

Robust Management of Motion Uncertainty in Intensity-Modulated Radiation Therapy

Thomas Bortfeld

Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School,
Boston, Massachusetts 02114, tbortfeld@partners.org

Timothy C. Y. Chan

Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139,
tcychan@alum.mit.edu

Alexei Trofimov

Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School,
Boston, Massachusetts 02114, atrofimov@partners.org

John N. Tsitsiklis

Laboratory for Information and Decision Systems, Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139, jnt@mit.edu

Radiation therapy is subject to uncertainties that need to be accounted for when determining a suitable treatment plan for a cancer patient. For lung and liver tumors, the presence of breathing motion during treatment is a challenge to the effective and reliable delivery of the radiation. In this paper, we build a model of motion uncertainty using probability density functions that describe breathing motion, and provide a robust formulation of the problem of optimizing intensity-modulated radiation therapy. We populate our model with real patient data and measure the robustness of the resulting solutions on a clinical lung example. Our robust framework generalizes current mathematical programming formulations that account for motion, and gives insight into the trade-off between sparing the healthy tissues and ensuring that the tumor receives sufficient dose. For comparison, we also compute solutions to a nominal (no uncertainty) and margin (worst-case) formulation. In our experiments, we found that the nominal solution typically underdosed the tumor in the unacceptable range of 6% to 11%, whereas the robust solution underdosed by only 1% to 2% in the worst case. In addition, the robust solution reduced the total dose delivered to the main organ-at-risk (the left lung) by roughly 11% on average, as compared to the margin solution.

Subject classifications: linear programming; applications; health care: treatment.

Area of review: Special Issue on Operations Research in Health Care.

History: Received December 2006; revision received March 2007; accepted April 2007.

1. Introduction

For the year 2006, the American Cancer Society (2006) estimated that there would be roughly 1.4 million new cases of cancer and 550,000 cancer deaths in the United States. Overall, cancer accounts for roughly 25% of all deaths in the United States. Even in comparison to other cancers, lung cancer stands out as particularly deadly and difficult to treat. Although the incidence of lung cancer is falling due to the decreasing prevalence of smoking, it is still by far the leading cause of cancer-related death in both men and women, accounting for roughly 29% of all cancer deaths. It outpaces the second leading causes of cancer death, which are prostate/colorectal cancer for men and breast cancer for women, by a factor of roughly three and two, respectively. As we will describe below, we will use some unique characteristics of tumors in the lung as motivation for our work.

Over half of all cancer patients receive radiation therapy at some point during their treatment. We will focus on

the case of external beam radiation therapy, where a linear accelerator mounted on a rotating gantry is used to treat a patient with high-energy photon beams. A photon beam deposits energy as it passes through tissue, damaging not only the tumor cells but also the healthy tissue in front of and behind the tumor. For this reason, radiation is delivered from many different angles so that each individual beam will deliver a small amount of dose to the healthy cells in its path, whereas the overlapping region will be an area of high dose, centered on the tumor.

The process for this type of treatment often starts with a planning session in which data on the internal anatomy are acquired. Using these data, along with objectives and constraints (or sometimes just rough guidelines) provided by a physician, a treatment planner determines a suitable treatment with the help of a software system. Once the treatment plan is approved, it is typically delivered on a daily basis (five days a week) over the course of several weeks (typically four to six). During this time, imaging

data are rarely reacquired, partly because tumor response to radiation is often substantially delayed.

In recent years, an advanced technique known as intensity-modulated radiation therapy (IMRT) has garnered much research and clinical attention, and is now in clinical use at most radiation oncology centers in the United States (Mell et al. 2005). We can think of IMRT as partitioning each beam into a large set of “beamlets” (e.g., 5 mm × 5 mm in size), which have individually adjustable intensities. The flexibility to deliver intensity that is nonuniform across the beam can enhance the conformity of the dose to the tumor and the sparing of healthy tissue, especially for complex shapes of the target volume (IMRT Collaborative Working Group 2001, Webb 2003, Bortfeld 2006). The calculation of the “intensity maps” for each beam is generally accomplished with “inverse” treatment-planning systems that employ optimization approaches in which the intensities of the individual beamlets are the variables. The delivery of intensity-modulated beams has been accomplished in a number of ways, but the multileaf collimator (MLC) is the most commonly used apparatus (IMRT Collaborative Working Group 2001).

In this paper, we consider IMRT treatment planning under uncertainty. IMRT is generally more vulnerable to uncertainties because of the steeper dose gradients (how quickly the dose falls off from areas of high dose to areas of low dose) it is able to produce. There are many sources of uncertainty that need to be taken into account in the course of the treatment-planning process. The segmentation of the tumor and the critical structures based on medical images is inherently uncertain and error prone. Furthermore, patient-positioning uncertainties are relevant because the patient needs to be set up in the same position every day over the course of the treatment. Motion effects constitute another class of uncertainty. In particular, there are two types of motion: interfraction and intrafraction motion. Interfraction motion refers to motion between treatment sessions (also known as “fractions”). An example of this type of motion comes from the variations in the location of a prostate tumor due to differences in rectal and bladder filling from day to day. Intrafraction motion refers to motion that occurs during a treatment session, such as breathing motion.

We draw motivation from lung cancer, and we will focus on breathing motion as the source of uncertainty. As a patient breathes during treatment, a tumor in the lung will be moving (possibly irregularly), and we would like to find a method that reliably delivers a sufficient amount of dose to the tumor, while sparing the healthy tissue as much as possible. Langen and Jones (2001) give a detailed review on the topic of motion in radiation therapy. The dosimetric effects of respiratory motion in IMRT were studied by Bortfeld et al. (2004). It was found that, to first order, the effects can be described as a “blurring” of the spatial dose distribution. In particular, steep dose gradients are washed out due to the motion, whereas secondary effects such as interplay between breathing motion and the motion of the

delivery device (MLC) can often be neglected. An important consequence of dose blurring due to motion is that, without compensation, motion leads to regions of underdose (“cold spots”) near the edge of the target volume. Mathematically, the dose-blurring effect can be modeled as a convolution of the static spatial dose distribution with a probability density function (p.d.f.) that describes the motion (Lujan et al. 1999, Engelsman et al. 2005, Bortfeld et al. 2002). This p.d.f. specifies the relative amount of time the tumor (and other tissue) spends in different locations or, in an alternative interpretation that we will use in this paper, the relative amount of time spent in different phases of the breathing cycle.

A number of strategies have been proposed to reduce the effects of motion (see Keall et al. 2006 for a comprehensive review). The methods fall into two main categories: (i) reduction of the amplitude of the motion relative to the treatment device (e.g., breath hold, gating, tumor tracking), and (ii) reduction of the dosimetric consequences of the motion. The former methods can be somewhat unreliable and difficult for the patient to tolerate (breath hold), technologically demanding (gating and tracking), or may prolong the treatment time (breath hold and gating). The most common method, which belongs to the latter category (ii), is the use of a *margin*, which is an expanded region around the nominal position of the tumor in which we are confident the tumor will remain. Then, by treating this margin area with a high uniform dose, we ensure that the tumor receives a sufficient amount of dose, no matter how it moves within this margin. Of course, the main disadvantage of using a margin is that the healthy tissue surrounding the tumor will receive more dose than necessary. We can think of the use of a margin as a worst-case approach, in which we use some bounds on the motion, but assume no additional information on how the tumor will move within these bounds. We show later that the margin is indeed the correct approach in this case, and define more precisely the worst case that the margin is protecting against.

A recent development that also belongs to category (ii) is that of “4D” optimization, which aims at undoing (deconvolving) the blurring effect of the motion. The basic idea is to hit the tumor edges harder (with higher intensities), and thus compensate for the blurring effect and possible dose cold spots (Trofimov et al. 2005, Zhang et al. 2004). These recent studies showed promising results, but relied on the assumption that the breathing motion realized during every treatment fraction was exactly the same as the one that was exhibited during the planning session. Hence, this approach assumes “perfect information” or “no uncertainty” on the motion p.d.f., and is appropriate if we can be certain that the patient will breathe in the same predictable way during each fraction. On the other hand, Sheng et al. (2006) showed that the quality of the resulting dose distribution can be seriously degraded if this assumption is violated. The subject of changes in respiratory patterns between treatment simulation and treatment delivery

was the very first in a list of topics recommended for further research in a recent report on the management of respiratory motion in radiation oncology by a group of experts from the American Association of Physicists in Medicine (Keall et al. 2006).

Our work in this paper strives to generalize and improve upon previous studies by relaxing the rather strong assumption of highly regular and reproducible breathing. In particular, we develop a model of uncertainty in the motion p.d.f., and incorporate it into a robust optimization framework that can produce treatment plans that are robust to this uncertainty—that is, plans that are guaranteed to be feasible under any realization of the uncertainty. Recent research has indeed highlighted the topic of robustness in radiation therapy optimization (Baum et al. 2006, Unkelbach and Oelfke 2004). The combination of the need for more rigorous optimization approaches to deal with clinical uncertainties, together with advances in deriving tractable robust counterparts for uncertain mathematical programs, has led to a fair amount of recent interest in the use of robust optimization methods in radiation therapy treatment planning (Chu et al. 2005, Ólafsson and Wright 2006). Some of our ideas on this subject were sketched in Chan et al. (2006), which included preliminary results on artificial, one-dimensional problems. More generally, the topic of (not necessarily robust) optimization in radiation therapy has received a lot of recent attention from the optimization community (Romeijn et al. 2006, Ferris et al. 2003, Lee et al. 2003, Preciado-Walters et al. 2004, Holder 2003).

This paper is organized into two major parts. In the first part, we develop a general approach to robust IMRT treatment planning in the presence of motion uncertainty, and provide a full discussion of the associated mathematical formulations. In particular, in §2, we present a model of uncertainty, and in §3, we develop our robust formulation. In the second part of this paper, we apply our approach to real patient data. In §4, we describe our patient data set and a method for incorporating the information in this data set into our formulation. In §5, we illustrate the results obtained from our robust formulation, and compare them to the results from the margin and motion p.d.f. approaches. We conclude in §6 with some discussion and final remarks.

2. Model of Uncertainty

The presence of motion during radiation delivery spreads out the dose to areas surrounding the target region. To compensate for this blurring effect, the “motion p.d.f.” approach aims to produce a dose distribution that when convolved with the assumed p.d.f. recovers the desired distribution. Naturally, this method relies on a priori knowledge of what the motion p.d.f. will be during treatment. If the realized motion is quite different from what was assumed, then convolving the realized motion p.d.f. with a dose distribution optimized for a different p.d.f. is likely to result in an unacceptable dose distribution with hot and cold spots. This

motivates the need for an approach that can mitigate the uncertainty in the p.d.f. to be realized during treatment.

Our framework involves a finite set X , which corresponds to the set of possible phases of the breathing cycle. A (*motion*) p.d.f. is defined as a nonnegative real function $f: X \rightarrow \mathbb{R}$, such that $\sum_{x \in X} f(x) = 1$. Our starting point is a *nominal* p.d.f., denoted by p , that is constructed on the basis of data taken during the planning session. Regarding the actual p.d.f., \tilde{p} , to be realized during treatment, we assume that it can differ from the nominal p.d.f., p , on a subset U of the domain X , and satisfies

$$p(x) - \underline{p}(x) \leq \tilde{p}(x) \leq p(x) + \bar{p}(x) \quad \forall x \in U,$$

where the functions \underline{p} and \bar{p} provide bounds on the difference between the nominal and the realized p.d.f. during treatment. Without loss of generality, we assume that $p - \underline{p} \geq 0$ and $p + \bar{p} \leq 1$. Thus, the set P_U , of possible realized p.d.f.s, called the *uncertainty set*, is defined as

$$P_U = \left\{ \tilde{p} \in \mathbb{R}^{|X|}: \tilde{p}(x) \in [p(x) - \underline{p}(x), p(x) + \bar{p}(x)] \quad \forall x \in U; \right. \\ \left. \tilde{p}(x) = p(x) \quad \forall x \in X \setminus U; \sum_{x \in X} \tilde{p}(x) = 1 \right\}. \quad (1)$$

Note that the set U is somewhat redundant because it can be accounted for by setting $\bar{p}(x) = \underline{p}(x) = 0$ for $x \in X \setminus U$, but we will keep it for clarity of exposition. We will refer to the upper and lower bounds on \tilde{p} as *error bars*.

A treatment plan will be said to be *robust* if all of the constraints in our subsequent formulation are satisfied (cf. §3), no matter which p.d.f. from the set P_U is realized. Thus, P_U is to be interpreted as the set of all p.d.f.s that we are protecting against in our optimization. If one views this approach as being overly conservative (because it permits certain implausible, highly oscillatory p.d.f.s), one can introduce simple (linear) “smoothness” constraints into the definition of P_U , such as

$$|\tilde{p}(x) - \tilde{p}(y)| \leq \epsilon \quad \text{if } |x - y| \leq \delta,$$

for suitable $\epsilon, \delta > 0$.

One of the main challenges here is to keep the set P_U large enough so that we can protect against realistic variations that the patient may exhibit, without losing all information about the patient’s particular breathing patterns, and just delivering a margin. The practical construction of this set from real data is discussed in §4.1.

3. Formulations

This section introduces the different formulations that will be investigated. In our formulations, the goal is to minimize the total dose delivered, while ensuring that the tumor receives an adequate dose. In §3.1, we start with the basic problem of finding the optimal intensities of the beamlets in the IMRT planning problem when there is no motion and

no uncertainty. We then continue with a nominal formulation (a linear program) that takes into account a nominal motion p.d.f. In §3.2, we present a robust formulation in which the possible tumor motion is modeled by an uncertainty set P_U , and demonstrate that it can also be cast as a linear program. Finally, in §3.3, we describe the classical margin approach and explore its relation to our robust formulation.

3.1. Nominal Formulation

We assume that the treatment will be delivered using beamlets belonging to a given set \mathcal{B} , and we let w_b be the weight or intensity of beamlet b (related to the amount of time the beamlet is “on”). We are interested in a *phantom*, represented by a set \mathcal{V} of *voxels*. A voxel is a small volume-element used to represent a specific location in the patient. A phantom is a physical model of the human body or any of its parts, and here it is synonymous with the set of voxels that represent the tumor and all healthy tissue. Let \mathcal{T} be the set of voxels in the tumor, and let \mathcal{N} be the set of voxels in the normal (nontumor) tissue. Furthermore, let $m_{\mathcal{T}} = |\mathcal{T}|$, $n = |\mathcal{B}|$, and $m_U = |U|$. Let $D_{v,b}$ be the dose that voxel v receives per unit of intensity of beamlet b , and let θ_v be the prescribed dose that voxel v should receive. The following formulation aims at minimizing the total delivered dose, in the absence of motion, and will be referred to as the *basic problem*:

$$\begin{aligned} & \underset{w}{\text{minimize}} && \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} D_{v,b} w_b \\ & \text{subject to} && \sum_{b \in \mathcal{B}} D_{v,b} w_b \geq \theta_v \quad \forall v \in \mathcal{T}, \\ & && w_b \geq 0 \quad \forall b \in \mathcal{B}. \end{aligned} \tag{2}$$

In practice, the prescribed dose is often uniform over the tumor, in which case we have $\theta_v = \theta$ for all $v \in \mathcal{T}$.

The next step is to incorporate motion (but not yet motion uncertainty) into (2), using the nominal p.d.f., p , described in §2. In the presence of motion, we need to fix some nominal reference frame in relation to which the voxels will be moving. For example, this reference frame could correspond to the position of the internal anatomy when a patient is in the full exhale phase of breathing. We also need to adjust the “matrix” D to account for the motion. For that purpose, we create a “matrix” Δ , with components $\Delta_{v,x,b}$ that describe the dose delivered to voxel v , when the anatomy is in breathing phase x , per unit intensity of beamlet b . Without loss of generality, we can assume that the nominal reference frame corresponds to breathing phase 0 and the position of the anatomy in the basic problem (i.e., in the absence of motion). Hence, $\Delta_{v,x,b} = D_{v,b}$ when $x = 0$. With this picture in mind, $p(x)$ is the probability that the patient is in breathing phase x . This definition facilitates the incorporation of nonrigid-body motion into the problem formulation because $\Delta_{v,x,b}$ is computed for snapshots of the anatomy in each phase (as opposed to

extrapolating based on rigid-body motion). Note that rigid-body motion can be cast as a special case of this definition, where x corresponds to displacement from a nominal position so that, again, $x = 0$ corresponds to the nominal position.

Adjusting for motion, as described by the nominal p.d.f., we obtain the following *nominal problem*:

$$\begin{aligned} & \underset{w}{\text{minimize}} && \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b \\ & \text{subject to} && \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b \geq \theta_v \quad \forall v \in \mathcal{T}, \\ & && w_b \geq 0 \quad \forall b \in \mathcal{B}. \end{aligned} \tag{3}$$

This formulation is a straightforward extension of (2), in which the coefficients $D_{v,b}$ are replaced with

$$\sum_{x \in X} \Delta_{v,x,b} p(x)$$

to account for the motion as described by p . Note that the p.d.f. used in the objective function and constraints does not correspond to just one fraction, but refers to the average p.d.f. over the entire course of treatment. Although it is true that during one fraction we do not expect the dose delivered to be well approximated by the expression in the objective function, we do expect convergence of the delivered dose to the expected delivered dose (represented by the p.d.f. p) after many treatment fractions (Bortfeld et al. 2002).

Formulation (3) is the appropriate formulation when there is “perfect information,” in the sense that the breathing-induced motion is accurately represented by p . However, if we use this formulation and the patient breathes according to a different p.d.f., the lower-bound constraints on the tumor dose are likely to be violated. This motivates the inclusion of the model of uncertainty, as described in §2, into formulation (3).

3.2. Robust Formulation

In the presence of uncertainty, we want the tumor to receive the required dose no matter which p.d.f. within the uncertainty set is realized during treatment. This desire is captured by requiring the lower-bound constraint in formulation (3) to be satisfied when p is replaced by \tilde{p} for every $\tilde{p} \in P_U$. This leads to the following *robust problem*:

$$\begin{aligned} & \underset{w}{\text{minimize}} && \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b \\ & \text{subject to} && \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} \tilde{p}(x) w_b \geq \theta_v \\ & && \forall v \in \mathcal{T}, \forall \tilde{p} \in P_U, \\ & && w_b \geq 0 \quad \forall b \in \mathcal{B}. \end{aligned} \tag{4}$$

Note that the uncertainty is in the p.d.f., and not in the Δ matrix. This corresponds to the assumption that the possible geometrical configurations of the anatomy (i.e., path

of the motion trajectory), due to motion, do not change between treatment planning and treatment delivery, but that the fraction of time spent in each configuration (i.e., breathing phase or location on the trajectory) is uncertain.

Furthermore, note that the uncertainty in p is accounted for in the constraints, but not in the objective. To gain some understanding of the implications, note that one could just minimize the total dose delivered to the anatomy in a particular phase x . However, the choice of x would have been arbitrary and subjective. Instead, our goal is to use a reasonable objective function that approximates well the total dose delivered to the healthy tissue. By using p in the objective, we are essentially averaging the dose delivered over the various phases because this objective function is a convex combination of the corresponding objective functions associated with individual phases. Different p in the objective will indeed redistribute the dose between voxels, but will not change the integral dose appreciably because all p.d.f.s integrate to one.

There are other types of objective functions that would require accounting for the uncertainty in the p.d.f. For example, if we are interested in an objective such as minimizing the maximum dose delivered to the healthy tissue, \mathcal{N} , we should replace the objective with an auxiliary variable z , and introduce to formulation (4) additional constraints of the form

$$z \geq \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} \tilde{p}(x) w_b \quad \forall v \in \mathcal{N}, \quad \forall \tilde{p} \in P_U. \quad (5)$$

These constraints have essentially the same structure as the lower-bound constraints in the robust problem, and the analysis provided in the rest of this section (i.e., the derivation of a finite LP) still applies to formulation (4) augmented with constraints (5).

Formulation (4) is not a linear program due to the infinite number of constraints. In what follows, we transform (4) into an equivalent linear program, and also develop some insights on its structure. The proof is straightforward and is omitted.

PROPOSITION 1. *Formulation (4) is equivalent to*

$$\begin{aligned} & \underset{w}{\text{minimize}} && \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b \\ & \text{subject to} && \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b + \beta_v(\mathbf{w}) \geq \theta_v \\ & && \forall v \in \mathcal{T}, \\ & && w_b \geq 0 \quad \forall b \in \mathcal{B}, \end{aligned} \quad (6)$$

where

$$\begin{aligned} \beta_v(\mathbf{w}) = \min_{\hat{p}} & \sum_{b \in \mathcal{B}} \sum_{x \in U} \Delta_{v,x,b} \hat{p}(x) w_b \\ & \text{subject to} \quad \sum_{x \in U} \hat{p}(x) = 0, \\ & -\underline{p}(x) \leq \hat{p}(x) \leq \bar{p}(x) \quad \forall x \in U. \end{aligned} \quad (7)$$

The quantity $\beta_v(\mathbf{w})$ arises because we are protecting against worst-case (in terms of underdosing the tumor) motion within our uncertainty set P_U . Next, we investigate the structure of the solution of problem (7).

PROPOSITION 2. *Let*

$$d_{v,x}(\mathbf{w}) := \sum_{b \in \mathcal{B}} \Delta_{v,x,b} w_b.$$

For any v and \mathbf{w} , we introduce an ordering $x(1), \dots, x(m_U)$ of the elements of X such that $d_{v,x(1)}(\mathbf{w}) \leq \dots \leq d_{v,x(m_U)}(\mathbf{w})$. The optimal value of (7) is equal to

$$\begin{aligned} \beta_v(\mathbf{w}) = & - \sum_{b \in \mathcal{B}} \sum_{j > j^*} (\Delta_{v,x(j),b} - \Delta_{v,x(j^*),b}) \underline{p}(x(j)) w_b \\ & - \sum_{b \in \mathcal{B}} \sum_{j < j^*} (\Delta_{v,x(j^*),b} - \Delta_{v,x(j),b}) \bar{p}(x(j)) w_b, \end{aligned} \quad (8)$$

where j^* satisfies the inequalities

$$\begin{aligned} \sum_{j \geq j^*} \underline{p}(x(j)) - \sum_{j < j^*} \bar{p}(x(j)) & \geq 0, \\ \sum_{j \geq j^*+1} \underline{p}(x(j)) - \sum_{j < j^*+1} \bar{p}(x(j)) & < 0. \end{aligned}$$

The expression for $\beta_v(\mathbf{w})$ in (8) gives insight into the structure of the robust dose distribution generated by the solution to the robust formulation (4). Let \mathbf{w} be an optimal solution to the nominal problem and consider the resulting dose delivered

$$\sum_{b \in \mathcal{B}} D_{v,b} w_b,$$

not to a particular anatomical voxel, but to the point in space where voxel v resides in its nominal position; this is the static dose distribution. Naturally, there may be regions of high static dose and regions of low static dose. Now imagine that a particular voxel v moves within this dose distribution, due to breathing motion. The solution \mathbf{w} assumes that v will move according to the nominal p , “acquiring” just the right amount of dose from both the high- and low-dose regions of the static dose distribution. However, if v moves according to a different p.d.f., we must protect against the scenario where v spends more time in the low-dose locations and less time in the high-dose locations than nominally assumed. By protecting against this situation, which is exactly what $\beta_v(\mathbf{w})$ represents in (8), we are ensuring feasibility over the entire uncertainty set. Practically, a robust dose distribution will be smoother (fewer peaks and valleys), as compared to a nominal dose distribution, to ensure that voxels that move according to a different p.d.f. still receive enough dose.

To create a tractable formulation that is equivalent to (4), we need the following proposition, which is essentially a *theorem of the alternative*.

PROPOSITION 3. For a given v and z , we have $\beta_v(\mathbf{w}) \geq z$ if and only if there exists q_v and \mathbf{r}_v such that

$$\begin{aligned} & \sum_{x \in U} \underline{p}(x)q_v - \sum_{x \in U} r_{v,x} - \sum_{b \in \mathcal{B}} \sum_{x \in U} \Delta_{v,x,b} \underline{p}(x)w_b \geq z, \\ & (\underline{p}(x) + \bar{p}(x))q_v - r_{v,x} \\ & \leq \sum_{b \in \mathcal{B}} \Delta_{v,x,b} (\bar{p}(x) + \underline{p}(x))w_b \quad \forall x \in U, \end{aligned} \tag{9}$$

q_v free,

$$r_{v,x} \geq 0 \quad \forall x \in U.$$

PROOF. The dual of (7) is

$$\begin{aligned} & \text{maximize}_{q_v, r_{v,x}, s_{v,x}} \sum_{x \in U} \bar{p}(x)r_{v,x} + \sum_{x \in U} \underline{p}(x)s_{v,x} \\ & \text{subject to } q_v + r_{v,x} + s_{v,x} = \sum_{b \in \mathcal{B}} \Delta_{v,x,b} w_b \quad \forall x \in U, \\ & q_v \text{ free,} \\ & r_{v,x} \leq 0 \quad \forall x \in U, \\ & s_{v,x} \geq 0 \quad \forall x \in U. \end{aligned} \tag{10}$$

Using the equality constraint to eliminate the $s_{v,x}$ variables, and replacing the variables $r_{v,x}$ with $-r_{v,x}/(\bar{p}(x) + \underline{p}(x))$, we arrive at an equivalent formulation

$$\begin{aligned} & \text{maximize}_{q_v, r_{v,x}} \sum_{x \in U} \underline{p}(x)q_v - \sum_{x \in U} r_{v,x} - \sum_{b \in \mathcal{B}} \sum_{x \in U} \Delta_{v,x,b} \underline{p}(x)w_b \\ & \text{subject to } (\bar{p}(x) + \underline{p}(x))q_v - r_{v,x} \\ & \leq \sum_{b \in \mathcal{B}} \Delta_{v,x,b} (\bar{p}(x) + \underline{p}(x))w_b \quad \forall x \in U, \tag{11} \\ & q_v \text{ free,} \\ & r_{v,x} \geq 0 \quad \forall x \in U. \end{aligned}$$

Because formulation (7) is feasible, the result follows from strong duality of linear programming. \square

We now use the above proposition to transform (4) into a linear program.

THEOREM 1. Formulation (6) is equivalent to the following linear program:

$$\begin{aligned} & \text{minimize}_{\mathbf{w}, \mathbf{q}, \mathbf{r}} \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x)w_b \\ & \text{subject to } \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x)w_b - \sum_b \sum_{x \in U} \Delta_{v,x,b} \underline{p}(x)w_b \\ & \quad + \sum_{x \in U} \underline{p}(x)q_v - \sum_{x \in U} r_{v,x} \geq \theta_v \quad \forall v \in \mathcal{V}, \\ & (\bar{p}(x) + \underline{p}(x))q_v - r_{v,x} \\ & \leq \sum_b \Delta_{v,x,b} (\bar{p}(x) + \underline{p}(x))w_b \\ & \quad \forall v \in \mathcal{V}, \forall x \in U, \\ & q_v \text{ free } \quad \forall v \in \mathcal{V}, \\ & r_{v,x} \geq 0 \quad \forall v \in \mathcal{V}, \forall x \in U, \\ & w_b \geq 0 \quad \forall b \in \mathcal{B}. \end{aligned} \tag{12}$$

Formulation (12) provides us with a linear program that takes into account the uncertainty set in a tractable manner. By adjusting the size of the uncertainty set in our model, we will affect the conservativeness of the resulting robust solution. We observe that whereas the nominal formulation (3) has $m_{\mathcal{T}}$ constraints and n variables, the robust formulation (12) has $m_{\mathcal{T}}m_U + m_{\mathcal{T}}$ constraints and $m_{\mathcal{T}}m_U + m_{\mathcal{T}} + n$ variables. Although the problem now has a larger size, it is still a linear program and, therefore, tractable for realistic values of m_U and $m_{\mathcal{T}}$. See §5 for the exact problem sizes and running times of our experiments.

3.3. Margin Formulation

Independent from the development of our robust methodology, one can formulate the problem of ensuring enough dose to the tumor in the absolute worst case, without any information on the motion except for bounds on the amplitude. This type of formulation is what we will call a *margin formulation*. This name comes from the medical physics community, where a margin (an expansion of the irradiated area) is used to account for uncertainty and ensure that the tumor receives a sufficient dose. It is important to note that other studies aimed at calculating optimal margin sizes (van Herk et al. 2000, Engelsman et al. 2005) implicitly assume some distribution for an unknown p.d.f. This results in smaller margin sizes and less conservative solutions as compared to the margin approach described here, which aims to deliver a uniform dose to the tumor in the absence of a particular distribution for the p.d.f. In this section, we illustrate the connection between the margin concept and our robust formulation.

The robust formulation proposed in the previous section allows us to control the robustness of our solution depending on the amount of uncertainty present. In one extreme, one assumes perfect knowledge of the motion p.d.f., which is equivalent to the nominal formulation. At the opposite extreme, one assumes that there is no information on the motion (other than the possible range or amplitude of motion, which is captured by the set X), and protects against the absolute worst case, which is exactly the margin approach. Intuitively, we expect to be able to replicate these formulations by using the robust formulation with an appropriate uncertainty set. We will show that the robust formulation builds a bridge between these two extreme solutions and gives us the flexibility to craft solutions anywhere in between.

First, we need to more precisely define a margin. In fact, there are two types of margins to consider: a *dosimetric margin* and a *geometric margin*. The difference between these two concepts reflects the difference between using anatomical voxels to account for the tumor and other structures versus voxels that are fixed in space (based on a coordinate system independent of the patient). A dosimetric margin is the minimal (in terms of total dose delivered) dose distribution that, when added voxel by voxel to the nominal dose distribution, delivers a sufficient amount of

dose to the tumor throughout each phase of the breathing cycle. A geometric margin (similar to an internal target volume (ITV) in the medical physics literature) is constructed by taking the union of the tumor voxels throughout the breathing cycle, and subtracting away the nominal tumor voxels (the voxels representing the tumor in its nominal position). In practice, however, the dosimetric margin is the more meaningful concept because it is able to account for the fact that the motion itself may affect the dose delivered to a point in space. For example, as a result of motion, the beam of radiation may pass through ribs or different amounts of air in the lung, which would affect the dose delivered. Hence, we focus solely on the dosimetric margin in this paper, whereas in Chan et al. (2006) the focus was on a geometric margin because the problem had a simplified one-dimensional geometry.

We can write the *dosimetric margin problem* as

$$\begin{aligned} & \underset{\mathbf{w}}{\text{minimize}} && \sum_{v \in \mathcal{T}} \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b \\ & \text{subject to} && \sum_{b \in \mathcal{B}} \Delta_{v,x,b} w_b \geq \theta_v \quad \forall v \in \mathcal{T}, \forall x \in X, \\ & && w_b \geq 0 \quad \forall b \in \mathcal{B}. \end{aligned} \quad (13)$$

This formulation ensures that every voxel in the tumor receives a sufficient dose throughout the entire breathing cycle. As mentioned before, the dose to the tumor is typically desired to be uniform, in which case we have $\theta_v = \theta$ for all $v \in \mathcal{T}$.

By considering the function of the model of uncertainty, it should be intuitive that the “smallest” uncertainty set replicates the nominal formulation situation, whereas the “largest” uncertainty set would give rise to a margin-like solution. The next proposition makes this intuition concrete, illustrating the fact that the robust formulation is a generalization of the margin and nominal formulations (the proof can be found in the appendix).

PROPOSITION 4. *If $U = \emptyset$ (or equivalently, $\bar{p} = \underline{p} = 0$), then the robust formulation is equivalent to the nominal formulation. Alternatively, if $U = X$, $\bar{p} = 1 - p$, $\underline{p} = p$, and $\theta_v = \theta$ for all $v \in \mathcal{T}$, then the robust formulation is equivalent to the dosimetric margin formulation.*

We can think of the solution of the nominal formulation as the most risky because it takes at face value the information contained in the motion p.d.f. A solution to the margin formulation is the most conservative because it must protect against the largest uncertainty set. By using the robust formulation with an uncertainty set appropriate for the patient at hand, the solution will automatically be driven toward a margin solution if the uncertainty set is big, a nominal solution if the uncertainty set is small, or something in between otherwise.

The robust formulation also provides some insight into why using a margin may be a poor method in many cases. The margin solution is only optimal when $U = X$,

$\bar{p} = 1 - p$, and $\underline{p} = p$; that is, only if there is a possibility that the tumor will spend 100% of its time in any phase of the breathing cycle. This is especially dubious for extreme phases of the breathing cycle such as inhale and exhale. Thus, by gathering information on a patient’s breathing and associated uncertainty set, we should always be able to improve upon a margin solution.

4. Patient Data and P.D.F.s

Probability density function data are not immediately available from the planning session, but must be mined from other data that are collected. One type of data that is collected for patients with significant motion is data that track the movement of the tumor. The best way to obtain data on internal motion is to implant a marker into the tumor and continuously image the patient, tracking the motion of the marker. Often, though, this method is not used due to possible complications associated with the implantation procedure or the inability of the facility to carry it out. A far less invasive alternative involves the use of an external marker placed on the patient’s abdomen. By recording the motion of this external marker, we obtain information about the patient’s breathing pattern and corresponding internal motion. Then, by estimating anatomical amplitudes through imaging, we can use this external data as a surrogate for internal motion data.

The patient data used in this study were acquired at the Massachusetts General Hospital. We used Varian’s (Varian Medical Systems, Inc., Palo Alto, CA) real-time position management (RPM) system to gather data from five patients by recording their breathing patterns in every treatment fraction over the course of their entire treatments. RPM data are recorded by an infrared camera mounted in the treatment room, tracking the movement of an optical marker placed on the chest or abdomen of the patient. This camera records the respiratory motion of the patient by observing the up and down motion of the marker block. Much research has focused on correlating the external signal to internal motion (Vedam et al. 2003, Gierga et al. 2005, Tsunashima et al. 2004), and we will assume, as many studies have shown to be the case, that the internal and external movement are highly correlated. In reality, there are uncertainties that need to be taken into account when using this method, but this is outside the scope of this paper. Ideally, internal motion data would be used instead of this external data. Accordingly, we develop our framework with the understanding that future advances in medical imaging will greatly enhance the accuracy and availability of anatomical motion tracking, and will give us data that are indeed representative of internal motion.

The recent advance of breathing-correlated “4D” computed tomography imaging (Pan et al. 2004, Low et al. 2003, Keall et al. 2004) in combination with deformable image registration (McInerney and Terzopoulos 1996, Brock et al. 2003, Lu et al. 2004) has made it possible

to visualize tissue motion trajectories within the scanned 4D-CT volume of a patient, and could in principle provide an alternative to RPM data. However, although 4D-CT data are a useful tool to study the spatial variability of motion, they do not reflect temporal variability, which has a stronger influence on the delivery outcome. Furthermore, we do not acquire this internal data for each treatment fraction due to time and resource constraints. Therefore, we focus on RPM data because they are easily obtained on different days and can be used to estimate the variability in breathing from fraction to fraction.

The amplitude of an RPM trace is not representative of any particular anatomical motion because its units depend on many external factors such as the exact position and tilt of the marker block on the patient’s abdomen, and the distance from the block to the camera. Still, the RPM data are useful for measuring variations in the breathing patterns. Therefore, we scaled the average RPM amplitude for each person to the average anatomical amplitude gathered from 4D-CT scans. Each RPM trace was used to create a single p.d.f. by integrating over the displacement axes of the RPM trace. This produces a function that describes the relative amount of time the tumor spends in various phases. A detailed explanation of the generation of a p.d.f. in this manner is described in Lujan et al. (1999).

To make these p.d.f.s useful for our optimization framework, we need to aggregate the RPM data into bins that are most appropriate for the x of our matrix $\Delta_{v,x,b}$. That is, we are given the matrices $\Delta_{v,x,b}$ for a fixed set of phases, and to use any p.d.f.s generated from this process, we must tailor the support of these p.d.f.s to match the set X of possible phases. The finer the set X is, the more flexibility we have to precisely specify a p.d.f. However, the set of breathing phases X is typically small, resulting in somewhat coarse p.d.f.s.

The error bars and associated uncertainty set are generated from a family of p.d.f.s as shown in the following section. The idea is that once we have some information about the patient at hand, in the form of a nominal p.d.f. from RPM data, we can use information on breathing patterns from past patients to form an uncertainty set for this patient.

4.1. Creating the Uncertainty Set from Patient Data

The error bars in our model are formed from a family of p.d.f.s as follows. Consider a family of p.d.f.s p_0, p_1, \dots, p_k obtained from patient data (e.g., according to the method described above), where $p_0 := p$ is the nominal p.d.f. of our current patient, p_1 is the nominal p.d.f. of a past patient, and p_2, \dots, p_k are the p.d.f.s of the past patient realized during his/her treatment fractions. We are interested in finding out how this past patient deviated from his/her nominal p.d.f. on subsequent days, and using this as a measure of the relative deviation that our current patient may exhibit from his/her nominal p.d.f. Of course, this

relationship needs to be tested in practice to be useful for clinical implementation. However, the underlying assumption seems reasonable, as long as we build the model of uncertainty based on data from past patients that are appropriately similar to the patient at hand. What is deemed “appropriately similar” should be studied in future work.

Let us generate error bars for this past patient as follows. For every $x \in X$,

$$\underline{p}_1(x) := p_1(x) - \min_{i=1, \dots, k} p_i(x)$$

and

$$\bar{p}_1(x) := \max_{i=1, \dots, k} p_i(x) - p_1(x).$$

We define the relative minimum deviation below the nominal p.d.f. p_1 to be \underline{p}_1/p_1 , and the relative maximum deviation above the nominal p.d.f. p_1 as $\bar{p}_1/(1 - p_1)$. Note that the definition for the maximum takes into account the fact that the realized p.d.f.s are bounded above by 1. Using these quantities, we define \bar{p} and \underline{p} for the current patient as

$$\underline{p}(x) := \frac{\underline{p}_1(x)}{p_1(x)} p(x)$$

and

$$\bar{p}(x) := \frac{\bar{p}_1(x)}{1 - p_1(x)} (1 - p(x))$$

for every $x \in X$. Finally, the upper and lower error bars of p can be written as

$$p(x) + \bar{p}(x) = p(x) + \frac{\bar{p}_1(x)}{1 - p_1(x)} (1 - p(x))$$

and

$$p(x) - \underline{p}(x) = p(x) - \frac{\underline{p}_1(x)}{p_1(x)} p(x),$$

respectively. It is straightforward to check that using these definitions for \bar{p} and \underline{p} , $0 \leq p - \underline{p} \leq p + \bar{p} \leq 1$. The natural extension of the above procedure is to use multiple patients $j = 1, \dots, m$, to create \underline{p} and \bar{p} :

$$\underline{p}(x) := \max_{j=1, \dots, m} \frac{\underline{p}_j(x)}{p_j(x)} p(x) \tag{14}$$

and

$$\bar{p}(x) := \max_{j=1, \dots, m} \frac{\bar{p}_j(x)}{1 - p_j(x)} (1 - p(x)), \tag{15}$$

where $p_j, \bar{p}_j, \underline{p}_j$ are indexed by patient j .

5. Results from a Clinical Case

In this section, we present the results of using the three formulations (nominal, robust, and dosimetric margin) on a clinical problem, involving a tumor in the lower left lung. Our formulations are driven by a simple model where the major internal structures of interest are a tumor and the surrounding healthy lung tissue, which is a common representation of a lung cancer case. However, before implementing our formulations on a clinical case, we made two slight modifications to account for realistic considerations.

First, in the objective function, instead of minimizing the dose delivered to the entire phantom, we minimize the total dose delivered to the healthy tissue, \mathcal{N} . Studies indicate that the mean dose delivered in the lung is well correlated with the probability of normal tissue complication associated with treatment (Theuws et al. 1998).

The other modification we made was to add upper-bound constraints to control the dosage fluctuations within the tumor. In particular, we constrained the maximum dose in the tumor to be no more than 10% above the minimum dose. Using a minimum dose of 72 Gy for the tumor meant that the maximum dose within the tumor did not exceed 79.2 Gy. This is slightly more stringent than the current standards, which require that the dose in the tumor be no more than 5% below or 7% above the prescribed dose (a gap of 12%, or 12.6% above the minimum). We found that these modifications produced acceptable solutions, so we did not further refine the formulation. Of course, for very complex cases, constraints that limit the dose to certain healthy structures will be needed. These types of constraints can be added to our model without major revisions of the theory we presented. The robust formulation used in the subsequent runs is shown in (16):

$$\begin{aligned}
 & \underset{\mathbf{w}}{\text{minimize}} && \sum_{v \in \mathcal{N}} \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b \\
 & \text{subject to} && \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} \tilde{p}(x) w_b \geq \theta_v \\
 & && \forall v \in \mathcal{T}, \forall \tilde{p} \in P_U, \quad (16) \\
 & && \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} \tilde{p}(x) w_b \leq \gamma \theta_v \\
 & && \forall v \in \mathcal{T}, \forall \tilde{p} \in P_U, \\
 & && w_b \geq 0 \quad \forall b \in \mathcal{B}.
 \end{aligned}$$

As mentioned above, $\gamma = 1.1$ and $\theta_v = 72$ for all $v \in \mathcal{T}$.

The anatomy was discretized into voxels of size 2.93 mm × 2.5 mm × 2.93 mm, resulting in a total of 110,275 voxels, 5,495 of which were in the tumor. Radiation was delivered from five angles at 0° (directly anterior), 52°, 104°, 156°, and 208°, which resulted in a total of 1,625 beamlets (of size 5 mm × 5 mm). We modeled the support of the motion p.d.f.s using five phases of the breathing cycle. Our uncertainty set allowed the value of the realized p.d.f. to differ from the nominal one in every phase x ,

which corresponded to setting $U = X$. The nominal p.d.f. was taken from the pretreatment RPM trace of the current patient.

We conducted a series of experiments using p.d.f. data derived from five patients. In each experiment, the p.d.f.s corresponding to four patients were used to construct the uncertainty set (in the notation of Equations (14) and (15), we have $m = 4$), and the p.d.f. data from the fifth patient was used to test the resulting optimized solutions. This is meant to simulate the realistic situation where the uncertainty set may be generated from past data, whereas the p.d.f.s realized during treatment may lie outside this uncertainty set. Each experiment used a different combination of four patients (out of the five total patients) to generate the uncertainty set. Below, we present two cases that represent the experiments that showed the least and most difference between the nominal and robust solutions in terms of tumor coverage, respectively.

For the first experiment, we used data from four patients totaling 250 traces (multiple RPM traces are taken during most fractions) to generate the uncertainty set according to Equations (14) and (15). Then, to measure the quality of the resulting optimized solutions, we used 78 p.d.f.s from the fifth patient. We compare the performance of the nominal, robust, and margin solutions below. The three formulations were optimized using the barrier algorithm of CPLEX 9 on a 3 GHz computer with 3 GB of memory. Table 1 shows the number of variables and constraints in each formulation, along with the associated solution times.

With an optimal solution, \mathbf{w}^* , of one of these formulations in hand, we can simulate its delivery by computing

$$\sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} \tilde{p}(x) w_b^*$$

for each voxel $v \in \mathcal{V}$, where \tilde{p} is a particular realized p.d.f. Different \tilde{p} in this equation lead to different delivered dose distributions. In the case of the robust formulation, for example, a \tilde{p} that lies within the uncertainty set will guarantee

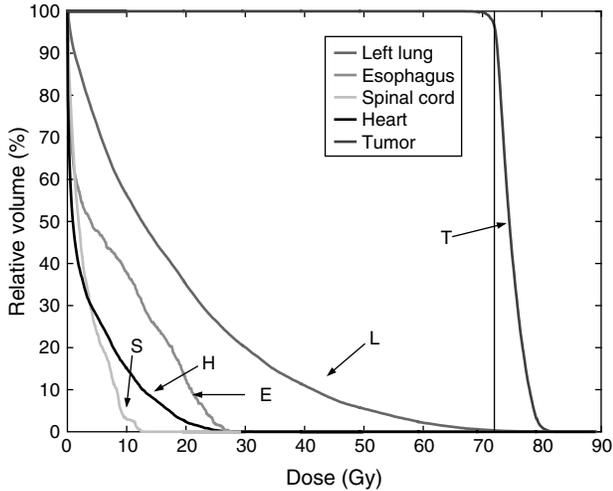
$$\sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} \tilde{p}(x) w_b^* \geq \theta_v \quad \forall v \in \mathcal{T}.$$

A useful way of visualizing three-dimensional dose distributions is through a dose-volume histogram (DVH). A DVH illustrates what fraction of a particular organ of interest receives (at least) a certain level of dose. Figures 1, 2, and 3 show the DVHs corresponding to the

Table 1. Basic information regarding the three formulations.

	Nominal	Robust	Margin
Number of variables	1,625	67,565	1,625
Number of constraints	10,990	65,940	54,950
Solution time (in min.)	3	13	9

Figure 1. Nominal solution DVHs for 78 realized p.d.f.s.



results of delivering the nominal, robust, and margin solutions, respectively, for the different realizations of \tilde{p} . The main organs of interest in this study are the tumor and left lung (in which the tumor resides), whereas secondary organs-at-risk (due to their relative remoteness, in this case) are the esophagus, spinal cord, and heart. These three figures overlay the 78 DVHs (one for each realized p.d.f.) on the same axes, resulting in blurred lines that highlight the range of possible DVHs achievable with these realizations of the uncertain p.d.f. The vertical line at 72 Gy in the figures illustrates the location of the minimum dose requirement for the tumor. We will adopt the terminology of Chu et al. (2005) and refer to these figures as *DVH clouds*. Note that although our DVH clouds are quite narrow, the nominal solution exhibits more “blurriness” in the tumor cloud near 70 Gy and 80 Gy. That is, there is more variability in what fraction of the tumor receives a sufficient dose

Figure 2. Robust solution DVHs for 78 realized p.d.f.s.

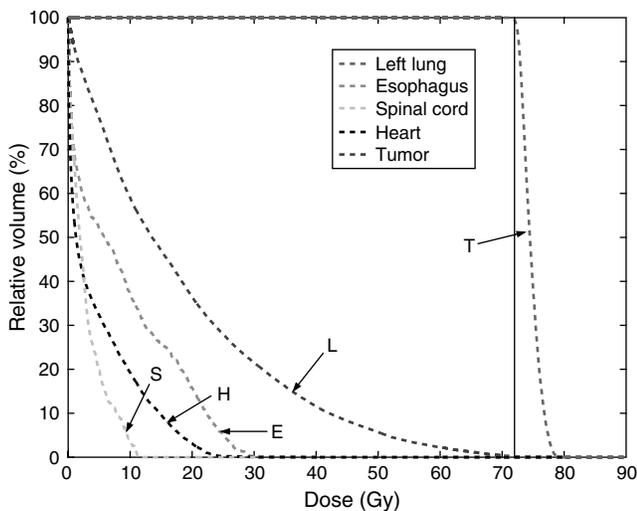
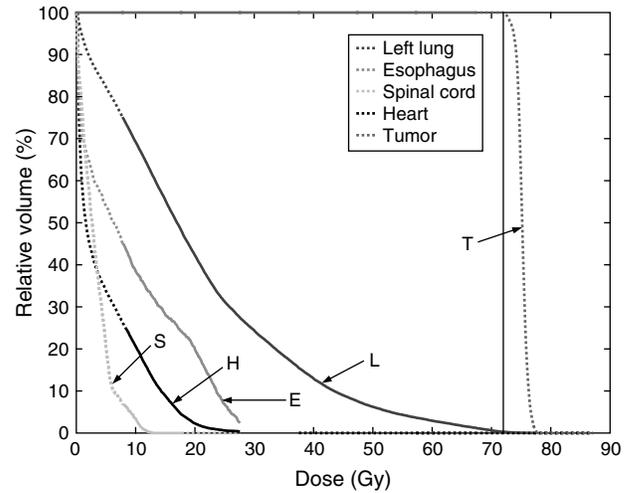
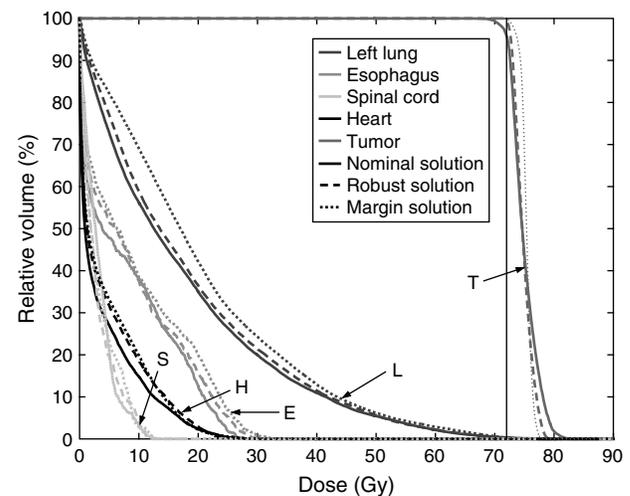


Figure 3. Margin solution DVHs for 78 realized p.d.f.s.



(equivalently, gets underdosed) and what fraction receives the most dose. The fact that the tumor DVH clouds for the robust and margin solutions exhibit sharper falloffs is indicative of their increased robustness to the uncertain p.d.f.s. For comparison, Figure 4 plots the three solutions for one particular realized p.d.f. on the same axes. It can be seen that the nominal and robust solutions typically deliver less dose than the margin solution to the healthy tissue, especially in the left lung, which is the main organ-at-risk here. However, the nominal solution produces an inferior dose distribution in the tumor, increasing both the fraction of the tumor that gets underdosed and overdosed, as compared to the robust and margin solutions. Also, it is interesting to note that although the margin solution delivers the most homogeneous dose to the tumor (as evidenced by the tumor DVH line that is closest to vertical), it is outperformed by the robust solution near the minimum dose requirement. That is, the robust solution exhibits a steeper

Figure 4. The nominal, robust, and margin solution DVHs for one realized p.d.f.



initial dropoff in its DVH line. This is due to the fact that the margin solution enforces tumor coverage in every phase of the breathing cycle, whereas the robust solution provides tumor coverage averaged over the phases (based on the family of p.d.f.s in the uncertainty set).

Numerical data from these three solutions are also summarized in Table 2.

Note that the robust solution essentially ensures that the tumor receives the required dose (only an average 1% underdose), while significantly reducing the dose delivered to the healthy tissue, especially to the left lung where the integral dose delivered is reduced by roughly 11%, as compared to the margin solution. In fact, less than 0.6% of the tumor volume, on average, receives less than 72 Gy, which makes this robust solution clinically acceptable. Although the nominal solution performs the best in terms of dose minimization, it also allows an unacceptable 6% average underdose to over 4% of the tumor volume. Furthermore, the worst-case realization of these 78 trials for the nominal solution results in an 8% underdose (66.30 Gy), whereas the worst-case realization for the robust solution results in only a 1% underdose (71.23 Gy).

The results for the experiment that showed the greatest improvement in tumor coverage of the robust solution over the nominal solution are summarized in Table 3. In this experiment, 276 p.d.f.s were used to generate the uncertainty set, and 52 p.d.f.s were used to test the optimized solutions.

In this experiment, note that the robust solution is able to maintain almost 100% tumor coverage (0.13% underdose to roughly 0.2% of the tumor) on average when tested with the 52 p.d.f.s, whereas the nominal solution underdoses by almost 11% on average (to roughly 6% of the tumor). It is important to recognize that the statistics related to healthy tissue and left lung dosage are very similar to those presented in Table 2. This is true of all the other experiments

Table 2. This table compares the nominal, robust, and margin formulations using the average of the 78 p.d.f.s. This experiment showed the least difference between the nominal and robust solutions in terms of tumor coverage.

	Nominal		Robust		Margin	
	Gy	%	Gy	%	Gy	%
Minimum dose delivered in tumor	67.72	94.06	71.40	99.17	72.04	100.06
Integral dose to left lung	17.40	85.29	18.23	89.36	20.40	100.00
Integral dose to normal tissue	9.04	88.98	9.44	92.91	10.16	100.00

Notes. The left subcolumns represent values in Gy, whereas the right subcolumns are percentages. The first-row percentages are relative to 72 Gy, which is the desired dose level, whereas the next two rows are quoted as a percentage of the dose delivered by the margin solution.

Table 3. This table compares the nominal, robust, and margin formulations using the average of 52 p.d.f.s. This experiment showed the greatest difference between the nominal and robust solutions in terms of tumor coverage.

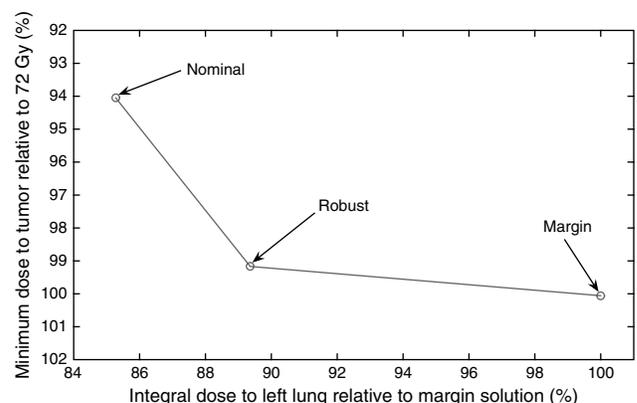
	Nominal		Robust		Margin	
	Gy	%	Gy	%	Gy	%
Minimum dose delivered in tumor	64.26	89.25	71.91	99.87	72.05	100.07
Integral dose to left lung	17.18	85.11	18.02	89.27	20.18	100.00
Integral dose to normal tissue	8.99	88.92	9.39	92.89	10.11	100.00

Notes. The left subcolumns represent values in Gy, whereas the right subcolumns are percentages. The first-row percentages are relative to 72 Gy, which is the desired dose level, whereas the next two rows are quoted as a percentage of the dose delivered by the margin solution.

as well. Overall, these experiments suggest that, unlike nominal solutions, robust solutions can maintain consistently acceptable tumor coverage when tested with different realized p.d.f.s, all other things being equal.

In accordance with the theory we presented, our results suggest that the robust solution combines the dose minimization capability of the nominal solution with the robustness of the margin solution. Viewed another way, the robust formulation is facilitating the trade-off between dose minimization and protection from uncertainty. With these two competing objectives in mind, Figure 5 shows the relevant data points from Table 2 to illustrate this trade-off. In fact, by computing the optimal solution of the robust formulation for various uncertainty sets, we may map out the Pareto surface of robust solutions. In effect, Figure 5 provides a rough bound on this surface using only three data points. By creating an appropriate uncertainty set for the patient at hand, solving the robust formulation will identify the point on the Pareto surface that best trades off dose minimization with uncertainty management.

Figure 5. An approximation of the Pareto surface generated by the robust formulation.



6. Conclusions

In this paper, we develop a robust approach to the management of motion uncertainty in the optimization of intensity-modulated radiation therapy treatments. We considered uncertainty in the probability density function describing motion and incorporated a suitable model of uncertainty into a linear optimization problem. Because the corresponding uncertainty set was a polyhedral set, the resulting robust counterpart remained a linear program. We showed that the robust formulation generalized current mathematical programming approaches to IMRT optimization, and compared the robust solution’s performance to a nominal (aggressive solution based on complete knowledge of the underlying motion) and margin (conservative solution based on minimal knowledge of the underlying motion) solution on a clinical lung case. Our experiments showed that whereas the nominal solution allowed anywhere between an unacceptable 6% to 11% underdose in the tumor, the robust solution consistently provided nearly the same level of protection as the margin solution, allowing around a 1% underdose on average. In addition, the robust solution also managed to consistently lower the dose delivered to the left lung (the main organ-at-risk) by almost 11% as compared to the margin solution. Finally, we positioned our robust approach as a method to rigorously trade off competing objectives, and sketched the associated Pareto surface.

In our numerical implementation, we assumed that variations of the breathing phases as seen in an external surrogate (the RPM signal) were highly correlated with internal breathing variations. As progress is made in image-guided treatments and more real-time data of internal anatomy movements are available, hopefully reliance on such an assumption will diminish. However, even image-guided treatments will not be error-free, so the need for robust methods to incorporate these measurements into the treatment-planning routine will still exist. Although we focused only on motion uncertainty in this paper, there are clearly other sources of uncertainty that can benefit from a similar robust approach.

Finally, we note that our approach essentially assumed that the radiation would be delivered in “one shot” as opposed to over many treatment fractions. By considering optimization on a day-to-day basis, it should be clear that our current “open-loop” policy will benefit from feedback acquired over the course of a treatment. Therefore, robustifying adaptive treatments based on data acquired throughout treatment is a relevant extension of the work presented here.

Appendix. Proofs

PROOF OF PROPOSITION 2. Let λ be the optimal dual variable corresponding to the equality constraint in (7). Then, there exists a \hat{p} that minimizes $\sum_{x \in U} \hat{p}(x)(d_{v,x}(\mathbf{w}) - \lambda)$ subject to the constraints $-\underline{p}(x) \leq \hat{p}(x) \leq \bar{p}(x)$ for all $x \in U$, which is also an optimal solution for the original

problem (7). In such an optimal solution, $\hat{p}(x)$ must be set to its lower (respectively, upper) bound $-\underline{p}(x)$ (respectively, $\bar{p}(x)$), for every x for which $d_{v,x}(\mathbf{w}) > \lambda$ (respectively, $d_{v,x}(\mathbf{w}) < \lambda$). The result follows by examining the value of the objective function at this optimal solution. \square

PROOF OF PROPOSITION 4. If $U = \emptyset$, then $P_U = p$ and the robust formulation is equivalent to the nominal formulation. Next, we will prove the equivalence of the dosimetric margin formulation (13) and the robust formulation (4) with

$$P_X = \left\{ \tilde{p} \in \mathbb{R}^{|X|}: 0 \leq \tilde{p}(x) \leq 1 \ \forall x \in X; \sum_{x \in X} \tilde{p}(x) = 1 \right\}$$

in place of P_U and θ in place of θ_v :

$$\begin{aligned} & \underset{\mathbf{w}}{\text{minimize}} \quad \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p(x) w_b \\ & \text{subject to} \quad \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} \tilde{p}(x) w_b \geq \theta \\ & \quad \quad \quad \forall v \in \mathcal{T}, \ \forall \tilde{p} \in P_X, \\ & \quad \quad \quad w_b \geq 0 \quad \forall b \in \mathcal{B}. \end{aligned} \tag{17}$$

Let \mathbf{w} be a feasible solution to the robust problem (17). Pick an arbitrary $\bar{x} \in X$ and define

$$p^*(x) = \begin{cases} 1 & \text{if } x = \bar{x}, \\ 0 & \text{otherwise.} \end{cases}$$

Then, for any $v \in \mathcal{T}$,

$$\sum_{b \in \mathcal{B}} \Delta_{v,\bar{x},b} w_b = \sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} p^*(x) w_b \geq \theta,$$

which demonstrates that \mathbf{w} is feasible for the dosimetric margin problem (13). For the converse, assume that \mathbf{w} is feasible for (13), and let $\tilde{p} \in P_X$. For any $v \in \mathcal{T}$,

$$\sum_{b \in \mathcal{B}} \sum_{x \in X} \Delta_{v,x,b} \tilde{p}(x) w_b \geq \sum_{x \in X} \theta \tilde{p}(x) = \theta. \quad \square$$

Acknowledgments

The authors thank David Craft and Benjamin Martin for their help with the implementation of this method in treatment-planning software. They also appreciate the discussions with Gregory Sharp about the patient data. This research was supported in part by the National Science Foundation under grant ECS-0312921, the Natural Sciences and Engineering Research Council of Canada, the United States–Israel Binational Science Foundation (BSF) under grant 2003275, and the National Cancer Institute under grant P01-CA21239 and grant R01-CA103904.

References

American Cancer Society. 2006. *Cancer Facts and Figures*. American Cancer Society, Atlanta.

- Baum, C., M. Alber, M. Birkner, F. Nusslin. 2006. Robust treatment planning for intensity modulated radiotherapy of prostate cancer based on coverage probabilities. *Radiotherapy Oncology* **78** 27–35.
- Bortfeld, T. 2006. IMRT: A review and preview. *Phys. Med. Biol.* **51** R363–R379.
- Bortfeld, T., S. B. Jiang, E. Rietzel. 2004. Effects of motion on the total dose distribution. *Seminars in Radiation Oncology* **14**(1) 41–51.
- Bortfeld, T., K. Jokivarsi, M. Goitein, J. Kung, S. B. Jiang. 2002. Effects of intra-fraction motion on IMRT dose delivery: Statistical analysis and simulation. *Phys. Med. Biol.* **47**(13) 2203–2220. <http://stacks.iop.org/0031-9155/47/2203>.
- Brock, K. M., J. M. Balter, L. A. Dawson, M. L. Kessler, C. R. Meyer. 2003. Automated generation of a four-dimensional model of the liver using warping and mutual information. *Med. Phys.* **30**(6) 1128–1133. <http://link.aip.org/link/?MPH/30/1128/1>.
- Chan, T. C. Y., T. Bortfeld, J. N. Tsitsiklis. 2006. A robust approach to IMRT optimization. *Med. Biol.* **51**(10) 2567–2583. <http://stacks.iop.org/0031-9155/51/2567>.
- Chu, M., Y. Zinchenko, S. G. Henderson, M. B. Sharpe. 2005. Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty. *Phys. Med. Biol.* **50**(23) 5463–5477. <http://stacks.iop.org/0031-9155/50/5463>.
- Engelsman, M., G. C. Sharp, T. Bortfeld, R. Onimaru, H. Shirato. 2005. How much margin reduction is possible through gating or breath hold? *Phys. Med. Biol.* **50** 477–490.
- Ferris, M. C., J. Lim, D. M. Shepard. 2003. An optimization approach for radiosurgery treatment planning. *SIAM J. Optim.* **13**(3) 921–937.
- Gierga, D. P., J. Brewer, G. C. Sharp, M. Betke, C. G. Willett, G. T. Y. Chen. 2005. The correlation between internal and external markers for abdominal tumors: Implications for respiratory gating. *Internat. J. Radiation Oncology Biol. Phys.* **61**(5) 1551–1558.
- Holder, A. 2003. Designing radiotherapy plans with elastic constraints and interior point methods. *Health Care Management Sci.* **6**(1) 5–16.
- IMRT Collaborative Working Group. 2001. Intensity-modulated radiotherapy: Current status and issues of interest. *Internat. J. Radiation Oncology Biol. Phys.* **51**(4) 880–914.
- Keall, P. J., G. Starkschall, H. Shukla, K. M. Forster, V. Ortiz, C. W. Stevens, S. S. Vedam, R. George, T. Guerrero, R. Mohan. 2004. Acquiring 4D thoracic CT scans using a multislice helical method. *Phys. Med. Biol.* **49**(10) 2053–2067. <http://stacks.iop.org/0031-9155/49/2053>.
- Keall, P. J., G. S. Mageras, J. M. Balter, R. S. Emery, K. M. Forster, S. B. Jiang, J. M. Kapatoes, D. A. Low, M. J. Murphy, B. R. Murray, C. R. Ramsey, M. B. Van Herk, S. S. Vedam, J. W. Wong, E. Yorke. 2006. The management of respiratory motion in radiation oncology report of AAPM task group 76. *Med. Phys.* **33**(10) 3874–3900. <http://link.aip.org/link/?MPH/33/3874/1>.
- Langen, K. M., D. T. Jones. 2001. Organ motion and its management. *Internat. J. Radiation Oncology Biol. Phys.* **50**(1) 265–278.
- Lee, E. K., T. Fox, I. Crocker. 2003. Integer programming applied to intensity-modulated radiation therapy treatment planning. *Ann. Oper. Res.* **119** 165–181.
- Low, D. A., M. Nystrom, E. Kalinin, P. Parikh, J. F. Dempsey, J. D. Bradley, S. Mutic, S. H. Wahab, T. Islam, G. Christensen, D. G. Polite, B. R. Whiting. 2003. A method for the reconstruction of four-dimensional synchronized CT scans acquired during free breathing. *Med. Phys.* **30**(6) 1254–1263. <http://link.aip.org/link/?MPH/30/1254/1>.
- Lu, W., M.-L. Chen, G. H. Olivera, K. J. Ruchala, T. R. Mackie. 2004. Fast free-form deformable registration via calculus of variations. *Phys. Med. Biol.* **49**(14) 3067–3087. <http://stacks.iop.org/0031-9155/49/3067>.
- Lujan, A. E., E. W. Larsen, J. M. Balter, R. K. Ten Haken. 1999. A method for incorporating organ motion due to breathing into 3D dose calculations. *Med. Phys.* **26**(5) 715–720.
- McInerney, T., D. Terzopoulos. 1996. Deformable models in medical image analysis: A survey. *Med. Image Anal.* **1**(2) 91–108.
- Mell, L. K., A. K. Mehrotra, A. J. Mundt. 2005. Intensity-modulated radiation therapy use in the U.S., 2004. *Cancer* **104**(6) 1296–1303.
- Ólafsson, A., S. J. Wright. 2006. Efficient schemes for robust IMRT treatment planning. *Phys. Med. Biol.* **51**(21) 5621–5642. <http://stacks.iop.org/0031-9155/51/5621>.
- Pan, T., T.-Y. Lee, E. Rietzel, G. T. Y. Chen. 2004. 4D-CT imaging of a volume influenced by respiratory motion on multi-slice CT. *Med. Phys.* **31**(2) 333–340. <http://link.aip.org/link/?MPH/31/333/1>.
- Preciado-Walters, F., R. Rardin, M. Langer, V. Thai. 2004. A coupled column generation, mixed integer approach to optimal planning of intensity modulated radiation therapy for cancer. *Math. Programming B* **101** 319–338.
- Romeijn, H. E., R. K. Ahuja, J. F. Dempsey, A. Kumar, J. G. Li. 2006. A new linear programming approach to radiation therapy treatment planning problems. *Oper. Res.* **54**(2) 201–216.
- Sheng, K., J. Cai, J. Brookeman, J. Molloy, J. Christopher, P. Read. 2006. A computer simulated phantom study of tomotherapy dose optimization based on probability density functions (pdf) and potential errors caused by low reproducibility of pdf. *Med. Phys.* **33**(9) 3321–3326. <http://link.aip.org/link/?MPH/33/3321/1>.
- Theuvs, J. C. M., S. L. S. Kwa, A. C. Wagenaar, Y. Seppenwoolde, L. J. Boersma, E. M. F. Damen, S. H. Muller, P. Baas, J. V. Lebesque. 1998. Prediction of overall pulmonary function loss in relation to the 3-D dose distribution for patients with breast cancer and malignant lymphoma. *Radiotherapy Oncology* **49**(3) 233–243.
- Trofimov, A., E. Rietzel, H.-M. Lu, B. Martin, S. Jiang, G. T. Y. Chen, T. Bortfeld. 2005. Temporo-spatial IMRT optimization: Concepts, implementation and initial results. *Phys. Med. Biol.* **50**(12) 2779–2798. <http://stacks.iop.org/0031-9155/50/2779>.
- Tsunashima, Y., T. Sakae, Y. Shioyama, K. Kagei, T. Terunuma, A. Nohtomi, Y. Akine. 2004. Correlation between the respiratory waveform measured using a respiratory sensor and 3D tumor motion in gated radiotherapy. *Internat. J. Radiation Oncology Biol. Phys.* **60**(3) 951–958.
- Unkelbach, J., U. Oelfke. 2004. Inclusion of organ movements in IMRT treatment planning via inverse planning based on probability distributions. *Phys. Med. Biol.* **49**(17) 4005–4029.
- van Herk, M., P. Remeijer, C. Rasch, J. V. Lebesque. 2000. The probability of correct target dosage: Dose-population histograms for deriving treatment margins in radiotherapy. *Internat. J. Radiation Oncology Biol. Phys.* **47**(4) 1121–1135.
- Vedam, S. S., V. R. Kini, P. J. Keall, V. Ramakrishnan, H. Mostafavi, R. Mohan. 2003. Quantifying the predictability of diaphragm motion during respiration with a noninvasive external marker. *Med. Phys.* **30**(4) 505–513.
- Webb, S. 2003. The physical basis of IMRT and inverse planning. *British J. Radiology* **76** 678–689.
- Zhang, T., R. Jeraj, H. Keller, W. Lu, G. H. Olivera, T. McNutt, T. Mackie, B. Paliwal. 2004. Treatment plan optimization incorporating respiratory motion. *Med. Phys.* **31**(6) 1576–1586.