Projections onto the Pareto surface in multicriteria radiation therapy optimization

Rasmus BOKRANTZ[*†‡] and Kaisa MIETTINEN[*§]

August 2015

Abstract

Purpose: To eliminate or reduce the error to Pareto optimality that arises in Pareto surface navigation when the Pareto surface is approximated by a small number of plans.

Methods: We propose to project the navigated plan onto the Pareto surface as a post-processing step to the navigation. The projection attempts to find a Pareto optimal plan that is at least as good as or better than the initial navigated plan with respect to all objective functions. An augmented form of projection is also suggested where dose-volume histogram constraints are used to prevent that the projection causes a violation of some clinical goal. The projections were evaluated with respect to planning for intensity modulated radiation therapy delivered by step-and-shoot and sliding window, and spot-scanned intensity modulated proton therapy. Retrospective plans were generated for a prostate and a head and neck case.

Results: The projections led to improved dose conformity and better sparing of organs at risk (OARs) for all three delivery techniques and both patient cases. The mean dose to OARs decreased by 3.1 Gy on average for the unconstrained form of the projection and by 2.0 Gy on average when dose-volume histrogram constraints were used. No consistent improvements in target homogeneity were observed.

Conclusions: There are situations when Pareto navigation leaves room for improvement in OAR sparing and dose conformity, for example if the approximation of the Pareto surface is coarse or the problem formulation has too permissive constraints. A projection onto the Pareto surface can identify an inaccurate Pareto surface representation and, if necessary, improve the quality of the navigated plan.

[*]Optimization and Systems Theory, Department of Mathematics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden.

[†]RaySearch Laboratories, Sveavägen 44, SE-103 65 Stockholm, Sweden.

[‡]E-mail: bokrantz@kth.se and rasmus.bokrantz@raysearchlabs.com.

[§]University of Jyvaskyla, Department of Mathematical Information Technology, FI-400 14 University of Jyvaskyla, Finland.

1 Introduction

Radiation therapy treatment planning is generally guided towards fulfilment of a set of physician-defined plan evaluation criteria. These criteria are sometimes incompatible, and the treatment planner is therefore asked to find a suitable tradeoff between the conflicting ones. The main tool that planners then have at their disposal are weights associated with the objective functions that drive the treatment plan optimization. These weights often need extensive tuning before the optimized plan meets approval [14,32], which can be time-consuming. Inefficient plan preparation is undesirable because it can cause a time lag between diagnosis and the first treatment fraction and poses a risk that plan quality is compromised in the interest of time.

Pareto surface navigation is an alternative planning technique that has recently entered clinical use, see, e.g., Craft [9] and references therein. This technique avoids a priori prioritization of the objectives. A representation of all the possible tradeoffs between the objectives is instead calculated, which the planner or physician can explore through linear interpolation of the precalculated plans' doses. Studies indicate that this form of navigation generally permits an acceptable plan to be identified within a time frame of tens of minutes or less [12, 30]. The mathematical basis for Pareto surface navigation is multicriteria optimization, meaning optimization with multiple objectives where any feasible solution such that no objective can be improved without deteriorating at least one of the others is consider optimal (Pareto optimal), see, e.g., Miettinen [21].

The benefits of navigation are at the cost of that interpolation between precalculated plans introduces an error to Pareto optimality. Algorithms exist that can bound the magnitude of this error [4,5,27], but the number of plans that are required to maintain a given error bound increases exponentially with the number of objectives in the worst case (because hypervolume grows exponentially with increasing dimension). To some relief, studies report that the relation between the required number of plans and the number of objectives is more benign for radiation therapy optimization [8, 10]. Nevertheless, Craft and Bortfeld [10] and Bokrantz [4] both observed approximation errors above 10 % for representations with less than about 20 plans, and that up to about 75 plans are needed to reduce the error below 5 %. These studies considered between five to ten objectives.

In view of these concerns, we present a technique that eliminates or reduces the error to Pareto optimality through the minimization of a projective distance between the navigated plan and the Pareto surface. We use a formulation of this projection as in Nakayama [23], which attempts to find a plan that is at least as good as or better than the navigated plan with respect to all objectives. We also suggest an augmented formulation where constraints are imposed on maintained dose-volume histogram (DVH) quality. We quantify the dosimetric benefit of the suggested technique by application to planning for step-and-shoot intensity-modulated radiation therapy (ss-IMRT), sliding window intensity-modulated radiation therapy (sw-IMRT), and spot-scanned intensity-modulated proton therapy (IMPT).

2 Methods

2.1 Pareto surface-based planning

We formulate treatment planning for radiation therapy as a multicriteria optimization problem with n objective functions $f_1, \ldots, f_n$ that are to be minimized with respect to the vector of variables x. The minimization occurs over a feasible set $\mathcal{X}$ that represents the physical limitations of the delivery method and, possibly, constraints on the planned dose, according to

$$\begin{array}{ll} \underset{x}{\text{minimize}} & f(x) = \Big[f_1(x) \cdots f_n(x) \Big]^T \\ \text{subject to} & x \in \mathcal{X}. \end{array} \tag{1}$$

We understand optimality to this formulation in a Pareto sense, meaning that a feasible x^* is Pareto optimal if there is no feasible x such that $f_i(x) \leq f_i(x^*)$ for $i = 1, \ldots, n$, with a strict inequality for at least one index i. A feasible x^* is called weakly Pareto optimal if there is no feasible x such that $f_i(x) < f_i(x^*)$ for $i = 1, \ldots, n$.

Formulation (1) has an infinite number of Pareto optimal solutions in general. To solve this problem from a practical perspective therefore entails to select the single, best preferred, Pareto optimal solution. We perform this selection by Pareto surface navigation, meaning that a representative set of Pareto optimal solutions $x_1, \ldots, x_m$ is first calculated and a convex combination $\bar{x} = \sum_{j=1}^{m} \lambda_j x_j$ of these solutions then selected, where the components of λ need to be nonnegative and sum to unity. The selection is guided by a navigation interface that permits the priorities of each objective to be continuously adjusted using associated slider bar controls. The sliders are coupled to an algorithm that updates λ accordingly, see, e.g., Craft et al. [11] and Monz et al. [22]. For the navigation to be valid, we assume that $\mathcal{X}$ is a nonempty and convex set and that all functions $f_1, \ldots, f_n$ are convex and bounded on $\mathcal{X}$.

2.2 Projection onto the Pareto surface

The navigated plan $\bar{x}$ is feasible thanks to the assumed convexity of formulation (1). It also has objective values that are bounded by Jensen's inequality for convex functions, i.e., $f_i(\sum_{j=1}^{m} \lambda_j x_j) \leq \sum_{j=1}^{m} \lambda_j f_i(x_j)$ for $i = 1, \ldots, n$, as illustrated in Figure 1 for the case $m = n = 2$. The figure also shows that the image of the Pareto optimal solutions in the objective space forms a connected surface in the boundary of a convex set [28, Proposition 2.3], but the surface is not convex itself, because convex surfaces are generally not convex sets. The navigated point $f(\bar{x})$ is therefore merely known to lie between the Pareto surface and the set of objective vectors that can be formed as convex combinations of the known points $f(x_1), \ldots, f(x_m)$.

To mitigate that the navigated plan $\bar{x}$ in general is not Pareto optimal, we propose to convert this plan into a Pareto optimal plan with at least as good or better performance in all objectives. Specifically, we propose to solve the following opti-

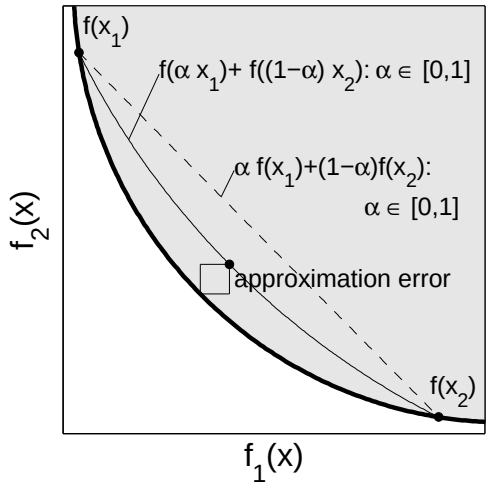

Figure 1: Continuous navigation between two Pareto optimal solutions x_1 and x_2. The shaded area indicates the feasible objective space $f(\mathcal{X})$, the thick solid line indicates the set of Pareto optimal solutions, the thin solid line indicates the set of convex combinations of x_1 and x_2, and the dashed line indicates the upper bound on the navigated objective values from Jensens' inequality. The componentwise error between a navigated plan and the Pareto surface is indicated by a square.

mization problem as a post-processing step to the navigation:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \max\limits_{i=1,\ldots,n} \left\{ \dfrac{f_i(x) - f_i(\bar{x})}{f_i(\bar{x}) - z_i^*} \right\} \\ \text{subject to} & x \in \mathcal{X}, \end{array} \tag{2}$$

where z^* is the ideal point, i.e., $z_i^* = \min_{j=1,\ldots,m} f_i(x_j)$ for $i = 1, \ldots, n$. This formulation is a variant of an achievement function suggested by Wierzbicki [31], which projects the navigated point $f(\bar{x})$ along a ray towards the ideal objective function vector z^*, see Figure 2. The particular achievement function that is used in (2) has been used previously by Nakayama [23]. Buchanan and Gardiner [7] observed that decision makers tend to prefer this form of minimization of the distance to the ideal point over maximization of the distance from the worst feasible point if the reference point is attainable, as is the case for $f(\bar{x})$.

The navigated plan $\bar{x}$ is a feasible solution to (2) with an objective value of zero. An optimal solution x^* to (2) therefore satisfies $f_i(x^*) \leq f_i(\bar{x})$ for $i = 1, \ldots, n$, meaning that x^* is as least as good as or better than the navigated plan with respect to all objectives. Further, x^* is weakly Pareto optimal because the achievement function in (2) is strictly increasing [21, Theorem 3.5.4]. If solutions that are weakly Pareto optimal but not Pareto optimal are to be avoided, it is possible to augment the objective function of (2) with the term $\rho \sum_{i=1}^{n} f_i(x)/(f_i(\bar{x}) - z_i^*)$ for some sufficiently small positive scalar ρ, as discussed in Miettinen [21, Section 5.8]. Addition of this term makes the objective function strongly increasing, which ensures Pareto optimality [21, Theorem 3.5.4]. With regard to numerical stability, observe that

the denominator in (2) can approach zero if the navigated point $f(\bar{x})$ becomes very close to the ideal point z^* in some component. The ideal point should therefore for numerical purposes be replaced with an utopian objective vector that is better by some small but numerically significant positive value [21, Definition 2.4.2].

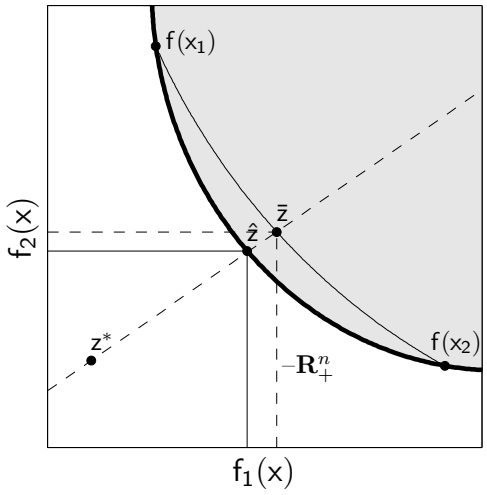

Figure 2: Projection of a navigated plan onto the Pareto surface. The point $\bar{z} = f(\bar{x})$ is shifted towards z^* along the dashed line until it intersects with the boundary of the feasible objective space. The projected point $\hat{z}$ is nondominated, as illustrated by that no other point is contained in the cone $-\mathbb{R}^n_+$ that emanates from this point.

2.3 Dose-volume histogram constraints

A commonly used formulation for radiation therapy optimization is penalization of the deviation from the desired dose to each anatomical structure, see, e.g., Oelfke and Bortfeld [26]. Objective functions of this type cannot capture all aspects of plan quality, for instance, they do not take the three-dimensional shape of the dose into account nor the biological effect of the irradiation. It is therefore possible for a clinician to judge a plan obtained from formulation (1) as worse than the initial navigated plan even though it is better as measured by all objectives $f_1, \ldots, f_n$. To mitigate any deterioration in plan quality not captured by the objectives, we consider an augmented version of formulation (2) that prevents deterioration with respect to clinical goals. We restrict ourselves to consider clinical goals that are related to the DVH distribution of the navigated plan. Consequently, we introduce constraints that require each DVH curve of the projected plan to lie between the corresponding DVH curve for the navigated plan and a vertical line that intersects the dose axis at the prescription level for targets and at zero for organs at risk (OARs), see Figure 3. These requirements ensure that a DVH criterion that is satisfied by the navigated plan cannot become violated after the projection.

The DVH requirements are implemented using functions that impose a one-sided penalty on the error between the DVH curves associated with the current dose d

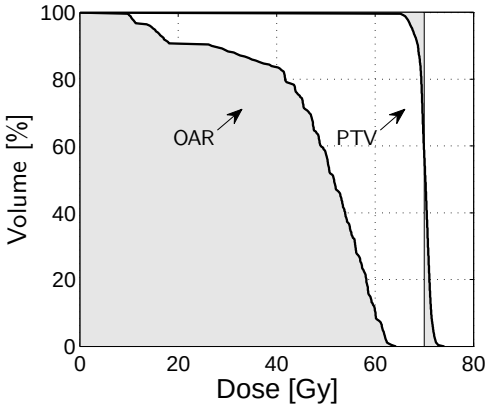

Figure 3: Feasible DVH region (shaded areas) for projection onto the Pareto surface under DVH constraints. The thick solid lines indicates the DVH of the navigated plan.

and the navigated dose $\bar{d}$. The one-sidedness prevents penalization of when d has better normal tissue sparing or target coverage than $\bar{d}$. To define the DVH functions mathematically, let $D(\cdot; d)$ be the function that parametrizes the DVH curve for some subvolume $\mathcal{V}$ along the cumulative volume axis, i.e.,

$$D(v; d) = \max \left\{ d' \in \mathbb{R} : \frac{|q \in \mathcal{V} : d(q) \geq d'|}{|\mathcal{V}|} \geq v \right\},$$

where v denotes cumulative volume in percent, $d(q)$ is the dose to some point $q \in \mathcal{V}$, and $|\mathcal{V}|$ is the volume of $\mathcal{V}$. In other words, $D(v; d)$ is the DVH point associated with the some dose d and cumulative volume v. Let also $\hat{d}$ denote the prescription level for targets and zero for OARs, $\hat{v}$ denote the cumulative volume such that $\hat{d} = D(\hat{v}, \bar{d})$, and $(\cdot)_+$ denote the positive part operator $\max\{\cdot, 0\}$. Then, a min reference DVH constraint takes the form

$$\int_{\hat{v}}^{1} \left(D(v; \bar{d}) - D(v; d) \right)_+ \, dv \leq 0,$$

while a max reference DVH constraint takes the form

$$\int_{0}^{\hat{v}} \left(D(v; d) - D(v; \bar{d}) \right)_+ \, dv \leq 0.$$

The requirements in Figure 3 are implemented by assignment of a max reference DVH constraint to each OAR, and assignment of a min reference DVH and max reference DVH constraint to each target. Similar reference DVH functions have been used previously in Bokrantz [3] and Fredriksson [13].

The use of DVH constraints leads to a nonconvex optimization formulation. An optimization solver can therefore not guarantee more than convergence to a locally optimal point. Some studies have, however, shown that the nonconvexity of DVH

functions does not lead to severe local optimality effects in practice [18, 20, 33], possibly because the nonconvexities disappear or become negligible due to the physical properties of radiation delivery [9]. Regardless of convergence, the introduction of additional constraints implies that the projected solution in general lies in the interior of the feasible objective space and not on the Pareto surface itself.

2.4 Computational study

The projections were evaluated with respect to three delivery techniques: ss-IMRT, sw-IMRT, and IMPT; and two tumor sites: prostate and head and neck. We first outline our numerical implementation, then describe the studied patient cases and delivery techniques, and finally summarize our set of evaluation criteria.

2.4.1 Numerical optimization

The projections were implemented as an add-on to the multicriteria optimization module in RayStation v2.4 (RaySearch Laboratories, Stockholm, Sweden) [1]. This treatment planning system's optimization solver (a quasi-Newton sequential quadratic programming method) requires a continuously differentiable objective and constraints. The nondifferentiability of the maximum in (2) was handled by the addition of $1+\epsilon$, with ϵ being a positive infinitesimal value, to the arguments of the maximization (so that they become positive, see also Nakayama [24, Remark 3.1]), followed by a substitution a smooth power mean function for the maximum operator according to

$$\max_{i=1,\ldots,n}\{x_i\} \approx \left(\frac{1}{n}\sum_{i=1}^{n} x_i^p\right)^{1/p}. \tag{3}$$

This approximation approaches the exact maximum of some positive $x_1, \ldots, x_n$ as $p \to \infty$. A parameter value of $p = 10$ was used for all numerical experiments in this paper, guided by the close relationship between (3) and an equivalent uniform dose (EUD) functions (the EUD is the power mean of the dose, with the power p determined by the EUD parameter a, cf. Niemerko [25]). An EUD function with an EUD parameter of about 10 is common for serial organs such as the spinal cord where the risk for complication highly depends on the maximum dose value, see, e.g., Thieke et al. [29].

The approximation (3) means that the results in Section 2.2 do not hold rigorously for our numerical implementation. In particular, it is possible that the projection can lead to a mild degradation in objective function value compared to the navigated point for some objectives. Nevertheless, formulation (3) amounts to minimization of a strongly increasing achievement function, and it therefore finds Pareto optimal points [21, Theorem 3.5.4]. If it is critical to maintain objective function values exactly, then an everywhere differentiable epigraph reformulation

of (2) according to

$$\begin{aligned} \underset{x,t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & \frac{f_i(x) - f_i(\bar{x})}{f_i(\bar{x}) - z_i^*} \leq t, \quad i = 1, \ldots, n, \\ & x \in \mathcal{X}, \end{aligned}$$

could be preferable to (3). The auxiliary nonlinear constraints of this formulation, however, makes it more computationally expensive to solve than formulation (2) combined with the approximation (3).

2.5 Patient data and machine model

Retrospective planning was performed with respect to the following two patient cases:

- A prostate cancer patient with a prescribed dose of 59.2 Gy to the prostate and seminal vesicles, with a simultaneous boost of 74 Gy to the prostate. Considered critical structures were the bladder and rectum.

- A head and neck cancer patient with a prescribed dose of 66 Gy, 60 Gy, and 50 Gy to the primary target, high risk nodal regions, and low risk nodal regions, respectively. Considered critical structures were the brainstem, parotid glands, and spinal cord.

Treatment planning for IMRT was performed with respect to a Varian 2100 linear accelerator (Varian Medical Systems, Palo Alto, California), with ten static segments per field for ss-IMRT and 320 dynamic control points per field for sw-IMRT. A coplanar five-field setup was used for the prostate case and a coplanar seven-field setup used for the head and neck case. Planning for IMPT was for the prostate case performed with respect to two coplanar and parallel-opposed fields and for the head and neck case performed with respect to two coplanar fields with a perpendicular setup. A dose grid resolution of $3 \times 3 \times 3\,\text{mm}^3$ was used for all calculations. The optimizations were performed with respect to least-squares penalties on the deviation in voxel dose or EUD from a scalar-valued reference level, see Appendix A for a complete list functions and Bokrantz [3, Appendix C] for mathematical definitions.

2.6 Treatment plan generation

A total of $2n$ Pareto optimal plans was generated per delivery technique and patient case. RayStation uses the algorithm in Bokrantz and Forsgren [3] for this calculation. A single plan was then selected by Pareto surface navigation, and the projection finally applied with or without DVH constraints. Deliverable plans were for comparative purposes generated without performing any projection. Minor differences between the studied delivery techniques are elaborated in the following three subsections.

2.6.1 Step-and-shoot IMRT

The plans in the Pareto surface representation were generated by fluence map optimization with respect to dose calculated using a singular value decomposition (SVD) of pencil beam kernels, similar to Bortfeld et al. [6]. The navigated plan was made deliverable by an optimization where the error in DVH due to the conversion was minimized, see Bokrantz [3]. This minimization was performed with leaf positions and segment weights as variables, see Hårdemark et al. [15], with the SVD dose augmented with intermediate and final dose calculations performed using a collapsed cone (CC) algorithm, see, e.g., Ahnesjö [2]. The projection was performed using the same set of variables and dose calculation algorithms.

2.6.2 Sliding window IMRT

The calculation of the Pareto surface representation and the projection were both performed with fluence as variables. The projected plan was converted to control points by sliding window conversion, see, e.g., Kamath [19]. All dose distributions were calculated by SVD.

2.6.3 IMPT

The calculation of the Pareto surface representation and the projection were both performed with spot weights as variables. All dose distributions were calculated using a pencil beam algorithm.

2.7 Evaluation criteria

Plan quality was assessed with respect to a selection of dose-volume statistics. The dose to OARs was assessed in terms of dose-to-volume levels V_x (the fractional volume of a structure that receives a dose greater than or equal to $x\,\mathrm{Gy}$), volume-to-dose levels D_x (the minimum dose such that the associated isodose volume contains $x\,\%$ of the volume of a structure), and mean dose levels $\bar{\mathrm{D}}$. The planned dose to target structures was assessed in terms of a homogeneity index (HI) [17] according to

$$\mathrm{HI} = (\mathrm{D}_2 - \mathrm{D}_{98}) \,/\, \mathrm{D}_{50},$$

and a conformity index (CI) [16] according to

$$\mathrm{CI} = \mathrm{V}^{\mathrm{External}}_{95\,\%} \,/\, \mathrm{V}^{\mathrm{PTV}},$$

where $\mathrm{V}_{95\,\%}$ is the volume contained within the isodose volume defined at $95\,\%$ of the prescription level and $\mathrm{V}^{\mathrm{PTV}}$ the total volume of all targets with prescription level greater than or equal to the prescription level of the structure to which the index refers.

3 Results

Our numerical results are summarized by the DVHs in Figure 4 and the dose statistics in Tables 1 and 2. Results are shown for plans subject to no projection, subject to a projection without DVH constraints, and subject to a projection under DVH constraints. Planning target volumes (PTVs) are designated by their prescription level in subscript.

Table 1: Dose statistics for the prostate case. Values where the projection resulted in a relative improvement of 5 % are indicated in bold.

Plan		PTV_{74}		$PTV_{59.2}$		Bladder		Rectum	
		HI [%]	CI [%]	HI [%]	CI [%]	D_{10} [Gy]	D [Gy]	D_{10} [Gy]	D [Gy]
ss-IMRT	No projection	8.2	117.6	27.9	111.1	55.6	26.6	56.1	31.9
	Projected	8.9	115.7	27.6	107.7	54.6	**24.4**	54.1	**26.4**
	DVH proj.	8.4	116.1	27.7	108.1	55.3	25.4	54.4	**28.2**
sw-IMRT	No projection	7.9	117.0	27.3	108.7	54.1	24.9	56.2	33.5
	Projected	9.2	117.0	27.3	109.0	54.7	**23.6**	54.7	**25.5**
	DVH proj.	7.8	116.3	27.1	108.0	54.1	24.2	55.7	**28.6**
IMPT	No projection	7.5	115.3	26.4	111.6	54.5	13.4	55.3	16.5
	Projected	8.6	114.1	27.0	**103.8**	53.7	**12.3**	55.9	**14.8**
	DVH proj.	7.7	115.1	26.8	**104.8**	54.1	13.0	55.2	**14.2**

Table 2: Dose statistics for the head and neck case. Values where the projection resulted in a relative improvement of 5 % or more are indicated in bold.

Plan		PTV_{66}		PTV_{60}		PTV_{50}		L Parotid		R Parotid	
		HI [%]	CI [%]	HI [%]	CI [%]	HI [%]	CI [%]	V_{30} [%]	D [Gy]	V_{30} [%]	D [Gy]
ss-IMRT	No projection	6.8	132.4	12.0	160.1	18.6	152.7	52.1	34.2	29.1	22.4
	Projected	6.6	127.4	**11.3**	**147.8**	18.8	**138.2**	**49.4**	**30.7**	30.0	21.9
	DVH proj.	**6.3**	127.2	11.6	**151.7**	18.5	**144.0**	51.1	**32.2**	28.7	21.9
sw-IMRT	No projection	6.4	124.3	10.4	145.0	19.3	136.5	52.2	35.6	28.6	22.5
	Projected	6.8	124.1	10.1	141.9	18.8	134.4	**46.9**	**30.1**	29.6	21.9
	DVH proj.	6.9	123.5	10.0	141.8	19.1	133.5	**47.4**	**30.3**	28.0	21.5
IMPT	No projection	9.2	132.4	12.7	132.3	16.8	127.8	45.7	26.8	21.3	18.3
	Projected	**8.4**	126.8	**11.2**	127.5	16.9	**117.5**	**38.2**	**22.5**	**20.0**	**15.2**
	DVH proj.	**8.4**	129.9	12.3	128.5	16.8	120.8	**44.1**	25.6	20.8	17.9

The projection of a navigated plan onto the Pareto surface led to improved OAR sparing and better dose conformity for all three delivery techniques and both patient cases. The improved OAR sparing was most pronounced in the low and moderate dose regions whereas the high dose regions generally did not improve much, see, e.g., D_{10} for the bladder and rectum of the prostate case in Table 1. The largest improvements in dose conformity occurred for low-dose targets, see, e.g., $PTV_{59.2}$ of the IMPT plan for the prostate case and PTV_{50} and PTV_{60} of the ss-IMRT and IMPT plan for the head and neck case. The improvements in target homogeneity were very minor, except for PTV_{66} of the IMPT plan of the head and neck case.

A projection without DVH constraints only in rare instances led to a deterioration in some dose statistics. Such occurrences are the reduced homogeneity for PTV_{74} of the ss-IMRT and sw-IMRT plan for the prostate case, the increase of D_{10} for the

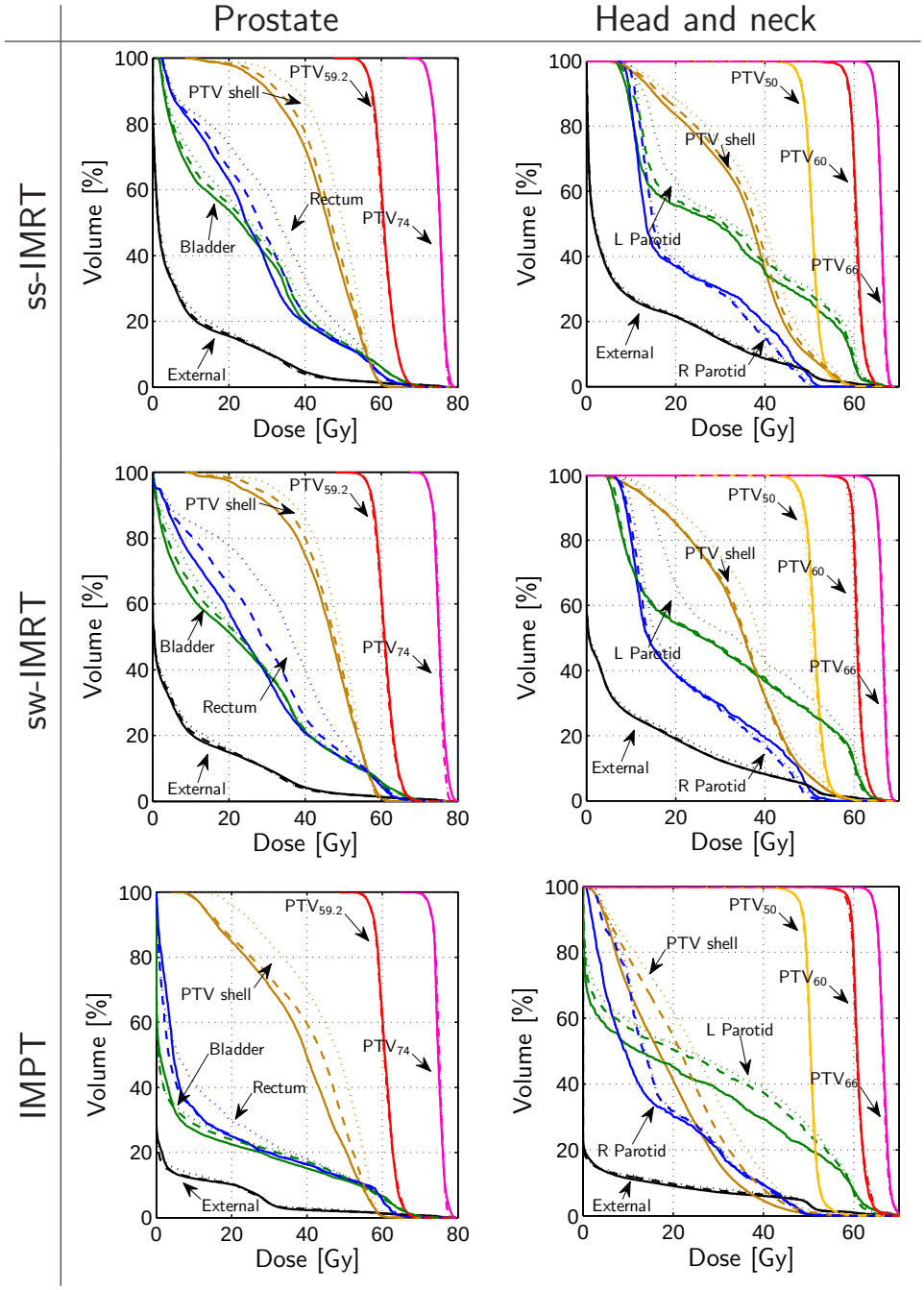

Figure 4: DVH results for projections onto Pareto surface. Plans projected without DVH constraints are indicated by solid lines, plans projected with DVH constraints indicated by dashed lines, and plans generated without performing the projection indicated by dotted lines.

rectum of the IMPT plan for the prostate case, and the decrease in homogeneity for PTV_{66} of the sw-IMRT plan for the head and neck case. These deteriorations did not occur when DVH constraints were used; however, the improvements where then smaller overall.

Table 3 further quantifies the main observed effect of the projections, namely that the OAR sparing was improved but at a mild dose increase in the high dose regions of OARs unless DVH constraints were used. The improved sparing of OARs is in the tabulated data quantified by two-sided DVH curve differences while violations of the navigated DVH are quantified by one-sided DVH curve differences. Clearly, a projection onto the Pareto surface poses a tradeoff between the degree of plan improvement and the acceptable violation of the navigated DVH. The average decrease in dose to OARs and the average violation of the navigated DVH was 3.13 Gy and 0.18 Gy, respectively. These figures should be contrasted to an average dose decrease for OARs at 1.98 Gy and a vanishing small violation of the navigated DVH for projection with DVH constraints.

Table 3: Mean two-sided and one-sided differences along the dose axis between DVH curves of projected plans and corresponding plans generated without performing any projection.

Patient		Prostate				Head and neck			
Plan		Bladder		Rectum		L Parotid		R Parotid	
		2-sided [Gy]	1-sided [Gy]	2-sided [Gy]	1-sided [Gy]	2-sided [Gy]	1-sided [Gy]	2-sided [Gy]	1-sided [Gy]
ss-IMRT	Projected	-2.22	0.00	-5.52	0.00	-3.56	0.00	-0.56	0.65
	DVH proj.	-1.20	0.00	-3.70	0.00	-2.01	0.00	-0.48	0.00
sw-IMRT	Projected	-1.39	0.07	-8.04	0.00	-5.54	0.00	-0.55	0.45
	DVH proj.	-0.78	0.01	-4.93	0.00	-5.32	0.00	-0.95	0.00
IMPT	Projected	-1.03	0.16	-1.75	0.25	-4.26	0.05	-3.13	0.03
	DVH proj.	-0.42	0.00	-2.28	0.00	-1.21	0.00	-0.44	0.00

2-sided $= \int_0^1 (D(v,d) - D(v,\bar{d}))\, dv$, 1-sided $= \int_0^1 (D(v,d) - D(v,\bar{d}))_+\, dv$.

4 Discussion

Our results show that there exist situations when a projection onto the Pareto surface can improve OAR sparing and dose conformity of the navigated plan. This conclusion holds true both in the unconstrained case and if DVH constraints are used to preserve clinical goals. The DVH constraints were found effective for prevention of a dose increase in the high-dose region of OARs, but dampened the magnitude of the overall improvements. We therefore only recommend such constraints for structures where it is critical to maintain the navigated DVH exactly, e.g., if the dose to the structure only barely meets the clinical acceptance criteria. Another option is to first perform an unconstrained projection and then perform a second optimization that includes DVH constraints if necessary.

The observed dose improvement due to the projections was of the order of several Grays. This magnitude is, however, a function of the approximation error to Pareto optimality for the navigate plan, which depends on the number of objectives, the

number of plans used to represent the Pareto surface, and how large dose variations that the set of constraints permit. Our study only covers a narrow spectrum of the possible values for these parameters: 8–10 objectives, 16–20 plans in the representations, and a single set of constraints per patient case. The results should therefore not be extrapolated to conclude that Pareto surface navigation yields plans that are sub-optimal in terms of OAR sparing in general. Rather, we envisage that the projections can be used as a learning tool (for example in the training of practitioners), which can identify situations when Pareto surface representations are inaccurate.

Finally, note that the projections are unnecessary if it is possible to calculate such dense Pareto surface representations that the navigated plan is nearly error-free. If this is the case, however, then it is likely that the plan optimization times are so short that they are perceived as occurring in real time. The advance of real-time optimization would permit Pareto surface navigation where the navigated plan is generated on-the-fly instead of being interpolated from a set prealculated plans.

5 Conclusions

We have presented a method that eliminates or reduces the error to Pareto optimality that arises during Pareto surface navigation. The error is removed through minimization of a projective distance to the ideal point in the objective function space. An augmented form of the projection was also suggested where the DVH distribution of the projected solution is required to be at least as good as that of the initial navigated plan. Empirical results with respect to two clinical cases and three delivery techniques show that the projections can lead to improved OAR sparing and better dose conformity at maintained, or slightly improved, target coverage. The main mechanism behind the improvements observed in this study was a reduction of the low to moderate dose to healthy structures.

A Optimization problem formulations

The optimization formulations that were used in the numerical experiments are summarized in Tables 4 and 5.

Table 4: Optimization formulation for the prostate case. The reference dose level of a function is denoted $\hat{d}$. The constraints used during proton and photon therapy planning are indicated by "Pr" and "Ph" in subscript, respectively.

Objectives			Constraints			
Structure	Function	$\hat{d}$ [Gy]	Structure	Function	$\hat{d}_{\mathrm{Ph}}$ [Gy]	$\hat{d}_{\mathrm{Pr}}$ [Gy]
PTV_{74}	Min dose	74.00	PTV_{74}	Min dose	66.60	68.00
	Uniform dose	74.00		Min 95 % DVH	71.78	72.52
$\mathrm{PTV}_{59.2}$	Min dose	59.20		Max dose	81.04	79.92
$\mathrm{PTV}_{59.2}$ - PTV_{74}	Uniform dose	59.20	$\mathrm{PTV}_{59.2}$	Min dose	53.28	53.28
Bladder	Max EUD $a = 2$	0.00		Min 95 % DVH	56.24	56.24
Rectum	Max EUD $a = 2$	0.00	$\mathrm{PTV}_{59.2}$ - PTV_{74}	Max 5 % DVH	66.50	65.71
PTV shell $[5, 15]$ mm	Max EUD $a = 2$	0.00	External	Max dose	81.04	79.92
External	Dose fall-off 2 cm	74.00				

Table 5: Optimization formulation for the head and neck case. The reference dose level of a function is denoted $\hat{d}$. The constraints used during proton and photon therapy planning are indicated by "Pr" and "Ph" in subscript, respectively.

Objectives			Constraints			
Structure	Function	$\hat{d}$ [Gy]	Structure	Function	$\hat{d}_{\mathrm{Ph}}$ [Gy]	$\hat{d}_{\mathrm{Pr}}$ [Gy]
PTV_{66}	Min dose	66.00	PTV_{66}	Min dose	59.40	60.72
	Uniform dose	66.00		Min 95 % DVH	62.70	63.36
PTV_{60}	Min dose	60.00		Max dose	72.60	72.60
	Uniform dose	60.00	PTV_{60}	Min dose	54.00	55.20
PTV_{50}	Min dose	50.00		Min 95 % DVH	57.00	57.60
	Uniform dose	50.00		Max 5 % DVH	66.00	66.00
L Parotid	Max EUD $a = 1$	0.00	PTV_{50}	Min dose	45.00	45.00
R Parotid	Max EUD $a = 1$	0.00		Min 95 % DVH	47.50	47.50
PTV shell $[5, 15]$ mm	Max EUD $a = 2$	0.00		Max 5 % DVH	57.50	55.00
External	Dose fall-off 2 cm	66.00	Brainstem	Max dose	52.00	52.00
			Spinal cord	Max dose	48.00	48.00
			External	Max dose	72.60	72.60

Acknowledgments

The authors thank Anders Forsgren and Björn Hårdemark for careful reading and constructive comments on earlier drafts. The second author's work was carried out during a visiting professorship at the KTH Royal Institute of Technology.

References

[1] Multi criteria optimization in RayStation. RaySearch Laboratories white paper, 2012.

[2] A. Ahnesjö. Collapsed cone convolution of radiant energy for photon dose calculation in heterogeneous media. *Med. Phys.*, 16(4):577–592, 1989.

[3] R. Bokrantz. Multicriteria optimization for volumetric-modulated arc therapy by decomposition into a fluence-based relaxation and a segment weight-based restriction. *Med. Phys.*, 39(11):6712–6725, 2012.

[4] R. Bokrantz. Distributed approximation of Pareto surfaces in multicriteria radiation therapy treatment planning. *Phys. Med. Biol.*, 58(11):3501–3516, 2013.

[5] R. Bokrantz and A. Forsgren. An algorithm for approximating convex Pareto surfaces based on dual techniques. *INFORMS J. Comput.*, 25(2):377–393, 2013.

[6] T. Bortfeld, W. Schlegel, and B. Rhein. Decomposition of pencil beam kernels for fast dose calculations in three-dimensional treatment planning. *Med. Phys.*, 20(2):311–318, 1993.

[7] J. Buchanan and L. Gardiner. A comparison of two reference point methods in multiple objective mathematical programming. *Eur. J. Oper. Res.*, 149(1):17–34, 2003.

[8] D. Craft. Calculating and controlling the error of discrete representations of Pareto surfaces in convex multi-criteria optimization. *Phys. Medica*, 26(4):184–191, 2010.

[9] D. Craft. Multi-criteria optimization methods in radiation therapy planning: a review of technologies and directions. *Preprint arXiv:1305.1546*, 2013.

[10] D. Craft and T. Bortfeld. How many plans are needed in an IMRT multi-objective plan database? *Phys. Med. Biol.*, 53(11):2785–2796, 2008.

[11] D. Craft, T. Halabi, H. Shih, and T. Bortfeld. An approach for practical multiobjective IMRT treatment planning. *Int. J. Radiat. Oncol.*, 69(5):1600–1607, 2007.

[12] D. Craft, T. Hong, H. Shih, and T. Bortfeld. Improved planning time and plan quality through multi-criteria optimization for intensity modulated radiation therapy. *Int. J. Radiat. Oncol.*, 82(1):e83–e90, 2012.

[13] A. Fredriksson. Automated improvement of radiation therapy treatment plans by optimization under reference dose constraints. *Phys. Med. Biol.*, 57(23):7799–7811, 2012.

[14] M. Hunt, C. Hsiung, S. Spirou, C. Chui, H. Amols, and C. Ling. Evaluation of concave dose distributions created using an inverse planning system. *Int. J. Radiat. Oncol.*, 54:953–962, Nov 2002.

[15] B. Hårdemark, A. Liander, H. R. H, and J. Löf. Direct machine parameter optimization with RayMachine in Pinnacle. RaySearch Laboratories white paper, 2003.

[16] ICRU. *Prescribing, Recording and Reporting Photon Beam Therapy (Supplement to ICRU Report 50)*. ICRU Report 62. International Commission on Radiation Units and Measurements, Bethesda, Maryland, 1999.

[17] ICRU. *Prescribing, Recording, and Reporting Photon-beam Intensity-modulated Radiation Therapy (IMRT)*. ICRU Report 83. International Commission on Radiation Units and Measurements, Bethesda, Maryland, 2010.

[18] R. Jeraj, C. Wu, and T. R. Mackie. Optimizer convergence and local minima errors and their clinical importance. *Phys. Med Biol.*, 48(17):2809–2827, 2003.

[19] S. Kamath, S. Sahni, J. Palta, and S. Ranka. Algorithms for optimal sequencing of dynamic multileaf collimators. *Phys. Med. Biol.*, 49(1):33–54, 2004.

[20] J. Llacer, J. Deasy, T. Bortfeld, T. Solberg, and C. Promberger. Absence of multiple local minima effects in intensity modulated optimization with dose volume constraints. *Phys. Med. Biol.*, 48(2):183–210, 2003.

[21] K. Miettinen. *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, Boston, Massachusetts, 1999.

[22] M. Monz, K.-H. Kufer, T. Bortfeld, and C. Thieke. Pareto navigation—algorithmic foundation of interactive multi-criteria IMRT planning. *Phys. Med. Biol.*, 53(4):985–998, 2008.

[23] H. Nakayama. Aspiration level approach to interactive multi-objective programming and its applications. In P. Pardalos, Y. Siskos, and C. Zopounidis, editors, *Advances in multicriteria analysis*, pages 147–174. Kluwer Academic Publishers, Dordrecht, Netherlands, 1995.

[24] H. Nakayama. Multi-objective optimization and its engineering applications. In J. Branke, K. Deb, K. Miettinen, and R. E. Steuer, editors, *Practical Approaches to Multi-Objective Optimization*, number 04461 in Dagstuhl Seminar Proceedings. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany, 2005.

[25] A. Niemierko. Reporting and analyzing dose distributions: a concept of equivalent uniform dose. *Med. Phys*, 24(1):103–110, 1997.

[26] U. Oelfke and T. Bortfeld. Inverse planning for photon and proton beams. *Med. Dosim.*, 26(2):113–124, 2001.

[27] G. Rennen, E. van Dam, and D. den Hertog. Enhancement of sandwich algorithms for approximating higher-dimensional convex Pareto sets. *INFORMS J. Comput.*, 23(4):493–517, 2011.

[28] E. Romeijn, J. Dempsey, and J. Li. A unifying framework for multi-criteria fluence map optimization models. *Phys. Med. Biol.*, 49(10):1991–2013, 2004.

[29] C. Thieke, T. Bortfeld, A. Niemierko, and S. Nill. From physical dose constraints to equivalent uniform dose constraints in inverse radiotherapy planning. *Med. Phys.*, 30(9):2332–2339, 2003.

[30] J. Wala, D. Craft, J. Paly, A. Zietman, and J. Efstathiou. Maximizing dosimetric benefits of IMRT in the treatment of localized prostate cancer through multicriteria optimization planning. *Med. Dosim.*, 38(3):298–303, 2013.

[31] A. Wierzbicki. A mathematical basis for satisficing decision making. *Math. Modelling*, 3(5):391–405, 1982.

[32] 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.*, 79(4):1241–1247, 2011.

[33] Q. Wu and R. Mohan. Multiple local minima in IMRT optimization based on dose-volume criteria. *Med. Phys.*, 29(7):1514–1527, 2002.