity-modulation radiotherapy using independent collimators: An algorithm study," *Med. Phys.* **26**, 2562–2570 (1999).

²⁵J. Dai and Y. Zhu, "Minimizing the number of segments in a delivery sequence for intensity-modulated radiation therapy with a multileaf collimator," *Med. Phys.* **28**, 2113–2120 (2001).

²⁶P. Xia and L. J. Verhey, "Multileaf collimator leaf sequencing algorithm for intensity modulated beams with multiple static segments," *Med. Phys.* **25**, 1424–1434 (1998).

²⁷K. Engel, "A new algorithm for optimal multileaf collimator field segmentation," *Discrete Appl. Math.* **152**(1–3), 35–51 (2005).

²⁸T. Kalinowski, "Reducing the number of monitor units in multileaf collimator field segmentation," *Phys. Med. Biol.* **50**, 1147–1161 (2005).

²⁹T. Kalinowski, "A duality based algorithm for multileaf collimator field segmentation with interleaf collision constraint," *Discrete Appl. Math.* **152**(1–3), 52–88 (2005).

³⁰G. J. Lim and J. Choi, "A two-stage integer programming approach for optimizing leaf sequence in IMRT," Technical Report No. IE0807-01 (University of Houston, 2007).

³¹Z. C. Taşkın, J. C. Smith, H. E. Romeijn, and J. F. Dempsey, "Optimal multileaf collimator leaf sequencing in IMRT treatment planning," *Oper.*

- Res. **58**(3), 674–690 (2010).
- ³²J. P. C. Van Santvoort and B. J. M. Heijmen, “Dynamic multileaf collimation without ‘tongue-and-groove’ underdosage effects,” *Phys. Med. Biol.* **41**, 2091–2105 (1996).
- ³³S. Webb, T. Bortfeld, J. Stein, and D. Convery, “The effect of stair-step leaf transmission on the ‘tongue-and-groove problem’ in dynamic radiotherapy with a multileaf collimator,” *Phys. Med. Biol.* **42**, 595–602 (1997).
- ³⁴W. Que, J. Kung, and J. Dai, “Tongue-and-groove effect in intensity modulated radiotherapy with static multileaf collimator fields,” *Phys. Med. Biol.* **49**(3), 399–405 (2004).
- ³⁵S. Kamath, S. Sahni, J. Palta, S. Ranka, and J. Li, “Optimal leaf sequencing with elimination of tongue-and-groove underdosage,” *Phys. Med. Biol.* **49**(3), N7–N17 (2004).
- ³⁶S. Kamath, S. Sahni, S. Ranka, J. Li, and J. Palta, “A comparison of step-and-shoot leaf sequencing algorithms that eliminate tongue-and-groove effects,” *Phys. Med. Biol.* **49**(14), 3137–3143 (2004).
- ³⁷D. M. Shepard, M. A. Earl, X. A. Li, S. Naqvi, and C. Yu, “Direct aperture optimization: A turnkey solution for step-and-shoot IMRT,” *Med. Phys.* **29**, 1007–1018 (2002).
- ³⁸G. Bednarz, D. Michalski, C. Houser, M. S. Huq, Y. Xiao, P. R. Anne, and J. M. Galvin, “The use of mixed-integer programming for inverse treatment planning with pre-defined field segments,” *Phys. Med. Biol.* **47**, 2235–2245 (2002).
- ³⁹F. Preciado-Walters, M. P. Langer, R. L. Rardin, and V. Thai, “Column generation for IMRT cancer therapy optimization with implementable segments,” *Ann. Operat. Res.* **148**(1), 65–79 (2006).
- ⁴⁰H. E. Romeijn, R. K. Ahuja, J. F. Dempsey, and A. Kumar, “A column generation approach to radiation therapy treatment planning using aperture modulation,” *SIAM J. Optim.* **15**, 838–862 (2005).
- ⁴¹C. Men, H. E. Romeijn, Z. C. Taskin, and J. F. Dempsey, “An exact approach to direct aperture optimization in IMRT treatment planning,” *Phys. Med. Biol.* **52**(24), 7333–7352 (2007).
- ⁴²M. Earl, D. Cao, and D. Shepard, “SU-FF-T-435: Tongue and groove effect in direct aperture optimization IMRT Plans,” *Med. Phys.* **33**, 2145 (2006).
- ⁴³B. E. Bjärngard and R. L. Siddon, “A note on equivalent circles, squares, and rectangles,” *Med. Phys.* **9**, 258–260 (1982).
- ⁴⁴J. R. Clarkson, “A note on depth doses in fields of irregular shape,” *Br. J. Radiol.* **14**, 265–268 (1941).
- ⁴⁵D. E. Sanz, “A mathematical study of the area over perimeter rule using the sector-integration equation,” *Med. Phys.* **27**(10), 2376–2379 (2000).
- ⁴⁶A. Boyer, P. Biggs, J. Galvin, E. Klein, T. LoSasso, D. Low, K. Mah, and C. Yu, “Basic applications of multileaf collimators,” AAPM Report No. 72 (Task Group No. 50: Radiation Therapy Committee, 2001).
- ⁴⁷Varian, Millenium MLC Specifications, www.varian.com.
- ⁴⁸H. E. Romeijn, J. F. Dempsey, and J. G. Li, “A unifying framework for multi-criteria fluence map optimization models,” *Phys. Med. Biol.* **49**, 1991–2013 (2004).
- ⁴⁹A. Ben-Tal and A. Nemirovski, “Robust convex optimization,” *Math. Oper. Res.* **23**(4), 769–805 (1998).
- ⁵⁰A. Ben-Tal and A. Nemirovski, “Robust optimization—Methodology and applications,” *Math. Program.* **92**(3), 453–480 (2002).
- ⁵¹D. M. Aleman, D. Glaser, H. E. Romeijn, and J. F. Dempsey, “Interior point algorithms: guaranteed optimality for fluence map optimization in IMRT,” *Phys. Med. Biol.* **55**, 5467–5482 (2010).
- ⁵²M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 2nd ed. (Wiley, New York, 1993).
- ⁵³J. B. Hiriart-Urruty, “On optimality conditions in nondifferentiable programming,” *Math. Program.* **14**, 73–86 (1978).
- ⁵⁴R. K. Ahuja, T. L. Magnanti, J. B. Orlin, and K. Weihe, *Network Flows: Theory, Algorithms, and Applications* (Prentice-Hall, Englewood Cliffs, NJ, 1993).