# Minimax optimization for handling range and setup uncertainties in proton therapy

Albin Fredriksson[a)]
*Department of Mathematics, Optimization and Systems Theory, Royal Institute of Technology (KTH), SE-100 44 Stockholm, Sweden and RaySearch Laboratories, Sveavägen 25, SE-111 34 Stockholm, Sweden*

Anders Forsgren
*Department of Mathematics, Optimization and Systems Theory, Royal Institute of Technology (KTH), SE-100 44 Stockholm, Sweden*

Björn Hårdemark
*RaySearch Laboratories, Sveavägen 25, SE-111 34 Stockholm, Sweden*

(Received 30 August 2010; revised 27 January 2011; accepted for publication 27 January 2011; published 1 March 2011)

**Purpose:** Intensity modulated proton therapy (IMPT) is sensitive to errors, mainly due to high stopping power dependency and steep beam dose gradients. Conventional margins are often insufficient to ensure robustness of treatment plans. In this article, a method is developed that takes the uncertainties into account during the plan optimization.
**Methods:** Dose contributions for a number of range and setup errors are calculated and a minimax optimization is performed. The minimax optimization aims at minimizing the penalty of the worst case scenario. Any optimization function from conventional treatment planning can be utilized by the method. By considering only scenarios that are physically realizable, the unnecessary conservativeness of other robust optimization methods is avoided. Minimax optimization is related to stochastic programming by the more general minimax stochastic programming formulation, which enables accounting for uncertainties in the probability distributions of the errors.
**Results:** The minimax optimization method is applied to a lung case, a paraspinal case with titanium implants, and a prostate case. It is compared to conventional methods that use margins, single field uniform dose (SFUD), and material override (MO) to handle the uncertainties. For the lung case, the minimax method and the SFUD with MO method yield robust target coverage. The minimax method yields better sparing of the lung than the other methods. For the paraspinal case, the minimax method yields more robust target coverage and better sparing of the spinal cord than the other methods. For the prostate case, the minimax method and the SFUD method yield robust target coverage and the minimax method yields better sparing of the rectum than the other methods.
**Conclusions:** Minimax optimization provides robust target coverage without sacrificing the sparing of healthy tissues, even in the presence of low density lung tissue and high density titanium implants. Conventional methods using margins, SFUD, and MO do not utilize the full potential of IMPT and deliver unnecessarily high doses to healthy tissues.  [DOI: 10.1118/1.3556559]

Key words: IMPT optimization, minimax optimization, robust planning, uncertainty

## I. INTRODUCTION

By enabling control of the depths at which the dose depositions peak, intensity modulated proton therapy (IMPT) allows for planned dose distributions that conform closely to the target volume while limiting the dose to surrounding healthy tissue. The Bragg peak positions are, however, highly affected by the densities and materials of the volume traversed by the incident protons. In combination with steep beam dose gradients, this makes IMPT susceptible to errors.

Two influential error sources in proton therapy are range and setup uncertainties. Range uncertainty may arise from inaccuracies in the computed tomography (CT) imaging and in the conversion from Hounsfield units to stopping power.[1,2] Setup uncertainty is due to such factors as errors in the positioning of the patient and mechanical inaccuracies in the delivery unit.[3] Failing to account for the uncertainties may result in a delivered dose inferior to the planned one.[2,4]

Commonly, uncertainties are handled by using margins: The clinical target volume (CTV) is expanded into a planning target volume (PTV) and planning is performed to irradiate the latter.[3] An underlying assumption when using margins to account for uncertainties is that the effects of the errors are well approximated by rigid shifts of the dose distribution. This assumption is inadequate for volumes of heterogeneous density, for which errors may drastically distort the dose distribution due to the high density dependency of proton beams. The modulated fluence that often results from IMPT optimization is especially affected. Motivated by this shortcoming, a method is developed to locate where to deposit dose in order to ensure robustness of the plan and thus make the expansion of the CTV to account for range and

setup uncertainties unnecessary. In this method, information about the uncertainties is incorporated in the problem formulation and a minimax optimization is performed. The minimax optimization minimizes the objective function in the worst case scenario and thus provides a bound on how much the plan quality can deteriorate due to the errors. The probability distributions of the uncertainties need not be known, but intervals of possible deviations must be specified.

Stochastic programming and robust optimization in IMPT have been previously used by Unkelbach *et al.*,[5] Chan,[6] and Unkelbach *et al.*[7] They use two formulations: A nonlinear optimization approach in which the expected value of the objective function is minimized and a linear programming approach in which the worst case dose deviation for each voxel considered independently is minimized. Pflugfelder *et al.*[8] add a term using lower and upper bound dose estimates, corresponding to delivering the worst case dose to each voxel independently, to the objective function to account for uncertainties in IMPT.

Building on linear programming formulations of the intensity modulated radiation therapy (IMRT) optimization problem, Chu *et al.*,[9] Chan *et al.*,[10] and Ólafsson and Wright[11] present robust approaches to account for geometric uncertainty. Chan *et al.* use the robust formulation of Bertsimas and Sim[12] to account for errors in the probability distribution of the motion. The other authors use the robust formulation of Ben-Tal and Nemirovski[13] to achieve a high probability of sufficient dose in each voxel considered independently.

Stochastic programming has been used to account for organ motion and setup errors in IMRT optimization by, e.g., Löf *et al.*[14,15] and Unkelbach and Oelfke.[16] More recently, Sobotta *et al.*[17] suggested a method to maximize the probability of finding the penalty function values within given intervals.

The minimax formulation presented in this paper is based on the general nonlinear treatment plan optimization problem. For instance, quadratic penalties to deviations from dose-volume criteria can be used, as described in Sec. II D. Since different voxels may receive their worst doses in different scenarios, taking the worst dose for each voxel independently leads to an overly conservative approximation of the true worst case scenario. In the minimax approach, this conservativeness is avoided by considering only physically realizable scenarios, i.e., evaluating the full objective function in a number of scenarios and considering the worst of these. Correlation between voxels is thereby taken into account. A minimax formulation for treatment plan optimization using linear programming is mentioned by Chan[6] as an alternative to his formulation with independent voxels. In the linear minimax formulation, several constraints are introduced for each voxel in each scenario, which makes the program intractable for clinically relevant cases. Our nonlinear programming minimax approach enables a reduction of the number of auxiliary constraints to the number of scenarios and we have a tractable program. It also allows for the nonlinear optimization functions typically used in treatment planning. We discuss how the minimax optimization method is related to the stochastic programming method and how the minimax stochastic programming formulation generalizes the methods and enables accounting for uncertainties in the probability distributions of the errors.

## II. METHODS

### II.A. Uncertainty models

Range and setup uncertainties are considered, while uncertainties due to organ motion are neglected in the present paper. The uncertainties are assumed to be mutually independent.

#### *II.A.1. Range uncertainty*

We model range errors by scaling the mass density of the treatment volume uniformly. Such scaling may result from measurement errors or errors in the conversion from Hounsfield units to stopping power. In effect, the range shifts of all spots are correlated in the considered uncertainty model. The model could be refined by taking into account that the uncertainty in the stopping powers is a function of the Hounsfield unit values.

Assuming a uniform density change in the volume traversed by a given spot, the radiological length of its shift will be proportional to its radiological depth. Its depth dose curve will consequently be stretched or contracted. The geometrical distance that the spot is shifted will depend on the material of the traversed volume. A small scaling of the density might move the spot several centimeters: If it is planned to hit the boundary of a lung tumor, a small scaling of the density might move it to the lung wall.

#### *II.A.2. Setup uncertainty*

Setup errors are modeled as shifts of the beam isocenters. The isocenters of all beams are assumed to be shifted equally. Such errors correspond to misalignments of the patient relative to the gantry.

In general, a shift of the beam isocenters leads to a nonrigid shift of the dose distribution. If a spot is shifted laterally so that it traverses a volume of different density than planned, it will also be shifted longitudinally and its dose distribution will be deformed. Small setup errors might thus lead to large displacements of spots that travel close to and in parallel with steep density gradients, such as along bone edges.

#### *II.A.3. Assessing the effects of the errors*

To estimate the effects of the errors, the set of possible scenarios is discretized and dose distributions are calculated for each of the scenarios. In the plan evaluation, the plans resulting from the optimization are evaluated by moving the beam isocenters, scaling the mass density, and recomputing the dose. So as not to bias the results, the evaluation encompasses scenarios other than those used in the optimization: Finer sampling grids are used both for the density errors and for the setup errors.

TABLE I. Sizes of the patient cases.

| Patient case | No. of contributing voxels | Spot spacing (mm) | No. of spots |
|---|---|---|---|
| Lung | 152 060 | 5 | 4 733 |
| Paraspinal | 183 913 | 5 | 17 721 |
| Prostate | 325 836 | 6 | 8 846 |

During the minimax optimization, the effects of density errors are computed by calculating dose for a number of density scalings. The effects of setup errors are approximated by moving the spot weights: The weights for all spots of a given beam are represented in a three-dimensional spot weight grid where the $x$- and $y$-coordinates specify the lateral position in the beam eye's view and the $z$-coordinate specifies the energy. All weights in a given layer along the $z$-axis consequently correspond to spots of the same energy, which might be located at different physical depths in the patient. Setup errors are approximated by moving the spot weights to adjacent positions within the energy layers in the spot weight grid and calculating the dose for the new weights using the same dose contribution matrix as in the nominal scenario. A shift of the beam isocenter at a distance equal to the spot spacing along the $x$-axis of the grid can thus be approximated by moving all spot weights one step along the $x$-axis in the grid. If the isocenter is shifted in the beam direction, the dose distribution is unchanged in this approximation. Since the same dose contribution matrix can be used for several shifts, only one matrix for each density error is required during the optimization.

When accounting for setup errors smaller than the spot line spacing, spots are precalculated on a finer grid than the one intended for delivery (see Unkelbach *et al.*[7]). If there are no spot positions corresponding to precisely the same isocenter shift for all beams, the positions are approximated by the nearest neighbors. A hexagonal scanning pattern is used, implying that each spot has six neighbors.

### II.B. Nominal optimization formulation

Let $x \in \mathbb{R}^n$ denote the vector of spot weights and let $P \in \mathbb{R}^{m\times n}$ denote the matrix mapping spot weights to dose, so that each row of $P$ corresponds to a voxel and contains elements that specify the dose contribution from each spot to the voxel under unit weight. Let $d(x) \in \mathbb{R}^m$ denote the dose distribution vector as a linear function of the spot weights, given by $d(x)=Px$. This will be used implicitly in the problem formulations below. The nominal treatment plan optimization problem is given by

$$\underset{x}{\text{minimize}} \quad f(d(x)) \tag{1a}$$

$$\text{subject to} \quad x \geq 0, \tag{1b}$$

where, typically, $f$ is a sum of optimization functions weighted by importance factors. This bound constrained program is convex whenever the objective function $f$ is (since $d$ is linear in $x$) and can be solved by gradient based methods. Using dose-volume optimization functions leads to a nonconvex problem, but in practice, there seems to be little difficulty with local minima.[18]

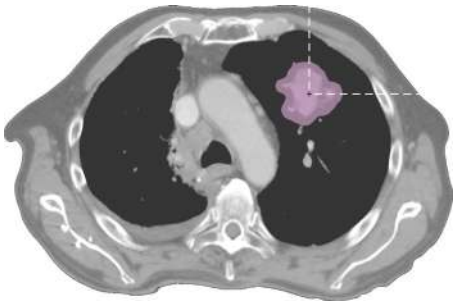


FIG. 1. A transversal slice of the lung case. The shaded structure is the CTV. The beam directions are indicated by dashed lines.

### II.C. Robust methods

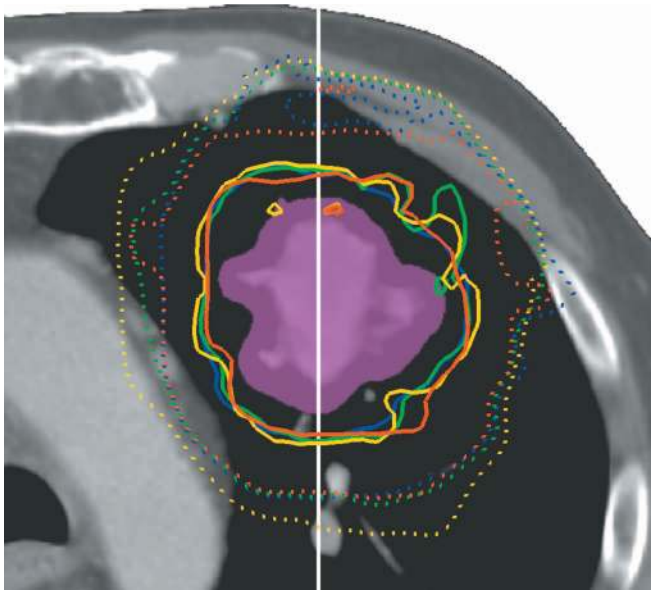
#### II.C.1. Conventional methods

In order to obtain robustness in conventional treatment planning, the CTV is expanded into a PTV and the optimizing is performed toward irradiation of the whole PTV. This can be combined with the single field uniform dose (SFUD) technique, in which the dose for each beam is enforced to be uniform. For tumors in low density regions such as the lung, a conventional PTV amounts to just a slight margin in radiological depth. By overriding the material of the low density volume surrounding the tumor and thereby planning as if this volume were made of normal tissue, a more effective PTV can be constructed. We refer to this as "material override" (MO).

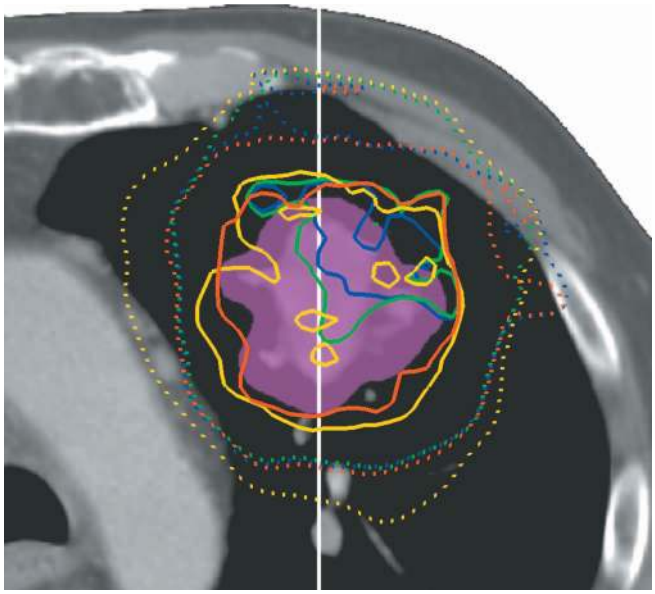
#### II.C.2. Minimax optimization formulation

We propose using minimax optimization for handling uncertainties in proton therapy. The optimization then aims at minimizing the penalty of the worst scenario. By considering only physically realizable scenarios, correlations between voxels are preserved and the extra conservativeness that results from independent handling of voxels is avoided. Since the worst case scenario is considered, there is no need to specify probabilities for the different scenarios. Defining intervals of the deviations for which robustness is demanded is similar to the desired effects of margins. However, since the effects of errors are accounted for in the minimax optimization, there is no need to add explicit margins to compensate for setup and density errors: The optimization settles where to deposit dose in order to compensate for the possible errors.

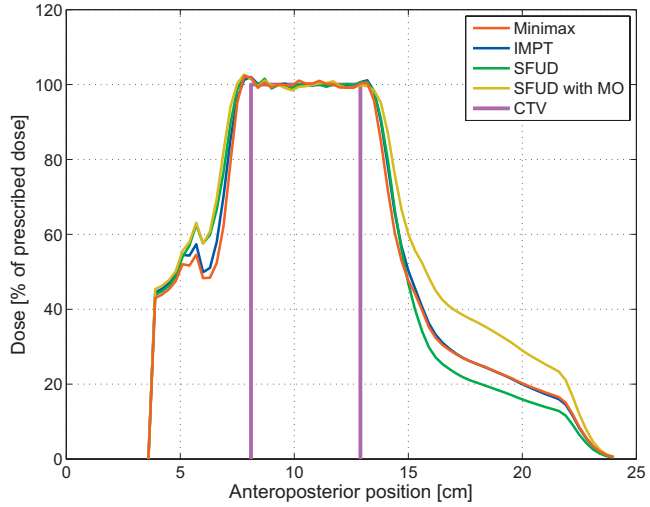
Since the minimax method optimizes the worst case scenario, including too improbable scenarios can compromise the plan quality. Therefore, the interval of the uncertain variable variations to include must be selected with caution, preferably to cohere with the errors that margins are used to take into account. Thus, the minimax method does not minimize the worst of all possible scenarios, but the worst sce-



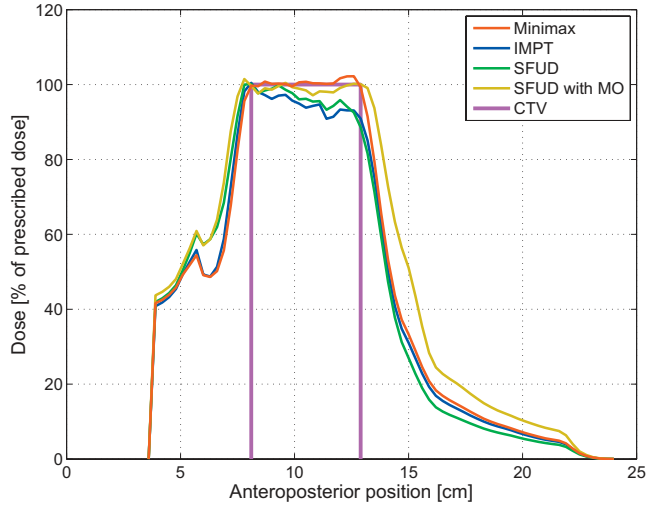
(a) Nominal scenario isodose curves



(b) Perturbed scenario isodose curves




(c) Nominal scenario line doses




(d) Perturbed scenario line doses

FIG. 2. [(a) and (c)] Isodose curves and line doses in the nominal and [(b) and (d)] a perturbed scenario in a transversal slice of the lung case. In the perturbed scenario, the density is 3% higher than planned and the isocenters are shifted 5 mm inferiorly. [(a) and (b)] Isodose curves for 98% (solid) and 55% (dotted) of the prescribed target dose. The curves correspond to minimax (red), IMPT (blue), SFUD (green), and SFUD with MO (yellow). The shaded structure is the CTV (magenta). The white line displays the trace of the line doses. [(c) and (d)] Line doses. The height of the CTV is the prescribed dose level.

nario within some interval. The set $\mathcal{S}$ indexing the scenarios included in the optimization is constructed accordingly. Different scenarios lead to different dose contributions, so let $P(s)$ denote the matrix mapping spot weights to dose in scenario $s \in \mathcal{S}$ and let $d(x,s)$ be the corresponding dose distribution given by $d(x,s)=P(s)x$. The minimax problem is formulated as

$$\underset{x}{\text{minimize}} \quad \max_{s\in\mathcal{S}}\{f(d(x,s))\}$$

$$\text{subject to} \quad x \geq 0.$$

The max-function preserves convexity, so this is a convex problem whenever the function $f$ is. Since the maximization is taken over a finite number of scenarios, by introducing the auxiliary variable $t \in \mathbb{R}$, the problem can be equivalently formulated as

$$\underset{x,t}{\text{minimize}} \quad t \tag{2a}$$

$$\text{subject to} \quad t \geq f(d(x,s)) \quad \forall\, s \in \mathcal{S}, \tag{2b}$$

$$x \geq 0, \tag{2c}$$

which is a nonlinearly constrained program, but still convex whenever $f$ is. The constraints increase the problem size as compared to the nominal formulation (B) and require an optimization solver that can handle nonlinear constraints. As in the case of the nominal formulation, dose-volume optimiza-

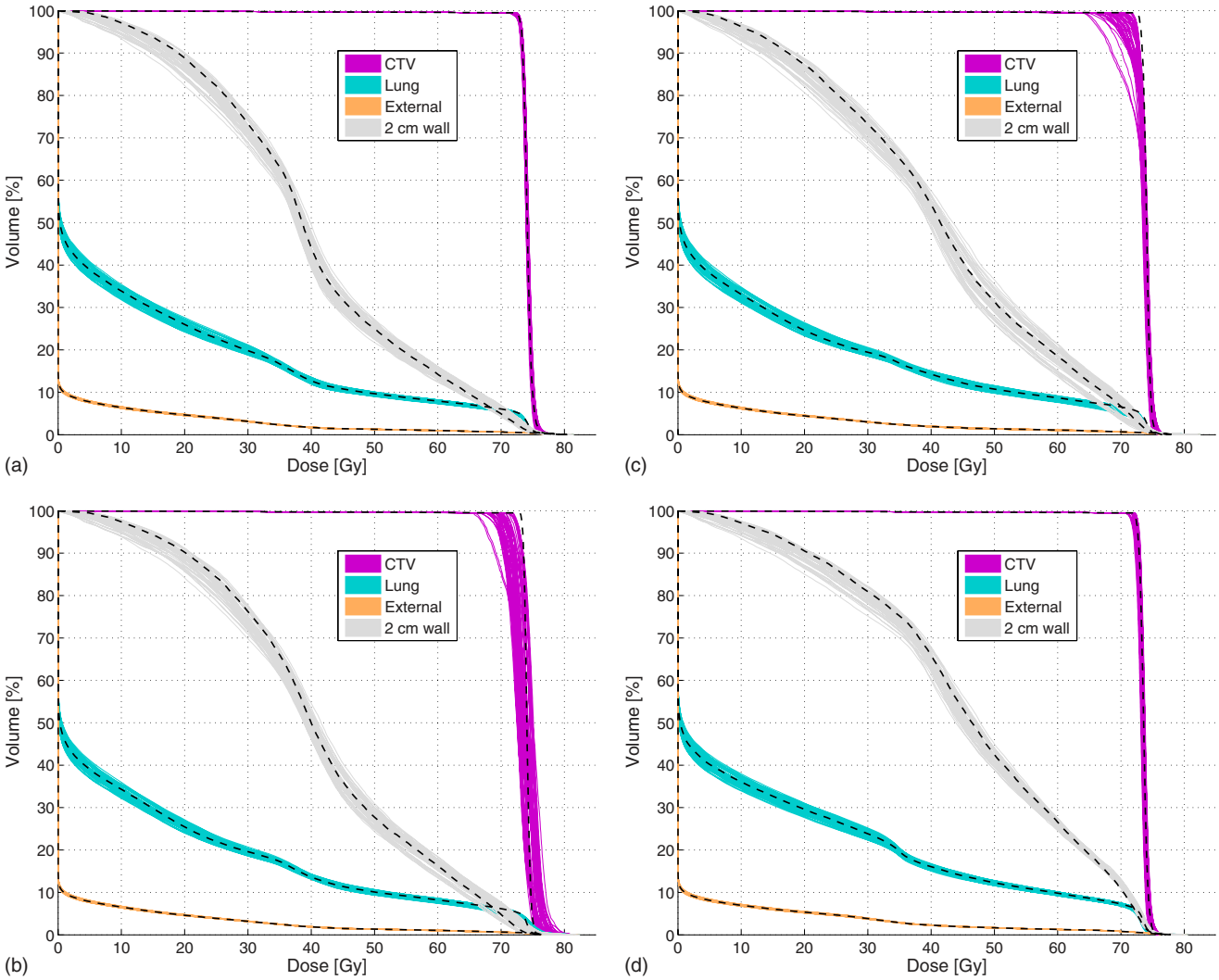



FIG. 3. DVHs for the lung case in the nominal scenario (dashed) and in 45 scenarios sampled from the surface of the sphere of radius 5 mm: Five for each of nine equispaced shifts in [0.97, 1.03] of the planned density. The 2 cm wall ROI surrounds the PTV and gives an indicative measure of target conformance.

tion functions are applicable in the minimax formulation although they make the problem nonconvex. Compared to a stochastic programming formulation, the minimax optimization is more computationally demanding due to the nonlinear constraints. Since both methods require dose distributions in the same scenarios, their memory requirements are similar. The memory requirement for the nonlinear constraints, one entry for each spot in each scenario, is small in comparison. The memory requirements for the dose contribution matrices are, however, similar for the methods, since both require calculating dose distributions for the same scenarios. The gradients of the nonlinear constraints of the minimax method increase the memory requirement slightly, but this increase is small compared to the requirement for the dose contribution matrices.

Note that the max-function in the minimax formulation can be applied to a subset of all ROIs, while the other ROIs are treated nominally. Moreover, one max-function can be introduced for each ROI as a heuristic method to introduce more independency into the model. This increases the number of constraints by a factor of the number of independent volumes, but is often computationally cheaper than introducing more independency in the uncertainty model.

The minimax formulation relates closely to the stochastic programming formulation. In fact, the two formulations are special cases of the more general minimax stochastic programming formulation, in which the probability distributions of the uncertain factors are themselves subject to uncertainty. The optimization then aims at minimizing the expected value of the worst case realization of the probability distributions, i.e., solving the problem

$$\underset{x}{\text{minimize}} \quad \max_{p\in\mathcal{P}}\left\{\sum_{s\in\mathcal{S}} p_s f(d(x,s))\right\}$$

$$\text{subject to} \quad x \geq 0,$$

where $p_s$ is the probability of scenario $s \in \mathcal{S}$ occurring and $\mathcal{P}$ is the set of probability distributions for which $a_s \leq p_s \leq b_s$ for the given parameters $0 \leq a_s \leq b_s \leq 1$ for $s \in \mathcal{S}$. When $a_s = 0$ and $b_s = 1$ for all $s \in \mathcal{S}$, this problem is equivalent to the

TABLE II. Dose statistics for the lung case. The doses are in Gy. Here, $d_x$ denotes minimum dose to $x$ % of the volume, $\bar{d}$ denotes the nominal scenario mean dose level in the volume, and a circumflex denotes dose in the worst scenario (i.e., minimum for $\hat{d}_{98}$ and maximum otherwise). $\text{External}_{0.5}$ is composed the voxels of the external ROI that receive 0.5 Gy or more and $v$ is its volume normalized to that of the $\text{External}_{0.5}$ ROI in the minimax plan. Values that differ from those of the minimax plan by 2 Gy or more are in bold.

| | CTV | | | Lung | | | $\text{External}_{0.5}$ |
|---|---|---|---|---|---|---|---|
| | $\hat{d}_{98}$ | $\hat{d}_2$ | $\bar{d}$ | $d_{10}$ | $\hat{d}_{10}$ | $\bar{d}$ | $\bar{d}\cdot v$ |
| Minimax | 72.5 | 76.2 | 74.2 | 48.3 | 50.5 | 13.8 | 21.7 |
| IMPT | **66.7** | **78.8** | 74.1 | **50.5** | **52.9** | 13.9 | 22.1 |
| SFUD | **65.3** | 75.9 | 74.0 | **53.3** | **56.3** | 13.8 | 21.6 |
| SFUD with MO | 71.5 | 75.4 | 75.4 | **59.4** | **60.6** | 15.6 | **24.8** |

minimax problem and when $a_s=b_s$ for all $s\in\mathcal{S}$, it is equivalent to the stochastic programming problem. Adjusting these parameters enables continuous scaling between minimizing the objective function in the worst case scenario and minimizing its expected value. As the problem is formulated above, it is difficult to solve, but linear programming duality can be used to avoid the difficulty. This yields an equivalent minimax stochastic programming formulation of the same computational size as the minimax formulation (2). For completeness, the derivation of the minimax stochastic programming formulation is included in Appendix A.

The optimization algorithms used in radiation therapy treatment planning typically strive toward satisfying the Karush–Kuhn–Tucker (KKT) conditions that are necessary for optimality, provided some regularity conditions are satisfied. Considering the KKT conditions for the stochastic programming problem and the minimax stochastic programming problem, one can verify that the Lagrange multipliers of the minimax stochastic programming problem (or, as a special case, of the minimax optimization problem) correspond precisely to a probability distribution in $\mathcal{P}$. This means that the solution to the minimax stochastic programming problem is equivalent to the solution to a stochastic programming problem with a specific probability distribution (see Shapiro and Ahmed[21]). This worst case probability distribution is unknown prior to solving the minimax problem, but after small parameter changes, it could be used for approximative replanning using a computationally cheaper stochastic programming formulation.

### II.D. Computational study

We apply the minimax method to three patient cases: A lung case, a paraspinal case, and a prostate case. The voxel sizes are set to $3\times3\times3$ mm$^3$ for all cases. The total number of nonzero elements in the dose contribution matrices of all scenarios was $10^8$–$10^9$ for the cases studied. Sizes of the patient cases are summarized in Table I. These cases present different obstacles for robustness: The lung case tumor is surrounded by a low density volume, the paraspinal case CTV surrounds the spinal cord and holds metal implants that cause density heterogeneities, while the prostate case is more homely. Resulting plans are compared to plans reached by conventional methods, in which margins, SFUD, and MO are used to account for uncertainties. In the presentation of the results, the conventional method using only margins to account for uncertainties is referred to as "IMPT," the method in which uniform beam doses are enforced is referred to as "SFUD," and the method using both SFUD and MO is referred to as "SFUD with MO." Margins are used in all conventional approaches.

Dose kernels are calculated using the pencil beam dose algorithm of the treatment planning system RayStation version 1.2 (RaySearch Laboratories, Stockholm, Sweden), which takes heterogeneities into account also within the cross sections of the spots. The optimization problems are solved in MATLAB version 7.9 (The MathWorks, Natick, MA) using SNOPT version 7.2,[22] which uses a sparse sequential quadratic programming method with limited-memory Broyden–Fletcher–Goldfarb–Shanno quasi-Newton updates of the approximation to the Hessian of the Lagrangian. The computations are performed on a single Intel Xeon 3 GHz processor core with multithreading disabled and with 32 GB of memory under 64-bit Linux. Resulting plans are imported into RayStation, where evaluation is performed by calculating dose distributions for perturbations of the density and the beam isocenters.

In all cases, the density uncertainty is assumed to be up to ±3%, which is similar to the uncertainty used by Lomax[2] to include the effects of conversion errors and CT data acquisition errors and also to the uncertainty Moyers *et al.*[23] use when calculating compensators for passive scattering. In the minimax optimization, three density scalings are included (nominal density and ±3%), whereas nine scalings are in-

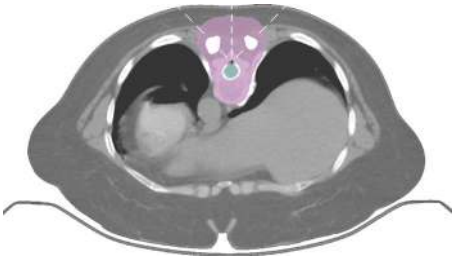


FIG. 4. A transversal slice of the paraspinal case. The shades structures are the CTV, the spinal cord and the titanium implants. The beam directions are indicated by dashed lines.

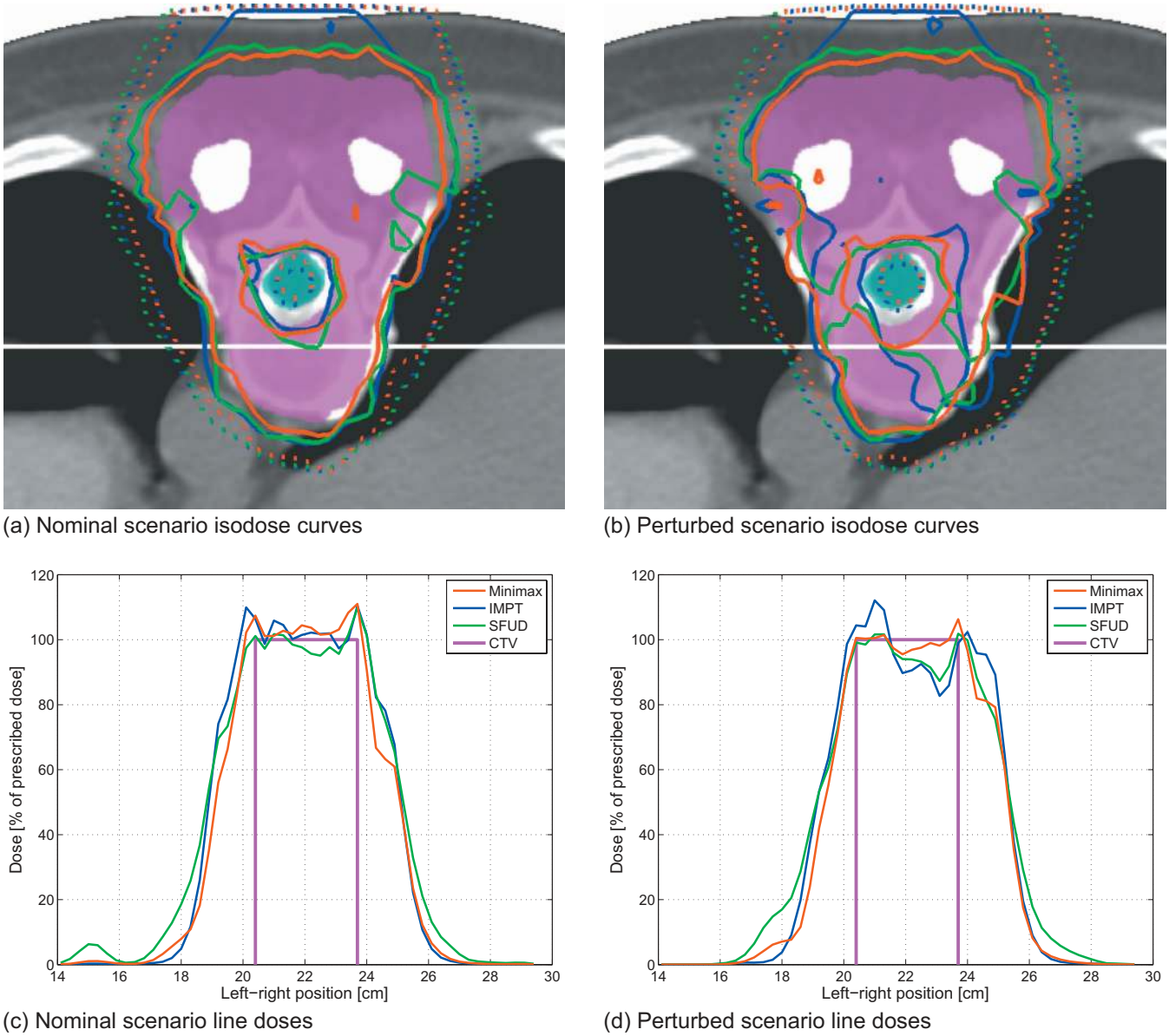



FIG. 5. [(a) and (c)] Isodose curves and line doses in the nominal and [(b) and (d)] a perturbed scenario in a transversal slice of the paraspinal case. In the perturbed scenario, the density is 3% higher than planned and the isocenters are shifted 2.5 mm diagonally in the directions anterior, inferior, and right. [(a) and (b)] Isodose curves for 95% (solid) and 55% (dotted) of the prescribed target dose. The curves correspond to minimax (red), IMPT (blue), and SFUD (green). The shaded structures are the CTV (magenta), the spinal cord (cyan), and the titanium implants (white). The 55% isodose curve of SFUD cannot be seen along the spinal cord since the doses are higher than 55% of the prescribed target dose. The white line displays the trace of the line doses. [(c) and (d)] Line doses. The height of the CTV is the prescribed dose level.

cluded in the plan evaluation. The setup uncertainty is assumed to be isotropic. Due to the hexagonal scanning pattern, each spot has six neighbors, implying that each beam has seven possible setup shift positions (including staying stationary) that can be used in the optimization.

The combinations of beam shifts used to approximate a setup shift in the optimization are chosen to cohere with physically realizable shifts. The number of coherent combinations depends on the beam directions: There are 17 combinations for the lung case, 21 combinations for the paraspinal case, and 7 combinations for the prostate case (see Appendix B for details). Each of these shifts is evaluated for all density shifts. In the plan evaluation, isocenter shifts are sampled from the sphere with a radius corresponding to the size of the setup uncertainty.

The optimization functions used are the minimum and maximum dose-volume histogram (DVH) functions commonly used in radiation therapy. A maximum DVH optimization function is applied to a ROI and has two parameters: A dose level $\hat{d}$ and a volume fraction $\hat{v}$. It can be interpreted as "at most a fraction $\hat{v}$ of the ROI volume should receive a dose exceeding $\hat{d}$ Gy." Let $\Delta v_i$ be the relative volume of voxel $i$. For a given voxel $j$, the penalty of the maximum DVH function is

$$g_j(d,\hat{v})\Delta v_j \max\{d_j-\hat{d},0\}^2,$$

where

$$g_j(d,\hat{v})=\begin{cases}0 & \text{if} \quad \sum_{i:d_i\geq d_j}\Delta v_i \leq \hat{v}\\ 1 & \text{if} \quad \sum_{i:d_i>d_j}\Delta v_i \geq \hat{v}\\ \left(\hat{v}-\sum_{i:d_i>d_j}\Delta v_i\right)/\left(\sum_{i:d_i=d_j}\Delta v_i\right) & \text{otherwise.}\end{cases}$$

The third option in the definition of the function $g_j$ ensures that the dose to precisely the volume $\hat{v}$ of voxels that receive the highest dose is allowed go without penalty by partly penalizing the voxel $j$ if $\Sigma_{i:d_i>d_j}\Delta v_i<\hat{v}<\Sigma_{i:d_i\geq d_j}\Delta v_i$. If $\hat{v}=0$, then $g_j(d,\hat{v})=1$ and the maximum DVH function becomes a maximum dose function. The definitions of minimum DVH and minimum dose functions are analogous to those of maximum DVH and maximum dose functions. By combining a minimum and a maximum dose function, a uniform dose function is obtained.

## III. RESULTS

A supplementary material is provided for transversal slices of the beam dose distributions of all cases.[26]

### III.A. Lung case

For the lung case, two perpendicular beams (0° and 90°) were used. The prescribed dose to the target was 74 Gy. Setup errors of up to 5 mm were accounted for, which corresponds to the margins of 5–10 mm recommended for lung cases in WPE protocol.[24] The margin used in the conventional approaches was constructed as a 5 mm expansion of the CTV, since the radiological depth of the tumor was so small that this encompassed also the effects of the density errors. A transversal slice of the lung case is shown in Fig. 1.

The minimax optimization took 1 h to run. Isodose curves for the nominal scenario dose distributions in a transversal slice, obtained using the minimax method, IMPT, SFUD, and SFUD with MO, are shown in Fig. 2(a). Line doses from the same slice are shown in Fig. 2(c). The figures show that the dose distribution of SFUD with MO is notably more spread out than those of the other methods. Isodose curves for the dose distributions in a perturbed scenario are displayed in Fig. 2(b) and the corresponding line doses are shown in Fig. 2(d). The perturbed 98% isodose curves of IMPT and SFUD are distorted and enclose less than half the CTV, while the curve of SFUD with MO is slightly distorted but encloses most of the CTV. The curve of the minimax method encloses almost the whole CTV. The line dose plots show how the perturbation results in underdosage for the IMPT and SFUD plans. DVHs are displayed in Fig. 3. The DVHs confirm that the plans of IMPT and SFUD are less robust than the plans of the minimax method and SFUD with MO, which display equivalent robustness with respect to target coverage. The figure also shows that the SFUD with MO plan delivers higher dose to healthy tissues than the plans of the other methods. Dose statistics are presented in Table II. Similar to the DVHs, they reflect that the minimax method provides robust target coverage and better sparing of healthy tissues than the other methods.

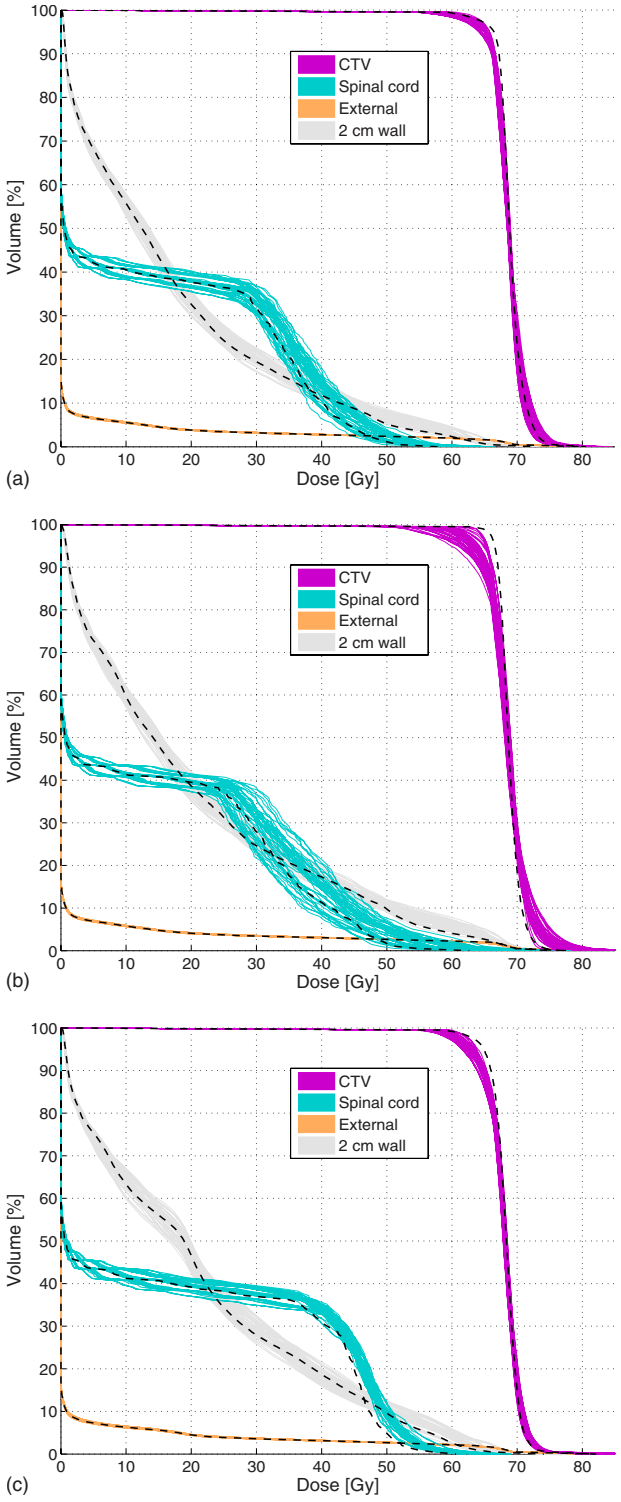



FIG. 6. DVHs for the paraspinal case in the nominal scenario (dashed) and in 45 scenarios sampled from the surface of the sphere of radius 2.5 mm: Five for each of nine equispaced shifts in [0.97, 1.03] of the planned density. The 2 cm wall ROI surrounds the PTV and gives an indicative measure of target conformance.

TABLE III. Dose statistics for the paraspinal case. The doses are in Gy. Here, $d_x$ denotes minimum dose to $x$ % of the volume, $\bar{d}$ denotes the nominal scenario mean dose level in the volume, and a circumflex denotes dose in the worst scenario (i.e., minimum for $\hat{d}_{98}$ and maximum otherwise). $\text{External}_{0.5}$ is composed the voxels of the external ROI that receive 0.5 Gy or more and $v$ is its volume normalized to that of the $\text{External}_{0.5}$ ROI in the minimax plan. Values that differ from those of the minimax plan by 2 Gy or more are in bold.

| | CTV | | | Spinal cord | | | $\text{External}_{0.5}$ |
|---|---|---|---|---|---|---|---|
| | $\hat{d}_{98}$ | $\hat{d}_2$ | $\bar{d}$ | $d_{10}$ | $\hat{d}_{10}$ | $\bar{d}$ | $\bar{d}\cdot v$ |
| Minimax | 60.7 | 75.1 | 68.9 | 40.3 | 44.6 | 14.3 | 23.4 |
| IMPT | **55.4** | **77.9** | 68.8 | 41.3 | **47.7** | 14.6 | **25.4** |
| SFUD | **58.7** | 73.4 | 68.2 | **47.3** | **50.9** | **17.8** | **26.6** |

### III.B. Paraspinal case

To study the effects of different approaches when the target volume surrounds an OAR and when density heterogeneities are present, a paraspinal case was constructed. A target surrounding the spinal cord was drawn on a patient case. Titanium implants were added with material override. The paraspinal case was designed to be similar to that studied by Unkelbach *et al.*[7] Three beams were used, located at 0°, 45°, and 315°. The prescribed dose to the target was 68.4 Gy. Setup errors of up to 2.5 mm were accounted for, which is the setup error used in the plan evaluation by Unkelbach *et al.* Since the spot spacing was 5 mm, additional spots were precalculated on a grid of 2.5 mm to enable the approximation of the effects of setup errors in the minimax optimization. The margin in the conventional approaches was 2.5 mm in the inferior, posterior, and superior directions and 5 mm in the anterior, left, and right directions to account for setup errors in combination with density errors. A transversal slice of the paraspinal case is shown in Fig. 4.

The minimax optimization took 8 h to run. Isodose curves for the nominal scenario dose distributions in a transversal slice, obtained using the minimax method, IMPT, and SFUD, are shown in Fig. 5(a). Line doses from the same slice are shown in Fig. 5(c). The 95% isodose curve of SFUD does not enclose the parts of the CTV that are behind the titanium implants. That of IMPT does not conform to the CTV posteriorly. The curves of the minimax method conform closer to the CTV than those of the other methods. Isodose curves for the dose distributions in a perturbed scenario are displayed in Fig. 5(b) and the line doses of the same scenario are displayed in Fig. 5(d). The perturbed 95% isodose curves of IMPT and SFUD leave gaps over the CTV behind the spinal cord. This is also notable in the line dose plots. The perturbed 95% isodose curve of the minimax method leaves a larger hole around the spinal cord than in the nominal scenario, but still encloses most of the CTV, also behind the spinal cord. DVHs are displayed in Fig. 6. It can be seen that the plans of IMPT and SFUD are less robust than the plan of the minimax method. For sparing of the spinal cord to be possible, the dose to the CTV in the SFUD plan declines earlier than in the minimax plan. Moreover, the DVHs show that the spinal cord generally receives higher maximum doses in the IMPT and SFUD plans than in the minimax plan. Dose statistics are presented in Table III. They confirm that the minimax method provides more robust target coverage and less dose to healthy tissues than the other methods. The SFUD plan yields higher dose to the spinal cord than the other methods, but is more robust than the IMPT plan.

### III.C. Prostate case

Two opposed fields (90° and 270°) were used for the prostate case. This is similar to the beam configuration used by Meyer *et al.*[25] The prescribed dose to the target was 70 Gy. Setup errors were assumed to be up to 6 mm, which is similar to the errors in the anteroposterior direction used by Meyer *et al.* The conventional methods therefore had a 6 mm margin surrounding the CTV. Effects of setup errors in the left-right direction are minor since it is parallel to both beams, but we still used a 6 mm margin to account for density errors in these directions, calculated as 3% of 20 cm. A transversal slice of the prostate case is shown in Fig. 7.

The minimax optimization took 4 h to run. Isodose curves for the nominal scenario dose distributions in a transversal slice, obtained using the minimax method, IMPT, and SFUD are shown in Fig. 8(a). Line doses from the same slice are shown in Fig. 8(c). The 98% isodose curves are similar in all methods, but that of the minimax method is slightly narrower than the others. The 55% isodose curve of the minimax method extends farther in the left and right directions but less in the anterior and posterior directions. Isodose curves for the dose distributions in a perturbed scenario are displayed in Fig. 8(b) and the corresponding line doses are shown in Fig. 8(d). Since the treatment volume has rather

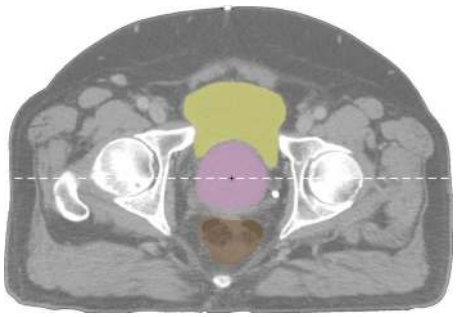
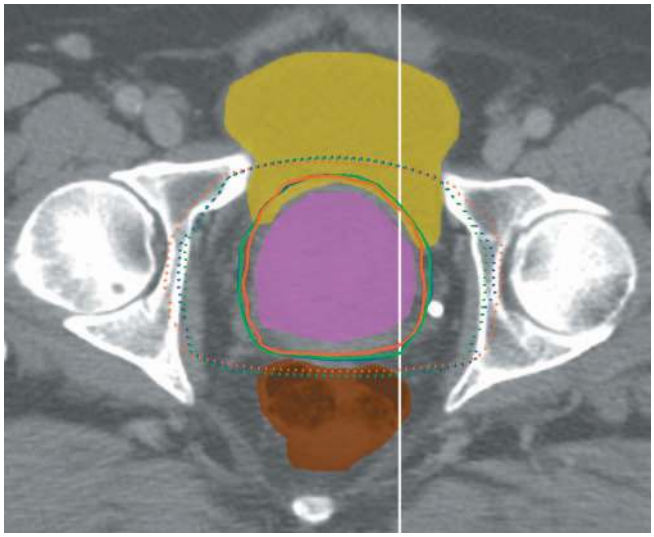
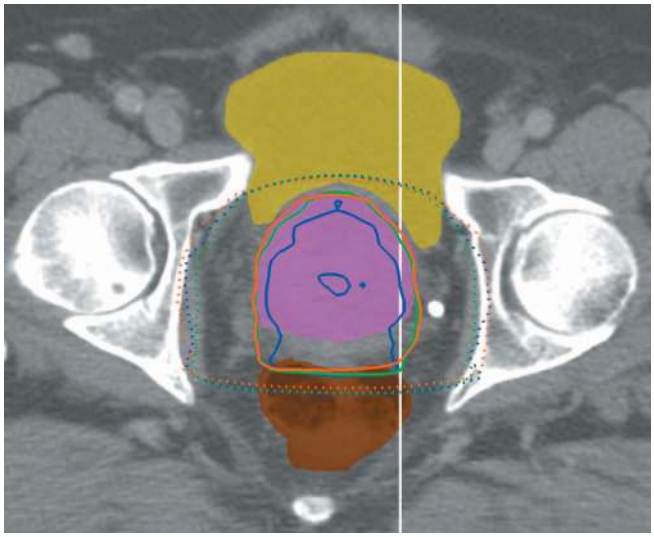


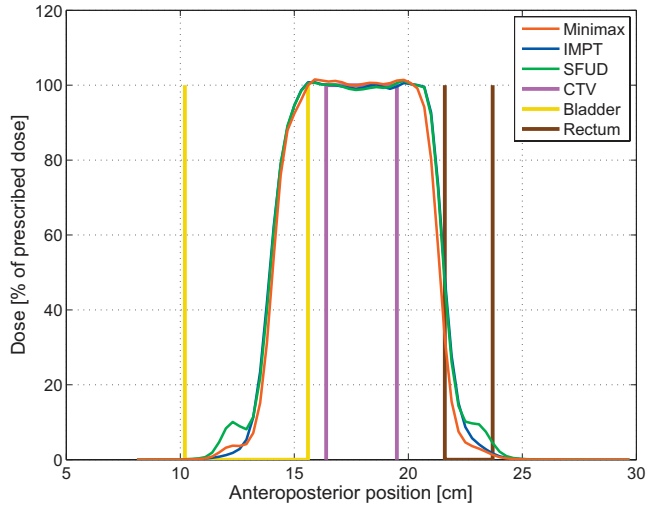
FIG. 7. A transversal slice of the prostate case. The shaded structures are the bladder, the CTV, and the rectum. The beam directions are indicated by dashed lines.



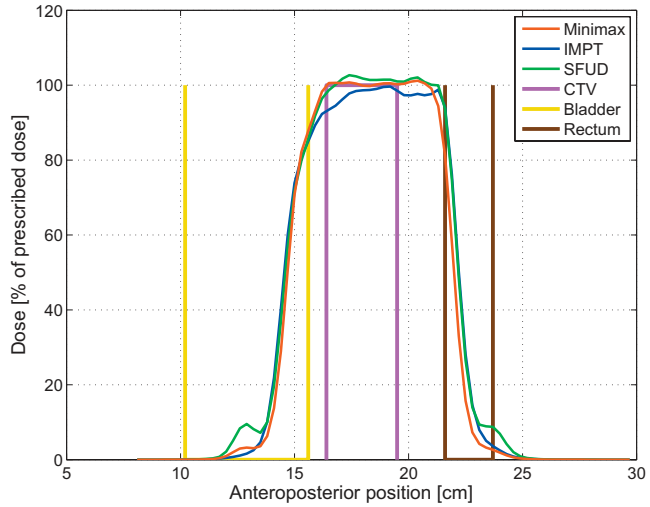
(a) Nominal scenario isodose curves



(b) Perturbed scenario isodose curves




(c) Nominal scenario line doses




(d) Perturbed scenario line doses

FIG. 8. [(a) and (c)] Isodose curves and line doses in the nominal and [(b) and (d)] a perturbed scenario in a transversal slice of the prostate case. In the perturbed scenario, the density is 3% higher than planned and the isocenters are shifted 6 mm anteriorly. [(a) and (b)] Isodose curves for 95% (solid) and 55% (dotted) of the prescribed target dose. The curves correspond to minimax (red), IMPT (blue), and SFUD (green). The shaded structures are the bladder (yellow), the CTV (magenta), and the rectum (brown). The white line displays the trace of the line doses. [(c) and (d)] Line doses. The height of the CTV is the prescribed target dose level.

homogeneous density, the dose distribution of the SFUD plan is robust to the perturbation. The perturbed 98% isodose curves of SFUD and minimax both enclose almost the entire CTV. The dose distribution of the IMPT plan is distorted and its curve no longer encloses the whole CTV. Also, its line dose shows that underdosage results. DVHs are displayed in Fig. 9. The DVHs show that the IMPT plan is less robust than the plans of the minimax method and SFUD, which display equivalent robustness with respect to target coverage. It is also seen that IMPT and SFUD delivers higher dose to the prostate than the minimax method. Dose statistics are presented in Table IV. They show that the robustness of the minimax method and SFUD are equivalent, but that the minimax method plan leads to better sparing of the OARs.

## IV. DISCUSSION

For all examined cases, the minimax method provided more robustness than IMPT with a margin to account for uncertainties, both with respect to target coverage and to the sparing of healthy tissues. For the lung case and the prostate case, methods using margins in combination with SFUD with MO and SFUD, respectively, provided similar target coverage robustness as the minimax method, but did so at the cost of higher doses to healthy tissues. For the paraspinal case, which was the most challenging case studied due to its geometry and density heterogeneities, methods using margins were insufficient to achieve the same target coverage robustness and sparing of OARs as the minimax method.

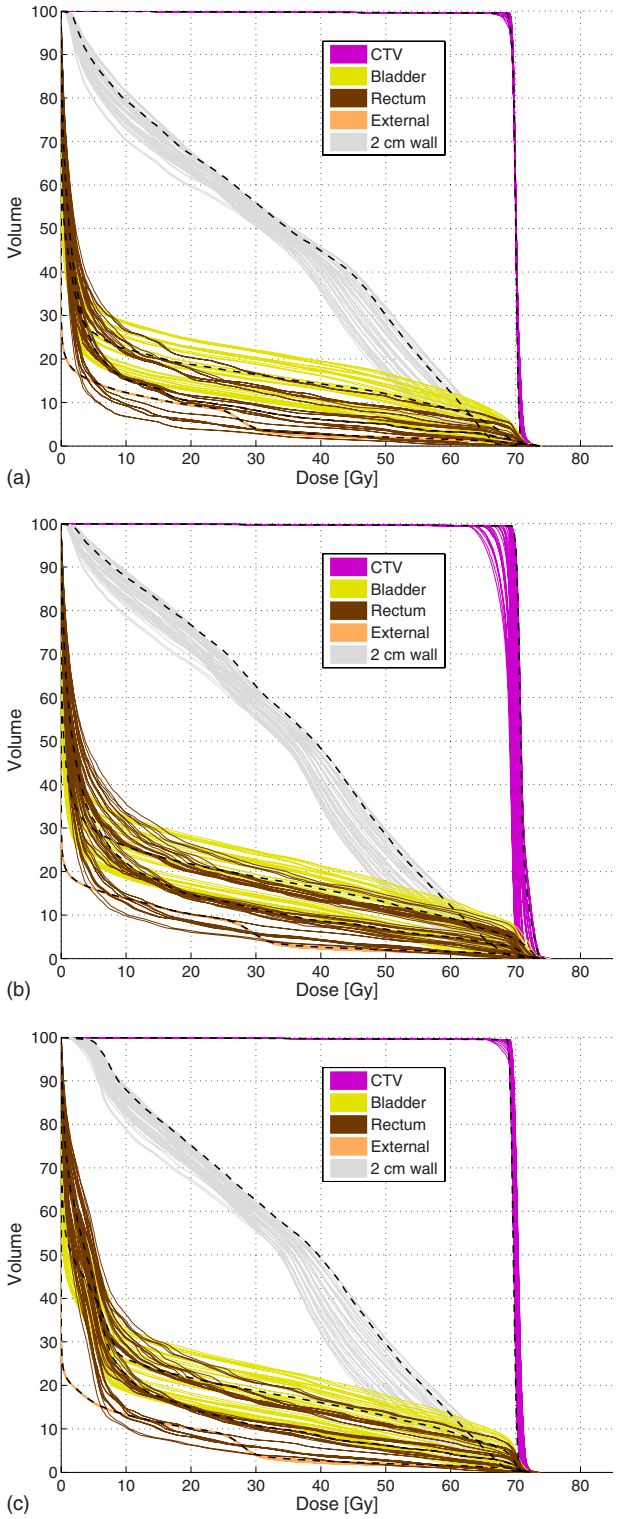



FIG. 9. DVHs for the prostate case in the nominal scenario (dashed) and in 45 scenarios sampled from the surface of the sphere of radius 6 mm: Five for each of nine equispaced shifts in [0.97, 1.03] of the planned density. The 2 cm wall ROI surrounds the PTV and gives an indicative measure of target conformance.

The use of margins may lead to an unnecessary increase in integral dose and in the dose to OARs close to the target. In many cases, margins alone do not render the robustness that they are designed for, but must be supplemented by other techniques, such as SFUD and MO. Since these techniques impose dispensable restrictions on the optimization, they compromise the plan quality unnecessarily. Incorporating more information in the problem formulation allows the optimizer to determine where to deposit dose in order to achieve robust plans. The dose to healthy tissues can thereby be reduced and the necessary dose depositions can be localized in a way that avoids OARs better than heuristic methods.

The minimax formulation is general in that it is neither restricted to certain types of uncertainties nor to certain optimization functions used in the treatment planning. In this paper, optimization functions penalizing deviations from dose-volume criteria were used, but biological optimization functions can also be managed using this method. Since the method accounts only for scenarios that are physically realizable, the correlation between voxels is preserved and unnecessary conservativeness is avoided.

Since the minimax method requires one nonlinear constraint per scenario, its optimization is more computationally demanding than that of the stochastic programming method. To get a time comparison between the minimax method and stochastic programming, we solved stochastic programming problems with the scenario probabilities taken as the Lagrange multipliers of the minimax method solutions, implying that the optimizations of both methods aimed for the same solution. We observed that the minimax optimization required a factor of 1–3 of the stochastic programming optimization times for the patient cases studied in this paper.

The paraspinal case studied in this paper is similar to the cases studied by Pflugfelder *et al.*[8] and Unkelbach *et al.*[7] As their methods, the minimax method avoids steep lateral dose gradients alongside the target and steep distal dose gradients in front of the target. This is to be expected from a robust method. The different approaches result in different plans, but since there are yet no generally accepted evaluation criteria for robustness, it is not clear how the differences should be measured. In this paper, we compare the minimax method to clinical practice using DVH families and dose statistics. Further studies should be employed before guidelines regarding which method to choose for which cases can be given.

## V. CONCLUSION

We have proposed using minimax optimization for handling uncertainties in IMPT. By incorporating information about the uncertainties into the optimization, the method enables the optimizer to determine where to deposit dose in order to achieve robust plans. This leads to better utilization of the modality and eliminates some of the problems associated with margins in IMPT. For three patient cases, it has been shown to provide more robust target coverage and better sparing of healthy tissues than methods using margins, SFUD, and MO to account for uncertainties.

TABLE IV. Dose statistics for the prostate case. The doses are in Gy. Here, $d_x$ denotes minimum dose to $x$ % of the volume, $\bar{d}$ denotes the nominal scenario mean dose level in the volume, and a circumflex denotes dose in the worst scenario (i.e., minimum for $\hat{d}_{98}$ and maximum otherwise). External$_{0.5}$ is composed the voxels of the external ROI that receive 0.5 Gy or more and $v$ is its volume normalized to that of the External$_{0.5}$ ROI in the minimax plan. Values that differ from those of the minimax plan 2 Gy or more are in bold. The $d_{10}$, $\hat{d}_{10}$, and $\bar{d}$ dose levels for the bladder of the different plans differed by less than 2 Gy, but those of the minimax plan were at least 1 Gy below those of the other plans.

| | CTV | | | Rectum | | | External$_{0.5}$ |
|---|---|---|---|---|---|---|---|
| | $\hat{d}_{98}$ | $\hat{d}_2$ | $\bar{d}$ | $d_{10}$ | $\hat{d}_{10}$ | $\bar{d}$ | $\bar{d}\cdot v$ |
| Minimax | 68.3 | 71.6 | 70.1 | 22.5 | 54.9 | 6.9 | 19.3 |
| IMPT | **64.0** | 73.5 | 70.0 | **33.6** | **64.0** | **9.0** | 21.1 |
| SFUD | 67.4 | 71.7 | 70.0 | **33.0** | **61.1** | **9.7** | 20.6 |

## ACKNOWLEDGMENTS

The authors thank Henrik Rehbinder for valuable discussions. This research was supported by the Swedish Research Council (VR).

## APPENDIX A: MINIMAX STOCHASTIC FORMULATION

The minimax optimization and the stochastic programming formulations are special cases of the more general minimax stochastic programming formulation (see, e.g., Shapiro and Ahmed[21] and references therein). In this approach, the expected value of the objective function is minimized, but the probability distributions of the uncertain factors are unknown or, as in Chan *et al.*,[10] are themselves subject to uncertainty. For completeness, we include the derivation for our problem. Thus, assume that the probability $p_s$ of scenario $s \in \mathcal{S}$ occurring is uncertain, but is known to lie within the interval $[a_s, b_s]$, where $0 \le a_s \le b_s$. Let $\mathcal{P}$ denote the set of all probability distributions obeying the bounds, given by

$$\mathcal{P} = \left\{ p \in \mathbb{R}^{|\mathcal{S}|}:\ p_s \in [a_s, b_s]\ \forall\, s \in \mathcal{S},\quad \sum_{s\in\mathcal{S}} p_s = 1 \right\}.$$

We want to solve the minimax stochastic programming problem

$$\underset{x}{\text{minimize}} \quad \max_{p\in\mathcal{P}} \left\{ \sum_{s\in\mathcal{S}} p_s f(d(x,s)) \right\} \tag{A1a}$$

$$\text{subject to} \quad x \ge 0. \tag{A1b}$$

Since the set $\mathcal{P}$ is infinite, the max-function in the objective is hard to evaluate as stated. However, using linear programming duality, it can be readily solved.[19,20] The max-function is then viewed as a linear program in which $f(d(x,s))$ and $s \in \mathcal{S}$, can be considered as known constants. The dual of this linear program is given by

$$\underset{\lambda,\mu,\nu}{\text{minimize}} \quad \lambda - \sum_{s\in\mathcal{S}} (\mu_s a_s - \nu_s b_s)$$

$$\text{subject to} \quad f(d(x,s)) - \lambda + \mu_s - \nu_s \le 0 \quad \forall\, s \in \mathcal{S},$$

$$\mu, \nu \ge 0,$$

By substituting this for the max-function in Eq. (A1), the minimax stochastic program can be equivalently written as

$$\underset{x,\lambda,\mu,\nu}{\text{minimize}} \quad \lambda - \sum_{s\in\mathcal{S}} (\mu_s a_s - \nu_s b_s)$$

$$\text{subject to} \quad f(d(x,s)) - \lambda + \mu_s - \nu_s \le 0 \quad \forall\, s \in \mathcal{S},$$

$$x, \mu, \nu \ge 0,$$

which is a nonlinear program with the same number of nonlinear constraints as the minimax formulation (2). It too is convex whenever the function $f$ is.

## APPENDIX B: SETUP SHIFTS USED IN THE OPTIMIZATION

Since each beam has seven possible setup shifts, there are $7^b$ possible combinations of beam shifts, where $b$ is the number of beams. Not all of these are physically realizable. The shift positions for each beam are arranged in three rows of, respectively, two, three, and two positions forming the corners of a hexagon and its central point. For two parallel opposed beams, exactly seven of the combinations may be physically realized since both beams must shift in the same direction; for two orthogonal beams, for each row of shifts, each shift of the first beam can be combined with each shift of the second beam, yielding $4+9+4=17$ shifts; and for three beams with 45° gantry spacing, introducing nearest neighbor approximation, the upper and lower row combinations of the orthogonal beams can be combined with the central beam positions to yield six possible combinations each, whereas each of the nine middle combinations can be approximated, totaling to $6+9+6=21$ shifts.

[a]Electronic addresses: albin.fredriksson@raysearchlabs.com and albfre@kth.se

[1] B. Schaffner and E. Pedroni, "The precision of proton range calculations in proton radiotherapy treatment planning: Experimental verification of the relation between CT-HU and proton stopping power," Phys. Med. Biol. **43**, 1579–1592 (1998).

[2] A. J. Lomax, "Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: The potential effects of calculational uncertain-

ties," Phys. Med. Biol. **53**, 1027–1042 (2008).
[3]ICRU, "Prescribing, recording, and reporting proton-beam therapy," ICRU Report No. 78 (Oxford University Press, Oxford, 2007), pp. 83–94.
[4]A. J. Lomax, "Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: The potential effects of inter-fraction and inter-field motions," Phys. Med. Biol. **53**, 1043–1056 (2008).
[5]J. Unkelbach, T. C. Y. Chan, and T. Bortfeld, "Accounting for range uncertainties in the optimization of intensity modulated proton therapy," Phys. Med. Biol. **52**, 2755–2773 (2007).
[6]T. C. Y. Chan, "Optimization under uncertainty in radiation therapy," Ph.D. thesis, Massachusetts Institute of Technology, 2007.
[7]J. Unkelbach, T. Bortfeld, B. C. Martin, and M. Soukup, "Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning," Med. Phys. **36**, 149–163 (2009).
[8]D. Pflugfelder, J. J. Wilkens, and U. Oelfke, "Worst case optimization: A method to account for uncertainties in the optimization of intensity modulated proton therapy," Phys. Med. Biol. **53**, 1689–1700 (2008).
[9]M. Chu, Y. Zinchenko, S. G. Henderson, and M. B. Sharpe, "Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty," Phys. Med. Biol. **50**, 5463–5477 (2005).
[10]T. C. Y. Chan, T. Bortfeld, and J. N. Tsitsiklis, "A robust approach to IMRT optimization," Phys. Med. Biol. **51**, 2567–2583 (2006).
[11]A. Ólafsson and S. J. Wright, "Efficient schemes for robust IMRT treatment planning," Phys. Med. Biol. **51**, 5621–5642 (2006).
[12]D. Bertsimas and M. Sim, "The price of robustness," Oper. Res. **52**, 35–53 (2004); http://or.journal.informs.org/cgi/reprint/52/1/35.pdf.
[13]A. Ben-Tal and A. Nemirovski, "Robust convex optimization," Math. Oper. Res. **23**, 769–805 (1998); http://mor.journal.informs.org/cgi/reprint/23/4/769.pdf.
[14]J. Löf, B. K. Lind, and A. Brahme, "Optimal radiation beam profiles considering the stochastic process of patient positioning in fractionated radiation therapy," Inverse Probl. **11**, 1189–1209 (1995).
[15]J. Löf, B. K. Lind, and A. Brahme, "An adaptive control algorithm for optimization of intensity modulated radiotherapy considering uncertainties in beam profiles, patient set-up and internal organ motion," Phys. Med. Biol. **43**, 1605–1628 (1998).
[16]J. Unkelbach and U. Oelfke, "Inclusion of organ movements in IMRT treatment planning via inverse planning based on probability distributions," Phys. Med. Biol. **49**, 4005–4029 (2004).
[17]B. Sobotta, M. Söhn, and M. Alber, "Robust optimization based upon statistical theory," Med. Phys. **37**, 4019–4028 (2010).
[18]J. Llacer, J. O. Deasy, T. R. Bortfeld, T. D. Solberg, and C. Promberger, "Absence of multiple local minima effects in intensity modulated optimization with dose volume constraints," Phys. Med. Biol. **48**, 183–210 (2003).
[19]J. H. B. Kemperman, "The general moment problem, a geometric approach," Ann. Math. Stat. **39**, 93–122 (1968).
[20]J. Dupačová, "Minimax stochastic programs with nonseparable penalties," in *Optimization Techniques*, Lecture Notes in Control and Information Sciences Vol. 22 (Springer-Verlag, Berlin, 1980), pp. 157–163.
[21]A. Shapiro and S. Ahmed, "On a class of minimax stochastic programs," SIAM J. Optim. **14**, 1237–1249 (2004).
[22]P. E. Gill, W. Murray, and M. A. Saunders, "SNOPT: An SQP algorithm for large-scale constrained optimization," SIAM Rev. **47**, 99–131 (2005).
[23]M. F. Moyers, D. W. Miller, D. A. Bush, and J. D. Slater, "Methodologies and tools for proton beam design for lung tumors," Int. J. Radiat. Oncol., Biol., Phys. **49**, 1429–1438 (2001).
[24]M. Stuschke, A. Kaiser, D. Geismar, and C. Pöttgen, "Definitive proton radiotherapy in combination with standard chemotherapy for locally advanced NSCLC Stage III A & B" (2009).
[25]J. Meyer, J. Bluett, R. Amos, L. Levy, S. Choi, Q.-N. Nguyen, X. R. Zhu, M. Gillin, and A. Lee, "Spot scanning proton beam therapy for prostate cancer: Treatment planning technique and analysis of consequences of rotational and translational alignment errors," Int. J. Radiat. Oncol., Biol., Phys. **78**, 428–434 (2010).
[26]See supplementary material at http://dx.doi.org/10.1118/1.3556559 E-MPHYA6-38-049103 for transversal slices of the beam dose distributions of all cases.