A DVH-guided IMRT optimization algorithm for automatic treatment planning and adaptive radiotherapy replanning

Masoud Zarepisheh
Department of Radiation Medicine and Applied Sciences and Center for Advanced Radiotherapy Technologies, University of California San Diego, La Jolla, California 92037-0843

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

Nan Li
Department of Radiation Medicine and Applied Sciences and Center for Advanced Radiotherapy Technologies, University of California San Diego, La Jolla, California 92037-0843

Zhen Tian
Department of Radiation Medicine and Applied Sciences and Center for Advanced Radiotherapy Technologies, University of California San Diego, La Jolla, California 92037-0843 and Department of Radiation Oncology, University of Texas Southwestern Medical Center, Dallas, Texas 75390-8542

H. Edwin Romeijn
Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan 48109-2117

Xun Jia and Steve B. Jiang[a)]
Department of Radiation Medicine and Applied Sciences and Center for Advanced Radiotherapy Technologies, University of California San Diego, La Jolla, California 92037-0843 and Department of Radiation Oncology, University of Texas Southwestern Medical Center, Dallas, Texas 75390-8542

Purpose: To develop a novel algorithm that incorporates prior treatment knowledge into intensity modulated radiation therapy optimization to facilitate automatic treatment planning and adaptive radiotherapy (ART) replanning.

Methods: The algorithm automatically creates a treatment plan guided by the DVH curves of a reference plan that contains information on the clinician-approved dose-volume trade-offs among different targets/organs and among different portions of a DVH curve for an organ. In ART, the reference plan is the initial plan for the same patient, while for automatic treatment planning the reference plan is selected from a library of clinically approved and delivered plans of previously treated patients with similar medical conditions and geometry. The proposed algorithm employs a voxel-based optimization model and navigates the large voxel-based Pareto surface. The voxel weights are iteratively adjusted to approach a plan that is similar to the reference plan in terms of the DVHs. If the reference plan is feasible but not Pareto optimal, the algorithm generates a Pareto optimal plan with the DVHs better than the reference ones. If the reference plan is too restricting for the new geometry, the algorithm generates a Pareto plan with DVHs close to the reference ones. In both cases, the new plans have similar DVH trade-offs as the reference plans.

Results: The algorithm was tested using three patient cases and found to be able to automatically adjust the voxel-weighting factors in order to generate a Pareto plan with similar DVH trade-offs as the reference plan. The algorithm has also been implemented on a GPU for high efficiency.

Conclusions: A novel prior-knowledge-based optimization algorithm has been developed that automatically adjust the voxel weights and generate a clinical optimal plan at high efficiency. It is found that the new algorithm can significantly improve the plan quality and planning efficiency in ART replanning and automatic treatment planning. © *2014 American Association of Physicists in Medicine*. [http://dx.doi.org/10.1118/1.4875700]

Key words: automatic treatment planning, adaptive radiotherapy re-planning, intensity modulation, optimization, Pareto surface

1. INTRODUCTION

Treatment planning is currently a patient specific, time-consuming, and resource-demanding task in intensity modulated radiation therapy (IMRT). Plan quality is subjective and highly dependent on the institution and the planners' skills and experiences.[1–3] Automatic treatment planning has been introduced to facilitate this procedure. Two general techniques have been followed in automatic treatment planning. The first technique is to develop a systematic algorithm to

automatically adjust the optimization model parameters for the new patient to generate a satisfactory plan that meets some predetermined clinical criteria.[4–7] The second technique is to employ a library of clinically approved and delivered plans of previously treated patients with similar medical characteristics in order to find a set of parameters for a new patient that produce a clinically desirable plan.[8–10] In this work, we propose an algorithm to automatically adjust the optimization model parameters to replicate a reference plan for a new patient's geometry with the reference plan being selected on patient similarity from a library of preciously delivered plans.

In conventional radiation therapy, a clinically approved plan is constructed based on the patient image data (planning CT) acquired at the beginning of the treatment process. Then, this plan will be delivered to the patient throughout the whole treatment course. The patient's geometry changes during the treatment course. This implies that the original plan built based on the patient's initial geometry is no longer an optimal plan for the patient's new geometry. Adaptive radiotherapy (ART) has been proposed to ease this issue by imaging the patient's changed geometry during the treatment course and modifying the original, or reference, plan accordingly.[11–20]

The new plan in either automatic treatment planning or ART replanning is supposed to inherit some clinically meaningful characteristics from the reference one. In this paper, we would like to consider dose-volume information in the form of the DVH as the common characteristic between the new and the reference plans. Therefore, we are trying to generate a new plan with the DVH similar to the reference ones. The flexibility of the DVH to be used on different geometries as well as its clinical relevance and importance inspired us to employ that as the shared characteristic between the reference and the new plan. In fact, as opposed to some evaluation criteria such as dose distribution, a DVH is quite flexible and can be simply applied to different geometries since it provides information about the amount of doses in different volumes of each organ which is irrespective to the geometry and the shape of the structures.

There are numerous research studies in the literature about developing an optimization model capable of generating a plan with the DVHs satisfying a set of partial DVH constraints or with the DVHs close to the given ones.[21–30] In this paper, we aim to develop an optimization model to handle the entire DVH curves rather than a few points on the curves. The goal is to minimize an appropriately defined metric designed to represent the distance between a plan's DVHs and the reference DVHs. Since the proposed metric is very complicated, nonconvex, and difficult to optimize, we employ a surrogate convex function along with some parameters in order to improve that metric by adjusting the parameters. It is a common optimization approach when our clinically relevant metric is very complicated.[4,31–37] In fact, by varying the parameters we are navigating among the set of plans known as Pareto-optimal plans (i.e., plans on the Pareto surface) to optimize the given metric. The Pareto surface includes a set of the plans which are Pareto optimal with respect to the surrogate function. We define our surrogate function based on the voxel-based optimization model that

has recently gotten a lot of attention in radiotherapy optimization due to its promising results in clinical studies.[4,31–37] In this model, there is a weight associated to each voxel in the objective function of the optimization model. We recently did a mathematical study on the voxel-based model that revealed the main advantages of this model.[38] However, the adjustment of a dramatically increased number of parameters (voxel weights) in the voxel-based model remained to be a big problem, which makes the trial-and-error approach of parameter selection impossible to apply clinically. We propose a systematic scheme to adjust the voxel weights automatically in order to approach the reference DVHs.

2. PARETO SURFACE NAVIGATION THROUGH THE VOXEL-WEIGHT ADJUSTMENT

2.A. Voxel-based optimization model

In this section, we briefly introduce the voxel-based model and compare it with the commonly used organ-based model. For details readers are referred to Ref. 38. Problems (1) and (2) illustrate the general form of the voxel-based and the organ-based optimization models, respectively,

$$\min_{x \geq 0} \sum_{\sigma \in S} \sum_{j \in v_\sigma} w_j F_j(D_j x), \tag{1}$$

$$\min_{x \geq 0} \sum_{\sigma \in S} w^\sigma G^\sigma(D^\sigma x), \tag{2}$$

where $S = T \cup C$ is the set of structures with T being the tumors and C critical organs and v_σ denotes the set of voxels that belong to the structure σ. w_j and w^σ are the weights corresponding to the voxel j and the organ σ. D denotes dose deposition matrix and its entry D_{jk} specifies the dose received by the voxel j from a beamlet k at its unit intensity. D_j is the jth row of the matrix D that corresponds to voxel j and D^σ is the submatrix of D corresponding to organ σ. F_j is a voxel penalty function and G^σ is an organ penalty function, and x is the fluence map intensity.

From the mathematical modeling point of view, the main difference between Problems (1) and (2) is that there is a weight associated with each voxel in Eq. (1), while all the voxels within each organ are weighted equally in Eq. (2). Both problems include a set of functions of the dose distribution along with their weighting factors. These functions are our clinically relevant criteria if those criteria can be represented as the tractable functions. Otherwise, a set of surrogate functions is employed and then the weights are adjusted to improve the clinical criteria.

Any optimal solution of Problem (1) is a Pareto optimal solution with respect to the given penalty function. It means, different penalty function G *usually* results in different Pareto surfaces in the organ-based model (some of them correspond to the same Pareto surface[39,40]). Clinical experiments have also verified a great impact of using different penalty functions in the organ-based model on the optimized plan quality,[41–43] and so far there is no general agreement on the choice of the objective function. In contrast, an objective function selection is no longer an issue in the voxel-based
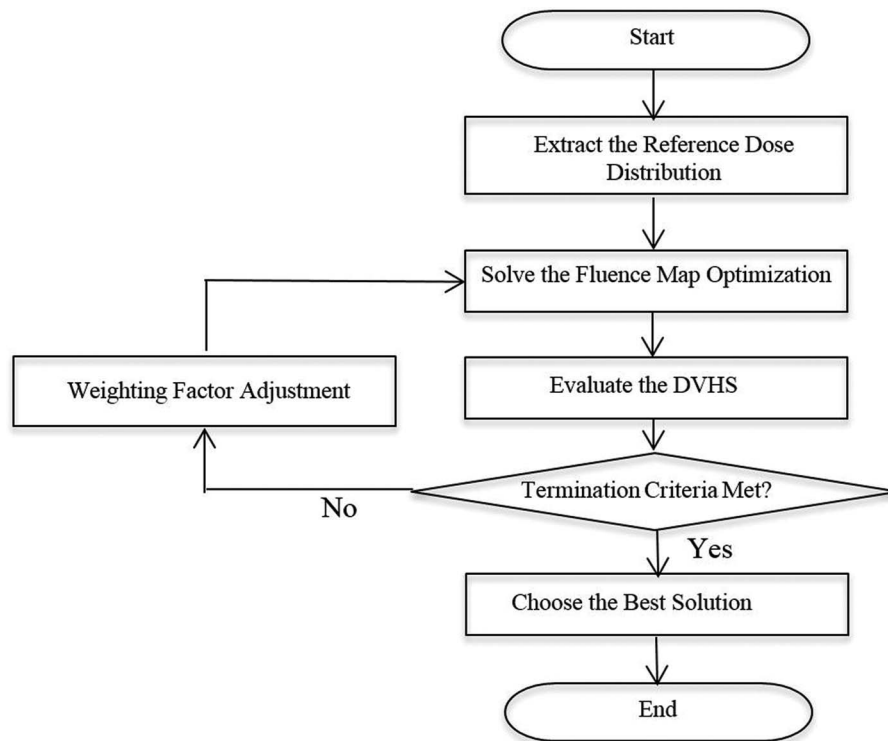
FIG. 1. Flowchart of the proposed algorithm.

model and a unique Pareto surface defined based on the dose distribution concept (called *dose distribution Pareto surface*) is generated by the voxel-based model as long as the penalty function F satisfies some desirable properties. The dose distribution Pareto surface is much larger than the organ-based Pareto surfaces, and in our previous work[38] we proved that it encompasses all the Pareto surfaces generated by the organ-based model using different penalty functions. We also proved that it includes all the plans which are Pareto optimal in terms of the DVH concept. Therefore, the voxel-based model can be employed as a surrogate optimization model especially when our clinically relevant metric is defined based on the DVH. In our previous work, it turns out that the following quadratic voxel-based model is capable of generating almost the entire dose distribution Pareto surface (the entire Pareto surface excluding the so-called *nonproperly* Pareto optimal points[44]). So, we employ the following model in this work due to its computational efficiency:

$$\min_{x \geq 0} \sum_{\sigma \in S} \sum_{j \in v_\sigma} w_j (D_j x - r^\sigma)^2, \tag{3}$$

where r^σ denotes the prescription dose for the structure σ. Despite the benefits of the voxel-based model, it has a big downside which is the huge number of the parameters. In this paper, we aim to tackle this problem by developing an algorithm which takes into account the prior DVH information. In fact, we would like to generate the similar DVHs to the reference ones, which already contain physician approved trade-off information by adjusting the voxel weights.

2.B. Algorithm

To facilitate reading, "the new patient" will refer to the patient for which we look for a plan to generate. That is, "the new patient" can represent the patient in automatic treatment planning or the existing patient with the new geometry in ART replanning. Also, let us refer to the dose distribution Pareto surface of the new patient simply as the "Pareto surface." We propose an algorithm which iteratively navigates the Pareto surface and projects the plan corresponding to the reference DVHs back on the Pareto surface.

Projecting a plan on the Pareto surface means finding the closest plan on the Pareto surface to the reference plan. It implies that, if the reference DVHs are feasible but not Pareto optimal for the new patient, then the projected plan on the Pareto surface would be even better than the reference one in terms of the DVHs; and if the reference DVHs are outside the feasible region of the new patient, meaning it is impossible to generate that plan, then the projected plan on the Pareto surface would be similar but not as good as the reference one. In the latter case, the closer the reference plan is to the Pareto surface, the more similar the projected plan would be to the reference one.

Figure 1 demonstrates the flow of the algorithm. At first, we find a particular dose distribution corresponding to the reference DVHs (Sec. 2.C), referred to as the reference dose distribution. Then, we solve the fluence map optimization problem (3) with some initial weights to get a Pareto optimal solution. This solution is evaluated by the metric defined to quantify the difference between its DVHs and the reference DVHs (Sec. 2.F). If the termination criteria are not met,
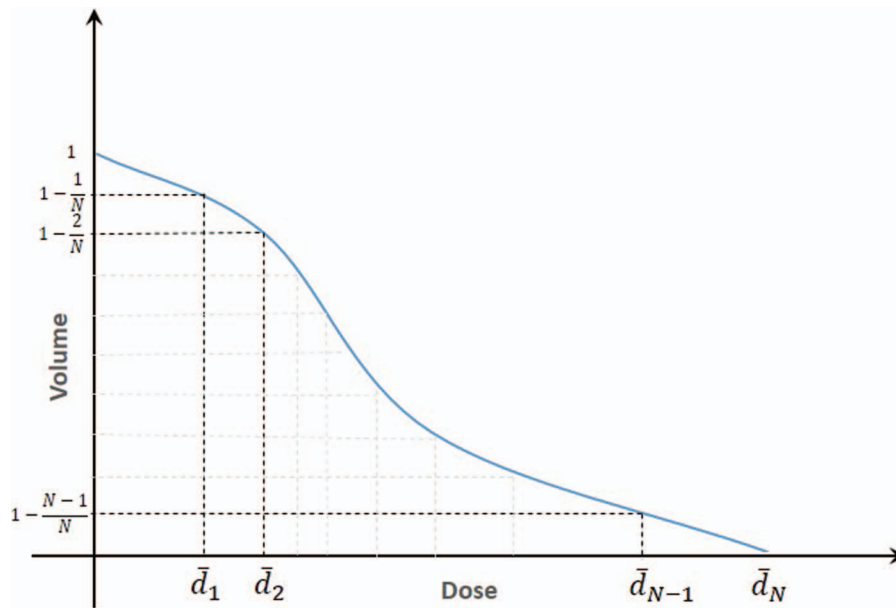
FIG. 2. Extracting the reference dose from the reference DVH for an organ with N voxels.

among all the permutations of the reference dose distribution that are corresponding to the reference DVHs, we target the closest one to the Pareto solution generated by solving Problem (3) (Sec. 2.D). We then project that reference dose distribution back onto the Pareto surface by adjusting the weighting factors (Sec. 2.E). We continue the algorithm by solving the fluence map optimization problem with the updated weighting factors. When the termination criteria are met, we pick the best solution, in terms of the metric defined to evaluate the DVHs (Sec. 2.F), among all the Pareto solutions generated by the algorithm.

2.C. Extracting a reference dose distribution from the reference DVHs

Given a set of reference DVHs, our algorithm needs to find out its corresponding dose distribution with the same number of the voxels for each structure as the new patient. A typical dose distribution can be reconstructed by discretizing the volume axis of the DVH uniformly by the number of the voxels and finding the corresponding dose values. Figure 2 illustrates the process for an organ with N voxels. In this example, we would have the similar DVH for the new patient's organ, by spreading the reference dose values $\bar{d}_1, \bar{d}_2, \ldots, \bar{d}_N$ among the voxels of the new patient's organ. It should be noted that DVH is insensitive to the voxel's spatial location meaning that the reference dose values can be spread among the voxels in any arbitrary way to generate the reference DVH. In another words, if we index the voxels of the new patient's organ, then any permutation of $\bar{d}$ can be assigned to the voxels to generate the reference DVH.

2.D. Nonconvexity issue

After extracting the reference dose distribution from the reference DVHs, we solve Problem (3) with arbitrary ini-

tial weighing factors to get the Pareto optimal solution $d_0 = Dx_0$. After obtaining a solution, the weights are updated to guide the DVHs of the current solution toward the reference DVHs. We do not have the DVH directly in our optimization model, but we have the voxel dose values in Problem (3). Therefore, we need to change the voxel dose values to make their corresponding DVHs close to the reference ones through updating the weights. The difficulty here is due to the insensitivity of the DVH to the permutation of the dose values in each organ that results in the existence of many paths through which we can change the dose values to get close to the reference DVHs. This problem is known as the nonconvexity problem in radiotherapy optimization. In fact, if $\pi(.) : \{1, \ldots, v_\sigma\} \rightarrow \{1, \ldots, v_\sigma\}$ denotes an arbitrary permutation of the dose values of organ σ, then $\pi(\bar{d}^\sigma)$ would also represent the same DVH and hence we can approach the reference DVH by driving d_0^σ toward $\pi(\bar{d}^\sigma)$. Now the question that arises here is: "which permutation of the reference dose should be targeted?" To answer this question, we need to consider three following possible scenarios regarding the position of the reference DVHs with respect to the feasible solution set of the new patient (the set of all deliverable dose distributions for the new patient).

(i)   All permutations of the reference dose are inside the feasible region.
(ii) All permutations of the reference dose are outside the feasible region.
(iii) Some of the permutations of the reference dose are inside the feasible region and the others are outside.

For (i), any permutation of the reference dose is possible to be generated since they are all inside the feasible region. They can be even improved by projecting them back onto the Pareto surface. The best plan in this case corresponds to the permutation whose projection provides the most improvement in the reference plan. Figure 3 (left) clarifies this situation for
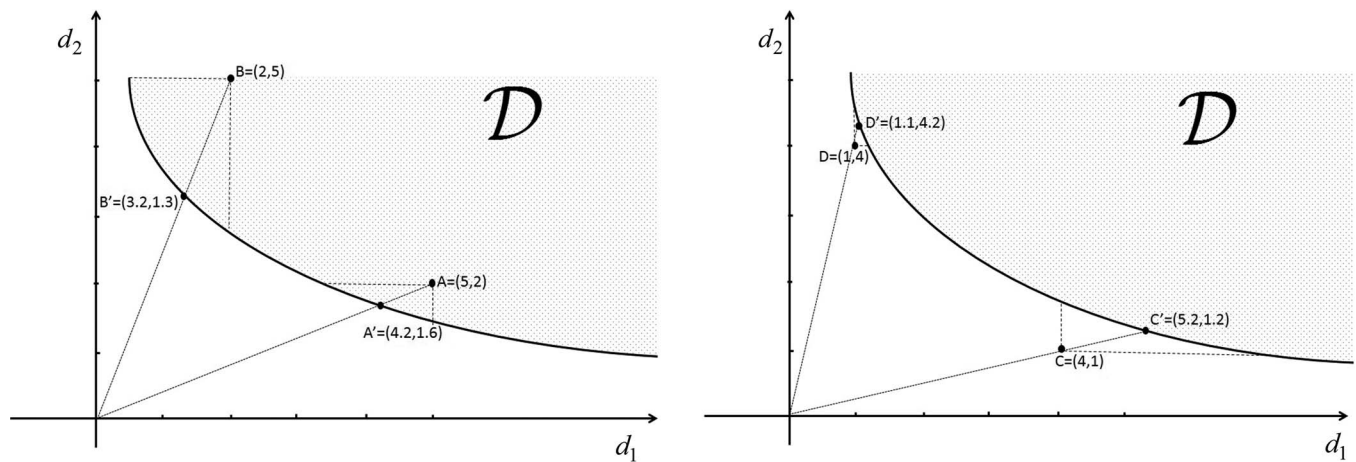
FIG. 3. Projection of different permutations of dose values onto the Pareto surface. (Left: permutations are inside the feasible region): Projecting B results in more improvement in the reference dose. (Right: permutations are outside the feasible region): Projecting D results in less deterioration in the reference dose.

an OAR with two voxels. Points A and B are two different permutations of the reference dose and considered to be the same in terms of the DVH, and A′ and B′ are their special projections. It can be seen that B′ is better than A′ in terms of the DVH which means the more improvement is achieved by projecting B onto the Pareto surface. For the second scenario, any permutation of the reference dose is impossible to be generated and they need to be projected back onto the Pareto surface in order to find the feasible and similar plans to the reference one. The projection in this case would deteriorate the reference dose to make it feasible, and the best plan corresponds to the permutation that provides the least deterioration. Figure 3 (right) explains this scenario. Points C and D correspond to different permutations of the reference dose, and C′ and D′ are their special projections on the Pareto surface. As can be seen, projecting D results in less deterioration and hence the plan with better DVH. It is obvious that in the third scenario the best plan is obtained by projecting the permutation of the reference dose that is inside the feasible region and results in the most improvement.

Unfortunately, the above procedure is very computationally expensive, if not prohibitive. We cope with this issue by iteratively choosing some particular permutation of the reference dose and projecting them back onto the Pareto surface. At each iteration of the algorithm, we target the permutation of the reference dose corresponding to the closest dose distribution to the current Pareto solution. Then, we project the targeted permutation of the reference dose on the Pareto surface to get a new Pareto optimal plan. We continue this process until either the maximum allowed number of iterations is reached or there is not enough improvement in the plan generation. At the end, we pick up the best Pareto optimal plan among all the projected plans on the Pareto surface generated throughout this iteration process, as each iteration is not guaranteed to produce a better plan than the last. In the rest of this section, we explain how we can find out the closest permutation of the reference dose to the current Pareto solution, and in Sec. 2.E we will elaborate on the projection of the targeted plan back on the Pareto surface. Section 2.F introduces a metric to pick up the best Pareto solution.

Let us consider the initial Pareto solution d_0 as a current solution. For structure $\sigma \in S$, if $\tilde{\pi}(.)$ denotes the permutation that sorts the vector d_0^σ in an ascending order [i.e., $\tilde{\pi}(d_0^\sigma) = \text{sort}(d_0^\sigma)$], then according to the following theorem, $\tilde{\pi}^{-1}(\bar{d}^\sigma)$ is the closest permutation of the reference dose to the current solution regardless which p-norm used to measure the distance.

Theorem 1: For $\sigma \in S$, let Π denote the set of all possible permutations $\pi(.): \{1, \ldots, v_\sigma\} \to \{1, \ldots, v_\sigma\}$. Then for every $p \in [1, \infty)$, we have

$$\left|\left|d_0^\sigma - \tilde{\pi}^{-1}(\bar{d}^\sigma)\right|\right|_p \le \left|\left|d_0^\sigma - \pi(\bar{d}^\sigma)\right|\right|_p \quad \forall \pi \in \Pi. \tag{4}$$

The proof of this theorem is given in Appendix A. Theorem 1 reveals that, for each structure $\sigma \in S$, among all the permutations $\pi(\bar{d}^\sigma)$ of the reference dose that corresponds to the same DVH, $\tilde{\pi}(d_0^\sigma)$ yields the closest one to the current solution. Now, the dose distribution that we would like to target at a certain iteration of the algorithm can be computed by performing the above procedure for each structure. With a slight abuse of notation, we denote the closest dose distribution for all structure by $\tilde{\pi}^{-1}(\bar{d})$. It should be noted that different permutations of the reference dose would be targeted at different iterations of the algorithm due to the dependency of the $\tilde{\pi}(.)$ on the current Pareto solution at each iteration.

2.E. Projection on the Pareto surface

The next step is to project the point $\tilde{\pi}^{-1}(\bar{d})$ back on the Pareto surface of Problem (3). The Pareto surface, here is denoted by $F(z) = 0$, is a part of the frontier of the set $Z = \{z = (Dx - p)^2 | x \ge 0\}$. The frontier of Z is piecewise quadratic since it is a quadratic transformation of the frontier of the transferred cone $(Dx - p)$ which is piecewise linear. The shape of the Pareto surface depends on the deposition matrix D and the prescription vector p. There does not exist an efficient way to obtain the Pareto surface, making the projection a challenging problem. We approach this issue by using a quadratic approximation of the Pareto surface.

The projection idea is based on the fact that each point on the Pareto surface corresponds to the set of weighting
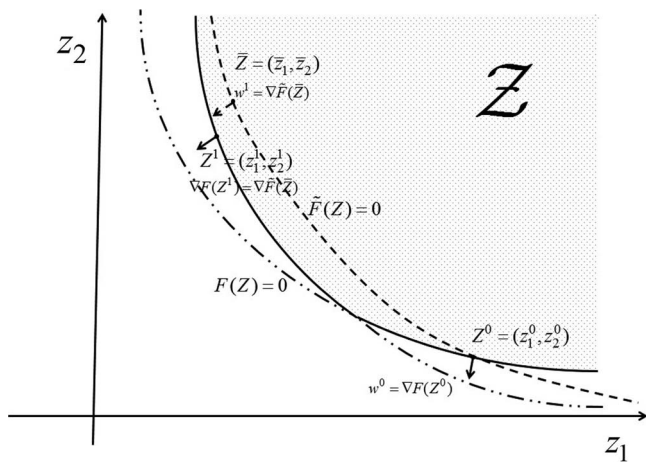
FIG. 4. Illustration of the projection method in a two-dimensional space. The solid line is the Pareto surface, the dashed line is the quadratic surface passing through the initial point and the reference point and its gradient at the reference point would be used as the new weighting factors in the optimization model.

factors that are the gradient of the Pareto surface ($\nabla F(z)$) at that point. Specifically, if we use $\nabla F(z)$ as the weighting factors in Problem (3), the optimal solution would be z. Now, we use the quadratic surface that passes through the current Pareto solution (d_0 at the first iteration) and $\tilde{\pi}^{-1}(\bar{d})$ as an approximation of the Pareto surface, and we set the weighting factors of Problem (3) equal to the gradient of this approximation of Pareto surface at $\tilde{\pi}^{-1}(\bar{d})$. Then, the optimal solution is expected to be on the projection area of the point $\tilde{\pi}^{-1}(\bar{d})$ on the Pareto surface. Figure 4 clarifies this process in a two-dimensional space where the point $\tilde{\pi}^{-1}(\bar{d})$ is inside the feasible region. The same behavior can be observed if the point that we want to project is outside the feasible region.

In Fig. 4, $\bar{Z} = (\bar{z}_1, \bar{z}_2)$ represents the point that is supposed to be projected on the Pareto surface. The Pareto surface, which is a border of the feasible region and is depicted by the solid line, is a piecewise quadratic surface and is denoted here by $F(Z) = 0$. The dashed-dotted lines are the remaining parts of the two quadratic surfaces that make the piecewise quadratic Pareto surface. The dashed line, here is denoted by $\tilde{F}(Z) = 0$, represents the quadratic surface passing through

the initial point $Z^0 = (z_1^0, z_2^0)$ and the reference point. The gradient of $\tilde{F}(Z)$ at $\bar{Z}$ is used as the weighting factors in Problem (3) to obtain the projection of $\bar{Z}$ on the Pareto surface. Since the weighting factors are corresponding to the gradient of the Pareto surface at the optimal solution, the optimal solution would be Z^1 where we have $\nabla F(Z^1) = \nabla \tilde{F}(\bar{Z})$. As can be observed in Fig. 4, Z^1 is located on the projection area of $\bar{Z}$ on the Pareto surface. Theorem 2 shows that $\nabla \tilde{F}(\bar{Z})$ can be easily approximated from the existing information.

Theorem 2: The new weighting factors utilized to project the closest permutation of the reference dose on the Pareto surface can be approximated as

$$w_j^1 = \nabla_j \tilde{F}(\bar{Z}) \cong w_j^0 \frac{|d_{0j} - p_j|}{\left| \tilde{\pi}_j^{-1}(\bar{d}) - p_j \right|} \quad \forall \sigma \in S, j \in v_\sigma.$$

The proof of this theorem is given in Appendix B. By setting the weighting factors equal to w^1 in Problem (3) and solving it, we would get the new dose distribution d^1. Then, the above procedure is carried out on d^1 as the current solution.

2.F. Evaluation of the DVHs

At each iteration of the algorithm, the weighting factors are updated to encourage the DVHs of the current solution (current DVHs) to approach the reference DVHs. However, we do not directly optimize the distance to the reference DVHs. We need to define an appropriate metric to evaluate the solution quality at each iteration and to pick the best solution among the generated solutions when the algorithm terminates.

We define the metric based on two areas between the obtained and the reference DVHs: the areas where the current DVHs are better than the reference DVHs, and the areas where the current DVHs are worse than the reference DVHs. Figure 5 illustrates these two types of areas for an organ at risk (left) and a PTV (right). The solid line represents the reference DVH and the dashed line represents the current DVH. The prescription dose for the PTV is shown by a solid vertical line in the right plot. For organ at risk, A depicts the area where the reference DVH is better than the current one and B depicts the area where the current DVH is better. Our first priority is to make A as small as possible, and our second
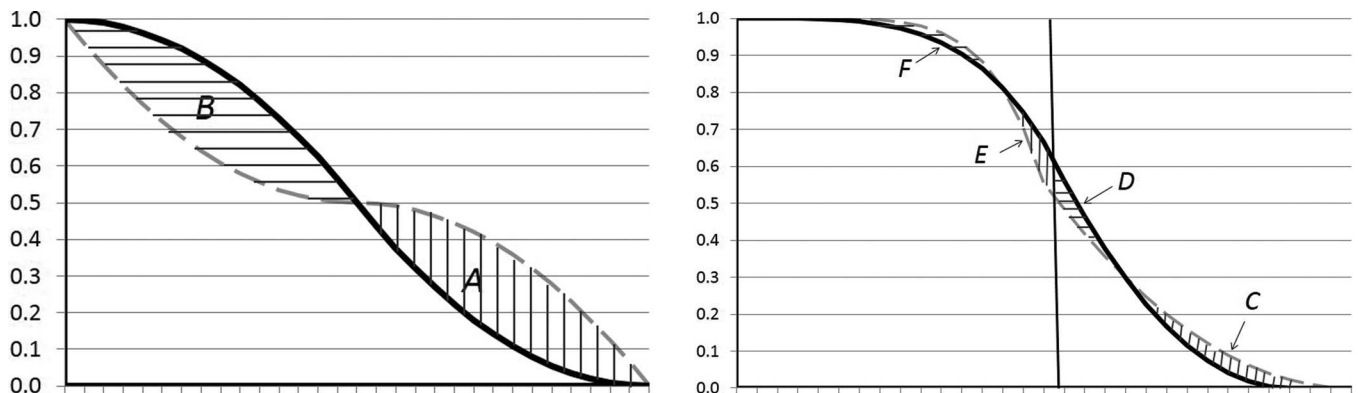

FIG. 5. Comparing the current DVH and the reference DVH for an OAR (left), and a PTV (right). The solid line represents the reference DVH and the dashed line represents the current DVH.

priority is to make B as big as possible. Therefore, the following metric is used to quantify the quality of the current DVH with respect to the reference one for this organ:

$$m = m_A - \epsilon m_B,$$

where ϵ is a small number, and m_A, m_B are the metrics corresponding to the areas A and B. Using a small number to reflect different levels of priorities among different objectives is a common approach in optimization.[45] For PTV, given that the closer dose to the prescription value is preferred, C and E represent the areas where the reference DVH is better, while D and F represent the areas where the current DVH is better. Therefore, we can define the metric for PTV as

$$m_\sigma = (m_C + m_E) - \epsilon(m_D + m_F).$$

Now, we need to combine all organs' metrics to come up with a single metric to quantify the current DVHs. While there are many ways to do that, we take the maximum value among all the organs' metrics

$$m = \max_{\sigma \in S} m_\sigma, \tag{5}$$

where m_σ is a metric corresponding to the structure σ. To some extent, this metric would choose the plan for which the deviation to the reference plan is uniformly distributed among all the organs. To have a better convergence behavior for the algorithm in terms of the above metric, it is better to reflect the desire of distributing the deviation uniformly among the organs in the algorithm. Toward this end, we penalize each organ weights proportional to its deviation m_σ to the reference DVH. It can be done through the following weight adjustment:

$$w_j^1 = w_j^0 \left(1 + \frac{m_\sigma}{\max_{\sigma \in S} |m_\sigma|}\right) \forall j \in v_\sigma, \sigma \in S.$$

3. EXPERIMENTS AND RESULTS

In this section, we provide three different case studies. All of them are IMRT cases with seven equispaced beam angles. The first case is designed to test the capability of the algorithm to project a reference plan on the Pareto surface, and we use Matlab as our computational framework for this case. The second and the third case studies represent the applications of the algorithm for adaptive radiotherapy replanning and automatic treatment planning. We use our inhouse GPU-based research planning system (SCORE) to run the algorithm and demonstrate the results for these two cases. We set beamlet size and voxel size to 0.5×0.2 and $0.2 \times 0.2 \times 0.25$ cm^3, respectively. We cannot afford solving the optimization problem including all the voxels and so we just pick out some of them by downsampling process. However, to make the second and the third case studies more realistic, the final dose calculation on all the voxels, including MLC transmission, will be performed after optimization. Moreover, since we aim to evaluate the performance of the proposed optimization algorithm, in the following we will just mention the number of the voxels used in optimization model and the time spent to solve the optimization problem.

3.A. A prostate case to test the algorithm

This is a small prostate case with 50 221 voxels and 943 beamlets. We implemented the algorithm in Matlab, and ran the code on a PC with Intel Core i7 3.40 GHz CPU and 12 GB RAM. Two distinct scenarios are investigated for this case: (1) the reference DVHs are inside the feasible region (feasible DVHs), (2) the reference DVHs are outside the feasible region (infeasible DVHs). We will show that the algorithm generates a plan with DVHs better than the reference ones for the first scenario, while it finds a plan with DVHs worse than, but similar to, the reference ones for the second scenario.

In order to get the feasible and infeasible sets of the DVHs, we first generated a plan by solving Problem (3) with a set of the weighting factors, which yields a plan that is Pareto optimal in terms of the DVHs,[38] and then we manipulated them. An infeasible plan can be obtained by improving the DVH curves of the Pareto optimal plan (shifting the OAR DVHs toward left and PTV DVH toward prescription dose), and a feasible plan can be generated by worsening the DVHs of the Pareto optimal plan.

Figures 6 and 7 are regarding the feasible reference DVHs shown by the solid lines. The dotted lines in Fig. 6 depict the result of the algorithm. As can be seen, the DVHs generated by the algorithm are better than the reference ones due to the feasibility of the reference DVHs. In fact, the algorithm projects the feasible reference plan on the Pareto surface resulting in a better plan. It took us 10.6 s to get this plan. Figure 7 investigates the nonconvexity issue that may lead to local optimality. For this purpose, we ran the algorithm 100 times with randomly generated initial weighting factors, and then we picked the best (dotted lines) and the worst (dashed lines) solutions with respect to our metric introduced in Eq. (5). It seems that the local optimality is not a big issue for this case, especially since both plans are better than the reference one. In fact, the algorithm projects the reference plan on the Pareto surface regardless the initial solution which implies that all 100 plans are Pareto optimal and on the projection area of the reference plan. However, given that the algorithm started from different points on the Pareto surface, it ends up in different parts of the projection area.

Figures 8 and 9 deal with the infeasible reference DVHs. Figure 8 shows that the algorithm is able to generate a plan that, although worse than the reference plan, is close to it. The computation time is 7.4 s in this case. Figure 9 examines the local optimality issue by running the algorithm for 100 times with randomly generated initial weights and illustrates the best and the worst plans. Again, the local optimality does not seem to be a big problem.

3.B. Adaptive radiotherapy replanning for a head and neck case

The application of the proposed algorithm in adaptive radiotherapy replanning is illustrated in this section with a head and neck case that has been treated previously at UCSD Moores Cancer Center. The number of the voxels and beamlets for this patient are 20 758 and 13 934, respectively. The
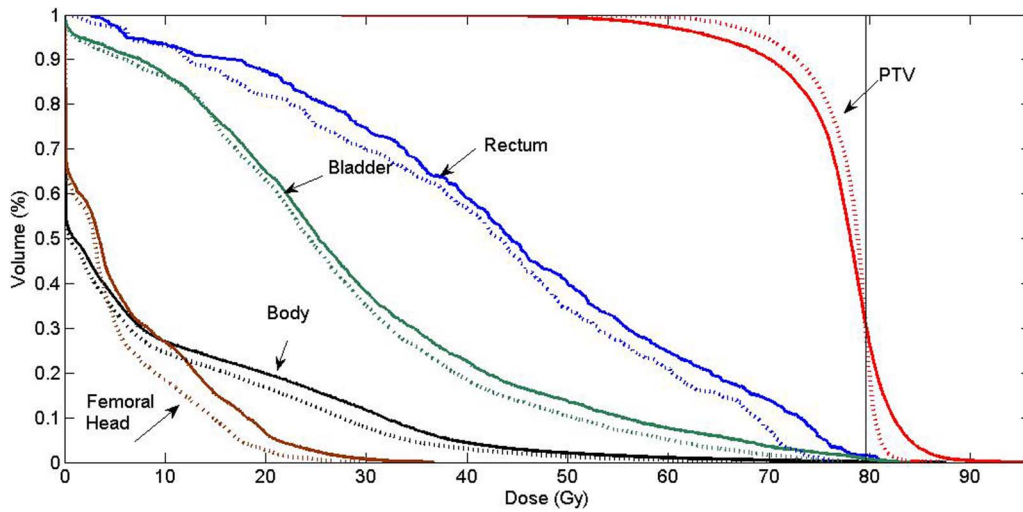
FIG. 6.   Dotted line: The result of the algorithm. Solid line: The feasible reference DVHs.
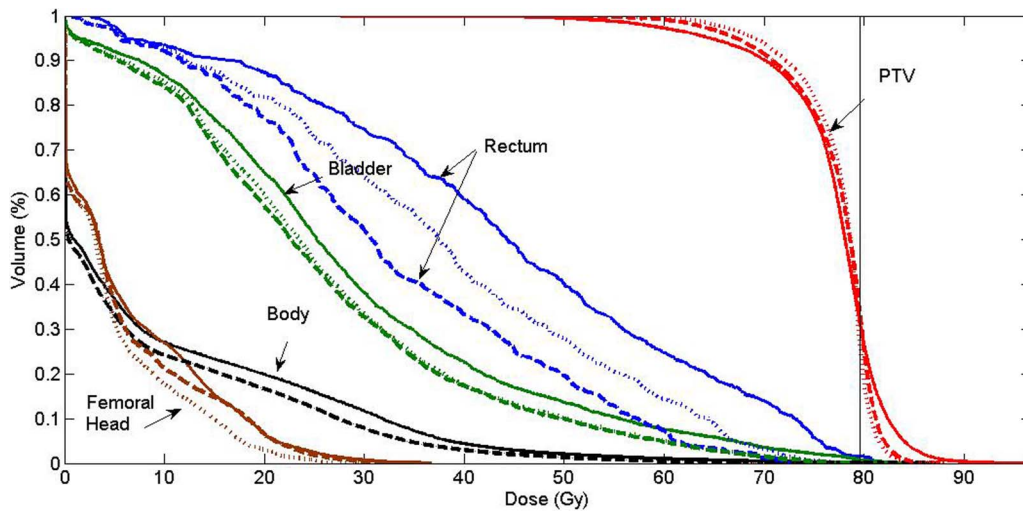

FIG. 7. Investigating the local optimality issue: The algorithm has been run 100 times with randomly generated starting points, and the dotted and the dashed lines represent the best and the worst solutions, respectively. Solid lines are the feasible reference DVHs.
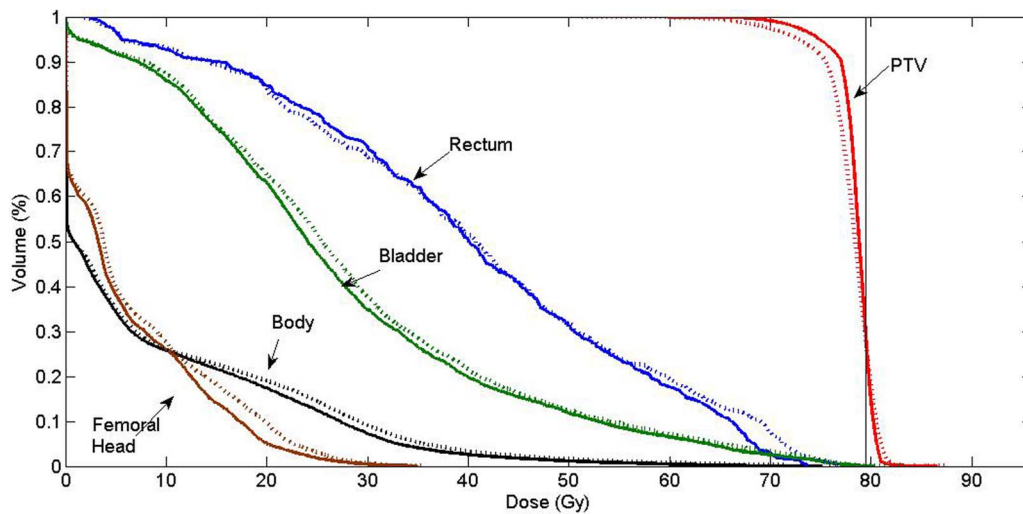

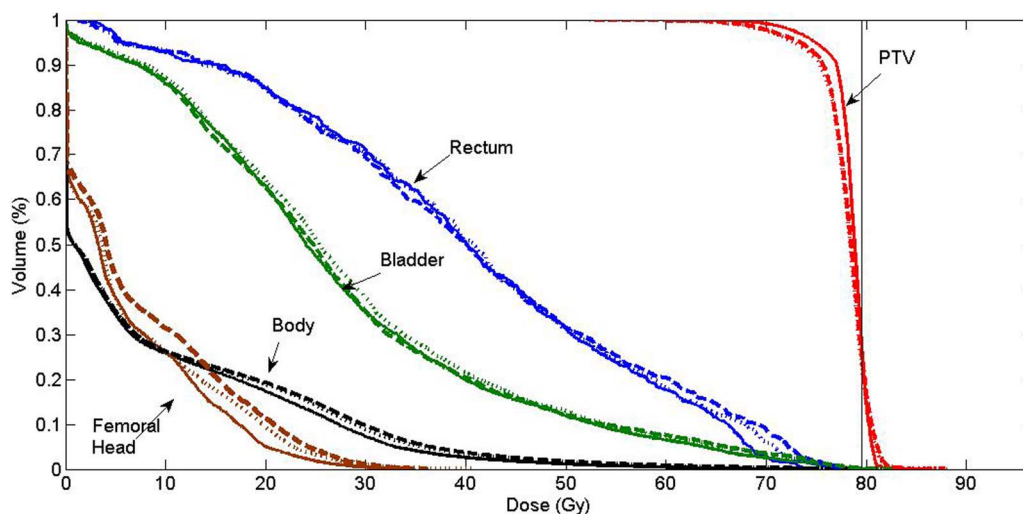FIG. 8.   Dotted line: The result of the algorithm. Solid line: The infeasible reference DVHs.

FIG. 9. Investigating the local optimality issue: The algorithm has been run 100 times with randomly generated starting points, and the dotted and the dashed lines represent the best and the worst solutions, respectively. Solid lines are the infeasible reference DVHs.

optimization algorithm took 11.6 s to run on our inhouse GPU-based research planning system (SCORE).

The solid line in Fig. 10 depicts the original clinical plan developed on the patient's original geometry. The dotted line demonstrates the result of our algorithm, after final dose calculation, when we used the original plan as a reference on the patient's new geometry. As can be seen, the optimal plan for the patient's new geometry is better than the original plan almost everywhere. We looked at the result of the algorithm right after the optimization, before final dose calculation, and we found out the result is even better. It means that the original plan is inside the feasible region corresponding to the patient's new geometry and so the algorithm is able to generate a better plan by projecting the original plan back on the Pareto surface. The significant improvement in plan quality in this figure can be explained with two main possible reasons: (1) the patient's geometry has been changed in a way that made

the original plan located far inside the feasible region of the new geometry leading to the substantial improvement by projecting that plan on the Pareto surface, (2) the proposed algorithm explores a much larger voxel-based Pareto surface of the new geometry that can help to find out a plan with better trade-offs. Another theoretically possible reason (probably unlikely) is that the original plan generated by the commercial software is not a Pareto optimal plan for the original geometry.

Figure 11 compares the original plan delivered on the patient's new geometry (solid line) with the new plan generated by the algorithm (dotted line). This figure verifies the benefit of using the proposed algorithm instead of delivering the original plan on the new geometry. If we do not consider the possible adverse effect of the final dose calculation and MLC transmission, since the original plan delivered on the new geometry is a feasible plan and always inside the feasible region,
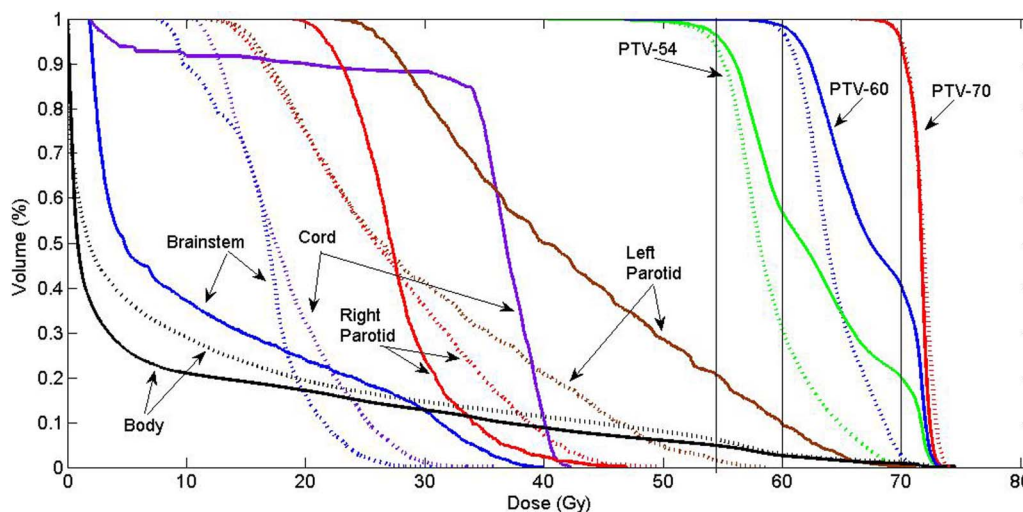

FIG. 10. Solid line: The original plan delivered on the patient's original geometry. Dotted line: The result of the algorithm.
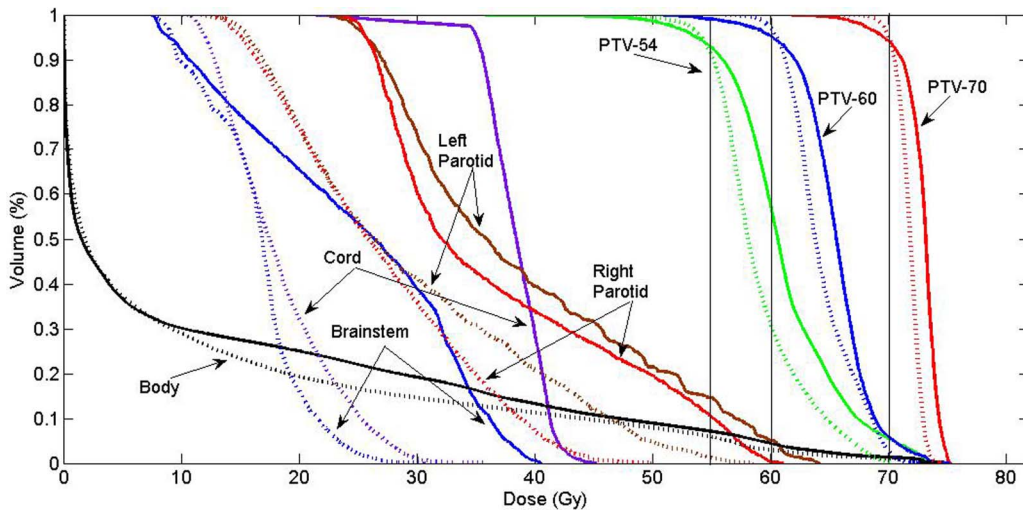
FIG. 11. Solid line: The original plan delivered on the patient's new geometry. Dotted line: The result of the algorithm.

the plan generated by the algorithm would be always better than this plan.

Figure 12 compares the dose distribution of the original plan delivered on the patient's new geometry (a) and (c) with the new plan generated by the algorithm (b) and (d). The dose distribution also verifies the improvement in plan quality in terms of both tumor coverage and healthy tissue sparing. However, since the dose distribution was not part of the optimization process, we cannot draw a solid conclusion about the quality of the dose distribution at this moment. This technique's effects on the dose distribution need either more ex-

perimental study or a change to explicitly consider dose distribution quality as part of the optimization model.

3.C. Automatic treatment planning for a prostate case

With this example we demonstrate how the proposed algorithm can be used for automatic treatment planning. We assume that there is a new patient that we want to treat by using a library of plans of already treated patients with similar medical conditions. We use a prostate case that has been treated previously at UCSD Moores Cancer Center as a new
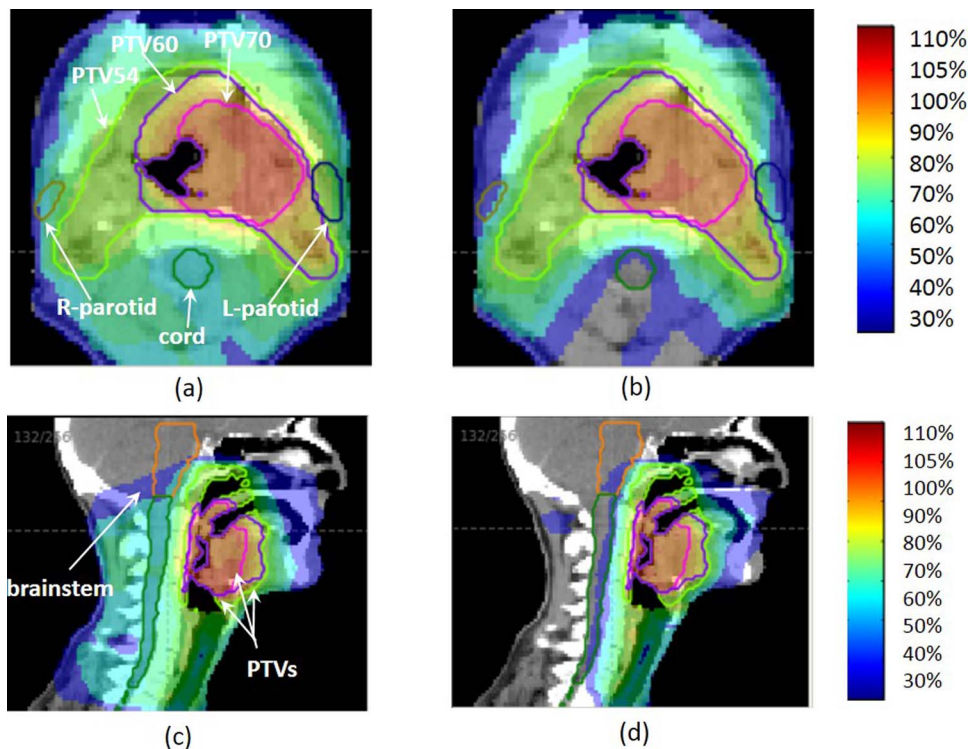

FIG. 12. (a) and (c) The original plan delivered on the patient's new geometry. (b) and (d) The result of the algorithm.
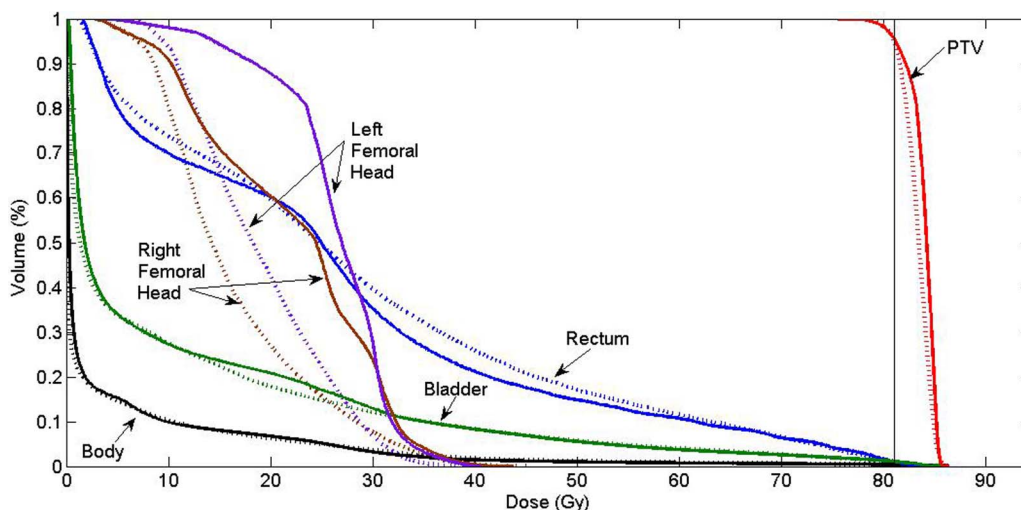
FIG. 13.   Solid line: The DVHs of the most similar patient as the reference DVHs. Dotted line: The result of the algorithm by using the solid line as reference.

patient. The number of the voxels and beamlets for this case are 22 229 and 3412, respectively. First, we went through the library of 99 treated prostate cases and found the most similar one to the new patient in terms of organs geometry.[46] This step is accomplished by employing some machine learning techniques to find out a similarity function that uses some geometric features as inputs (e.g., organ/PTV volume, organ/PTV overlapping volume, Dice index, and mutual information between a pair of two organs/PTVs/intersections) and DVH similarity as an output. Then, we used the DVHs of the most similar patient as the reference DVHs in the algorithm to generate a plan for the new patient.

The solid line in Fig. 13 illustrates the reference DVHs and the dotted line represents the result of the algorithm for the new patient. It can be seen that the result is better that the reference DVHs everywhere except for rectum which is due to the final dose calculation and MLC transmission. It means that the reference DVHs is inside the feasible region of the new patient. It took us 7.2 s to solve the optimization problem

and find the plan for the new patient in our inhouse planning system.

Figure 14 compares the clinical plan that has been delivered for patient treatment (solid line) to the result of the automatic treatment planning (dotted line). These two plans are not exactly the same, but they both provide clinically acceptable trade-offs between different organs.

4. DISCUSSION AND CONCLUSION

The benefits of using the voxel-based model have been discussed before by many authors, however, determining the numerous parameters in the model continues to be a challenge in practice. In this paper, we studied how the prior DVH information that exists in automatic treatment planning and ART replanning can be employed as guidance to tune treatment plan optimization parameters.

We proposed an algorithm that is capable of adjusting voxel weights automatically and navigating the dose
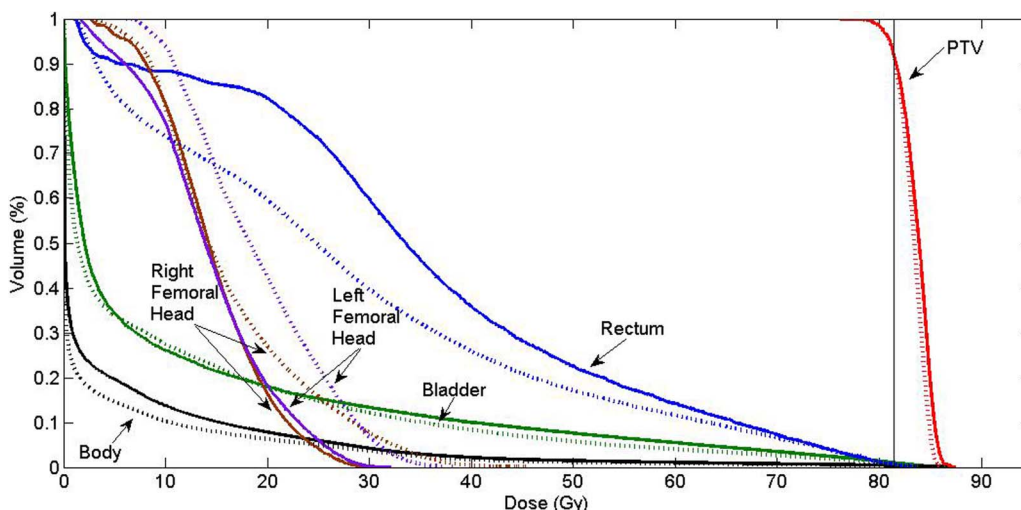

FIG. 14.   Solid line: The original clinical plan delivered for patient treatment. Dotted line: The result of the algorithm by using the automatic treatment planning technique.

distribution Pareto surface to approach a plan with DVHs similar to the given DVHs. The proposed algorithm iteratively approximates the piecewise quadratic Pareto surface with a quadratic one to project the given DVHs back on the Pareto surface. The nonconvexity issue that comes from the insensitivity of the DVH to the spatial properties of the voxels has been handled by iteratively picking some particular permutations of the dose values resulting in the same DVH. The local optimality turned out to be not a big problem based on our limited experience; however, more experimental study needs to be done in this direction to draw a solid conclusion.

We discussed the potential applications of the proposed algorithm in automatic treatment planning and ART replanning. However, the natural question that arises here is "what if the plan generated by the algorithm for automatic treatment planning or ART replanning does not satisfy some of the clinician's criteria in terms of the DVH or dose distribution?" In our next work, we will show how the clinician can interactively tune the DVH curves and dose distribution of the generated plan to make the plan satisfactory.

The capability of the algorithm to project a feasible and infeasible reference plan to the Pareto surface may beg the question of "what if we put an ideal DVH (zero dose in normal tissues and prescribed dose in tumor) or a very lousy DVH as a reference?" The answer to this question lies in the fact that, while the algorithm tries to project a reference plan on the projection area, it cannot guide the reference plan toward a particular part of the projection area. In other words, when the reference needs to be improved (feasible reference) or compromised (infeasible reference) to make the projection, we do not have too much control on the way that improvement or compromise is done. This is not going to be a problem if the reference plan is close to the Pareto surface since the projection area would be relatively small. However, if the reference plan is far from the Pareto surface, like the ideal or very lousy reference, then the projection area would be quite big and the result of the algorithm could be unpredictable. This is actually the main reason that for ART and automatic treatment planning we pick some reference plans that are expected to be close to the Pareto surface.

Finally, we would like to mention that we picked the DVH to guide the optimization process due to its flexibility and independence to the geometry of the patient; however, there are other clinically relevant indices that physicians care about and need to be taken into account in optimization process in order to produce a fully satisfactory plan. We believe that this will be a good direction for our future work.

APPENDIX A: PROOF OF THEOREM 1

First, we prove the following inequality:

$$||\text{sort}(a) - \text{sort}(b)||_p \leq ||a - b||_p, \tag{A1}$$

where $a, b \in \mathbb{R}^n$. The proof is by induction on n. For $n = 2$, the proof is intuitive and we just need to investigate different relative positions of a and b. Assume that the inequality holds for $n = k$ where $k \geq 2$. For $n = k + 1$, let us denote $\text{sort}(a)$

by $c = (c_1, \ldots, c_{k+1})$ and $\text{sort}(b)$ by $d = (d_1, \ldots, d_{k+1})$. Since $(k + 1) \geq 3$, it can be readily shown that there exist $i \in \{1, \ldots, k + 1\}$ such that $a_i > c_1$ and $b_i > d_1$. In fact, a_i and b_i are considered not to be the smallest components of a and b. Now, we have

$$||a - b||_p^p = ||(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{k+1})$$
$$- (b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_k)||_p^p + |a_i - b_i|^p, \tag{A2}$$

$$\geq ||\text{sort}(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_{k+1})$$
$$- \text{sort}(b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_k)||_p^p + |a_i - b_i|^p \tag{A3}$$

$$= |c_1 - d_1|^p + ||w - v||_p^p. \tag{A4}$$

Inequality (A3) is obtained from Eq. (A2) by applying the induction assumption. To get Eq. (A4) from Eq. (A3), we just need to consider that c_1 and d_1 are the smallest components of a and b, and they are not equal to a_i and b_i. w and v are the names for the remaining elements. Now, since w and v have k components, we can employ the induction assumption for $||w - v||_p^p$ and have

$$|c_1 - d_1|^p + ||w - v||_p^p \geq |c_1 - d_1|^p + ||\text{sort}(w)$$
$$- \text{sort}(v)||_p^p = ||\text{sort}(a) - \text{sort}(b)||_p^p. \tag{A5}$$

Inequality (A1) can be obtained from Eqs. (A2) to (A5) and by applying the function $f(x) = x^{\frac{1}{p}}$ to the both sides of the inequality.

Now, the original inequality (4) can be proved by using inequality (A1). For an arbitrary $\pi \in \Pi$, we set $a = d_0^\sigma$ and $b = \pi(\bar{d}^\sigma)$. Applying Eq. (A1) and considering that $\bar{d}^\sigma$ is a sorted vector and $\tilde{\pi}(\bar{d}_0^\sigma) = \text{sort}(\bar{d}_0^\sigma)$, we have

$$\left|\left|\tilde{\pi}\left(d_0^\sigma\right) - \bar{d}^\sigma\right|\right|_p \leq ||\bar{d}_0 - \pi(\bar{d}^\sigma)||_p.$$

Given that the norm is not sensitive to the permutation, we have

$$\left|\left|d_0^\sigma - \tilde{\pi}^{-1}(\bar{d}^\sigma)\right|\right|_p = \left|\left|\tilde{\pi}\left(d_0^\sigma\right) - \bar{d}^\sigma\right|\right|_p$$

which completes the proof.

APPENDIX B: PROOF OF THEOREM 2

Since the quadratic function is an increasing function, the Pareto surface of Problem (3) is exactly the quadratic transformation of the Pareto surface of the following problem:

$$\min_{x \geq 0} w.|Dx - r|. \tag{B1}$$

It means, z belongs to the Pareto surface of Problem (3) if and only if $\sqrt{z}$ belongs to the Pareto surface of Problem (B1). On the other hand, the Pareto surface of Problem (B1) is piecewise linear. Therefore, if $F(z) = 0$ denotes the Pareto surface of Problem (3), then there exists a set of vectors $\{\alpha_i\}_{i=1}^n$ for which we have $F(z) = \alpha.\sqrt{z} - 1$ for $\alpha \in \{\alpha_i\}_{i=1}^n$. Now, we approximate $F(z)$ with $\tilde{F}(z) = \tilde{\alpha}.\sqrt{z} - 1$ where

$\tilde{F}(z)$ passes through the points z^0 and $\bar{z}$. By using the appropriately defined vector multiplication and vector division, we have

$$\nabla \tilde{F}(\bar{z}) = \nabla \tilde{F}(z^0) . \sqrt{\frac{z^0}{\bar{z}}}.$$

Using the approximation $\nabla \tilde{F}(z^0) \cong \nabla F(z^0) = w^0$ and taking into account that $z^0 = (d_0 - p)^2$ and $\bar{z} = (\tilde{\pi}^{-1}(\bar{d}) - p)^2$ would complete the proof.

[a]Author to whom correspondence should be addressed. Electronic mail: Steve.Jiang@UTSouthwestern.edu

[1]M. J. Williams, M. Bailey, D. Forstner, and P. E. Metcalfe, "Multicentre quality assurance of intensity-modulated radiation therapy plans: A precursor to clinical trials," Australas. Radiol. **51**, 472–479 (2007).

[2]I. J. Das, C.-W. Cheng, K. L. Chopra, R. K. Mitra, S. P. Srivastava, and E. Glatstein, "Intensity-modulated radiation therapy dose prescription, recording, and delivery: Patterns of variability among institutions and treatment planning systems," J. Natl. Cancer Inst. **100**, 300–307 (2008).

[3]H. T. Chung, B. Lee, E. Park, J. J. Lu, and P. Xia, "Can all centers plan intensity-modulated radiotherapy (IMRT) effectively? an external audit of dosimetric comparisons between three-dimensional conformal radiotherapy and IMRT for adjuvant chemoradiation for gastric cancer," Int. J. Radiat. Oncol., Biol., Phys. **71**, 1167–1174 (2008).

[4]S. Breedveld, P. R. M. Storchi, M. Keijzer, A. W. Heemink, and B. J. M. Heijmen, "A novel approach to multi-criteria inverse planning for IMRT," Phys. Med. Biol. **52**, 6339–6353 (2007).

[5]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**, 7199–7209 (2009).

[6]S. Breedveld, P. R. M. Storchi, P. W. J. Voet, and B. J. M. Heijmen, "iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans," Med. Phys. **39**, 951–963 (2012).

[7]P. W. J. Voet, M. L. P. Dirkx, S. Breedveld, D. Fransen, P. C. Levendag, and B. J. M. Heijmen, "Toward fully automated multicriterial plan generation: A prospective clinical study," Int. J. Radiat. Oncol., Biol., Phys. **85**, 866–872 (2013).

[8]X. Zhang, X. Li, E. M. Quan, X. Pan, and Y. Li, "A methodology for automatic intensity-modulated radiation treatment planning for lung cancer," Phys. Med. Biol. **56**, 3873–3893 (2011).

[9]V. Chanyavanich, S. K. Das, W. R. Lee, and J. Y. Lo, "Knowledge-based IMRT treatment planning for prostate cancer," Med. Phys. **38**, 2515–2522 (2011).

[10]B. Wu, F. Ricchetti, G. Sanguineti, M. Kazhdan, P. Simari, R. Jacques, R. Taylor, and T. McNutt, "Data-driven approach to generating achievable dose–volume histogram objectives in intensity-modulated radiotherapy planning," Int. J. Radiat. Oncol., Biol., Phys. **79**, 1241–1247 (2011).

[11]D. Yan, F. Vicini, J. Wong, and A. Martinez, "Adaptive radiation therapy," Phys. Med. Biol. **42**, 123–132 (1997).

[12]R. Mohan, X. Zhang, H. Wang, Y. Kang, X. Wang, H. Liu, K. K. Ang, D. Kuban, and L. Dong, "Use of deformed intensity distributions for online modification of image-guided IMRT to account for interfractional anatomic changes," Int. J. Radiat. Oncol., Biol., Phys. **61**, 1258–1266 (2005).

[13]L. E. Court, L. Dong, A. K. Lee, R. Cheung, M. D. Bonnen, J. O'Daniel, H. Wang, R. Mohan, and D. Kuban, "An automatic CT-guided adaptive radiation therapy technique by online modification of multileaf collimator leaf positions for prostate cancer," Int. J. Radiat. Oncol., Biol., Phys. **62**, 154–163 (2005).

[14]Y. Feng, C. Castro-Pareja, R. Shekhar, and C. Yu, "Direct aperture deformation: An interfraction image guidance strategy," Med. Phys. **33**, 4490–4498 (2006).

[15]L. E. Court, R. B. Tishler, J. Petit, R. Cormack, and L. Chin, "Automatic online adaptive radiation therapy techniques for targets with significant shape change: A feasibility study," Phys. Med. Biol. **51**, 2493–2501 (2006).

[16]A. Mestrovic, M.-P. Milette, A. Nichol, B. G. Clark, and K. Otto, "Direct aperture optimization for online adaptive radiation therapy," Med. Phys. **34**, 1631–1646 (2007).

[17]E. E. Ahunbay, C. Peng, G.-P. Chen, S. Narayanan, C. Yu, C. Lawton, and X. A. Li, "An on-line replanning scheme for interfractional variations," Med. Phys. **35**, 3607–3615 (2008).

[18]E. E. Ahunbay, C. Peng, A. Godley, C. Schultz, and X. A. Li, "An on-line replanning method for head and neck adaptive radiotherapy," Med. Phys. **36**, 4776–4790 (2009).

[19]Q. J. Wu, D. Thongphiew, Z. Wang, B. Mathayomchan, V. Chankong, S. Yoo, W. R. Lee, and F.-F. Yin, "On-line re-optimization of prostate IMRT plans for adaptive radiation therapy," Phys. Med. Biol. **53**, 673–691 (2008).

[20]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).

[21]M. Langer and J. Leong, "Optimization of beam weights under dose-volume restrictions," Int. J. Radiat. Oncol., Biol., Phys. **13**, 1255–1260 (1987).

[22]M. Langer, R. Brown, M. Urie, J. Leong, M. Stracher, and J. Shapiro, "Large scale optimization of beam weights under dose-volume restrictions," Int. J. Radiat. Oncol., Biol., Phys. **18**, 887–893 (1990).

[23]Y. Censor, M. D. Altschuler, and W. D. Powlis, "On the use of Cimmino's simultaneous projections method for computing a solution of the inverse problem in radiation therapy treatment planning," Inverse Probl. **4**, 607–623 (1988).

[24]S. M. Morrill, R. G. Lane, J. A. Wong, and I. I. Rosen, "Dose-volume considerations with linear programming optimization," Med. Phys. **18**, 1201–1210 (1991).

[25]P. S. Cho, S. Lee, R. J. M. Ii, S. Oh, S. G. Sutlief, and M. H. Phillips, "Optimization of intensity modulated beams with volume constraints using two methods: Cost function minimization and projections onto convex sets," Med. Phys. **25**, 435–443 (1998).

[26]S. V. Spirou and C. S. Chui, "A gradient inverse planning algorithm with dose-volume constraints," Med. Phys. **25**, 321–321 (1998).

[27]D. Michalski, Y. Xiao, Y. Censor, and J. M. Galvin, "The dose–volume constraint satisfaction problem for inverse treatment planning with field segments," Phys. Med. Biol. **49**, 601–616 (2004).

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

[29]Y. Zhang and M. Merritt, "Dose-volume-based IMRT fluence optimization: A fast least-squares approach with differentiability," Linear Algebra Appl. **428**, 1365–1387 (2008).

[30]M. Zarepisheh, M. Shakourifar, G. Trigila, P. S. Ghomi, S. Couzens, A. Abebe, L. Noreña, W. Shang, S. B. Jiang, and Y. Zinchenko, "A moment-based approach for DVH-guided radiotherapy treatment plan optimization," Phys. Med. Biol. **58**, 1869–1887 (2013).

[31]C. Cotrutz and L. Xing, "Using voxel-dependent importance factors for interactive DVH-based dose optimization," Phys. Med. Biol. **47**, 1659–1669 (2002).

[32]C. Cotrutz and L. Xing, "IMRT dose shaping with regionally variable penalty scheme," Med. Phys. **30**, 544–544 (2003).

[33]C. Wu, G. H. Olivera, R. Jeraj, H. Keller, and T. R. Mackie, "Treatment plan modification using voxel-based weighting factors/dose prescription," Phys. Med. Biol. **48**, 2479–2491 (2003).

[34]Y. Yang and L. Xing, "Inverse treatment planning with adaptively evolving voxel-dependent penalty scheme," Med. Phys. **31**, 2839–2839 (2004).

[35]Z. Shou, Y. Yang, C. Cotrutz, D. Levy, and L. Xing, "Quantitation of the a priori dosimetric capabilities of spatial points in inverse planning and its significant implication in defining IMRT solution space," Phys. Med. Biol. **50**, 1469–1482 (2005).

[36]C. Holdsworth, M. Kim, J. Liao, and M. Phillips, "The use of a multiobjective evolutionary algorithm to increase flexibility in the search for better IMRT plans," Med. Phys. **39**, 2261–2274 (2012).

[37]N. Li, M. Zarepisheh, A. Uribe-Sanchez, K. Moore, Z. Tian, X. Zhen, Y. J. Graves, Q. Gautier, L. Mell, L. Zhou, X. Jia, and S. Jiang, "Automatic treatment plan re-optimization for adaptive radiotherapy guided with the initial plan DVHs," Phys. Med. Biol. **58**, 8725–8738 (2013).

[38]M. Zarepisheh, A. Uribe-Sanchez, N. Li, X. Jia, and S. Jiang, "A multi-criteria framework with voxel-dependent parameters for radiotherapy treatment plan optimization," Med. Phys. **41**, 041705 (10pp.) (2014).

[39]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).

[40] A. L. Hoffmann, D. den Hertog, A. Y. D. Siem, J. H. A. M. Kaanders, and H. Huizenga, "Convex reformulation of biologically-based multi-criteria intensity-modulated radiation therapy optimization including fractionation effects," Phys. Med. Biol. **53**, 6345–6362 (2008).

[41] D. Craft, T. Halabi, and T. Bortfeld, "Exploration of tradeoffs in intensity-modulated radiotherapy," Phys. Med. Biol. **50**, 5857–5857 (2005).

[42] M. L. Kessler, D. L. McShan, M. A. Epelman, K. A. Vineberg, A. Eisbruch, T. S. Lawrence, and B. A. Fraass, "Costlets: A generalized approach to cost functions for automated optimization of IMRT treatment plans," Optim. Eng. **6**, 421–448 (2005).

[43] P. Xia, N. Yu, L. Xing, X. Sun, and L. J. Verhey, "Investigation of using a power function as a cost function in inverse planning optimization," Med. Phys. **32**, 920–920 (2005).

[44] M. Ehrgott, *Multicriteria optimization* (Springer Verlag, Berlin, 2005).

[45] M. Zarepisheh and E. Khorram, "On the transformation of lexicographic nonlinear multiobjective programs to single objective programs," Math. Methods Operations Res. **74**, 217–231 (2011).

[46] M. Karimi, N. Li, M. Zarepisheh, L. Cervino, X. Jia, K. Moore, and S. Jiang, "Selecting reference patients for automatic treatment planning using multiple geometrical features," Med. Phys. **40**, 379 (2013) (Abstract).