Robust, Non-stationary, and Adaptive Fractionation in Radiotherapy

Ali Ajdari

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Archis Ghate, Chair

Linda Ng Boyle

Shan Liu

Hamed Mamani (GSR)

Program Authorized to Offer Degree:
Industrial & Systems Engineering

University of Washington

Abstract

Robust, Non-stationary, and Adaptive Fractionation in Radiotherapy

Ali Ajdari

Chair of the Supervisory Committee:
Professor Archis Ghate
Industrial & Systems Engineering

In external beam radiotherapy for cancer, high-energy radiation is passed through the patient's body from an outside source to kill tumor cells. The challenge is that radiation also damages healthy tissue and organs-at-risk (OAR) in its path. The objective therefore is to devise treatment plans that maximize tumor-damage while protecting healthy anatomies. Treatment planners attempt two separate methods to attain this goal: spatial and biological.

The spatial side focuses on the geometry and physics of the problem. The key consideration here is the location of the tumor relative to the nearby healthy regions as seen in an anatomical image, and the dose (energy absorbed per unit mass) deposition properties of the radiation beam. The treatment planner prescribes a high dose to the tumor and puts upper limits on the doses delivered to the healthy regions. Intensity Modulated Radiation Therapy (IMRT) technology is then employed to tune the profile (fluence-map) of the radiation beam to administer a dose that is as close as possible to this tumor-conforming prescription. Several mathematical optimization models and solution algorithms for this problem have been developed and embedded into treatment planning systems.

The biological side of planning exploits the difference between the dose-response characteristics of tumors and healthy anatomies. For example, healthy cells are believed to possess better damage repair capabilities than tumor cells. Thus, treatment is delivered over multiple sessions to give healthy tissue some time to recover between sessions. This is called

fractionation. Fractionation also gives the tumor some time to re-oxygenate, which increases its sensitivity to radiation. Tumors, however, proliferate during the treatment course, and hence, too long a treatment course may not be ideal. One key question on this biological side is to determine the optimal number of treatment sessions. This is called the fractionation problem. Existing optimization research on the fractionation problem relies on the linear-quadratic (LQ) model of dose-response with tumor- and OAR-specific parameters to approximately capture the behavior of the complex biological system involved.

Recent studies have suggested that an integrated approach that simultaneously tackles the spatial and biological sides of the problem may lead to a higher tumor-damage as compared to tackling the two aspects separately. The goal in such integrated formulations is to simultaneously find the fluence-map and the number of sessions that maximize tumor-damage while limiting toxic effects of dose on the healthy anatomies. Emerging advances in quantitative functional imaging technologies are enabling planners to observe the tumor's actual dose-response over the treatment course. This provides additional opportunities for better-utilizing the LQ model by dynamically adapting treatment plans to further improve outcomes. The challenge, however, is that spatiobiologically integrated formulations based on the LQ model typically yield nonconvex quadratically constrained quadratic programming problems, which are computationally difficult to solve exactly.

The research objective of this dissertation is to develop efficient convex, robust, and dynamic optimization methods to formulate and approximately solve different nonadaptive and adaptive versions of the spatiobiologically integrated fractionation problem within the LQ framework. Chapter 1 briefly describes state-of-the-art literature on spatiobiologically integrated fractionation. Each subsequent chapter is motivated by a distinct limitation of an existing formulation of the spatiobiologically integrated fractionation problem from this literature.

Chapter 2: The solutions offered by existing formulations of the fractionation problem

crucially depend on the assumed values of the dose-response parameters of the LQ model. Unfortunately, "true" values of these parameters are unknown. Consequently, a solution of the fractionation problem may not be feasible in practice. This concern is addressed in Chapter 2 via a robust formulation, whose solution remains feasible as long as the dose-response parameters belong to an interval. An efficient solution method rooted in a convex, finite-dimensional reformulation of the resulting infinite-dimensional problem is proposed. The price of robustness and feasibility properties of the robust solution are quantified via numerical experiments.

Chapter 3: Existing spatiobiologically integrated formulations of the fractionation problem assume that the fluence-map is not changed across treatment sessions. From a computational viewpoint, this simplifies the problem significantly. Chapter 3 relaxes this assumption, and proposes an efficient solution method that allows the fluence-maps to vary across sessions. The quality of the time-variant solutions produced by this method is compared against traditional time-invariant solutions via numerical experiments.

Chapter 4: Adaptive spatiobiologically integrated fractionation attempts to alter fluence-maps according to the observed evolution of tumor cell density in functional images. Chapter 4 proposes a formulation and solution method that also determine the length of the remaining treatment course adaptively. Potential benefits of such adaptive treatment-length planning are investigated through numerical experiments.

Chapter 5: Adaptive fluence-map planning methods assume that the treatment planner knows the probability distribution of the uncertainty in the tumor's dose-response parameters. In contrast to this "clairvoyant" approach, Chapter 5 proposes an alternative formulation, where the treatment planner learns this distribution from tumor-response information observed in functional images over the treatment course while also adaptively optimizing fluence-maps. This yields a Bayesian stochastic control formulation whose exact solution is impossible to derive. The chapter proposes a simple approximate solution method rooted

in certainty equivalent control, and compares its performance agains a clairvoyant certainty equivalent control scheme and a "no learning" approach via numerical experiments.

Finally, Chapter 6 outlines limitations of this dissertation work and describes two directions for future research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to sincerely thank Dr. Archis Ghate, who has been there for me not only as an advisor, but a friend. I'd like to thank Dr. Linda Boyle who has kindly guided me through my days at UW, literally from the moment I set foot in it. I'd also like to express my deepest appreciations to all my friends (you know who you are) who have been there for me through the sunshines and rains. I am forever indebted to all of them.

DEDICATION

to the warm memory of my mother, whispering softly in my ears
to the ever-lasting smiles of my father, waltzing on through the glass
to all the cradles distraught, the cries unconsoled, the lullabies unsung
to Aylan, Omran, Avaa
and one more.

They lived and laughed and loved and left...

Chapter 1

INTRODUCTION

In external beam radiotherapy for cancer, high-energy radiation is passed through the patient's body from an outside source to kill tumor cells. Unfortunately, radiation also damages nearby organs-at-risk (OAR). The goal therefore is to maximize tumor-damage while maintaining toxic effects on the OAR within tolerable limits. Efforts at attaining this objective focus on two components: spatial (also called anatomical) and biological (also called temporal).

The goal on the spatial side is to tune the radiation intensity profile (called the fluence-map) so that it delivers a high dose (energy per unit mass) to the tumor while keeping dose to the OAR under tolerable limits. This can be done via the now prevalent Intensity Modulated Radiation Therapy (IMRT) technology. Treatment planners attempt to find a fluence-map that delivers a dose that is as close as possible to an institutionally-approved dose-prescription that conforms to the shape of the tumor. This is often formulated as an optimization problem. Numerous increasingly sophisticated mathematical formulations of this problem and corresponding solution algorithms have been devised over the last two decades [16, 18, 20, 21, 42, 51, 52, 58, 72, 77]. Some of these are embedded into planning systems at cancer treatment centers around the world [72]. As such, research on the mathematical optimization aspects of this spatial side is well-matured and has been transferred to practice.

The biological side exploits the difference between how tumor cells and OAR respond to radiation. For instance, healthy cells are believed to possess better damage-repair capabilities than tumor cells. Thus, radiotherapy is often administered in multiple treatment sessions. This is called fractionation, and it gives healthy cells some time to recover between sessions. Fractionation typically enables the treatment planner to administer a higher total dose to

the tumor than what would be possible (without exceeding OAR tolerance limits) in a single session. However, tumors can proliferate between sessions and hence it might not be a good idea to employ a long treatment course. One central question on the biological side therefore is to find an optimal number of sessions. This is called the fractionation problem [50].

There is a large body of theoretical literature on the fractionation problem. This literature utilizes the linear-quadratic (LQ) model of dose-response from radiobiology [33].

The prevalent idea in theoretical formulations of the fractionation problem based on the LQ model is to find a prescription dose that maximizes tumor cell-kill while limiting the biologically effective dose (BED) to the OAR [28]. Spatial optimization is then employed to find a fluence-map that delivers a dose close to this prescription that was derived via biological considerations. This is called the spatiobiologically *separated* method, as it solves two different optimization problems: one biological and one spatial [3, 11, 15, 23, 26, 29, 34, 35, 48, 54, 55, 67, 75]. These separated formulations in essence assume that a fluence-map that is obtained via standard spatial optimization methods embedded in ubiquitous IMRT planning systems can be appropriately scaled to administer a biologically optimized dose [54]. It is partly for this reason that such separated formulations are believed to have the potential to be one day transferred to clinical practice.

Recent theoretical work has focused on the spatiobiologically *integrated* case. There, a fluence-map that maximizes tumor cell-kill subject to BED limits on OAR is found directly [1, 57, 65, 67, 68, 69]. Integrated formulations are computationally more demanding, but, on the positive side, Saberian et al. [57] found via numerical experiments that they could lead to a higher tumor-damage than separated formulations. Integrated formulations typically do not employ the concept of a prescription dose that is familiar to practitioners. It is perhaps for this reason that they are viewed to be further removed from current practice as compared to separated formulations.

There are several limitations of this theoretical literature on the fractionation problem. Each chapter in this dissertation attempts to answer a research question motivated by a distinct limitation as briefly described here.

- **Chapter 2:** Existing spatiobiologically separated as well as integrated formulations of the fractionation problem assume that the treatment planner knows the values of the parameters of the LQ dose-response model. Feasibility of the resulting "nominal" dosing plan crucially depends on these assumed parameter values. Consequently, in practice, a dosing plan that is presumed optimal might not even be tolerable by some of the OAR. Chapter 2 proposes a conservative approach to tackle this challenge. Specifically, it formulates *robust* spatiobiologically separated and integrated versions of the fractionation problem. These robust formulations have the property that their solutions remain feasible as long as the LQ dose-response parameters lie in pre-specified intervals. At first glance, these formulations appear computationally intractable, partly because they include an uncountably infinite number of constraints. The chapter develops efficiently solvable, equivalent, finite-dimensional convex reformulations of these robust problems. Numerical experiments are performed to quantify the price of robustness and to compare the feasibility properties of robust and nominal solutions.

- **Chapter 3:** The aforementioned spatiobiologically integrated formulations of the fractionation problem assume that the treatment planner does not vary the fluence-maps across sessions. That is, fluence-maps are assumed to be stationary. The reasoning behind this is twofold. Firstly, it is not known whether or not nonstationary fluence-maps could lead to a higher tumor-damage within the LQ dose-response framework. Secondly, exact solution of the non-stationary fractionation problem is computationally intractable. An approximate solution method that could achieve a higher tumor-damage as compared to that attained by a stationary solution is currently not known. Chapter 3 proposes one solution method that appears to possess this property based on numerical experiments.

- **Chapter 4:** Recent theoretical research on spatiobiologically integrated fractionation has considered adaptive planning of fluence-maps [38, 39, 56]. There, the planner alters the fluence-maps based on the evolution of cell density in different regions of the

tumor as *observed* in functional images over the treatment course. Chapter 4 proposes a solution method that adaptively determines the length of the treatment course as well as the fluence-maps. Specifically, after observing the tumor cell densities at the beginning of a session, the treatment planner solves a group of convex optimization problems to determine an optimal number of remaining treatment sessions, and a corresponding optimal fluence-map for each of these sessions. The objective is to minimize the total number of tumor cells remaining (TNTCR) at the end of this proposed treatment course, subject to upper limits on the biologically effective dose delivered to the organs-at-risk. This fluence-map is administered in future sessions until the next image is available, and then the number of sessions and the fluence-map are re-optimized based on the latest cell density information. The chapter demonstrates via computer simulations on five head-and-neck test cases that such adaptive treatment-length and fluence-map planning reduces the TNTCR and increases the biological effect on the tumor while employing shorter treatment courses, as compared to only adapting fluence-maps and using a pre-determined treatment course length based on one-size-fits-all guidelines.

- **Chapter 5:** Adaptive fluence-map planning methods assume that the treatment planner knows the probability distribution of the uncertainty in the tumor's dose-response parameters. Chapter 5 proposes an alternative formulation where the treatment planner learns this distribution from tumor-response information observed in functional images over the treatment course while adaptively optimizing fluence-maps. This yields a Bayesian stochastic control formulation whose exact solution is impossible to derive. An algorithm rooted in certainty equivalent control is devised to simultaneously learn this probability distribution while adapting fluence-maps based on dose-response data collected from functional images over the treatment course. This algorithm's performance is compared via numerical simulations with two other solution procedures that are also rooted in certainty equivalent control. The first one is a clairvoyant

method. This assumes that the treatment planner knows the probability distribution, and hence serves as an idealized gold-standard. The other one uses a fixed value of the dose-response parameter as available from the literature, and hence provides a natural benchmark without learning. The tumor-kill achieved by the learning algorithm is statistically indistinguishable from the clairvoyant approach, whereas it is higher than the no learning benchmark. Both these conclusions bode well for individualized spatiobiologically integrated radiotherapy using functional images.

The dissertation concludes by outlining its weaknesses and by describing opportunities for future work.

Chapter 2
ROBUST FRACTIONATION

This chapter studies robust counterparts of the spatiobiologically separated (Section 2.1) and the spatiobiologically integrated (Section 2.2) formulations of the fractionation problem.

2.1 Robust spatiobiologically separated fractionation

2.1.1 Review of the nominal formulation

Key parameters of the LQ dose-response model include the so-called α/β ratios for the OAR. Optimization research that uses this LQ model has evolved from single-OAR formulations, to two-OAR formulations, and, more recently, to models with multiple OAR. All of these formulations belong to the class of non-convex quadratically constrained quadratic programs (QCQPs) — problems known to be computationally difficult in general. A closed-form optimal solution is available for the single OAR case (see, for example, [15, 26, 29, 34, 48, 67] and references therein). One paper provided an optimal dosing scheme using Karush-Kuhn-Tucker conditions for the two-OAR case for a fixed number of sessions [11]. A simulated annealing heuristic was applied to a two-OAR formulation in [75].

The most recent multiple-OAR formulation of this problem (see [54, 55]) is given by

$$\text{(FRAC)} \quad \max_{\vec{d},\ N} \quad \alpha_0 \sum_{t=1}^{N} d_t + \beta_0 \sum_{t=1}^{N} d_t^2 - \tau(N) \tag{2.1}$$

$$\sum_{t=1}^{N} d_t + \rho_m \sum_{t=1}^{N} d_t^2 \leq \text{BED}_m, \ m \in \mathcal{M}, \tag{2.2}$$

$$\vec{d} \geq 0, \tag{2.3}$$

$$1 \leq N \leq N_{\max}, \ \text{integer}. \tag{2.4}$$

The goal in this problem is to find the number of fractions N and a corresponding sequence $\vec{d} = (d_1, d_2, \ldots, d_N)$ of doses so as to maximize tumor-damage while ensuring that the OAR can safely tolerate these doses. Here, α_0, β_0 are the tumor's dose-response parameters as per the LQ model. The term $\tau(N)$ in the objective function accounts for tumor proliferation and is given by

$$\tau(N) = \frac{[(N-1) - T_{\text{lag}}]^+ \ln 2}{T_{\text{double}}}, \tag{2.5}$$

where $[(N-1) - T_{\text{lag}}]^+$ is defined as $\max\{0, (N-1) - T_{\text{lag}}\}$. Here, T_{lag} is the time-lag (in days) after which tumor proliferation starts after treatment initiation; and T_{double} (in days) denotes the doubling time for the tumor. This proliferation term assumes that a single fraction is administered every day; it can be generalized to accommodate other fractionation schemes as described in [54]. The objective function equals the biological effect (BE) of $\vec{d}$ on the tumor, which is to be maximized. In constraints (2.2), $\mathcal{M} = \{1, 2, \ldots, M\}$ is the set of $M \geq 1$ OAR. The parameter $\rho_m = \beta_m/\alpha_m$ is the aforementioned (inverse) ratio of dose-response parameters for OAR $m \in \mathcal{M}$. The left hand side of each constraint equals the biologically effective dose (BED) administered to the corresponding OAR. The term on the right hand side is given by $\text{BED}_m = D_m + \rho_m D_m^2/N_m$. It equals the BED of a conventional treatment schedule that administers a total dose of D_m in N_m equal-dosage fractions and that OAR m is known to tolerate. Thus, each of these constraints ensures that, for each OAR, the BED of $\vec{d}$ is no more than what is safe for that OAR. In constraint (2.4), $N_{\max}$ is the maximum number of fractions that is logistically feasible in the treatment protocol. In the sequel, we will often refer to (FRAC) as the nominal problem.

Note that the nominal formulation in [54, 55] employed the concept of sparing factors to model the doses delivered to the various OAR. For instance, if dose d_t is given to the tumor in fraction t, then the dose to OAR $m \in \mathcal{M}$ equals $s_m d_t$; here, s_m is a non-negative sparing factor. In this chapter, we do not use such sparing factors because they are distracting to the main message of this work. They can be easily incorporated into our formulation as explained later in Section 2.1.7.

2.1.2 Optimal solution of the nominal formulation

An optimal solution for this multiple-OAR case was provided in [54]; this solution works either when $\alpha_0/\beta_0 \leq \min_{m\in\mathcal{M}} (\alpha_m/\beta_m)$ or when $\alpha_0/\beta_0 \geq \max_{m\in\mathcal{M}} (\alpha_m/\beta_m)$. The first provably optimal solution that works irrespective of the ordering of these ratios for the multiple-OAR case was recently derived in [55] based on the doctoral dissertation of Saberian [53]. This solution was obtained by equivalently reformulating (FRAC) for each fixed N as a *two-variable* linear program (LP) with $n+2$ linear constraints and non-negativity constraints on the two variables. The two variables in this LP are $x = \sum_{t=1}^{N} d_t$ and $y = \sum_{t=1}^{N} d_t^2$ and the LP is given by

$$(\text{2VARLP}) \quad \max_{x,y} \ \alpha_0 x + \beta_0 y - \tau(N) \tag{2.6}$$

$$x + \rho_m y \leq \text{BED}_m, \ m \in \mathcal{M}, \tag{2.7}$$

$$y \leq \gamma^* x, \tag{2.8}$$

$$c^* x \leq y, \tag{2.9}$$

$$x \geq 0, \tag{2.10}$$

$$y \geq 0, \tag{2.11}$$

where $\gamma^* = \min_{m\in\mathcal{M}} b_m(1)$ and $c^* = \min_{m\in\mathcal{M}} b_m(N)$ with $b_m(N) = \frac{-1+\sqrt{1+4\rho_m \text{BED}_m/N}}{2\rho_m}$ for $m \in \mathcal{M}$ and for all $N \geq 1$. Specifically, for each fixed N, if x^*, y^* is an optimal solution of this LP, then the dosing schedule $(q, \underbrace{p, p, \ldots, p}_{N-1 \text{ times}})$, where

$$p = \frac{x^*}{N}\left[1 - \sqrt{1 - \left(1 - \frac{y^*}{(x^*)^2}\right)\left(\frac{N}{N-1}\right)}\right], \tag{2.12}$$

$$q = x^* - (N-1)p, \tag{2.13}$$

is optimal. Moreover, it can be shown that there are only three possibilities for x^* and y^*. The first is where $\sqrt{y^*} = x^*$ and then $p = 0$ (this is called a single-dosage solution); the second is where $\sqrt{Ny^*} = x^*$ and then $p = q$ (this is called an equal-dosage solution); and the third is where $\sqrt{y^*} < x^* < \sqrt{Ny^*}$ and then $0 \neq p \neq q \neq 0$ (this is called an unequal-dosage solution) (see [53, 55] for details). An optimal number of fractions can then be found by substituting a dosing schedule so obtained into the objective function in (FRAC) for each $N \in \{1, 2, \ldots, N_{\max}\}$ and picking the one that yields the largest tumor BE. Consequently, (FRAC) is solved by solving exactly $N_{\max}$ two-variable LPs.

2.1.3 Limitations of the nominal formulation

One drawback of the nominal formulation is that the values of ρ_m are not known. Thus, a dosing schedule derived using estimated or "nominal" values of these parameters may not even be feasible in practice. In an unpublished manuscript [5], Badri et al., independently of our work, attempted to remedy this by studying a robust formulation of the above fractionation problem. In their formulation, the treatment planner derives a robust solution by assuming that the ρ_m values vary within a known non-negative interval. However, the crucial dependence of the right hand side BED_m on ρ_m in constraints (2.2) was ignored in that manuscript. This meant that an optimal solution to their robust formulation was obtained by replacing ρ_m on the left hand side in (2.2) by its largest possible value. This implied that the robust solution is derived simply by solving the two-variable LP in [53, 55]. Unfortunately, since the right hand side in constraints (2.2) in fact explicitly depends on ρ_m, such a simplified solution might not be robust in practice. Again, independently from our work, Badri et al. [4] corrected this later and developed a computationally demanding chance constrained approach to accommodate uncertainty in the α/β ratios in the separated model [6].

In this chapter, we present a robust optimization approach to overcome the aforementioned limitations of existing research on optimal fractionation based on the LQ model. Specifically, we formulate our robust fractionation problem assuming that the unknown pa-

rameters ρ_m belong to known intervals. We show that, for each fixed N, an optimal solution to the non-convex robust problem can be obtained by solving $n+1$ two-variable LPs. Consequently, the robust fractionation problem is solved by solving $(n+1)N_{\max}$ two-variable LPs; each of these LPs includes $n+4$ linear constraints and non-negativity constraints on the two variables. We perform extensive sensitivity analyses with respect to the values of T_{lag} and T_{double} currently available in the clinical literature to numerically quantify the price of robustness. We also provide a comparison between the nominal optimal fractionation schedules and the robust optimal fractionation schedules derived from our approach.

2.1.4 A robust formulation

We refer the reader to [8] for a textbook and to [10] for a survey on robust optimization. We employ a standard interval uncertainty model from these existing works to construct a robust counterpart of the nominal problem (FRAC). Specifically, we use $\tilde{\rho}_m$ to denote the "true" unknown value of ρ_m, for $m = 1, \ldots, n$. We assume that this unknown value belongs to a known interval of values $[\rho_m^{\min}, \rho_m^{\max}]$; here $0 < \rho_m^{\min} \leq \rho_m^{\max} < \infty$. We wish to find an $N, \vec{d}$ pair that is feasible to BED constraints (2.2) for all $m \in \mathcal{M}$ no matter what true values $\tilde{\rho}_m$ are realized (as long as they belong to the aforementioned intervals). The resulting robust counterpart of (FRAC) is given by

$$\max_{\vec{d},\ N} \ \alpha_0 \sum_{t=1}^{N} d_t + \beta_0 \sum_{t=1}^{N} d_t^2 - \tau(N) \tag{2.14}$$

$$\sum_{t=1}^{N} d_t + \tilde{\rho}_m \left(\sum_{t=1}^{N} d_t^2 - \frac{D_m^2}{N_m} \right) \leq D_m, \quad m \in \mathcal{M},\ \forall \tilde{\rho}_m \in [\rho_m^{\min}, \rho_m^{\max}], \tag{2.15}$$

$$\vec{d} \geq 0, \tag{2.16}$$

$$1 \leq N \leq N_{\max},\ \text{integer}. \tag{2.17}$$

Note here that, for simplicity of exposition, our formulation does not consider uncertainty in the values of α_0 and β_0 for the tumor. However, this can be easily incorporated by

maximizing the worst-case value of the objective function. This approach of not including uncertainty in the objective function coefficients is standard in robust optimization.

To write a compact robust formulation, it is convenient to normalize the intervals $[\rho_m^{\min}, \rho_m^{\max}]$, for $m \in \mathcal{M}$. That is, we express $\tilde{\rho}_m$ as $\rho_m^{\text{mean}} + \eta_m \rho_m^{\text{range}}$, where $\eta_m \in [-1, 1]$ is a real number; $\rho_m^{\text{mean}} = (\rho_m^{\max} + \rho_m^{\min})/2$ and $\rho_m^{\text{range}} = (\rho_m^{\max} - \rho_m^{\min})/2$. This allows us to rewrite constraint (2.15) as follows. Fix any OAR $m \in \mathcal{M}$. Then,

$$
\begin{aligned}
&\sum_{t=1}^{N} d_t + \tilde{\rho}_m \left(\sum_{t=1}^{N} d_t^2 - \frac{D_m^2}{N_m} \right) \leq D_m, \ \forall \tilde{\rho}_m \in [\rho_m^{\min}, \rho_m^{\max}] \\
&\Leftrightarrow \sum_{t=1}^{N} d_t + (\rho_m^{\text{mean}} + \eta_m \rho_m^{\text{range}}) \left(\sum_{t=1}^{N} d_t^2 - \frac{D_m^2}{N_m} \right) \leq D_m, \ \forall \eta_m \in [-1, 1] \\
&\Leftrightarrow \sum_{t=1}^{N} d_t + \rho_m^{\text{mean}} \left(\sum_{t=1}^{N} d_t^2 - \frac{D_m^2}{N_m} \right) + \eta_m \rho_m^{\text{range}} \left(\sum_{t=1}^{N} d_t^2 - \frac{D_m^2}{N_m} \right) \leq D_m, \ \forall \eta_m \in [-1, 1] \\
&\Leftrightarrow \sum_{t=1}^{N} d_t + \rho_m^{\text{mean}} \left(\sum_{t=1}^{N} d_t^2 - \frac{D_m^2}{N_m} \right) + \rho_m^{\text{range}} \left| \sum_{t=1}^{N} d_t^2 - \frac{D_m^2}{N_m} \right| \leq D_m \\
&\Leftrightarrow \sum_{t=1}^{N} d_t + \rho_m^{\text{mean}} \sum_{t=1}^{N} d_t^2 + \rho_m^{\text{range}} \left| \sum_{t=1}^{N} d_t^2 - \frac{D_m^2}{N_m} \right| \leq D_m + \rho_m^{\text{mean}} \frac{D_m^2}{N_m}.
\end{aligned}
$$

Thus, by defining the shorthand notation $\text{RC}_m = D_m + \rho_m^{\text{mean}} \frac{D_m^2}{N_m}$, and putting the above pieces together, we can rewrite the robust counterpart (2.14)-(2.17) as

$$
\text{(RFRAC)} \ f^* = \max_{\vec{d},\ N} \ \alpha_0 \sum_{t=1}^{N} d_t + \beta_0 \sum_{t=1}^{N} d_t^2 - \tau(N) \tag{2.18}
$$

$$
\sum_{t=1}^{N} d_t + \rho_m^{\text{mean}} \sum_{t=1}^{N} d_t^2 + \rho_m^{\text{range}} \left| \sum_{t=1}^{N} d_t^2 - \frac{D_m^2}{N_m} \right| \leq \text{RC}_m, \ m \in \mathcal{M}, \tag{2.19}
$$

$$
\vec{d} \geq 0, \tag{2.20}
$$

$$
1 \leq N \leq N_{\max}, \ \text{integer}. \tag{2.21}
$$

As in the nominal problem, in order to solve this robust problem, we first solve the problems obtained by fixing N at $1, 2, \ldots, N_{\max}$. For each fixed N, let $\vec{d}^*(N) = (d_1^*(N), \ldots, d_N^*(N))$

denote the corresponding optimal dosing sequence. We then compare the objective values of these N dosing sequences and pick the best. Thus, the problem we need to solve for each fixed $N \in \{1, 2, \ldots, N_{\max}\}$ is given by

$$\text{(RFRAC(N))} \; f^*(N) = \max_{\vec{d}} \;\; \alpha_0 \sum_{t=1}^{N} d_t + \beta_0 \sum_{t=1}^{N} d_t^2 - \tau(N) \tag{2.22}$$

$$\sum_{t=1}^{N} d_t + \rho_m^{\text{mean}} \sum_{t=1}^{N} d_t^2 + \rho_m^{\text{range}} \left| \sum_{t=1}^{N} d_t^2 - \frac{D_m^2}{N_m} \right| \leq \text{RC}_m, \; m \in \mathcal{M}, \tag{2.23}$$

$$\vec{d} \geq 0. \tag{2.24}$$

Note that when $\rho_m^{\min} = \rho_m^{\max} = \rho_m$, for $m = 1, 2, \ldots, n$, that is, when there is no uncertainty in these dose-response parameters, (RFRAC(N)) reduces to the nominal QCQP (FRAC) with N fixed as presented in Section 2.1.2, and which was solved recently as a two-variable LP in [53, 55]. Note, however, that the objective function as well as the constraints in (RFRAC(N)) are non-convex, and the problem is at least as hard as the nominal QCQP. The objective function in the nominal QCQP is identical in form to what we have in (RFRAC(N)), but the convex, quadratic constraints in the nominal QCQP do not include the absolute value term that appears in the corresponding constraints in (RFRAC(N)). Specifically, it is this absolute value term that makes the robust counterpart harder to solve as compared to the nominal problem. To overcome this challenge, we decompose the feasible region of (RFRAC(N)) into $n + 1$ subregions in a way such that the problem over each subregion can be solved via a two-variable LP. The details of this procedure are discussed in the next section.

2.1.5 Optimal solution of the robust formulation

To handle the absolute value term on the left hand side in constraints (2.23), we decompose the non-negative orthant $\{\vec{d} \in \Re^N | \vec{d} \geq 0\}$ as follows. For each OAR $m \in \mathcal{M}$, consider two possibilities: the first is where $\sum_{t=1}^{N} d_t^2 \geq D_m^2/N_m$ and the second is $\sum_{t=1}^{N} d_t^2 < D_m^2/N_m$. Suppose, in the rest of this section, without loss of generality that $D_1^2/N_1 \leq D_2^2/N_2 \leq \ldots \leq D_n^2/N_n$.

Then, if there is a $\vec{d} \geq 0$ and an OAR $m \in \mathcal{M}$ such that $\sum_{t=1}^{N} d_t^2 \geq D_m^2/N_m$, then for this $\vec{d}$, we have that $\sum_{t=1}^{N} d_t^2 \geq D_{m'}^2/N_{m'}$ for all $m' < m$. Similarly, if there is a $\vec{d} \geq 0$ and an OAR $m \in \mathcal{M}$ such that $\sum_{t=1}^{N} d_t^2 < D_m^2/N_m$, then for this $\vec{d}$, we have that $\sum_{t=1}^{N} d_t^2 < D_{m'}^2/N_{m'}$ for all $m' > m$. This means that the non-negative orthant $\{\vec{d} \in \Re^N | \vec{d} \geq 0\}$ is partitioned into $n+1$ subregions indexed by $k = 0, 1, 2, \ldots, n$. In the kth region, $\sum_{t=1}^{N} d_t^2 \geq D_m^2/N_m$ for the ***first*** k OAR and $\sum_{t=1}^{N} d_t^2 < D_m^2/N_m$ for the ***last*** $n-k$ OAR. Let $\mathrm{RC}^+ = D_m + \rho_m^{\max} \frac{D_m^2}{N_m}$ and $\mathrm{RC}^- = D_m + \rho_m^{\min} \frac{D_m^2}{N_m}$. Then, simple algebra reveals that for all $\vec{d}$ in the kth subregion, constraint (2.23) reduces to $\sum_{t=1}^{N} d_t + \rho_m^{\max} \sum_{t=1}^{N} d_t^2 \leq \mathrm{RC}_m^+$ for OAR $m = 1, 2, \ldots, k$ when $k \neq 0$; and it reduces to $\sum_{t=1}^{N} d_t + \rho_m^{\min} \sum_{t=1}^{N} d_t^2 \leq \mathrm{RC}_m^-$ for OAR $k+1, k+2, \ldots, n$ when $k \neq n$.

As a result of the above discussion, (RFRAC(N)) is solved by solving $n+1$ subproblems and then picking a dosing schedule with the largest tumor BE from the resulting $n+1$ solutions. The kth subproblem is given by

$$(\mathrm{kSub(N)}) \quad \max_{\vec{d}} \quad \alpha_0 \sum_{t=1}^{N} d_t + \beta_0 \sum_{t=1}^{N} d_t^2 - \tau(N) \tag{2.25}$$

$$\sum_{t=1}^{N} d_t + \rho_m^{\max} \sum_{t=1}^{N} d_t^2 \leq \mathrm{RC}_m^+, \ m = 1, 2, \ldots, k, \ k \neq 0, \tag{2.26}$$

$$\sum_{t=1}^{N} d_t + \rho_m^{\min} \sum_{t=1}^{N} d_t^2 \leq \mathrm{RC}_m^-, \ m = k+1, k+2, \ldots, n, \ k \neq n, \tag{2.27}$$

$$\sum_{t=1}^{N} d_t^2 \geq \frac{D_m^2}{N_m}, \ m = 1, 2, \ldots, k, \ k \neq 0, \tag{2.28}$$

$$\sum_{t=1}^{N} d_t^2 < \frac{D_m^2}{N_m}, \ m = k+1, k+2, \ldots, n, \ k \neq n, \tag{2.29}$$

$$\vec{d} \geq 0. \tag{2.30}$$

In addition, owing to the fact that $D_1^2/N_1 \leq D_2^2/N_2 \leq \ldots \leq D_n^2/N_n$, the group of n

constraints in (2.28)-(2.29) reduces to at most two constraints: $\sum_{t=1}^{N} d_t^2 \geq \frac{D_k^2}{N_k}$ when $k \neq 0$ and $\sum_{t=1}^{N} d_t^2 < \frac{D_{k+1}^2}{N_{k+1}}$ when $k \neq n$. After replacing this second strict inequality with a non-strict inequality[1], this simplifies the kth subproblem to

$$\text{(kSub(N))}\ f^*(N;k) = \max_{\vec{d}}\ \alpha_0 \sum_{t=1}^{N} d_t + \beta_0 \sum_{t=1}^{N} d_t^2 - \tau(N) \tag{2.31}$$

$$\sum_{t=1}^{N} d_t + \rho_m^{\max} \sum_{t=1}^{N} d_t^2 \leq \text{RC}_m^+,\ m = 1, 2, \ldots, k,\ k \neq 0, \tag{2.32}$$

$$\sum_{t=1}^{N} d_t + \rho_m^{\min} \sum_{t=1}^{N} d_t^2 \leq \text{RC}_m^-,\ m = k+1, k+2, \ldots, n,\ k \neq n, \tag{2.33}$$

$$\sum_{t=1}^{N} d_t^2 \geq \frac{D_k^2}{N_k},\ k \neq 0, \tag{2.34}$$

$$\sum_{t=1}^{N} d_t^2 \leq \frac{D_{k+1}^2}{N_{k+1}},\ k \neq n, \tag{2.35}$$

$$\vec{d} \geq 0. \tag{2.36}$$

The objective function and the constraints (2.32)-(2.33) in this subproblem are identical in form to that in the nominal problem (FRAC) with N fixed. Thus, the only difference between this subproblem and the nominal problem is the appearance of the additional constraints (2.34)-(2.35). Fortunately, as in the nominal problem studied in [53, 55], we are still able to use the variable transformation $x = \sum_{t=1}^{N} d_t$ and $y = \sum_{t=1}^{N} d_t^2$ to convert this subproblem into an equivalent two-variable LP. To write this LP compactly, we first introduce additional

[1]This can be rigorously justified by proving that if there is a feasible dosing schedule that satisfies $\sum_{t=1}^{N} d_t^2 = \frac{D_{k+1}^2}{N_{k+1}}$ in the kth subproblem with a non-strict inequality, then this dosing schedule is feasible to the $k+1$st subproblem with a strict inequality; consequently, using non-strict inequalities does not alter optimality in our overall group of $n+1$ subproblems with strict inequalities. We omit the details of this proof for brevity.

notation. Let

$$\mathrm{RC}_m^k = \begin{cases} \mathrm{RC}_m^+, \text{ for } m = 1, 2, \ldots, k, \ k \neq 0, \\ \mathrm{RC}_m^-, \text{ for } m = k+1, k+2, \ldots, n, \ k \neq n; \end{cases} \tag{2.37}$$

and similarly,

$$\rho_m^k = \begin{cases} \rho_m^{\max}, \text{ for } m = 1, 2, \ldots, k, \ k \neq 0, \\ \rho_m^{\min}, \text{ for } m = k+1, k+2, \ldots, n, \ k \neq n. \end{cases} \tag{2.38}$$

Moreover, let

$$c_k = \min_{m \in \mathcal{M}} \ \frac{-1 + \sqrt{1 + 4\rho_m^k \mathrm{RC}_m^k / N}}{2\rho_m^k}, \text{ and} \tag{2.39}$$

$$\gamma_k = \min_{m \in \mathcal{M}} \ \frac{-1 + \sqrt{1 + 4\rho_m^k \mathrm{RC}_m^k}}{2\rho_m^k}. \tag{2.40}$$

Then, the two-variable LP can be written as

$$(\text{2VARLPkSub(N)}) \quad \max_{x,y} \ \alpha_0 x + \beta_0 y - \tau(N) \tag{2.41}$$

$$x + \rho_m^k y \leq \mathrm{RC}_m^k, \ m \in \mathcal{M}, \tag{2.42}$$

$$y \geq \frac{D_k^2}{N_k}, \ k \neq 0, \tag{2.43}$$

$$y \leq \frac{D_{k+1}^2}{N_{k+1}}, \ k \neq n, \tag{2.44}$$

$$y \leq \gamma_k x, \tag{2.45}$$

$$c_k x \leq y, \tag{2.46}$$

$$x \geq 0, \tag{2.47}$$

$$y \geq 0. \tag{2.48}$$

The reason why this transformation works is the same as that provided for the nominal problem in [53, 55] — the constraints $x = \sum_{t=1}^{N} d_t$ and $y = \sum_{t=1}^{N} d_t^2$ that *should* be present

in the transformed problem can be replaced by the two linear constraints (2.45)-(2.46) without altering optimality. Furthermore, as stated in Section 2.1.2 for the nominal problem, if x^*, y^* is optimal to (2VARLPkSub(N)), a corresponding optimal dosing schedule $\vec{d}^*(N) = (d_1^*(N), d_2^*(N), \ldots, d_N^*(N)) = (q, \underbrace{p, p, \ldots, p}_{N-1 \text{ times}})$ is recovered by formulas (2.12)-(2.13). The discussion in this section thus provides a procedure, summarized in Algorithm 1 below, to obtain an exact solution of the non-convex robust counterpart (RFRAC).

Algorithm 1 An algorithm for finding an optimal solution to (RFRAC)

1: **for** $(N = 1, 2, \ldots, N_{\max})$ **do**
2: **for** $(k = 0, 1, 2, \ldots, n)$ **do**
3: Solve (2VARLPkSub(N)) to obtain its optimal solution (x_k^*, y_k^*).
4: Recover $\vec{d}^*(N; k) = (q_k, \underbrace{p_k, p_k, \ldots, p_k}_{N-1 \text{ times}})$ from (x_k^*, y_k^*) using formulas (2.12)-(2.13).
5: Compute the objective value $f^*(N; k)$ of $\vec{d}^*(N; k)$ in problem (kSub(N)).
6: **end for**
7: Find the optimal value in problem (RFRAC(N)) as $f^*(N) = \max\limits_k \ f^*(N; k)$ and the corresponding optimal dosing schedule $\vec{d}^*(N) = \vec{d}^*(N; k(N))$, where $k(N) = \operatorname{argmax}\limits_k \ f^*(N; k)$.
8: **end for**
9: Find the optimal value in problem (RFRAC) as $f^* = \max\limits_N \ f^*(N)$ and the corresponding optimal dosing schedule $\vec{d}^* = \vec{d}^*(N^*)$, where $N^* = \operatorname{argmax}\limits_N \ f^*(N)$.
10: Report $N^*, \vec{d}^*$ as an optimal solution to (RFRAC).

2.1.6 Numerical experiments

2.1.6.1 Description of the head-and-neck cancer case

We evaluate the preformance of our robust counterpart on a representative head-and-neck cancer case with spinal cord, brainstem, left and right parotids as the OARs. The nominal values for OAR-response parameter ρ (Gy^{-1}) for spinal cord, brainstem, left and right parotids were fixed at $1/3, 1/4, 1/5$ and $1/6$; respectively. These were based on the values

reported in [23, 25, 75]. We used different values of nominal ratios to fully explore the various possibilities that could arise in a robust formulation with multiple OAR.

Tumor-response parameters were set at $\alpha_0 = 0.35$ Gy^{-1} and $\beta_0 = 0.035$ Gy^{-2} based on [23, 24, 25, 26]. Maximum dose constraints of 45 Gy and 50 Gy were included for spinal cord and brainstem, respectively. The tolerance dose for mean dose constraints was set at 28 Gy for left and right parotids. The conventional number of fractions N_m was fixed at 35 days for all OAR. These tolerance dose and conventional fractionation numbers were similar to the standard QUANTEC treatment protocol [47]. $N_{\max}$ was set to 100 days. The reason behind using this unrealistically high number was to explore the full range of possibilities that could arise in our sensitivity analysis. All experiments were conducted in MATLAB on a laptop with 2.20 GHz Intel Core2 Duo CPU and 2 GB of memory, running a Microsoft Windows 8.1 operating system.

2.1.6.2 Price of robustness

The uncertainty intervals were parameterized as $\tilde{\rho}_m \in [(1-\delta)\rho_m, (1+\delta)\rho_m]$, where $\delta \in [0, 1]$. This allowed us to quantify the price of robustness as a function of a single uncertainty parameter δ. We varied δ from 0 (corresponding to the nominal case) to 1 (to represent 100% uncertainty) in increments of 0.1. T_{lag} values were set to $7, 14, 21, 28, 35$ days based on [25] and T_{double} values were set to $2, 8, 10, 20, 40, 50, 80, 100$ days based on [23, 25, 49, 75]. We are aware that the using 35 days as T_{lag} and 80 and 100 for T_{double} are unrealistically high for head-and-neck cancer. We nevertheless decided to include these values to fully explore possible trends that could arise in our numerical analysis. The results of our experiments are reported in Tables 2.1 and 2.2. In the remainder of this section, we only provide the results for $T_{\text{lag}} = 7$ and 14 to save space and avoid overcomplicating the tables.

Table 2.1 shows, as expected, that the price of robustness increases with increasing δ for each T_{double}. Overall, the price of robustness seems to be quite small in most experiments with an average of 1.27% over all 400 experiments. The first, second, and third quartiles were 0.12%, 0.47%, and 1.44%; respectively.

For each δ and T_{lag},in Table 2.1, the price of robustness first decreases with increasing T_{double}, reaches the smallest value when $T_{\text{double}} = 50$ days and then increases. This trend is consistent with the corresponding trend in the difference between $N_m = 35$ and N^* that can be inferred from Table 2.2. Specifically, for each T_{lag}, δ combination, the magnitude of $N_m - N^*$ decreases with increasing T_{double}, reaches about a day or two when $T_{\text{double}} = 50$, and then increases. In fact, as we can see in Figure 2.1(d), when $N = N_m = 35$, the price of robustness is exactly zero; more strongly, we found that this held true irrespective of the values of δ, T_{lag}, and T_{double}. A detailed algebraic proof of this fact can be developed, but is omitted here for brevity. Roughly speaking, the key idea in this proof is that when $N = N_m$, the BED constraints reduce to total dose constraints; this eliminates the dependence of the BED constraint on ρ_m and hence an equal-dosage solution that splits the tolerance dose across N fractions is optimal to the nominal *and* the robust problem. Consequently, the price of robustness is zero. Finally, for any combination of δ and T_{double} in Table 2.1, the price of robustness decreases as T_{lag} increases. Again, this is also consistent with the corresponding trend in the magnitude of $N_m - N^*$.

A closer look at Table 2.2 reveals that the evolution of N^* with δ for various fixed combinations of T_{double} and T_{lag} does not exhibit a universal trend. For instance, when $T_{\text{lag}} = 7$ days and $T_{\text{double}} = 2$ days, $N^* = 8$ for all δ (also see Figure 2.1 (a)). However, N^* increases with increasing δ when $T_{\text{lag}} = 7$ days and $T_{\text{double}} = 10$ days (Figure 2.1 (b)). On the other hand, N^* decreases as δ increases when $T_{\text{lag}} = 7$ days and $T_{\text{double}} = 100$ days (Figure 2.1 (c)). Consistent with this observation, Table 2.2 also shows that for each fixed combination of T_{double} and T_{lag}, optimal doses do not exhibit a universal qualitative trend as a function of δ.

Finally, our robust solutions continue to exhibit qualitative trends that are well-established in the nominal case (see [54] and references therein). For instance, N^* increases with increasing T_{double} for any fixed δ, T_{lag} combination; similarly, N^* also increases as T_{lag} increases for any fixed δ, T_{double} combination.

2.1.6.3 Infeasibility tests

The nominal solution is only guaranteed to be feasible when the realized value of ρ equals the nominal ρ. To quantify the frequency and the amount of infeasibility, we tested the nominal solution where the realized values of ρ were different from the nominal. We tried five different realized values of ρ at grid points within the uncertainty interval corresponding to each δ. The five values were $\tilde{\rho}_i^m = \rho_{\min}^m + 2i\rho_m^i\delta/5$, where $i = 1, 2, \ldots, 5$, and ρ_m^i denotes the nominal value ($1/3$ Gy^{-1}) for constraint m. Subsequently, the infeasibility test was performed for all combination of T_{lag}, T_{double}, and δ. The results of the test showed that when ρ the nominal solution was infeasible in over 75% of the experiments. Recall that the robust solution is feasible in all these cases (since the realized values were chosen within the uncertainty interval). The amount of this infeasibility in some cases is as high as 50%.

So far, everywhere in our formulations and numerical experiments, we have assumed that the ρ parameters for each constraint belonged to a presumed uncertainty set and all the realizations of those parameters come from within that set. The robust solution obtained based on this assumption is therefore guaranteed to be feasible only as long as this assumption holds. To quantify the frequency and the amount of infeasibility of the robust and nominal solutions when this assumption is violated, we repeated the above experiments after a slight modification to the choice of realized parameters. Specifically, for each uncertainty level δ and constraint m, we chose five different grid-points located at $\tilde{\rho}_m = (1 + \delta + \gamma)\rho_m$; as well as another five grid-points at $\tilde{\rho}_m = (1 - \delta - \gamma)\rho_m$, where $\gamma \in \{0.1, 0.2, \ldots, 0.5\}$ and ρ_m denotes the nominal value of $1/3$ Gy^{-1}. After performing numerical experiments for all combinations of T_{lag}, T_{double}, and δ, the nominal solution turned out to be infeasible in 65% of the experiments, while the robust solutions were infeasible in 43% of the cases. The difference between the amount of infeasibility of the robust and nominal solutions was deemed significant at the 95% confidence level by a pairwise $t-$test.

2.1.6.4 Uncertainty in tumor parameters

Throughout this chapter, we have assumed that there is no uncertainty in the values of the tumor-response parameters α_0 and β_0. To numerically quantify the effect of the uncertainty in the tumor-response parameters, we assume that both $\tilde{\alpha}_0$ and $\tilde{\beta}_0$ belong to uncertainty sets $\tilde{\alpha}_0 \in [\alpha_0^{\min}, \alpha_0^{\max}]$, $\tilde{\beta}_0 \in [\beta_0^{\min}, \beta_0^{\max}]$. In this case, the robust counterpart (with respect to uncertainty in $\tilde{\alpha}_0$ and $\tilde{\beta}_{00}$) is obtained simply by replacing $\tilde{\alpha}_0$ with $\alpha_0^{\min}$ and $\tilde{\beta}_0$ with $\beta_0^{\min}$.

To characterize the interval uncertainty in both tumor-response parameters, we used a single parameter $\theta \in 0.1, 0.2, \ldots, 0.9$ for simplicity. That is, the uncertainty is written as $\tilde{\alpha}_0 \in [(1-\theta)\hat{\alpha}_0, (1+\theta)\hat{\alpha}_0]$ and $\tilde{\beta}_0 \in [(1-\theta)\hat{\beta}_0, (1+\theta)\hat{\beta}_0]$, where the nominal values $\hat{\alpha}_0$ and $\hat{\beta}_0$ were set to 0.35 Gy^{-1} and 0.035 Gy^{-2}, respectively. Table 2.3 shows the effect of this uncertainty on the robust solutions for all combinations of T_{lag}, T_{double} and θ.

2.1.7 Conclusions

Existing research on robust optimization in cancer radiotherapy focuses on incorporating uncertainty in the actual dose delivered to various anatomical regions of interest via intensity modulated radiation therapy (IMRT) and other treatment methods. Causes of this uncertainty include patient movement, say due to breathing, or setup errors at the time of treatment delivery (see, for instance, [14, 19, 44, 45, 46, 66] and references therein). We provided a robust formulation of the fractionation problem in radiotherapy. Perhaps more importantly, we also provided a simple method for exact solution of this robust formulation. Although our robust formulation is, at first glance, inevitably at least as hard as a non-convex QCQP, we were able to show that it can be solved to optimality by solving a few two-variable LPs with a few constraints. Our numerical experiments provided insights into the behavior of robust optimal dosing schedules and also quantified the price of robustness. Overall, it appears that the price of robustness is small.

We emphasize that our solution procedure in Section 2.1.5 would work even if sparing factors were included in the nominal model. More strongly, our solution method would work

even if the true values of these sparing factors were unknown but were instead assumed to belong to a non-negative interval. This can be done simply by using the largest values of these sparing factors in our robust formulation.

$T_{\text{lag}} = 7$	T_{double}							
δ	2	8	10	20	40	50	80	100
0.1	1.74	1.44	1.21	0.56	0.04	0.01	0.25	0.36
0.2	3.34	2.67	2.23	0.98	0.04	0.01	0.39	0.62
0.3	4.80	3.74	3.10	1.31	0.04	0.01	0.39	0.70
0.4	6.14	4.66	3.85	1.56	0.04	0.01	0.39	0.70
0.5	7.38	5.48	4.51	1.74	0.04	0.01	0.39	0.70
0.6	8.53	6.22	5.09	1.87	0.04	0.01	0.39	0.70
0.7	9.61	6.88	5.61	1.95	0.04	0.01	0.39	0.70
0.8	10.61	7.47	6.07	2.00	0.04	0.01	0.39	0.70
0.9	11.54	8.01	6.47	2.02	0.04	0.01	0.39	0.70
1	12.42	8.50	6.84	2.02	0.04	0.01	0.39	0.70
$T_{\text{lag}} = 14$	T_{double}							
δ	2	8	10	20	40	50	80	100
0.1	0.92	0.92	0.92	0.54	0.04	0.01	0.25	0.36
0.2	1.78	1.78	1.78	0.96	0.04	0.01	0.38	0.61
0.3	2.59	2.59	2.59	1.28	0.04	0.01	0.39	0.70
0.4	3.34	3.34	3.30	1.52	0.04	0.01	0.39	0.70
0.5	4.05	4.05	3.93	1.69	0.04	0.01	0.39	0.70
0.6	4.71	4.71	4.48	1.82	0.04	0.01	0.39	0.70
0.7	5.34	5.34	4.97	1.90	0.04	0.01	0.39	0.70
0.8	5.93	5.89	5.41	1.95	0.04	0.01	0.39	0.70
0.9	6.49	6.41	5.80	1.96	0.04	0.01	0.39	0.70
1	7.02	6.87	6.14	1.96	0.04	0.01	0.39	0.70

Table 2.1: The price of robustness (%) for different combinations of T_{double}, and δ for $T_{\text{lag}} = 7$ and 14. The price of robustness equals $(\frac{g^* - f^*}{g^*}) \times 100\%$, where f^* and g^* are the optimal values of the robust and the nominal formulations, respectively.

$T_{\text{lag}} = 7$	T_{double}							
δ	2	8	10	20	40	50	80	100
0	(2.49, 8)	(2.10,10)	(1.82,12)	(1.20,20)	(0.80,32)	(0.71,37)	(0.55,49)	(0.49,56)
0.1	(2.46, 8)	(1.93,11)	(1.69,13)	(1.11,22)	(0.74,35)	(0.74,35)	(0.62,43)	(0.55,49)
0.2	(2.42, 8)	(1.79,12)	(1.59,14)	(1.02,24)	(0.74,35)	(0.74,35)	(0.71,37)	(0.63,42)
0.3	(2.39, 8)	(1.67,13)	(1.49,15)	(0.96,26)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
0.4	(2.36, 8)	(1.65,13)	(1.41,16)	(0.92,27)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
0.5	(2.34, 8)	(1.55,14)	(1.34,17)	(0.87,29)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
0.6	(2.31, 8)	(1.46,15)	(1.27,18)	(0.82,31)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
0.7	(2.29, 8)	(1.45,15)	(1.21,19)	(0.80,32)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
0.8	(2.27, 8)	(1.38,16)	(1.21,19)	(0.76,34)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
0.9	(2.25, 8)	(1.31,17)	(1.16,20)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
1	(2.23, 8)	(1.30,17)	(1.11,21)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
$T_{\text{lag}} = 14$	T_{double}							
δ	2	8	10	20	40	50	80	100
0	(1.53,15)	(1.53,15)	(1.53,15)	(1.20,20)	(0.80,32)	(0.71,37)	(0.55,49)	(0.49,56)
0.1	(1.51,15)	(1.51,15)	(1.51,15)	(1.11,22)	(0.74,35)	(0.74,35)	(0.62,43)	(0.55,49)
0.2	(1.50,15)	(1.50,15)	(1.50,15)	(1.02,24)	(0.74,35)	(0.74,35)	(0.71,37)	(0.63,42)
0.3	(1.49,15)	(1.49,15)	(1.49,15)	(0.96,26)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
0.4	(1.48,15)	(1.48,15)	(1.41,16)	(0.92,27)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
0.5	(1.47,15)	(1.47,15)	(1.34,17)	(0.87,29)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
0.6	(1.46,15)	(1.46,15)	(1.27,18)	(0.82,31)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
0.7	(1.45,15)	(1.45,15)	(1.21,19)	(0.80,32)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
0.8	(1.45,15)	(1.38,16)	(1.21,19)	(0.76,34)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
0.9	(1.44,15)	(1.31,17)	(1.16,20)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)
1	(1.43,15)	(1.30,17)	(1.11,21)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)	(0.74,35)

Table 2.2: The optimal solution (d^*, N^*) of the robust and nominal models for different combinations of T_{double}, and δ for $T_{\text{lag}} = 7, 14$. The first row in each section of the table shows the solution of the nominal case ($\delta = 0$). The cases marked with an asterisk ($*$) yield unequal-dosage solutions characterized by two dose values (q, p) (recall formulas (2.12)-(2.13)) and an N^* value in that order.

$T_{double} = 8$											
$T_{lag} = 7$	θ										
δ	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	(2.10,10)	(2.28, 9)	(2.49, 8)	(2.49, 8)	(2.49, 8)	(2.49, 8)	(2.49, 8)	(2.49, 8)	(2.49, 8)	(2.49, 8)	(8.91, 8)
0.1	(1.93,11)	(2.08,10)	(2.25, 9)	(2.46, 8)	(2.46, 8)	(2.46, 8)	(2.46, 8)	(2.46, 8)	(2.46, 8)	(2.46, 8)	(8.75, 8)
0.2	(1.79,12)	(1.91,11)	(2.05,10)	(2.22, 9)	(2.42, 8)	(2.42, 8)	(2.42, 8)	(2.42, 8)	(2.42, 8)	(2.42, 8)	(8.72, 8)
0.3	(1.67,13)	(1.89,11)	(2.03,10)	(2.19, 9)	(2.39, 8)	(2.39, 8)	(2.39, 8)	(2.39, 8)	(2.39, 8)	(2.39, 8)	(8.60, 8)
0.4	(1.65,13)	(1.75,12)	(1.87,11)	(2.01,10)	(2.36, 8)	(2.36, 8)	(2.36, 8)	(2.36, 8)	(2.36, 8)	(2.36, 8)	(8.56, 8)
0.5	(1.55,14)	(1.64,13)	(1.74,12)	(1.99,10)	(2.15, 9)	(2.34, 8)	(2.34, 8)	(2.34, 8)	(2.34, 8)	(2.34, 8)	(8.52, 8)
0.6	(1.46,15)	(1.63,13)	(1.73,12)	(1.84,11)	(2.13, 9)	(2.31, 8)	(2.31, 8)	(2.31, 8)	(2.31, 8)	(2.31, 8)	(8.48, 8)
0.7	(1.45,15)	(1.53,14)	(1.62,13)	(1.83,11)	(1.96,10)	(2.29, 8)	(2.29, 8)	(2.29, 8)	(2.29, 8)	(2.29, 8)	(0.00, 8)
0.8	(1.38,16)	(1.45,15)	(1.61,13)	(1.70,12)	(1.94,10)	(2.27, 8)	(2.27, 8)	(2.27, 8)	(2.27, 8)	(2.27, 8)	(0.00, 8)
0.9	(1.31,17)	(1.44,15)	(1.51,14)	(1.69,12)	(1.93,10)	(2.25, 8)	(2.25, 8)	(2.25, 8)	(2.25, 8)	(2.25, 8)	(0.00, 8)
1	(1.30,17)	(1.36,16)	(1.50,14)	(1.68,12)	(1.79,11)	(2.06, 9)	(2.23, 8)	(2.23, 8)	(2.23, 8)	(2.23, 8)	(0.00, 8)
$T_{double} = 10$											
$T_{lag} = 7$	θ										
δ	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	(1.82,12)	(1.95,11)	(2.10,10)	(2.28, 9)	(2.49, 8)	(2.49, 8)	(2.49, 8)	(2.49, 8)	(2.49, 8)	(2.49, 8)	(8.91, 8)
0.1	(1.69,13)	(1.80,12)	(1.93,11)	(2.08,10)	(2.46, 8)	(2.46, 8)	(2.46, 8)	(2.46, 8)	(2.46, 8)	(2.46, 8)	(8.75, 8)
0.2	(1.59,14)	(1.68,13)	(1.79,12)	(2.05,10)	(2.22, 9)	(2.42, 8)	(2.42, 8)	(2.42, 8)	(2.42, 8)	(2.42, 8)	(8.72, 8)
0.3	(1.49,15)	(1.57,14)	(1.67,13)	(1.89,11)	(2.03,10)	(2.39, 8)	(2.39, 8)	(2.39, 8)	(2.39, 8)	(2.39, 8)	(8.60, 8)
0.4	(1.41,16)	(1.48,15)	(1.65,13)	(1.75,12)	(2.01,10)	(2.17, 9)	(2.36, 8)	(2.36, 8)	(2.36, 8)	(2.36, 8)	(8.56, 8)
0.5	(1.34,17)	(1.40,16)	(1.55,14)	(1.64,13)	(1.86,11)	(2.15, 9)	(2.34, 8)	(2.34, 8)	(2.34, 8)	(2.34, 8)	(8.52, 8)
0.6	(1.27,18)	(1.39,16)	(1.46,15)	(1.63,13)	(1.84,11)	(1.97,10)	(2.31, 8)	(2.31, 8)	(2.31, 8)	(2.31, 8)	(8.48, 8)
0.7	(1.21,19)	(1.32,17)	(1.45,15)	(1.53,14)	(1.71,12)	(1.96,10)	(2.29, 8)	(2.29, 8)	(2.29, 8)	(2.29, 8)	(0.00, 8)
0.8	(1.21,19)	(1.26,18)	(1.38,16)	(1.52,14)	(1.70,12)	(1.94,10)	(2.27, 8)	(2.27, 8)	(2.27, 8)	(2.27, 8)	(0.00, 8)
0.9	(1.16,20)	(1.25,18)	(1.31,17)	(1.44,15)	(1.60,13)	(1.80,11)	(2.25, 8)	(2.25, 8)	(2.25, 8)	(2.25, 8)	(0.00, 8)
1	(1.11,21)	(1.20,19)	(1.30,17)	(1.43,15)	(1.59,13)	(1.79,11)	(2.06, 9)	(2.23, 8)	(2.23, 8)	(2.23, 8)	(0.00, 8)

Table 2.3: Optimal dosage and optimal number of fractions (d^*, N^*) in the presence of uncertainty in tumor parameters (denoted by ψ) and uncertainty in ρ parameters (denoted by δ).

2.2 Robust spatiobiologically integrated fractionation

In this section, we introduce a robust counterpart of the spatiobiologically *integrated* fractionation problem. We assume that the unknown parameters of the LQ dose-response model belong to known intervals. The goal in the robust problem is then to ensure that its solution remains feasible for all values of the dose-response parameters from these intervals. We demonstrate that the resulting robust optimization problem can be equivalently reformulated as a group of problems with convex quadratic constraints. It is thus efficiently solvable. We remark as an aside that this reformulation method is more general than and hence applicable to the robust spatiobiologically separated formulation in Section 2.1.4. We numerically quantify the price of robustness for a representative test case for head-and-neck cancer from [57], and also compare the frequency and amount of infeasibility suffered by the nominal and the robust solutions. Our results suggest that the price of robustness is small and the robust solution remains statistically more feasible than the nominal solution even when dose-response parameters vary outside the presumed intervals of uncertainty.

2.2.1 Nominal and robust formulations

We first briefly recall the nominal spatiobiologically integrated formulation from [57]. The objective in this formulation is to choose the number of sessions and the fluence-map in each session to maximize tumor-BE subject to maximum dose and mean dose constraints on various OAR[2]. The formulation also includes smoothness constraints on the fluence-map to ensure that it is deliverable in practice using IMRT. Here, we only discuss minimal, crucial ingredients of this formulation; the reader is referred to [57] for an extensive discussion of its derivation, advantages, and limitations.

[2]Dose-volume constraints can also be handled using the approach described in [57]; we omit those here for brevity.

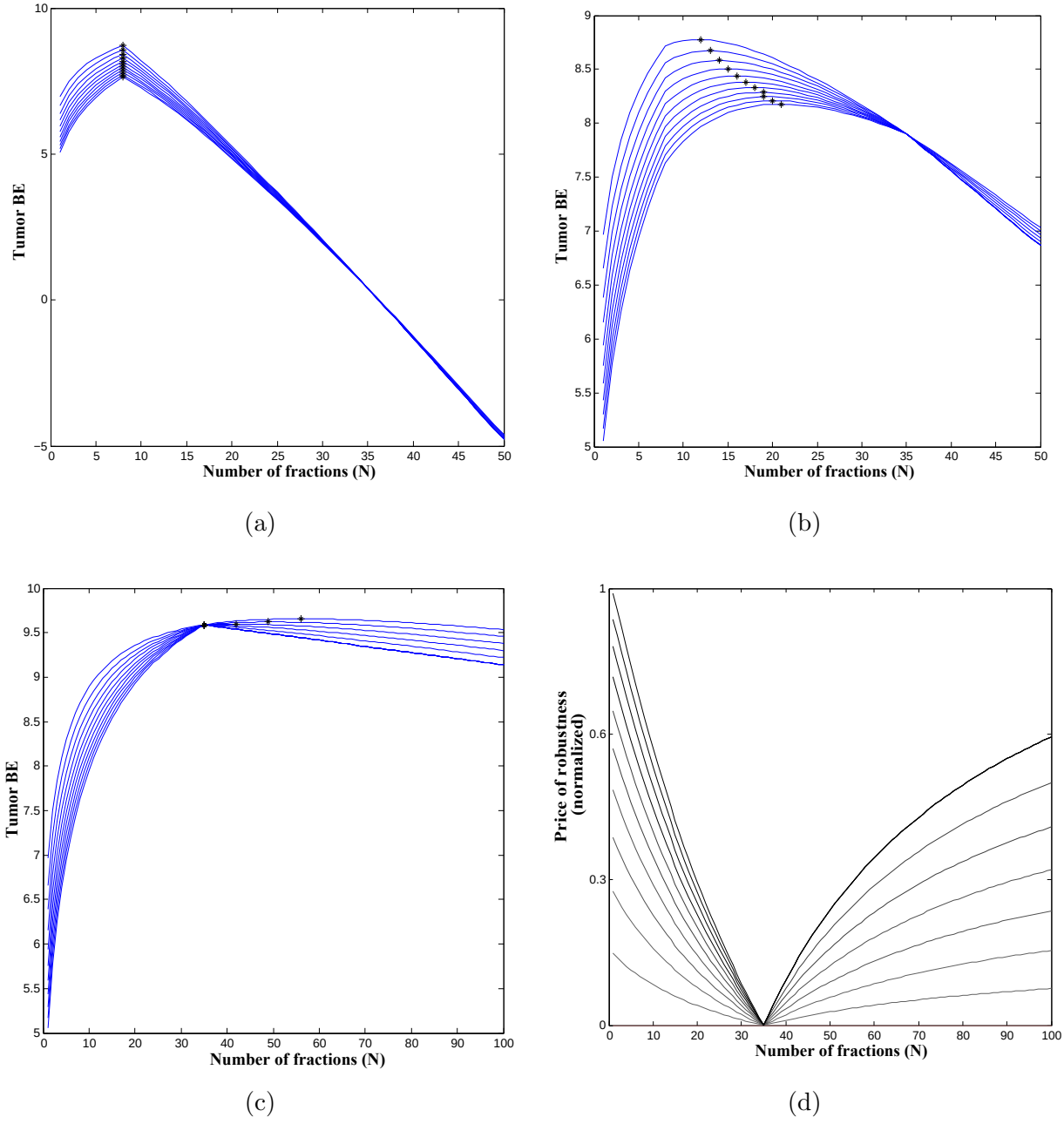


Figure 2.1: (a), (b), (c): The value of the objective as a function of N for $T_{\text{lag}} = 7$. The data points denoted by $*$ show the points (N^*, f^*) in each graph. The uppermost line in each set of graphs (a), (b), and (c) represents the nominal case ($\delta = 0$) and the other lines correspond to $\delta = \{0.1, 0.2, \ldots, 1\}$, respectively, from top to bottom. (a): $T_{\text{double}} = 2$ days and $N^* = 8$ for all δ. (b): $T_{\text{double}} = 10$ days. In this case, N^* increases with increasing δ. (c): $T_{\text{double}} = 100$ days. In this case, N^* decreases with increasing δ (from top to bottom). Moreover, note that in some cases (for example in (c)), some of these points lie on top of each other as the optimal solution for their corresponding δ are equal. (d): Normalized price of robustness as a function of N when $T_{\text{lag}} = 7$, $T_{\text{double}} = 2$. The uppermost line represents the most uncertain case ($\delta = 1$) and other lines represent $\delta = \{0.9, 0.8, \ldots, 0.1\}$, respectively, from top to bottom.

2.2.1.1 Review of the nominal formulation from [57]

Let n denote the number of tumor voxels. The radiation field is discretized into small segments called beamlets. Let k denote the number of beamlets. Let $u \in \Re_+^k$ denote the k-dimensional fluence-map employed in each session. The number of treatment sessions is denoted by N, with $N_{\max}$ being the maximum number of sessions allowed in the protocol. Let A denote the $n \times k$ nonnegative tumor dose-deposition matrix; A_i denotes its ith row, which corresponds to the ith tumor voxel. According to the linear dose-deposition model, $A_i u$ is the dose delivered to the ith tumor voxel in each session. Moreover, let $\bar{A}u \triangleq \sum_{i=1}^{n} A_i u/n$ be the average dose over all tumor voxels in each session. Let S denote the matrix employed in writing smoothness constraints on u.

Let α_0 and β_0 denote the tumor-response parameters. A tumor proliferation term is also included in our model; it is defined by $\tau(N) \triangleq ([(N-1) - T_{\text{lag}}]^+ \ln 2)/T_{\text{double}}$, where $[(N-1) - T_{\text{lag}}]^+ = \max((N-1) - T_{\text{lag}}, 0)$. In this formula, T_{lag} represents the time (in days) after which the tumor starts proliferating following the start of treatment and T_{double} is the tumor doubling time (in days).

The set of OAR is denoted by $\mathcal{M} \triangleq \mathcal{M}_1 \cup \mathcal{M}_2$; here, $\mathcal{M}_1$ and $\mathcal{M}_2$ are mutually exclusive sets of OAR with maximum dose and mean dose constraints, respectively. We use the index m to denote quantities related to OAR $m \in \mathcal{M}$. The set of voxels in OAR m is denoted by $\mathcal{N}_m \triangleq \{1, 2, \ldots, n_m\}$. Let $\rho_m \triangleq \beta_m/\alpha_m$ denote the inverse alpha-over-beta ratio of the α and β parameters of the LQ dose-response model for OAR m. Suppose for OAR $m \in \mathcal{M}_1$ that a total dose $D_{\max}^m$ is known to be tolerated by each voxel if administered in N_{conv}^m equal-dose fractions. Similarly, suppose for OAR $m \in \mathcal{M}_2$ that total mean dose D_{mean}^m is known to be tolerated if administered in N_{conv}^m equal-dose fractions. Let $\text{BED}_{\square}^m = D_{\square}^m + \rho_m (D_{\square}^m)^2 / N_{\text{conv}}^m$ be the BED of total dose $D_{\square}^m$ if administered in N_{conv}^m equal-dose fractions, where $\square$ represents either max or mean.

The nominal spatiobiologically integrated formulation is written in [57] as

$$(P)\ F^* = \max_{N,u}\ N\alpha_0(\bar{A}u) + N\beta_0(\bar{A}u)^2 - \tau(N), \tag{2.49}$$

$$N(A_j^m u) + N\rho_m(A_j^m u)^2 \leq \mathrm{BED}_{\max}^m,\ \forall j \in \mathcal{N}_m,\ m \in \mathcal{M}_1, \tag{2.50}$$

$$N\sum_{j=1}^{n_m}(A_j^m u) + N\rho_m\sum_{j=1}^{n_m}(A_j^m u)^2 \leq n_m \mathrm{BED}_{\mathrm{mean}}^m,\ m \in \mathcal{M}_2, \tag{2.51}$$

$$Su \leq 0, \tag{2.52}$$

$$u \geq 0,\ 1 \leq N \leq N_{\max},\ \text{integer}. \tag{2.53}$$

The objective in (2.49) is to maximize the BE of average tumor dose. Constraints (2.50) enforce that the BED to each OAR voxel is no more than the conventional BED; these are called the maximum dose constraints. Constraints (2.51) imply that the average BED of doses administered to different voxels is bounded above by the conventional BED; these are the mean dose constraints. Constraint (2.52) ensures, via an appropriately constructed smoothness matrix S, that the relative absolute difference between geometrically adjacent components of u is within a range that IMRT can attain.

2.2.1.2 Robust counterpart

Let $\tilde{\rho}_m$ represent the "true" unknown value of ρ_m for each OAR $m \in \mathcal{M}$. We assume that $\tilde{\rho}_m$ belongs to the uncertainty interval $[\rho_m^{\min}, \rho_m^{\max}]$, where $0 \leq \rho_m^{\min} \leq \rho_m^{\max}$. Thus, by replacing ρ_m with $\tilde{\rho}_m$ and expanding $\mathrm{BED}_{\square}^m$ in the nominal formulation, the robust counterpart of (P) is

$$(RP)\ \max_{N,u}\ N\alpha_0(\bar{A}u) + N\beta_0(\bar{A}u)^2 - \tau(N), \tag{2.54}$$

$$N(A_j^m u) + \tilde{\rho}_m\left(N(A_j^m u)^2 - \frac{(D_{\max}^m)^2}{N_{\mathrm{conv}}^m}\right) \leq D_{\max}^m,\ \forall j \in \mathcal{N}_m,\ m \in \mathcal{M}_1,\ \rho_m^{\min} \leq \tilde{\rho}_m \leq \rho_m^{\max}, \tag{2.55}$$

$$N\sum_{j=1}^{n_m}(A_j^m u)+\tilde{\rho}_m\left(N\sum_{j=1}^{n_m}(A_j^m u)^2-\frac{n_m(D_{\text{mean}}^m)^2}{N_{\text{conv}}^m}\right)\leq n_m D_{\text{mean}}^m,\ m\in\mathcal{M}_2,\ \rho_m^{\min}\leq\tilde{\rho}_m\leq\rho_m^{\max},$$

(2.56)

$$(2.52)-(2.53).\tag{2.57}$$

Here, inequalities (2.55)-(2.56) ensure that the maximum dose and the mean dose constraints are respected for every possible realization of $\tilde{\rho}_m$ in the interval $[\rho_m^{\min},\rho_m^{\max}]$. Consequently, they include an uncountably infinite number of constraints, and thus this robust problem appears intractable at first glance. Fortunately, it can be equivalently rewritten in a simple form by applying the general method proposed in Section 2.2 of [10] for robust optimization under polyhedral uncertainty. Below we convey the main idea in this reformulation using the mean dose constraints (2.56).

For any fixed u and N, the mean dose constraint holds for every realization of $\tilde{\rho}_m$ if and only if the maximum of the left hand side in inequality (2.56) over all $\tilde{\rho}_m\in[\rho_m^{\min},\rho_m^{\max}]$ is no more than the right hand side $n_m D_{\text{mean}}^m$. For any fixed u and N, the problem of maximizing this left hand side is an LP in variables $\tilde{\rho}_m$ with only two constraints: $\tilde{\rho}_m\leq\rho_m^{\max}$ and $\tilde{\rho}_m\geq\rho_m^{\min}$. By attaching variables q_m and p_m, respectively, with these two constraints, we can write the dual of this LP. This dual is a minimization problem, and by strong duality its optimal cost equals the maximum value in the primal LP. In other words, For any fixed u and N, the mean dose constraint holds for every realization of $\tilde{\rho}_m$ if and only if the optimal cost in the dual LP is no more than $n_m D_{\text{mean}}^m$. This allows us to replace the robust mean dose constraints with another equivalent group of constraints. This idea can be expressed mathematically as follows:

$$\max_{\rho_m^{\min}\leq\tilde{\rho}_m\leq\rho_m^{\max}}\left\{N\sum_{j=1}^{n_m}(A_j^m u)+\tilde{\rho}_m\left(N\sum_{j=1}^{n_m}(A_j^m u)^2-n_m\frac{(D_{\text{mean}}^m)^2}{N_{\text{conv}}^m}\right)\right\}\leq n_m\ D_{\text{mean}}^m$$

$$\Leftrightarrow \left(\begin{array}{c} \min\limits_{q_m,p_m} \; N\sum\limits_{j=1}^{n_m}(A_j^m u) + \rho_m^{\max}\; q_m + \rho_m^{\min}\; p_m, \\ q_m + p_m \geq N\sum\limits_{j=1}^{n_m}(A_j^m u)^2 - n_m \frac{(D_{\text{mean}}^m)^2}{N_{\text{conv}}^m}, \\ q_m \geq 0,\; p_m \leq 0. \end{array} \right) \leq n_m\; D_{\text{mean}}^m.$$

Applying this to maximum dose constraints as well, the robust optimization problem becomes

$$(RP)\;\; F^* = \max_{N,u,\text{q,p}} \; N\alpha_0(\bar{A}u) + N\beta_0(\bar{A}u)^2 - \tau(N), \tag{2.58}$$

$$\text{s.t. } N(A_j^m u) + \rho_m^{\max}\; q_m^j + \rho_m^{\min}\; p_m^j \leq D_{\max}^m,\; \forall j \in \mathcal{N}_m,\; m \in \mathcal{M}_1, \tag{2.59}$$

$$N(A_j^m u)^2 - q_m^j - p_m^j \leq \frac{(D_{\max}^m)^2}{N_{\text{conv}}^m},\; \forall j \in \mathcal{N}_m,\; m \in \mathcal{M}_1, \tag{2.60}$$

$$N\sum_{j=1}^{n_m}(A_j^m u) + \rho_m^{\max}\; q_m + \rho_m^{\min}\; p_m \leq n_m\; D_{\text{mean}}^m,\; \forall m \in \mathcal{M}_2, \tag{2.61}$$

$$N\sum_{j=1}^{n_m}(A_j^m u)^2 - q_m - p_m \leq n_m \frac{(D_{\text{mean}}^m)^2}{N_{\text{conv}}^m},\; \forall m \in \mathcal{M}_2, \tag{2.62}$$

$$p \leq 0,\; q \geq 0,\; (2.52) - (2.53). \tag{2.63}$$

Here, q and p are the vectors of all dual variables.

For each fixed N, (RP) maximizes a convex function of u. This difficulty can be addressed by further simplifying (RP) as in the nominal formulation from [57]. First, we propose to solve the robust problem for each fixed integer value of N in the range $1 \leq N \leq N_{\max}$. We call these problems $(RP(N))$ and denote their optimal values by $F^*(N)$. The optimal value and solution for (RP) can then be found simply by choosing the best among all $(RP(N))$. Furthermore, for each fixed N, the objective in $(RP(N))$ is increasing in $\bar{A}u$. Thus, maximizing this objective is equivalent to maximizing $\bar{A}u$. We utilize this in solving $(RP(N))$. Consequently, we note that each $(RP(N))$ includes an objective that is linear in u and that does not include variables p, q. Finally, constraints in each $(RP(N))$ are convex quadratic in u and linear in p, q. As such, problems $(RP(N))$ can be solved efficiently in practice using off-the-shelf software.

2.2.2 Numerical results

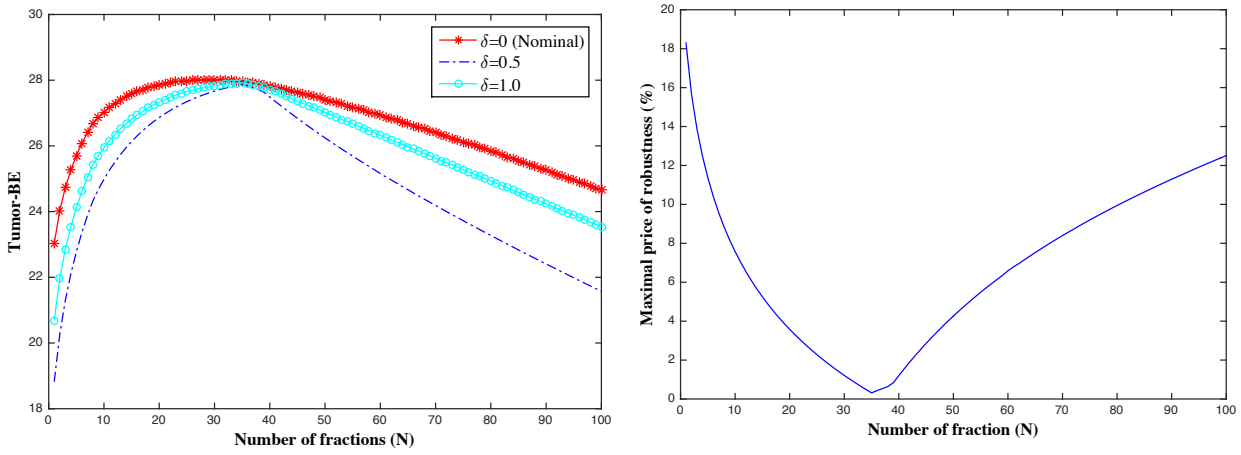
2.2.2.1 Description of a head-and-neck cancer case

We demonstrate our numerical results on a representative head-and-neck cancer case from [57]. The case included four OAR: spinal cord, brainstem, left and right parotids. The nominal values of ρ for all OAR were fixed at $\rho = 1/3$ Gy^{-1} based on [23, 75]. We followed a treatment protocol similar to QUANTEC [47]. Specifically, the conventional number of sessions N_{conv} was fixed at 35. Maximum dose constraints of 45 Gy and 50 Gy were included for spinal cord and brainstem, respectively. Mean dose constraints with a tolerance dose of 28 Gy were included for left and right parotids. $N_{\max}$ was set to 100 days. We are aware that this value is somewhat unrealistically high; we nevertheless used it to bring forth the full range of sensitivity behaviors in our problem. Tumor-response parameters were fixed at $\alpha_0 = 0.35$ Gy^{-1} and $\beta_0 = 0.035$ Gy^{-2} based on [23, 26]. A maximum dose of 90 Gy was enforced on all tumor voxels to encourage dose-uniformity. The fluence-map u consisted of 3,910 beamlets and the total number of constraints in the nominal formulation was 27,450. All computer simulations were performed on a 3.1 GHz iMac desktop with 16 GB RAM using the convex optimization toolbox CVX [31] in MATLAB.

2.2.2.2 Price of robustness

Price of robustness is defined as the relative decrease in the optimal tumor BE in the robust formulation compared to the nominal formulation. In our numerical experiments, uncertainty intervals were parameterized as $[(1-\delta)/3, (1+\delta)/3]$, with $\delta \in \{0, 0.1, \ldots, 1\}$, where $(1/3)$ Gy^{-1} is the nominal value of ρ for all OAR. Here, $\delta = 0$ corresponds to the nominal case, $\delta = 1$ represents 100% uncertainty. This allowed us to quantify the price of robustness as a function of the single uncertainty parameter δ.

Tables 2.4 and 2.5 summarize the results of 200 experiments for different values of T_{lag}, T_{double}, and δ. In these experiments, T_{lag} values were set to $7, 14, 21, 28$ days based on [26] and T_{double} values were set to $2, 10, 20, 40, 50$ days based on [23, 26, 49, 75].



(a) (b)

Figure 2.2: (a): Tumor-BE as a function of N. Three lines show the tumor-BE for three cases: no uncertainty ($\delta = 0$), medium uncertainty ($\delta = 0.5$), and maximum uncertainty ($\delta = 1$). (b): maximal price of robustness (%) as a function of N. This values was calculated as $100 * (\text{BE}_{\text{nom}} - \text{BE}_{\delta=1})/\text{BE}_{\text{nom}})$.

Table 2.4 shows, as expected, that the price of robustness increases with increasing δ for each T_{lag}, T_{double} combination. Overall, the price of robustness was small with an average of 1.29%, with first, second, and third quartiles of 0.48%, 0.98%, and 1.66%, respectively.

For all T_{lag}, δ combinations in Table 2.4, the price of robustness first decreases with increasing T_{double}, reaches its smallest value, and then increases. This trend is consistent with the corresponding trend in the difference between $N^m_{\text{conv}} = 35$ and the optimal number of fractions (N^*) that can be inferred from Table 2.5. Specifically, for most cases reported in Table 2.5, the magnitude of difference between N^m_{conv} and N^* ($|N^m_{\text{conv}} - N^*|$) decreases with increasing T_{double}, reaches its minimum, and then increases.

Figure 2.2a shows the evolution of tumor-BE with N for $T_{\text{lag}} = 7$ days and $T_{\text{double}} = 10$ days as an example. The price of robustness (the vertical distance between lines) is not monotone with increasing N. It is increasing at first, reaches its maximum, decreases afterwards until it reaches zero when $N = N^m_{\text{conv}} = 35$ and increases thereafter. The price of robustness equals zero at 35 because when $N = N^m_{\text{conv}}$, ρ_m is eliminated from the BED

constraints thus reducing them to total dose constraints. In other words, the nominal and the robust problems become identical. Figure 2.2b shows the maximal price of robustness (%) between the nominal ($\delta = 0$) and the most uncertain case ($\delta = 1$) as a function of N for the same case. The trend described above for the price of robustness is more visible in this figure.

As can be deducted from Table 2.4, there is no consistent trend for how the price of robustness changes as T_{lag} increases. For example, when $T_{\text{double}} = 2$ and $\delta = 0.1$, the price of robustness decreases as T_{lag} increases; whereas when $T_{\text{double}} = 2$ and $\delta = 0.7$, the price of robustness first decreases from $T_{\text{lag}} = 7$ to 14, decreases when it $T_{\text{lag}} = 21$, and then decreases again. However, in all cases, the change in the price of robustness is consistent with the trend in $|N^m_{\text{conv}} - N^*|$, as reported in Table 2.5.

A closer look at Table 2.5 reveals that the evolution of N^* with δ does not exhibit a universal trend. For instance, N^* increases with increasing δ when $T_{\text{double}} = 2$ and 10 days and decreases as δ increases when $T_{\text{double}} = 20,\ 40$ and 50 days.

Our robust solutions continue to exhibit qualitative trends that are well-established in the nominal case. For instance, N^* increases with increasing T_{double} for any fixed δ, T_{lag} combination. Similarly, N^* also increases as T_{lag} increases for any fixed δ, T_{double} combination.

2.2.2.3 Infeasibility tests

The nominal solution is guaranteed to remain feasible only when the realized value of the ρ parameters equal their nominal values. To quantify the frequency and the amount of infeasibility of the nominal solutions when this is not the case, we computed the fluence-map and the number of sessions that would be optimal if the dose-response parameters equaled their nominal values, and tested whether or not such a nominal solution would remain feasible if the true dose-parameters took some other values from the uncertainty interval. For each uncertainty level δ and for each constraint m, five different grid-points were selected to calculate the realized values of dose-parameters. The five values were $\tilde{\rho}^m_i = \rho^m_{\min} + 2i\rho^i_m\delta/5$, where $i = 1, 2, \ldots, 5$, and ρ^i_m denotes the nominal value (1/3 Gy^{-1}) of the dose-parameter

for constraint m. We performed this experiment for all combinations of T_{lag}, T_{double}, and δ. The nominal solution was infeasible in over 91% of the cases. The amount of maximum infeasibility itself was as large as 70% in some cases, with an average of 10.11% and first, second, third quartiles of 3.20%, 7.12%, and 15.56%. In all cases, the robust solutions of course remained feasible.

The fundamental assumption in developing our robust formulation was that the true values of the dose-parameters, although uncertain, belonged to some presumed uncertainty set. This assumption however might not hold. Therefore, to quantify the frequency and amount of infeasibility for robust and nominal solutions when dose-response parameters varied outside the uncertainty intervals, we repeated the above experiments by changing parameter values as follows: for each uncertainty level δ and constraint m, five grid-points were chosen at $\tilde{\rho}_m = (1+\delta+\gamma)\rho_m$ and five grid-points at $\tilde{\rho}_m = (1-\delta-\gamma)\rho_m$, where $\gamma \in \{0.1, 0.2, \ldots, 0.5\}$ and ρ_m denotes the nominal value of $1/3$ Gy^{-1}. We performed numerical experiments for all combinations of T_{lag}, T_{double}, and δ. The nominal solution was infeasible in over 76% of the cases, while the robust solution was infeasible in 58% of the cases. Furthermore, the average amount of maximum infeasibility was less for the robust solution than for the nominal (6.84% versus 12.19%). A pairwise t-test revealed that this difference was significant at the 95% level. This suggests that the robust solution is statistically less infeasible than the nominal.

2.2.2.4 Uncertainty in tumor parameters α_0 and β_0

We assumed for simplicity that there is no uncertainty in the tumor-response parameters α_0 and β_0. It is easy to relax this assumption while writing the robust counterpart of (P). This is because, under interval uncertainty for parameters α_0 and β_0, the worst-case objective value is obtained simply by inserting the worst values of these parameters. Specifically, we now assume that true values $\tilde{\alpha}_0$ and $\tilde{\beta}_0$ of these parameters belong to known intervals $[\alpha_0^{\min}, \alpha_0^{\max}]$ and $[\beta_0^{\min}, \beta_0^{\max}]$, respectively. Then, the objective function of the robust problem is given by replacing $\tilde{\alpha}_0$ with $\alpha_0^{\min}$ and $\tilde{\beta}_0$ with $\beta_0^{\min}$. Recall that the treatment protocol described in Section 2.1.6.1 also included maximum dose constraints on tumor voxels. The robust

counterpart of these constraints under interval uncertainty was derived in our numerical experiments below by applying the procedure described in Section 2.2 for maximum dose constraints on OAR.

We characterized the interval uncertainty in tumor-response parameters via a single parameter by setting $\alpha_0^{\min} = (1-\theta)0.35$ Gy^{-1}; $\alpha_0^{\max} = (1+\theta)0.35$ Gy^{-1}; $\beta_0^{\min} = (1-\theta)0.035$ Gy^{-2}; and $\beta_0^{\max} = (1+\theta)0.035$ Gy^{-2}, for $\theta \in \{0.1, 0.2, \ldots, 0.9\}$. Table 2.6 shows the results for $T_{\text{lag}} = 7$ days and $T_{\text{double}} = 21$ days as a representative example. For each fixed θ, the qualitative trend in the evolution of N^* was similar to the previous section. By dividing the objective function in the robust problem by $(1-\theta)$, we see that $(1-\theta)$ scales T_{double} in the denominator of the proliferation term $\tau(N)$. Consequently, a larger θ has the same effect on BE as a faster growing tumor (that is, a smaller T_{double}). Thus, a larger θ should lead to a shorter treatment course. This trend can be observed in each row of Table 2.6. Finally, the price of robustness (not listed in the table for brevity) was insensitive to θ, and remained small with an average of 0.87% over all δ, θ combinations.

2.2.2.5 Conclusions

There is a large body of literature on robust spatial optimization in radiotherapy, where the goal is to compensate for intra-fraction and inter-fraction errors made in dose-delivery due to patient movement and patient setup. Here, we applied robust optimization to a different problem — that of incorporating uncertainty in dose-response parameters in deciding spatiobiologically integrated fractionation schedules. Although the robust counterpart of this fractionation problem first appeared intractable, we were able to equivalently reformulate it as a simple convex problem that can be solved efficiently. Our numerical experiments on a representative head-and-neck test case suggest that the resulting price of robustness is small and that the robust solutions are statistically less infeasible than nominal. This bodes well for potential clinical implementation of our robust optimization methodology.

$T_{\text{lag}} = 7$	T_{double}					$T_{\text{lag}} = 14$	T_{double}				
δ	2	10	20	40	50	δ	2	10	20	40	50
0.1	1.47	0.17	0.1	0.51	0.61	0.1	0.61	0.27	0.1	0.5	0.61
0.2	2.8	0.26	0.16	0.92	1.15	0.2	1.15	0.43	0.16	0.92	1.14
0.3	4.01	0.3	0.21	1.17	1.54	0.3	1.54	0.52	0.21	1.16	1.54
0.4	5.11	0.33	0.27	1.24	1.66	0.4	1.66	0.56	0.26	1.23	1.66
0.5	6.12	0.37	0.32	1.31	1.73	0.5	1.73	0.6	0.32	1.3	1.73
0.6	7.01	0.4	0.39	1.39	1.81	0.6	1.81	0.63	0.39	1.38	1.81
0.7	7.81	0.43	0.45	1.46	1.9	0.7	1.9	0.66	0.45	1.46	1.89
0.8	8.54	0.46	0.48	1.54	1.97	0.8	1.97	0.69	0.48	1.53	1.97
0.9	9.19	0.49	0.51	1.63	2.06	0.9	2.06	0.72	0.51	1.62	2.06
1	9.8	0.51	0.54	1.71	2.17	1	2.17	0.75	0.53	1.71	2.16
$T_{\text{lag}} = 21$	T_{double}					$T_{\text{lag}} = 28$	T_{double}				
δ	2	10	20	40	50	δ	2	10	20	40	50
0.1	0.47	0.27	0.1	0.5	0.61	0.1	0.21	0.21	0.1	0.5	0.6
0.2	0.89	0.43	0.16	0.92	1.14	0.2	0.41	0.37	0.16	0.91	1.14
0.3	1.29	0.51	0.21	1.16	1.53	0.3	0.59	0.45	0.21	1.15	1.53
0.4	1.65	0.55	0.26	1.23	1.65	0.4	0.75	0.49	0.26	1.22	1.65
0.5	1.99	0.59	0.32	1.3	1.72	0.5	0.91	0.53	0.32	1.29	1.72
0.6	2.31	0.62	0.38	1.38	1.8	0.6	1.05	0.56	0.38	1.37	1.8
0.7	2.6	0.65	0.44	1.45	1.89	0.7	1.19	0.59	0.44	1.44	1.88
0.8	2.88	0.68	0.48	1.53	1.96	0.8	1.32	0.62	0.47	1.52	1.96
0.9	3.13	0.71	0.5	1.62	2.05	0.9	1.44	0.65	0.5	1.61	2.04
1	3.38	0.74	0.53	1.7	2.15	1	1.55	0.67	0.52	1.69	2.14

Table 2.4: Price of robustness (%) for different T_{lag} (days), T_{double} (days), and uncertainty levels δ.

$T_{\text{lag}} = 7$	T_{double}				
δ	2	10	20	40	50
0	(6.03,8)	(2.46,28)	(1.68,44)	(1.12,71)	(0.98,82)
0.1	(5.97,8)	(2.35,29)	(1.85,39)	(1.25,62)	(1.09,73)
0.2	(5.91,8)	(2.17,32)	(1.85,39)	(1.43,53)	(1.26,61)
0.3	(5.86,8)	(2.07,34)	(1.89,38)	(1.78,41)	(1.58,47)
0.4	(5.81,8)	(2.02,35)	(1.89,38)	(1.85,39)	(1.85,39)
0.5	(5.76,8)	(2.02,35)	(1.89,38)	(1.85,39)	(1.85,39)
0.6	(5.30,9)	(2.02,35)	(1.89,38)	(1.85,39)	(1.85,39)
0.7	(5.27,9)	(2.02,35)	(1.98,36)	(1.89,38)	(1.89,38)
0.8	(4.89,10)	(2.02,35)	(2.02,35)	(1.89,38)	(1.89,38)
0.9	(4.86,10)	(2.02,35)	(2.02,35)	(1.89,38)	(1.89,38)
1	(4.84,10)	(2.02,35)	(2.02,35)	(2.02,35)	(1.88,38)
$T_{\text{lag}} = 14$	T_{double}				
δ	2	10	20	40	50
0	(4.29,14)	(2.57,26)	(1.68,44)	(1.12,71)	(0.98,82)
0.1	(4.29,14)	(2.35,29)	(1.85,39)	(1.25,62)	(1.09,73)
0.2	(4.29,14)	(2.17,32)	(1.85,39)	(1.43,53)	(1.26,61)
0.3	(4.29,14)	(2.07,34)	(1.89,38)	(1.78,41)	(1.58,47)
0.4	(4.29,14)	(2.02,35)	(1.89,38)	(1.85,39)	(1.85,39)
0.5	(4.29,14)	(2.02,35)	(1.89,38)	(1.85,39)	(1.85,39)
0.6	(3.88,15)	(2.02,35)	(1.89,38)	(1.85,39)	(1.85,39)
0.7	(3.88,15)	(2.02,35)	(1.98,36)	(1.89,38)	(1.89,38)
0.8	(3.88,15)	(2.02,35)	(2.02,35)	(1.89,38)	(1.89,38)
0.9	(3.88,15)	(2.02,35)	(2.02,35)	(1.89,38)	(1.89,38)
1	(3.88,15)	(2.02,35)	(2.02,35)	(2.02,35)	(1.88,38)
$T_{\text{lag}} = 21$	T_{double}				
δ	2	10	20	40	50
0	(2.92,22)	(2.57,26)	(1.68,44)	(1.12,71)	(0.98,82)
0.1	(2.91,22)	(2.35,29)	(1.85,39)	(1.25,62)	(1.09,73)
0.2	(2.90,22)	(2.17,32)	(1.85,39)	(1.43,53)	(1.26,61)
0.3	(2.89,22)	(2.07,34)	(1.89,38)	(1.78,41)	(1.58,47)
0.4	(2.88,22)	(2.02,35)	(1.89,38)	(1.85,39)	(1.85,39)
0.5	(2.87,22)	(2.02,35)	(1.89,38)	(1.85,39)	(1.85,39)
0.6	(2.86,22)	(2.02,35)	(1.89,38)	(1.85,39)	(1.85,39)
0.7	(2.86,22)	(2.02,35)	(1.98,36)	(1.89,38)	(1.89,38)
0.8	(2.85,22)	(2.02,35)	(2.02,35)	(1.89,38)	(1.89,38)
0.9	(2.84,22)	(2.02,35)	(2.02,35)	(1.89,38)	(1.89,38)
1	(2.84,22)	(2.02,35)	(2.02,35)	(2.02,35)	(1.88,38)

Table 2.5: Optimal average dose per session (Gy) and the optimal number of sessions (N^*) for $T_{\text{lag}} = 7, 14, 21$ days.

	Tumor uncertainty level (θ)								
δ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0	(1.78,41)	(1.94,37)	(2.07,34)	(2.29,30)	(2.57,26)	(3.02,21)	(3.54,17)	(4.85,11)	(6.03, 8)
0.1	(1.85,39)	(1.89,38)	(2.03,35)	(2.12,33)	(2.35,29)	(2.72,24)	(3.25,19)	(4.27,13)	(5.97, 8)
0.2	(1.89,38)	(1.89,38)	(2.02,35)	(2.02,35)	(2.17,32)	(2.55,26)	(3.00,21)	(4.03,14)	(5.91, 8)
0.3	(1.89,38)	(1.93,37)	(2.02,35)	(2.02,35)	(2.07,34)	(2.34,29)	(2.89,22)	(3.81,15)	(5.86, 8)
0.4	(1.89,38)	(1.98,36)	(2.02,35)	(2.02,35)	(2.02,35)	(2.22,31)	(2.70,24)	(3.62,16)	(5.81, 8)
0.5	(1.89,38)	(2.02,35)	(2.02,35)	(2.02,35)	(2.02,35)	(2.17,32)	(2.61,25)	(3.46,17)	(5.76, 8)
0.6	(1.98,36)	(2.02,35)	(2.02,35)	(2.02,35)	(2.02,35)	(2.07,34)	(2.46,27)	(3.31,18)	(5.30, 9)
0.7	(2.02,35)	(2.02,35)	(2.02,35)	(2.02,35)	(2.02,35)	(2.02,35)	(2.39,28)	(3.17,19)	(5.27, 9)
0.8	(2.02,35)	(2.02,35)	(2.02,35)	(2.02,35)	(2.02,35)	(2.02,35)	(2.33,29)	(3.05,20)	(4.89,10)
0.9	(2.02,35)	(2.02,35)	(2.02,35)	(2.02,35)	(2.02,35)	(2.02,35)	(2.27,30)	(2.94,21)	(4.86,10)
1	(2.02,35)	(2.02,35)	(2.02,35)	(2.02,35)	(2.02,35)	(2.02,35)	(2.21,31)	(2.84,22)	(4.84,10)

Table 2.6: Optimal average dose (Gy) per session and optimal number of sessions (N^*) for various tumor uncertainty levels (θ) and OAR uncertainty levels (δ) for $T_{\text{lag}} = 7$ days and $T_{\text{double}} = 21$ days.

Chapter 3

NONSTATIONARY FRACTIONATION

The spatiobiologically integrated formulation in Section 2.2.1.1 assumed that the fluence-maps across all sessions were identical, that is, stationary. This exposes the following two research questions: (1) could nonstationary fluence-maps be superior to stationary ones in clinically realistic test-cases for spatiobiologically integrated fractionation when dose-response parameters do not vary over time? and (2) can we devise an efficient algorithm to discover such nonstationary fluence-maps? In this chapter, we answer these questions in the affirmative.

3.1 Problem description

This chapter employs the same notation as Section 2.2. The only difference is that we now denote the fluence-map in session t by $u^t \in \Re^k_+$. This is to emphasize that the treatment planner could choose different fluence-maps in distinct sessions. The goal is to maximize the tumor-BE of average dose for the sequence of fluence-maps $(u^1; u^2; \ldots; u^t; \ldots; u^N)$. This tumor-BE objective is given by

$$\sum_{t=1}^{N}(\alpha_0 \bar{A}u^t + \beta_0(\bar{A}u^t)^2) - \tau(N). \tag{3.1}$$

This yields the following variation of the spatiobiologically integrated fractionation formulation from Section 2.2:

$$(P)\ F^* = \max \sum_{t=1}^{N}(\alpha_0 \bar{A}u^t + \beta_0(\bar{A}u^t)^2) - \tau(N), \tag{3.2}$$

$$\text{subject to } \sum_{t=1}^{N}(A_j^m u^t) + \sum_{t=1}^{N}\rho_m(A_j^m u^t)^2 \leq \text{BED}_{\max}^m,\ \forall j \in \mathcal{N}_m,\ m \in \mathcal{M}_1, \tag{3.3}$$

$$\sum_{t=1}^{N}\sum_{j=1}^{n_m}(A_j^m u^t) + \sum_{t=1}^{N}\rho_m \sum_{j=1}^{n_m}(A_j^m u^t)^2 \leq n_m \text{BED}_{\text{mean}}^m,\ m \in \mathcal{M}_2, \tag{3.4}$$

$$Su^t \leq 0,\ t = 1, 2, \ldots, N, \tag{3.5}$$

$$u^t \geq 0,\ t = 1, 2, \ldots, N. \tag{3.6}$$

Since $\tau(N)$ is fixed and it does not affect optimal solutions of (P), we will ignore it in the rest of this chapter. All constraints in (P) are convex in the concatenated fluence-map vector $u \triangleq (u^1; u^2; \ldots; u^N) \in \Re^{k \times N}$. The objective is also convex in u, but since we wish to maximize this function, (P) is *not* a convex problem. Specifically, it is a nonconvex QCQP. If we restricted consideration to stationary fluence-maps by setting $u^1 = u^2 = \ldots = u^N \triangleq v$ as in [57] (and as in Section 2.2), the objective function becomes monotone in $\bar{A}v$ and thus the problem can be reformulated as an equivalent convex one. In this chapter, we compare our nonstationary solutions against an optimal (stationary) solution to that convex problem.

To get a sense of the large scale of (P), note that for head-and-neck cancer, the typical number of beamlets, k, is about 3000 and the number of sessions N is 35. Thus, the dimension of u is about a 100,000. The number of voxels in a serial OAR can be about a 1000. As such, the total number of constraints in (P) can be as high as several thousand. Thus, efficient exact solution of this nonconvex QCQP is computationally difficult in practice.

First note that the objective and constraint functions in (P) are symmetric with respect to permutations over t. Second, standard approaches for approximate solution of large-scale nonconvex QCQPs call for convexification [43]. For (P), since the objective function is the only source of nonconvexity, these approximation methods would amount to replacing the objective with a concave function that is additively separable and symmetric over t. This would create a symmetric, convex optimization problem. Such problems are known to possess symmetric, that is, stationary optimal solutions [71]. Thus, standard convexification methods for (P) would not produce nonstationary fluence-maps. We thus borrow an approach

called model predictive control (MPC) from the literature on finite-horizon, discrete-time, deterministic, constrained, non-linear control theory [9].

3.2 Model predictive control

The idea in MPC, for a problem with N sessions such as (P), is simple. Beginning with the initial "state" of the problem, we first solve an N-session problem assuming stationary fluence-maps. Suppose an optimal stationary sequence of fluence-maps for this problem is $(\underbrace{u_*^1; u_*^1; \ldots; u_*^1}_{N \text{ times}})$. Implement fluence-map u_*^1 in the first session only. This transforms the state of the problem to a new state at the beginning of the second session. Now solve an $N-1$-session problem to obtain an optimal stationary sequence of fluence-maps $(\underbrace{u_*^2; u_*^2; \ldots; u_*^2}_{N-1 \text{ times}})$ and implement the fluence-map u_*^2 in the second session only. Repeat this process until the last session, where a single-period problem is solved and the resulting optimal fluence-map is implemented.

To implement MPC on (P), we first need to define the "state" of the problem at the beginning of sessions $t \geq 1$. Toward this end, for each serial OAR $m \in \mathcal{M}_1$ and for each voxel $j \in \mathcal{N}_m$ in this OAR, let $z_j^{t,m}$ denote the total BED administered in the first $t-1$ treatment sessions. Let $z^{t,m} \triangleq (z_1^{t,m}, z_1^{t,m}, \ldots, z_{n_m}^{t,m})$, and z^t be the vector formed by concatenating vectors $z^{t,m}$ for all $m \in \mathcal{M}_1$. Similarly, for each parallel OAR $m \in \mathcal{M}_2$, let $w^{t,m}$ denote the total average (over all voxels in $\mathcal{N}_m$) BED administered in the first $t-1$ sessions. Let w^t be the vector formed by concatenating vectors $w^{t,m}$ for all $m \in \mathcal{M}_2$. Then we define the state as $[z^t, w^t]$. We will need the optimization problem

$$
\begin{aligned}
(P_t) \quad \max \ & \sum_{t=1}^{N} (\alpha_0 \bar{A} u^t + \beta_0 (\bar{A} u^t)^2) \\
& \sum_{l=t}^{N} (A_j^m u^l + \rho_m (A_j^m u^l)^2) \leq \text{BED}_{\max}^m - z_j^{t,m}, \ \forall j \in \mathcal{N}_m, \ m \in \mathcal{M}_1, \\
& \sum_{j=1}^{n_m} \sum_{l=t}^{N} ((A_j^m u^l) + \rho_m (A_j^m u^t)^2) \leq n_m \text{BED}_{\text{mean}}^m - n_m w^{t,m}, \ m \in \mathcal{M}_2,
\end{aligned}
$$

$$Su^l \leq 0,\ l = t, t+1, \ldots, N,$$
$$u^l \geq 0,\ l = t, t+1, \ldots, N$$

in our precise listing of the MPC algorithm below.

The MPC Algorithm

1. Set $t = 1$, and begin with the initial state $[z^1, w^1] = [\vec{0}, \vec{0}]$.

2. **WHILE** $t \leq N$

   A. Let $u^t = u^{t+1} = \ldots = u^N$ in (P_t) and solve it. Let $(\underbrace{u_*^t; u_*^t; \ldots; u_*^t}_{N-t+1 \text{ times}})$ denote a stationary optimal solution of (P_t) obtained in this manner.

   B. Implement fluence-map u_*^t in session t and update the state as

$$z_j^{t+1,m} = z_j^t + A_j^m u_*^t + \rho_m (A_j^m u_*^t)^2,\ j = 1, 2, \ldots, n_m,\ m \in \mathcal{M}_1,$$
$$w^{t+1,m} = w^{t,m} + \frac{\sum_{j=1}^{n_m} \left[A_j^m u_*^t + \rho_m (A_j^m u_*^t)^2 \right]}{n_m},\ m \in \mathcal{M}_2.$$

   C. set $t = t + 1$ and go back to Step 2.

3. **END WHILE**

This algorithm delivers the sequence of fluence-maps $(u_*^1, u_*^2, \ldots, u_*^N)$ that is feasible to problem (P). It seems plausible that this sequence could be *nonstationary*. The next theorem dashes this hope.

Theorem 3.2.1. *Consider an optimization problem of the form*

$$\max \sum_{t=1}^{N} (\alpha_0 \bar{A} u^t + \beta_0 (\bar{A} u^t)^2),$$

$$g_m(u^1, u^2, \ldots, u^N) \leq b_m, \ m = 1, 2, \ldots, M.$$

Suppose that functions g_m are convex and symmetric. When MPC is implemented on this problem, it cannot *deliver a nonstationary solution that is strictly better than all stationary solutions. In particular, since the constraint functions in (P) are convex and symmetric, this stationarity property holds for our MPC implementation on problem (P).*

Proof. We provide a proof by contradiction. Suppose that MPC delivers a nonstationary solution of the form $\mathbf{u_N} = (\underbrace{u^1, \ldots, u^1}_{K_1 \text{ times}}, \underbrace{u^2, \ldots, u^2}_{K_2 \text{ times}}, \ldots, \underbrace{u^N, \ldots, u^N}_{K_N \text{ times}})$, where $u^i \neq u^j$ for $i \neq j$. Here, $K_1 + K_2 + \ldots, K_N = N$. Then MPC must have returned

$$\mathbf{u_{N\text{-}1}} = (\underbrace{u^1, \ldots, u^1}_{K_1 \text{ times}}, \underbrace{u^2, \ldots, u^2}_{K_2 \text{ times}}, \ldots, \underbrace{u^{N-1}, \ldots, u^{N-1}}_{K_{N-1}+K_N \text{ times}}),$$

when $t = 1 + K_1 + K_2 + \ldots + K_{N-2}$. Continuing backwards, MPC must have returned $\mathbf{u}_2 = (\underbrace{u^1, \ldots, u^1}_{K_1 \text{ times}}, \underbrace{u^2, \ldots, u^2}_{K_2+K_3\ldots+K_N \text{ times}})$ when $t = 1 + K_1$. Finally, MPC must have delivered $\mathbf{u}_1 = (\underbrace{u^1, \ldots, u^1}_{N \text{ times}})$ when $t = 1$. We define the notation $F^t(u^1, u^2, \ldots, u^N) = \sum_{l=t}^{N} f(u^l)$, where $f(u^l) = \alpha_0 \bar{A} u^l + \beta_0 (\bar{A} u^l)^2$ for $l = 1, 2, \ldots, N$. Now consider the following three cases.

Case 1: $\bar{A}u^1 > \bar{A}u^2$. In this case, $f(u^1) > f(u^2)$. Then $F^{1+K_1}(\mathbf{u_1}) = (K_2 + \ldots + K_N) f(u^1) > (K_2 + K_3 \ldots + K_N) f(u^2) = F^{1+K_1}(\mathbf{u_2})$. Since $\mathbf{u_1} = (\underbrace{u^1, u^1, \ldots, u^1}_{N \text{ times}})$ is a feasible solution to the N-session problem, this strict inequality contradicts the optimality of $(\underbrace{u^2, \ldots, u^2}_{K_2+K_3\ldots+K_N \text{ times}})$ in session $t = 1 + K_1$.

Case 2: $\bar{A}u^1 < \bar{A}u^2$. Let $v = \frac{K_1 u^1 + (K_2 + K_3 \ldots K_N) u^2}{N}$, and consider the alternative solution $\mathbf{u_{alt}} = (\underbrace{v, v, \ldots, v}_{N \text{ times}})$. Now, since constraints g_m are convex and symmetric, Jensen's inequality can be employed to show that this alternative solution is feasible to the N-session problem. Moreover, $\bar{A}v = \frac{\bar{A}(K_1 u^1 + (K_2 + \ldots + K_N) u^2)}{N} = \frac{K_1 \bar{A} u^1 + (K_2 + \ldots + K_N) \bar{A} u^2}{N} > \frac{K_1 \bar{A} u^1 + (K_2 + \ldots + K_N) \bar{A} u^1}{N} = \bar{A} u^1$.

Therefore, $f(v) > f(u^1)$. Consequently, $F(\mathbf{u_{alt}}) = \sum_{t=1}^{N} f(v) > \sum_{t=1}^{N} f(u^1) = F(\mathbf{u_1})$. This contradicts the optimality of $\mathbf{u_1}$ when $t = 1$.

Case 3: $\bar{A}u^1 = \bar{A}u^2$. In this case, $f(u^1) = f(u^2)$ and hence $F(\mathbf{u_1}) = F(\mathbf{u_2})$. Therefore, we can replace $\mathbf{u_2}$ with $\mathbf{u_1}$ when $t = 1 + K_1$.

Applying this argument recursively, we see that MPC cannot return a nonstationary solution that is strictly better than all stationary solutions. □

In view of this theorem, in order to force MPC to return a nonstationary sequence of fluence-maps with a high tumor-BE than any stationary fluence-map, we need to use a surrogate objective for which the conclusion of the theorem fails. We propose the total number of tumor cells remaining (TNTCR) objective from [38, 56] for this purpose. TNTCR is given by $\sum_{i=1}^{n} x_i^{N+1}$, where x_i^t denotes the total number of remaining tumor cells at the beginning of the tth session with dynamics $x_i^{t+1} = x_i^t \exp(-\alpha_0(A_i u^t) - \beta_0(A_i u^t)^2)$ for $i = 1, 2, \ldots, n$. In the next section, we study whether or not this surrogate objective induces MPC to return nonstationary fluence-maps with a higher tumor-BE than the best stationary fluence-map.

3.3 Numerical results

We conducted numerical experiments on five different head-and-neck test cases from [57]. These cases included four OAR: spinal cord (serial), brainstem (serial), left and right parotids (parallel). Parameters ρ_m were fixed at 1/3 for all OAR as is common in the clinical literature [23, 24, 25, 26]. All tolerance doses were set as in [57], which were in turn taken from a standard head-and-neck treatment protocol [47]. The number of sessions was fixed at $N = 35$. We also included a maximum BED constraint (corresponding to a total dose of 90 Gy in 35 sessions) on the tumor to facilitate dose homogeneity. We fixed $\alpha_0 = 0.35$ Gy^{-1} and $\beta_0 = 0.035$ Gy^{-2} as is standard in the clinical literature [23, 24, 25, 26]. The initial cell

density was assumed to be homogeneous over all tumor voxels. All computer simulations were performed on a 3.1 GHz iMac desktop with 16 GB RAM using MATLAB.

MPC was able to discover nonstationary solutions better than stationary solutions in three of the five cases. The percentage improvements attained by a nonstationary solution over the best stationary solution were 5.2%, 3.2%, and 2.1% for these three cases. In the other two cases, stationary solutions turned out to be superior than the nonstationary solutions returned by MPC by 3.2% and 0.92%. Figure 3.1 shows the tumor-BE for every voxel averaged over 35 sessions delivered by the nonstationary and best stationary solution for the first test-case. It illustrates that the nonstationary tumor-BE seems higher than the best stationary tumor-BE in this case. Figure 3.2 shows the nonstationary intensity profiles administered by MPC for ten sample tumor voxels. Figure 3.3 shows the nonstationary doses delivered to 500 sample tumor voxels by MPC. It also illustrates the average dose over all tumor voxels for this nonstationary solution (red line) as well as for the best stationary one (light blue line).

In summary, we conclude that MPC may be able to discover nonstationary fluence-maps with a higher tumor-BE than stationary ones. An interesting direction for future research would be to optimize the number of treatment sessions N by using nonstationary fluence-maps.

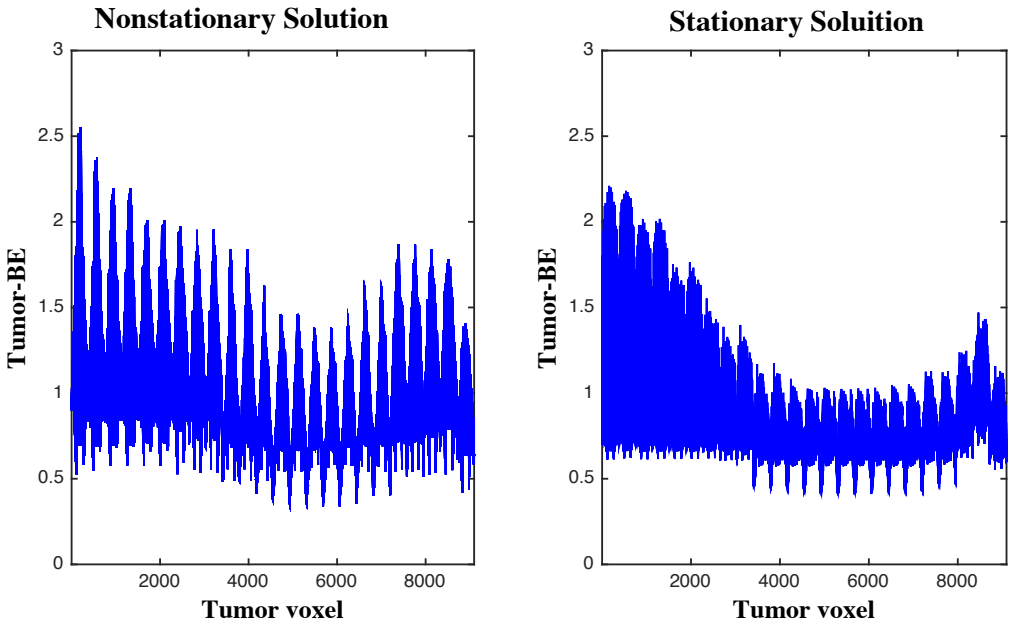


Figure 3.1: Tumor-BE averaged over 35 sessions administered by the nonstationary and best stationary solutions for test-case 1.

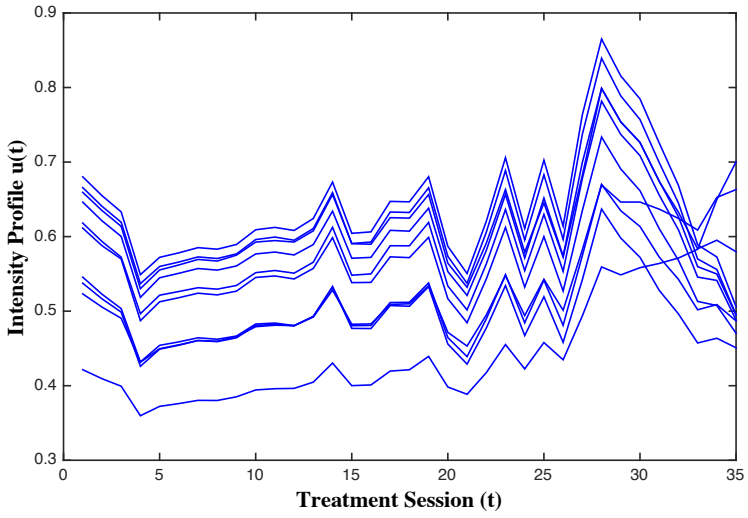


Figure 3.2: Nonstationary intensity profiles for ten sample tumor voxels.

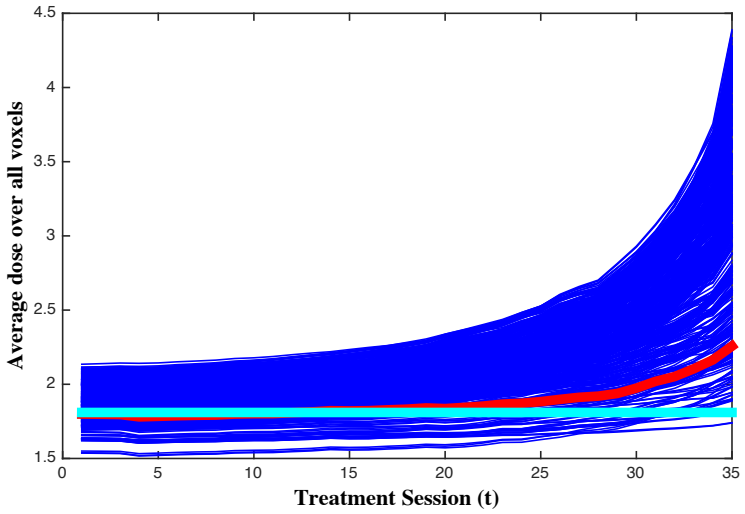


Figure 3.3: Average tumor doses delivered over 35 treatment sessions by the stationary (light blue line) and the nonstationary (red line) solutions. The dark blue lines show the doses administered by the nonostationary solution for 500 sample tumor voxels.

Chapter 4

ADAPTIVE TREATMENT-LENGTH OPTIMIZATION

4.1 Introduction

A recent direction of theoretical research on the spatiobiologically integrated side focuses on adaptive planning of fluence-maps [30, 36, 38, 39]. Unlike the traditional static-deterministic approach of adhering to a fluence-map that was calculated at the beginning of the treatment course, adaptive methods acknowledge that there is uncertainty in how tumors respond to radiation, and propose to recompute fluence-maps based on the *observed* evolution of the tumor's biological condition over multiple treatment sessions. Interest in this research stream is partly fueled by emerging advances in functional imaging techniques such as positron emission tomography (PET) and magnetic resonance spectroscopic imaging (MRSI). Functional images, at least in principle, allow the treatment planner to noninvasively and quantitatively track biological entities such as tumor cell density and oxygen pressure (which is known to affect radiosensitivity) over the treatment course [2, 7, 12, 13, 41, 60, 59, 63, 74, 75]. The idea is to obtain a new fluence-map by solving a new instance of a spatiobiologically integrated optimization problem using the most recent quantitative information about the tumor's biological condition as input, after each functional image is acquired. The standard objective in this spatiobiologically integrated optimization problem is to minimize the total number of tumor cells remaining (TNTCR) at the end of the treatment course subject to BED constraints on OAR. This adaptive methodology, although currently only theoretical, is intended as a small step toward patient-specific radiotherapy planning. One limitation of this research, however, is that the number of treatment sessions is assumed to be fixed at the beginning of the entire treatment course based on one-size-fits-all guidelines [47]. In particular, the number of treatment sessions is *not* determined adaptively. This paper proposes

a theoretical framework to addresses this limitation.

The primary research objective of this chapter is to test the hypothesis that adaptive treatment-length planning, when combined with adaptive fluence-map planning, decreases the TNTCR as compared to only using adaptive fluence-map planning. We test this hypothesis by (1) developing a spatiobiolgically integrated optimization formulation that attempts to minimize the TNTCR by adaptively selecting both the number of treatment sessions and the fluence-map, and (2) comparing the resulting TNTCR via computer simulations against an existing formulation from [36, 38] that also minimizes TNTCR by adaptively selecting only the fluence-map. These computer simulations are conducted on five well-studied phantom head-and-neck cancer cases from [54, 56, 57], each with three different tumor proliferation rates. The hypothesis is settled in the affirmative: adaptive treatment-length planning does reduce TNTCR. We also employ these computer simulations to compare two other secondary performance metrics of interest: the biological effect (BE) on the tumor [28, 33] and the length of the treatment course. These comparisons also favor adaptive treatment length planning, which attains higher tumor-BEs with shorter treatment courses. These results bode well for adaptive treatment-length planning, at least in theory, as the two methods compared have identical imaging requirements in each session.

4.2 Problem formulation and solution method

The mathematical notations in this chapter are standard in the literature on spatiobiologically integrated optimization [36, 38, 56, 57] and are similar to the ones given in previous chapters. It is redefined here nonetheless, for completeness.

Let $T_{\max}$ denote the maximum possible number of treatment sessions that the treatment planner is willing to administer (based on logistical or other clinical considerations). We think of $T_{\max}$ as a perhaps loose upper bound on the optimal number of treatment sessions. Treatment sessions are indexed by $t = 1, 2, \ldots, T_{\max}$ and are assumed to be administered once daily.

Let k denote the number of beamlets in the radiation field. The fluence-map in one

session is denoted by a nonnegative vector $u \in \Re_+^k$.

The tumor includes n equal-volume voxels indexed by $i = 1, 2, \ldots, n$. Each voxel has volume ν. Let A denote the nonnegative dose deposition matrix of size $n \times k$ for the tumor. Let A_i denote its ith row, which corresponds to the ith voxel. According to the standard linear dose deposition model, $A_i u$ is the dose deposited in voxel i in one session.

Let x_i^t denote the tumor cell density in the ith voxel as measured in a functional image acquired at the beginning of the tth treatment session. Then, according to the LQ model of dose-response, the tumor cell density at the end of the tth session is given by

$$x_i^{t+1} = x_i^t \exp(-\alpha_i^t[(A_i u) + r(A_i u)^2] + \ln 2/T_{\text{double}}),\ i = 1, 2, \ldots, n. \tag{4.1}$$

Here, α_i^t and $\beta_i^t = r\alpha_i^t$ are the (treatment planner's estimates of) linear and quadratic dose-response parameters at the beginning of session t. Thus, in this notation, $1/r$ is the so-called α/β ratio in radiobiology [24, 28, 73]. The term $\ln 2/T_{\text{double}}$ models tumor repopulation, where T_{double} is (an estimated value of) the doubling time in days [25, 28, 33]. As we shall see in more detail in Section 4.2, the formulations in this chapter do not make any explicit assumptions about the uncertainty in the tumor's response to radiation. They simply use cell densities observed in functional images to plan treatment using the above LQ model of dose-response with estimated parameter values. The chapter also does not take any position on how these parameter estimates are (or should be) obtained, but rather, leaves this to the treatment planner.

Let $O_1, O_2, \ldots, O_M$ denote the OAR in the irradiated region. Let $\mathcal{M} = \{1, 2, \ldots, , M\}$ be the set of indices of these OAR. For each $m \in \mathcal{M}$, n_m denotes the number of voxels in O_m. These voxels are indexed by $j = 1, 2, \ldots, n_m$. We assume A^m to be the $n_m \times K$ dose deposition matrix for O_m. Let A_j^m be the jth row of this matrix; this is the row corresponding to the jth voxel in O_m. That is, $A_j^m u$ is the dose delivered to the jth voxel in O_m in one session. Let α_m and β_m be the parameters of the LQ model for OAR O_m and define $\rho_m = 1/(\alpha_m/\beta_m)$ for brevity. Then, the BED delivered to the jth voxel in O_m in the

tth session is given by $A_j^m u + \rho_m (A_j^m u)^2$.

Two types of OAR constraints are considered: maximum dose (for serial OAR) and mean dose (for parallel OAR)[1]. As is standard in spatiobiologically integrated formulations, constraints are expressed in terms of BED equivalents as described next.

Let $\mathcal{M}_1 \subseteq \mathcal{M}$ be the set of indices of serial OAR with maximum dose constraints. These are the OAR whose function is impeded even when a single voxel is damaged by radiation. Examples include spinal cord and brainstem in head-and-neck cancer. We assume that for $m \in \mathcal{M}_1$, a total dose $D_{\max}^m$ can be tolerated by each voxel in serial OAR O_m if it is delivered in N_{conv}^m equal-dose sessions. The BED corresponding to this schedule equals

$$\text{BED}_{\max}^m = D_{\max}^m (1 + \rho_m (D_{\max}^m / N_{\text{conv}}^m)). \tag{4.2}$$

Let $z_j^{t,m}$ denote the BED delivered to voxel j in O_m in the first $t-1$ sessions. Then, a dose of $(A_j^m u)$ can be tolerated by this voxel in session t if

$$z_j^{t,m} + A_j^m u + \rho_m (A_j^m u)^2 \leq \text{BED}_{\max}^m, \;\; j = 1, 2, \ldots, m. \tag{4.3}$$

This implies that in order to verify the feasibility of u in session t, the treatment planner must know $\vec{z}^{t,m} = (z_1^{t,m}, \ldots, z_{n_m}^{t,m})$. In other words, the treatment planner must track $\vec{z}^{t,m}$ as treatment evolves, via the update formula

$$z_j^{t+1,m} = z_j^{t,m} + A_j^m u + \rho_m (A_j^m u)^2. \tag{4.4}$$

Similarly, let $\mathcal{M}_2 \subseteq \mathcal{M}$ be the set of indices of parallel OAR with mean dose constraints. These OAR can continue to function even if a sufficiently small proportion of voxels is damaged. Examples include parotid glands in head-and-neck cancer. Suppose that for $m \in M$, mean dose D_m can be tolerated by parallel OAR O_m, if it is delivered in N_{conv}^m

[1]A third type of OAR constraints, namely the dose-volume constraints, are omitted here for simplicity. They can, however, be incorporated by using the constraint-generation method from [56].

equal-dose sessions. The BED corresponding to this mean dose equals

$$\text{BED}^m_{\text{mean}} = D^m_{\text{mean}}(1 + \rho_m(D^m_{\text{mean}}/N^m_{\text{conv}})). \tag{4.5}$$

Let $v^{t,m}$ denote the total mean BED delivered over all voxels in O_m in the first $t-1$ sessions. Then, doses $(A^m_j u)$ delivered in session t to voxels j in O_m can be tolerated by O_m if

$$v^{t,m} + \frac{\sum_{j=1}^{n_m}(A^m_j u) + \rho_m \sum_{j=1}^{n_m}(A^m_j u)^2}{n_m} \leq \text{BED}^m_{\text{mean}}. \tag{4.6}$$

This implies that in order to verify the feasibility of u in session t, the treatment planner must know $\vec{v}^{t,m} = (v^{t,m}_1, \ldots, v^{t,m}_{n_m})$. Thus, the treatment planner must track $\vec{v}^{t,m}$ as treatment evolves, via the update formula

$$v^{t+1,m} = v^{t,m} + \frac{\sum_{j=1}^{n_m}(A^m_j u) + \rho_m \sum_{j=1}^{n_m}(A^m_j u)^2}{n_m}. \tag{4.7}$$

Our formulation also includes smoothness constraints on fluence-maps. This ensures that the fluence-maps obtained by our model can be delivered in practice using a multi-leaf collimator. These constraints bound the absolute relative difference between intensities of each pair of nearest neighbor beamlets by a fraction ϵ. This can be written as $Su \leq 0$, where S is a block diagonal matrix with entries $(1+\epsilon), (1-\epsilon), 0, +1$ at appropriate locations (see [57]).

After observing tumor cell densities x^t_i in voxels $i = 1, 2, \ldots, n$ in a functional image at the beginning of session t, the treatment planner chooses the remaining number of sessions N^t_{rem} and a corresponding fluence-map u for each of these sessions by solving the problem

$$(P^0_t) \quad \min_{u, N^t_{\text{rem}}} \sum_{i=1}^{n} \nu x^t_i \exp\left(-N^t_{\text{rem}}(\alpha^t_i[(A_i u) + r(A_i u)^2] + \ln 2/T_{\text{double}})\right) \tag{4.8}$$

$$z^{\ell+1,m}_j = z^{\ell,m}_j + (A^m_j u) + \rho_m (A^m_j u)^2, \ j = 1, 2, \ldots, n_m, \ m \in \mathcal{M}_1, \ \ell = t, \ldots, t + N^t_{\text{rem}} - 1, \tag{4.9}$$

$$z_j^{\ell+1,m} \leq \text{BED}_{\max}^m,\ j = 1, 2, \ldots, n_m,\ m \in \mathcal{M}_1,\ \ell = t, \ldots, t + N_{\text{rem}}^t - 1, \tag{4.10}$$

$$v^{\ell+1,m} = v^{\ell,m} + \frac{\sum_{j=1}^{n_m}(A_j^m u) + \rho_m \sum_{j=1}^{n_m}(A_j^m u)^2}{n_m},\ m \in \mathcal{M}_2,\ \ell = t, \ldots, t + N_{\text{rem}}^t - 1, \tag{4.11}$$

$$v^{\ell+1,m} \leq \text{BED}_{\text{mean}}^m,\ m \in \mathcal{M}_2,\ \ell = t, \ldots, t + N_{\text{rem}}^t - 1, \tag{4.12}$$

$$Su \leq 0, \tag{4.13}$$

$$u \geq 0, \tag{4.14}$$

$$u \in \Re^K, \tag{4.15}$$

$$N_{\text{rem}}^t \in \{1, 2, \ldots, T_{\max} - t + 1\}. \tag{4.16}$$

The objective in this problem is to minimize the TNTCR at the end of N_{rem}^t additional treatment sessions. Note that the expression in this objective is calculated by a repeated application (N_{rem}^t times) of the cell density update formula (4.1), and then summing the cell densities over all voxels. Equation (4.9) is simply the update formula (4.4) for sessions $t, t+1, \ldots, t + N_{\text{rem}}^t - 1$. Inequality (4.10) is the maximum dose constraint as explained in (4.3), written for sessions $t, t+1, \ldots, t + N_{\text{rem}}^t - 1$. Equation (4.11) is the update formula (4.7) for sessions $t, t+1, \ldots, t + N_{\text{rem}}^t - 1$. Inequality (4.12) is the mean dose constraint as explained in (4.6), written for sessions $t, t+1, \ldots, t + N_{\text{rem}}^t - 1$. Inequality (4.13) is the smoothness constraint.

Suppose $u^*, N_{\text{rem}*}^t$ is an optimal solution to this problem. Then, the treatment planner administers fluence-map u^* in session t; and employs formulas (4.4) and (4.7) with $u = u^*$ to update OAR BED values. The tumor then responds to u^* and the cell densities evolve to their new values by the beginning of the next session. The treatment planner observes these cell densities in an image in the next session, and this process of imaging-optimizing-updating repeats until either $N_{\text{rem}*}^t = 1$ is discovered to be optimal at some t or the upper limit of $T_{\max}$ sessions is reached. Specifically, the treatment planner does not "commit" to administering $N_{\text{rem}*}^t$ sessions at the beginning of session t. The number $N_{\text{rem}*}^t$ nevertheless plays an important role in planning through its effect on the choice of an optimal u^* in

problem (P_t^0). Also note that this approach does not rely on any assumptions about the uncertainty in the tumor's response to radiation. It simply uses observed cell densities to plan treatment using the LQ model with estimated values of dose-response parameters.

It is worthwhile to distinguish this formulation from the optimal stopping approach that was originally envisioned in the doctoral dissertation of our collaborator Dr. Kim's [36]. In that formulation, the treatment planner (at the beginning of each session) decided whether or not to stop treatment. It turned out that it was always optimal to continue treatment in their formulation as long as there was room in the OAR constraints. That is, as long as the OAR could tolerate additional dose. The treatment planner decided to stop only when faced with a "cost" of continuing treatment. Problem (P_t^0) pursues a different approach in that the treatment planner attempts to answer a different question: what the treatment length should be at this particular session; in other words, would it be better to continue treatment for one more session, or two more sessions, or three more sessions, and so on. As such, in problem (P_t^0), the treatment planner does not solve an optimal stopping problem, but rather, solves a treatment-length planning problem.

Problem (P_t^0) can be simplified as follows. Consider any $m \in \mathcal{M}_1$ and any voxel j in O_m. Equation (4.9), and nonnegativity of the dose-deposition matrices and the fluence-maps imply that $z_j^{\ell+1,m} \geq z_j^{\ell,m}$, for $\ell = t, t+1, \ldots, t+N_{\text{rem}}^t - 1$. Also note that the right hand side $\text{BED}_{\max}^m$ of constraint (4.10) does not depend on ℓ. Therefore, constraint (4.10) holds for $\ell = t, t+1, \ldots, t+N_{\text{rem}}^t - 1$ if and only if it holds when $\ell = t + N_{\text{rem}}^t - 1$. Consequently, (4.9) and (4.10) can be replaced by a single constraint

$$z_j^{t,m} + N_{\text{rem}}^t \left((A_j^m u) + \rho_m (A_j^m u)^2 \right) \leq \text{BED}_{\max}^m, \ j = 1, \ldots, n_m, \ m \in \mathcal{M}_1. \tag{4.17}$$

By a similar logic, (4.11) and (4.12) can be replaced by a single constraint

$$v^{t,m} + N_{\text{rem}}^t \frac{\left(\sum_{j=1}^{n_m} (A_j^m u) + \rho_m \sum_{j=1}^{n_m} (A_j^m u)^2 \right)}{n_m} \leq \text{BED}_{\text{mean}}^m. \tag{4.18}$$

Furthermore, since $(A_j^m u) \geq 0$, inequality (4.17) can be equivalently rewritten as the linear constraint

$$A_j^m u \leq \frac{-1 + \sqrt{1 + 4\rho_m(\mathrm{BED}_{\max}^m - z_j^{t,m})/N_{\mathrm{rem}}^t}}{2\rho_m},\ j = 1, 2, \ldots, n_m,\ m \in \mathcal{M}_1, \tag{4.19}$$

by solving a quadratic equation. This discussion allows us to rewrite problem (P_t^0) as

$$(P_t) \min_{u, N_{\mathrm{rem}}^t} \sum_{i=1}^{n} v x_i^t \exp\left(-N_{\mathrm{rem}}^t(\alpha_i^t[(A_i u) + r(A_i u)^2] + \ln 2/T_{\mathrm{double}})\right), \tag{4.20}$$

$$A_j^m u \leq \frac{-1 + \sqrt{1 + 4\rho_m(\mathrm{BED}_{\max}^m - z_j^{t,m})/N_{\mathrm{rem}}^t}}{2\rho_m},\ j = 1, 2, \ldots, n_m,\ m \in \mathcal{M}_1, \tag{4.21}$$

$$\sum_{j=1}^{n_m}(A_j^m u) + \rho_m \sum_{j=1}^{n_m}(A_j^m u)^2 \leq n_m \left(\frac{\mathrm{BED}_{\mathrm{mean}}^m - v^{t,m}}{N_{\mathrm{rem}}^t}\right),\ m \in \mathcal{M}_2, \tag{4.22}$$

$$Su \leq 0, \tag{4.23}$$

$$u \geq 0, \tag{4.24}$$

$$N_{\mathrm{rem}}^t \in \{1, 2, \ldots, T_{\max} - t + 1\}. \tag{4.25}$$

This problem is convex if $\alpha_i^t r \geq 2$ for all i (see [38, 56] for proof). This condition is believed to be met in practice for most tumors including head-and-neck and lung [73]. In that case, the problem can be solved efficiently via a log-barrier interior point algorithm with Newton's method and backtracking line search as described in detail in [17, 56].

We performed computer simulations to compare the performance of this adaptive treatment-length planning method against a method that does not adapt the treatment length. Specifically, this base-case method solves problem (P_t) in each session, but with one crucial difference: the minimization is not performed over N_{rem}^t; instead, the total number of treatment sessions is fixed (before the treatment course begins) at a value N_{guide} recommended by existing one-size-fits-all guidelines [47]. Thus, at the beginning of session t, a treatment planner

implementing this non-adaptive approach to treatment-length planning solves the problem

$$(Q_t)\min_{u}\ \sum_{i=1}^{n} v x_i^t \exp\left(-(N_{\text{guide}}-t+1)(\alpha_i^t[(A_i u)+r(A_i u)^2]+\ln 2/T_{\text{double}})\right), \tag{4.26}$$

$$A_j^m u \le \frac{-1+\sqrt{1+4\rho_m(\text{BED}_{\max}^m - z_j^{t,m})/(N_{\text{guide}}-t+1)}}{2\rho_m},\ j=1,2,\ldots,n_m,\ m\in\mathcal{M}_1, \tag{4.27}$$

$$\sum_{j=1}^{n_m}(A_j^m u)+\rho_m\sum_{j=1}^{n_m}(A_j^m u)^2 \le n_m\left(\frac{\text{BED}_{\text{mean}}^m - v^{t,m}}{N_{\text{guide}}-t+1}\right),\ m\in\mathcal{M}_2, \tag{4.28}$$

$$Su \le 0, \tag{4.29}$$

$$u \ge 0. \tag{4.30}$$

Note that both methods have identical functional imaging requirements because they both adapt the fluence-maps u to the tumor cell density observed at the beginning of each session. Numerical results are presented in the next section.

4.3 Numerical results

We tested our approach on five phantom head-and-neck test cases from [54, 56, 57]. All cases used seven equally spaced coplanar beams. The beamlet resolution was 5×5 mm^2. All voxels were $5\times 5\times 5$ mm^3, leading to voxel volume $\nu = 125$mm^3. All cases included two serial OAR (spinal cord and brainstem), and two parallel OAR (left and right parotids). The total number of voxels in the head-and-neck target and the total number of beamlets are shown in Table 4.1.

The number of conventional sessions N_{conv}^m was fixed at 35 for all OAR $m\in\mathcal{M}$. For the spinal cord and brain stem, the maximum tolerable doses $D_{\max}$ were set to 45 Gy and 50 Gy, respectively. The tolerable mean doses D_{mean} for the left and right parotids were set to 28 Gy. The α/β ratios for all OARs were fixed at 3 Gy and the α/β for tumor was fixed at $r=10$. We executed the algorithms for three different $T_{\text{double}} = 2, 10, 20$ days, which represent

Table 4.1: Description of five head-and-neck cancer cases.

Case #	# of beamlets (k)	# of tumor voxels (n)
1	1680	9096
2	1485	6384
3	1500	7248
4	1584	7350
5	1746	8640

fast, medium, and slowly proliferating tumors, respectively. All these values of problem parameters were based on the existing clinical literature [22, 24, 26, 27, 28, 33, 47, 73]. The maximum possible number of sessions $T_{\max}$ was fixed at 35, the same as the conventional number of sessions. This choice embodies the idea that the treatment planner does not want to employ a treatment course that is longer than conventional. The number of sessions recommended by one-size-fits-all guidelines, which is denoted N_{guide}, was also fixed at the conventional value of 35 [22, 33, 47].

The values of α_i^t were sampled from a truncated normal distribution in the range $[0.1, 0.6]$ with mean 0.35 Gy^{-1} and standard deviation 0.08 Gy^{-1}, respectively. These choices were motivated by typical values reported in the literature for head-and-neck tumors [27, 73]

For each T_{double} value in each of the five cases, we ran the adaptive and the non-adaptive algorithms for 30 independent simulation runs. The resulting mean TNTCR was estimated by averaging the TNTCR values achieved in these simulations. Each run used an independent sample of size n (the number of tumor voxels) from the aforementioned normal distribution in each treatment session, to simulate the tumor's response to radiation.

Using a two-sided t-test at the 95% confidence level, we examined the null hypothesis H_0 that the mean TNTCR for the adaptive method μ_1 is equal to that of the non-adaptive method μ_2. That is, the tested hypothesis was $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$. The percentage relative difference in average TNTCR attained by these two procedures over 30 independent

simulations was calculated as

$$\text{RelDiff}_{\text{TC}}^{\square} = \frac{\sum_{\text{sim}=1}^{30} \text{TNTCR}_{\text{Non-adaptive}}^{\square}(\text{sim}) - \sum_{\text{sim}=1}^{30} \text{TNTCR}_{\text{Adaptive}}^{\square}(\text{sim})}{\sum_{\text{sim}=1}^{30} \text{TNTCR}_{\text{Adaptive}}^{\square}} \times 100, \tag{4.31}$$

where TNTCR(sim) is the TNTCR for simulation sim $= 1, 2, \ldots, 30$, and $\square$ denotes each combination of case and T_{double}. We have used TC (tumor cells) as a shorthand for TNTCR in the subscript of RelDiff. The non-adaptive method is also referred to as the "Fixed N" method in some of our figures below.

Table 4.2 displays this percentage relative difference for all five cases and all T_{double} values. The symbol $\times$ in that table marks scenarios where the null hypothesis H_0 was rejected. This implies that in those scenarios, the mean TNTCR for the adaptive and non-adaptive methods was not equal. Similarly, the symbol ✓ marks scenarios where there was no significant difference between the mean TNTCR of the adaptive and non-adaptive algorithms (H_0 was not rejected). Out of the 15 different combinations tested, H_0 was rejected in 10, and the TNTCR obtained by the adaptive algorithm was superior to that achieved by the non-adaptive method. Note that in the 5 scenarios where H_0 was not rejected, the mean TNTCR of both algorithms are deemed equal. This implies that in all combinations tested, the adaptive method was at least as good as the non-adaptive method. The percentage relative difference in average TNTCR ranged from $[-2.45\%, 19.63\%]$, and the average over all 15 scenarios was 4.94%.

Note that there is a noticeable improvement in the performance of the adaptive method for Cases 3 and 5 as compared to Cases 1, 2, and 4. The most dominant case turned out to be Case 5 in which the average $\text{RelDiff}_{\text{TC}}$ (over all values of T_{double}) was 11.42%. This could be rooted in the fact that the anatomy of Cases 3 and 5 is a bit different from that of Cases 1, 2, and 4 (see [56]). The spinal cord and brainstem are closer to the tumor in Cases 1, 2, and 4 than in Cases 3 and 5. Thus, Cases 3 and 5 offer more leeway for improvement via

Table 4.2: Percentage relative difference in TNTCR between two methods. The second column under each T_{double} shows whether the difference between the means of two methods was statistically significant (×: significant; ✓: not significant).

Case	T_{double} (days)					
	2		10		20	
	$\text{RelDiff}_{\text{TC}}$	$H_0 : \mu_1 = \mu_2$	$\text{RelDiff}_{\text{TC}}$	$H_0 : \mu_1 = \mu_2$	$\text{RelDiff}_{\text{TC}}$	$H_0 : \mu_1 = \mu_2$
1	-2.45%	✓	5.32%	×	5.43%	×
2	2.71%	✓	-1.54%	✓	5.71%	×
3	3.73%	×	5.32%	×	9.38%	×
4	-1.78%	✓	2.45%	✓	4.74%	×
5	4.20%	×	10.43%	×	19.63%	×

better planning.

Another interesting trend in Table 4.2 is that the improvement achieved by the adaptive algorithm is larger for larger T_{double} values. This is better illustrated in Figure 4.1 for Cases 3 and 5, where the vertical separation between the TNTCR attained by the two methods increases as T_{double} increases. One possible explanation is that it is generally better to employ a longer treatment course for slower-growing tumors, that is, those with larger T_{double} values (see the literature on the fractionation problem reviewed in Section 4.1). The adaptive method thus has more opportunities to re-plan for these tumors. This can be verified by comparing the $N^t_{\text{rem}*}$ values computed by the adaptive algorithm for different T_{double} values, as depicted in Figure 4.2. This figure illustrates how the algorithm adapts at each session. For example, as can be seen in Figure 4.2a, when $T_{\text{double}} = 20$ days, $N^t_{\text{rem}*} = 8$ for $t = 1$. That is, at the start of the treatment course, the treatment planner administers a fluence-map u^* that would be optimal if there were 8 more sessions in the course. At the start of session 2, after observing the cell densities in the latest functional image, it turns out that $N^2_{\text{rem}*} = 9$. The treatment planner then administers a u^* that would be optimal if there were 9 more sessions in the course. This continues until session 15, where it is decided that treatment should be stopped at the end of that session. The figure also shows, as expected, that the eventual treatment-length increases with increasing T_{double}.

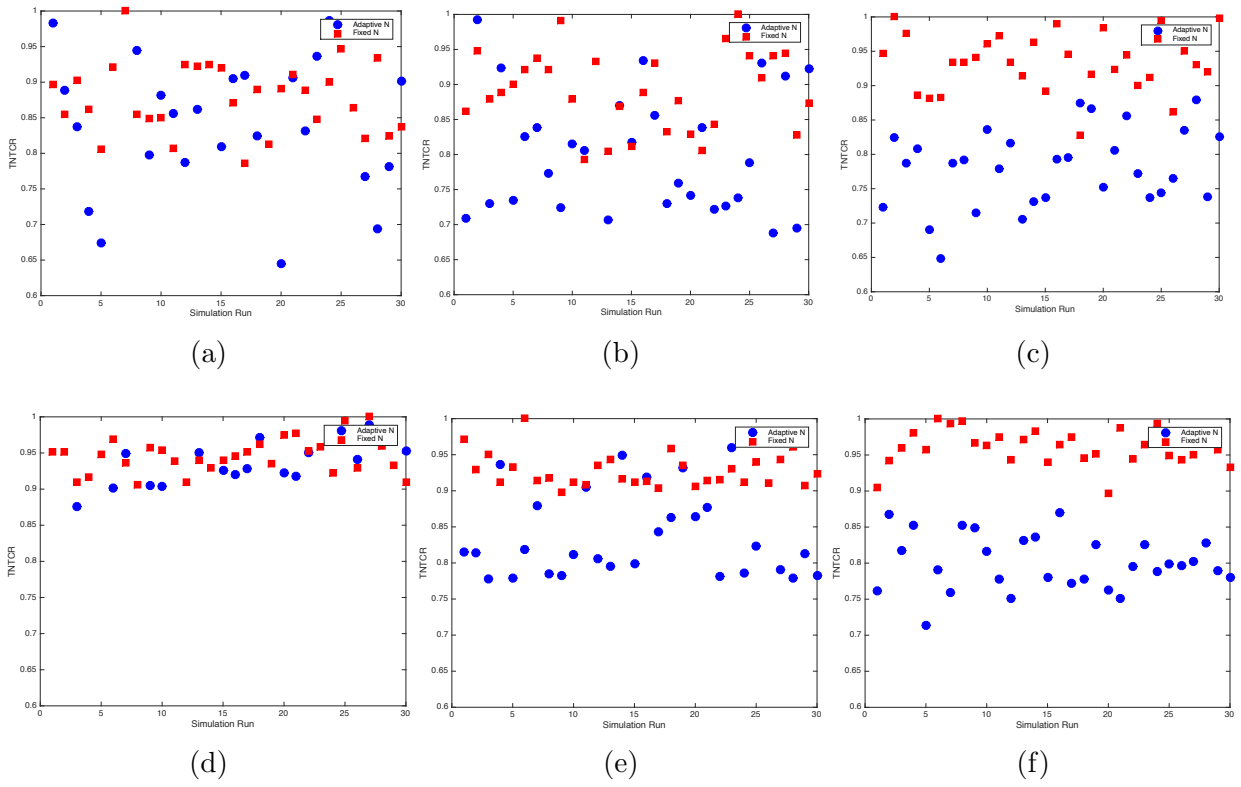


Figure 4.1: Normalized TNTCR in each simulation run for Case 3 and Case 5 in rows and $T_{\text{double}} = 2, 10, 20$ days in columns.

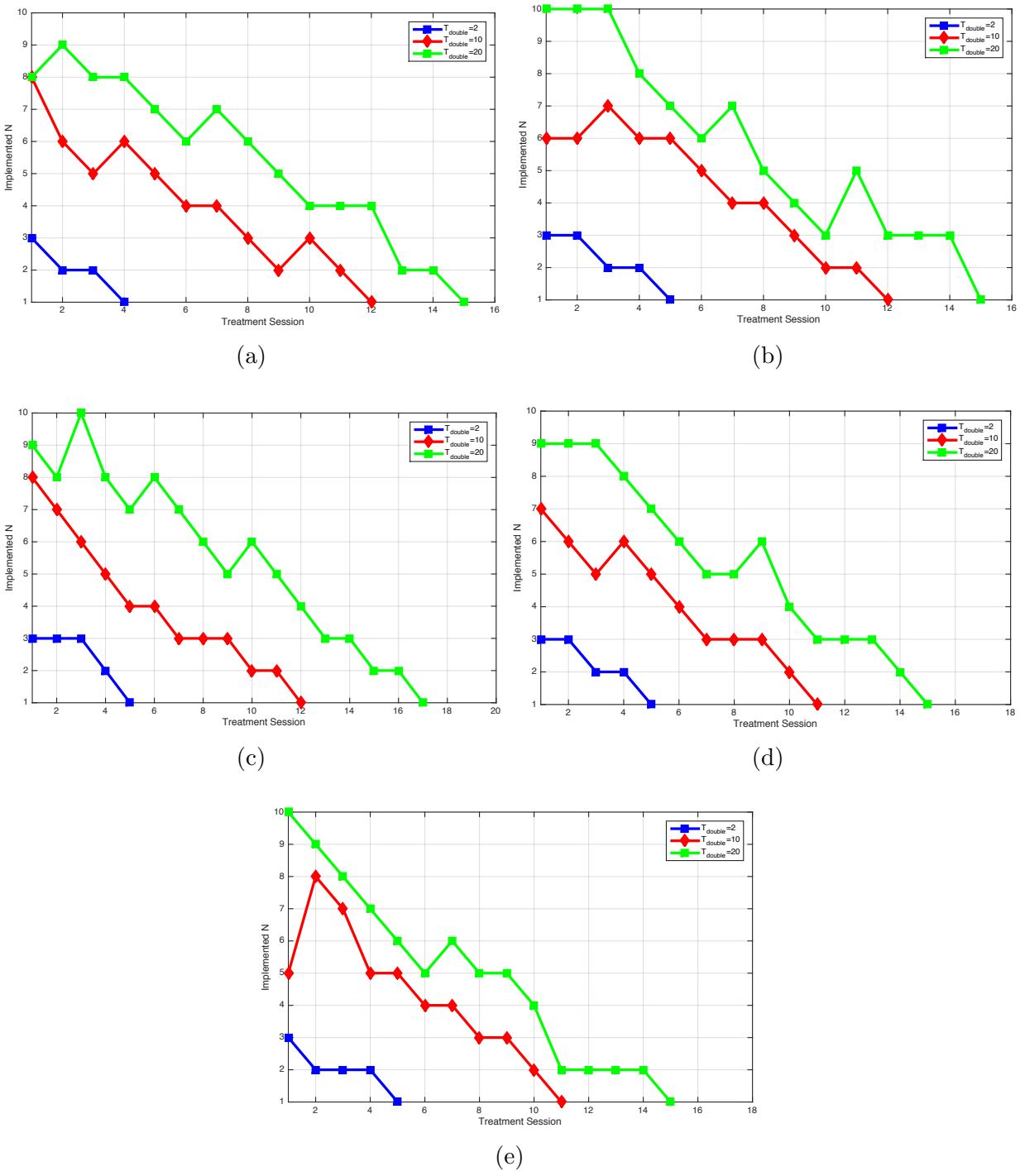


Figure 4.2: Evolution of $N^t_{\mathrm{rem}*}$ (called "implemented N" on the Y axis) over the treatment course for one representative simulation for Cases 1-5 in (a)-(e) for different T_{double} values in days.

Table 4.3: Number of sessions ($N99$) required by both algorithms to kill at least 99% of the total tumor cells.

Case	T_{double} (days)					
	2		10		20	
	Adaptive	Non-adaptive	Adaptive	Non-adaptive	Adaptive	Non-adaptive
1	2	10	4	10	4	10
2	2	10	3	12	4	12
3	2	8	3	9	4	11
4	3	11	4	13	4	16
5	2	10	3	10	4	12

Table 4.4: Ultimate treatment length by the adaptive algorithm.

Case	$T_{\textbf{double}}$		
	2	10	20
1	4	11	15
2	5	12	17
3	5	12	16
4	5	11	15
5	5	11	14

The number of sessions needed for each algorithm to kill at least 99% of the tumor cells (denoted $N99$) is reported in Table 4.3. The adaptive algorithm consistently needs between $[2, 4]$ sessions (average 3.2), whereas the non-adaptive algorithm requires between $[8, 16]$ sessions (average 10.9). Ultimate treatment length of the adaptive algorithm is reported in Table 4.4. The adaptive algorithm consistently needs between 4-5 sessions to terminate the treatment when T_{double} is 2 days. These numbers for $T_{\text{double}} = 10$ and 20 are 11-12 and $[14, 17]$. Note that all these treatment lengths are much shorter than the 35 sessions of the nonadaptive method.

Another performance metric that might be of interest is the total average biological effect

Table 4.5: Percentage relative difference in Tumor-BE between two methods. The second column under each T_{double} shows whether the difference between the means of two methods was statistically significant (×: significant; ✓: not significant).

Case	T_{double} (days)					
	2		10		20	
	$\text{RelDiff}_{\text{BE}}$	$H_0 : \mu_1 = \mu_2$	$\text{RelDiff}_{\text{BE}}$	$H_0 : \mu_1 = \mu_2$	$\text{RelDiff}_{\text{BE}}$	$H_0 : \mu_1 = \mu_2$
1	-2.57%	✓	4.01%	×	10.94%	×
2	7.93%	×	7.49%	×	10.32%	×
3	0.34%	✓	7.49%	×	12.51%	×
4	0.77%	✓	-1.50%	✓	5.87%	×
5	13.34%	×	10.68%	×	14.44%	×

(BE) of the dose delivered to the tumor. This metric can be calculated as follows:

$$\text{BE} = \sum_{t=1}^{T} \frac{\sum_{i=1}^{n} \alpha_i^t [A_i u^t + r(A_i u^t)^2]}{n}. \tag{4.32}$$

Although this metric was not explicitly optimized by the two algorithms, it is compared here to gain further insight. We again tested the hypothesis H_0 that the mean BE for both methods are equal. The result of these tests are reported in Table 4.5. Similar to Table 4.2, the adaptive method is able to outperform the non-adaptive method in 11 out of the 15 scenarios. A formula similar to (4.31) was employed to calculate the percentage relative difference in BE (denoted by $\text{RelDiff}_{\text{BE}}$) for both methods. The average percentage relative difference ranged from $[-2.57\%, 14.44\%]$ and the average over all 15 scenarios was 6.80%. The trends discussed for Table 4.2 are also true in this table, which is not surprising, given the close relationship between BE and TNTCR. Figure 4.3 shows the total BE delivered to each tumor voxel over the treatment course. In a majority of the voxels, the total tumor BE delivered by the adaptive algorithm is larger compared to that by the non-adaptive method.

In summary, despite terminating treatment much earlier, the adaptive method was able to match or better the performance of the non-adaptive method, both in terms of TNTCR and BE, in all cases.

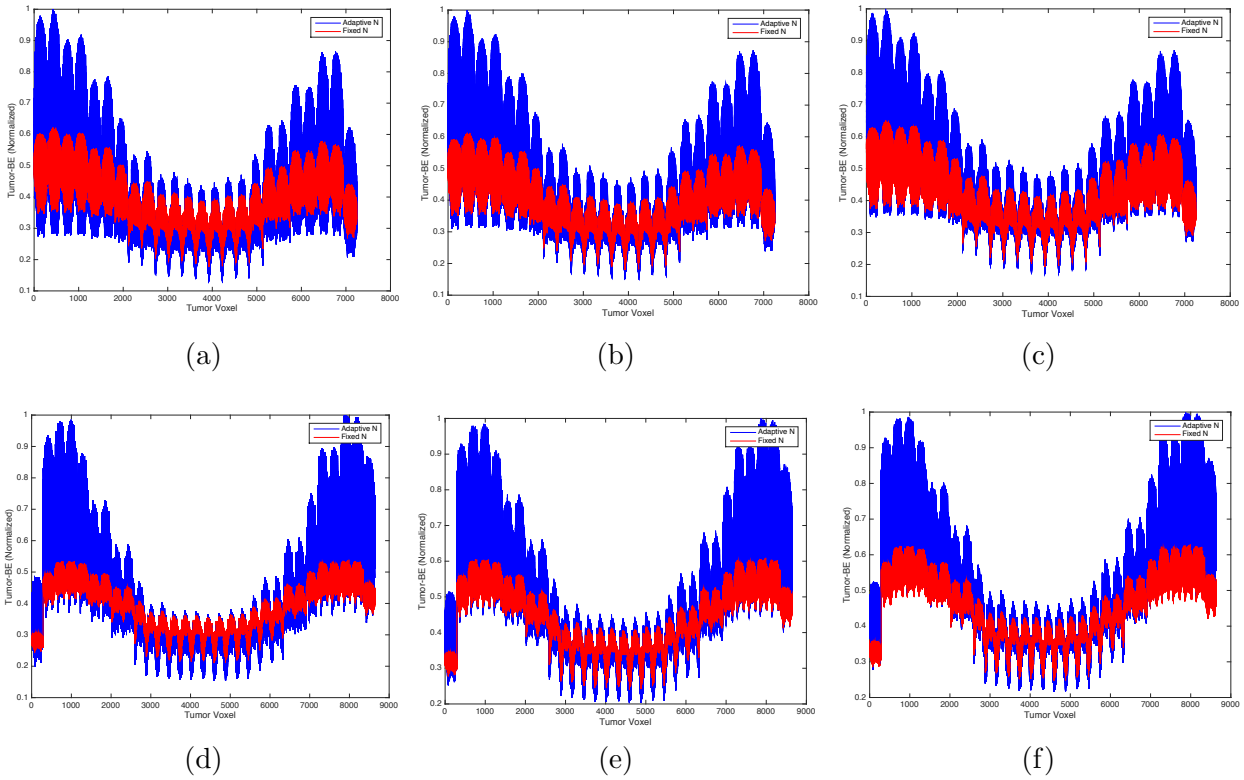


Figure 4.3: Normalized total tumor-BE over the treatment course in each voxel averaged over 30 independent simulations for Case 3 and Case 5 in rows and $T_{\text{double}} = 2, 10, 20$ days in columns.

4.4 Conclusions

In the existing optimization literature on spatiobiologically integrated radiotherapy, the length of the treatment course is determined and fixed before the course begins. Recent advances in functional imaging have paved the way for quantitatively tracking tumor cell densities in each voxel and adapting the treatment plan accordingly. Adaptive optimization models in spatiobiologically integrated radiotherapy have thus far focused on methods that employ such functional images to tune fluence-maps over the treatment course. The hope in these models is to deliver the right dose to the right patient at the right time, thereby bringing us closer to patient-response-specific radiotherapy planning.

In this chapter, we proposed a method to adapt treatment-length (in addition to adapting fluence-maps) based on the observed evolution of tumor cell densities. Performance of this adaptive method was benchmarked against a non-adaptive method that does not tune the treatment-length (but does adapt fluence-maps). The two methods were compared via computer simulations on five phantom head-and-neck cancer cases for three different tumor proliferation rates. Our simulations suggest that the adaptive method outperformed the non-adaptive method both in terms of TNTCR and tumor BE. This superior performance was despite the fact that the adaptive method utilized much shorter treatment courses.

Chapter 5

LEARNING TUMOR DOSE-RESPONSE UNCERTAINTY

5.1 Introduction

Existing theoretical studies investigates three types of spatiobiologically integrated formulations: deterministic [1, 57, 69, 67, 67, 65, 68], robust (Section 2.2 of Chapter 2), and stochastic dynamic (Chapter 4 and [38, 39, 56]). In deterministic formulations, dose-response parameters of the LQ model are assumed to be known, tumor- and OAR-specific constants. Point-estimates for these parameter values are assumed to be available from previous population-based clinical studies. One limitation of such deterministic formulations is of course that they entirely ignore the uncertainty in these estimates. Furthermore, such estimates are not specific to the individual patient. Robust formulations attempt to address this by instead assuming that the parameter values are only known to belong to some uncertainty sets, for instance, intervals. The treatment planner then finds the best fluence-map that remains feasible no matter which specific values the parameters take from these uncertainty sets. This approach is often viewed as too conservative. Owing to recent advances in non-invasive functional imaging techniques such as Positron Emission Tomography (PET) or Magnetic Resonance Spectroscopic Imaging (MRSI) that at least in principle can monitor a tumor's biological response over the treatment course, deterministic and robust formulations are categorized as non-adaptive. That is, they do not adapt the fluence-map to the evolution of the tumor's biological response as *observed* in functional images acquired repeatedly over the treatment course. Stochastic dynamic formulations overcome this limitation by modeling the uncertainty in tumor-response via a probability distribution. This results in a stochastic control model, where the states include the tumor's biological condition as observed in functional images and the control variables correspond to fluence-maps. One limitation

of such formulations is that the probability distribution of tumor-response uncertainty is assumed to be known at the beginning of the treatment course. This distribution would need to be estimated from population-based studies and hence may not represent an individual patient's dose-response behavior. To address this limitation, a Bayesian framework that attempts to learn an individual patient's dose-response while simultaneously calculating adaptive fluence-maps over the treatment course was briefly envisioned in the doctoral dissertation of our collaborator Dr. Saberian [53]. This chapter fully executes that vision and provides extensive numerical results to gain insight into the potential benefits of learning tumor-response.

Specifically, the stochastic control formulation in [38, 56] provides a starting point for this chapter. The treatment planner begins with a prior belief about the probability distribution of an uncertain tumor-response parameter. This prior is updated based on the uncertain evolution of tumor cell density as observed in functional images over the treatment course. The prior is chosen to be a conjugate to the dose-response probability distribution. This leads to a Bayesian stochastic control formulation wherein the state includes (hyper)parameters that characterize the prior and that can be updated using a simple formula. Exact solution of this problem is computationally intractable owing to its high-dimensional and continuous state-space and action-space. An approximate learning and optimization algorithm rooted in Certainty Equivalent Control (CEC) [9] is therefore proposed. This algorithm's performance is evaluated against two other methods: (i) a "clairvoyant" CEC method from [38, 56], where the treatment planner is assumed to know the probability distribution of dose-response; and (ii) a "no learning" CEC approach from [38], which uses a fixed value of dose-response over the treatment course. Since the clairvoyant CEC method is currently the best known approach for approximate adaptive optimization of fluence-maps in stochastic dynamic spatiobiologically integrated problems, it serves as an idealized, hypothetical "gold-standard" for the learning approach that does not know (and hence must estimate) the dose-response distribution on the fly while simultaneously and adaptively optimizing fluence-maps. On the other extreme, the no learning approach is a simple CEC method for approximately

solving spatiobiologically integrated problems. It is something a practitioner without knowledge of stochastic control would implement, and hence serves as a natural benchmark. The three methods have identical functional imaging requirements. Numerical simulations using three dose-response probability distributions (symmetric, left-skewed, and right-skewed) are presented. The main conclusion in all three cases is that (i) there is a statistically significant difference between the tumor-kill achieved by the learning method and the no learning method; in particular, the learning method achieves a higher tumor-kill; and (ii) there is no statistically significant difference between the learning and the clairvoyant methods. This bodes well for the learning method, at least in theory.

5.2 Problem formulation

Spatiobiologically integrated mathematical formulations for adaptively optimizing fluence-maps are presented in this section. Section 5.2.1 focuses on a clairvoyant formulation, where the treatment planner is assumed to know the probability distribution of tumor-response uncertainty. A simpler version of this clairvoyant formulation was first envisioned in [38], and a more complicated version specifically tailored for adapting fluence-maps to the uncertain evolution of tumor hypoxia (lack of oxygen, which is a known cause of tumor radioresistance [59, 60, 61, 62]) was developed in [56]. Although the clairvoyant formulation is not the main topic of this chapter, its description is needed to fully understand the Bayesian learning extension in Section 5.2.2, where the treatment planner does not know the probability distribution of tumor-response uncertainty.

5.2.1 Clairvoyant stochastic control

The conceptual idea in the stochastic control formalism in [38, 56] is to view the tumor and OAR as a dynamic system whose state evolution over the treatment course is influenced by the control variables chosen by the treatment planner. The goal is to optimize some quantitative measure of treatment efficacy. Details of this formalism are presented here briefly for the clairvoyant case using notation borrowed from [56].

Consider a treatment course with T sessions indexed by $t = 1, 2, \ldots, T$. The number of beamlets in the radiation field is k. The fluence-map, that is, the beamlet intensity vector, chosen by the treatment planner in the tth treatment session is denoted by $u^t \in \Re_+^k$. These u^t, for $t = 1, 2, \ldots, T$, are the control variables.

The tumor consists of n equal-volume voxels (each with volume ν) indexed by $i = 1, 2, \ldots, n$. Let A denote the $n \times k$ tumor dose-deposition coefficient matrix and let A_i denote its ith row corresponding to the ith tumor voxel. According to the standard linear dose-deposition model, $A_i u^t$ is the dose deposited in the ith tumor voxel in the tth session. Let α_i^t and β_i^t denote the tumor's dose-response parameters for voxel i in session t as per the LQ model. Since it is common in the radiobiology literature to report α/β ratios as tissue-dependent constants (see [73], for instance), it will be convenient to express $\beta_i^t = r\alpha_i^t$. Here, $1/r$ is the α/β ratio. As in [38], α_i^t are modeled as independent and identically distributed (iid) random variables across i and t. Suppose that the probability distribution of these random variables is categorical with J possible values $\{b_1, b_2, \ldots, b_J\}$. Further, suppose that the probability that α_i^t equals b_ℓ is p_ℓ, for $\ell = 1, 2, \ldots, J$. This distribution is denoted by categorical$(\vec{p}, \vec{b})$. This specific choice for the distribution of α_i^t is not essential to describe the clairvoyant formulation in this section. The categorical distribution is the most general discrete distribution with finite support. It is therefore utilized here for concreteness. A continuous distribution could also be used with minimal changes to notation. In the clairvoyant formulation, it is assumed that the treatment planner knows the probabilities $\vec{p} = (p_1, p_2, \ldots, p_J)$ and the corresponding parameter values $\vec{b} = (b_1, b_2, \ldots, b_J)$.

The tumor state for voxel i equals the cell density x_i^t in that voxel as observed in a functional image taken prior to treatment session t. According to the LQ model, the tumor cell dynamics during session t are then given by[1]

$$x_i^{t+1} = x_i^t \exp(-\alpha_i^t[(A_i u^t) + r(A_i u^t)^2]),\ i = 1, 2, \ldots, n. \tag{5.1}$$

[1]A tumor proliferation term is omitted here for notational simplicity. It can be easily incorporated without any changes to the methodology.

The cell density vector $\vec{x}^t = (x_1^t, x_2^t, \ldots, x_n^t) \in \Re_+^n$ defines the state of the tumor at the beginning of session t. As in [38, 56], the treatment planner's objective is to minimize the expected total number of tumor cells remaining (TNTCR) at the end of T sessions, where $\text{TNTCR} = \sum_{i=1}^n \nu x_i^{T+1}$.

Let $O_1, O_2, \ldots, O_M$ denote the M different OAR under consideration. Let $\mathcal{M} = \{1, 2, \ldots, , M\}$ be the set of indices of these OAR. For each $m \in \mathcal{M}$, n_m denotes the number of voxels in O_m. These voxels are indexed by $j = 1, 2, \ldots, n_m$. Let A^m be the $n_m \times K$ dose deposition matrix for O_m. Let A_j^m be the jth row of this matrix; this is the row corresponding to the jth voxel in O_m. That is, $A_j^m u^t$ is the dose delivered to the jth voxel in O_m in session t. Let α_m and β_m be the parameters of the LQ model for OAR O_m and define $\rho_m = 1/(\alpha_m/\beta_m)$ for brevity. Then, the BED delivered to the jth voxel in O_m in the tth session is given by $A_j^m u^t + \rho_m (A_j^m u^t)^2$.

Two types of OAR constraints are considered: maximum dose (for serial OAR) and mean dose (for parallel OAR)[2]. Since this chapter is motivated by biological considerations, these constraints are expressed in terms of BED equivalents.

Let $\mathcal{M}_1 \subseteq \mathcal{M}$ be the set of indices of serial OAR for which maximum dose constraints are needed. The functioning of these OAR is hindered even when a small region is damaged by radiation (e.g. spinal cord in head-and-neck cancer). Assume that a total dose $D_{\max}^m$ is known to be tolerated by each voxel in O_m for $m \in \mathcal{M}_1$, if administered in N_{conv}^m equal-dose sessions. The BED corresponding to this schedule equals $\text{BED}_{\max}^m = D_{\max}^m (1 + \rho_m (D_{\max}^m / N_{\text{conv}}^m))$. Let $z_j^{t,m}$ denote the BED delivered to voxel $j \in O_m$ in the first $t-1$ sessions. Then, a dose of $(A_j^m u^t)$ can be tolerated by the jth voxel in O_m in session t if

$$z_j^{t,m} + A_j^m u^t + \rho_m (A_j^m u^t)^2 \leq \text{BED}_{\max}^m, \;\; j = 1, 2, \ldots, m. \tag{5.2}$$

This implies that in order to test the feasibility of u^t in session t, the treatment planner must

[2]Dose-volume constraints are omitted here for simplicity; they can be incorporated by following the methodology in [56].

know $\vec{z}^{t,m} = (z_1^{t,m}, \ldots, z_{n_m}^{t,m})$. Thus, for each $m \in \mathcal{M}_1$, $\vec{z}^{t,m}$ is defined as the state of O_m at the beginning of session t.

Let $\mathcal{M}_2 \subseteq \mathcal{M}$ be the set of indices of parallel OAR for which mean dose constraints are needed. For these OAR, a sufficiently small portion can be damaged without affecting the organ's function (e.g. parotid glands in head-and-neck cancer). Suppose that mean dose D_m can be tolerated by O_m, for $m \in M$, if administered in N_{conv}^m equal-dose sessions. The BED corresponding to this mean dose is given by $\text{BED}_{\text{mean}}^m = D_{\text{mean}}^m(1 + \rho_m(D_{\text{mean}}^m/N_{\text{conv}}^m))$. Let $v^{t,m}$ denote the mean BED delivered in the first $t-1$ sessions over all voxels in O_m. Then, doses $(A_j^m u^t)$ delivered to voxels $j \in N_m$ in session t can be tolerated by O_m if

$$v^{t,m} + \frac{\sum_{j=1}^{n_m}(A_j^m u^t) + \rho_m \sum_{j=1}^{n_m}(A_j^m u^t)^2}{n_m} \leq \text{BED}_{\text{mean}}^m. \tag{5.3}$$

Therefore, for each $m \in \mathcal{M}_2$, $v^{t,m}$ is defined as the state of O_m at the beginning of session t.

A smoothness constraint is put on each radiation field. This bounds the absolute relative difference between intensities of each pair of nearest neighbor beamlets by a fraction ϵ. This can be written as $Su^t \leq 0$, where S is a block diagonal matrix with entries $(1+\epsilon), (1-\epsilon), 0, +1$ at appropriate locations (see [57]). This is done to ensure that the intensity profile is deliverable in practice using a multi-leaf collimator.

Let $|\mathcal{M}_1|$ and $|\mathcal{M}_2|$ denote the cardinalities of sets $\mathcal{M}_1$ and $\mathcal{M}_2$, respectively. Let $\vec{\vec{z}}^t = (\vec{z}^{t,1}, ..., \vec{z}^{t,|\mathcal{M}_1|})$ and $\vec{v}^t = (v^{t,1}, ..., v^{t,|\mathcal{M}_2|})$. Thus, the combined state of all OAR is $[\vec{\vec{z}}^t; \vec{v}^t]$. The set of fluence-maps $u^t \in \Re_+^K$ that satisfy the maximum dose, mean dose, and smoothness constraints is denoted by $\mathcal{U}^t([\vec{\vec{z}}^t; \vec{v}^t])$. In stochastic control parlance, a feasible policy assigns a fluence-map from the set $\mathcal{U}^t([\vec{\vec{z}}^t; \vec{v}^t])$ to every possible tumor and OAR states $[\vec{x}^t; \vec{\vec{z}}^t; \vec{v}^t]$ in each treatment session $t = 1, 2, \ldots, T$. The set of feasible policies is denoted by $\mathcal{P}$, and generic policies in this set are written as π.

Note that in the initial state $[\vec{x}^1; \vec{\vec{z}}^1; \vec{v}^1]$, all components of $\vec{\vec{z}}^1$ and $\vec{v}^1$ are zero since no dose is delivered prior to the first session. Let $J_\pi^1([\vec{x}^1; \vec{\vec{z}}^1; \vec{v}^1])$ denote the expected TNTCR at the end of the treatment course if policy $\pi \in \mathcal{P}$ is implemented in sessions $t = 1, 2, \ldots, T$,

starting in state $[\vec{x}^1; \vec{\vec{z}}^1; v^1]$ at the beginning of the first session. The problem of minimizing the TNTCR can then be formulated as the stochastic control problem

$$J_*^1([\vec{x}^1; \vec{\vec{z}}^1; \vec{v}^1]) = \min_{\pi \in \mathcal{P}} J_\pi^1([\vec{x}^1; \vec{\vec{z}}^1; \vec{v}^1]). \tag{5.4}$$

The standard approach for solving such problems, at least in principle, is to employ Bellman's backward recursive algorithm of dynamic programming [9]. That is explained next.

Let $J^t([\vec{x}^t; \vec{\vec{z}}^t; \vec{v}^t])$, for all possible states $[\vec{x}^t; \vec{\vec{z}}^t; \vec{v}^t]$ and for $t = 1, 2, \ldots, T+1$, denote the optimal cost-to-go functions. These cost-to-go functions can be obtained by solving, in the backward order $t = T, T-1, \ldots, 1$, the non-linear stochastic optimization problems

$$(P_t) \quad J^t([\vec{x}^t; \vec{\vec{z}}^t; \vec{v}^t]) = \min_{u^t \in \mathcal{U}^t([\vec{\vec{z}}^t; \vec{v}^t])} E_{\vec{p}} \left(J^{t+1}([\vec{x}^{t+1}; \vec{\vec{z}}^{t+1}; \vec{v}^{t+1}]) \right), \tag{5.5}$$

subject to

$$x_i^{t+1} = x_i^t \exp(-\alpha_i^t[(A_i u^t) + r(A_i u^t)^2]),\ i = 1, 2, \ldots, n, \tag{5.6}$$

$$z_j^{t+1,m} = z_i^{t,m} + (A_j^m u^t) + \rho_m (A_j^m u^t)^2,\ j = 1, 2, \ldots, n_m,\ m \in \mathcal{M}_1, \tag{5.7}$$

$$v^{t+1,m} = v^{t,m} + \frac{\sum_{j=1}^{n_m} (A_j^m u^t) + \rho_m \sum_{j=1}^{n_m} (A_j^m u^t)^2}{n_m},\ m \in \mathcal{M}_2, \tag{5.8}$$

$$Su^t \leq 0, \tag{5.9}$$

$$u^t \geq 0, \tag{5.10}$$

for all possible states $[\vec{x}^t; \vec{\vec{z}}^t; \vec{v}^t]$, starting with the boundary condition

$$J^{T+1}([\vec{x}^{T+1}; \vec{\vec{z}}^{T+1}; \vec{v}^{T+1}]) = \sum_{i=1}^{n} v x_i^{T+1}. \tag{5.11}$$

The optimal objective value $J_*^1([\vec{x}^1; \vec{\vec{z}}^1; \vec{v}^1])$ of the clairvoyant stochastic control problem (5.4) equals the optimal cost-to-go $J^1([\vec{x}^1; \vec{\vec{z}}^1; \vec{v}^1])$ as delivered by the above backward recursion algorithm.

Unfortunately, this backward recursion algorithm is not implementable because it re-

quires that the minimization in (5.5) be performed for an uncountable number of states. Discretization of states and/or actions is not a practical option either, owing to their high dimensions[3]. An approximate solution method that does not require discretization and exploits the structure of the clairvoyant stochastic control problem is described in Section 5.3.1.

5.2.2 Stochastic control with Bayesian learning of tumor-response uncertainty

The clairvoyant formulation assumes that the treatment planner knows the categorical distribution $\vec{p} = (p_1, p_2, \ldots, p_J)$ of the parameter values $\vec{b} = (b_1, b_2, \ldots, b_J)$. This assumption is relaxed here.

The treatment planner instead begins with a Dirichlet prior, with a hyperparameter vector $\vec{a}^1 = (a_1^1, a_2^1, \ldots, a_J^1)$, on the unknown categorical distribution $\vec{p}$. The Dirichlet distribution is conjugate to the categorical distribution. This means that the posterior distribution, after observing each tumor-voxel's response, is also Dirichlet. Furthermore, the posterior hyperparameters are obtained via a simple update formula. In particular, if the hyperparameters at the beginning of session t are $\vec{a}^t = (a_1^t, a_2^t, \ldots, a_J^t)$, and if the realized values of the random variables α_i^t in that session are such that c_ℓ^t of the values belong to category ℓ, for $\ell = 1, 2, \ldots, J$, then the posterior hyperparameters are given by

$$\vec{a}^{t+1} = (a_1^t + c_1^t, a_2^t + c_2^t, \ldots, a_J^t + c_J^t). \tag{5.12}$$

Note that the realized values of the random variables α_i^t are not observed, but rather are *calculated* by inverting Equation (5.1), given x_i^{t+1}, x_i^t, and u^t. That is,

$$\alpha_i^t = \frac{-\ln(x_i^{t+1}/x_i^t)}{A_i u^t + r(A_i u^t)^2}. \tag{5.13}$$

In this calculation process, that is called "information update", cell densities x_i^{t+1} and x_i^t are

[3]The number of voxels could be in the tens of thousands and the number of beamlets could be several thousand.

observed in functional images before and after session t, and u^t is chosen by the treatment planner in session t. Thus, by appending the hyperparameter vector $\vec{a}^t$ (called the information state) to the (biological) state $[\vec{x}^t; \vec{z}^t; \vec{v}^t]$ of the clairvoyant problem, the Bayesian stochastic control problem is formulated as

$$J_*^1([\vec{a}^1; \vec{x}^1; \vec{\vec{z}}^1; \vec{v}^1]) = \min_{\pi \in \mathcal{P}} J_\pi^1([\vec{a}^1; \vec{x}^1; \vec{\vec{z}}^1; \vec{v}^1]). \tag{5.14}$$

This generalization of the clairvoyant stochastic control problem (5.4), can in theory be solved via Bellman's backward recursive algorithm of dynamic programming as explained next.

In particular, let $J^t([\vec{a}^t; \vec{x}^t; \vec{z}^t; \vec{v}^t])$, for all possible states $[\vec{a}^t; \vec{x}^t; \vec{z}^t; \vec{v}^t]$ and for $t = 1, 2, \ldots, T+1$, denote the optimal cost-to-go functions. These cost-to-go functions can be obtained by solving, in the backward order $t = T, T-1, \ldots, 1$, the non-linear stochastic optimization problems

$$J^t([\vec{a}^t; \vec{x}^t; \vec{\vec{z}}^t; \vec{v}^t]) = \min_{u^t \in \mathcal{U}^t([\vec{\vec{z}}^t; \vec{v}^t])} E_{\vec{a}^t}\left(J^{t+1}([\vec{a}^{t+1}; \vec{x}^{t+1}; \vec{\vec{z}}^{t+1}; \vec{v}^{t+1}])\right), \tag{5.15}$$

subject to (5.16)

$$x_i^{t+1} = x_i^t \exp(-\alpha_i^t[(A_i u^t) + r(A_i u^t)^2]),\ i = 1, 2, \ldots, n, \tag{5.17}$$

$$\vec{a}^{t+1} = \vec{a}^t + \vec{c}^t, \tag{5.18}$$

$$z_j^{t+1,m} = z_i^{t,m} + (A_j^m u^t) + \rho_m (A_j^m u^t)^2,\ j = 1, 2, \ldots, n_m,\ m \in \mathcal{M}_1, \tag{5.19}$$

$$v^{t+1,m} = v^{t,m} + \frac{\sum_{j=1}^{n_m}(A_j^m u^t) + \rho_m \sum_{j=1}^{n_m}(A_j^m u^t)^2}{n_m},\ m \in \mathcal{M}_2, \tag{5.20}$$

$$Su^t \leq 0, \tag{5.21}$$

$$u^t \geq 0, \tag{5.22}$$

for all possible states $[\vec{a}^t; \vec{x}^t; \vec{\vec{z}}^t; \vec{v}^t]$, starting with the boundary condition

$$J^{T+1}([\vec{a}^{T+1}; \vec{x}^{T+1}; \vec{\vec{z}}^{T+1}; \vec{v}^{T+1}]) = \sum_{i=1}^{n} v x_i^{T+1}. \tag{5.23}$$

Again, the optimal objective value $J_*^1([\vec{a}^1; \vec{x}^1; \vec{\vec{z}}^1; \vec{v}^1])$ of the Bayesian stochastic control problem (5.14) equals the optimal cost-to-go $J^1([\vec{a}^1; \vec{x}^1; \vec{\vec{z}}^1; \vec{v}^1])$ as obtained by this backward recursion. Unfortunately, this backward recursion is even harder than the clairvoyant case (P_t) owing to the multi-dimensional information state $\vec{a}^t$. In particular, similar to the clairvoyant case, the Bayesian stochastic control problem also cannot be solved exactly in practice. An approximate solution method is described in Section 5.3.2.

5.3 Solution methods

Approximate solution methods for the clairvoyant stochastic control problem (5.4), and the stochastic control problem with Bayesian learning (5.14) are described in this section. Both methods are rooted in certainty equivalent control for problems with T stages [9]. The idea in CEC is to replace all future uncertain quantities in a stochastic control problem with some nominal values, after observing the state at the beginning of stage t. This results in a deterministic optimization problem with $T - t + 1$ stages. This deterministic problem is solved (approximately or exactly) to obtain control variable values for stages $t, t+1, \ldots, T$. Only the control variable for stage t is implemented and the rest are discarded. The system then evolves stochastically to a new state, and the process is repeated until the end of stage T. Thus, in one complete run of CEC, a total of T deterministic optimization problems are solved; the first one includes T sessions, the second includes $T-1$ sessions, and ultimately, the last one involves only one session. CEC is particularly suitable when the original stochastic control problem has a continuous state-space and a continuous action-space, especially if the resulting deterministic optimization problems are convex. CEC has been shown to be effective in numerical experiments for various clairvoyant stochastic control problems in spatiobiologically integrated radiotherapy [30, 38, 56]. It is generalized in Section 5.3.2 to

the stochastic control problem with Bayesian learning.

5.3.1 Clairvoyant CEC [idealized gold-standard]

After observing state $[\vec{x}^t; \vec{\hat{z}}^t; \vec{v}^t]$ at the beginning of session t, the uncertainty in the clairvoyant stochastic control problem (5.4) lies in the future values $\alpha_i^t, \alpha_i^{t+1}, \ldots, \alpha_i^T$, for voxels $i = 1, 2, \ldots, n$. There are several potential ways to implement CEC, depending on what nominal values replace these uncertain quantities. The typical approach involves replacing these random variables with their expected values. This approach was implemented in [38, 56]. In preliminary numerical experiments for this chapter, a slightly different approach was found to perform better. Specifically, in the CEC algorithm in this section, the treatment planner samples n iid numbers $\hat{\alpha}_i$, for $i = 1, 2, \ldots, n$, (from the categorical$(\vec{p}, \vec{b})$ distribution), and replaces the future uncertain values $\alpha_i^t, \alpha_i^{t+1}, \ldots, \alpha_i^T$ with $\hat{\alpha}_i$. This yields the non-linear, nonconvex deterministic optimization problem:

$$(Q_t^0) \min_{u^t,\ldots,u^T} \sum_{i=1}^{n} \nu x_i^t \prod_{l=t}^{T} \exp\left(-\hat{\alpha}_i[(A_i u^l) + r(A_i u^l)^2]\right) \tag{5.24}$$

subject to

$$z_j^{\ell+1,m} = z_i^{\ell,m} + (A_j^m u^\ell) + \rho_m (A_j^m u^\ell)^2,\ j = 1, 2, \ldots, n_m,\ m \in \mathcal{M}_1,\ \ell = t, t+1, T, \tag{5.25}$$

$$z_j^{\ell+1,m} \leq \text{BED}_{\max}^m,\ j = 1, 2, \ldots, n_m,\ m \in \mathcal{M}_1,\ \ell = t, t+1, T, \tag{5.26}$$

$$v^{\ell+1,m} = v^{\ell,m} + \frac{\sum_{j=1}^{n_m} (A_j^m u^\ell) + \rho_m \sum_{j=1}^{n_m} (A_j^m u^\ell)^2}{n_m},\ m \in \mathcal{M}_2,\ \ell = t, t+1, T, \tag{5.27}$$

$$v^{\ell+1,m} \leq \text{BED}_{\text{mean}}^m,\ m \in \mathcal{M}_2,\ \ell = t, t+1, T, \tag{5.28}$$

$$Su^\ell \leq 0,\ \ell = t, t+1, \ldots, T, \tag{5.29}$$

$$u^\ell \geq 0,\ \ell = t, t+1, \ldots, T. \tag{5.30}$$

This problem is very large in practice (tens of thousands of linear and quadratic constraints, and upwards of a hundred thousand variables [38, 56, 57]). As such, it is computationally intractable. A well-established, efficient, and effective approach from [38, 56, 57] for approximate solution is described next.

The idea in this approximation approach is to search within the class of time-invariant fluence-maps. That is, to set $u^t = u^{t+1} = \ldots = u^T = u$, and thus, search for a solution of the form $(\underbrace{u, u, \ldots, u}_{(T-t+1)\text{times}})$, where $u \in \Re_+^K$. After some algebraic simplifications, this reduces the above problem to

$$(Q_t^1) \quad \min_u \sum_{i=1}^{n} v x_i^t \exp\left(-(T-t+1)(\hat{\alpha}_i[(A_i u) + r(A_i u)^2])\right), \tag{5.31}$$

subject to

$$A_j^m u \leq \frac{-1 + \sqrt{1 + 4\rho_m(\text{BED}_{\max}^m - z_j^{t,m})/(T-t+1)}}{2\rho_m}, \ j = 1, 2, \ldots, n_m, \ m \in \mathcal{M}_1, \tag{5.32}$$

$$\sum_{j=1}^{n_m}(A_j^m u) + \rho_m \sum_{j=1}^{n_m}(A_j^m u)^2 \leq n_m \left(\frac{\text{BED}_{\text{mean}}^m - v^{t,m}}{T-t+1}\right), \ m \in \mathcal{M}_2, \tag{5.33}$$

$$Su \leq 0, \tag{5.34}$$

$$u \geq 0. \tag{5.35}$$

This problem is convex if $\hat{\alpha}_i r \geq 2$ for all i (see [38, 56] for proof). This condition is believed to be met in practice for most tumors including head-and-neck and lung. In that case, the problem can be solved efficiently via a log-barrier interior point algorithm with Newton's method and backtracking line search as described in detail in [56]. Its description is omitted here for brevity, but a high-level pseudocode for the resulting clairvoyant CEC procedure is listed in Algorithm 2 below.

In the numerical experiments in Section 5.4, this approach is termed *clairvoyant* for brevity.

Algorithm 2 Clairvoyant CEC

1: **Initilization:** Set $t = 1$; begin with a given initial state $[\vec{x}^1; \vec{\vec{z}}^1; \vec{v}^1]$, and a fixed value r.
2: **while** $t \leq T$ **do**
3: Sample iid $\hat{\alpha}_i$, for $i = 1, \ldots, n$, from the categorical$(\vec{p}, \vec{b})$ distribution;
4: Solve (Q_t^1) using the log-barrier interior point algorithm from [56] to obtain fluence-map u_*;
5: Administer $u^t = u^*$; observe tumor state $\vec{x}^{t+1}$; calculate OAR state $[\vec{\vec{z}}^{t+1}; \vec{v}^{t+1}]$ from (5.7)-(5.8);
6: Set $t \leftarrow t + 1$.
7: **end while**

This procedure is generalized to the stochastic control problem with Bayesian learning (5.14) next.

5.3.2 CEC with Bayesian learning

There are two sources of uncertainty in the Bayesian learning problem described in Section 5.2.2. The higher-level uncertainty arises because the treatment planner does not know the categorical probabilities $\vec{p}$. The lower-level uncertainty, which also exists in the clairvoyant problem, is in the realized values of the tumor's dose-response parameters. The CEC approach in this section replaces both these uncertain quantities with their nominal values. In particular, at the beginning of session t, the unknown probabilities $p_1, \ldots, p_J$ are replaced by the componentwise expectations of the corresponding Dirichlet posterior: $\hat{p}_\ell^t = \frac{a_\ell^t}{\sum_{\ell=1}^{J} a_\ell^t}$. The treatment planner then samples n iid numbers $\bar{\alpha}_i$, for $i = 1, 2, \ldots, n$, (from the categorical$(\vec{\hat{p}}^t, \vec{b})$ distribution), and replaces the future uncertain values $\alpha_i^t, \alpha_i^{t+1}, \ldots, \alpha_i^T$ with $\bar{\alpha}_i$. This yields the problem

$$(B_t^0) \ \min_u \ \sum_{i=1}^{n} \nu x_i^t \prod_{l=t}^{T} \exp\left(-\bar{\alpha}_i[(A_i u^l) + r(A_i u^l)^2]\right), \tag{5.36}$$

$$\text{subject to } (5.25) - (5.30),$$

which is identical in form to (Q_t^0). Consequently, (B_t^0) is also computationally intractable for reasons outlined in the context of (Q_t^0). Similar to (Q_t^0), this problem is solved approximately by searching for time-invariant solutions. This yields the problem

$$(B_t^1) \quad \min_u \sum_{i=1}^{n} v x_i^t \exp\left(-(T-t+1)(\bar{\alpha}_i[(A_i u) + r(A_i u)^2])\right), \tag{5.37}$$
$$\text{subject to } (5.32) - (5.35),$$

which is identical in form to (Q_t^1). Thus, similar to (Q_t^1), the log-barrier interior point algorithm with Newton's method and backtracking line search can be applied. A pseudocode for the overall CEC procedure for the Bayesian learning problem is described in Algorithm 3 below.

Algorithm 3 CEC with Bayesian learning

Initilization: Set $t = 1$; begin with a given initial state $[\vec{x}^1; \vec{\vec{z}}^1; \vec{v}^1]$, and a fixed value r; choose a hyperparameter vector $\vec{a}_1$ for the initial Dirichlet prior.
while $t \leq T$ **do**
  Let $\hat{p}_\ell = \frac{a_\ell^t}{\sum_{\ell=1}^{J} a_\ell^t}$, for $\ell = 1, \ldots, J$;
  Sample $\bar{\alpha}_i$, for $i = 1, \ldots, n$, from the categorical$(\vec{\hat{p}}^t, \vec{b})$ distribution;
  Solve (B_t^1) using the log-barrier interior point algorithm from [56] to obtain fluence-map u_*;
  Administer $u^t = u^*$; observe tumor state $\vec{x}^{t+1}$; calculate OAR state $[\vec{\vec{z}}^{t+1}; \vec{v}^{t+1}]$ from (5.7)-(5.8);
  Calculate the "realized" α_i^t from formula (5.13);
  Let c_ℓ^t be the number of realized α_i^t that equal b_ℓ, for $\ell = 1, 2, \ldots, J$;
  Update the hyperparameter vector $\vec{a}_{t+1}$ using formula (5.12);
  Set $t \leftarrow t + 1$.
end while

In the numerical experiments in Section 5.4, this is called the *learning* approach for brevity.

5.3.3 CEC with a fixed value of tumor α [existing benchmark]

The CEC method with Bayesian learning above will be compared against a simpler version of CEC that can be viewed as an existing benchmark. Here, the treatment planner simply uses a fixed value of tumor α (say something currently available in the literature) as the nominal value in the objective function of (Q_t^1). After the simplification outlined in the context of (Q_t^1), this yields

$$(F_t^1) \quad \min_u \sum_{i=1}^{n} v x_i^t \exp\left(-(T-t+1)(\alpha[(A_i u) + r(A_i u)^2])\right), \tag{5.38}$$
$$\text{subject to } (5.32) - (5.35),$$

and the procedure listed in Algorithm 4.

Algorithm 4 CEC with a fixed value of tumor α [existing benchmark]

1: **Initilization:** Set $t = 1$; begin with a given initial state $[x^1; z^1; v^1]$, and fixed values α, r.
2: **while** $t \leq T$ **do**
3: Set $\alpha_i^\ell = \alpha$, for $i = 1, \ldots, n$ and $\ell = t, t+1, \ldots, T$;
4: Solve (F_t^1) using the log-barrier interior point algorithm from [56] to obtain fluence-map u_*;
5: Administer $u^t = u^*$; observe tumor state x^{t+1}; calculate OAR state $[z^{t+1}; v^{t+1}]$ from (5.7)-(5.8);
6: Let $t \leftarrow t + 1$.
7: **end while**

In the numerical experiments in Section 5.4, this is called the *no learning* approach for brevity.

5.4 Numerical results

The performance of the clairvoyant, learning, and no learning methods is compared here via numerical simulations.

A phantom head-and-neck cancer case with four OAR: brain stem, spinal cord, left and

right parotid glands, was used. Data for this case was obtained from [56]. This case includes seven radiation beams with the total number of beamlets $k = 1680$. The total number of tumor voxels was 9096. The number of conventional sessions N_{conv} was 35. For the spinal cord and brain stem, the maximum tolerable doses $D_{\max}$ were set to 45 Gy and 50 Gy, respectively. The tolerable mean doses D_{mean} for the left and right parotids were set to 28 Gy. The α/β ratios for all OARs were fixed at 3 Gy and the α/β for tumor was fixed at $r = 10$. The tumor's α_i^t values followed a categorical distribution with $J = 51$ possible values: $\{0.1, 0.11, \ldots, 0.6\}$ Gy^{-1}; that is, $\vec{b} = \{0.1, 0.11, \ldots, 0.6\}$. These values of the conventional treatment course length, tolerance doses, and radiobiological dose-response parameters were based on the existing clinical literature [22, 24, 26, 27, 33, 47, 73]. The number of sessions T was set to equal 35, the same as N_{conv}. For Algorithm 4, α was fixed at 0.35 since this is the commonly reported value for head-and-neck cancer in the aforementioned clinical literature.

The algorithms were compared on three different shapes of the categorical distribution. All three shapes were derived from truncated, discretized normal distributions over the range $[0.1, 0.6]$. The first was a symmetric distribution derived from a normal distribution with mean and standard deviation of 0.35 and 0.25/3, respectively. This is denoted $\mathbf{N}(0.35, 0.25/3)$. The second was skewed to the right and was derived from $\mathbf{N}(0.25, 0.05)$. The third was skewed to the left and was derived from $\mathbf{N}(0.45, 0.05)$. Figure 5.1a shows the three distributions.

The initial Dirichlet prior hyperparameters were set to $\vec{a} = (\underbrace{1, 1, \ldots, 1}_{51 \text{ times}})$ for all three categorical distribution shapes. This corresponds to an uninformed prior, which assigns equal probabilities of $1/51 = 0.02$ to each possible value in $\{0.1, 0.11, \ldots, 0.6\}$. These initial uniform probabilities $\hat{p}_\ell^1 = 0.02$ for $b_\ell \in \{0.1, 0.11, \ldots, 0.6\}$ are shown by a flat line of blue squares in Figures 5.1b-5.1d for the three categorical distribution shapes. The evolution of probabilities $\hat{p}_\ell^t$, for $t = 2, 3, \ldots, T$, in one simulation of the learning algorithm is also illustrated in those figures. Squares of distinct colors depict distinct sessions. The figures show that the estimated probabilities $\hat{p}_\ell^t$ quickly emulate the true probabilities p_ℓ within

about 5 sessions in all three cases.

The performance of the three algorithms was compared via two sets of null hypotheses. The first set of null hypotheses $H_0^{\square}$ was that $\text{TNTCR}^{\square}_{\text{clairvoyant}} = \text{TNTCR}^{\square}_{\text{learning}}$, where $\text{TNTCR}^{\square}_{\blacksquare}$ denotes the expected TNTCR achieved by method $\blacksquare \in \{\text{clairvoyant}, \text{learning}\}$ when $\square$ belongs to the set $\{\text{symmetric}, \text{right-skewed}, \text{left-skewed}\}$. A t-test at a .05 significance level using 30 independent simulations of the two algorithms did not reject this null hypothesis for any of the three categorical distributions. This suggests that the learning algorithm performed as well as the clairvoyant, which is an idealized gold-standard. This is not unexpected, given that the categorical distribution estimated by the learning algorithm quickly approaches the true distribution as shown in Figures 5.1b-5.1d. The second set of hypotheses $G_0^{\square}$ was $\text{TNTCR}^{\square}_{\text{no learning}} = \text{TNTCR}^{\square}_{\text{learning}}$, where $\square$ belongs to the set $\{\text{symmetric}, \text{right-skewed}, \text{left-skewed}\}$. Again, a t-test at a .05 significance level using 30 independent simulations of the two algorithms rejected this null hypothesis for each of the three categorical distributions. This suggests that the learning algorithm performed differently than the no learning method. The percentage relative difference in average TNTCR attained by these two procedures over 30 independent simulations was calculated as

$$\text{RelDiff}^{\square}_{\text{TNTCR}} = \frac{\sum_{\text{sim}=1}^{30} \text{TNTCR}^{\square}_{\text{no learning}}(\text{sim}) - \sum_{\text{sim}=1}^{30} \text{TNTCR}^{\square}_{\text{learning}}(\text{sim})}{\sum_{\text{sim}=1}^{30} \text{TNTCR}^{\square}_{\text{learning}}} \times 100, \qquad (5.39)$$

where TNTCR(sim) is the TNTCR for simulation sim $= 1, 2, \ldots, 30$, and $\square$ belongs to the set $\{\text{symmetric}, \text{right-skewed}, \text{left-skewed}\}$. Table 5.1 displays this percentage relative difference for all three categorical distributions.

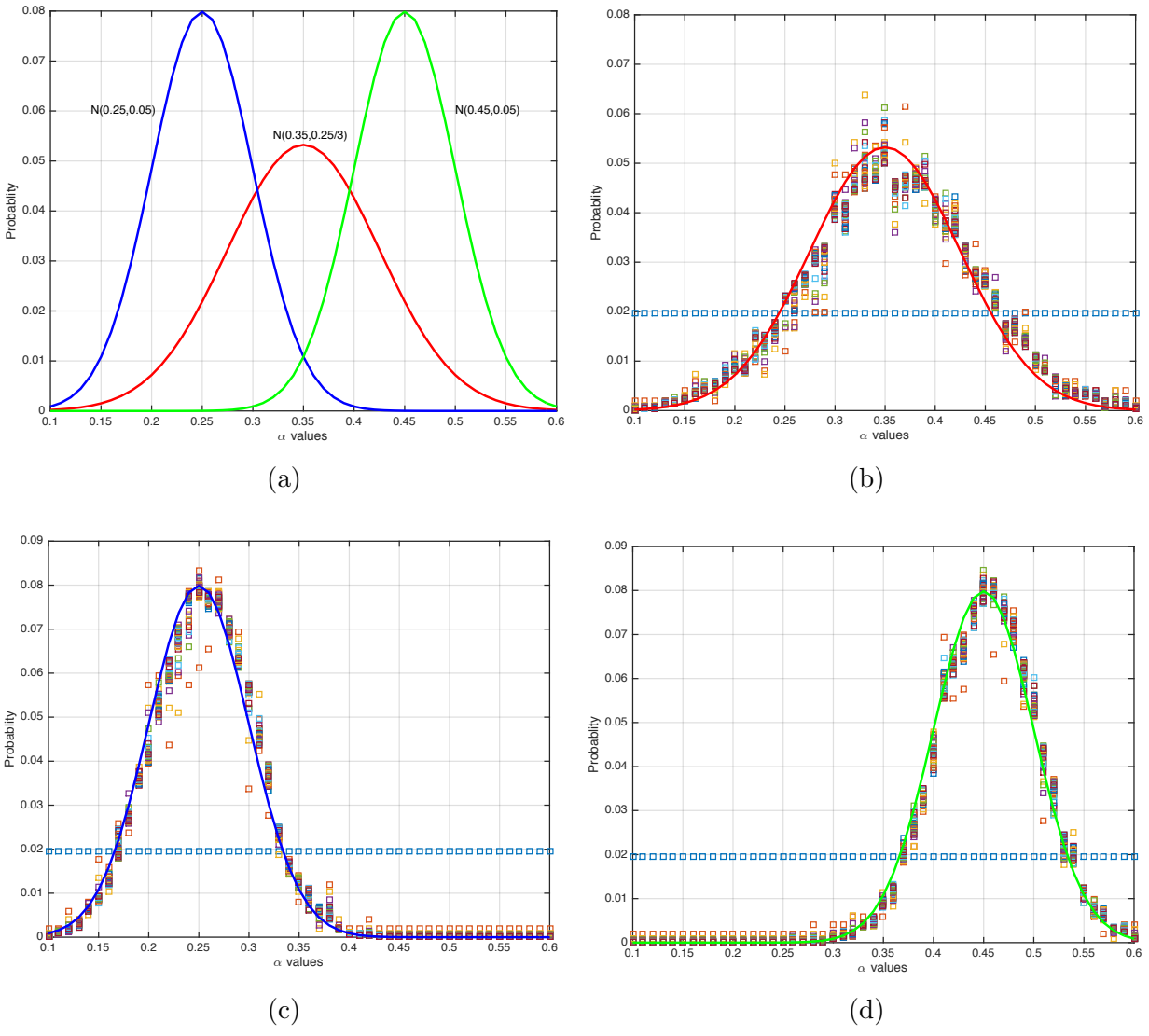


Figure 5.1: (a) Three probability distributions used in the algorithm; Learning the probability vector over 35 treatment sessions for the (b) Symmetric distribution $\mathbf{N}(0.35, 0.25/3)$, (c) right-skewed $\mathbf{N}(0.25, 0.05)$ and (d) Left-skewed $\mathbf{N}(0.45, 0.05)$ distributions.

Distribution	$\mathbf{RelDiff_{TNTCR}}$
Symmetric	23.32%
Right-skewed	20.45%
Left-skewed	18.19%

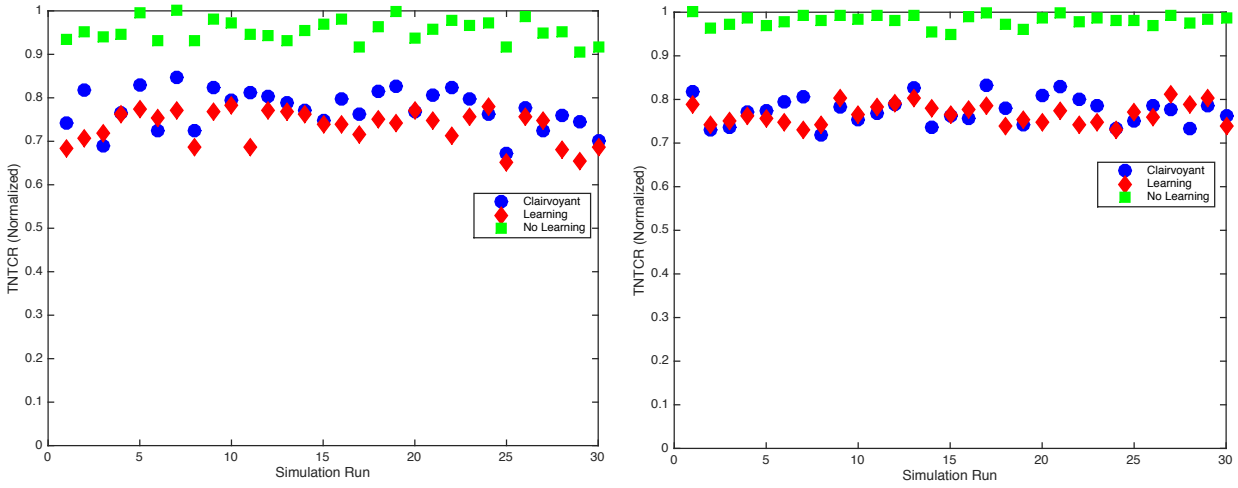
Table 5.1: Percentage relative difference between average TNTCR for the no learning and learning methods per formula (5.39).

Figure 5.2 shows the TNTCR attained in each of the 30 simulations for all three algorithms and all three categorical distributions. In all cases, the TNTCR values achieved by the learning method appear comparable with those of the clairvoyant method (consistent with the t-test for the first set of hypotheses). The TNTCR reached by the no learning method appears higher than that with the learning method in every single simulation.

Two other performance metrics related to the TNTCR are perhaps also of interest to researchers in this area. The first one is the expected total (over T treatment sessions) average (over all tumor voxels) dose delivered to the tumor. The second one is the expected (with respect to the true categorical distribution of the iid dose-response parameters α_i^t) total average biological effect (BE) of the dose delivered to the tumor. Here, this total average BE is defined as

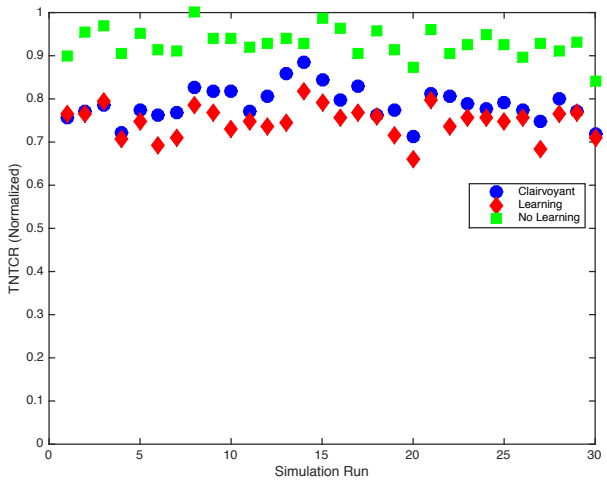
$$\text{BE} = \sum_{t=1}^{T} \frac{\sum_{i=1}^{n} \alpha_i^t [A_i u^t + r(A_i u^t)^2]}{n}. \tag{5.40}$$

Although these two metrics were not explicitly optimized by the three solution methods, they are compared here to gain further insight. In particular, the total average dose and BE were averaged over 30 independent simulations. The percentage relative difference between these averages for the no learning and learning methods were calculated using formulas similar to (5.39). The results are summarized in Table 5.2. The table shows that the learning method administers higher dose and BE than the no learning method in all three cases. A t-test as before confirmed that this difference between the two methods was statistically significant



(a)

(b)



(c)

Figure 5.2: TNTCR in 30 independent simulations for (a) Symmetric distribution derived from $\mathbf{N}(0.35, 0.25/3)$, (b) right-skewed distribution derived from $\mathbf{N}(0.25, 0.05)$, and (d) left-skewed distribution derived from $\mathbf{N}(0.45, 0.05)$.

in all three cases. In fact, Figures 5.3 and 5.4 show that the learning method administers a higher tumor-dose and tumor-BE in every simulation in all three cases. As is perhaps apparent from these figures, the difference between the learning and the clairvoyant methods was not statistically significant according to a t-test, and hence it is not reported in Table 5.2. These observations are consistent with Table 5.1 and Figure 5.2.

Distribution	**RelDiff$_{\mathbf{Dose}}$**	**RelDiff$_{\mathbf{BE}}$**
Symmetric	8.91%	10.21%
Right-skewed	11.83%	12.57%
Left-skewed	6.43%	5.43%

Table 5.2: Percentage relative difference between average (over 30 simulations) total (over 35 sessions) average (over all tumor voxels) dose and BE delivered by the no learning and learning methods.

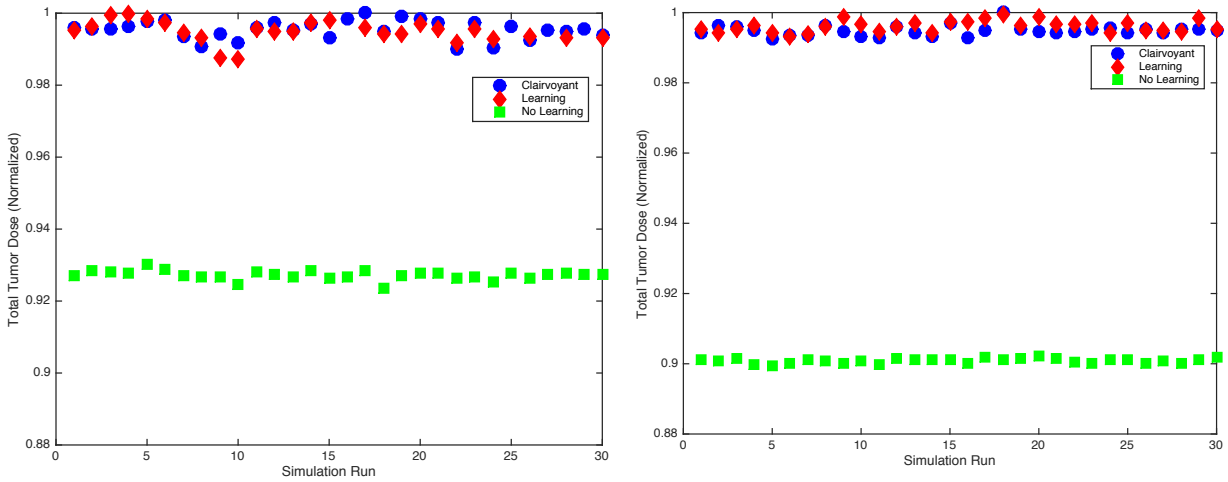
To gain further insight into the three methods, Figure 5.5 displays the average (over voxels) tumor-BE in each treatment session, averaged over 30 independent simulations. This figure shows that the learning method closely tracks the tumor-BE administered by the clairvoyant method, whereas the no learning method does not. The difference between the learning/clairvoyant pair and the no learning method is more stark in the symmetric and the right-skewed cases than in the left-skewed case. In fact, this observation applies to all tables and figures in this chapter. The precise reason for this is not entirely clear, but one reason is because the tumor-dose response parameter α is generally higher in the left-skewed case (recall Figure 5.1). That is, the tumor is more radiosensitive, and hence a naive planning approach such as the one without learning could work somewhat satisfactorily. In other words, learning might help more in the case of patient- and tumor-specific spatiotemporal heterogeneity in radioresistance, although more research would be needed to verify this conjecture.

5.5 Conclusions

This chapter provides a theoretical framework for learning the probability distribution of a tumor's dose-response uncertainty in spatiobiologically integrated individualized radiotherapy. This framework is rooted in the field of Bayesian stochastic control for simultaneous learning of problem parameters and optimization of decision variables. Since exact solution of the resulting formulation is computationally intractable, an approximation scheme based on CEC was proposed. This was termed CEC with Bayesian learning, and it was compared with two other "extreme" algorithms. On the one extreme was a clairvoyant CEC algorithm, where the treatment planner was assumed to know the probability distribution of the tumor's dose-response uncertainty. This served as an idealized gold-standard for the learning algorithm. On the other extreme was a CEC algorithm, where the treatment planner simply uses a typical value of the dose-response parameter from the literature. This is probably what practitioners would do today, if they were to administer spatiobiologically integrated radiotherapy. This therefore served as a benchmark. Numerical simulations with a phantom head-and-neck case showed that in terms of the TNTCR objective, CEC with Bayesian learning was comparable to clairvoyant CEC, and better than the benchmark CEC.

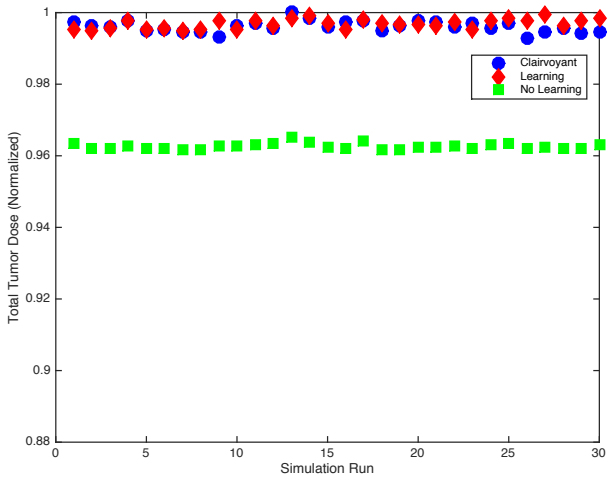
The formulation in this chapter has some limitations on the theoretical side. The first is the assumption that α_i^t are iid across i and t. This assumption can be relaxed. For instance, the tumor could be divided into sub-regions, with different probability distributions for dose-response parameters in different sub-regions. The Bayesian framework can be extended to this case. The assumption of independence across voxels can also be relaxed by introducing spatial correlation among dose-response parameters. This was done in the clairvoyant formulation for adapting to hypoxia in [56]. That clairvoyant model in [56] also incorporated temporal correlations in dose-response parameters using a first order vector autoregressive process. A similar approach could be combined with Bayesian learning here. These extensions would call for the estimation of more parameters with fewer data points. This would slow down learning, which might cause the TNTCR for the learning method to drift higher

than the clairvoyant one.



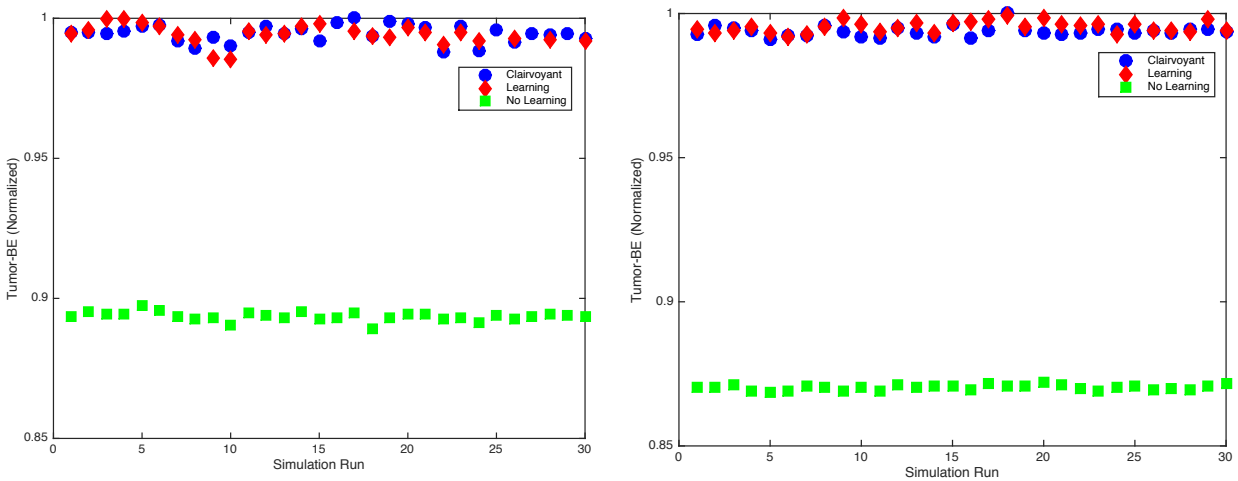
(a)

(b)



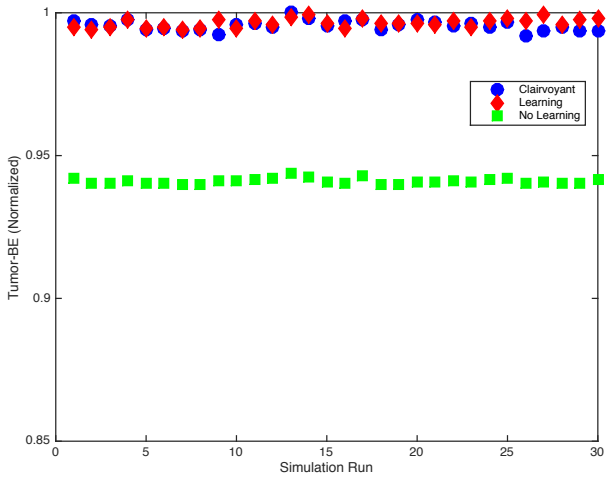
(c)

Figure 5.3: Total (in 35 sessions) average (over all voxels) tumor-BE in 30 independent simulations for (a) Symmetric distribution derived from $\mathbf{N}(0.35, 0.25/3)$, (b) right-skewed distribution derived from $\mathbf{N}(0.25, 0.05)$, and (d) left-skewed distribution derived from $\mathbf{N}(0.45, 0.05)$.



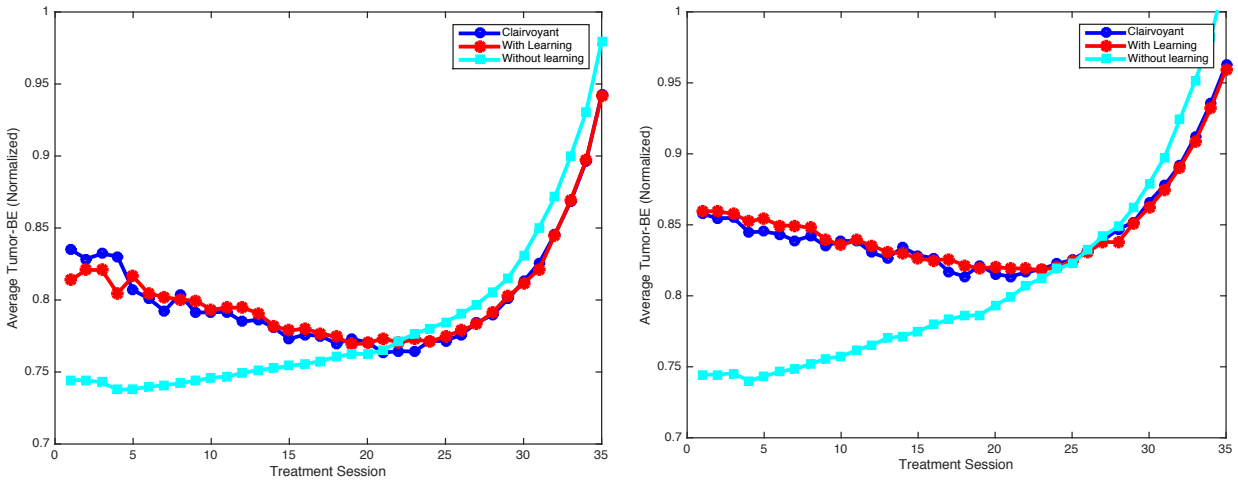
(a)

(b)

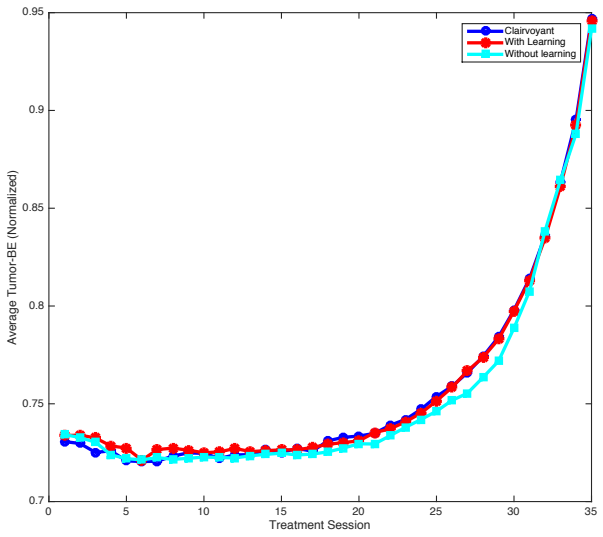


(c)

Figure 5.4: Total (in 35 sessions) average (over all voxels) tumor-BE in 30 independent simulations for (a) Symmetric distribution derived from $\mathbf{N}(0.35, 0.25/3)$, (b) right-skewed distribution derived from $\mathbf{N}(0.25, 0.05)$, and (d) left-skewed distribution derived from $\mathbf{N}(0.45, 0.05)$.



(a) (b)



(c)

Figure 5.5: Average (over voxels) tumor-BE across 35 treatment sessions, averaged over 30 independent simulations for (a) Symmetric distribution derived from $\mathbf{N}(0.35, 0.25/3)$, (b) right-skewed distribution derived from $\mathbf{N}(0.25, 0.05)$, and (d) left-skewed distribution derived from $\mathbf{N}(0.45, 0.05)$.

Chapter 6

FUTURE WORK

Similar to all existing optimization formulations for spatiobiologically separated or integrated radiotherapy, one fundamental limitation of this entire dissertation is its reliance on the LQ dose-response model. Although the LQ model is currently the standard radiobiological framework for mathematically evaluating dose-response, some practitioners are skeptical about its utility in clinical decision-making [28, 33]. Until a better-accepted radiobiological framework becomes available, however, theoretical research will likely continue to rely on the LQ model. If another mathematical model proves more accurate in the future, the general frameworks and solution approaches offered in this thesis will still remain valid, as the core of the proposed algorithms does not depend on the specific mathematical model being used. However, in such a case, other technical issues such as the convexity of the model and appropriate algorithms to discover the solutions may need to be re-examined.

There are currently no prevalent imaging techniques that can quantitatively measure damage to the OAR within the LQ framework. Thus, the research in Chapters 4 and 5 cannot be extended to adapt fluence-maps to the observed dose-response of the OAR within the LQ context.

The formulations in Chapters 4 and 5 assumed that a functional image is acquired at the beginning of every session. This assumption can be relaxed via simple changes to our notation and formulation without any substantial changes to our methodology as discussed in [30, 36]. In particular, our adaptive methods can be easily extended to the case where functional images are acquired only at a few time-points in the treatment course. Since the methods compared in these chapters employ identical imaging frequencies, the reported conclusions of the hypothesis tests are unlikely to change.

The quality, availability, safety, and cost of (repeated) functional images are currently not favorable for a clinical implementation of adaptive spatiobiologically integrated radiotherapy as discussed in [37, 38]. Some practitioners remain hopeful that this might change [63, 64, 70, 76]. More research and clinical studies would be needed before functional imaging techniques can be successfully combined with mathematical optimization models such as the ones presented in Chapters 4 and 5.

In developing our models in Chapters 4 and 5, we assumed that the tumor-response parameters are independent and identically distributed across voxels. This assumption can be relaxed by dividing the tumor volume into separate regions, each with its own set of tumor-response parameter-distributions. This change will increase the burden on the proposed algorithms. For instance, in Chapter 5, several distributions will have to be learned by the algorithm.

There are several potentially fruitful directions for extending this dissertation. Some of these are outlined below.

6.1 Uncertainty in dose deposition matrix

This dissertation ignores all uncertainties in the components a_{ij} of the dose-deposition matrices A. Recall that each component a_{ij} denotes the dose deposited in voxel i by beamlet j of unit intensity. The values of these components depend on the interactions of many complex factors and are in fact not known to the treatment planner. These include the patient's setup position (which can vary from fraction-to-fraction and within fraction), internal organ movements, and approximate physics-based simulations of the dose- deposition profile of each beamlet inside the patient's body. As a result, treatment planners, as we also did, usually employ a 'nominal' dose-deposition matrix based on intricate dose calculations for treatment planning purposes [32].

One possible way to explicitly incorporate such uncertainties in the components of various dose-deposition matrices could be as follows. We could model a dose-deposition matrix as a random matrix with correlated entries and then enforce that each BED constraint be met

with a sufficiently high probability. This would lead to a computationally demanding chance constrained formulation that might provide opportunities for future research.

6.2 Adaptive fractionation with imperfect information

Chapters 4 and 5 assumed that a "perfect" map of tumor cell density is available to the treatment planner via quantitative functional imaging. This does not hold in practice. It would be interesting in the future to generalize the original stochastic control formulations in [36, 53] for adaptive spatiobiologically integrated radiotherapy, and the work in Chapters 4 and 5 to the case of "imperfect information" [9]. This will lead to a partially observable Markov decision process [40] formulation of the adaptive fractionation problem and could provide an interesting avenue for future research.

BIBLIOGRAPHY

[1] A Adib and E Salari. Spatiotemporally integrated radiotherapy plan optimization. In *AAPM 59th Annual Meeting*, Denver, CO, 2017.

[2] M Alber, F Paulsen, S M Eschmann, and H J Machulla. On biologically conformal boost dose optimization. *Physics in Medicine and Biology*, 48(2):N31–N35, 2003.

[3] C I Armpilia, R G Dale, and B Jones. Determination of the optimum dose per fraction in fractionated radiotherapy when there is delayed onset of tumour repopulation during treatment. *The British Journal of Radiology*, 77(921):765–767, 2004.

[4] H Badri, Y Watanabe, and K Leder. Robust and probabilistic optimization of dose schedules in radiotherapy. available online at `arxiv.org/pdf/1503.00399v2.pdf`, June 2015.

[5] H Badri, Y Watanabe, and K Leder. Robust optimization of dose schedules in radiotherapy. available online at `arxiv.org/pdf/1503.00399v1.pdf`, March 2015.

[6] H Badri, Y Watanabe, and K Leder. Optimal radiotherapy dose schedules under parametric uncertainty. *Physics in Medicine and Biology*, 61(1):338, 2016.

[7] M Baumann, M Krause, J Overgaard, J Debus, S Bentzen, J Daartz, C Richter, D Zips, and T Bortfeld. Radiation oncology in the era of precision medicine. *Nature Reviews Cancer*, 16(4):234–250, 2016.

[8] A Ben-Tal, L El Ghaoui, and A Nemirovski. *Robust Optimization*. Princeton University Press, Princeton, NJ, USA, 2009.

[9] D P Bertsekas. *Dynamic Programming and Optimal Control, vol. 1 and 2*. Athena Scientific, Nashua, NH, 3rd edition, 2007.

[10] D Bertsimas, D B Brown, and C Caramanis. Theory and applications of robust optimization. *SIAM Review*, 53(3):464–501, 2011.

[11] A Bertuzzi, C Bruni F Papa, and C Sinisgalli. Optimal solution for a cancer radiotherapy problem. *Journal of Mathematical Biology*, 66(1-2):311–349, 2013.

[12] P Bhatnagar, M Subesinghe, C Patel, R Prestwich, and A F Scarsbrook. Functional imaging for radiation treatment planning, response assessment, and adaptive therapy in head and neck cancer. *RadioGraphics*, 33(7):1909–1929, 2013.

[13] A I Blanco, L A Meyer, G Verghese, B S Teh, A Rios, K Ferachi, M Rodriguez, A Gonzalez, and J Dalrymple. The use of modern imaging technologies in radiation therapy of cervical cancer. *Journal of Radiation Oncology*, 4(1):1–10, 2015.

[14] T Bortfeld, T C Y Chan, A Trofimov, and J N Tsitsiklis. Robust management of motion uncertainty in intensity modulated radiation therapy. *Operations Research*, 56(6):1461–1473, 2008.

[15] T Bortfeld, J Ramakrishnan, J N Tsitsiklis, and J Unkelbach. Optimization of radiation therapy fractionation schedules in the presence of tumor repopulation. *INFORMS Journal on Computing*, 27(4):788–803, 2015.

[16] J J Boutilier, T Craig, M B Sharpe, and T C Y Chan. Sample size requirements for knowledge-based treatment planning. *Medical Physics*, 43(3):1212–1221, 2016.

[17] S Boyd and L Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.

[18] C Burman, C-S Chui, G Kutcher, S Leibel, M Zelefsky, T LoSasso, S Spirou, Q Wu, J Yang, J Stein, R Mohan, Z Fuks, and C C Ling. Planning, delivery, and quality

assurance of intensity-modulated radiotherapy using dynamic multileaf collimator: A strategy for large-scale implementation for the treatment of carcinoma of the prostate. *International Journal of Radiation Oncology Biology Physics*, 39(4):863 – 873, 1997.

[19] T C Y Chan, T Bortfeld, and J Tsitsiklis. A robust approach to IMRT optimization. *Physics in Medicine and Biology*, 51:2567–2583, 2006.

[20] M Ehrgott, C Guler, H W Hamacher, and L Shao. Mathematical optimization in intensity modulated radiation therapy. *4OR*, 6(3):199–262, 2008.

[21] A Eisbruch. Intensity-modulated radiotherapy of head-and-neck cancer: encouraging early results. *International Journal of Radiation Oncology Biology Physics*, 53(1):1–3, 2002.

[22] B Emami, J Lyman, A Brown, L Coia, M Goiten, J E Munzenride, and et al. Tolerance of normal tissue to therapeutic radiation. *International Journal of Radiation Oncology Biology Physics*, 21(1):109–122, 1991.

[23] J F Fowler. How worthwhile are short schedules in radiotherapy?: A series of exploratory calculations. *Radiotherapy and Oncology*, 18(2):165–181, 1990.

[24] J F Fowler. Biological factors influencing optimum fractionation in radiation therapy. *Acta Oncologica*, 40(6):712–717, 2001.

[25] J F Fowler. Is there an optimal overall time for head and neck radiotherapy? a review with new modeling. *Clinical Oncology*, 19(1):8–27, 2007.

[26] J F Fowler. Optimum overall times II: Extended modelling for head and neck radiotherapy. *Clinical Oncology*, 20(2):113–126, 2008.

[27] J F Fowler. Sensitivity analysis of parameters in linear-quadratic radiobiologic modeling. *International Journal of Radiation Oncology Biology Physics*, 73(5):1532–1537, 2009.

[28] J F Fowler. 21 years of Biologically Effective Dose. *British Journal of Radiology*, 83(991):554–568, 2010.

[29] J F Fowler and M A Ritter. A rationale for fractionation for slowly proliferating tumors such as prostatic adenocarcinoma. *International Journal of Radiation Oncology Biology Physics*, 32(2):521–529, 1995.

[30] A Ghate. Dynamic optimization in radiotherapy. In J Geunes, editor, *Tutorials in Operations Research*, chapter 4, pages 60–74. Institute for Operations Research and the Management Sciences, Hanover, Maryland, USA, 2011.

[31] M Grant and S Boyd. CVX: MATLAB software for disciplined convex programming (web page and software), 2009.

[32] X Gu, D Choi, C Men, H Pan, A Majumdar, and S B Jiang. GPU-based ultra-fast dose calculation using a finite pencil beam model. *Physics in Medicine and Biology*, 54(20):6287–6297, 2009.

[33] E J Hall and A J Giaccia. *Radiobiology for the Radiologist*. Lippincott Williams & Wilkins, Philadelphia, Pennsylvania, USA, 2005.

[34] B Jones, L T Tan, and R G Dale. Derivation of the optimum dose per fraction from the linear quadratic model. *The British Journal of Radiology*, 68(812):894–902, 1995.

[35] H Keller, G Meier, A Hope, and M Davison. Fractionation schedule optimization for lung cancer treatments using radiobiological and dose distribution characteristics. *Medical Physics*, 39(6):3811–3811, 2012.

[36] M Kim. *A Mathematical Framework for Spatiotemporal Optimality in Radiation Therapy*. PhD thesis, University of Washington, Seattle, Washington, USA, 2010.

[37] M Kim, D L Craft, and C G Orton. Within the next five years, most radiotherapy treatment schedules will be designed using spatiotemporal optimization. *Medical Physics*, 43(5):2009–2012, 2016.

[38] M Kim, A Ghate, and M Phillips. A stochastic control formalism for dynamic biologically conformal radiation therapy. *European Journal of Operational Research*, 219(3):541–556, 2012.

[39] M Kim and M Phillips. A feasibility study of dynamic adaptive radiotherapy for non-small cell lung cancer. *Medical Physics*, 43(5):2153, 2016.

[40] V Krishnamurthy. *Partially observed Markov decision processes*. Cambridge University Press, Cambridge, United Kingdom, 1st edition, 2016.

[41] K Langen, J Polf, and R Schulte. Chapter 7: Imaging in Proton Therapy. In I J Das and H Paganetti, editors, *Principles and practice of proton beam therapy*, pages 191–209. Medical Physics Pub Corp., Madison, WI, 1st edition, 2015.

[42] M Langer, E K Lee, J O Deasy, R L Rardin, and J A Deye. Operations research applied to radiotherapy, an NCI-NSF-sponsored workshop February 7-9, 2002. *International Journal of Radiation Oncology Biology Physics*, 57(3):762–768, 2003.

[43] Z-Q Luo, W-K Ma, A M C So, and Y Ye. Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Processing Magazine*, 27(3):20–34, 2010.

[44] H Mahmoudzadeh, J Lee, T C Y Chan, and T G Purdie. Robust optimization methods for cardiac sparing in tangential breast IMRT. *Medical Physics*, 42(5):2212, 2015.

[45] H Mahmoudzadeh, T G Purdie, and T C Y Chan. Constraint generation methods for robust optimization in radiation therapy. forthcoming in Operations Research for Health Care, June 2015.

[46] P A Mar and T C Y Chan. Adaptive and robust radiation therapy in the presence of drift. *Physics in Medicine and Biology*, 60(9):3599–3615, 2015.

[47] L B Marks, E D Yorke, A Jackson, R K Ten Haken, L S Constine, A Eisbruch, S M Bentzen, J Nam, and J O Deasy. Use of normal tissue complication probability models in the clinic. *International Journal of Radiation Oncology Biology Physics*, 76(3):S10–S19, 2010.

[48] M Mizuta, S Takao, H Date, N Kishimoto, K L Sutherland, R Onimaru, and H Shirato. A mathematical study to select fractionation regimen based on physical dose distribution and the linear-quadratic model. *International Journal of Radiation Oncology Biology Physics*, 84(3):829 – 833, 2012.

[49] X S Qi, Q Yang, S P Lee, X A Li, and D Wang. An estimation of radiobiological parameters for head-and-neck cancer cells and the clinical implications. *Cancers*, 4:566–580, 2012.

[50] S Rockwell. Experimental radiotherapy: a brief history. *Radiation Research*, 150(Supplement):S157–S169, November 1998.

[51] H E Romeijn, R K Ahuja, J F Dempsey, and A Kumar. A new linear programming approach to radiation therapy treatment planning problems. *Operations Research*, 54(2):201–216, 2006.

[52] H E Romeijn and J F Dempsey. Intensity modulated radiation therapy treatment plan optimization. *TOP*, 16(2):215–243, 2008.

[53] F Saberian. *Convex and Dynamic Optimization with Learning for Adaptive Biologically Conformal Radiotherapy*. PhD thesis, University of Washington, Industrial & Systems Engineering, May 2015.

[54] F Saberian, A Ghate, and M Kim. Optimal fractionation in radiotherapy with multiple normal tissues. *Mathematical Medicine and Biology*, 33(2):211–252, 2015.

[55] F Saberian, A Ghate, and M Kim. A two-variable linear program solves the standard linear–quadratic formulation of the fractionation problem in cancer radiotherapy. *Operations Research Letters*, 43(3):254 – 258, 2015.

[56] F Saberian, A Ghate, and M Kim. A theoretical stochastic control framework for adapting radiotherapy to hypoxia. *Physics in Medicine and Biology*, 61(19):7136–7161, 2016.

[57] F Saberian, A Ghate, and M Kim. Spatiotemporally optimal fractionation in radiotherapy. *INFORMS Journal on Computing*, 29(3):422–437, 2017.

[58] D M Shepard, M C Ferris, G H Olivera, and T R Mackie. Optimizing the delivery of radiation therapy to cancer patients. *SIAM Review*, 41(4):721–744, 1999.

[59] C P South, P M Evans, and M Partridge. Dose prescription complexity versus tumor control probability in biologically conformal radiotherapy. *Medical Physics*, 36(10):4379–4388, 2009.

[60] C P South, M Partridge, and P M Evans. A theoretical framework for prescribing radiotherapy dose distributions using patient-specific biological information. *Medical Physics*, 35(10):4599–4611, 2008.

[61] A Sovik, E Malinen, O S Bruland, S M Bentzen, and D R Olsen. Optimization of tumor control probability in hypoxic tumors by radiation dose redistribution: a modeling study. *Physics in Medicine and Biology*, 52(2):499–513, 2007.

[62] A Sovik, E Malinen, H K Skogmo, Soren M Bentzen, O S Bruland, and D R Olsen. Radiotherapy adapted to spatial and temporal variability in tumor hypoxia. *International Journal Radiation Oncology Biology Physics*, 68(5):1496–1504, 2007.

[63] R D Stewart and X A Li. BGRT: Biologically guided radiation therapy — the future is fast approaching. *Medical Physics*, 34(10):3739–3751, 2007.

[64] D Thorwarth. Functional imaging for radiotherapy treatment planning: current status and future directions—a review. *British Journal of Radiology*, 88(1051):20150056, 2015.

[65] J Unkelbach. Non-uniform spatiotemporal fractionation schemes in photon radiotherapy. In *IFMBE Proceedings: World Congress on Medical Physics and Biomedical Engineering*, pages 401–404, Toronto, Canada, 2015.

[66] J Unkelbach, T C Y Chan, and T Bortfeld. Accounting for range uncertainties in the optimization of intensity modulated proton therapy. *Physics in Medicine and Biology*, 52:2755–2773, 2007.

[67] J Unkelbach, D Craft, E Saleri, J Ramakrishnan, and T Bortfeld. The dependence of optimal fractionation schemes on the spatial dose distribution. *Physics in Medicine and Biology*, 58(1):159–167, 2013.

[68] J Unkelbach and D Papp. The emergence of nonuniform spatiotemporal fractionation schemes within the standard BED model. *Medical Physics*, 42(5):2234–2241, 2015.

[69] J Unkelbach, C Zeng, and M Engelsman. Simultaneous optimization of dose distributions and fractionation schemes in particle radiotherapy. *Medical Physics*, 40(9):091702, 2013.

[70] W V Vogel, M G E H Lam, F A Pameijer, U A van der Heide, J B van den Kamer, M E Philippens, M van Vulpen, and M Verheij. Functional imaging in radiotherapy in the Netherlands: availability and impact on clinical practice. *Clinical Oncology*, 28(12):e206–e215, 2016.

[71] W C Waterhouse. Do symmetric problems have symmetric solutions? *The American Mathematical Monthly*, 90(6):378–387, 1983.

[72] S Webb. *Contemporary IMRT: Developing Physics and Clinical Implementation*. IOP Publishing, Bristol, UK, 2010.

[73] M V Williams, J Denekamp, and J F Fowler. A review of alpha/beta ratios for experimental tumors: implications for clinical studies of altered fractionation. *International Journal of Radiation Oncology Biology Physics*, 11(1):87–96, 1985.

[74] T Yamamoto, S Kabus, M Balc, P Kealld, S Benedict, and M Dalya. The first patient treatment of computed tomography ventilation functional image-guided radiotherapy for lung cancer. *Radiotherapy and Oncology*, 118(2):227–231, feb 2016.

[75] Y Yang and L Xing. Optimization of radiotherapy dose-time fractionation with consideration of tumor specific biology. *Medical Physics*, 32(12):3666–3677, 2005.

[76] Y Yang and L Xing. Towards biologically conformal radiation therapy (BCRT): selective IMRT dose escalation under the guidance of spatial biology distribution. *Medical Physics*, 32(6):1473–1484, 2005.

[77] M Zarepisheh, T Long, N Li, Z Tian, H E Romeijn, X Jiang, and S B Jiang. A DVH-guided IMRT optimization algorithm for automatic treatment planning and adaptive radiotherapy replanning. *Medical Physics*, 41(6Part1):061711, 2014.