Multistage Stochastic Fractionated Intensity Modulated Radiation Therapy Planning

Merve Bodur* Mucahit Cevik † Andre A. Cire ‡ Mark Ruschin §
Juyoung Wang*

Abstract

Intensity modulated radiation therapy (IMRT) is a widely used cancer treatment technique designed to target malignant cells. To enhance its effectiveness on tumors and reduce side effects, radiotherapy plans are usually divided into consecutive treatments, or fractions, that are delivered over multiple weeks. However, typical planning approaches have focused on finding the full sequence of radiation intensities prior to the treatment, or were restricted to a single treatment session. In this work, we investigate a fractioned variant of the IMRT planning problem that accounts for geometric motion-related uncertainty during treatment. We propose a novel multistage stochastic programming (MSP) modeling framework that incorporates the sequential decision-making nature of the problem and prevailing stochasticity in cancer treatment. The model is solved via a sample average approximation based on stochastic dual dynamic programming considering a variety of risk measures. We conduct computational experiments on five test cases that are generated based on clinical data. Through extensive simulations, we show that our MSP model generates higher quality treatment plans compared to deterministic and two-stage program counterparts based on multiple performance measures. In particular, our model leads to a higher rate of tumor coverage and a lower rate of radiation exposure for healthy tissues. Accordingly, the proposed MSP framework can greatly contribute to the clinical practice in fractionated IMRT.

Keywords: Intensity modulated radiation therapy (IMRT); multistage stochastic programming; fractionation

1 Introduction

Intensity modulated radiation therapy (IMRT) is a cancer treatment technique that consists of shaping and delivering highly complex therapeutic radiation dose distributions to tumor cells [30]. Although complete tumor control is the primary goal in IMRT, it is also of great importance to minimize radiation exposure to healthy tissues located near the tumor cells, referred to as organs-at-risk (OARs) [2].

In many cases, treatment of a patient involves multiple sessions called *fractions*. Such a practice has several benefits to the treatment outcomes as opposed to a single dose [1, 11, 24]. First,

*Department of Mechanical and Industrial Engineering, University of Toronto, Emails: bodur@mie.utoronto.ca, juyoung.wang@mail.utoronto.ca

†Department of Mechanical and Industrial Engineering, Toronto Metropolitan University, Email: mcevik@ryerson.ca

‡Department of Management, University of Toronto Scarborough & Rotman School of Management, Canada, Email: andre.cire@utoronto.ca

§Department of Radiation Oncology, Sunnybrook Health Sciences Centre, University of Toronto, Ontario, Canada, Email: mark.ruschin@sunnybrook.ca

the radiosensitivity of cells depends on their stage in the cell cycle. During a treatment session, each tumor cell is at a different phase of the cell cycle. Hence, delivering radiation in one single session might not be effective for some groups of tumor cells. Second, cells with high oxygen levels are likely damaged first, causing irregular tumor destruction throughout the treatment sessions. Fractionation provides time for hypoxic cells to improve their oxygen level for subsequent fractions. Third, the time between fractions allows OARs to repair from previously received radiation damage. Since OAR cells have significantly better DNA restoration ability than tumor cells, they benefit from fractionation.

Excessive radiation exposure to healthy tissues can negatively impact patients' health conditions [37]. In general, it is difficult to completely avoid damage to OARs due to multiple factors, such as proximity to organs. Another important factor, which is the focus of this study, is associated with motion uncertainty. More specifically, there exist two primary sources of geometric uncertainties that affect treatment outcomes [23], namely *interfractional motion* and *intrafractional motion*. The former is defined to be the motion seen between the treatment fractions. Such a movement can be caused, e.g., by patient misalignment on the treatment table and patient weight change, and it can be detected based on computational (e.g., computed tomography) images of organs taken on consecutive fractions. The intrafractional motion occurs during the radiation beam delivery. For example, in the case of lung cancer treatment, lung volume changes during the treatment due to respiration activity.

Different strategies have been proposed in the literature to hedge against the detrimental effects of geometric uncertainty [3]. Nonetheless, the standard practice suggested by the International Commission on Radiation Units and Measurement [17] is the *planning target volume* (PTV) approach. The PTV is a geometrical construct designed to ensure that the clinical tumor volume (CTV), the region that needs to be adequately treated, receives the intended prescription dose in the face of uncertainties. To this end, the PTV is a sufficiently large volume that ensures full tumor coverage but may capture some degree of health tissue. While this approach addresses the primary IMRT objective – full tumor exposure to radiation – it may cause significant damage to OARs and tissue surrounding the tumor, e.g., as investigated in Samuelsson et al. [32].

To ensure tumor dose coverage while minimizing damage to healthy tissue, recent studies have attempted to reduce PTV values by explicitly accounting for geometric uncertainty, in particular via stochastic programming and robust optimization. However, such models have been primarily concerned with generating non-adaptive improved robust treatment plans, either for a single treatment session, or to make all the fraction decisions here-and-now. That is, they do not allow any recourse decisions which consider the impact a fraction's decisions has on the subsequent fractions. Moreover, they do not incorporate fraction-wise performance measures, including those related to intermediate cumulative dose amounts a tumor receives during treatment.

In this paper, we propose a multistage stochastic programming (MSP) framework in order to capture the sequential decision-making nature of the fractionated IMRT planning problem under geometric uncertainty. Our contributions are as follows:

- We propose a novel MSP model for the fractionated IMRT treatment planning problem, which allows fraction-wise performance measures in the objective function.
- We apply sample average approximation and solve the obtained model exactly via stochastic dual dynamic programming (SDDP) algorithm [28]. We describe how our model can be reduced to an SDDP and become amenable to state-of-the-art solvers such as the `SDDP.jl` package [14].

- We conduct an extensive computational study using five clinical instances, where we compare the MSP policies traditionally used deterministic policies with a folding horizon framework, considering a variety of risk measures. Our simulations show that the proposed policies significantly outperform the deterministic counterparts in terms of multiple clinically accepted performance metrics.

- We also evaluate the performance of policies obtained from a two-stage stochastic programming approximation to the MSP model. To the best of our knowledge, this approximation constitutes the first proper two-stage stochastic program of the problem in the literature. We also develop an enhanced Benders decomposition algorithm to solve the resulting approximation.

The remainder of the paper is organized as follows. In Section 2, we provide a detailed review of the existing approaches to IMRT planning under uncertainty. In Section 3, we formally define the fractionated IMRT problem and introduce our MSP model, followed by our solution methodology in Section 4. Regarding the computational experiments, we first present the setup in Section 5, and then provide numerical results and obtained insights in Section 6. Lastly, in Section 7, we provide a brief summary of our work and discuss future research directions.

2 Literature review

Reducing negative side effects of radiotherapy caused by geometric uncertainties is a critical task. To this day, the large majority of techniques are based on deterministic PTV approaches [32]. Romeijn et al. [30] proposed a linear programming formulation for the IMRT planning problem, minimizing a linear combination of over- and under-dose penalty functions. Cevik et al. [4] proposed a mixed-integer linear programming model for the radiosurgery treatment planning problem, where different than the model of Romeijn et al. [30], they additionally minimized the total number of isocenters (i.e., target intersection points for the radiation beams) to be used for the treatment session, with a goal of reducing the treatment time.

Saberian et al. [31] developed a deterministic optimization model, where they simultaneously optimized fluence maps (i.e., the resulting beam dose distribution observed in the patient) and number of fractions. Their model maximized the biological effect of average dose over tumor voxels subject to biologically effective dose constraints on normal tissue. They assumed a linear-quadratic tumor proliferation model for tumors during the treatment sessions. Although the model computed beamlet intensity vectors over multiple fractions, it did not assume fraction-wise penalty or different intensity vectors for each fraction. Kishimoto and Yamashita [18] focused on hotspots (i.e., voxels with excessive radiation exposure), and proposed a deterministic successive hotspot elimination procedure for the single-fraction fluence map optimization problem. Their results indicate significantly improved treatment plans in terms of damage to healthy tissues, without noticeable loss in the quality of treatment outcome.

When there are many nearby structures to the target volume, it has been shown that large PTV values can end up causing excessive damage to OARs, or underdosage of the tumor [32]. Moreover, performing treatment over highly radiation-sensitive organs such as brain and spine can result in damage to these OARs and might cause severe side effects. To cope with the disadvantages of large PTV values, and to more properly incorporate geometric uncertainty into treatment planning, several studies have been conducted in the realm of robust optimization and stochastic programming.

A large number of studies employed robust optimization techniques for various radiotherapy problems. Chu et al. [8] developed a robust optimization model for the IMRT problem, considering patient motions over the course of treatment as the main source of uncertainty. While considering the multi-fraction setup, independent random variables were assumed for each fraction. They developed a robust optimization model with an ellipsoidal uncertainty set, which was shown to be equivalent to an LP model, under certain probabilistic assumptions. Although their model optimized fluence map for multiple fractions, they did not consider fraction-wise target dose or fraction-wise intensity vectors. Instead, they assumed one common beamlet intensity vector across the multiple fractions in their model, which simplifies their model to a single-stage robust optimization model. Chan et al. [6] proposed a linear robust optimization model for the single-fraction IMRT problem. They built an interval uncertainty set for intrafractional motion uncertainty, using the probability mass function derived from the histogram obtained via clinical data, which leads to a tractable linear programming reformulation of the robust model. Under the assumption that the possible movements of organs can be completely described by the constructed probability mass function, they showed that their model delivers significantly less radiation to healthy tissues, compared to the deterministic counterpart (so-called nominal) model. Bortfeld et al. [3] extended the work of Chan et al. [6] by considering a more complex margin type, namely the dosimetric margin, as opposed to the geometric margin due to the simplified one-dimensional geometry. Later, Chan and Mišić [5] proposed an adaptive optimization framework for the fractionated IMRT planning problem. They solved the single-fraction robust optimization model of Bortfeld et al. [3] at the beginning of each fraction with a new uncertainty set obtained by using the information collected prior to the treatment session. They showed that improved treatment plans can be obtained compared to the non-adaptive robust optimization approach.

Chan et al. [7] proposed the first single-stage fluence map optimization model which combined robust optimization and average value at risk (AVaR) frameworks. The uncertainty stemmed from breathing motion, which caused the chest to move unpredictably during treatment. They made their problem tractable by applying duality-based reformulation. Motivated by Chan et al. [7]'s model, which has a large number of robust constraints, Mahmoudzadeh et al. [20] developed a decomposition-based solution method for the same problem.

Robust optimization has also been used to incorporate some other types of uncertainties into the treatment planning. For instance, Fredriksson [16] investigated various sources of geometric uncertainties that can occur during the treatment, including systematic range errors and setup errors. Orvehed Hiltunen [26] and Engwall et al. [15] made use of robust optimization to account for the interplay effect, the interference between the tumor motion and the time-structure of the radiation delivery. Lim et al. [19] proposed an adaptive robust optimization framework similar to Chan and Mišić [5] for treatment planning under tumor shrinkage uncertainty, however, they updated the uncertainty set after multiple fractions (35 fractions in their experiments), after a noticeable change is observed in tumor volume. Lastly, Dabadghao and Roy [9] considered radioresistivity, the responsiveness of the tumor to the radiation, depending on cell oxygenation status, as the uncertain factor. Defining a time-dependent uncertainty set, they proposed a multistage robust optimization model for a simplified version of the problem, where they only decide on the average dose delivered to the tumor at the current and future time periods.

Lastly, Ripsman et al. [29] proposed the first robust mixed-integer programming model for direct aperture optimization, where breathing motion was considered as the main source of intrafractional uncertainty. In their model, the authors performed both fluence map optimization and the discovery of the optimal combination of beam angles, apertures, and intensity values of the multi-leaf collimators at the same time. These robust optimization models fundamentally differ from our

modeling framework, as they seek to obtain optimized worst-case treatment outcomes. As opposed to the robust optimization framework, we adopt a stochastic programming approach, where we can consider both average and worst-case performance, or any middle-ground performance measures by adopting different risk measures.

A few other studies employed stochastic programming to account for uncertainties in IMRT. Men et al. [22] introduced a *one-stage* stochastic programming model for the IMRT planning problem under interfractional motion uncertainty, which is solved via sample average approximation. Although the authors motivated the problem by mentioning the fractionated IMRT planning, they developed a mathematical model that primarily focuses on the expected performance of one fraction. More specifically, in the proposed model, the same treatment plan is assumed to be applied at each fraction, i.e., only one set of beamlet intensity decisions is introduced. Moreover, the same set of scenarios are used for each fraction to estimate fraction-based performance criteria, whereas the total effects of the proposed static treatment plan is only captured by using the expected delivered dose for each voxel which is computed by multiplying the beamlet intensities with the expected value of the associated random variables, i.e., under the assumption that the total dose delivered over a treatment can be accurately approximated by the expected dose delivered over all the fractions. On the other hand, Sir et al. [35] proposed a stochastic programming model for the fractionated IMRT planning problem. They applied sample average approximation using the multiple instance geometry approximation (MIGA) proposed by McShan et al. [21], as we also do in this study. However, rather than using an exact solution method, or considering the whole scenario tree, they derived solutions to their model via two heuristics that reduce the model complexity significantly: either replacing the realization of random variables with their nominal value, which yields a deterministic model, or approximating the multistage decision-making problem via a *one-stage* stochastic program, i.e., making the treatment decisions associated with all the fractions here-and-now decisions. Furthermore, they only use performance measures that are solely functions of the total dose received at the end of the treatment, i.e., do not allow incorporating any fraction-wise measures. Thanks to this assumption in addition to the assumption that *all* the uncertainties are revealed at once, they could simplify the underlying stochastic process significantly, ignoring the order of the realized uncertainties. This means that different permutations of a sample path (scenario vector) are treated as the same. This would be a highly limiting assumption in a more realistic setting as ours, since the fraction decisions are indeed the functions of the observed history of the realizations and in practice we have the flexibility to adapt our decisions according to the observed history, that is, make different decisions in the cases where different permutations of the realization history is revealed. Also, the order of the realizations directly impact the fraction-wise performance measures, which we incorporate, thus need to be preserved.

Table 1 provides a high-level summary of the most relevant studies from the literature, and shows the relative positioning of our paper with respect to these studies. We specify whether multiple fractions are considered, which modeling paradigm is employed, and what type of uncertainty is considered, if any. Studies that consider interfractional or intrafractional uncertainty as organ movement more generally, such as this work, are simply referred to as geometric uncertainty in the table.

3 MSP approach to the fractionated IMRT planning problem

In this section, we formally describe the stochastic version of the fractionated IMRT planning problem, present our sequential framework and the method to represent uncertainty, and finally introduce our MSP model.

Table 1: Summary of the most relevant studies in the literature.

Reference	Multi-fraction	Modeling paradigm	Uncertainty
Chu et al. [8]	✓	1RO	geometric uncertainty
Chan et al. [6]	–	1RO	breathing (intra)
Bortfeld et al. [3]	–	1RO	breathing (intra)
Sir et al. [35]	✓	1SP & DO	geometric uncertainty
Men et al. [22]	–	1SP	patient setup (inter)
Chan and Mišić [5]	–	1RO	breathing (intra)
Chan et al. [7]	–	1RO	breathing (intra)
Mahmoudzadeh et al. [20]	–	1RO	breathing (intra)
Saberian et al. [31]	✓	DO	–
Kishimoto and Yamashita [18]	–	DO	–
Cevik et al. [4]	–	DO	–
Lim et al. [19]	–	1RO	geometric uncertainty
Ripsman et al. [29]	–	1RO	breathing (intra)
Our work	✓	MSP & 2SP	geometric uncertainty

RO: Robust Optimization, DO: Deterministic Optimization, SP: Stochastic Programming

M: Multi-stage, 2: Two-stage, 1: One-stage

inter: interfractional uncertainty, intra: intrafractional uncertainty, ✓: considered, –: not considered

3.1 Problem description and MSP framework

The IMRT treatment planning aims to determine a radiation beam distribution to be delivered that best conforms to the tumor shape in the sense that the tumor voxels (volumetric pixels) receive the radiation dose prescribed by clinicians whereas the surrounding tissue voxels have minimal radiation exposure. The planning process usually consists of three phases which respectively determine a set of beam angles, a radiation intensity profile for each beam angle, and a set of aperture shapes and their intensities [39]. In this paper, we focus on the optimization problem associated with the second phase, also known as the *fluence map optimization*. The problem asks for the intensities to be applied to a discretized set of beamlets (fluences) to generate a fluence map, i.e., a dose distribution delivery. The fluence map must observe the treatment goals related to tumor coverage while also reducing dosage to healthy tissues as much as possible.

The fractionated IMRT planning problem extends the IMRT planning problem into multiple-period decision epochs. That is, the goal is to define the intensities for tumors and OARs at each fraction separately, which now must also account for quality-of-treatment associated with the whole treatment plan. For example, the plan is flexible to increase or decrease radiation doses to tissues during any fraction, as they may have received larger or lower radiation than planned, respectively, due to geometric uncertainties.

Incorporating uncertainty leads to a sequential decision-making problem, which we model via MSP framework. In fractionated IMRT planning under geometric uncertainty, specifically in the so-called the offline adaptive radiation therapy framework [35], beamlet intensity decisions are made at the beginning of each fraction, and then the actual delivered doses are observed once all the uncertainties associated with the fraction is revealed, before making the next fraction decisions. As such, decisions are adaptive, e.g., a higher dose amount can be delivered to potentially reach a target dose level at the end of a fraction if an insufficient total amount has been received by the tumor during the fraction due to uncertainty. However, they should be non-anticipative, i.e., can

depend only on the history of observations, not future uncertainty realizations. Then, the goal is to find a policy (i.e., a solution to an MSP model) that is a function mapping the history of decisions and uncertainty observations to a set of dose delivery decisions.

As decisions are made before observing a realization of geometric uncertainty, the sequential decision-making process of the stochastic fractionated IMRT planning falls into a specific class of MSP, namely the decision-hazard framework [13]. The process for F fractions is illustrated in Figure 1.

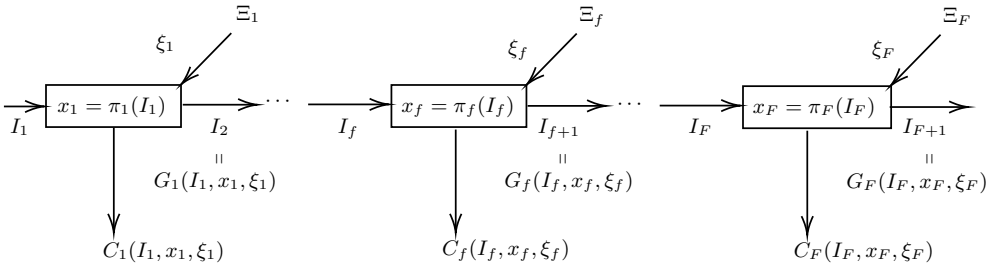


Figure 1: Decision-hazard decision-making framework [13]

At each fraction f, the system state information I_f (e.g., the cumulative amount of dose received) is taken as input, fluence map decisions x_f suggested by policy π_f are implemented, before observing the impact of geometric uncertainties represented by random variable ξ_f. Depending on the realization of ξ_f, from a set of possible realizations Ξ_f, the system state is updated to I_{f+1} via transition function G_f for the sake of the next decision-making step, and the fraction-level performance measure (e.g., underdose and overdose penalty cost) is quantified by function C_f. Next, we present a model for incorporating geometric uncertainty during treatment.

3.2 Uncertainty representation

For each fraction f, we let ξ_f denote the random variable defining the organ status at fraction f, where we represent the vector of random variables associated with fractions up to fraction f (inclusive) by ξ^f, i.e., $\xi^f = (\xi_1, \xi_2, \ldots, \xi_f)$. We denote by $\mathbb{P}$ the probability distribution associated with the stochastic process, i.e., ξ^F.

Considering the fact that tissues can move in a continuous range, a natural approach to represent organ motion is to use a continuous distribution $\mathbb{P}$. However, since this usually yields notoriously difficult optimization problems, a common approach is to use discrete distributions with finite support as an approximation [22, 35], most notably the MIGA, where geometric uncertainties are approximated by multiple instances representing patient's anatomy [21]. Although the MIGA approach and in turn our proposed MSP model can capture geometric uncertainty more generally, i.e., accommodate intrafraction and interfraction uncertainties simultaneously, the challenge would be in obtaining an accurate uncertainty model to generate the scenarios from. Thus, in our numerical analysis, we leverage the well-accepted margin approach for intrafractional uncertainty, and incorporate the scenarios for interfractional uncertainty.

As there is interruption between treatment sessions (e.g., they typically happen in different days or weeks), we can assume that random variables associated with different fractions are independent. This is indeed a common assumption in the literature, e.g., in [1, 40], among others. Then, under the MIGA scheme, for each fraction f, we can generate a finite set of possible realizations, denoted by Ξ_f, for random patient geometry ξ_f, and estimate their probability of occurrence. As illustrated

in Figure 2, the first of the possible outcomes correspond to the nominal patient geometry, whereas the others are relative to the nominal one.

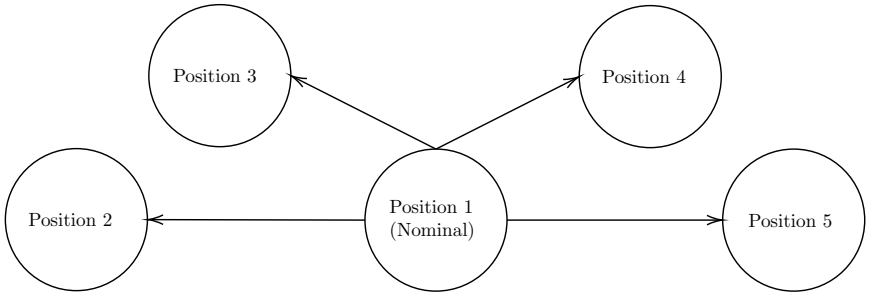


Figure 2: MIGA scheme [21] example

In what follows, we present our mathematical model and solution methodology using the general probability distribution $\mathbb{P}$, whereas we employ the MIGA approach in our numerical study.

Table 2: Notation used in the multistage stochastic fractionated IMRT planning model

Sets:	
$\mathcal{C}$	Set of OARs
$\mathcal{T}$	Set of tumors, $t \in \mathcal{T}$
$\mathcal{B}$	Set of beamlets, $b \in \mathcal{B}$
$\mathcal{L}$	Set of tissues, $\ell \in \mathcal{L} = \mathcal{C} \cup \mathcal{T}$
$\mathcal{V}_\ell$	Set of voxels of $\ell \in \mathcal{L}$, $v \in \mathcal{V}_\ell$ (t and ℓ interchangeably used for tumors)
Parameters:	
Deterministic:	
F	Total number of fractions
$T^+_{\ell v}$	Maximum cumulative dose allowed for voxel $v \in \mathcal{V}_\ell$ of tissue $\ell \in \mathcal{L}$ at the end of all treatment plan
T^-_{tv}	Minimum cumulative dose required for voxel $v \in \mathcal{V}_t$ of tumor $t \in \mathcal{T}$ at the end of all treatment plan
$R^+_{\ell v f}$	Maximum dose allowed for voxel $v \in \mathcal{V}_\ell$ of tissue $\ell \in \mathcal{L}$ at fraction $f \in [F]$
R^-_{tvf}	Minimum dose required for voxel $v \in \mathcal{V}_t$ of tumor $t \in \mathcal{T}$ at fraction $f \in [F]$
α^+_ℓ	Penalty for tissue $\ell \in \mathcal{L}$ unit overdose at the end of all treatment plan
α^-_t	Penalty for tumor $t \in \mathcal{T}$ unit underdose at the end of all treatment plan
$\beta^+_{\ell f}$	Unit overdose penalty for tissue $\ell \in \mathcal{L}$ for fraction $f \in [F]$
β^-_{tf}	Unit underdose penalty for tumor $t \in \mathcal{T}$ for fraction $f \in [F]$
Uncertain:	
$D_{\ell v b}$	Dose received by voxel $v \in \mathcal{V}_\ell$ of tissue $\ell \in \mathcal{L}$ when unit-dose beam is shot through beamlet $b \in \mathcal{B}$
Decision variables:	
x_{bf}	Intensity of beam to shoot through beamlet $b \in \mathcal{B}$ at fraction $f \in [F]$
$z_{\ell v f}$	Dose received by voxel $v \in \mathcal{V}_\ell$ of structure $\ell \in \mathcal{L}$ at the end of fraction $f \in [F]$
$I_{\ell v f}$	Cumulative amount of dose received by the beginning of fraction $f \in [F+1]$
$\gamma^+_{\ell v}$	Overdose amount of voxel $v \in \mathcal{V}_\ell$ of structure $\ell \in \mathcal{L}$ at the end of all treatment plan
γ^-_{tv}	Underdose amount of voxel $v \in \mathcal{V}_t$ of tumor $t \in \mathcal{T}$ at the end of of all treatment plan
$\theta^+_{\ell v f}$	Overdose amount of voxel $v \in \mathcal{V}_\ell$ of structure $\ell \in \mathcal{L}$ at the end of fraction $f \in [F]$
θ^-_{tvf}	Underdose amount of voxel $v \in \mathcal{V}_t$ of tumor $t \in \mathcal{T}$ at the end of fraction $f \in [F]$

3.3 MSP model description

For notational convenience, for any natural number N, we let $[N] = \{1, 2, \ldots, N\}$, where $[0] = \emptyset$. Similarly, for $a, b \in \mathbb{N}$ such that $b \geq a$, we let $[a, b]_{\mathbb{N}} = [b] \setminus [a-1]$. We also define $\mathbb{I}_{\text{condition}}$ as the indicator function which returns value 1 if the condition provided as the subscript is satisfied, and 0 otherwise.

We formulate the fractionated IMRT problem as a multistage stochastic program, which determines beamlet intensities at each fraction given the number of fractions, tumor minimum cumulative dose prescriptions and their desired breakdown to fractions, as well as maximum cumulative and fraction-wise allowed doses for tissues. We provide a detailed description of the model parameters and decision variables in Table 2, where the dependence of uncertain parameters and decision variables to underlying random variables $\{\xi_f\}_{f\in[F]}$ is omitted.

We note that dependence on the history of random variable observations is made explicit by parametrizing the decision variables and uncertain parameters, and letting for convenience $\xi_0 = 1$ (meaning that any variable that depends on ξ_0 is deterministic). Accordingly, an MSP model for fractionated IMRT treatment planning can be obtained as follows.

$$\min \rho\Bigg(\sum_{f\in[F]} \sum_{\ell\in\mathcal{L}} \sum_{v\in\mathcal{V}_\ell} \frac{\beta^+_{\ell f}}{|\mathcal{V}_\ell|} \theta^+_{\ell v f}(\xi^f) + \sum_{f\in[F]} \sum_{t\in\mathcal{T}} \sum_{v\in\mathcal{V}_t} \frac{\beta^-_{tf}}{|\mathcal{V}_t|} \theta^-_{tvf}(\xi^f) + \sum_{\ell\in\mathcal{L}} \sum_{v\in\mathcal{V}_\ell} \frac{\alpha^+_\ell}{|\mathcal{V}_\ell|} \gamma^+_{\ell v}(\xi^F) + \sum_{t\in\mathcal{T}} \sum_{v\in\mathcal{V}_t} \frac{\alpha^-_t}{|\mathcal{V}_t|} \gamma^-_{tv}(\xi^F) \Bigg)$$

s.t. $\mathbb{P}$-a.s.:

$$\begin{aligned}
& z_{\ell v f}(\xi^f) = \sum_{b\in\mathcal{B}} D_{\ell v b}(\xi^f) x_{bf}(\xi^{f-1}) && \forall\, \ell \in \mathcal{L}, v \in \mathcal{V}_\ell, f \in [F] \\
& I_{\ell v f}(\xi^{f-1}) = I_{\ell v, f-1}(\xi^{f-2}) + z_{\ell v, f-1}(\xi^{f-1}) && \forall\, \ell \in \mathcal{L}, v \in \mathcal{V}_\ell, f \in [2, F+1] \\
& \theta^+_{\ell v f}(\xi^f) \geq O_\Pi(z_{\ell v f}(\xi^f) - R^+_{\ell v f}) && \forall\, \ell \in \mathcal{L}, v \in \mathcal{V}_\ell, f \in [F] \\
& \theta^-_{tvf}(\xi^f) \geq O_\Pi(R^-_{tvf} - z_{tvf}(\xi^f)) && \forall\, t \in \mathcal{T}, v \in \mathcal{V}_t, f \in [F] \\
& \gamma^+_{\ell v}(\xi^F) \geq O_\Pi(I_{\ell v, F+1}(\xi^F) - T^+_{\ell v}) && \forall\, \ell \in \mathcal{L}, v \in \mathcal{V}_\ell \\
& \gamma^-_{tv}(\xi^F) \geq O_\Pi(T^-_{tv} - I_{tv, F+1}(\xi^F)) && \forall\, t \in \mathcal{T}, v \in \mathcal{V}_t \\
& x_{bf}(\xi^{f-1}) \geq 0 && \forall\, b \in \mathcal{B}, f \in [F] \\
& z_{\ell v f}(\xi^f) \geq 0 && \forall\, \ell \in \mathcal{L}, v \in \mathcal{V}_\ell, f \in [F] \\
& I_{\ell v, 1}(\xi^0) = 0 && \forall\, \ell \in \mathcal{L}, v \in \mathcal{V}_\ell \\
& I_{\ell v f}(\xi^{f-1}) \geq 0 && \forall\, \ell \in \mathcal{L}, v \in \mathcal{V}_\ell, f \in [2, F+1] \\
& \theta^+_{\ell v f}(\xi^f) \geq 0 && \forall\, \ell \in \mathcal{L}, v \in \mathcal{V}_\ell, f \in [F] \\
& \theta^-_{tvf}(\xi^f) \geq 0 && \forall\, t \in \mathcal{T}, v \in \mathcal{V}_t, f \in [F] \\
& \gamma^+_{\ell v}(\xi^F) \geq 0 && \forall\, \ell \in \mathcal{L}, v \in \mathcal{V}_\ell \\
& \gamma^-_{tv}(\xi^F) \geq 0 && \forall\, t \in \mathcal{T}, v \in \mathcal{V}_t
\end{aligned}$$

The objective function minimizes a risk measure $\rho(\cdot)$ of the averaged penalties for tissue overdose and tumor underdose, both at each fraction and at the end of all treatment plan. All the constraints need to be satisfied almost surely with respect to the joint probability distribution $\mathbb{P}$. The first set of constraints defines $z_{\ell vf}(\xi^f)$ variables that quantify the dose actually delivered to each voxel at a fraction, observed after the uncertainty associated with the fraction (and accordingly the dose deposition values $D_{\ell vb}(\xi^f)$) is revealed, thus are functions of the history ξ^f. On the other hand, note that due to the decision-hazard framework, the beamlet intensity decisions $x_{bf}(\xi^{f-1})$ need to be made before the uncertainty is observed, as such depend on ξ^{f-1} rather than ξ^f. The second set of constraints defines the state variables which store the cumulative amount of dose delivered to each tissue, where we use zero as the initial (before the patient starts any treatment) state values (i.e., for the $I_{\ell v,1}$ variables). Third to sixth constraint sets quantify penalty terms for fraction-wise overdosage of tissues, fraction-wise underdosage of tumors, end-of-horizon overdosage of tissues, and end-of-horizon underdosage of tumors, respectively. The function $O_\Pi(\cdot)$ defining these penalties is a monotonically increasing convex function in $\mathbb{R}_+$ that can be selected based on domain expertise. The remaining sets of constraints provide domains of the decision variables, which are random variables themselves due to geometric uncertainty.

Our MSP model has relatively complete recourse since at any fraction, given any feasible set of decisions to previous fractions and history of random variable realizations, we can make feasible decisions (e.g., by setting all beamlet intensity and the received dose amount variables to zero, keeping the state variable values the same as in the previous fraction, and setting the overdose and underdose amounts accordingly).

In order to solve the model efficiently, we choose to keep it linear, by assuming $O_\Pi(x) = x$, i.e., an identity function. We note that by using the expected value as the risk measure, taking $O_\Pi(x) = x^2$, $F = 1$, $\alpha_\ell^+ = \alpha_t^- = 0$, and imposing random variables associated with tissues of choice to be replaced by their average value, we obtain the hybrid stochastic programming model proposed by Men et al. [22]. For the objective function, in our numerical study, we experiment with five coherent nested[1] risk measures, namely,

(1) $\rho_E(\cdot) := \mathbb{E}[\cdot]$, expected value,

(2) $\rho_A(\cdot) := \mathsf{AV@R}_{0.8}(\cdot)$, average value-at-risk,

(3) $\rho_W(\cdot) := \mathsf{AV@R}_{1.0}(\cdot)$, worst case,

(4) $\rho_{EA}(\cdot) := 0.5 \cdot \rho_E(\cdot) + 0.5 \cdot \rho_A(\cdot)$,

(5) $\rho_{EW}(\cdot) := 0.5 \cdot \rho_E(\cdot) + 0.5 \cdot \rho_W(\cdot)$,

which are all evaluated based on a discrete approximation of the probability distribution $\mathbb{P}$ in the form of a scenario tree. Throughout the paper, the model with $\rho = \rho_E$ is referred to as the risk-neutral model, and when any other risk measure is considered, the model is called risk-averse.

4 Solution methodology

MSP models are known to be notoriously difficult, as such they are typically solved by means of approximations. In that regard, a variety of decision-rule and scenario tree-based methodolo-

[1] A nested risk measure is defined in [12] as one that takes the risk of the scenario tree according to a recursive definition, and is usually preferred, over the alternative so-called end-of-horizon risk measures, as a modeling choice because it leads to computationally efficient algorithms. Please see our nested MSP formulation in Section 4.1 that clearly demonstrates the use of such a risk measure.

gies have been proposed. The former approaches restrict the form of the MSP policy (i.e., $\pi_f(\cdot)$ functions) in Figure 1, whereas the latter approaches approximate the true probability distribution, $\mathbb{P}$, via a scenario tree. The commonly used MIGA scheme yields a scenario-tree model of the geometric uncertainty. However, the size of the scenario tree increases exponentially in the number of fractions, the well-known phenomenon called the curse of dimensionality, making even scenario tree-based approximations of MSP models intractable. On the other hand, some classes of MSP models can be solved efficiently under the stage-wise independence assumption, i.e., when the random variables associated with different decision-making stages are independent. We note that this is a natural assumption for the fractionated IMRT planning problem, under which the linear version of our MSP model can be solved efficiently via the SDDP algorithm. In what follows, we explain the adaptation of SDDP to our model, and then how our model can be transformed to become amenable to employ a state-of-the-art implementation of the SDDP algorithm.

4.1 SDDP algorithm

SDDP, introduced by Pereira and Pinto [28], is regarded as the most successful solution technique for linear MSP models (with right-hand-side uncertainty and stage-wise independence), having been employed in a wide range of practical applications. Under the stage-wise independence assumption, it overcomes the difficulty arising from the exponential growth of scenario tree size with increase of the number of decision-making stages. The SDDP algorithm is a nested decomposition algorithm where scenarios are sampled from the scenario tree. Under mild sampling assumptions, it converges to an optimal policy almost surely in a finite number of iterations [33].

Before we describe how SDDP algorithm works, in particular on our MSP model, we transform our model into a nested form, bringing it into the dynamic programming framework. In that regard, we follow the system dynamics illustrated in Figure 1 as well as the notation used therein and in the original MSP formulation. The proposed MSP model for the fractionated IMRT problem can be written in a nested form as follows (given the initial state of the system described by $(I_{\ell v 1})_{\ell \in \mathcal{L}, v \in \mathcal{V}_\ell}$):

$$\min_{x_1 \in \mathbb{R}_+^{|\mathcal{B}|}} \rho_{\xi_1} \Bigg(C_1(I_1, x_1, \xi_1) + \min_{x_2 \in \mathbb{R}_+^{|\mathcal{B}|}} \rho_{\xi_2} \Bigg(C_2(I_2, x_2, \xi_2) + \min_{x_3 \in \mathbb{R}_+^{|\mathcal{B}|}} \rho_{\xi_3} \Bigg(\cdots + \min_{x_F \in \mathbb{R}_+^{|\mathcal{B}|}} \rho_{\xi_F} \Big(C_F(I_F, x_F, \xi_F) \Big) \Bigg) \Bigg) \Bigg)$$

where we added the random vector associated with each fraction as a subscript to the risk measure. Note that it is sufficient to use the random vector ξ_f for fraction f rather than the set of random vectors associated with the fractions up to f, i.e., ξ^f, thanks to the stage-wise independence

property. The fraction-wise cost functions, for all $f \in [F]$, are defined as follows:

$$\begin{aligned} C_f(I_f, x_f, \xi_f) = &\sum_{\ell \in \mathcal{L}} \sum_{v \in \mathcal{V}_\ell} \frac{\beta^+_{\ell f}}{|\mathcal{V}_\ell|} \max\left\{0, \sum_{b \in \mathcal{B}} D_{\ell v b}(\xi_f) x_{bf} - R^+_{\ell v f}\right\} + \\ &\sum_{t \in \mathcal{T}} \sum_{v \in \mathcal{V}_t} \frac{\beta^-_{tf}}{|\mathcal{V}_t|} \max\left\{0, R^-_{tvf} - \sum_{b \in \mathcal{B}} D_{\ell v b}(\xi_f) x_{bf}\right\} + \\ &\mathbb{I}_{f=F} \cdot \sum_{\ell \in \mathcal{L}} \sum_{v \in \mathcal{V}_\ell} \frac{\alpha^+_\ell}{|\mathcal{V}_\ell|} \max\left\{0, I_{\ell v f} + \sum_{b \in \mathcal{B}} D_{\ell v b}(\xi_f) x_{bf} - T^+_{\ell v}\right\} + \\ &\mathbb{I}_{f=F} \cdot \sum_{t \in \mathcal{T}} \sum_{v \in \mathcal{V}_t} \frac{\alpha^-_t}{|\mathcal{V}_t|} \max\left\{0, T^-_{tv} - I_{\ell v f} + \sum_{b \in \mathcal{B}} D_{\ell v b}(\xi_f) x_{bf}\right\}, \end{aligned}$$

the last of which, for fraction F, includes the end-of-treatment cumulative dose, overdose and underdose penalties. We can now rewrite the nested formulation as the dynamic programming formulation as

$$\min_{x_1 \in \mathbb{R}_+^{|\mathcal{B}|}} \mathcal{Q}_1(I_1, x_1)$$

by recursively defining the cost-to-go functions (also known as the value functions) through

$$\mathcal{Q}_f(I_f, x_f) = \rho_{\xi_f}\left(Q_f(I_f, x_f, \xi_f)\right) \qquad \forall\, f \in [F]$$

where, for all $f \in [F-1]$,

$$Q_f(I_f, x_f, \xi_f) = C_f(I_f, x_f, \xi_f) + \min_{x_{f+1} \in \mathbb{R}_+^{|\mathcal{B}|}} \mathcal{Q}_{f+1}(G_f(I_f, x_f, \xi_f), x_{f+1})$$

with $G_f(I_f, x_f, \xi_f)$ representing the cumulative dose delivered to each voxel at the end of fraction f, i.e., I_{f+1}.

Given the dynamic programming formulation of our proposed MSP model, we now discuss how such a problem can be iteratively solved via SDDP. We provide a graphical illustration of the algorithm in Figure 3. The algorithm works in a backward and forward manner. It builds and sequentially refines outer approximations of the (convex) cost-to-go functions, $\bar{\mathcal{Q}}_f, f \in [F]$, by generating subgradient-based cutting planes in the backward steps, using the information gathered in the forward steps.

More specifically, in the forward pass, a set of scenarios are randomly sampled from the scenario tree, and for each selected scenario $\bar{\xi}^F$, associated subproblems are solved in a sequential manner, for $f = 1, 2, \ldots, F$, to obtain optimal candidate beamlet intensity decisions as

$$\bar{x}_1 \in \arg\min_{x_1 \in \mathbb{R}_+^{|\mathcal{B}|}} \bar{\mathcal{Q}}_1(\bar{I}_1, x_1), \tag{1a}$$

$$\bar{x}_{f+1} \in \arg\min_{x_{f+1} \in \mathbb{R}_+^{|\mathcal{B}|}} \bar{\mathcal{Q}}_{f+1}(G_f(\bar{I}_f, \bar{x}_f, \bar{\xi}_f), x_{f+1}) \quad \forall\, f \in [F-1], \tag{1b}$$

along with optimal candidate cumulative dose (state) decisions

$$\begin{aligned} &\bar{I}_1 = 0 \text{ (initialization)}, \\ &\bar{I}_f = G_f(\bar{I}_{f-1}, \bar{x}_{f-1}, \bar{\xi}_{f-1}) \quad \forall\, f \in [2, F+1]. \end{aligned}$$

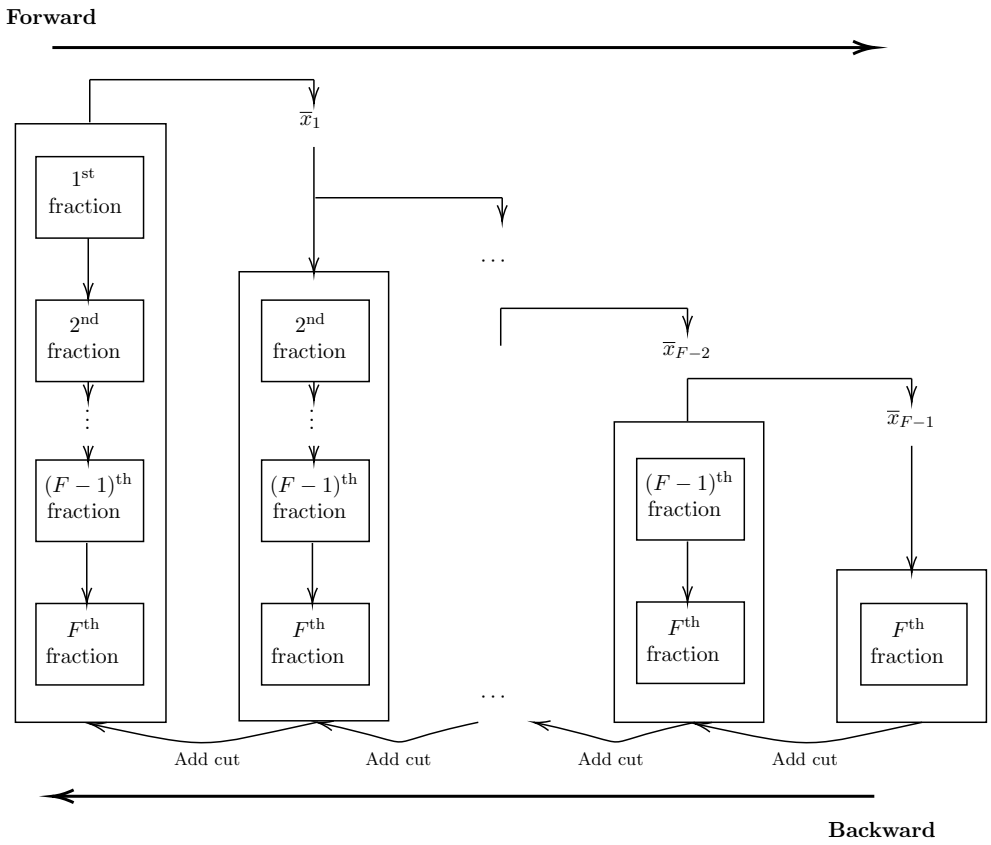


Figure 3: SDDP algorithm applied to fractionated IMRT [28]

In the backward pass, given the sequence of candidate decisions obtained during each forward step, the algorithm uses $\bar{x}_f$ to generate a cutting plane to refine the description of the cost-to-go function approximation $\bar{\mathcal{Q}}_f$. Such an approximation and its combination with the other forward information $\bar{x}_{f-1}$ is again used to generate a cutting plane to refine the description of $\bar{\mathcal{Q}}_{f-1}$. This process is repeated until the outer approximation of $\bar{\mathcal{Q}}_1$ is updated. Using the convexity of the cost-to-go functions, cuts are generated based on their subgradient information at candidate solutions, yielding a piece-wise linear approximation.

The forward and backward procedures are performed in a cyclic manner until the problem is *sufficiently* well-approximated according to convergence criteria. In that regard, the optimality gap measure is commonly used for convergence. Thanks to the outer approximations of the cost-to-go functions, a valid lower bound on the optimal value of the MSP model is obtained as a by-product at the end of every backward iteration. On the other hand, since only a subset of all possible scenarios are evaluated in a forward pass, only statistical valid upper bounds, i.e., those in the form a confidence interval, are obtained in the course of the algorithm. For further details of the SDDP algorithm and enhancement strategies, we kindly refer readers to [12, 33, 34].

4.2 Adaptation to hazard-decision framework

Our proposed MSP model for the fractionated IMRT planning problem fits into the decision-hazard framework, as illustrated in Figure 1, where radiation delivery decisions are made before observing realizations of the random variables associated with the patient geometry during the treatment.

In other words, decisions at a fraction should be non-anticipative with respect to the randomness occurring in that fraction. However, as also mentioned in [38], traditional computational tools use SDDP under another modeling framework known as hazard-decision. Unlike the decision-hazard framework, the hazard-decision framework assumes that decision-makers observe the realizations of random variables corresponding to the current decision stage before making their decisions. Therefore, it is not appropriate to use the hazard-decision framework to model the fractionated IMRT problem since the use of the hazard-decision scheme is essentially equivalent to solving a one-step-ahead anticipative problem. Figure 4 illustrates how the dynamics of a fraction would have been under the hazard-decision framework; we refer to such a building block as a *node*.

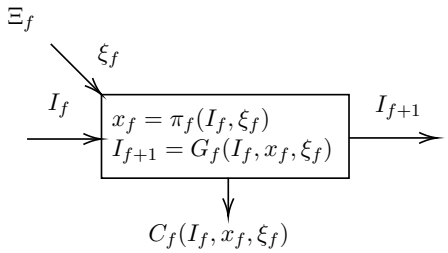


Figure 4: A node in hazard-decision decision-making framework [13]

More precisely, in a hazard-decision node both the cost $C_f(\cdot)$ and transition functions $G_f(\cdot)$ can be fully determined in terms of the state I_f, the control x_f, and the current-stage realization ξ_f. That is, there is sufficient information to calculate the immediate cost and state transition observed when applying x_f. In contrast, the decision-hazard nature of our model implies that, to calculate the cost and state transition associated with the intensity x_f applied at a fraction f, one must necessarily *wait* until ξ_f is observed. That is, the cost (and subsequent state) resulting from a control x_f is not known until the uncertainty is realized.

In order to be able to use a state-of-the-art SDDP implementation, we transform our decision-hazard-based MSP model into an hazard-decision-based MSP model. This necessitates modifying the fraction-level subproblems. We can achieve this by splitting each decision-hazard node into two hazard-decision nodes as shown in Figure 5. At the first hazard-decision node, beamlet intensity

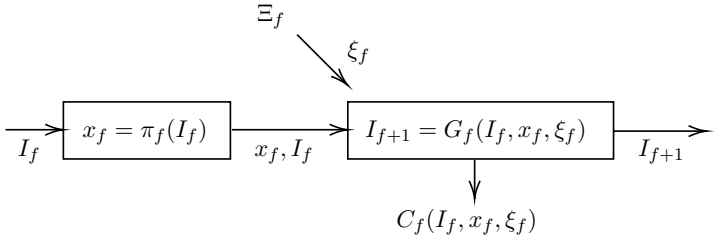


Figure 5: Two hazard-decision-nodes representing a decision-hazard-node [38]

decision variables are optimized in a deterministic manner, i.e., there is no associated uncertainty. Then, in the second hazard-decision node, the cumulative dose variables are updated, according to the previously made decisions and the newly revealed patient geometry observation. In terms of mathematical modeling, such a transformation can be achieved by reformulating the one-stage subproblems in the SDDP framework, given in equations (1), into two-stage stochastic programs. As the obtained subproblems would be potentially computationally difficult to solve, an alternative strategy of state space expansion can be applied in order to obtain deterministic subproblems. We

next explain these two strategies, which have been also previously applied to a multistage stochastic hydrothermal scheduling problem by Street et al. [38].

4.2.1 Two-stage stochastic programming reformulation of subproblems

Below, we provide the two-stage stochastic programming model associated with fraction $f \in [F]$ in our hazard-decision-based MSP model's dynamic programming formulation, where the cost-to-go function approximation for the subsequent stage is represented by $\tilde{Q}_{f+1}(\cdot)$. Recall from Section 3.2 that Ξ_f is the set of possible uncertainty realizations associated with fraction f. Then, given the cumulative doses delivered by the beginning of fraction f, $\tilde{I}_f$, the two-stage model is obtained as follows:

$$\min \ \rho_{\xi_f}\left(C_f\left(\tilde{I}_f, x_f^{\texttt{First}}, \xi_f\right) + \tilde{\mathcal{Q}}_{f+1}\left(I_{f+1}^{\texttt{Second}}(\xi_f)\right)\right) \tag{2a}$$

$$\text{s.t. } I_{f+1}^{\texttt{Second}}(\xi_f) = G_f\left(\tilde{I}_f, x_f^{\texttt{First}}, \xi_f\right), \qquad \forall\, \xi_f \in \Xi_f \tag{2b}$$

$$x_f^{\texttt{First}} \in \mathbb{R}_+^{|\mathcal{B}|}. \tag{2c}$$

The first-stage decisions are the beamlet intensities of the considered fraction f, denoted by $x^{\texttt{First}}$. The initial state variables of the subsequent stage $f+1$, i.e., cumulative doses delivered denoted by $I_{f+1}^{\texttt{Second}}(\xi_f)$, are the second-stage variables since they can be obtained only after the uncertainty associated with fraction f is observed. Lastly, note that the cost-to-go approximation is a function of those state variables.

4.2.2 Deterministic reformulation of subproblems via state space expansion

The subproblems in the form of two-stage stochastic program would be computationally expensive to solve in the forward pass of the SDDP algorithm. Moreover, such subproblems prevent scenario decomposition in the backward pass of the SDDP algorithm. In order to overcome these difficulties and in turn improve the efficiency of the algorithm, we follow the alternative approach employed in [38].

The idea is to make the $x_f^{\texttt{First}}$ variables that appear in the two-stage model (2) state variables, that are henceforth denoted by $\tilde{x}_f$, accordingly add them as input to the approximate cost-to-go functions, i.e., use

$$\tilde{\mathcal{Q}}_f(\tilde{I}_f, \tilde{x}_f) = \rho_{\xi_f}\left(\tilde{Q}_f(\tilde{I}_f, \tilde{x}_f, \xi_f)\right),$$

where $\tilde{Q}_f(\tilde{I}_f, \tilde{x}_f, \xi_f)$ is the optimal objective value of the following deterministic subproblem corresponding to a single realization ξ_f:

$$\begin{aligned} \min \ & C_f(\tilde{I}_f, \tilde{x}_f, \xi_f) + \tilde{\mathcal{Q}}_{f+1}(I_{f+1}, x_f^{\texttt{next}}) \\ \text{s.t. } & I_{f+1} = G(\tilde{I}_f, \tilde{x}_f, \xi_f), \\ & x_f^{\texttt{next}} \in \mathbb{R}_+^{|\mathcal{B}|}. \end{aligned}$$

Since the beamlet intensities have become state variables, a new local copy of such variables, $x_f^{\texttt{next}}$, are added to the subproblem to represent the predetermined subsequent stage decisions. Note that they are only linked to the approximate cost-to-go function, as such to the next fraction.

5 Experimental setup

In this section, we explain our experimental design, baseline parameters, and describe the standard metrics used for the evaluation of the models and proposed treatments.

Target volume generation and uncertainty We extend the setup by Men et al. [22] for determining the PTV value and the interfractional uncertainties, as our MSP model generalizes their one-stage stochastic program. Specifically, previous deterministic works consider a fixed conservative PTV value of 4 mm to hedge against both intrafractional and interfractional motion uncertainties. In our numerical analysis, similar to [22], we consider cases with a lower margin of 2 and 3 mm around the CTV, referred to as CTV+, since the proposed formulations represent interfractional motion uncertainties explicitly via the MIGA approach (i.e., in the form of scenarios in our models, as explained in Section 3.2). Further, we consider five fractions and a branching factor of 20 when generating the scenario tree via the MIGA approach (i.e., $|\Xi_f| = 20$ for all $f = 1, 2, \ldots, 5$), resulting in 3.2 million scenarios in total.

Random voxel sampling Given that the models become difficult to solve in the presence of tissues with a large number of voxels (e.g., due to the increase in the number of variables), we incorporate the same strategy as [22] and employ a random voxel downsampling. More precisely, we sample voxels uniformly at random without replacement for each tumor or OAR for constructing our models. However, we use the original number of voxels for the deterministic counterpart as the underlying models remain computationally tractable. The performance evaluation (via simulations) also considers the original number for all deterministic and stochastic models. We also include a sensitivity analysis of the impact of distinct sample sizes in Section 6.

Downsampling the voxels to achieve faster runtimes has been frequently adopted in previous studies. For instance, in a recent study, Sjölund et al. [36] employed a downsampling strategy to select a small number of voxels to be used in their linear programming model for radiosurgery treatment planning. We also refer to the work of Oreshkin and Arbel [25] for alternate random voxel selection strategies.

Policy simulation We employ the folding-horizon framework to evaluate the treatment plans suggested by the tested models. The procedure first samples scenarios from the full scenario tree, and simulates the policy for each of the sampled scenarios to obtain statistical bounds on the quality metrics discussed below. More specifically, consider a scenario $\hat{\xi}^F$ and the notation provided in Section 4.1. Starting with the initial cumulative dose values $I_1 = 0$, we first implement the beamlet intensity decisions x_1 suggested by the policy, and calculate the cumulative dose values I_2 and the observation $\hat{\xi}_1$. Then, we fold the horizon and apply the same procedure for the problem with $F - 1$ fractions, i.e., starting with I_2, we implement the suggested x_2 and calculate I_3. We repeat the same procedure until the end of fraction F. The MSP policies obtained from the SDDP algorithm are simulated directly via the associated functions in the `SDDP.jl` package. We consider 200 sample scenarios for each tested case.

Implementation details For all optimization problems, we use the Julia language v1.0.5 with JuMP v0.20.1 and Gurobi tunnel v0.7.2 (i.e., the Julia API for Gurobi solver) packages, linked with Gurobi solver v9.0.0. We perform all the experiment on a Debian-OS machine with Intel(R) Core (TM) i9-9900K CPU 3.60GHz, 128 GB of RAM, and a single thread. Our analysis also impose a limit on the total solution time (i.e., *runtime*) of the SDDP approach. We consider a baseline limit of 12 hours, but also devote a section below to evaluate the impact of smaller and larger runtime limits. We note that the SDDP time is largely dominated by the cutting plane generation. Once the cuts are generated, simulating a policy takes negligible time for all cases (20-30 seconds per scenario), as they are tractable linear programs.

We use a state-of-the-art implementation of the SDDP algorithm, `SDDP.jl`, to derive solutions to our MSP model. `SDDP.jl`[2] is a Julia package developed by Dowson and Kapelevich [14]. The library receives the decision stage-level subproblems of an MSP model. Using the subproblem information, it iteratively executes the forward and backward steps of the SDDP algorithm without requiring their explicit implementation from users. The package supports a variety of risk measures, such as $\mathbb{E}[\cdot]$, $\mathsf{AV@R}_\alpha(\cdot)$, and their convex combinations. As `SDDP.jl` operates with the hazard-decision framework, we pass it the deterministic extended reformulations of the subproblems described in Section 4.2.2.

Data We considered five clinical instances whose characteristics are provided in Table 3. The cases are synthetically generated with the specifications largely overlapping with prostate cancer instances provided in [22]. The last row of the table, "Std. for tumor shift (mm)," refers to the standard deviation used to simulate the movement of organs according to the MIGA scheme. As in [22], we obtained the organ shifting vector from a multivariate Gaussian distribution with 0 mean vector and $\sigma\mathcal{I}$ standard deviation, where $\mathcal{I}$ is the identity matrix. The original number of voxels of each tissue is also provided in Table 3 with the number of downsampled voxels in parenthesis. Note that the number of voxels for OAR_2 (last row) was small and no downsampling was applied. Lastly, the MSP model parameters are given in Table 4. The same parameter value is used for all the indices that are not explicitly stated.

Table 3: Clinical case characteristics (Case 5 does not have a nearby OAR)

	Case 1	**Case 2**	**Case 3**	**Case 4**	**Case 5**
Number of OARs	2	2	2	2	0
Number of tumors	1	1	1	1	1
Target volume (cm^3)	4.52	1.35	2.60	2.52	3.45
CTV+ margin size (mm)	4	2	2	2	3
PTV margin size (mm)	8	4	4	4	6
Std. for tumor shift (mm)	5	3	3	3	4
Number of beamlets	576	456	480	792	384
Number of voxels for CTV+	4814 (160)	3423 (228)	3916 (130)	3961 (132)	6472 (431)
Number of voxels for OAR_1	9682 (193)	5750 (115)	8035 (160)	7622 (127)	-
Number of voxels for OAR_2	386 (386)	94 (94)	73 (73)	21 (21)	-

Evaluation criteria We report on the following two standard metrics to evaluate the quality of the simulated treatment plans:

- *Coverage*: It is the percent ratio between the number of tumor voxels that received more than the minimum prescribed dose and the total number of tumor voxels. Thus, it indicates the degree of tumor incineration. We denote this metric by $\mathbb{C}$.
- *Hotspot*: It is the percent ratio between the number of tissue voxels receiving more than 110% of the maximum prescribed dose and the total number of tissue voxels. Thus, it indicates the degree of excessive damage received by tissues. We use $\mathbb{H}_\ell$ to denote the hotspot of tissue $\ell \in \mathcal{L}$.

[2] https://github.com/odow/SDDP.jl

Table 4: MSP model parameters

	Case 1	**Case 2**	**Case 3**	**Case 4**	**Case 5**
F	5	5	5	5	5
$\alpha^{-}_{\text{Tumor}}$	16	16	14	26	14
$\alpha^{+}_{\text{Tumor}}$	2	2	2	2	2
$\alpha^{+}_{\text{OAR\_1}}$	1	1	1	1	
$\alpha^{+}_{\text{OAR\_2}}$	1	1	1	1	
$\beta^{-}_{\text{Tumor},f}$	16	16	14	26	14
$\beta^{+}_{\text{Tumor},f}$	2	2	2	2	2
$\beta^{+}_{\text{OAR\_1},f}$	1	1	1	1	
$\beta^{+}_{\text{OAR\_2},f}$	1	1	1	1	
$R^{-}_{\text{Tumor},vf}$	9.33	8.33	8	8.33	12
$R^{+}_{\text{Tumor},vf}$	18.66	16.66	16	16.66	24
$R^{+}_{\text{OAR\_1},vf}$	10	10	10	10	
$R^{+}_{\text{OAR\_2},vf}$	10	10	7.66	4.66	
$T^{-}_{\text{Tumor},v}$	46.66	41.66	40	41.66	60
$T^{+}_{\text{Tumor},v}$	93.33	83.33	80	83.33	120
$T^{+}_{\text{OAR\_1},v}$	50	50	50	50	
$T^{+}_{\text{OAR\_2},v}$	50	50	38.33	23.33	

We also use the dose-volume histogram (DVH), a widely used clinical evaluation criterion, to graphically illustrate the radiation dose delivered to a target volume. For any dose level $d \geq 0$ given in Gy, the DVH curve function value, DVH(d), is defined as:

$$100 \cdot \frac{\text{Number of tissue voxels that receive more than } d \text{ Gy of dose}}{\text{Total number of tissue voxels}}$$

As we perform multiple simulations in our evaluation process, we report the 90% prediction interval of each performance metric. To ensure the percentages are meaningful values, we write the minimum between 100% and the upper interval bound, and analogously the maximum between 0% and the lower interval bound.

6 Experimental results

This section presents the numerical results of the proposed solution methods. We analyze the impact of different risk measures on the quality of MSP policies and the SDDP algorithm performance (Section 6.1), compare treatment plans obtained from our MSP model with those of the standard PTV-based deterministic model (Section 6.2), perform a sensitivity analysis on the number fractions (Section 6.3), test different margin choices for the PTV-based approach (Section 6.4), evaluate the quality of solutions when more time is devoted to the SDDP solution approach (Section 6.5), evaluate the impact of increasing the number of sampled voxels (Section 6.6), and finally compare our method with a simplified two-stage stochastic model.

For the ease of exposition, we present the detailed experimental results only for Case 3 for most of the analyses, which can be taken as the representative data instance. We note that the effectiveness of the deterministic and stochastic models for generating treatment plans that are highly conformal to the specifications (e.g., maximum OAR dose limits) are highly impacted by instance characteristics such as tumor size and shape, and proximity of OARs to the tumors. Similarly, relative success of the stochastic treatment plan over its deterministic counterpart can be noticeably impacted by instance characteristics. Overall, we observe consistent improvements

brought up by the stochastic models across five distinct cases. We also provide detailed experimental results in the Online Supplement for each case.

6.1 Risk measures

We begin by reporting the performance of the MSP policies obtained via the SDDP algorithm with respect to the five risk measures provided in Section 3.3. Table 5 summarizes the prediction interval results for Case 3 when limiting the SDDP solution time to 12 hours, as larger cut-off times did not reflect significant quality gains (see Section 6.5 for details). All policies achieve an acceptable coverage of at least 97.6% and hotspot tumor values of zeros, reflecting good-quality treatments overall. With regards to the organs, only ρ_A (average value-at-risk) and ρ_W (worst-case) have achieved zero hotspot values for OAR_2 with certainty, while also presenting relatively low prediction interval upper bounds for OAR_1 (at most 8.6%).

Table 5: Risk measure comparison for Case 3 with 12-hour runtime limit.

	ρ_{EA}	ρ_{EW}	ρ_A	ρ_E	ρ_W
$\mathbb{C}$	[97.6, 100.0]	[98.0, 100.0]	[97.7, 100.0]	[97.8, 100.0]	[97.9, 99.9]
$\mathbb{H}_{\text{Tumor\_1}}$	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]
$\mathbb{H}_{\text{OAR\_1}}$	[0.0, 6.151]	[0.0, 7.767]	[0.0, 5.726]	[0.0, 5.546]	[0.0, 8.595]
$\mathbb{H}_{\text{OAR\_2}}$	[0.0, 0.337]	[0.0, 0.337]	[0.0, 0.0]	[0.0, 0.168]	[0.0, 0.0]

For the purposes of our analysis, the remainder of this section focuses on the worst-case metric ρ_W because of its relatively similar performance with respect to the other metrics, except for the marginally higher prediction interval upper bound for the hotspot value of OAR_1. In particular, we wish to observe how changing other parameters (e.g., runtime) impacts the quality of the policy.

6.2 Value of stochastic solution

We next compare the quality of the MSP policies to the deterministic policies, i.e., the PTV-based procedures consisting of adding a margin across tissue volumes. Our objective is to assess the value of incorporating uncertainty into the fractionated IMRT planning, specifically considering the proposed multi-stage perspective and the worst-case risk metric (e.g., risk-averse policies). Table 6 summarizes the prediction interval values for all cases considering a 12-hour runtime limit and the worst-case risk measure ρ_W for the stochastic approach.

Table 6: Comparison of the deterministic (*Det.*) and the multi-stage stochastic approach (*MSP*) for all cases with 12-hour runtime limit and risk measure ρ_W.

	Case 1		Case 2		Case 3		Case 4		Case 5	
	Det.	MSP	Det.	MSP	Det.	MSP	Det.	MSP	Det.	MSP
$\mathbb{C}$	[94, 100]	[98, 100.0]	[86, 100]	[100, 100]	[97.5, 100]	[97.9, 99.9]	[99, 100]	[99, 100]	[95, 100]	[97, 100]
$\mathbb{H}_{\text{Tumor\_1}}$	[0, 20.1]	[0, 0.4]	[0, 3.7]	[0, 0.5]	[5.97, 36.58]	[0.0, 0.0]	[9.1, 39.1]	[0, 0.0]	[1.1, 65.3]	[10.1, 47.4]
$\mathbb{H}_{\text{OAR\_1}}$	[42.7, 62.2]	[31.4, 54.3]	[0, 15.1]	[0.6, 13.3]	[0, 12.40]	[0.0, 8.595]	[0, 17.6]	[0, 11.2]	-	-
$\mathbb{H}_{\text{OAR\_2}}$	[0, 0.0]	[0, 0.0]	[0, 90.2]	[0, 9.2]	[0, 74.50]	[0.0, 0.0]	[0, 100]	[0, 71.7]	-	-

The results in Table 6 suggests that the multi-stage stochastic policy provides significant benefits with respect to the deterministic policy. For both the tumor and the two OARs, the width of the prediction intervals is significantly smaller for the stochastic policy, which is highly desirable from a clinical perspective. Moreover, for both OARs, the number of voxels receiving more than the

maximum prescribed dose is significantly lower in the treatment plan obtained using the stochastic model. This indicates that the plan obtained from the stochastic model is potentially safer, an essential property in cancer treatment plans especially for tumors near sensitive organs, such as the brain.

We also provide the DVH curves associated with the two policies in Figures 6a and 6b for Case 3 (see Online Supplement for detailed results for other cases). The figures are obtained as a result of 200 simulations and illustrate the means of the DVH values and the 90% prediction intervals, respectively, in solid lines with markers and shaded areas. The critical thresholds, more precisely the maximum and minimum dose amounts, are indicated with solid vertical lines with markers at their endpoints. In particular, the deterministic model presents a higher variability in received dosages (i.e., larger shaded areas) across the three tissues. Dose amounts in the deterministic models are also typically higher than their prescribed minimum, and significantly higher than the maximum dose with respect to the tumor. The figures suggest that the damage risk to healthy tissues caused by the deterministic plan could be potentially more prominent than the damage caused by the stochastic plan.

Lastly, we present a comparison in terms of total dose penalties. More specifically, let Γ_ℓ^+ and Θ_ℓ^+ represent the total fraction-wise and end-of-treatment overdose penalty, respectively, incurred by a treatment plan, and analogously Γ_ℓ^- and Θ_ℓ^- for the underdose. That is,

$$\Theta_\ell^+ = \frac{1}{|\mathcal{V}_\ell|}\sum_{v\in\mathcal{V}_\ell}\sum_{f\in[F]}\max\{0, z_{\ell vf} - R^+_{\ell vf}\}, \quad \Theta_\ell^- = \frac{1}{|V_l|}\sum_{v\in\mathcal{V}_\ell}\sum_{f\in[F]}\max\{0, R^-_{\ell vf} - z_{\ell vf}\}$$

$$\Gamma_\ell^+ = \frac{1}{|\mathcal{V}_\ell|}\sum_{v\in\mathcal{V}_\ell}\max\{0, I_{\ell i,F+1} - T^+_{\ell v}\}, \quad \Gamma_\ell^- = \frac{1}{|\mathcal{V}_\ell|}\sum_{i\in\mathcal{V}_\ell}\max\{0, T^-_{\ell v} - I_{\ell v,F+1}\}$$

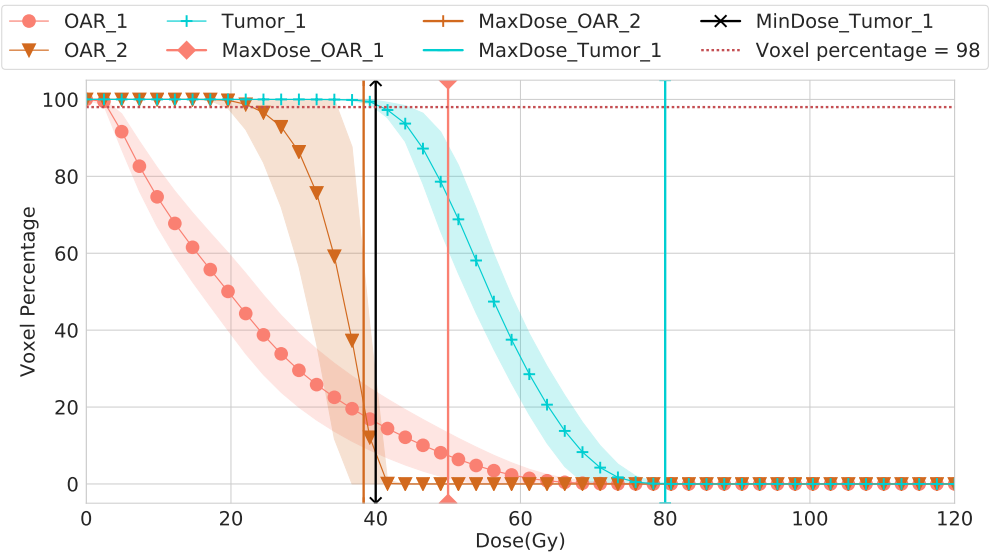
Table 7 presents the resulting prediction interval overdose and underdose total penalties for all tissues. The symbols "+" and '−" next to the tissue names represent overdose and underdose values, respectively; note that organs are only associated with overdose penalties. As indicated in the previous DVH analysis, the prediction interval upper bound on the maximum and minimum dose penalties for the stochastic plan are potentially one order of magnitude less in comparison to the deterministic plans, providing further indication as to its benefits with respect to the deterministic plan.

Table 7: Total dose penalties for Case 3 with 12-hour runtime limit and risk measure ρ_W.

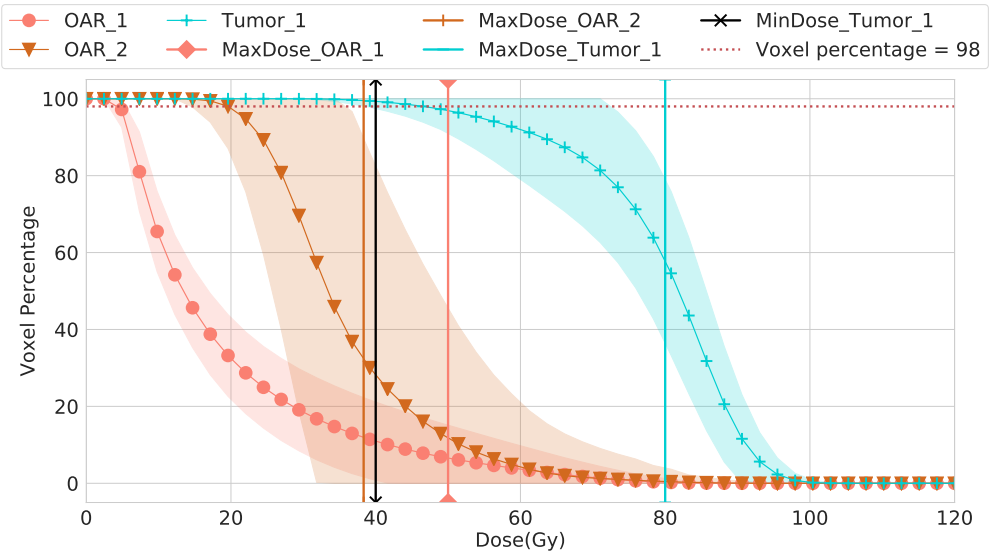
		OAR_1+	OAR_2+	Tumor+	Tumor-
Deterministic	Θ	[0, 5.42]	[0, 17.41]	[15.87, 25.90]	[0, 1.83]
	Γ	[0, 2.25]	[0, 12.53]	[1.74, 6.18]	[0, 0.10]
MSP	Θ	[0.07, 0.39]	[0.03, 0.42]	[0.0, 0.04]	[0.01, 0.05]
	Γ	[0.0, 1.24]	[0.0, 1.27]	[0.0, 0.0]	[0.0, 0.07]

6.3 Number of fractions

We next evaluate the impact of the number of fractions F on treatment quality. To this end, we maintain the end-of-treatment target values of T^+ and T^- as in Table 4, but re-set the fraction-wise targets R^+ and R^- to T^+/F and T^-/F, respectively. For example, in Case 3 we consider $T^-_{\text{Tumor},v} = 80$ and $R^-_{\text{Tumor},v} = 40$ for $F = 2$. The underlying reasoning is that fewer fractions reflect shorter treatment plans, and fraction-wise doses must be higher to achieve the desired treatment goals.



(a) MSP model



(b) Deterministic model

Figure 6: DVHs of the treatment plans for Case 3

Table 8 provides the performance metrics of the treatment plans for $F \in \{3, 4, 5, 6, 7\}$, risk measure ρ_W, and a runtime of 12 hours. We observe that in general, the larger F is, the better the coverage prediction interval lower bound is. This suggests that, in the presence of more fractions, there is a higher chance for the model to fine-tune the treatment plan and achieve a higher level of robustness against dosage excess or shortage. However, we note that this observation is not always valid. For instance, the hotspot prediction interval upper bound for OAR_1 is not the best in the case of $F = 7$. This can potentially attributed to the use of a statistical stopping criterion for the SDDP algorithm. Moreover, larger number of fractions also impact solution times; with $F = 7$, the scenario tree used by the SDDP has now $20^7 = 25.6$ billion scenarios as opposed to 3.2 million with $F = 5$. In such cases, it is likely that larger runtimes would be needed to achieve the same

hotspot values for OAR_1.

Table 8: Number of fractions comparison for Case 3 with 12-hour runtime limit and risk measure ρ_W.

	$F = 3$	$F = 4$	$F = 5$	$F = 6$	$F = 7$
$\mathbb{C}$	[97.1, 100.0]	[97.7, 100.0]	[97.9, 99.9]	[98.1, 99.9]	[98.2, 99.9]
$\mathbb{H}_{\text{Tumor\_1}}$	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]
$\mathbb{H}_{\text{OAR\_1}}$	[0.0, 10.556]	[0.0, 9.395]	[0.0, 8.595]	[0.0, 7.655]	[0.0, 7.196]
$\mathbb{H}_{\text{OAR\_2}}$	[0.0, 3.565]	[0.0, 0.168]	[0.0, 0.0]	[0.0, 0.168]	[0.0, 0.242]

6.4 Margin analysis

Next, we analyze the impact of the margin size on the PTV approaches, again using Case 3 as reference. We test the performance of the deterministic model by considering 0 (mm) to 4 (mm) extra margin to the CTV. We compare this with respect to our current CTV+ baseline, i.e., obtained obtained by adding 2 (mm) margin to the CTV (see Section 5). Our objective is to assess whether it is possible for the deterministic models to obtain similar performance as the stochastic model, in particular with respect to coverage, by increasing the target volume margins.

Table 9 presents the numerical results for our experiments with different margin values. We observe that higher margins do impact the coverage of the deterministic models as expected. However, increasing the CTV margin also increases organ exposure; notice that the prediction interval upper bound for the OAR_2 hotspot is 0 for the MSP model, whereas the hotspot value upper bounds are between 36.36 and 74.50 for the deterministic model. Overall, the width of the prediction intervals for all performance metrics is consistently narrower when the stochastic model is used.

Table 9: Margin analysis for Case 3. The stochastic model considers a 12-hour runtime limit and risk measure ρ_W.

	Deterministic					MSP
	0 (mm)	1 (mm)	2 (mm)	3 (mm)	4 (mm)	2 (mm)
$\mathbb{C}$	[89, 100]	[94, 100]	[96, 100]	[97, 100]	[98, 100]	[98, 100]
$\mathbb{H}_{\text{Tumor\_1}}$	[0, 1.25]	[0, 8.25]	[0, 17.38]	[0, 18.72]	[5.97, 36.58]	[0.00, 0.00]
$\mathbb{H}_{\text{OAR\_1}}$	[0, 6.79]	[0, 8.49]	[0, 10.52]	[0, 11.40]	[0, 12.40]	[0, 8.60]
$\mathbb{H}_{\text{OAR\_2}}$	[0, 36.36]	[0, 49.93]	[0, 58.09]	[0, 64.77]	[0, 74.50]	[0.00, 0.00]

6.5 Runtime

Increasing the SDDP runtime limit leads to additional cuts and hence more accurate cost-to-go approximations. Thus, a natural question is how larger times correlate with the quality of the proposed treatments. Table 10 reports the prediction interval values for Case 3 considering runtimes of four up to 24 hours, with simulations considering increments of four hours each. Results suggest that there is little incentive in larger runtimes for coverage and tumor hotspot values, since such metrics are at least 97.8% and surely 0.0, respectively, starting at the four-hour runtime. However,

there are noticeable differences in OAR_2 hotspot value, which has a lower prediction interval upper bound starting at 12 hours.

Table 10: Runtime comparison for Case 3 with risk measure ρ_W.

time	4 hs	8 hs	12 hs	16 hs	20 hs	24 hs
$\mathbb{C}$	[98.0, 99.9]	[97.8, 100.0]	[97.9, 99.9]	[97.9, 99.9]	[98.0, 99.9]	[97.9, 100.0]
$\mathbb{H}_{\text{Tumor\_1}}$	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]
$\mathbb{H}_{\text{OAR\_1}}$	[0.0, 7.714]	[0.0, 8.408]	[0.0, 8.595]	[0.0, 8.388]	[0.0, 8.236]	[0.0, 8.638]
$\mathbb{H}_{\text{OAR\_2}}$	[0.0, 0.886]	[0.0, 1.503]	[0.0, 0.0]	[0.0, 0.538]	[0.0, 0.381]	[0.0, 0.0]

6.6 Voxel downsampling ratio

While we consider the original number of voxels of each tissue (Table 3) for evaluating performance, intuitively the quality of the MSP treatment plans could be improved if we add more voxels when solving the SDDP, since our downsampling is significant to ensure the underlying stochastic models are tractable. We evaluate if this is the case for Cases 1 (the largest of the instances) and 3, considering two and four times more voxels than in our baseline model. For instance, in Case 1, we re-solve and simulate our MSP with 772 voxels (2x) and 1,544 voxels (4x) selected uniformly at random for OAR_2, in comparison to the baseline of 386 voxels.

The simulation results are presented in Tables 11 and 12, respectively, for Cases 1 and 3. We consider the same runtime limit of 12 hours and risk measure ρ_W. The baseline is denoted by a ratio of "1x." We remark that Case 1 represents a scenario where the tumor envelops OAR_1, hence the exposure of such organ is expected to be higher than in previous cases. The results suggest that there are minor differences when doubling the number of voxels with respective to the baseline, except for a slight increase in coverage for both cases and a reduction of the prediction interval upper limit for the hotspot value of Case 3, organ OAR_1. However, the performance for 4x voxel ratio is generally worse in terms of hotspot values. This occurs because the SDDP models are significantly larger and therefore more difficult to solve, highlighting the trade-off between model tractability and treatment quality for such a fixed time limit. Thus, while our treatment plans are clinically reasonable and improvements with respect to the PTV-based deterministic approaches for these cases, similar trade-off analysis should be carried out given the organ sizes and a budget of runtime limits.

Finally, we note in passing that the tested downsampling is uniform. We refer to [25] for other voxel selection strategies.

Table 11: Voxel ratio analysis for Case 1 with a 12-hour runtime limit and risk measure ρ_W.

number of voxels ratio	1x	2x	4x
$\mathbb{C}$	[97.5, 100.0]	[98.0, 100.0]	[97.6, 100.0]
$\mathbb{H}_{\text{Tumor\_1}}$	[0.0, 0.419]	[0.0, 0.48]	[0.0, 2.245]
$\mathbb{H}_{\text{OAR\_1}}$	[31.446, 54.342]	[30.659, 55.207]	[31.221, 56.346]
$\mathbb{H}_{\text{OAR\_2}}$	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]

Table 12: Voxel ratio analysis for Case 3 with a 12-hour runtime limit and risk measure ρ_W.

number of voxels ratio	1x	2x	4x
$\mathbb{C}$	[97.9, 99.9]	[98.5, 100.0]	[99.1, 100.0]
$\mathbb{H}_{\text{Tumor\_1}}$	[0.0, 0.0]	[0.0, 0.0]	[0.0, 0.0]
$\mathbb{H}_{\text{OAR\_1}}$	[0.0, 8.595]	[0.0, 6.833]	[0.0, 6.739]
$\mathbb{H}_{\text{OAR\_2}}$	[0.0, 0.0]	[0.0, 0.898]	[0.0, 5.568]

6.7 Comparison with the two-stage variant of the MSP model

Finally, we compare the SDDP approach with a simpler two-stage stochastic model as an additional benchmark. The two-stage model only imposes non-anticipative constraints in the first stage, i.e., once first-fraction doses are fixed, the model decomposes per scenario and becomes conceptually more tractable. We particularly address it using a classical Benders decomposition with an enhanced master problem initialization; we refer to the Online Supplement for model details and the Benders implementation. In particular, the underlying two-stage linear programs are formulated by sampling 50 scenarios uniformly at random for tractability.

To evaluate the two-stage stochastic programming model, we consider a folding horizon analogously to the deterministic and MSP settings, with the resulting prediction interval performance evaluated as above. We also remark that the same 200 sampled scenarios are used for both the two-stage stochastic programming model and MSP for consistency.

Table 13 reports the prediction intervals for Case 3 considering the deterministic, two-stage stochastic programming, and the MSP methods, the latter with a runtime limit of 12 hours. Tables for the remaining cases are included in the Online Supplement. However, Case 3 represents the two typical characteristics that are observed in all instances. First, the stochastic models significantly outperform the deterministic approach, generating much smaller hotspot prediction interval upper bounds for all tissues. Second, the MSP treatment plans are of relative higher treatment quality, with slight differences in coverage and/or hotspot values. The same trend was generally observed in all cases, except for Case 5 where the tumor hotspots for the MSP were generally higher.

Nonetheless, we also observed that the MSP provided further computational advantages. First, the SDDP algorithm incorporates the full exponentially sized scenario tree as input, deriving lower and upper bounds on the objective function value of the obtained solution as a result. Thus, when the SDDP algorithm stops, we obtain performance guarantees that are of practical relevance when assessing a treatment. Further, simulating a policy with SDDP takes negligible time; between 20 and 30 seconds per sample path. On the other hand, the two-stage stochastic programming solution requires simulating a subset of scenarios extracted from the scenario tree. The simulation of each individual sample path, in turn, amounts to solving F-many two-stage stochastic programming models, which took around 30 minutes on average in our experiments despite using our enhanced decomposition algorithm. We simulated 200 scenarios and ended up spending significantly more computational time even for the simplified, two-stage model. Overall, however, both models significantly outperformed the deterministic setting, motivating further research into their theoretical and computational aspects.

7 Conclusion

In this paper, we proposed the first multistage stochastic model (MSP) for the fractionated IMRT planning problem. We described a transformation method to convert our decision-hazard-based

Table 13: Comparison across the three methods for Case 3 with risk measure ρ_W. The runtime limit for the MSP is set to 12 hours.

	Deterministic	Two-stage	MSP
$\mathbb{C}$	[97.5, 100]	[97.6, 100.0]	[97.9, 99.9]
$\mathbb{H}_{\text{Tumor\_1}}$	[5.97, 36.58]	[0.0, 0.0]	[0.0, 0.0]
$\mathbb{H}_{\text{OAR\_1}}$	[0, 12.40]	[0.0, 9.454]	[0.0, 8.595]
$\mathbb{H}_{\text{OAR\_2}}$	[0, 74.50]	[0.0, 3.398]	[0, 0.00]

model into an hazard-decision-based model, making it amenable to state-of-the-art SDDP solution approaches such as the `SDDP.jl` library. We have also developed a deterministic reformulation of the SDDP via a state-space expansion, which is key to model tractability. We compared the quality of treatment plans obtained by solving the proposed MSP model and its deterministic counterpart for distinct risk measures within a fixed runtime limit. When particularly evaluating the worst-case risk measure in depth, our MSP model generated better treatment plans than those obtained using its deterministic counterparts, in terms of multiple clinically accepted numerical performance metrics. Besides the numerical performance metrics, we also used DVH to show potential benefits over treatment quality obtained by the proposed MSP model. We showed that our model generates safe treatment plans compared to its deterministic counterparts with PTV of any margin size. We have also investigated the sensitivity of our solution to changes in key model parameters, such number of fractions, PTV margins, and number of downsampled voxels. Further, we have also compared our model with a simplification of the MSP as a two-stage stochastic program solved via an enhanced Benders decomposition. In particular, our numerical analysis suggested that our MSP model produces quality treatment plans and outperforms the deterministic setting in all configurations. The results indicate that, despite the curse of dimensionality, the MSP has potential to be considered into clinical practice.

The main limitation of our study is the use synthetically generated instances in our numerical experiments. In this regard, our results constitute a proof of concept for the impact of directly incorporating uncertainty in treatment planning for fractionated IMRT. While our data instances were carefully generated with the input of domain experts, it would be important to test the performance of our proposed approach with real clinical instances in our future works. Furthermore, we only utilized random voxel sampling technique as a data size reduction strategy, allowing incorporation of higher number of scenarios into the proposed MSP model. However, we recognize that adding extra samples is one of many ways to reduce the variance of outcome quality. For instance, Parpas et al. [27] showed that Monte Carlo Markov Chain-based importance sampling technique can greatly contribute to the sample size reduction. Accordingly, combining a complex sampling technique with our proposed model has potential to produce higher quality treatment plans. Lastly, recent developments such as Neural SDDP [10] that leverages deep-learning within an SDDP framework can be employed to increase the scalability of the proposed approach.

References

[1] Michelle Böck, Kjell Eriksson, and Anders Forsgren. On the interplay between robustness and dynamic planning for adaptive radiation therapy. *Biomedical Physics & Engineering Express*, 5(4):045004, 2019.

[2] Thomas Bortfeld, Steve B Jiang, and Eike Rietzel. Effects of motion on the total dose distribution. In *Seminars in Radiation Oncology*, volume 14, pages 41–51. Elsevier, 2004.

[3] Thomas Bortfeld, Timothy CY Chan, Alexei Trofimov, and John N Tsitsiklis. Robust management of motion uncertainty in intensity-modulated radiation therapy. *Oper. Res.*, 56(6):1461–1473, 2008.

[4] Mucahit Cevik, Dionne Aleman, Young Lee, Azamat Berdyshev, H Nordström, Stella Riad, Arjun Sahgal, and Mark Ruschin. Simultaneous optimization of isocenter locations and sector duration in radiosurgery. *Phys. Med. Biol.*, 64(2):025010, 2019.

[5] Timothy CY Chan and Velibor V Mišić. Adaptive and robust radiation therapy optimization for lung cancer. *European J. Oper. Res.*, 231(3):745–756, 2013.

[6] Timothy CY Chan, Thomas Bortfeld, and John N Tsitsiklis. A robust approach to IMRT optimization. *Phys. Med. Biol.*, 51(10):2567, 2006.

[7] Timothy CY Chan, Houra Mahmoudzadeh, and Thomas G Purdie. A robust-CVaR optimization approach with application to breast cancer therapy. *European J. Oper. Res.*, 238(3):876–885, 2014.

[8] Millie Chu, Yuriy Zinchenko, Shane G Henderson, and Michael B Sharpe. Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty. *Physics in Medicine & Biology*, 50(23):5463, 2005.

[9] Shaunak Dabadghao and Arkajyoti Roy. Optimal interventions for adaptive robust optimization under time-dependent uncertainty with application to radiotherapy. *Available at SSRN 3624421*, 2020.

[10] Hanjun Dai, Yuan Xue, Zia Syed, Dale Schuurmans, and Bo Dai. Neural stochastic dual dynamic programming. *arXiv preprint arXiv:2112.00874*, 2021.

[11] Mellar P Davis, Petra Feyer, Petra Ortner, and Camilla Zimmermann. *Supportive Oncology E-Book*. Elsevier Health Sciences, 2011.

[12] Oscar Dowson. *Applying Stochastic Optimisation to the New Zealand Dairy Industry*. PhD thesis, University of Auckland, 2018.

[13] Oscar Dowson. The policy graph decomposition of multistage stochastic programming problems. *Networks*, 76(1):3–23, 2020.

[14] Oscar Dowson and Lea Kapelevich. SDDP.jl: A Julia package for stochastic dual dynamic programming. *INFORMS J. Comput.*, 33(1):27–33, 2021.

[15] Erik Engwall, Albin Fredriksson, and Lars Glimelius. 4D robust optimization including uncertainties in time structures can reduce the interplay effect in proton pencil beam scanning radiation therapy. *Med. Phys.*, 45(9):4020–4029, 2018.

[16] A Fredriksson. *Robust optimization of radiation therapy accounting for geometric uncertainty*. PhD thesis, KTH Royal Institute of Technology, 2013.

[17] V Grégoire and TR Mackie. State of the art on dose prescription, reporting and recording in intensity-modulated radiation therapy (ICRU report no. 83). *Cancer/Radiothérapie*, 15(6-7):555–559, 2011.

[18] Shogo Kishimoto and Makoto Yamashita. A successive LP approach with C-VaR type constraints for IMRT optimization. *Operations Research for Health Care*, 17:55–64, 2018.

[19] Gino J Lim, Laleh Kardar, Saba Ebrahimi, and Wenhua Cao. A risk-based modeling approach for radiation therapy treatment planning under tumor shrinkage uncertainty. *European J. Oper. Res.*, 280 (1):266–278, 2020.

[20] Houra Mahmoudzadeh, Thomas G Purdie, and Timothy CY Chan. Constraint generation methods for robust optimization in radiation therapy. *Oper. Res. for health care*, 8:85–90, 2016.

[21] DL McShan, ML Kessler, K Vineberg, and BA Fraass. Inverse plan optimization accounting for random geometric uncertainties with a multiple instance geometry approximation (MIGA). *Med. Phys.*, 33(5): 1510–1521, 2006.

[22] Chunhua Men, H Edwin Romeijn, Anneyuko Saito, and James F Dempsey. An efficient approach to incorporating interfraction motion in IMRT treatment planning. *Comput. and Oper. Res.*, 39(7): 1779–1789, 2012.

[23] Andrea Michalski, John Atyeo, Jennifer Cox, and Marianne Rinks. Inter-and intra-fraction motion during radiation therapy to the whole breast in the supine position: a systematic review. *J. Med. Imaging. Radiat. Oncol.*, 56(5):499–509, 2012.

[24] Andrew Murphy and Edward Chmiel. Fractionation (radiation therapy), 2019. https://radiopaedia.org/articles/fractionation-radiation-therapy.

[25] Boris N Oreshkin and Tal Arbel. Uncertainty driven probabilistic voxel selection for image registration. *IEEE transactions on medical imaging*, 32(10):1777–1790, 2013.

[26] E Orvehed Hiltunen. Robust optimization of radiotherapy treatment plans considering time structures of the delivery, 2018.

[27] Panos Parpas, Berk Ustun, Mort Webster, and Quang Kha Tran. Importance sampling in stochastic programming: A markov chain monte carlo approach. *INFORMS J. Comput.*, 27(2):358–377, 2015.

[28] Mario VF Pereira and Leontina MVG Pinto. Multi-stage stochastic optimization applied to energy planning. *Math. Program.*, 52(1-3):359–375, 1991.

[29] Danielle A Ripsman, Thomas G Purdie, Timothy CY Chan, and Houra Mahmoudzadeh. Robust direct aperture optimization for radiation therapy treatment planning. *arXiv preprint arXiv:2111.04847*, 2021.

[30] H Edwin Romeijn, Ravindra K Ahuja, James F Dempsey, and Arvind Kumar. A new linear programming approach to radiation therapy treatment planning problems. *Oper. Res.*, 54(2):201–216, 2006.

[31] Fatemeh Saberian, Archis Ghate, and Minsun Kim. Spatiotemporally optimal fractionation in radiotherapy. *INFORMS Journal on Computing*, 29(3):422–437, 2017.

[32] Anna Samuelsson, Claes Mercke, and Karl-Axel Johansson. Systematic set-up errors for IMRT in the head and neck region: effect on dose distribution. *Radiother. Oncol.*, 66(3):303–311, 2003.

[33] A. Shapiro. Analysis of stochastic dual dynamic programming method. *European J. Oper. Res.*, 209 (1):63–72, 2011.

[34] Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on stochastic programming: modeling and theory*. SIAM, 2014.

[35] Mustafa Y Sir, Marina A Epelman, and Stephen M Pollock. Stochastic programming for off-line adaptive radiotherapy. *Ann. Oper. Res.*, 196(1):767–797, 2012.

[36] Jens Sjölund, Stella Riad, Marcus Hennix, and Håkan Nordström. A linear programming approach to inverse planning in Gamma Knife radiosurgery. *Med. Phys.*, 46(4):1533–1544, 2019.

[37] Helen B Stone, C Norman Coleman, Mitchell S Anscher, and William H McBride. Effects of radiation on normal tissue: consequences and mechanisms. *Lancet Oncol.*, 4(9):529–536, 2003.

[38] Alexandre Street, Davi Valladão, André Lawson, and Alexandre Velloso. Assessing the cost of the hazard-decision simplification in multistage stochastic hydrothermal scheduling. *Appl. Energy*, 280: 115939, 2020.

[39] Z Caner Taşkın and Mucahit Cevik. Combinatorial benders cuts for decomposing imrt fluence maps using rectangular apertures. *Comput. and Oper. Res.*, 40(9):2178–2186, 2013.

[40] Jan Unkelbach and Uwe Oelfke. Inclusion of organ movements in IMRT treatment planning via inverse planning based on probability distributions. *Phys. Med. Biol.*, 49(17):4005, 2004.

Online supplement for "Multistage Stochastic Fractionated Intensity Modulated Radiation Therapy Planning"

M. Bodur[1], M. Cevik[2], A.A. Cire[3], M. Ruschin[4], J. Wang[1]

[1] Department of Mechanical and Industrial Engineering, University of Toronto, ON, Canada

[2] Department of Mechanical and Industrial Engineering, Toronto Metropolitan University, ON, Canada

[3] Department of Management, University of Toronto Scarborough & Rotman School of Management, ON, Canada

[4] Department of Radiation Oncology, Sunnybrook Health Sciences Center, ON, Canada

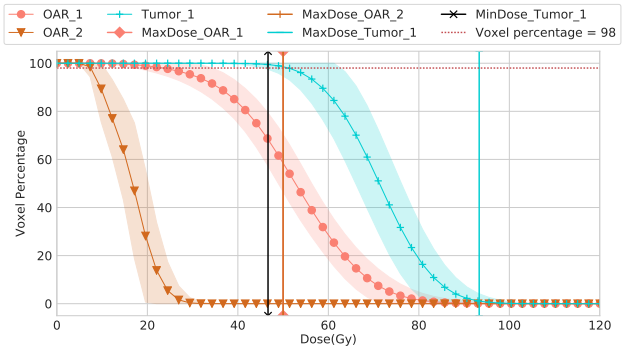
Author to whom correspondence should be addressed. email: mcevik@ryerson.ca

1. Detailed comparative results for clinical cases

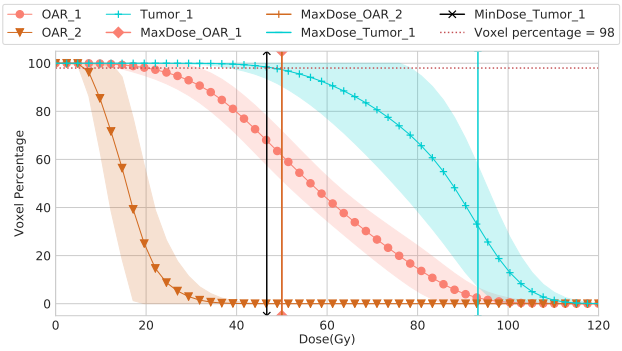
We first provide detailed results for four clinical cases (i.e., except Case 3 that is discussed in detail in the paper), which constitute a diverse test bed for our models. Because these instances vary in terms of size and shape of the tumors as well as proximity of OARs to the tumor volumes, success of the stochastic and deterministic treatment plans vary across the cases. Regardless, in all the cases, we find that stochastic treatment plans substantially outperform the deterministic counterparts.

Figures A.1a and A.1b show the DVHs generated by stochastic and deterministic treatment plans for Case 1, respectively, which has one tumor and two nearby OARs. First, we note that OAR_1 is very close to tumor in this case, which impacts its dose distribution considerably, leading to a noticeable OAR_1 overdose. The deterministic model presents a slightly higher variability in received dosages for both OARs, and also significantly higher variability for the tumor volume. The delivered dose for the tumor volume also significantly exceeds the corresponding maximum dose threshold for the deterministic model.

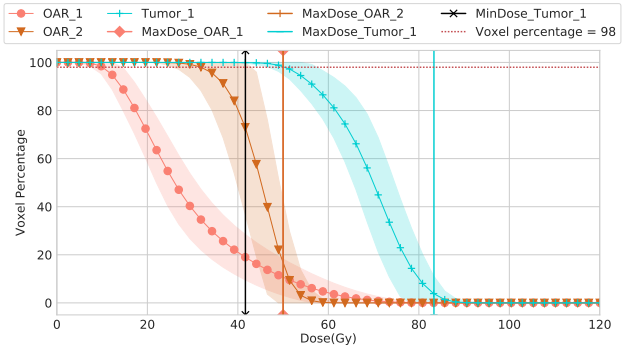
The DVH plots for Case 2 (see figures A.1c and A.1d) and Case 4 (see figures A.1e and A.1f) attribute a more significant advantage to stochastic plans over deterministic ones compared to Case 1. Specifically, we note that stochastic treatment plans show substantially lower variability in radiation dose received for the tumor and two OARs in each case. Case 5 does not include a nearby OAR, and the prescribed dose for tumor is highest for this case. Similar to other cases, figures A.1e and A.1f show that the prediction interval for tumor dose is much narrower in the stochastic treatment plan, showing the capability of our stochastic model in incorporating uncertainty to treatment planning.



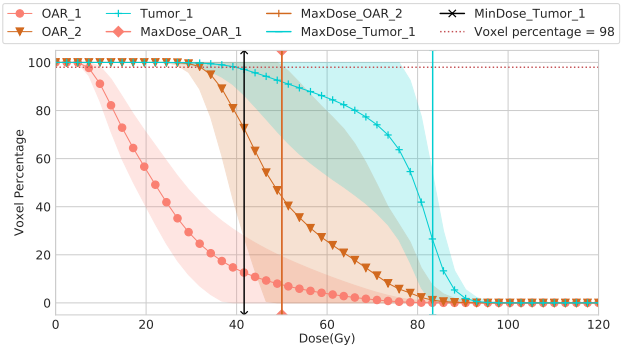
(a) Case 1 - Stochastic plan



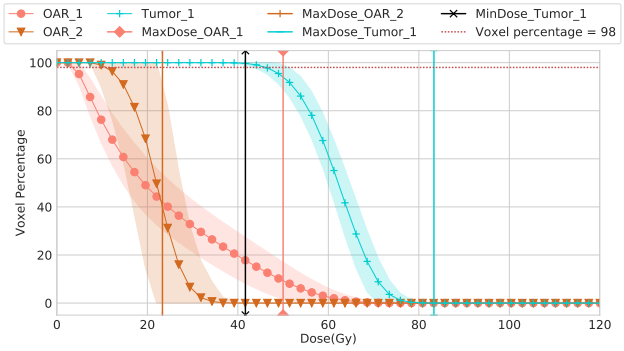
(b) Case 1 - Deterministic plan



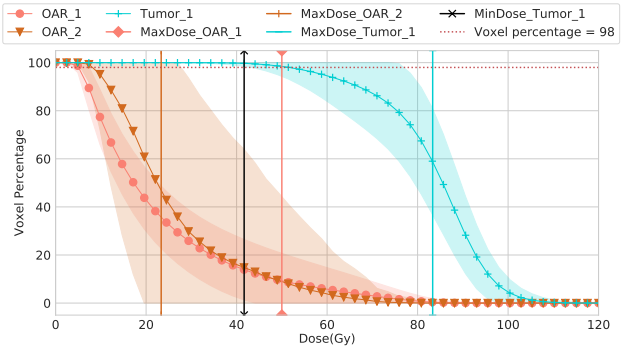
(c) Case 2 - Stochastic plan



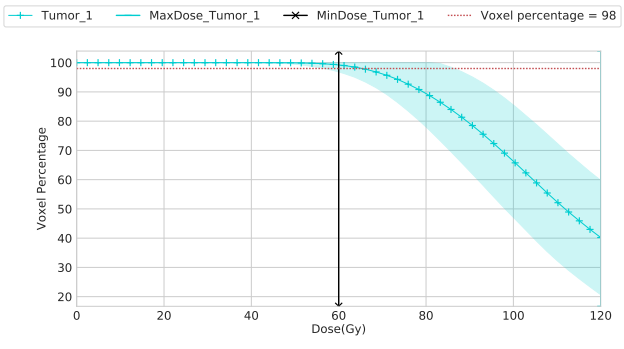
(d) Case 2 - Deterministic plan



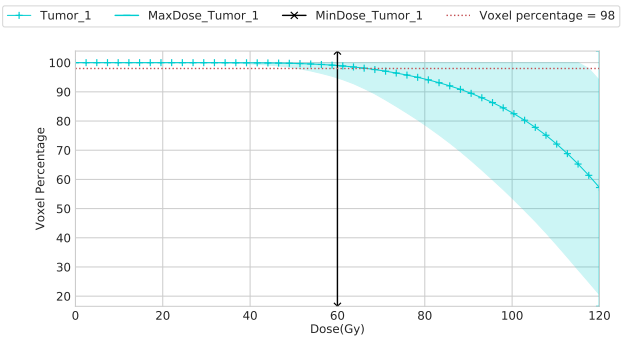
(e) Case 4 - Stochastic plan



(f) Case 4 - Deterministic plan



(g) Case 5 - Stochastic plan



(h) Case 5 - Deterministic plan

Figure A.1: DVHs of the treatment plans (means of the DVH values and the 90% prediction intervals are plotted)

Table A.1 show the prediction intervals of stochastic and deterministic models for overdose and underdose total penalties for all four cases. We observe that, for the stochastic treatment plans, the prediction interval upper bound on the maximum and minimum dose penalties are consistently substantially lower accros these cases, indicating strong benefits over the deterministic plans.

Table A.1: Dose penalty tables for different cases (Γ and Θ represent the total fraction-wise and end-of-treatment penalties, respectively, and "+" and "-" signs next to volumes signify overdose and underdose, respectively).

(a) Case 1 - Stochastic plan

	OAR_1+	OAR_2+	Tumor+	Tumor-
Θ	[1.69, 2.44]	[0, 0.00]	[0.07, 0.43]	[0.03, 0.24]
Γ	[4.02, 9.28]	[0, 0.00]	[0.0, 0.21]	[0.0, 0.16]

(b) Case 1 - Deterministic plan

	OAR_1+	OAR_2+	Tumor+	Tumor-
Θ	[15.73, 21.72]	[0, 2.19]	[12.85, 30.09]	[0, 4.37]
Γ	[8.85, 15.29]	[0, 0.00]	[0, 4.40]	[0, 0.39]

(c) Case 2 - Stochastic plan

	OAR_1+	OAR_2+	Tumor+	Tumor-
Θ	[0.1, 0.71]	[0.08, 0.81]	[0.11, 0.34]	[0.0, 0.07]
Γ	[0.0, 2.22]	[0.0, 1.61]	[0.0, 0.31]	[0.0, 0.01]

(d) Case 2 - Deterministic plan

	OAR_1+	OAR_2+	Tumor+	Tumor-
Θ	[0, 7.05]	[0, 26.29]	[10.98, 30.98]	[0, 4.66]
Γ	[0, 2.57]	[0, 18.97]	[0, 1.79]	[0, 0.69]

(e) Case 4 - Stochastic plan

	OAR_1+	OAR_2+	Tumor+	Tumor-
Θ	[0.12, 0.5]	[0.13, 1.95]	[0.0, 0.06]	[0.01, 0.07]
Γ	[0.0, 1.59]	[0.0, 5.8]	[0, 0.00]	[0.0, 0.07]

(f) Case 4 - Deterministic plan

	OAR_1+	OAR_2+	Tumor+	Tumor-
Θ	[0, 7.85]	[0, 29.63]	[18.32, 28.21]	[0.25, 2.18]
Γ	[0, 3.63]	[0, 23.81]	[2.14, 7.18]	[0, 0.07]

(g) Case 5 - Stochastic plan

	Tumor+	Tumor-
Θ	[2.62, 7.58]	[0.02, 0.85]
Γ	[2.58, 27.4]	[0.0, 0.33]

(h) Case 5 - Deterministic plan

	Tumor+	Tumor-
Θ	[13.62, 42.96]	[0, 7.20]
Γ	[0.61, 16.46]	[0, 0.43]

2. Two-stage stochastic programming approximation

A multistage stochastic program (MSP) can be approximated by a two-stage stochastic program (2SP). There are two classical such approximations: (i) two-stage restriction of MSP, where all the state variables are made first-stage variables as such they become fully nonanticipative, i.e., deterministic, and (ii) two-stage relaxation of MSP, where where the nonanticipativity constraints on all except first-stage variables are removed, as such the decisions associated with the second to the end stage are all moved to the second stage. For our fractionated IMRT problem, the 2SP restriction would be indeed equivalent to a *one-stage* SP since the determination of all the state variables, namely the cumulative doses received at the beginning of each fraction, leaves no other real treatment related decisions, uniquely implying the values of doses actually received at every fraction and penalties incurred for each scenario. This model would be the analog of the one-stage SP proposed in [1]. On the other hand, the two-stage relaxation of MSP remains as a proper

two-stage model, i.e., there are actual recourse decisions, namely the beamlet intensities associated with all the stages except the first one. The decision-making process of the MSP and the aforementioned 2SP approximations are provided in Figures A.2, A.3, and A.4, respectively. Lastly, we note that due to the special structure in our problem, the two-stage restriction of MSP (i.e., the one-stage SP) is also a restriction of the two-stage relaxation of MSP (i.e., the two-stage SP), as such we propose it as an alternative model of the problem, and compare it with our MSP. In what follows, we explain the two-stage model along with a decomposition algorithm for its efficient solution.

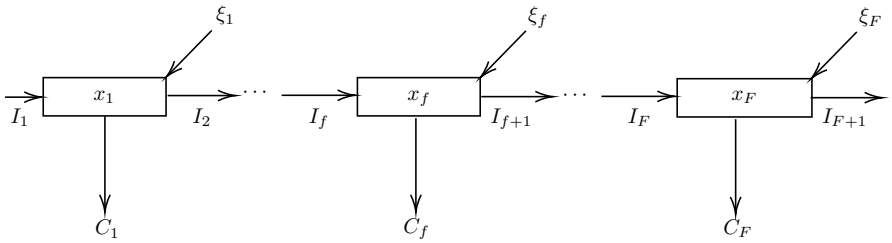


Figure A.2: MSP framework

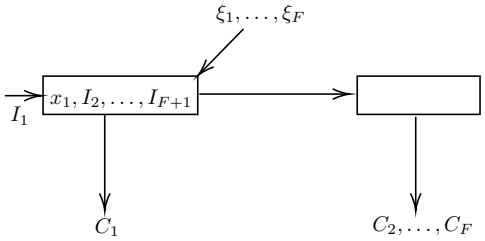


Figure A.3: One-stage SP framework (MSP restriction)

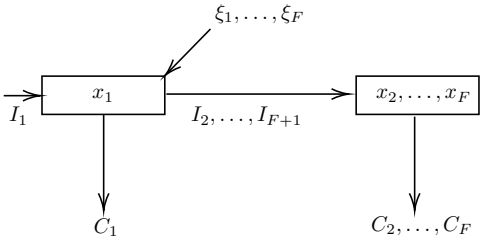


Figure A.4: Two-stage SP framework (MSP relaxation)

The two-stage relaxation of our MSP, obtained by removing the nonanticipativity constraints on all except first-stage variables, as such moving the decisions associated with the second to the end fraction to the second stage, is as follows[1]:

$$\min_{x_1 \in \mathbb{R}_+^{|\mathcal{B}|}} \rho_{\xi^F} \left(Q^{\texttt{AllFuture}}(I_1, x_1, \xi^F) \right) \tag{1}$$

where the initial state vector I_1 is taken as all zeros, and the second-stage value function is

$$Q^{\texttt{AllFuture}}(I_1, x_1, \xi^F) = \min \sum_{f \in [F]} C_f(I_f, x_f, \xi_f) \tag{2a}$$

$$\text{s.t. } I_{f+1} = G(I_f, x_f, \xi_f), \quad \forall f \in [F] \tag{2b}$$

$$x_2, \ldots, x_F \in \mathbb{R}_+^{|\mathcal{B}|}. \tag{2c}$$

[1]We employ the end-of-horizon risk measures rather than the nested ones (whose definitions can be found in [2]) for the ease of exposition in text as well as improving the computational efficiency of the decomposition algorithm to solve the two-stage model.

We solve the sample average approximation of the two-stage model (1), obtained by considering a sample of (equally likely) scenarios $\{\xi_s^F\}_{s\in\mathcal{S}} \subseteq \Xi_1 \times \cdots \times \Xi_F$, via the multi-cut version of the Benders decomposition algorithm [3, 4]. The algorithm decomposes the problem into a master problem and a set of subproblems, one per scenario. It iteratively solves these problems, and at every iteration refines the master problem with the addition of Benders cuts according to the subproblems solutions.

The *master problem* decides on the first-fraction intensities, x_1, using the outer approximation of the second-stage value function which is obtained via the auxiliary variables $\eta(\xi_s^F)$ denoting the second-stage objective under-estimate for each scenario and the generated Benders cuts whose coefficient vector (`coeff`) and constant (`const`) are stored in the set $\texttt{Cuts}(\xi_s^F)$:

$$\min\ \rho_{\xi^F}\left(\eta(\xi_s^F)\right) \tag{3a}$$

$$\text{s.t. } x_1 \in \mathbb{R}_+^{|\mathcal{B}|} \tag{3b}$$

$$\eta(\xi_s^F) \geq \texttt{coeff}^\top x_1 + \texttt{const}, \qquad \forall s \in \mathcal{S},\ (\texttt{coeff}, \texttt{const}) \in \texttt{Cuts}(\xi_s^F) \tag{3c}$$

$$\eta(\xi_s^F) \geq 0, \qquad \forall s \in \mathcal{S} \tag{3d}$$

$$\frac{1}{|\mathcal{S}|}\sum_{s\in\mathcal{S}} \eta(\xi_s^F) \geq \sum_{f\in[F]} C_f(I_f, x_f^{\texttt{Exp}}, \xi_f^{\texttt{Exp}}) \tag{3e}$$

$$x_f^{\texttt{Exp}} \in \mathbb{R}_+^{|\mathcal{B}|}, \qquad \forall f \in [F] \tag{3f}$$

$$I_{f+1}^{\texttt{Exp}} = G(I_f^{\texttt{Exp}}, x_f^{\texttt{Exp}}, \xi_f^{\texttt{Exp}}), \qquad \forall f \in [F] \tag{3g}$$

$$I_1^{\texttt{Exp}} = I_1 \tag{3h}$$

In order to strengthen the master problem, we initialize it with nonnegativity constraints (3d) as the scenario objectives corresponding to deviation penalty cannot be negative, as well as the new set of variables with superscript `Exp` corresponding to the expected scenario denoted by $\xi^{\texttt{Exp}}$ and constraints (3e)-(3h) linking those new variables to the η variables. More specifically, the latter is valid thanks to the Jensen's inequality, which implies (3e) stating the expected cost of decisions should be greater than or equal to the cost of decisions under the expected scenario; this initialization is also in the literature, e.g., see [5].

Given a master problem canditate solution $(\hat{x}_1, \hat{\eta})$, the subproblem for scenario s computes the correct cost of the candidate first-stage solution, $Q^{\texttt{AllFuture}}(I_1, \hat{x}_1, \xi_s^F)$ via the model given in (2), and returns a Benders optimality cut to the master problem if the correct cost is strictly larger than the estimated cost $\hat{\eta}(\xi_s^F)$. We omit the details of deriving the standard Benders optimality cuts, which relies on linear programming duality. We note that as our subproblems are always feasible, there is no need for Benders feasibility cuts.

As mentioned in the main manuscript, we evaluate the two-stage approximation policy in a folding-horizon framework and compare against the deterministic and MSP policies. The supplementary results for this discussion are provided in Table A.2.

Table A.2: Comparison of the two-stage model (2SP) and the multi-stage stochastic approach (MSP) for all cases with 12-hour runtime limit and risk measure ρ_W.

	Case 1		Case 2		Case 3		Case 4		Case 5	
	2SP	MSP	2SP	MSP	2SP	MSP	2SP	MSP	2SP	MSP
$\mathbb{C}$	[98.1, 100]	[98, 100.0]	[99.7, 100]	[100, 100]	[97.6, 100]	[97.9, 99.9]	[99.1, 100]	[99, 100]	[98, 100]	[97, 100]
$\mathbb{H}_{\text{Tumor\_1}}$	[0, 0.353]	[0, 0.4]	[0, 2.244]	[0, 0.5]	[0, 0]	[0, 0]	[0, 0]	[0, 0]	[0.154, 32.042]	[10.1, 47.4]
$\mathbb{H}_{\text{OAR\_1}}$	[33.302, 55.638]	[31.4, 54.3]	[0.042, 11.962]	[0.6, 13.3]	[0, 9.454]	[0.0, 8.595]	[0.171, 12.546]	[0, 11.2]	-	-
$\mathbb{H}_{\text{OAR\_2}}$	[0, 0]	[0, 0]	[0, 13.642]	[0, 9.2]	[0, 3.398]	[0, 0]	[0, 95.846]	[0, 71.7]	-	-

References

[1] M. Y. Sir, M. A. Epelman, S. M. Pollock, Stochastic programming for off-line adaptive radiotherapy, Ann. Oper. Res. 196 (2012) 767–797.

[2] O. Dowson, Applying Stochastic Optimisation to the New Zealand Dairy Industry, Ph.D. thesis, University of Auckland, 2018.

[3] J. F. Benders, Partitioning procedures for solving mixed-variables programming problems, Numerische Mathematik 4 (1962) 238–252.

[4] R. Rahmaniani, T. G. Crainic, M. Gendreau, W. Rei, The Benders decomposition algorithm: A literature review, European J. Oper. Res. 259 (2017) 801–817.

[5] M. Bodur, J. R. Luedtke, Mixed-integer rounding enhanced benders decomposition for multiclass service-system staffing and scheduling with arrival rate uncertainty, Management Sci. 63 (2017) 2073–2091.