Sensitivity analysis for lexicographic ordering in radiation therapy treatment planning

T. Long[a)]
Department of Industrial and Operations Engineering, University of Michigan, 1205 Beal Avenue, Ann Arbor, Michigan 48109-2117

M. Matuszak[b)] and M. Feng[c)]
Department of Radiation Oncology, University of Michigan Health System, 1500 East Medical Center Drive, Ann Arbor, Michigan 48109-5010

B. A. Fraass[d)]
Department of Radiation Oncology, University of Michigan Health System, 1500 East Medical Center Drive, Ann Arbor, Michigan 48109-5010 and Department of Radiation Oncology, Cedars-Sinai Medical Center, 8700 Beverly Boulevard, Los Angeles, California 90048

R. K. Ten Haken[e)]
Department of Radiation Oncology, University of Michigan Health System, 1500 East Medical Center Drive, Ann Arbor, Michigan 48109-5010

H. E. Romeijn[f)]
Department of Industrial and Operations Engineering, University of Michigan, 1205 Beal Avenue, Ann Arbor, Michigan 48109-2117

(Received 11 February 2012; revised 5 April 2012; accepted for publication 28 April 2012; published 25 May 2012)

Purpose: To introduce a method to efficiently identify and calculate meaningful tradeoffs between criteria in an interactive IMRT treatment planning procedure. The method provides a systematic approach to developing high-quality radiation therapy treatment plans.
Methods: Treatment planners consider numerous dosimetric criteria of varying importance that, when optimized simultaneously through multicriteria optimization, yield a Pareto frontier which represents the set of Pareto-optimal treatment plans. However, generating and navigating this frontier is a time-consuming, nontrivial process. A lexicographic ordering (LO) approach to IMRT uses a physician's criteria preferences to partition the treatment planning decisions into a multistage treatment planning model. Because the relative importance of criteria optimized in the different stages may not necessarily constitute a strict prioritization, the authors introduce an interactive process, sensitivity analysis in lexicographic ordering (SALO), to allow the treatment planner control over the relative sequential-stage tradeoffs. By allowing this flexibility within a structured process, SALO implicitly restricts attention to and allows exploration of a subset of the Pareto efficient frontier that the physicians have deemed most important.
Results: Improvements to treatment plans over a LO approach were found by implementing the SALO procedure on a brain case and a prostate case. In each stage, a physician assessed the tradeoff between previous stage and current stage criteria. The SALO method provided critical tradeoff information through curves approximating the relationship between criteria, which allowed the physician to determine the most desirable treatment plan.
Conclusions: The SALO procedure provides treatment planners with a directed, systematic process to treatment plan selection. By following a physician's prioritization, the treatment planner can avoid wasting effort considering clinically inferior treatment plans. The planner is guided by criteria importance, but given the information necessary to accurately adjust the relative importance at each stage. Through these attributes, the SALO procedure delivers an approach well balanced between efficiency and flexibility.
[http://dx.doi.org/10.1118/1.4720218]

Key words: radiation therapy treatment planning, multicriteria optimization, Pareto surface navigation

I. INTRODUCTION

When addressing a radiation therapy case, a physician generally presents the treatment planner with a number of dosimetric goals of varying importance. While the general objective is to deliver a prescribed radiation dose to the target(s) while simultaneously sparing critical structures, a major challenge remains how to make the unavoidable tradeoffs between these conflicting goals. The literature on radiation therapy treatment planning as well as clinical treatment planning systems

contains a multitude of evaluation criteria that can be used to quantify various properties of a treatment plan. Because treatment planning is generally a time-consuming endeavor which has to be performed for individual patients, providing a treatment planner with tools that allow for an efficient assessment of the interplay and tradeoffs between conflicting treatment plan evaluation criteria is essential to an efficient and effective treatment planning process.

Traditionally, radiation therapy treatment planning is based on optimization models containing a single objective function to be optimized subject to a set of hard constraints on the treatment plan. The objective function is typically a simple weighted sum of individual treatment plan evaluation criteria (see Ref. 1). Since there is no formal basis for choosing a priori values for these weights, their values are usually updated manually by the treatment planner in an iterative fashion in order to arrive at a clinically desirable treatment plan. Occasionally this method yields acceptable results quickly, but in general this approach is inefficient and may lead to inferior treatment plans.

A modern technique for exploring the tradeoffs between treatment plan evaluation criterion is based on multicriteria optimization (MCO) (see, e.g., Ref. 2 for a recent overview of this area). In this approach, the goal is to approximate the Pareto surface containing all efficient treatment plans, i.e., treatment plans with the property that it is not possible to improve the plan with respect to one of the criteria without deteriorating the plan with respect to at least one other. While there are many methods for generating this surface (see, e.g., Ref. 3), a common technique is to solve a sequence of single-objective optimization problems, each using an appropriately chosen set of weights for the individual criteria. When all criteria are convex functions of the dose distribution delivered to the patient, each of the corresponding solutions will represent a point on the Pareto surface. If the number of such solutions is large enough to allow the Pareto frontier to be accurately approximated (typically using interpolation), the treatment planner can assess the tradeoffs between competing objectives by navigating the frontier and use this information to select a treatment plan. Using MCO as a means of quantifying tradeoffs is conceptually attractive, in the sense that it provides the treatment planner with complete and comprehensive tradeoff information on all criteria. However, the number of competing criteria can be large (say on the order of 10–25 in a typical clinical setting), which means that the Pareto frontier is embedded in a correspondingly high-dimensional space. Many solutions may then be required to accurately approximate the Pareto frontier, which reduces its efficiency (see Ref. 4). Moreover, visualizing and interpreting the plethora of tradeoffs can prove difficult (see Ref. 5). Of course a reduction in the number of criteria or data reduction in the form of a coarser representation of patient geometry and/or capabilities of the delivery equipment may mitigate these drawbacks, but this may affect the accuracy of the frontier or the quality of the tradeoff information (see, e.g., Ref. 6).

A key observation is that the full Pareto frontier identified by MCO will likely contain many tradeoff regions that are clinically unacceptable or irrelevant. This not only complicates the navigation process as outlined above but it also means that a large amount of time may be spent identifying such uninteresting tradeoffs. It therefore seems appropriate to explicitly incorporate better *a priori* clinical information on priorities associated with the different criteria into the treatment plan optimization process. One such approach is lexicographic optimization (LO), which is sometimes also referred to as prioritized optimization (see Refs. 7–9). This is a multistage approach that is based on a complete ranking or prioritization of treatment planning goals. In its purest form it starts by optimizing the highest ranked criterion. The optimal value to this problem is then used to constrain the value of the corresponding criterion in subsequent optimization models. In particular, in the following stage the second criterion on the prioritized list is optimized subject to the value of the first criterion being optimal. This approach is then repeated for each criterion on the list, and the solution to the final optimization problem in the sequence is the optimal treatment plan with respect to the prioritized list of criteria. LO is computationally efficient and provides a clear, systematic approach. In contrast with MCO, LO does not rely on interaction with the treatment planner (once the prioritization is fixed). However, much flexibility is sacrificed in the wake of the computational and structural benefits. In particular, a notable drawback of using an LO approach is that the treatment planner may be unaware of opportunities that may exist to improve a treatment plan. In terms of MCO, the LO approach can be interpreted as confining the treatment planner's view to a specific extreme solution on the full Pareto frontier of inter-criterion tradeoffs. If a minor sacrifice in high-priority criteria could yield meaningful benefits with respect to lower-priority criteria, the pure LO approach would not recognize or identify this opportunity. In order to introduce some flexibility into the process one might relax the optimality constraint on high-priority criteria and instead require previously optimized criteria to remain "near-optimal." Since tradeoffs are not characterized and assessed explicitly, it is not clear how to quantify the concept of near-optimality nor how to predict the consequences of allowing a deviation from optimality. In contrast, our method will provide an interactive way for the user to select the relaxation based on a formal sensitivity analysis.

In this paper, we propose a systematic approach, sensitivity analysis in lexicographic ordering (SALO), which combines the benefits of MCO (flexibility and comprehensiveness) and LO (efficiency and clinical focus) while avoiding their pitfalls. Similar to LO, it incorporates clinical information through a prioritized list of treatment plan evaluation criteria. However, in contrast with LO, it uses this information to, in an interactive and iterative fashion, efficiently navigate the clinically interesting and relevant segment of the Pareto efficient frontier. In Sec. II of this paper, we will provide a formal and detailed description of the SALO approach. In Sec. III, we will then illustrate the approach on two clinical cases and discuss SALO. In Sec. IV, we will discuss some implementation characteristics and conclude the paper.

II. METHODS AND MATERIALS

The goal of the SALO approach is to provide local information on the shape of the Pareto frontier to treatment planners for use as a decision making aid, based on clinical preferences represented via a prioritized list of treatment plan evaluation criteria. This local information takes the form of a two-dimensional Pareto frontier that, in each stage, characterizes the tradeoff between two consecutive criteria while (i) constraining higher priority criteria to values that have been established earlier in the process and (ii) temporarily ignoring lower priority criteria. The treatment planner can then examine this tradeoff curve and select a point that appropriately captures the tradeoff between the two criteria currently under consideration. This point then defines an upper bound for the criterion that has the higher priority.

II.A. Notation and model

Optimization models for radiation therapy treatment planning are usually classified as "beamlet-based" (yielding an optimal fluence map, which subsequently needs to be converted into a deliverable plan in a leaf-sequencing stage) and "aperture-based." We have chosen to use the latter, direct aperture optimization (DAO), approach (see, e.g., Refs. 10–13) since it not only eliminates the need for a leaf-sequencing stage but can also allow for a more efficient implementation and solution since an instance of the DAO model is typically much smaller and hence can be solved more rapidly than an instance of a beamlet-based fluence map optimization (FMO) problem. This is particularly important since many of these problems will need to be solved during the course of the SALO procedure. However, if desired the general SALO approach can be based on a more traditional FMO model with only minor modifications.

As is traditionally done in radiation therapy treatment planning, we discretize the relevant patient geometry into a finite set of voxels $\mathcal{V}$. Moreover, we assume that there is a finite set of deliverable apertures $\mathcal{K}$. The dose delivered to voxel $j \in \mathcal{V}$ from aperture $k \in \mathcal{K}$ at unit intensity is given by D_{kj}, which we will refer to as aperture-based dose deposition coefficients. The decision variables are the aperture intensities, denoted by y_k $(k \in \mathcal{K})$. For convenience, we will add as decision variables the dose z_j delivered to each voxel $j \in \mathcal{V}$. For convenience we will let $y = (y_k : k \in \mathcal{K})$ and $z = (z_j : j \in \mathcal{V})$ denote the corresponding vectors. Finally, let Z denote a convex set that excludes all clinically unacceptable treatment plans. Convexity of this set is important for tractability of our approach, and we usually expect this set to contain only simple lower and upper bound constraints on the individual voxel doses. In principle other hard constraints on treatment plan evaluation criteria could be included as well, although we envision those tradeoffs to be made in the actual SALO procedure rather than by *a priori* excluding certain dose distributions.

The treatment plan evaluation criteria are given as functions of the dose distribution: $G_\ell : \mathbb{R}^{|\mathcal{V}|} \to \mathbb{R}\,(\ell = 1, \ldots, L+1)$, where we assume that the criteria are indexed in order of decreasing priority. For mathematical convenience we will assume that these criteria are such that smaller values are preferred to larger values. Of course the essence of our approach could be generalized to cases where this assumption is violated, and in fact our examples in Sec. III will include criteria for which larger values are preferred. In addition, we will generally assume that they are all convex functions. Note that, in a multi-criteria context, many common treatment plan evaluation criteria, such as voxel-based penalty functions, (generalized) equivalent uniform dose (EUD, gEUD), tumor control probabilities (TCP), normal-tissue complication probabilities (NTCP), or conditional value-at-risk (CVaR), are either convex or can equivalently be replaced by convex ones (see Ref. 14). Our proposed approach could in principle be generalized to accommodate a nonconvex set Z and/or truly nonconvex criteria, such as traditional dose-volume histogram (DVH) constraints, albeit at the expense of computational efficiency. The last criterion, G_{L+1}, is typically chosen in order to minimize total dose delivered to the patients while maintaining treatment plan quality with respect to all previously considered criteria. The SALO approach then interactively searches for a treatment plan by solving a sequence of bicriteria optimization models of the following form (for $\ell = 1, \ldots, L$) referred to as *stages* of the procedure:

$$
\begin{aligned}
&\text{minimize } \{G_\ell(z),\ G_{\ell+1}(z)\} \\
&\text{subject to} \qquad\qquad (\mathrm{P}^{(\ell)}) \\
&z_j = \sum_{k \in \mathcal{K}} D_{kj} y_k \quad \text{for} \quad j \in \mathcal{V} \\
&G_{\ell'}(z) \le \bar{G}_{\ell'} \quad \text{for} \quad \ell' = 1, \ldots, \ell - 1 \\
&y_k \ge 0 \quad \text{for} \quad k \in \mathcal{K} \\
&z \in Z,
\end{aligned}
$$

where $\bar{G}_{\ell'}$ is an upper bound on treatment plan evaluation criterion $G_{\ell'}$ that is set by solving the prior bicriteria optimization problem $(\mathrm{P}^{(\ell')})$ (for $\ell' = 1, \ldots, \ell - 1$) and making the corresponding tradeoff.

Due to the convexity of the criterion functions, the solution to the bicriteria optimization problem $(\mathrm{P}^{(\ell')})$ can be found by solving single-criterion optimization problems with an objective function of the form

$$\alpha G_\ell(z) + (1 - \alpha) G_{\ell+1}(z) \tag{1}$$

for all $\alpha \in [0, 1]$.

II.B. Sensitivity analysis in lexicographic optimization

If the set of all deliverable apertures is manageable, we could directly apply the approach outlined above. Unfortunately, in general the cardinality of the set $\mathcal{K}$ is very large and the optimization problems $(\mathrm{P}^{(\ell')})$ cannot be solved explicitly. One potential approach would be to generate high-quality apertures "on the fly" according to a column generation approach that has been proposed for solving single-criterion DAO problems (see, e.g., Refs. 11–13). However, this would mean that the set of apertures considered in later stages of the algorithm is different from (in fact, larger than) the set of apertures allowed in earlier stages. This means that the tradeoffs between the higher priority,

and hence clinically more important, criteria are based made on a more limited set of apertures. Intuitively it would seem more attractive to base the more important (or, in fact, and all) tradeoff decisions on the most accurate representation of the optimization model rather than the least accurate, which makes a straightforward application of this idea undesirable.

In order to address this issue we propose to start the SALO procedure with an initial phase in which a high-quality pool of apertures is generated, which is then kept fixed throughout the L stages of the SALO procedure. This does not only improve the computational efficiency of the approach, but also ensures that all decisions are made based on consistent input and information. However, it is clear that, in this process, the tradeoff decisions are not made with respect to the full information regarding all deliverable apertures. We therefore also propose a final phase in which a full DAO model is solved to identify a new set of apertures that minimizes the last criterion, G_{L+1}, subject to all bounds imposed on criteria $G_1, \ldots, G_L$. Clearly, this final phase could also take other considerations, such as treatment delivery efficiency, into account. In summary, we propose a SALO procedure that proceeds in three phases:

Phase 1. Generation of a clinically relevant aperture pool of computationally manageable cardinality.

Phase 2. Generation of patient-specific treatment planning goals $\bar{G}_\ell$ $(\ell = 1, \ldots, L)$ by solving a sequence of bicriteria optimization problems $(\mathrm{P}^{(1)}), \ldots, (\mathrm{P}^{(L-1)})$.

Phase 3. Generation of final treatment plan that satisfies the patient-specific treatment planning goals while minimizing an overall single objective function.

In the remainder of this section, we will discuss these three stages in more detail.

II.B.1. Aperture pool generation

We generate an aperture pool by solving a traditional single-criterion treatment plan optimization model based on the treatment plan evaluation criteria

$$\begin{aligned} &\text{minimize} \sum_{\ell=1}^{L} \alpha_\ell G_\ell(z) \\ &\text{subject to} \qquad\qquad (\mathrm{P}) \\ &z_j = \sum_{k \in \mathcal{K}} D_{kj} y_k \quad \text{for} \quad j \in \mathcal{V} \\ &y_k \geq 0 \qquad\quad \text{for} \quad k \in \mathcal{K} \\ &z \in Z, \end{aligned}$$

where $\alpha_\ell \geq 0$ $(\ell = 1, \ldots, L)$ are nonnegative criterion weights. The set of weights used in the aperture generation phase could either be based on experience with other, similar, patient cases. Alternatively, we could use a sequence of criterion weights, allowing for the generation of apertures that are attractive with respect to a variety of tradeoffs. For convenience, we will denote the set of apertures in the pool by $\overline{\mathcal{K}}$.

II.B.2. Solving a bicriteria optimization problem

We use the so-called Sandwich Algorithm (see, e.g., Refs. 2, 15, and 16) to approximate the Pareto frontier at a given stage of the SALO procedure. This algorithm, which tries to balance clinical accuracy and computational efficiency, applies when all treatment plan evaluation criteria are convex, and is particularly efficient in the bicriteria case. The idea behind this algorithm is to approximate the entire Pareto frontier by constructing both an upper (conservative) and a lower (optimistic) bound on the frontier based on a finite set of points on the frontier. This is done by solving a sequence of optimization problems of the form $(\mathrm{P}^{(\ell)})$ with objective function of the form (Sec. II.A) for different values of α. The optimal solutions to these problems yield points on the Pareto frontier. For convenience, let $z^*(\alpha; \ell)$ denote an optimal solution to $(\mathrm{P}^{(\ell)})$ when parameter α is used. Then, let

$$\begin{aligned} &G^*_\ell(\alpha; \ell) = G_\ell(z^*(\alpha; \ell)), \\ &G^*_{\ell+1}(\alpha; \ell) = G_{\ell+1}(z^*(\alpha; \ell)), \\ &G^*_{\ell,\ell+1}(\alpha; \ell) = \alpha G^*_\ell(\alpha; \ell) + (1-\alpha) G^*_{\ell+1}(\alpha; \ell). \end{aligned}$$

The Sandwich algorithm then determines upper and lower bounds on the Pareto frontier as follows.

II.B.2.a. Upper bound. Using simple linear interpolation of a set of Pareto efficient solutions, we obtain a piecewise-linear and convex function which is well-known to form an upper bound on the Pareto frontier. This follows immediately from the fact that the line segment connecting any two points of the form $\left(G^*_\ell(\alpha; \ell), G^*_{\ell+1}(\alpha; \ell)\right)$ for different values of α is guaranteed to be entirely on or above the Pareto frontier. More formally, such a line segment can be characterized as

$$\begin{aligned} \{&(\lambda G^*_\ell(\alpha; \ell) + (1-\lambda) G^*_\ell(\alpha'; \ell), \lambda G^*_{\ell+1}(\alpha; \ell) \\ &+ (1-\lambda) G^*_{\ell+1}(\alpha'; \ell)) : \lambda \in [0,1]\}, \end{aligned} \tag{2}$$

where $0 < \alpha \neq \alpha' < 1$. It is interesting to note that we can find an even better bound on the Pareto frontier by, instead of interpolating the pairs of optimal objective function values for different values of α, interpolating the optimal treatment plans (or, equivalently, and optimal dose distributions) for different values of α. In other words, the curve of the form

$$\begin{aligned} \{&(G_\ell(\lambda z^*(\alpha; \ell) + (1-\lambda) z^*(\alpha'; \ell)) G_{\ell+1}(\lambda z^*(\alpha; \ell) \\ &+ (1-\lambda) z^*(\alpha'; \ell))) : \lambda \in [0,1]\}, \end{aligned}$$

where $0 < \alpha \neq \alpha' < 1$ is guaranteed to not only be entirely on or above the Pareto frontier, but also entirely on or below the curve in Eq. (1).

II.B.2.b. Lower bound. A lower bound can be determined by observing that the line given by

$$\alpha g_\ell + (1-\alpha) g_{\ell+1} = G^*_{\ell,\ell+1}(\alpha; \ell) \tag{3}$$

is entirely on or below the Pareto frontier [where $(g_\ell, g_{\ell+1})$ denotes a point in $\mathbb{R}^2$]. This follows since (i) the Pareto frontier is convex by convexity of the criteria functions G_ℓ and $G_{\ell+1}$ and (ii) the point $\left(G^*_\ell(\alpha; \ell), G^*_{\ell+1}(\alpha; \ell)\right)$ lies both on the efficient frontier and the line (2). This means that the upper envelope of these lines over a collection of different values of $\alpha \in (0,1)$ is a piecewise-linear convex function that is entirely on or below the Pareto frontier as well.

II.B.2.c. Choosing objective function weights α. There are different ways in which the set of values for α to be used at a particular SALO stage can be determined. In an interactive implementation, the treatment planner could indicate which value to use with the goal of refining the approximation of the Pareto frontier in the clinically most relevant or interesting areas. In an automated setting this can be done by measuring the discrepancy between the upper and lower bounds, and choosing that value of α where the discrepancy is largest. Since the bounds are themselves curves, different discrepancy measures can be used, and each of them will yield a different sequence of values for α and a different bounding of the frontier. However, when designed carefully, we can ensure that the lower and upper bounds both converge to the Pareto frontier as the number of values for α increases (see Ref. 17).

II.B.3. Final treatment plan optimization

This full model clearly has a feasible solution (by construction), i.e., by generating a new set of apertures from scratch we know it will be possible to achieve all previously identified treatment planning goals. The DAO column generation procedure can therefore be initialized with the final solution obtained by the SALO procedure. However, if there is an additional goal related to treatment plan delivery efficiency, such as limiting the beam-on-time or number of apertures, it may be preferable to discard the original aperture pool and start the DAO algorithm from scratch. Since a feasible solution is needed to start the procedure, the algorithm is then started by first optimizing an auxiliary problem of the following form:

$$\begin{aligned} &\text{minimize} \sum_{\ell=1}^{L} \max\{G_\ell(z) - \bar{G}_\ell, 0\} \\ &\text{subject to} \qquad\qquad\qquad\qquad\qquad\qquad (\text{I}) \\ &z_j = \sum_{k\in\mathcal{K}} D_{kj} y_k \qquad \text{for} \quad j \in \mathcal{V} \\ &y_k \geq 0 \qquad\qquad\quad \text{for} \quad k \in \mathcal{K} \\ &z \in Z. \end{aligned}$$

Any feasible solution to this problem with objective function value 0 is a feasible solution to the actual problem from which the DAO algorithm can be started.

II.C. Treatment plan evaluation criteria

For our experiments we have chosen to use a single measure of generalized equivalent uniform dose (gEUD) for each target and major critical structure as our main treatment plan evaluation criteria. Letting $\mathcal{S}$ denote the set of structures and $\mathcal{V}_s$ the set of voxels in structure $s \in \mathcal{S}$, the gEUD corresponding to the dose distribution in structure $s \in \mathcal{S}$ is given by

$$\text{gEUD}_s(z; a_s) = \left(\frac{1}{|\mathcal{V}_s|} \sum_{j\in\mathcal{V}_s} z_j^{a_s} \right)^{\frac{1}{a_s}},$$

where $1 \leq a_s \leq \infty$ if s is a critical structure while $-\infty \leq a_s \leq 1$ if s is a target, and where

$$\text{gEUD}_s(z; a_s) = \begin{cases} \max_{j\in\mathcal{V}_s} z_j & \text{if} \quad a_s = \infty \\ \frac{1}{|\mathcal{V}_s|} \sum_{j\in\mathcal{V}_s} z_j & \text{if} \quad a_s = 1 \\ \min_{j\in\mathcal{V}_s} z_j & \text{if} \quad a_s = -\infty \end{cases}$$

(see, e.g., Refs. 18 and 19). For the sake of computational efficiency, we have chosen to use an approximation of gEUD given by a convex combination of mean and maximum dose for critical structures and of mean and minimum dose for targets (see Refs. 20 and 21). In particular, we choose treatment plan evaluation criteria functions G_ℓ of the form

$$\begin{aligned} &\gamma_s \text{gEUD}_s(z; 1) + (1 - \gamma_s)\text{gEUD}_s(z; \infty) \\ &\quad = \gamma_s \frac{1}{|\mathcal{V}_s|} \sum_{j\in\mathcal{V}_s} z_j + (1 - \gamma_s) \max_{j\in\mathcal{V}_s} z_j, \end{aligned}$$

if s is a critical structure, and

$$\begin{aligned} &\gamma_s \text{gEUD}_s(z; 1) + (1 - \gamma_s)\text{gEUD}_s(z; -\infty) \\ &\quad = \gamma_s \frac{1}{|\mathcal{V}_s|} \sum_{j\in\mathcal{V}_s} z_j + (1 - \gamma_s) \min_{j\in\mathcal{V}_s} z_j, \end{aligned}$$

if s is a target, where in both cases $\gamma_s \in [0, 1]$. (Note that, when s is a target, the criterion described above is a function for which larger values are preferred to smaller values, which means that in the optimization problems as described earlier in this paper, we actually use the negative of this criterion.) The advantage of using these approximations is that our optimization problems can be formulated and solved as linear programs.

As our final criterion we have chosen to minimize the sum of all voxel doses

$$G_{L+1}(z) = \sum_{j\in\mathcal{V}} z_j.$$

Moreover, we assume that no tradeoff takes place between gEUD-criterion G_L and this final criterion G_{L+1}, so that, in problem $(\text{P}^{(L)})$, we limit ourselves to $\alpha = 1$ in the corresponding objective function (Sec. II.A).

II.D. Data and computations

We illustrate our SALO procedure on two clinical cases from different sites: brain and prostate. The brain cancer case has 8 beams and 575 beamlets while the prostate cancer case had 7 beams and 796 beamlets. In both cases, the beamlets were of dimension 5×5 mm^2, and only beamlets whose primary trajectory intersected with the target(s) were included in the model. We chose the set Z to be of the form

$$Z = \left\{ z \in \mathbb{R}^{|\mathcal{V}|} : \underline{z}_s \leq z_j \leq \bar{z}_s, j \in \mathcal{V}_s, s \in \mathcal{S} \right\}.$$

The weighting parameters γ_s in the gEUD-approximation described in Sec. II.C were found by evaluating both the gEUD and its approximation on a clinically acceptable dose distribution, where suitable values of the gEUD parameters a_s were taken from the works of Burman *et al.*,[22] Lawrence *et al.*,[23] Mayo *et al.*,[24] Michalski *et al.*,[25] Roach *et al.*,[26] Viswanathan *et al.*,[27] and clinical practice at the University of Michigan, Department of Radiation Oncology. Table I

TABLE I. Structures with corresponding number of voxels and gEUD-parameters for the two clinical test cases.

Site	Priority A	Priority B	Structure (s)	$\lvert\mathcal{V}_s\rvert$	a_s	γ_s	$\bar{z}_s$	$\underline{z}_s$
Brain	1	2	PTV	6318	−15	0.9	63	56
	2	1	Chiasm	216	10	0.38	57	—
	3	3	Brainstem	1836	10	0.5	60	—
	4	4	Optic nerve (contralateral)	218	10	0.54	63	—
	5	5	Optic nerve (ipsilateral)	247	10	0.33	63	—
			Left eye	363	—	—	63	—
			Right eye	345	—	—	63	—
			Left lens	167	—	—	63	—
			Right lens	136	—	—	63	—
			Normal tissue ring 1	6723	—	—	62	—
			(0–1.5 cm from PTV)		—	—	—	—
			Normal tissue ring 2	4652	—	—	57	—
			(1.5–3 cm from PTV)		—	—	—	
			Normal tissue ring 3	13 037	—	—	45	
			(>3 cm from PTV)		—	—	—	—
			Total	34 258	—	—	—	—
Prostate	1	2	PTV	3586	−5	0.3	85.5	73
	2	1	Rectum	8766	8	0.4	78	—
	3	3	Bladder	5373	2	0.85	78	—
	4	4	Penile bulb	294	1	1	85.5	—
	5	5	Femora	7049	4	0.8	85.5	—
			Normal tissue ring 1	2700	—	—	83	—
			(0–1.5 cm from PTV)	—	—	—		—
			Normal tissue ring 2	7203	—	—	77	—
			(1.5–3 cm from PTV)		—	—	—	—
			Normal tissue ring 3	9419	—	—	65	—
			(>3 cm from PTV)	—	—	—	—	—
			Total	44 390	—	—	—	—

provides, for each of the two cases, the structures and two prioritization scenarios, as well as the number of voxels $\lvert\mathcal{V}_s\rvert$ in each structure $s \in \mathcal{S}$, the gEUD parameters a_s and γ_s, and the dose upper and lower bounds $\bar{z}_s$ and $\underline{z}_s$ (where the latter is 0 if omitted). The structures without prioritization values will be addressed in the final treatment plan optimization, but not in the interactive portion of the algorithm.

The optimization problems were all optimized on a Mac Pro 4.1 with a single 2.93 GHz Quad-Core Intel Xeon processor and 12 GB DDR3 memory at 1066 MHz. All model generation code was written in C++ and executed in Xcode, and the primal simplex method of CPLEX 12.2 was used as the solver. Since the number of coefficients D_{kj} are too numerous to be precomputed and stored, and since the column generation relies on an efficient representation of these coefficients, we make the common assumption that these coefficients can be expressed in terms of so-called beamlet-based dose deposition coefficients.

$$D_{kj} = \sum_{i \in \mathcal{A}_k} d_{ij},$$

where $\mathcal{A}_k \subseteq \mathcal{B}$ is the subset of beamlets that is exposed in deliverable aperture $k \in \mathcal{K}$, $\mathcal{B}$ is a set of beamlets that discretizes the beams used for treatment, and d_{ij} is the dose delivered to voxel $j \in \mathcal{V}$ from beamlet $i \in \mathcal{B}$. Storing the (nonzero) coefficients d_{ij} for all $j \in \mathcal{V}$ and $i \in \mathcal{B}$ is manageable in a sparse format. Finally, in Phase 1 of the SALO procedure we generated a pool of $\lvert\overline{\mathcal{K}}\rvert = 100$ apertures.

III. RESULTS AND DISCUSSION

In this section, we will describe in detail how the treatment planning process based on the SALO procedure would proceed in the two clinical cases. We will illustrate the SALO procedure by going through all steps that a treatment planner would take when developing a treatment plan. For both cases, we will show two examples of potential a priori clinical priority lists, along with a potential sequence of decisions made by the treatment planner, for a total of four SALO applications. Furthermore, for ease of exposition we limit ourselves in both cases to a relatively small set of criteria. The tradeoff decisions made during the SALO process were made by radiation oncologists. Note that, in addition to the tradeoff curves, in our experiments the treatment planners were also provided with summary dose distribution information for the different structures (such as minimum, maximum, and mean dose, as well as DVH endpoints) during the process.

III.A. Brain case

For the brain case, we distinguish $L = 5$ major gEUD criteria, one for each of the first five structures listed in Table I. We consider the two alternative priority scenarios A and B as indicated in that table. Figures 1 and 2 show the $L - 1$ stages of Phase 2 of the SALO procedure for these two scenarios, while Fig. 3 shows the final sets of DVH curves. These curves show how scenario A compares to (a) LO and (b) scenario B.

Consider scenario A for the brain case in Fig. 1. When the treatment planner begins the process, Fig. 1(a) is generated and presented to the planner. The treatment planner then uses this information to assess the relationship between the gEUDs delivered to the PTV and the chiasm. In this instance, the treatment planner used this information to choose a lower bound on the gEUD to the PTV of 53.96 Gy, as indicated by the dot. From the graph we can then also conclude that this means that the gEUD to the chiasm will have to be at least equal to 47.94 Gy. The optimization model then adds the lower bound on the gEUD to the PTV to the set of constraints and generates Fig. 1(b). As we can see, by slightly increasing gEUD to the chiasm it is possible to reduce the gEUD to the brainstem by a meaningful amount. However, as we allow more dose to the chiasm, the benefit to the brainstem lessens. Without this accurate information, the planner would not be able to identify the clinically most beneficial tradeoff between these two criteria. The treatment planner then follows this procedure for all other stages.

After optimizing the final criterion (i.e., minimizing the sum of all voxel doses and generating a new set of apertures given the chosen gEUD bounds) we obtain a treatment plan whose DVHs are shown in Fig. 3(a). In addition, the DVHs obtained by using pure LO are shown as well. We conclude that, by accepting a minor reduction in PTV dose, all other

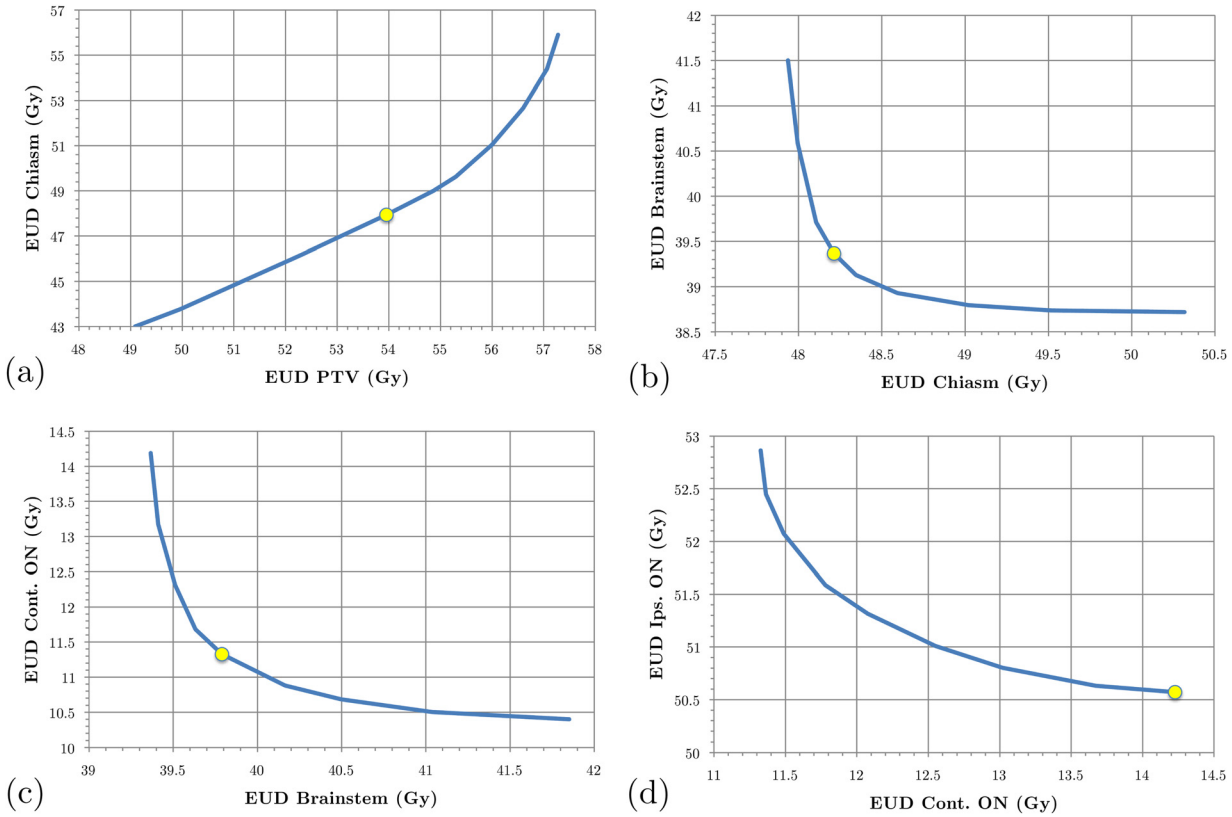

FIG. 1. SALO progression for the brain case, scenario A: (a) stage 1, (b) stage 2, (c) stage 3, and (d) stage 4.

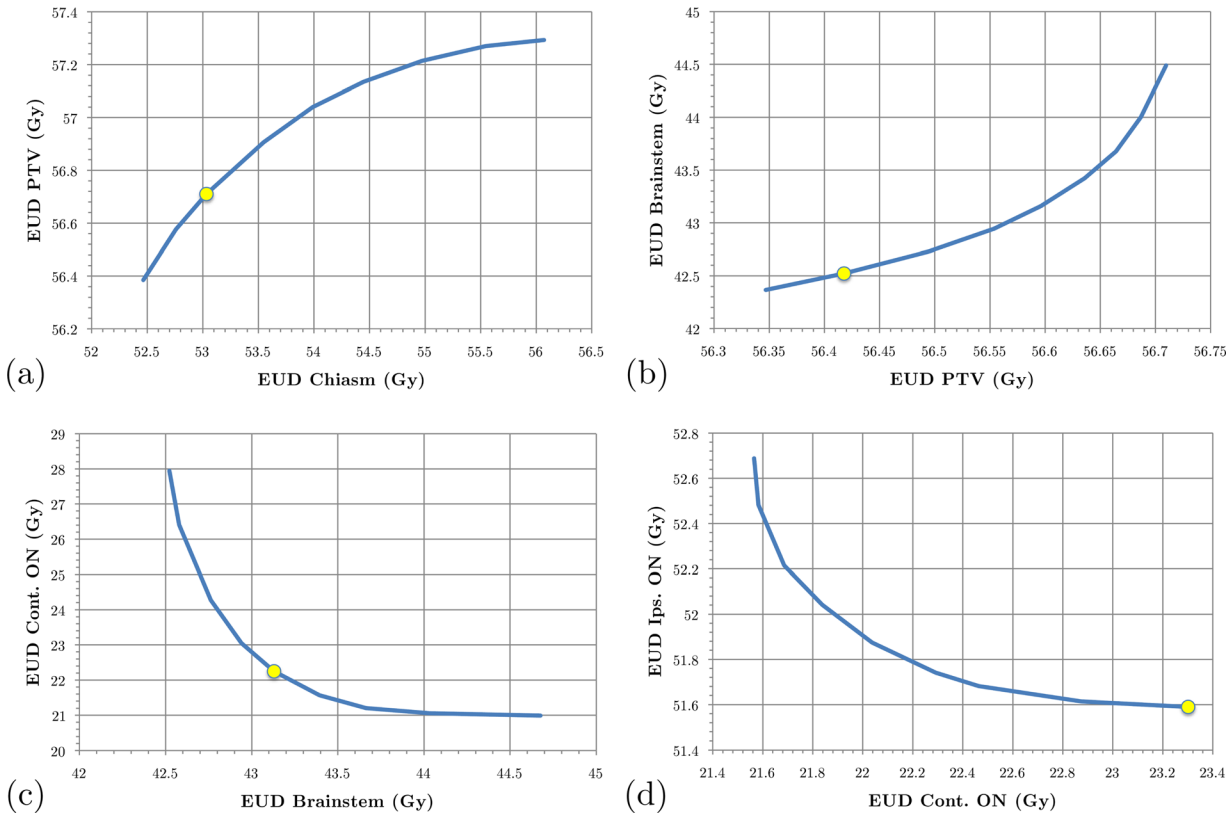

FIG. 2. SALO progression for the brain case, scenario B: (a) stage 1, (b) stage 2, (c) stage 3, and (d) stage 4.

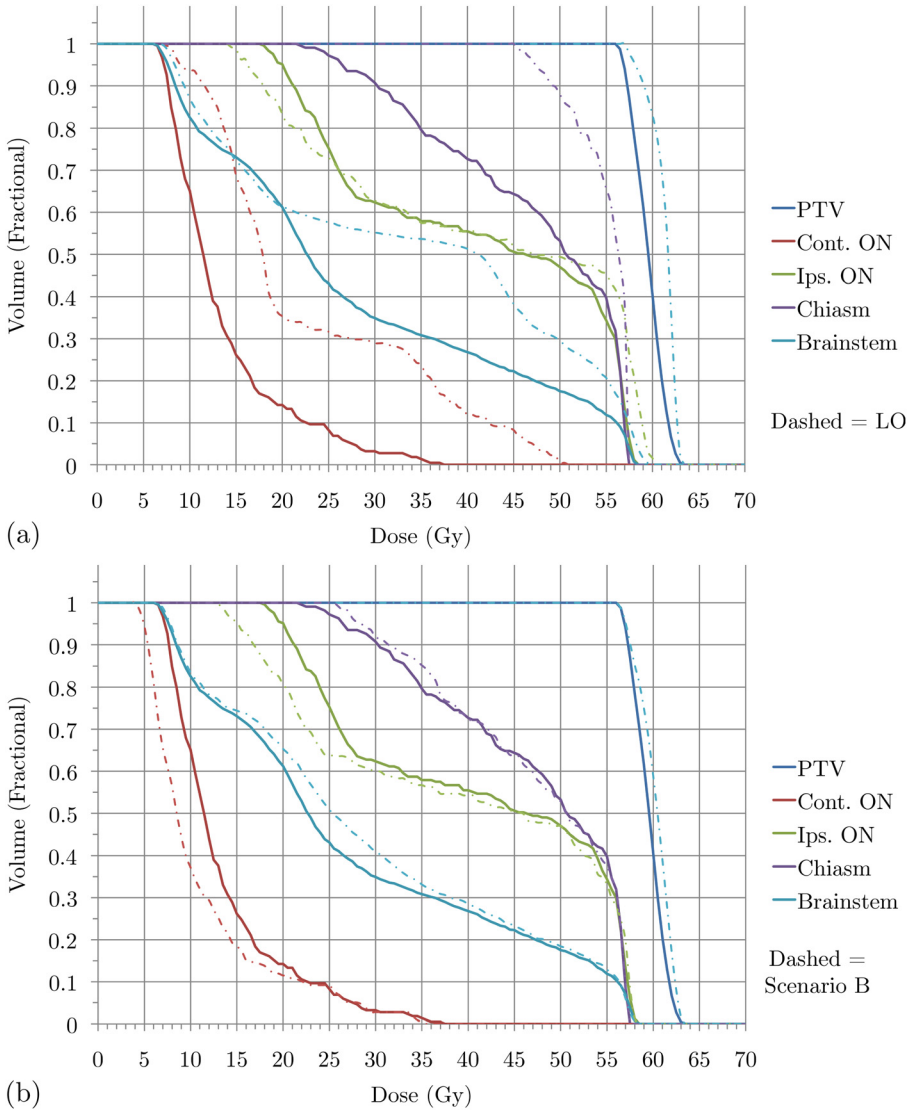

FIG. 3. Brain case: DVHs for treatment plan for (a) scenario A versus LO and (b) scenario A versus scenario B.

high priority structures receive improved dose distributions, particularly the chiasm. Our treatment planners consider the plan generated by the SALO procedure to be superior to the one created using pure LO. This is consistent with expectations, for if the LO plan were more desirable, then the treatment planner would have selected the extreme points on the tradeoff curves (representing strict prioritization). Scenario B for the brain case provides an alternate prioritization for the criteria, and the choices made by the treatment planner are shown in Fig. 2. In this scenario, the chiasm is of higher importance than the PTV and is constrained before the gEUD to the PTV. Finally, Fig. 3(b) shows that the difference between the two scenarios is relatively small, indicating a level of robustness of the procedure with respect to interchanging the priorities of PTV and chiasm.

III.B. Prostate case

For the prostate case we also distinguished $L = 5$ major gEUD criteria, again one for each of the first five structures listed in Table I. We consider the two alternative priority scenarios A and B as indicated in that table. Figures 4 and 5 show the $L - 1$ stages of Phase 2 of the SALO procedure for these two scenarios, while Fig. 6 shows the final set of DVH curves. These curves show how scenario A compares to (a) LO and (b) scenario B.

For the prostate case, the procedure progresses in a similar fashion as for the brain case. However, in this case most of the clinically desired treatment planning goals was more easily satisfied. Therefore, instead searching for a clinically feasible treatment plan, the SALO process as applied here primarily focused on finding the most desirable treatment plan. For scenario A the first tradeoff is presented in Fig. 4(a). Instead of just focusing on meeting treatment goals, the treatment planner can decide how aggressively they wish to treat the PTV. For the next sets of tradeoffs, a similar line of reasoning is used, and Fig. 6(a) allows a comparison of the DVHs obtained by the SALO procedure and pure LO. As in the brain case, a minor reduction in PTV dose allowed for significant reductions in dose to critical structures, especially the rectum.

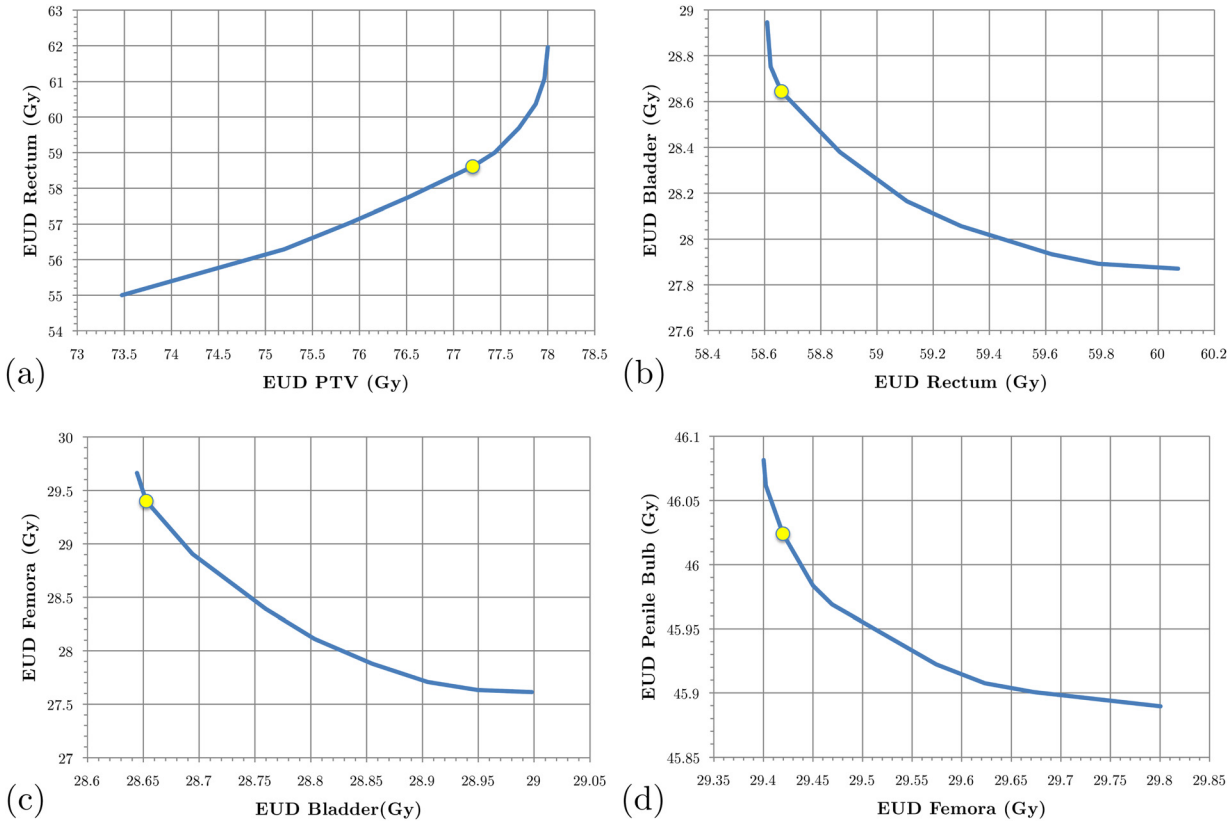

FIG. 4. SALO progression for the prostate case, scenario A: (a) stage 1, (b) stage 2, (c) stage 3, and (d) stage 4.

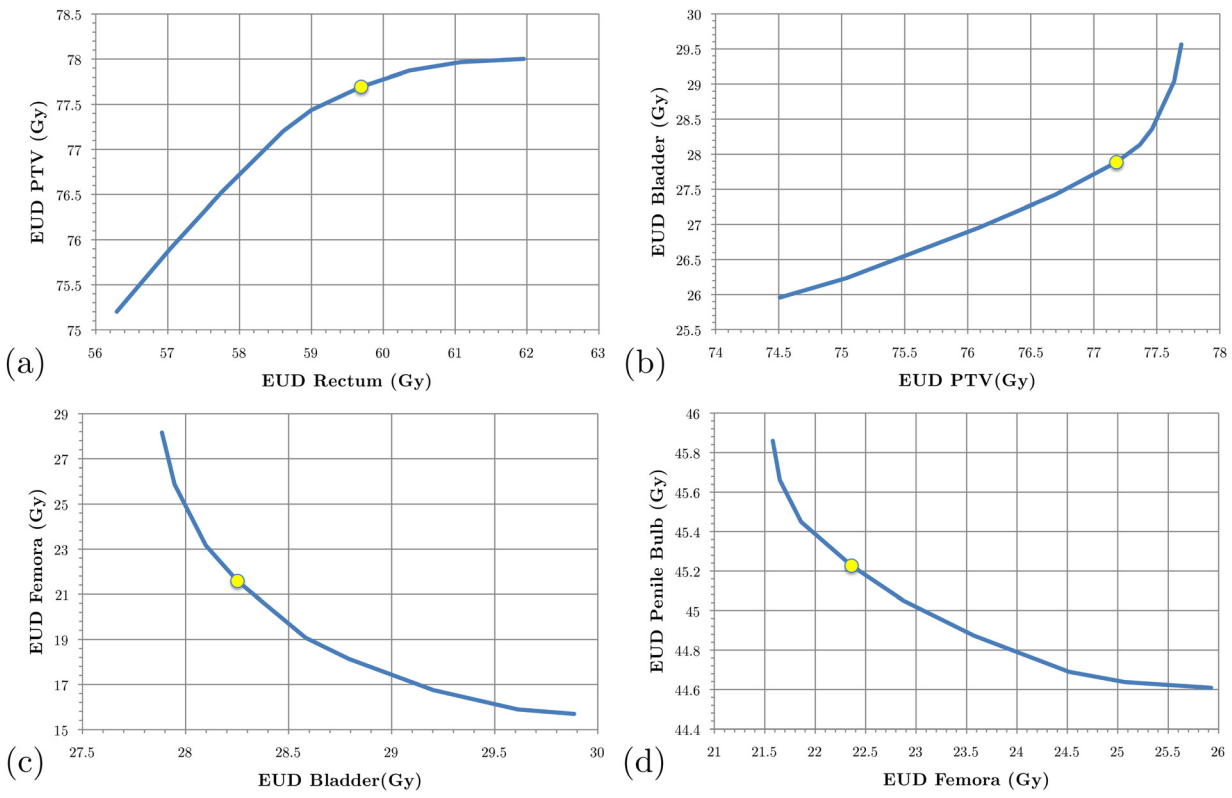

FIG. 5. SALO progression for the prostate case, scenario B: (a) stage 1, (b) stage 2, (c) stage 3, and (d) stage 4.

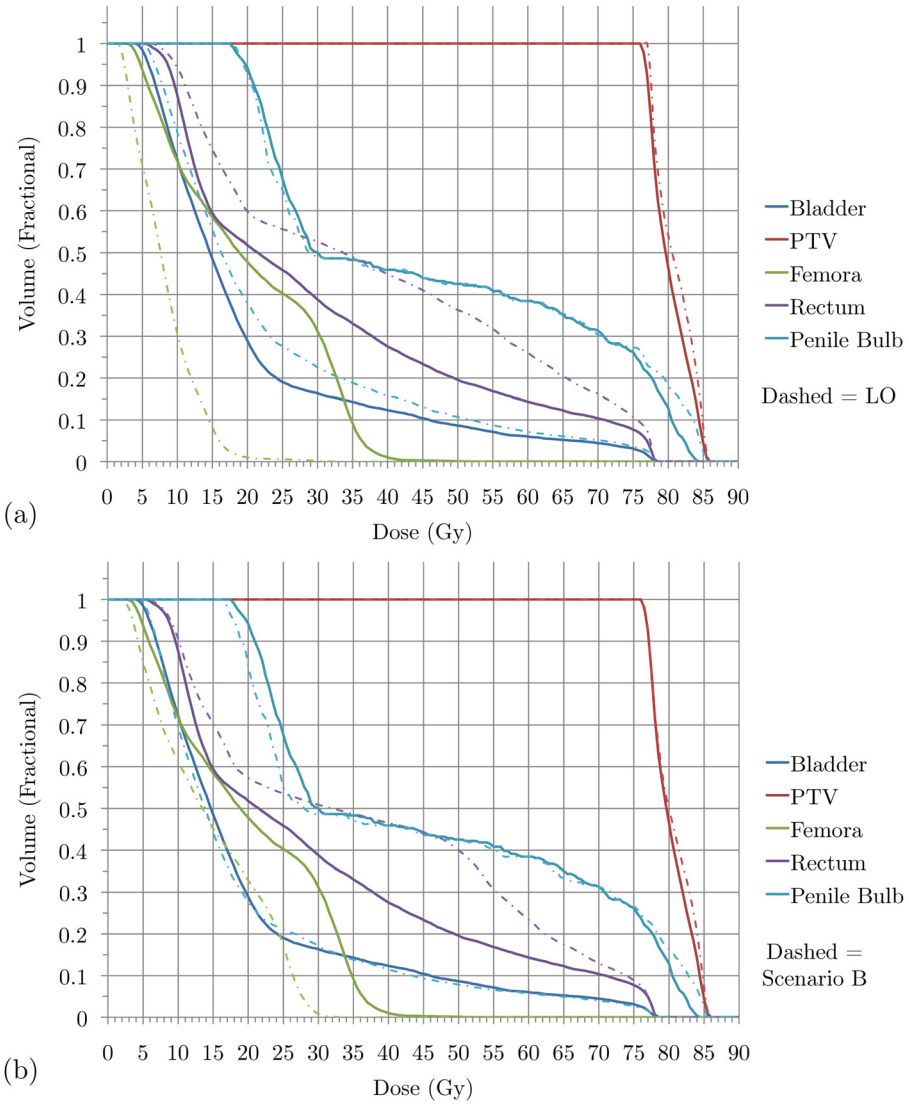

FIG. 6. Prostate case: DVHs for treatment plan for (a) scenario A versus LO and (b) scenario A versus scenario B.

When selecting desired tradeoffs on the plots, it is critical to pay close attention to the scales on the axes. For example, in Fig. 4(c) it can be seen that small changes in the bladder yields relatively large improvements in the femora. In contrast, in Fig. 4(d) the absolute differences are very small in magnitude and clinically insignificant due to the model being very tightly constrained at that stage in the optimization.

In scenario B for the prostate case the rectum received the highest priority. In this case, the treatment planner decided to be slightly more aggressive with respect to the PTV. As seen in Fig. 6(b), the rectum receives more dose using the priorities in scenario B, while the femora receives less. However, these changes are less dramatic than those between scenario A and pure LO.

III.C. General discussion

Through these processes, we can see how the prioritization aspect of LO has been integrated with the interactive nature of MCO. By combining these two characteristics, clinically desirable treatment plans were generated systematically and efficiently.

Without a treatment planning system with the flexibility to automate the SALO procedure, this analysis would not be clinically feasible. For treatment planning systems that allow plug-ins, implementing the SALO procedure is a straightforward process, and the clinical benefits could be realized quite easily. That is, the procedure can be implemented without changing a clinic's current treatment plan solver. Programmers need only to set up some background data structures and a coherent user interface.

The main downside to this type of implementation is that the usability heavily depends on efficiently approximating the tradeoff curve between criteria. If the solver is too slow, the treatment planner could be wasting time waiting for the tradeoff generation. One way to quicken the solving process is to solve a model that benefits from previous solution information. For the linear program applied to the brain and prostate cases in Sec. III, the model used the most recent solving iteration's solution to initialize the solver for the next point on the tradeoff curve. Another way to speed up the process

is to design a solver to run on a graphics processing unit (GPU). When properly designed and coded, models solved using GPUs allow for significant increases in speed (see, e.g., Ref. 28).

In practice, it might be beneficial to supplement the SALO procedure with other dose distribution information. Dose distribution statistics and DVHs for points along the tradeoff curves can be generated will little extra computational effort and would bolster the information presented treatment planners. Because all calculations up to the first stage tradeoff assessment can be computed without treatment planner interaction, different first stage scenarios can be generated to influence decisions on the full prioritization. That is, a treatment planner can assess multiple initial tradeoffs before deciding the relative importances of criteria. Finally, for the treatment planners interested in the final relative weights between the different criteria, these values can be recreated after the SALO approach (see Ref. 1 for this method).

IV. CONCLUSIONS

The SALO procedure provides treatment planners with a directed, systematic process to treatment plan selection. By following a physician's prioritization, the treatment planner can avoid wasting effort considering clinically inferior treatment plans. The planner is guided by criteria importance, but given the information necessary to accurately assess the tradeoff between criteria each stage. When applied to clinical cases, the SALO procedure efficiently generated desirable treatment plans. As treatment planning becomes more individualized and complex with new techniques and models, methods that efficiently guide the treatment planner towards desirable plans will be necessary to implement these advances at the clinical level.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Christina Tsien, MD, for her involvement in selecting tradeoffs in the SALO implementation. They also would like to thank Dr. James Balter, Ph.D., Dr. Marc Kessler, Ph.D., Dr. Dan McShan, Ph.D., and Dr. Jean Moran, Ph.D. of the Department of Radiation Oncology at the University of Michigan for their insights and contributions. This work was supported in part by Grant No. NIH-P01-CA59827.

[a]Electronic mail: troylong@umich.edu
[b]Electronic mail: marthamm@med.umich.edu
[c]Electronic mail: maryfeng@med.umich.edu
[d]Electronic mail: benedick.fraass@cshs.org
[e]Electronic mail: rth@med.umich.edu
[f]Electronic mail: romeijn@umich.edu

[1]S. Breedveld, P. R. M. Storchi, and B. J. M. Heijmen, "The equivalence of multi-criteria methods for radiotherapy plan optimization," Phys. Med. Biol. **54**, 123–135 (2009).
[2]K.-H. Küfer, M. Monz, A. Scherrer, P. Süss, F. Alonso, A. S. A. Sultan, T. Bortfeld, and C. Thieke, "Multicriteria optimization in intensity modulated radiotherapy planning," in *Handbook of Optimization in Medicine*, edited by P. M. Pardalos and H. E. Romeijn (Springer, New York, 2009), Chap. V, pp. 123–167.
[3]S. Ruzika and M. M. Wiecek, "Approximation methods in multiobjective programming," J. Optim. Theory Appl. **126**(3), 473–501 (2005).
[4]T. S. Hong, D. L. Craft, F. Carlsson, and T. R. Bortfeld, "Multicriteria optimization in intensity-modulated radiation therapy treatment planning for locally advanced cancer of the pancreatic head," Int. J. Radiat. Oncol., Biol., Phys. **72**(5), 1208–1214 (2008).
[5]D. Craft and M. Monz, "Simultaneous navigation of multiple Pareto surfaces, with an application to multicriteria IMRT planning with multiple beam angle configurations," Med. Phys. **37**(2), 736–741 (2010).
[6]D. Craft and T. Bortfeld, "How many plans are needed in an IMRT multi-objective plan database?," Phys. Med. Biol. **53**, 2785–2796 (2008).
[7]J. J. Wilkens, J. R. Alaly, K. Zakarian, W. L. Thorstad, and J. O. Deasy, "IMRT treatment planning based on prioritizing prescription goals," Phys. Med. Biol. **52**, 1675–1692 (2007).
[8]K. Jee, D. McShan, and B. A. Fraass, "Lexicographic ordering: Intuitive multicriteria optimization for IMRT," Phys. Med. Biol. **52**, 1845–1861 (2007).
[9]V. H. Clark, Y. Chen, J. Wilkens, J. R. Alaly, K. Zakaryan, and J. O. Deasy, "IMRT treatment planning for prostate cancer using prioritized prescription optimization and mean-tail-dose functions," Numer. Linear Algebra Appl. **428**(5–6), 1345–1364 (2008).
[10]D. 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**(6), 1007–1018 (2002).
[11]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," Math. Program. **101**(2), 319–338 (2004).
[12]H. 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. Control Optim. **15**(3), 838–862 (2005).
[13]C. Men, H. E. Romeijn, Z. C. Taşkin, and J. F. Dempsey, "An exact approach to direct aperture optimization in IMRT treatment planning," Phys. Med. Biol. **52**(24), 7333–7352 (2007).
[14]H. 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).
[15]A. L. Hoffmann, A. Y. D. Siem, D. den Hertog, J. H. A. M. Kaanders, and H. Huizenga, "Derivative-free generation and interpolation of convex Pareto optimal IMRT plans," Phys. Med. Biol. **51**, 6349–6369 (2006).
[16]M. Ehrgott, Ç. Güler, H. W. Hamacher, and L. Shao, "Mathematical optimization in intensity modulated radiation therapy," Ann. Oper. Res. **175**, 309–365 (2010).
[17]D. Craft, T. F. Halabi, H. A. Shih, and T. Bortfeld, "Approximating convex pareto surfaces in multiobjective radiotherapy planning," Med. Phys. **33**(9), 3399–3407 (2006).
[18]A. Niemierko, "A generalized concept of equivalent uniform dose," Med. Phys. **26**(6), 1100 (1999).
[19]B. Choi and J. O. Deasy, "The generalized equivalent uniform dose function as a basis for intensity-modulated treatment planning," Phys. Med. Biol. **47**, 3579–3589 (2002).
[20]D. Craft, T. Halabi, and T. Bortfeld, "Exploration of tradeoffs in intensity-modulated radiotherapy," Phys. Med. Biol. **50**, 5857–5868 (2005).
[21]C. Thieke, T. Bortfeld, and K.-H. Küfer, "Characterization of dose distributions through the max and mean dose concept," Acta Oncol. **41**(2), 158–161 (2002).
[22]C. Burman, G. J. Kutcher, B. Emami, and M. Goitein, "Fitting of normal tissue tolerance data to analytical function," Int. J. Radiat. Oncol., Biol., Phys. **21**(1), 7199–7209 (1991).
[23]Y. R. Lawrence, X. A. Li, I. El Naqa, C. A. Hahn, L. B. Marks, T. E. Merchant, and A. P. Dicker, "Radiation dose-volume effects in the brain," Int. J. Radiat. Oncol., Biol., Phys. (Suppl. 3), S20–S27 (2010).
[24]C. Mayo, E. Yorke, and T. E. Merchant, "Radiation associated brainstem injury," Int. J. Radiat. Oncol., Biol., Phys. **76**(Suppl. 3), S36–S41 (2010).
[25]J. M. Michalski, H. Gay, A. Jackson, S. L. Tucker, and J. O. Deasy, "Radiation dose-volume effects in radiation-induced rectal injury," Int. J. Radiat. Oncol., Biol., Phys. **76**(Suppl. 3), S123–S129 (2010).
[26]M. Roach III, J. Nam, G. Gagliardi, I. El Naqa, J. O. Deasy, and L. B. Marks, "Radiation dose-volume effects and the penile bulb," Int. J. Radiat. Oncol., Biol., Phys. **76**(Suppl. 3), S130–S134 (2010).
[27]A. N. Viswanathan, E. D. Yorke, L. B. Marks, P. J. Eifel, and W. U. Shipley, "Radiation dose-volume effects of the urinary bladder," Int. J. Radiat. Oncol., Biol., Phys. **76**(Suppl. 3), S116–S120 (2010).
[28]C. Men, X. Jia, and S. B. Jiang, "GPU-based ultra-fast direct aperture optimization for online adaptive radiation therapy," Phys. Med. Biol. **55**, 4309–4319 (2010).