Optimization Formulations and Algorithms for Cancer Therapy

Kelsey Lynn Maass

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

Doctor of Philosophy

University of Washington

2021

Reading Committee:

Aleksandr Aravkin, Chair

Minsun Kim

Anne Greenbaum

Program Authorized to Offer Degree:
Applied Mathematics

University of Washington

Abstract

Optimization Formulations and Algorithms for Cancer Therapy

Kelsey Lynn Maass

Chair of the Supervisory Committee:
Associate Professor Aleksandr Aravkin
Department of Applied Mathematics

Underlying all cancer therapy protocols are the competing objectives of maximizing tumor control and minimizing normal-tissue complications. As such, we can formulate many aspects of the cancer treatment planning workflow as optimization problems, enabling the development of mathematically rigorous treatment planning methods. In this dissertation, we present three novel optimization approaches to problems in cancer treatment planning: 1) a Markov decision process approach for optimizing multi-modality cancer therapy that balances the trade-off between tumor control and normal-tissue complication, 2) a nonconvex relaxation for the fluence map optimization problem for intensity-modulated radiation therapy that is well adapted to handle nonconvex dose–volume constraints, and 3) a hyperparameter optimization formulation for stereotactic body radiation therapy that has the potential to improve treatment plan quality and reduce the time needed to create a clinically acceptable treatment plan. We demonstrate the feasibility and potential benefit of each approach through numerical examples using synthetic and clinical cancer patient datasets. All project data and code are made openly available on GitHub.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

D2CM: Dose delivered within 2 cm of the PTV

DVH: Dose–volume histogram

EBRT: External-beam radiation therapy

FMO: Fluence map optimization

IMRT: Intensity-modulated radiation therapy

KBP: Knowledge-based planning

MDP: Markov decision process

MLC: Multileaf collimator

OAR: Organ-at-risk (normal tissue)

PTV: Planning target volume (tumor)

ROI: Region of interest

SBRT: Stereotactic body radiation therapy

TPS: Treatment planning system

ACKNOWLEDGMENTS

I would like to express my heartfelt gratitude to all of the people who have made this dissertation possible. First and foremost, thank you to my advisors, Aleksandr Aravkin and Minsun Kim, for introducing me to the worlds of optimization and radiation oncology. Your enthusiasm, support, and guidance have been indispensable throughout my time in graduate school. I would also like to thank the other members of my supervisory committee, Anne Greenbaum and Maryam Fazel, for their feedback and encouragement, and for taking the time to serve on my committee. Finally, thank you to Loyce Adams for giving me my first research opportunity as a master's student, and for encouraging me to pursue a PhD.

The Department of Applied Mathematics at the University of Washington has been an incredibly welcoming and supportive community. I would first like to thank the faculty for pushing me to become a better mathematician, and the staff for their important administrative support. In particular, thank you to Bernard Deconinck and Lauren Lederer for always having the time to listen to students and answer our many questions (and of course all the delicious baked goods). Next, I would like to thank my fellow students whose friendship and advice has kept me sane: my master's and PhD cohorts, research group, and office mates; everyone involved in the UW chapter of the Society for Industrial and Applied Mathematics, the Diversity Committee, the Women in Applied Mathematics Mentorship program, and Applied Knot Theory; and the many older students who have inspired me with their accomplishments. I would especially like to thank my counterpart, Kelsey Marcinko, for being a true friend and role model, and Ben Liu, for his genuine and insightful conversations.

Finally, none of my achievements would have been possible without the love and support

of my family. An enormous thank you to my parents, Mark and Robin, who have always believed in my abilities and encouraged me to pursue my interests. Thank you also to Whitney, Tim, and Eben, for cheering me on and providing much needed food and laughs. Last but not least, thank you Sid for your love, encouragement, and advice, and for your flexibility during these last few months of graduate school.

The work presented in this dissertation was supported in part by funding from the University of Washington Department of Applied Mathematics, the University of Washington Graduate School, a Boeing Fellowship in Applied Mathematics, a Seattle ARCS Foundation Fellowship (sponsored by the family of Dorothy Lewis Simpson), and National Science Foundation grants CMMI 1560476, DGE 1258485, and DMS 1839291.

DEDICATION

for my teachers

Chapter 1

INTRODUCTION

1.1 Cancer Therapy

Cancer is one of the leading causes of death worldwide, accounting for nearly 1 in every 6 deaths [158], with approximately 38.4% of the population expected to develop cancer at some point in their lives [79]. By 2040, the global cancer burden is projected to reach 27.5 million new cases and 16.2 million deaths [158]. This can be attributed in part to growing and aging populations, with more increases expected due to the adoption of behaviors and lifestyles associated with economic development and urbanization [156], a trend already observed in economically transitioning countries [157]. In the United States, cancer is the second leading cause of death; in fact, it is the leading cause of death in many states and amongst individuals younger than 80 [153].

Underlying all cancer therapy protocols are the competing objectives of maximizing tumor control and minimizing normal-tissue complications. As such, we can formulate many aspects of the cancer treatment planning workflow as optimization problems, enabling the development of mathematically rigorous treatment planning methods. In this dissertation, we present three novel optimization approaches to problems in cancer treatment planning: 1) a Markov decision process approach for optimizing multi-modality cancer therapy that balances the trade-off between tumor control and normal-tissue complication, 2) a nonconvex relaxation for the fluence map optimization problem for intensity-modulated radiation therapy that is well adapted to handle nonconvex dose–volume constraints, and 3) a hy-

perparameter optimization formulation for stereotactic body radiation therapy that has the potential to improve treatment plan quality and reduce the time needed to create a clinically acceptable treatment plan. We demonstrate the feasibility and potential benefit of each approach through numerical examples using synthetic and clinical cancer patient datasets.

1.1.1 Multi-modality treatment planning

There are several modalities currently used to treat cancer, including surgery, chemotherapy, and radiotherapy [157, 158]. Most patients receive treatment using two or more modalities, often sequentially, in the course of managing their cancer. With numerous patient characteristics, treatment modalities, and possible outcomes, making treatment decisions tailored to individual patients is extremely complex. It may no longer be practical to make decisions based solely on clinician experience and empirical intuition. In Chapter 2, we present a novel mathematical framework to optimize multi-modality treatment policies for cancer management using a finite-horizon Markov decision process approach, demonstrating the feasibility of the proposed model through numerical examples [104]. Our model has the potential to be personalized to individual patients via utility functions based on patient preferences, and it can also serve as a tool to explain an expected treatment course to patients.

1.1.2 Radiotherapy treatment planning

Many cancer therapies target the weakened DNA damage response of cancer cells; in particular, radiotherapy uses high-energy ionizing radiation to damage cancer cell DNA [81, 124]. Radiation has been used to treat cancer since the discovery of X-rays in 1895 [24], and it is used in up to two thirds of cancer cases in current practice [160]. The most common form of radiotherapy is external-beam radiation therapy, where photon beams generated by a linear accelerator are directed at the patient from a rotating gantry [80]. By surrounding the tumor with beams from multiple angles, physicians can design treatment plans that target tumors

with a variety of geometries while avoiding a large dose to critical structures.

In Chapter 3, we focus on the fluence map optimization problem for intensity-modulated radiation therapy, which can be formulated as a large-scale inverse problem with multiple objectives and constraints on the tumors and organs-at-risk [105]. Unfortunately, clinically relevant dose–volume constraints are nonconvex, so convex formulations and algorithms cannot be directly applied to the problem. We propose a novel approach to handle dose–volume constraints while preserving their nonconvexity, as opposed to previous efforts which focused on convex approximations. We develop efficient, provably convergent algorithms based on partial minimization, and show how to adapt them to handle maximum-dose constraints and infeasible problems. We demonstrate our approach using the CORT dataset [49, 50], and show that it is easily adaptable to radiation treatment planning with dose–volume constraints for multiple tumors and organs-at-risk.

Inverse planning to solve the fluence map optimization problem often requires the treatment planner to modify multiple parameters in the objective function to produce clinically relevant plans. Due to the manual steps in this process, the quality of plans can vary widely depending on the planning time available and the treatment planner's skills. In Chapter 4, we explore the feasibility of using hyperparameter optimization to improve treatment plan quality and efficiency for stereotactic body radiation therapy. Specifically, we tune objective function parameters using Bayesian optimization, comparing automatically generated plans for peripherally located lung tumors with plans created by medically certified dosimetrists at the University of Washington using the RayStation treatment planning system [28]. Our approach can be used in conjunction with any treatment planning system with a scripting interface, does not require access to previously optimized plans, and can improve upon manual plans by exploring a larger parameter space than what is typically considered in clinical settings.

1.2 Mathematical Preliminaries

Many interesting and important questions in science, engineering, and mathematics can be expressed in terms of minimizing or maximizing a function f relative to a set $\mathcal{C}$. A general optimization problem has the form

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) \leq b_i, \quad i = 1, \ldots, m, \\ & h_i(x) = c_i \quad i = 1, \ldots, n, \end{aligned} \tag{1.1}$$

with decision variable x, objective function f, constraint functions g_i and h_i, and constraint limits b_i and c_i. Applications of optimization abound, including problems in optimal control, signal and image processing, data analysis and statistics, finance, medicine, and more recently machine learning and artificial intelligence. We give a brief overview of some important aspects of mathematical optimization, including problem classes, solution methods, and optimality conditions, with a focus on concepts related to our model, algorithm, and convergence analysis in Chapter 3.

1.2.1 Problem classes

In any optimization problem, the form of the objective and constraints can impact the difficulty of the problem, the choice of solution method, and the criteria used to identify optimal solutions. For example, distinctions can be made between constrained vs. unconstrained, smooth vs. nonsmooth, and convex vs. nonconvex problems. The class of convex optimization problems is particularly significant, with theoretical results, state-of-the-art approaches, and commercial implementations readily available. A convex problem has objective and constraint functions that satisfy the inequality

$$f[\lambda x + (1 - \lambda)y] \leq \lambda f(x) + (1 - \lambda) f(y) \tag{1.2}$$

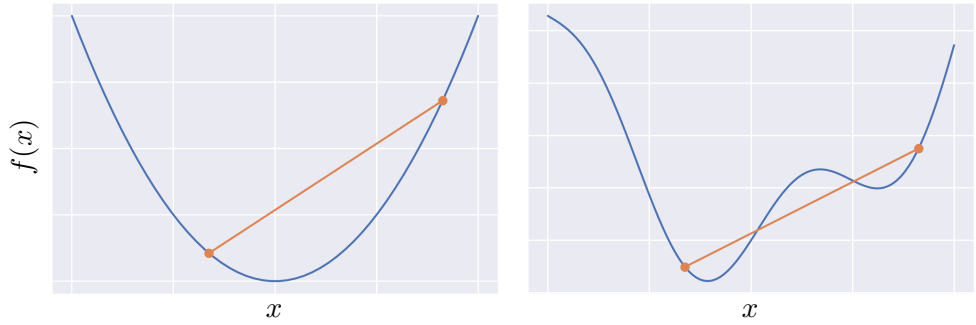

Figure 1.1: An example of convex and nonconvex functions. (Right) A convex function, where any line connecting two points $[x, f(x)]$ and $[y, f(y)]$ lies above the function. The local minimum is also the global minimum. (Left) A nonconvex function with two local minima and one global minimum.

for all vectors x and y on their respective domains and for all scalars $\lambda \in [0, 1]$. Graphically, this means that any line connecting two points $[x, f(x)]$ and $[y, f(y)]$ must lie above the function, as illustrated in Figure 1.1.

One important property of convex problems is that all local minima are also global. This result can be made stronger with additional assumptions; for example, if there exists a scalar $\mu > 0$ such that for all vectors x and y in the domain of f we have

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{\mu}{2} \|y - x\|_2^2, \tag{1.3}$$

where $\| \cdot \|_2^2$ denotes the Euclidean norm, the function is said to be μ-strongly convex and the optimal point is unique. While any differentiable convex function can be lower bounded by its tangent lines, that is

$$f(y) \geq f(x) + \nabla f(x)^T (y - x), \tag{1.4}$$

a strongly convex function can be lower bounded by a quadratic (see Figure 1.2). In contrast, if a function can be upper bounded by a quadratic, satisfying the inequality

$$f(y) \leq f(x) + \nabla f(x)^T (y - x) + \frac{L}{2} \|y - x\|_2^2 \tag{1.5}$$

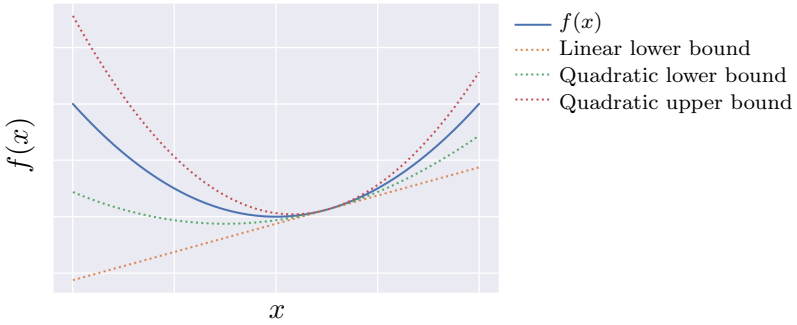

Figure 1.2: A graphical illustration of some useful inequalities for convex functions. While any differentiable convex function can be lower bounded by its tangent lines, a μ-strongly convex and L-Lipschitz smooth function can be lower and upper bounded by a quadratic, respectively.

for some scalar $L > 0$, it is said to be L-Lipschitz smooth. These inequalities provide bounds for how slowly or quickly a function can grow, which is useful for the convergence analysis of many iterative methods.

1.2.2 Solution methods

Very few optimization problems have a closed-form solution, so most problems are solved using iterative methods. One fundamental iterative method to minimize smooth functions is gradient descent, where vectors are refined using the update

$$x^{(k+1)} = x^{(k)} - \lambda \nabla f\left(x^{(k)}\right), \tag{1.6}$$

illustrated for a 1-D and 2-D problem in Figure 1.3. Gradient descent is often the first algorithm one encounters in the study of optimization, and it can be modified to work for a variety of problem classes.

For example, the gradient update cannot be applied to nonsmooth functions. In this case, we can use smoothing in the form of the Moreau envelope,

$$M_{\lambda f}(x) := \inf_y \left\{ f(y) + \frac{1}{2\lambda}\|y - x\|_2^2 \right\}, \tag{1.7}$$

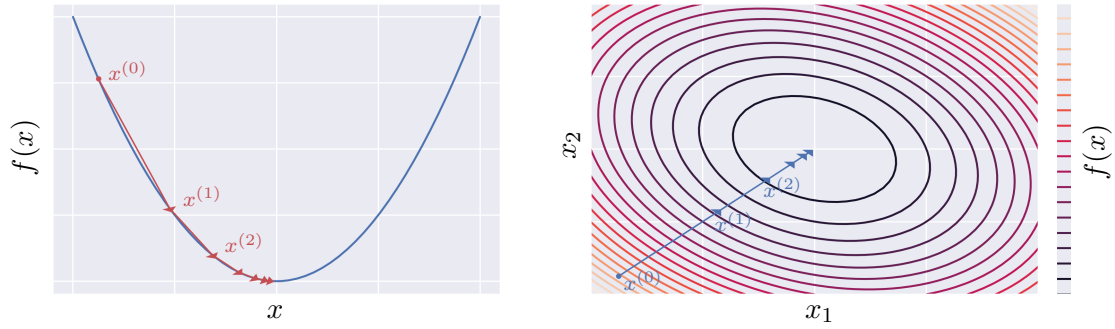

Figure 1.3: The gradient descent algorithm demonstrated on a 1-D and 2-D problem.

which is continuously differentiable and has the same set of minimizers as f [125]. This function is closely associated with the proximal operator

$$\text{prox}_{\lambda f}(x) := \arg\min_y \left\{ f(y) + \frac{1}{2\lambda}\|y - x\|_2^2 \right\}, \tag{1.8}$$

which can be interpreted as a generalized gradient descent update for f:

$$\text{prox}_{\lambda f}(x) = x - \lambda \nabla M_{\lambda f}(x). \tag{1.9}$$

One application of the proximal operator is in machine learning, where many objective functions are composed of a loss function g and a regularization function h, i.e.,

$$f(x) = g(x) + \lambda h(x). \tag{1.10}$$

While the loss function tends to be smooth, many popular regularization functions, such as the ℓ_1-norm or total-variation regularization, are not differentiable everywhere on their domains. A useful algorithm in this setting is proximal gradient descent,

$$x^{(k+1)} = \text{prox}_{\lambda h}\left[x^{(k)} - \lambda \nabla g\left(x^{(k)}\right)\right], \tag{1.11}$$

which takes a gradient step in g and a proximal step in h.

The proximal operator can also be viewed as a generalization of the projection operator, which can be used to solve constrained optimization problems. Specifically, using the

indicator function

$$\delta_{\mathcal{C}}(x) := \begin{cases} 0 & \text{if } x \in \mathcal{C}, \\ +\infty & \text{otherwise,} \end{cases} \tag{1.12}$$

we can express constrained problems of the form

$$\underset{x \in \mathcal{C}}{\text{minimize}} \quad f(x) \tag{1.13}$$

as unconstrained problems by moving the constraint into the objective function:

$$\underset{x}{\text{minimize}} \quad f(x) + \delta_{\mathcal{C}}(x). \tag{1.14}$$

The proximal gradient method applied to this problem is equivalent to projected gradient descent

$$x^{(k+1)} = \text{proj}_{\mathcal{C}}\left[x^{(k)} - \lambda \nabla f\left(x^{(k)}\right)\right], \tag{1.15}$$

where

$$\text{proj}_{\mathcal{C}}(x) := \underset{y \in \mathcal{C}}{\arg\min} \|y - x\|_2^2 = \underset{y}{\arg\min} \left\{ \delta_{\mathcal{C}}(y) + \frac{1}{2}\|y - x\|_2^2 \right\} = \text{prox}_{\delta_{\mathcal{C}}}(x). \tag{1.16}$$

In Chapter 3, we use a form of projected gradient descent to solve the fluence map optimization problem for intensity-modulated radiation therapy with nonconvex dose–volume constraints.

Gradient-based methods, such as the three mentioned in this section, are only one class of the many types of iterative methods used to solve optimization problems. For example, we use the backward induction algorithm for dynamic programming to solve a cancer treatment planning problem with multiple modalities in Chapter 2, and we use a Bayesian method for global optimization to solve a hyperparameter optimization formulation for radiotherapy treatment planning in Chapter 4.

1.2.3 Optimality conditions

The first steps of analyzing a new optimization problem often include deriving conditions to identify optimal solutions. Perhaps the most well-known optimality condition, familiar to calculus students, is the first-order necessary condition for unconstrained problems. Specifically, if x^* is a local minimizer of f, and f is differentiable at x^*, then $\nabla f(x^*) = 0$. If we also know that f is convex, this becomes a sufficient condition for global optimality at x^*. Not only can these criteria help to verify if a solution is locally or globally optimal, they can also be used to motivate algorithm design. For example, we can apply root-finding algorithms, such as Newton's method, to locate a point where the gradient is equal to zero. In this case, vectors are updated using information about the curvature at a point,

$$x^{(k+1)} = x^{(k)} - \left[\nabla^2 f\left(x^{(k)}\right)\right]^{-1} \nabla f\left(x^{(k)}\right), \tag{1.17}$$

which opens the door to many other second-order methods. Alternatively, the first-order condition implies that local minimizers are stationary points of (1.6), which makes the size of the gradient an intuitive stopping condition for the gradient descent algorithm.

For constrained optimization problems, a similar necessary condition can be derived based on the angle between the gradient at the optimal point and all points within the feasible set. For example, suppose we want to minimize a smooth function f over a convex set $\mathcal{C}$. Then if x^* is a local minimizer of f on $\mathcal{C}$, for all $x \in \mathcal{C}$ we have

$$\nabla f(x^*)^T (x - x^*) \geq 0. \tag{1.18}$$

Once again, if we also assume that f is convex, this becomes a sufficient condition for global optimality at x^*. Geometrically, this means that the negative gradient defines a supporting hyperplane to $\mathcal{C}$ at x^*, as illustrated on the left side of Figure 1.4. Furthermore, we see that the projection of $x^* - \lambda \nabla f(x^*)$ onto $\mathcal{C}$ is just x^*, so the local minimizer x^* is a fixed point of the projected gradient descent update in (1.15).

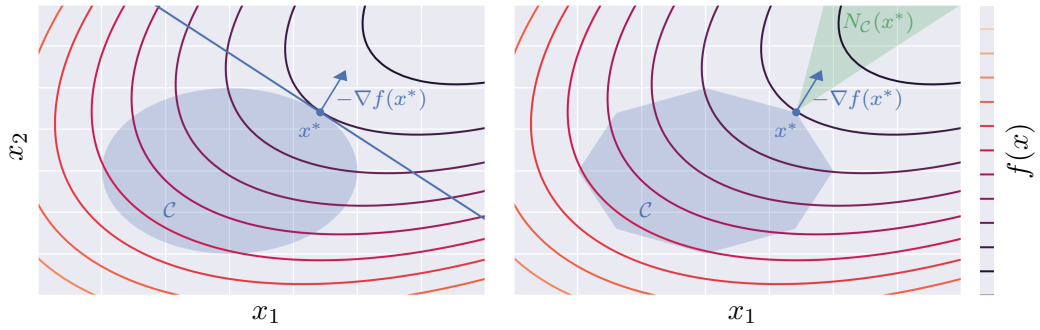

Figure 1.4: Illustration of local minima for constrained optimization problems. (Left) If a point x^* is a local minimizer of a smooth function f over a convex set $\mathcal{C}$, then the negative gradient at x^* defines a supporting hyperplane to $\mathcal{C}$ at x^*. Furthermore, if the boundary of $\mathcal{C}$ is smooth at x^*, then $-\nabla f(x^*) = N_{\mathcal{C}}(x^*)$. (Right) If the boundary of $\mathcal{C}$ at x^* is not smooth, then $-\nabla f(x^*) \in N_{\mathcal{C}}(x^*)$.

We can generalize this optimality condition using a set called the normal cone. For a convex set $\mathcal{C}$, the normal cone of $\mathcal{C}$ at a point x is defined as

$$N_{\mathcal{C}}(x) := \left\{ v : v^T(y - x) \leq 0 \text{ for all } y \in \mathcal{C} \right\}. \tag{1.19}$$

Using this definition, the first-order necessary condition in (1.18) is equivalent to

$$-\nabla f(x^*) \in N_{\mathcal{C}}(x^*), \tag{1.20}$$

as illustrated on the right side of Figure 1.4. This condition also holds when the set $\mathcal{C}$ is nonconvex, but in this case we consider a more general definition of the normal cone. Specifically, we say that a vector v is normal to $\mathcal{C}$ in a general sense, i.e., $v \in N_{\mathcal{C}}(x)$, if there exists a sequence $x^{(k)} \to_{\mathcal{C}} x$ and $v^{(k)} \to v$ such that

$$\left(v^{(k)}\right)^T \left(y - x^{(k)}\right) \leq o\left(\left\|y - x^{(k)}\right\|_2\right) \text{ for all } y \in \mathcal{C}, \tag{1.21}$$

where $o\left(\|y - x^{(k)}\|_2\right) / \|y - x^{(k)}\|_2 \to 0$ when $y \to_{\mathcal{C}} x^{(k)}$ with $y \neq x^{(k)}$. In particular, we use the optimality condition in (1.20) for our convergence analysis in Chapter 3, where we minimize a smooth function over a nonconvex set.

The optimality conditions discussed here are only a few of the many necessary and sufficient conditions that exist for local and global optimality, and we have only provided a small glimpse of the variety of problem classes and solution methods studied within the field of mathematical optimization. For a more in-depth overview of optimization problems, methods, optimality conditions, and applications, see [25, 34, 113, 120, 121, 130, 132, 135].

1.3 Bibliographic Notes

This dissertation is based upon the following publications.

1. Kelsey Maass and Minsun Kim. A Markov decision process approach to optimizing cancer therapy using multiple modalities. *Mathematical Medicine and Biology: A Journal of the IMA*, 37(1):22–39, 2020. (Reproduced by permission of Oxford University Press).

2. Kelsey Maass, Minsun Kim, and Aleksandr Aravkin. A nonconvex optimization approach to IMRT planning with dose–volume constraints. *arXiv preprint arXiv:1907.10712*, 2020. (Submitted for publication).

3. Kelsey Maass, Minsun Kim, and Aleksandr Aravkin. A hyperparameter optimization approach for automated radiotherapy treatment planning. (In preparation).

Chapter 2

A MARKOV DECISION PROCESS APPROACH FOR MULTI-MODALITY CANCER THERAPY

2.1 Introduction

There are many treatment modalities currently used for cancer therapy, including surgery, chemotherapy, radiotherapy, hormone therapy, immunotherapy, and targeted therapy [157, 158]. Because "*cancer treatments using a single therapeutic agent often result in limited clinical outcomes due to tumor heterogeneity and drug resistance,*" while "*combination therapies using multiple modalities can synergistically elevate anti-cancer activity while lowering doses of each agent, hence, reducing side effects*" [84], most patients receive treatment using two or more modalities, often sequentially, in the course of managing their cancer. Some examples of multi-modality cancer therapy include combining gene therapy and chemotherapy with nanotechnology [84]; using surgery, chemotherapy, and radiotherapy for head and neck cancers [57]; combining immunotherapy with the more traditional surgery, radiotherapy, chemotherapy, and targeted therapy [144]; treating glioblastomas with a combination of surgery, radiotherapy, and systemic therapy [162]; treating brain metastases using a combination of surgery, radiotherapy, and symptomatic care [109]; treating localized rectal cancer using a combination of surgery, chemotherapy, radiotherapy, and adjuvant cytotoxic therapy [12]; and chemoradiotherapy plus surgery for esophageal cancer [149].

In current practice, multi-modality treatment decisions rely predominantly upon individual clinicians' experiences, where the optimality of the resulting treatment plans is unclear. Considering the variety of treatment options available in modern medicine and the numerous possible outcomes associated with each treatment course, depending upon intuition or a

heuristic search for optimal multi-modality treatment policies can be costly and inefficient [21, 145]. There have been limited efforts to model optimal multi-modality treatments using a mathematical approach. For example, Beil and Wein study the optimal sequencing of surgery, chemotherapy, and radiotherapy, using ordinary differential equations to describe the behavior of the primary tumor and metastases [19]. They suggest two novel treatment sequences from their study; however, a potential weakness of their model is the requirement of an accurate, *a priori* estimate of 14 different equation parameters. Alternatively, Hathout et al. investigate the optimal combination of radiotherapy and surgery in the treatment of glioblastomas, using a reaction–diffusion partial differential equation to simulate tumor proliferation and response to radiation [68]. Their model provides answers regarding when surgical resection, in combination with radiotherapy, adds survival benefits.

In this chapter, we propose a novel mathematical framework to optimize multi-modality treatment policies for cancer therapy using a finite-horizon Markov decision process (MDP) approach, and we demonstrate the feasibility and potential of the proposed model with numerical examples. An MDP with a finite planning horizon is a mathematical framework for optimizing a sequence of actions within a stochastic system, where the state of the system is given at the beginning of each decision epoch, and the goal is to maximize the expected reward at the end of the planning horizon. MDPs have been used to model problems in various industries including robotics [86, 139] and economics [37, 140], and more recently they have been applied to problems in medicine. For example, MDPs have been used to find optimal treatments for sperocytosis [108] and ischemic heart disease [69], to determine the optimal dose in sequential radiation treatment periods [85], and to decide whether to accept or reject an available kidney or liver for transplantation [5, 6, 145]. In addition, a combination of MDPs and dynamic decision networks has been used to develop a general-purpose artificial intelligence framework that can "think like a doctor" for personalized medicine, improving patient outcomes and decreasing costs [21]. In our MDP model, we define the system state

as a combination of the patient's degree of tumor progression (or tumor control) and normal-tissue side effect. The action space consists of three different treatment modalities categorized by repetition, tumor reduction, and risk to normal tissue. At each decision epoch, the clinician observes the patient's state and chooses an optimal treatment modality to maximize the expected terminal reward, which is a function of the patient's final state.

The rest of this chapter is organized as follows. We introduce our model formulation in Section 2.2, with details on the treatment planner's actions (treatment modalities), patient states, state transition probabilities, and reward functions, and we review the backward induction algorithm for computing optimal policies. In Section 2.3, we present numerical examples using various intuitive reward functions to show the structural insights of optimal policies and the potential benefits of using a mathematically rigorous model for optimal multi-modality cancer therapy. Specifically, we demonstrate that changes in the reward functions and state transition probabilities result in changes to the optimal policy that correspond with clinical intuition. Finally, we conclude the chapter and discuss possible extensions to our model in Section 2.4.

2.2 Problem Formulation

Consider a treatment course with T treatment periods, where a patient seeks an optimal treatment decision at each period. To calculate the patient's optimal treatment policy, we define an MDP model with four components: the treatment planner's actions (Section 2.2.1), the patient state (Section 2.2.2), the state transition probabilities (Section 2.2.3), and the intermediate and terminal reward functions (Section 2.2.4). Finally, we review the backward induction algorithm to solve the recursive Bellman equations in Section 2.2.5.

2.2.1 Treatment planner's actions

We denote the action space as $\mathcal{A} = \{M_1, M_2, M_3\}$, where M_1, M_2, and M_3 represent the Type 1, 2, and 3 modalities, respectively. The definition of each modality type is as follows:

- Type 1: Treatment modalities with a high risk (increasing side effect) and high reward (decreasing tumor progression). The frequency of administering Type 1 modalities is restricted.

- Type 2: Treatment modalities with a lower risk and lower reward than Type 1. May be repeated without restriction in frequency.

- Type 3: No treatment (surveillance). Has a higher probability of reducing normal-tissue side effect and increasing tumor progression than the Type 1 and Type 2 modalities. May be repeated without restriction in frequency.

We categorize treatment modalities into three types acknowledging that certain modalities may be more effective but limited in the frequency of their administration. For example, whole-brain radiotherapy to manage brain metastases is only done once during a patient's lifetime due to its associated normal-tissue side effect, whereas gamma knife and partial-brain external-beam radiotherapy can be administered multiple times until normal-tissue toxicity has reached its tolerance level [119]. For such cases, whole-brain therapy is categorized as Type 1, while gamma knife and external-beam radiotherapy are categorized as Type 2 modalities.

In general, M_1 and M_2 may each represent a set of modalities, but to simplify notation we first consider the case where there is only one modality of each type. (We consider a case with two modalities of Type 2 in Section 2.3.5.) Additionally, we restrict the Type 1 modality to one-time use during the course of treatment, though this is certainly not a requirement of our model. It is also noted that our model does not have any restriction

on the order of administering treatment modalities, allowing full flexibility in selecting an optimal modality for a given patient state.

2.2.2 Patient state

Let the total number of possible states for normal-tissue side effect and tumor progression be $m+1$ and $n+1$, respectively. Let $s_t = (h_t, \phi_t, \tau_t) \in \mathcal{S}$ denote the patient's state in treatment period t with $t = 1, 2, \ldots, T$. Each state variable is defined as follows:

- $h_t \in \mathcal{H} = \{0, 1\}$: History of Type 1 modality in period t such that $h_t = 1$ if the Type 1 modality has been used in periods $1, 2, \ldots,$ or $t-1$. Otherwise $h_t = 0$.

- $\phi_t \in \Phi = \{0, 1, \ldots, m\}$: Normal-tissue side effect due to treatment observed in period t. Φ is ordered such that $\phi_t = 0$ represents no side effect while $\phi_t = m$ represents the worst possible side effect (patient death due to normal-tissue side effect).

- $\tau_t \in \mathcal{T} = \{0, 1, \ldots, n\}$: Tumor progression observed in period t. $\mathcal{T}$ is ordered such that $\tau_t = 0$ represents the best patient state (tumor remission) and $\tau_t = n$ represents the worst possible patient state (patient death due to tumor progression).

We denote the full patient state space by the Cartesian product $\mathcal{S} = \mathcal{H} \times \Phi \times \mathcal{T}$.

2.2.3 State transition probabilities

When treatment modality $a_t \in \mathcal{A}$ is implemented in period t for a patient in state s_t, the patient's state in period $t+1$ is assumed to be s_{t+1} with probability $P_t(s_{t+1} \mid s_t, a_t)$. We assume that the transition probabilities for each state variable are conditionally independent of one another, that is,

$$P_t(s_{t+1} \mid s_t, a_t) = P_t^{\mathcal{H}}(h_{t+1} \mid h_t, a_t) \times P_t^{\Phi}(\phi_{t+1} \mid \phi_t, a_t) \times P_t^{\mathcal{T}}(\tau_{t+1} \mid \tau_t, a_t). \tag{2.1}$$

This assumption is in line with our intuition that tumor progression and normal-tissue side effects depend on treatment type only and not on each other. The classification of M_1 as having a higher risk and higher reward than M_2 can then be written in terms of state transition probabilities as follows:

$$\text{Higher risk in side effects:} \quad P_t^{\Phi}(\phi_{t+1} \mid \phi_t, M_1) \geq P_t^{\Phi}(\phi_{t+1} \mid \phi_t, M_2) \text{ for } \phi_{t+1} > \phi_t, \tag{2.2}$$

$$\text{Higher reward in tumor control:} \quad P_t^{\mathcal{T}}(\tau_{t+1} \mid \tau_t, M_1) \geq P_t^{\mathcal{T}}(\tau_{t+1} \mid \tau_t, M_2) \text{ for } \tau_{t+1} < \tau_t. \tag{2.3}$$

We define the transition probabilities for the history variable, h_t, deterministically to restrict the number of administrations of M_1 during the course of treatment:

$$P_t^{\mathcal{H}}(1 \mid h_t, M_1) = 1 \text{ for all } h_t \in \mathcal{H}, \tag{2.4}$$

$$P_t^{\mathcal{H}}(h_t \mid h_t, a_t) = 1 \text{ for } a_t \in \{M_2, M_3\}. \tag{2.5}$$

The restriction that Type 1 modalities may only be used once during the course of treatment is then imposed through the transition probabilities in the following manner:

$$P_t(1, m, n \mid 1, \phi_t, \tau_t, M_1) = 1 \text{ for all } \phi_t \in \Phi \text{ and } \tau_t \in \mathcal{T}. \tag{2.6}$$

This means that the patient will transition to the worst possible state, i.e., $s_{t+1} = (1, m, n)$, if M_1 is chosen when the history of M_1 use is positive.

We impose absorbing boundary conditions to simulate either the death of the patient (when either side effect or tumor progression reaches its maximum value) or tumor remission (when tumor progression reaches zero). We note that in the remission state only tumor progression is fixed, allowing side effect to improve in subsequent treatment periods.

$$\text{Death due to side effect:} \quad P_t(h_t, m, \tau_t \mid h_t, m, \tau_t, a_t) = 1 \text{ for all } a_t \in \mathcal{A}, \tag{2.7}$$

$$\text{Death due to tumor progression:} \quad P_t(h_t, \phi_t, n \mid h_t, \phi_t, n, a_t) = 1 \text{ for all } a_t \in \mathcal{A}, \tag{2.8}$$

$$\text{Tumor remission:} \quad P_t^{\mathcal{T}}(0 \mid 0, a_t) = 1 \text{ for all } a_t \in \mathcal{A}. \tag{2.9}$$

2.2.4 Reward functions

We denote the real-valued terminal reward function $r_{T+1}(s)$, which quantifies the patient's utility of being in state s at the end of their treatment course, i.e., $t = T + 1$. After each treatment period the patient may also receive an intermediate reward, $r_t(s_t, a_t, s_{t+1})$, which is associated with the action chosen in period t, such as the cost of using treatment a_t or the expected outcome in the patient's next state, s_{t+1}. While our model will work with any reward function, in general patient utility corresponding to better states should be at least as large as patient utility corresponding to worse states. This means that

$$r(\phi, \tau) \geq r(\phi', \tau) \text{ for } \phi \leq \phi', \text{ and} \tag{2.10}$$

$$r(\phi, \tau) \geq r(\phi, \tau') \text{ for } \tau \leq \tau', \tag{2.11}$$

because states are ordered such that smaller states represent better patient conditions. Some commonly used utility measures in medicine are the quality-adjusted life year, the disability-adjusted life year, and the healthy-years equivalent [110, 117, 143, 176].

2.2.5 Bellman equations and backward induction

For each patient state $s_t \in \mathcal{S}$ and treatment period $t \in \{1, 2, \ldots, T\}$, our goal is to maximize the expected patient utility at the end of the treatment course, that is,

$$E\left[\sum_{t=1}^{T} r_t(s_t, a_t, s_{t+1}) + r_{T+1}(s_{T+1})\right]. \tag{2.12}$$

Bellman's recursive equations to solve this problem are given by

$$V_t(s) = \sum_{s' \in \mathcal{S}} P_t(s' \mid s, a)\Big[r_t(s, a, s') + V_{t+1}(s')\Big] \text{ for } t = 1, 2, \ldots, T, \tag{2.13}$$

with boundary condition $V_{T+1}(s) = r_{T+1}(s)$. The optimal policy can be solved recursively for all $s_t \in \mathcal{S}$ and $t = 1, 2, \ldots, T$ with the well-known backward induction algorithm [130].

Algorithm 1 Backward induction [130].

Set $V_{T+1}(s) = r_{T+1}(s)$ for all $s \in \mathcal{S}$

for $t = T, T-1, \ldots, 1$ **do**

$$V_t(s) = \max_{a \in \mathcal{A}} \sum_{s' \in \mathcal{S}} P_t(s' \mid s, a) \Big[r_t(s, a, s') + V_{t+1}(s') \Big]$$

$$a_t(s) = \arg\max_{a \in \mathcal{A}} \sum_{s' \in \mathcal{S}} P_t(s' \mid s, a) \Big[r_t(s, a, s') + V_{t+1}(s') \Big]$$

end for

With increasing dimensions in the state space $\mathcal{S}$ and action space $\mathcal{A}$, the problem faces the "curses of dimensionality" and may be solved using approximate dynamic programming [129].

2.3 Numerical Examples

In this section we present numerical examples to illustrate the structure of optimal multi-modality treatment policies generated with our MDP model. In Section 2.3.1 we present a base case of the state transition probabilities and reward functions used in our simulations, including the general assumptions used to make them clinically relevant. Next, in Sections 2.3.2 and 2.3.3 we explore how changes in the terminal and intermediate reward functions affect the optimal policies. The effect of changes in transition probabilities on the optimal policies is presented in Section 2.3.4. Finally, we explore a scenario with multiple Type 2 treatment modalities in Section 2.3.5. We note that the general assumptions made in this section are specific to our numerical examples and not necessary for solving optimal multi-modality treatment policies using our model. In practice, transition probabilities can be estimated from correlations between treatment modalities and patient outcomes from the clinical literature, which is beyond the scope of this chapter.

2.3.1 Base case

We define a base case with one modality of each type, where $\mathcal{A} = \{M_1, M_2, M_3\}$, and without intermediate rewards, letting $r_t = 0$ for $t = 1, 2, \ldots, T$. In all of our numerical examples, we use 11 states for ϕ and τ, i.e., $m = n = 10$, and three treatment periods, that is, $T = 3$. First, in Section 2.3.1.1 we specify our assumptions used to assign state transition probabilities, and we introduce our reward functions derived from clinical intuition and practice in Section 2.3.1.2. This is followed by the optimal policies resulting from these base transition probabilities and rewards in Section 2.3.1.3.

2.3.1.1 State transition probabilities

For our numerical examples, we utilize stationary transition probabilities that depend only on the changes between states, rather than on the actual value of the state. This means that

$$P_t(s_{t+1} \mid s_t, a_t) = P(s_{t+1} \mid s_t, a_t) \text{ for } t = 1, 2, \ldots, T, \tag{2.14}$$

$$P^{\Phi}(\phi_{t+1} \mid \phi_t, a_t) = P^{\Phi}(\phi'_{t+1} \mid \phi'_t, a_t) \text{ whenever } \phi_{t+1} - \phi_t = \phi'_{t+1} - \phi'_t, \text{ and} \tag{2.15}$$

$$P^{\mathcal{T}}(\tau_{t+1} \mid \tau_t, a_t) = P^{\mathcal{T}}(\tau'_{t+1} \mid \tau'_t, a_t) \text{ whenever } \tau_{t+1} - \tau_t = \tau'_{t+1} - \tau'_t. \tag{2.16}$$

For simplicity, we assume that state variables can only change by one increment between two successive treatment periods, and that tumor progression only improves after treatment (M_1 or M_2), while side effect only improves after surveillance (M_3). Therefore, tumor progression can either stay the same or get better (decrease by one) between two successive treatment periods after M_1 or M_2 is chosen, while side effect can either stay the same or get worse (increase by one). This implies that

$$P^{\mathcal{T}}(0 \mid 0, a_t) = 1 \text{ and } P^{\Phi}(m \mid m, a_t) = 1 \text{ for } a_t \in \{M_1, M_2\}, \tag{2.17}$$

which is consistent with our absorbing boundary conditions. When M_3 (surveillance) is chosen, tumor progression can either stay the same or get worse between two successive

Table 2.1: State transition probabilities $P^{\Phi}(\phi_{t+1} \mid \phi_t, a_t)$ and $P^{\mathcal{T}}(\tau_{t+1} \mid \tau_t, a_t)$ for the base case.

Modality	**Side Effect** (ϕ_{t+1})			**Tumor Progression** (τ_{t+1})		
(a_t)	$\phi_t - 1$	ϕ_t	$\phi_t + 1$	$\tau_t - 1$	τ_t	$\tau_t + 1$
M_1	0.0	0.4	0.6	0.7	0.3	0.0
M_2	0.0	0.6	0.4	0.6	0.4	0.0
M_3	0.6	0.4	0.0	0.0	0.3	0.7

treatment periods, while side effect can either stay the same or get better. This implies that

$$P^{\mathcal{T}}(n \mid n, M_3) = 1 \text{ and } P^{\Phi}(0 \mid 0, M_3) = 1. \tag{2.18}$$

For our numerical examples, we use the state transition probabilities for non-boundary states defined in Table 2.1. The transition probabilities for the history variable h are as defined in (2.4) and (2.5).

2.3.1.2 Reward functions

The reward (patient utility) of being in state s depends on the patient's side effect (ϕ) and tumor progression (τ). We define our patient utility using additively separable reward functions, where f measures the utility of the normal-tissue side effect being in state ϕ, and g measures the utility of tumor progression being in state τ. Let d_ϕ and d_τ be the parameters associated with functions f and g. Then the total reward is defined as

$$r(\phi, \tau) = c_\phi f(\phi; d_\phi) + c_\tau g(\tau; d_\tau), \tag{2.19}$$

where c_ϕ and c_τ are the weighting factors of f and g, which represent the relative importance of the side effect and tumor progression, respectively.

We use concave reward functions for f and g, where improvements made in worse patient states are more appreciated than improvements made in healthier states. These relationships

can often be found in studies that examine patient utility as a function of clinical states [13, 42, 53, 112]. Thus, the functions f and g can be written as

$$\text{Side effect:} \quad f(\phi; d_\phi) = \frac{100}{m^{d_\phi}} \left(m^{d_\phi} - \phi^{d_\phi} \right), \tag{2.20}$$

$$\text{Tumor progression:} \quad g(\tau; d_\tau) = \frac{100}{n^{d_\tau}} \left(n^{d_\tau} - \tau^{d_\tau} \right), \tag{2.21}$$

where $d_\phi, d_\tau \geq 1$ so that f and g are concave. (Note that $d_\phi, d_\tau = 1$ indicates a linear function.) The functions f and g are normalized to 100 so that the reward ranges between 0 (minimum) and 100 (maximum). We note that the concave shape of the reward function is not necessary to our model, but is used in our examples to simulate clinically relevant scenarios as previously demonstrated in [85].

Finally, we use a stationary intermediate reward function that does not depend on the action, i.e., $r_t(s_t, a_t, s_{t+1}) = r(s_t, s_{t+1})$ for $t = 1, 2, \ldots, T$. For the base case, we use exponents $d_\phi, d_\tau = 2$ and weights $c_\phi, c_\tau = 1/2$ in the terminal reward function (r_{T+1}), without intermediate reward functions, that is, $r(s_t, s_{t+1}) = 0$ for $t = 1, 2, \ldots, T$.

2.3.1.3 Base case

Figure 2.1 contains the optimal policy for the base case, with transition probabilities from Table 2.1 and rewards functions described in Section 2.3.1.2. On the left we see the isolines of the terminal reward function, with the patient state defined in terms of tumor progression on the horizontal axis and normal-tissue side effect on the vertical axis. With lighter regions corresponding to larger rewards, the lower-left corner is the patient's best state, i.e., zero side effect and tumor remission, while the upper-right corner is the patient's worst state, i.e., maximum side effect and tumor progression. On the right we show the optimal policy, again with states defined in terms of tumor progression on the horizontal axis and normal-tissue side effect on the vertical axis. The top row contains the optimal policy when the Type 1 modality has never been used ($h = 0$), and the bottom row shows the optimal policy for

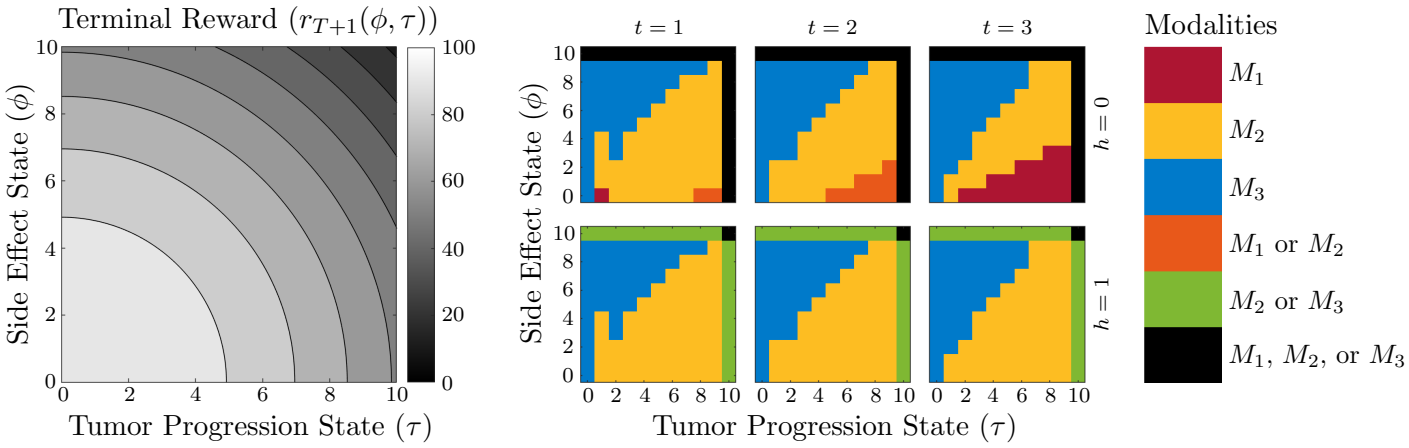

Figure 2.1: (Left) Terminal reward function $r_{T+1}(\phi, \tau) = 1/2 f(\phi; 2) + 1/2 g(\tau; 2)$. (Right) Optimal treatment policy for $T = 3$ and $\mathcal{A} = \{M_1, M_2, M_3\}$ with state transition probabilities given in Table 2.1.

$h = 1$. Each column represents a distinct treatment period.

The results agree with clinical intuition that, in general, when the patient's normal-tissue side effect is more detrimental to their overall health than their tumor progression (upper-left corner), the optimal action is surveillance (M_3). On the other hand, when the patient's current state is in the lower-right corner, i.e., severe tumor progression with minimal side effects, the optimal action is to use more aggressive treatment modalities (M_1 if it has not been used). At the absorption states where $\phi = m$ or $\tau = n$, there is no difference among modalities when M_1 has not been used. However, since using M_1 when $h = 1$ deterministically brings the patient's next state to the worst possible state, the optimal policy with $h = 1$ does not include M_1 for any states or treatment periods other than the worst state, $s_t = (1, m, n)$. We note that with $h = 0$, the Type 1 modality tends to be saved for later treatment periods, so M_2 is used more often in the beginning of the treatment course. This observation will be compared with the case in Section 2.3.3, where the patient's state during intermediate treatment periods contributes to the total reward via an intermediate reward function, unlike the base case where the total reward depends only on the patient's final state.

2.3.2 Effect of terminal reward function on optimal policy

In this section we demonstrate how changes in the terminal reward function affect optimal policies. Specifically, we consider various shapes of the terminal reward function in Section 2.3.2.1 and the relative importance of side effect and tumor progression in the terminal reward function in Section 2.3.2.2.

2.3.2.1 Shape of terminal reward function

We examine how the shape of the terminal reward function, represented by the exponents d_ϕ and d_τ, affects the optimal policy when side effect and tumor progression are weighted equally, with $c_\phi, c_\tau = 1/2$. Using the same exponent for f and g, letting $d_\phi, d_\tau = d$, we compare the optimal policies generated using the following terminal reward function:

$$r_{T+1}(\phi, \tau) = \frac{1}{2} f(\phi, d) + \frac{1}{2} g(\tau, d), \tag{2.22}$$

$$= \frac{50}{10^d}\Big[(10^d - \phi^d) + (10^d - \tau^d)\Big], \tag{2.23}$$

$$\text{where } d \in \{3/2, 2 \text{ (base case)}, 3\}. \tag{2.24}$$

The exponent $d = 1$ corresponds to a linear function, where the reward of decreasing tumor progression or side effect is identical in all patient states. When $d > 1$, the reward received from decreasing tumor progression or side effect in worse patient states is larger than the reward received in better states. As d increases, the reward at each patient state also increases, except at the best and worst states where rewards are fixed at zero and 100. This in turn results in a steeper slope and faster transition from "bad" states to "good" states. The optimal policies resulting from $d = 3/2$ and $d = 3$ are shown in Figures 2.2 and 2.3, respectively, which can be compared with the base case in Figure 2.1.

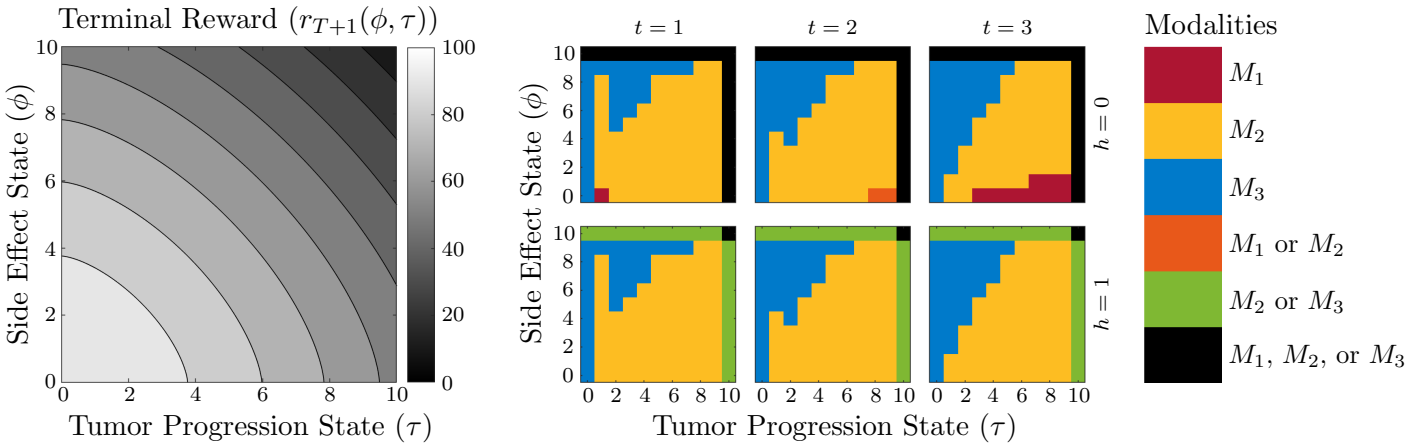

Figure 2.2: (Left) Terminal reward function $r_{T+1}(\phi, \tau) = 1/2 f(\phi; 3/2) + 1/2 g(\tau; 3/2)$. (Right) Optimal treatment policy for $T = 3$ and $\mathcal{A} = \{M_1, M_2, M_3\}$ with state transition probabilities given in Table 2.1.

As d approaches one, improvements made in all patient states (better or worse) are rewarded equally, so less emphasis is put on getting patients out of the worst possible states. This is demonstrated in the optimal policy included in Figure 2.2, where d is decreased from 2 in the base case to $3/2$, with less surveillance in states with high side effect (upper-left corner) and less of the M_1 modality in states with high tumor progression (lower-right corner). Alternatively, as we increase the exponent d to 3 in Figure 2.3, improvements made near the worst patient states (the worst tumor progression or worst side effect) are rewarded higher than improvements made in other states. This produces treatment policies that prioritize getting the patient out of these states, with both more surveillance in states with high side effect (upper-left corner) and more of the M_1 modality in states with worse tumor progression (lower-right corner). For the remainder of our numerical examples, we use quadratic terminal reward functions, letting $d = 2$, in order to compare with the base case.

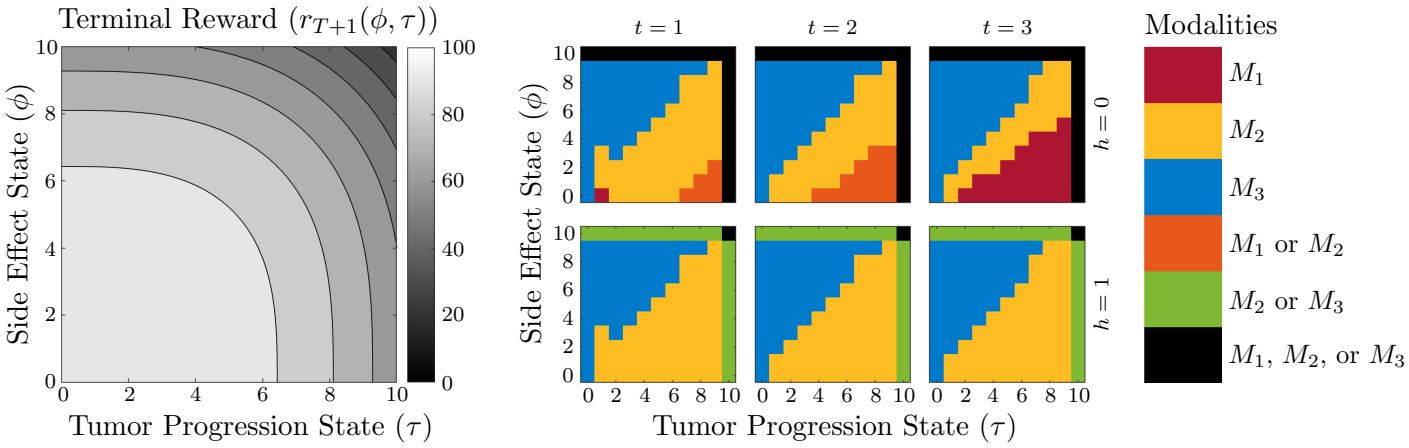

Figure 2.3: (Left) Terminal reward function $r_{T+1}(\phi, \tau) = 1/2 f(\phi; 3) + 1/2 g(\tau; 3)$. (Right) Optimal treatment policy for $T = 3$ and $\mathcal{A} = \{M_1, M_2, M_3\}$ with state transition probabilities given in Table 2.1.

2.3.2.2 Relative importance of side effect and tumor progression

Next we look at the effect of the relative importance of side effect and tumor progression, with $c_\phi = c$ and $c_\tau = (1 - c)$, in the concave terminal reward function given by

$$r_{T+1}(\phi, \tau) = cf(\phi; 2) + (1 - c)g(\tau; 2), \tag{2.25}$$

$$= c\left(100 - \phi^2\right) + (1 - c)\left(100 - \tau^2\right), \tag{2.26}$$

where $c \in \{1/3, 1/2 \text{ (base case)}, 2/3\}$. The optimal policies with $c = 1/3$ and $2/3$ are shown in Figures 2.4 and 2.5, respectively. As we weigh tumor progression more heavily than side effect in the terminal reward function, with $c = 1/3$ in Figure 2.4, the optimal policies become more aggressive. Specifically, decreasing c produces policies that recommend treatment over surveillance in more states than in the base case, with an increased frequency of using M_1 when $h = 0$ and M_2 when $h = 1$. On the other hand, when side effect has a higher weight than tumor progression, with $c = 2/3$ in Figure 2.5, the optimal policy becomes less aggressive, with surveillance chosen for more states than in the base case.

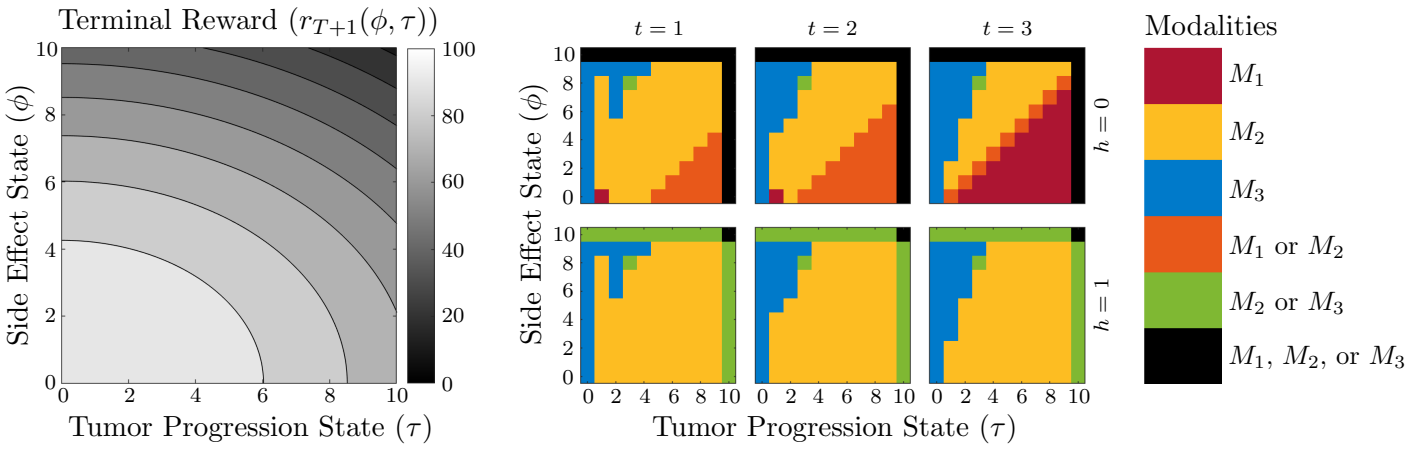

Figure 2.4: (Left) Terminal reward function $r_{T+1}(\phi, \tau) = 1/3 f(\phi; 2) + 2/3 g(\tau; 2)$. (Right) Optimal treatment policy for $T = 3$ and $\mathcal{A} = \{M_1, M_2, M_3\}$ with state transition probabilities given in Table 2.1.

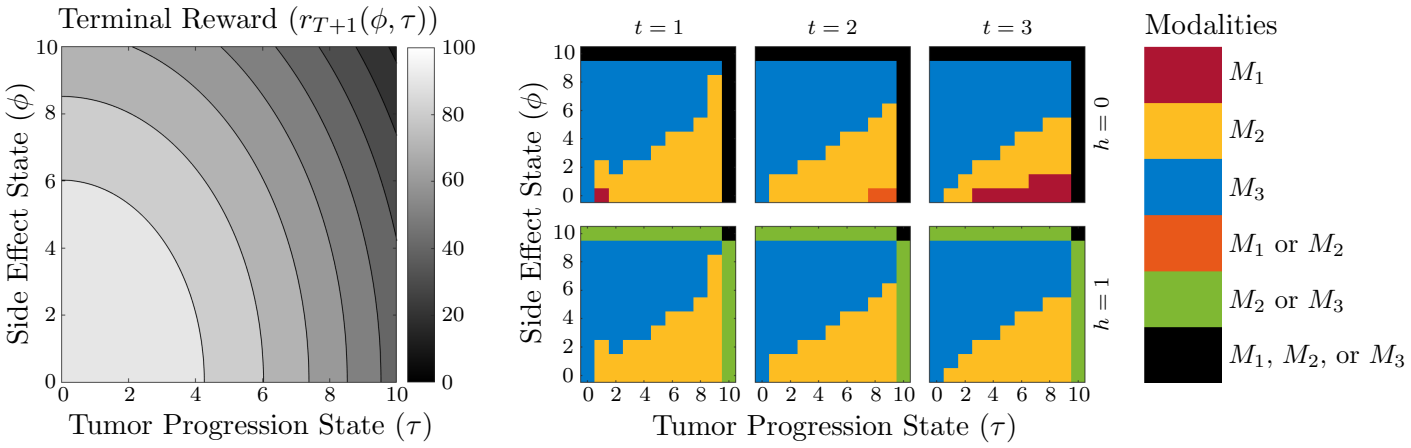

Figure 2.5: (Left) Terminal reward function $r_{T+1}(\phi, \tau) = 2/3 f(\phi; 2) + 1/3 g(\tau; 2)$. (Right) Optimal treatment policy for $T = 3$ and $\mathcal{A} = \{M_1, M_2, M_3\}$ with state transition probabilities given in Table 2.1.

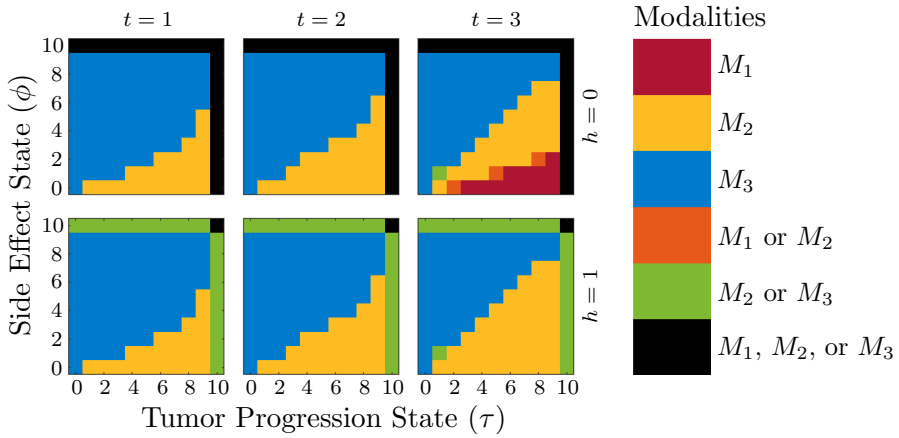

Figure 2.6: Optimal treatment policy for $T = 3$ and $\mathcal{A} = \{M_1, M_2, M_3\}$ with terminal reward function $r_{T+1}(\phi, \tau) = 1/2 f(\phi; 2) + 1/2\, g(\tau; 2)$, intermediate reward function $r_\phi(\phi) = 1/4 f(\phi; 2)$, and state transition probabilities given in Table 2.1.

2.3.3 Effect of intermediate rewards on optimal policy

In this section we explore the effect of adding an intermediate reward function on the optimal policies. We consider two cases, where we collect rewards for 1) reducing side effect during the treatment course by adding r_ϕ, and 2) reducing tumor progression during the treatment course by adding r_τ. Using quadratic rewards ($d_\phi, d_\tau = 2$), the intermediate reward functions r_ϕ and r_τ are defined as

$$r_\phi(s_t, s_{t+1}) = c_m f(\phi_{t+1}; 2), \quad t = 1, 2, \ldots, T, \text{ and} \tag{2.27}$$

$$r_\tau(s_t, s_{t+1}) = c_m g(\tau_{t+1}; 2), \quad t = 1, 2, \ldots, T, \tag{2.28}$$

where we let $c_m = 1/4$. The optimal policies using r_ϕ and r_τ are presented in Figures 2.6 and 2.7, respectively.

In Figure 2.6, we see that the optimal policy using r_ϕ tends to select actions that reduce side effect during the treatment course; therefore, the policy has a higher proportion of surveillance (M_3), a lower proportion of M_1 and M_2 with $h = 0$, and a lower proportion of M_2 with $h = 1$ in all treatment periods as compared to the base case in Figure 2.1. We note that M_1, which can only be used once, is saved for the last treatment period ($t = T$)

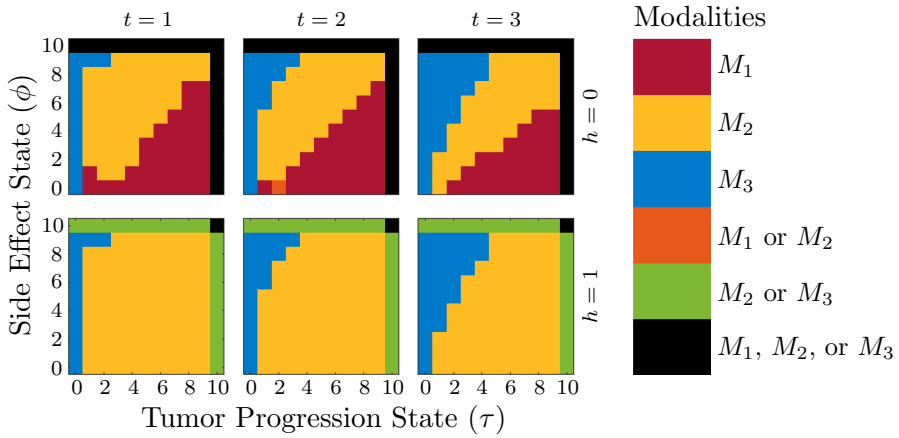

Figure 2.7: Optimal treatment policy for $T = 3$ and $\mathcal{A} = \{M_1, M_2, M_3\}$ with terminal reward function $r_{T+1}(\phi, \tau) = 1/2 f(\phi; 2) + 1/2 g(\tau; 2)$, intermediate reward function $r_\tau(\tau) = 1/4 g(\tau; 2)$, and state transition probabilities given in Table 2.1.

when r_ϕ is used. This can be compared to the policy in Figure 2.5, where the importance of reducing side effect is larger in the terminal reward than the importance of reducing tumor progression. When c_ϕ is larger than c_τ, without intermediate rewards, the optimal policy still includes M_1 for certain states in earlier treatment periods.

On the other hand, the optimal policy using r_τ in Figure 2.7 tends to select actions that reduce tumor progression during the treatment course. As compared to the base case in Figure 2.1, the proportion of M_3 decreases in all treatment periods, the proportion of M_1 and M_2 with $h = 0$ increases, and the proportion of M_2 with $h = 1$ increases. Therefore the preference for treatment over surveillance is higher. We note the difference in the optimal policy resulting from adding r_τ to that of Figure 2.4, where c_τ is increased in the terminal reward. The proportion of M_1, which can only be used once, is higher in all treatment periods when r_τ is used. This is different from using a larger c_τ than c_ϕ in the terminal reward function, where M_1 is used in a larger portion of the states in the last treatment period than in the earlier periods.

Table 2.2: State transition probabilities $P^{\Phi}(\phi_{t+1} \mid \phi_t, a_t)$ and $P^{\mathcal{T}}(\tau_{t+1} \mid \tau_t, a_t)$ with the assumption that M_1 is more effective than the base case in terms of decreasing tumor progression.

Modality	**Side Effect** (ϕ_{t+1})			**Tumor Progression** (τ_{t+1})		
(a_t)	$\phi_t - 1$	ϕ_t	$\phi_t + 1$	$\tau_t - 1$	τ_t	$\tau_t + 1$
M_1	0.0	0.4	0.6	**0.8**	**0.2**	0.0
M_2	0.0	0.6	0.4	0.6	0.4	0.0
M_3	0.6	0.4	0.0	0.0	0.3	0.7

2.3.4 Effect of transition probabilities on optimal policy

In this section we show how changes in transition probabilities affect the optimal treatment policy. First we explore a case where the treatment modality M_1 is more effective in reducing tumor progression than in the base case, where we increase the probability that the tumor progression will decrease after treatment, that is,

$$P^{\mathcal{T}}(\tau_t - 1 \mid \tau_t, M_1) = 0.8 \quad \text{and} \quad P^{\mathcal{T}}(\tau_t \mid \tau_t, M_1) = 0.2. \tag{2.29}$$

The optimal policy computed using the transition probabilities in Table 2.2 is shown in Figure 2.8. It shows a more aggressive treatment policy that uses M_1 in a higher proportion of the states as compared to the base case. We also note an interesting effect of the boundary condition at $\tau = 0$; specifically, M_1 is chosen for several states where $\tau = 1$ and $\phi < 4$ (without severe side effect) at $t = 1$ due to the fact that tumor remission in earlier treatment periods $(t = 1, 2)$ is now more likely if M_1 is chosen.

Next we explore a case where the treatment modality M_2 is less risky in terms of increasing side effect than in the base case, where we decrease the probability that the side effect will increase after treatment, that is,

$$P^{\Phi}(\phi_t \mid \phi_t, M_2) = 0.7 \quad \text{and} \quad P^{\Phi}(\phi_t + 1 \mid \phi_t, M_2) = 0.3. \tag{2.30}$$

The state transition probabilities in this case are shown in Table 2.3, and the resulting

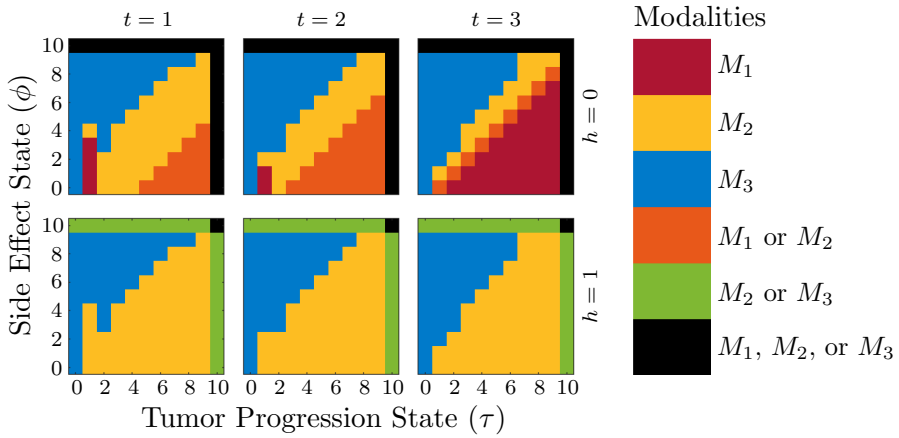

Figure 2.8: Optimal treatment policy for $T = 3$ and $\mathcal{A} = \{M_1, M_2, M_3\}$ with terminal reward function $r_{T+1}(\phi, \tau) = {}^1\!/_2 f(\phi, 2) + {}^1\!/_2\, g(\tau, 2)$ and state transition probabilities given in Table 2.2, modified to make M_1 more effective than the base case in decreasing tumor progression.

Table 2.3: State transition probabilities $P^{\Phi}(\phi_{t+1} \mid \phi_t, a_t)$ and $P^{\mathcal{T}}(\tau_{t+1} \mid \tau_t, a_t)$ with the assumption that M_2 is less risky than the base case in terms of increasing side effect.

Modality	**Side Effect** (ϕ_{t+1})			**Tumor Progression** (τ_{t+1})		
(a_t)	$\phi_t - 1$	ϕ_t	$\phi_t + 1$	$\tau_t - 1$	τ_t	$\tau_t + 1$
M_1	0.0	0.4	0.6	0.7	0.3	0.0
M_2	0.0	**0.7**	**0.3**	0.6	0.4	0.0
M_3	0.6	0.4	0.0	0.0	0.3	0.7

optimal policy is shown in Figure 2.9. We see that this change makes M_2 more favorable and produces an optimal policy that uses M_2 in a higher proportion of the states in all treatment periods as compared to the base case.

Finally, we investigate a case where the surveillance modality M_3 is less likely to increase tumor progression than the base case, where we change probabilities in the fourth row of Table 2.1 to

$$P^{\mathcal{T}}(\tau_t \mid \tau_t, M_3) = 0.7 \quad \text{and} \quad P^{\mathcal{T}}(\tau_t + 1 \mid \tau_t, M_3) = 0.3. \tag{2.31}$$

The state transition probabilities in this case are shown in Table 2.4, and the resulting

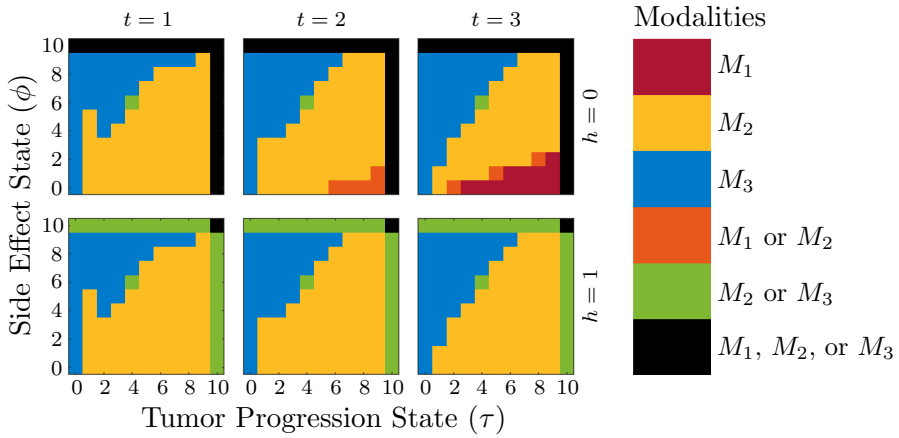

Figure 2.9: Optimal treatment policy for $T = 3$ and $\mathcal{A} = \{M_1, M_2, M_3\}$ with terminal reward function $r_{T+1}(\phi, \tau) = {}^1\!/_2 f(\phi, 2) + {}^1\!/_2\, g(\tau, 2)$ and state transition probabilities given in Table 2.3, modified to make M_2 less risky than the base case in increasing side effect.

Table 2.4: State transition probabilities $P^{\Phi}(\phi_{t+1} \mid \phi_t, a_t)$ and $P^{\mathcal{T}}(\tau_{t+1} \mid \tau_t, a_t)$ with the assumption that M_3 is less risky than the base case in terms of increasing tumor progression.

Modality	**Side Effect** (ϕ_{t+1})			**Tumor Progression** (τ_{t+1})		
(a_t)	$\phi_t - 1$	ϕ_t	$\phi_t + 1$	$\tau_t - 1$	τ_t	$\tau_t + 1$
M_1	0.0	0.4	0.6	0.7	0.3	0.0
M_2	0.0	0.6	0.4	0.6	0.4	0.0
M_3	0.6	0.4	0.0	0.0	**0.7**	**0.3**

optimal policy is shown in Figure 2.10. We see an increase in the amount of surveillance suggested by the policy in all treatment periods. We also note that M_3 is optimal in more states with higher side effects as compared to the base case. This policy may be appropriate in the case of a slowly growing tumor, where the probability of increasing tumor progression by surveillance is relatively low for a given treatment period.

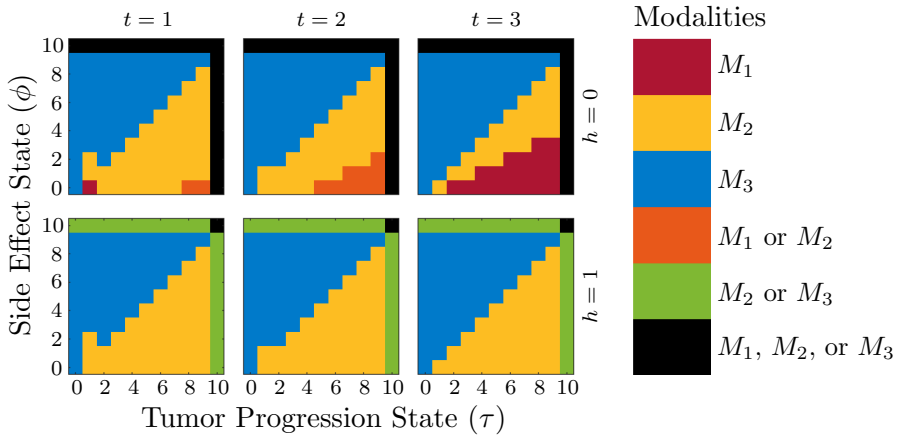

Figure 2.10: Optimal treatment policy for $T = 3$ and $\mathcal{A} = \{M_1, M_2, M_3\}$ with terminal reward function $r_{T+1}(\phi, \tau) = {}^1\!/_2 f(\phi, 2) + {}^1\!/_2 g(\tau, 2)$ and state transition probabilities given in Table 2.4, modified to make M_3 less risky than the base case in increasing tumor progression.

2.3.5 Multiple Type 2 modalities

In this section we demonstrate our model with multiple modalities of a particular treatment type. Specifically, we consider a case where there are two modalities of Type 2, i.e., action space $\mathcal{A} = \{M_1, M_{2a}, M_{2b}, M_3\}$. We order the index of the treatment modalities based on risk and effectiveness, that is,

$$P^{\Phi}(\phi_{t+1} \mid \phi_t, M_1) \geq P^{\Phi}(\phi_{t+1} \mid \phi_t, M_{2a}) \geq P^{\Phi}(\phi_{t+1} \mid \phi_t, M_{2b}) \geq P^{\Phi}(\phi_{t+1} \mid \phi_t, M_3) \tag{2.32}$$

for $\phi_{t+1} > \phi_t$, and

$$P^{\mathcal{T}}(\tau_{t+1} \mid \tau_t, M_1) \geq P^{\mathcal{T}}(\tau_{t+1} \mid \tau_t, M_{2a}) \geq P^{\mathcal{T}}(\tau_{t+1} \mid \tau_t, M_{2b}) \geq P^{\mathcal{T}}(\tau_{t+1} \mid \tau_t, M_3) \tag{2.33}$$

for $\tau_{t+1} < \tau_t$. The transition probabilities used in this case are presented in Table 2.5, and the resulting optimal policy is shown in Figure 2.11. The trend is qualitatively similar to the previous cases with three modalities. We note that there are only ties between consecutive treatment modalities; for example, M_1 never ties with M_{2b} or M_3 due to the ordering in the action space.

Table 2.5: State transition probabilities $P^{\Phi}(\phi_{t+1} \mid \phi_t, a)$ and $P^{\mathcal{T}}(\tau_{t+1} \mid \tau_t, a)$ for four treatment modalities $\mathcal{A} = \{M_1, M_{2a}, M_{2b}, M_3\}$.

Modality	**Side Effect** (ϕ_{t+1})			**Tumor Progression** (τ_{t+1})		
(a_t)	$\phi_t - 1$	ϕ_t	$\phi_t + 1$	$\tau_t - 1$	τ_t	$\tau_t + 1$
M_1	0.0	0.4	0.6	0.7	0.3	0.0
M_{2a}	0.0	0.5	0.5	0.6	0.4	0.0
M_{2b}	0.0	0.6	0.4	0.5	0.5	0.0
M_3	0.6	0.4	0.0	0.0	0.3	0.7

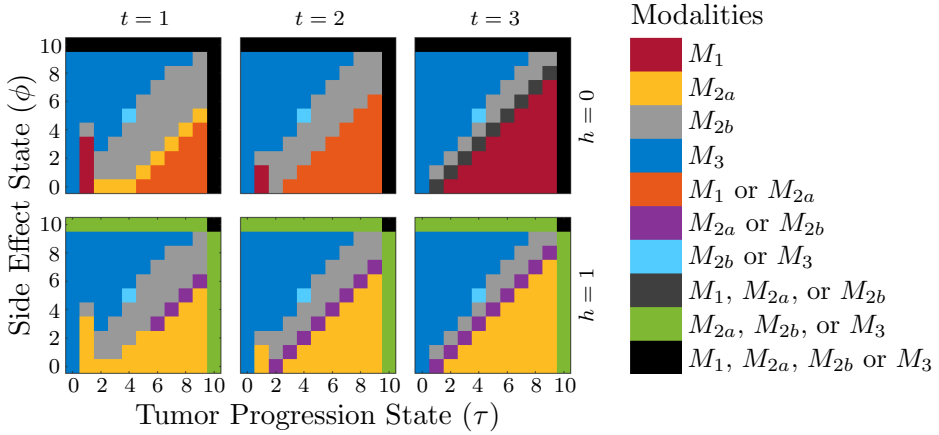

Figure 2.11: Optimal treatment policy for $T = 3$ and $\mathcal{A} = \{M_1, M_{2a}, M_{2b}, M_3\}$ with probabilities given in Table 2.5 and terminal reward function $r_{T+1}(\phi, \tau) = {}^1\!/_2 f(\phi; 2) + {}^1\!/_2 g(\tau; 2)$.

2.4 Discussion

Most patients receive treatment using two or more modalities, often sequentially, in the course of managing their cancer. However, with diverse patient characteristics, numerous available treatment modalities, and various possible outcomes, making multi-modality treatment decisions tailored to each individual patient is extremely complex, and it may no longer be practical to make optimal decisions based solely on individual clinicians' experiences and empirical intuition. We approached this problem using a novel mathematical framework to model optimal treatment policies for cancer therapy using a finite-horizon Markov decision process (MDP). Through numerical examples using simplified patient states and clinically intuitive reward functions, we have shown the potential application of our model to support treatment decision-making or to serve as a tool to explain an expected treatment course to patients.

Using state transition probabilities obtained from treatment–outcome clinical data, which could potentially be non-stationary over time, our model can assist clinicians in making optimal decisions for a particular patient's current state. Our model has the potential to be further personalized to individual patients via custom utility functions based on the patient's own preferences, with the addition of state variables such as age, specific side effects, and tumor types. However, as we incorporate more realistic and detailed attributes into the model, the problem becomes computationally intractable due to increased state and outcome spaces. We leave these high dimensional problems, which require an approximate dynamic programming approach, for future work. Other potential formulations for future investigation include infinite-horizon MDPs leading to stationary optimal solutions. An infinite horizon formulation is particularly useful when the length of a patient's expected treatment course is unpredictable.

Chapter 3

A NONCONVEX OPTIMIZATION APPROACH FOR IMRT PLANNING WITH DOSE–VOLUME CONSTRAINTS

3.1 Introduction

Many cancer therapies target the weakened DNA damage response of cancer cells; in particular, radiotherapy uses high-energy ionizing radiation to damage cancer cell DNA [81, 124]. Unfortunately, this radiation also damages the healthy cells in its path, so the main goal of radiation treatment planning is to maximize the differential between damage to the tumor and to nearby healthy tissue. The most common form of radiotherapy is external-beam radiation therapy (EBRT), where photon beams generated by a linear accelerator are directed at a patient from a rotating gantry [80]. By surrounding a tumor with beams from multiple angles, physicians can design treatment plans that target tumors with a variety of geometries while avoiding a large dose to critical structures.

One way to limit damage to healthy tissue is to require that treatment plans satisfy various constraints on the amount of radiation delivered to organs-at-risk (OARs). For example, it is often clinically relevant to impose maximum-dose constraints on a fraction of an organ, known as a *dose–volume constraint*. Unfortunately, these constraints are nonconvex, so convex formulations and algorithms cannot be directly applied to the treatment planning problem. We propose a novel approach to handle dose–volume constraints for intensity-modulated radiation therapy (IMRT) that preserves their nonconvex structure, as opposed to previous efforts that focused on convex approximations. The proposed method is amenable to efficient algorithms based on partial minimization and naturally adapts to handle maximum-dose constraints and cases of infeasiblity. We demonstrate our approach using the CORT

dataset [49, 50], and we show that it easily adapts to radiation treatment planning with dose–volume constraints for multiple tumors and OARs. We start with a brief overview of EBRT, including IMRT, and the specific challenges posed by dose–volume constraints.

3.1.1 External-beam radiation therapy

Conventional EBRT began soon after the discovery of X-rays in 1895 and was significantly impacted by technological innovations such as advanced imaging and multileaf collimators (MLC) [31]. Specifically, the invention of computed tomography (CT) in the 1970s enabled accurate three-dimensional mapping of a patient's internal anatomy, while the introduction of the MLC to radiotherapy in the 1990s provided more control over beam shape and fluence. For example, three-dimensional conformal radiation therapy (3D-CRT) uses the sliding leaves of the MLC to shape beam outlines to match tumor geometries, reducing radiation exposure to healthy tissue, while IMRT uses MLCs to not only shape beam outlines, but also modulate beam fluence, giving additional control over the patient dose distribution [175].

In the early history of radiotherapy, treatment plans were created using a trial-and-error approach called *forward planning*, where physicians would calculate dose distributions for different sets of beam parameters until an acceptable treatment plan was found. However, modulating beam fluence for IMRT involves discretizing each beam into subunits, called beamlets or bixels, which dramatically increases the number of variables under consideration, rendering forward planning infeasible [30, 56]. In the 1980s, the idea of *inverse planning* for IMRT was introduced by Brahme [36, 35], where mathematical optimization was used to calculate beam parameters to deliver a specified dose distribution.

Within radiotherapy inverse planning, there are a variety of different paradigms regarding dose objectives, mathematical models, and optimization methods. For instance, while our proposed model and many others seek to optimize dose distributions defined in terms of *physical criteria* such as dose and volume [30, 45], it is also possible to consider *biological*

criteria such as tumor control probability (TCP) and normal-tissue complication probability (NTCP) [9], equivalent uniform dose (EUD) [51, 122, 163], and biologically effective dose (BED) [141, 142]. Additionally, many early approaches to inverse planning used analytical *direct methods* inspired by the similarities between IMRT planning and CT image reconstruction [32, 47, 73], but the trend has since moved towards approximate *iterative methods* [185, 152, 184].

Amongst iterative methods, the presence of nonconvexity within many IMRT models has fueled the debate between *stochastic methods* such as simulated annealing [107, 173, 174] and genetic algorithms [3, 183], which can escape local minima but tend to have higher computational costs, versus *gradient-based methods* using techniques such as Newton's method [32, 181], L-BFGS [89, 128], and conjugate gradient methods [8, 159], which tend to be faster but can get trapped in local minima. Our method uses a projected gradient approach to handle issues related to nonconvexity, which we discuss in Section 3.1.3. For a more detailed history of radiotherapy and IMRT, see [30, 31, 45, 56, 151, 175].

3.1.2 Intensity-modulated radiation therapy

In current practice, the patient undergoes imaging (e.g., CT) before IMRT treatment, producing a three-dimensional model of their internal anatomy. This model is discretized into volume units (voxels) and labeled by structure. Next, the physician specifies the prescription, often in terms of a uniform target to the tumor with various constraints on critical structures. The goal of the treatment planner is then to determine the number of beams, beam angles, and beamlet intensities that will deliver the prescribed dose of radiation to the tumor while keeping doses to healthy tissue low. Inverse planning for IMRT typically involves three steps [56]:

1. Beam angles: Determine the number of beams and their orientations.

2. Fluence map: Calculate beamlet intensities to deliver the desired dose distribution.

3. Segmentation: Design MLC sequences that achieve the optimized fluence map.

These steps require solving optimization problems that can be approached separately or in combination [4, 26, 93, 188]. In this chapter, we focus on the fluence map optimization (FMO) problem of assigning beamlet intensities for a given set of beam angles.

The FMO problem has been approached using a number of different mathematical frameworks including linear programming [87, 92, 186], piecewise linear–quadratic models [44, 59, 136], quadratic programming [48, 159, 181], and other nonlinear programming models [9, 97, 152, 163]. Using multi-objective optimization, many studies have generated solutions along the Pareto frontier to explore the trade-off between competing tumor and healthy-tissue objectives and constraints [51, 65, 89]. Other approaches include formulating the FMO problem as a dynamical system [74] or eigenvalue problem [52], using a database of reference dose–volume histograms to guide optimization [187], and clustering beams and voxels to handle large-scale problems [168]. We develop a new approach that is well adapted to handle nonconvex dose–volume constraints, in addition to convex maximum-dose and mean-dose constraints. We first explain all three constraint types.

Radiation dose is measured in gray (Gy), where one Gy is defined as the absorption of one joule of energy per kilogram of matter. Although lower doses are always better for healthy tissue, tolerable doses for many tissues are known empirically and depend on the risk level that a patient and physician are willing to accept. In practice, tolerable doses for different tissues are formulated using the following three types of constraints:

- Maximum dose: No voxel in the organ receives more than $d^{\max}$ Gy.
- Mean dose: Average radiation per voxel does not exceed d^{mean} Gy.
- Dose–volume: At most $p\%$ of the organ volume receives more than d^{dv} Gy.

Serial structures such as the spinal cord and brainstem lose functionality if any of their subvolumes are damaged, so maximum-dose constraints are a good indicator of tissue damage.

Maximum-dose and mean-dose constraints define a *convex* feasible region for the decision variables. The associated optimization problems can be solved at scale; for example, inverse planning with linear inequality constraints can be formulated as a convex quadratic program, with state-of-the-art approaches and commercial implementations readily available (see e.g., [61]). On the other hand, parallel structures such as the liver and lungs are composed of semi-independent units. In this case, a fraction of the parallel organ can be damaged without loss of functionality, so dose–volume constraints are appropriate [159]. Assigning maximum-dose constraints on partial volumes to achieve the desired dose–volume criteria was first introduced by Langer and Leong in 1987 [92] and brought to the attention of the mathematical community by Shepard et al. in 1999 [151].

3.1.3 Challenges of dose–volume constraints

In contrast to maximum-dose and mean-dose constraints, dose–volume constraints are *combinatorial*, since the choice of the $p\%$ of organ voxels that may receive more than d^{dv} Gy is left to the planner. In early IMRT research, using only a small number of beamlets and voxels, it was possible to solve the treatment planning problem for all possible voxel combinations satisfying a dose–volume constraint and then choose the best resulting plan [92]. This approach becomes infeasible as the problem size increases to clinically relevant situations: for an organ with n voxels and a dose–volume constraint imposed on m of these voxels, there are ${}^{n!}/_{(n-m)!m!}$ possible voxel combinations that would satisfy the constraint [159].

To illustrate the nonconvex geometry of a feasible beamlet region induced by a dose–volume constraint, we recreate an example from Wu et al. [179] using the CORT dataset described in Section 3.4. We consider three body voxels, where voxel u belongs to a prostate tumor and voxels v_1 and v_2 belong to the rectum (left panel of Figure 3.1). Using two beamlets x_1 and x_2, we aim to deliver a uniform dose of 81 Gy to the tumor while ensuring that *no more than 50% of the rectum volume (one voxel in this case) exceeds 20 Gy.*

We calculate the optimal beamlet intensities by minimizing the objective function

$$\frac{1}{2}\|A_1 x - 81\|_2^2 + \frac{\lambda}{2}\|x\|_2^2 \tag{3.1}$$

subject to the dose–volume constraint

$$\|(A_2 x - 20)_+\|_0 \leq 1. \tag{3.2}$$

The matrices $A_1 \in \mathbb{R}^{1\times 2}$ and $A_2 \in \mathbb{R}^{2\times 2}$ map beamlet intensities to prostate and rectum voxel doses, the vector $x \in \mathbb{R}^2$ contains the beamlet intensities, and the regularization term controlled by $\lambda = 5 \times 10^{-6}$ is added to stabilize beamlet solutions, amortizing the ill-conditioned beamlet-to-voxel map. The dose–volume constraint (3.2) forces the number of rectum voxels receiving a dose greater than 20 Gy to be less than or equal to 1:

$$\|(A_2 x - 20)_+\|_0 = \|\max\{0, A_2 x - 20\}\|_0 = \text{number of voxels s.t. } A_2 x > 20, \tag{3.3}$$

where the ℓ_0-norm returns the number of nonzero elements in a vector (e.g., [127]). In Figure 3.1 (right), we plot the feasible beamlet region along with the contours of the objective function. Since negative beamlet intensities are impossible, we restrict the beamlet values to the first quadrant.

There are two possibilities for satisfying the dose–volume constraint: either the first rectum voxel is allowed to exceed 20 Gy (Figure 3.1, $v_2 \leq 20$ region), or the second rectum voxel is allowed to exceed 20 Gy (Figure 3.1, $v_1 \leq 20$ region). The union of these two regions forms an L-shaped nonconvex region with two local minima at points A and B. Depending on the beamlet initialization and algorithm used to minimize the objective function, it is possible to converge to either of these two local minima, and there is no guarantee that the global minimum at B will be reached. In general, the feasible beamlet region imposed by a dose–volume constraint comprises the nonconvex union of regions satisfying maximum-dose constraints for different voxel choices [54].

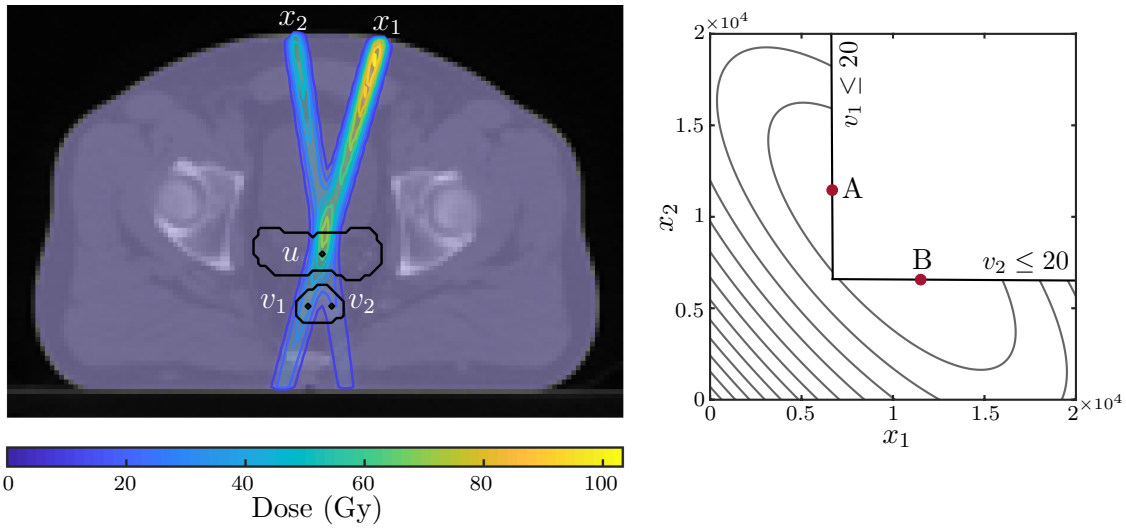

Figure 3.1: (Left) Three body voxels, one belonging to a prostate tumor (u) and two belonging to the rectum (v_1 and v_2), are irradiated by two beamlets (x_1 and x_2). The beamlet intensities shown correspond to the global minimum at point B. (Right) The nonconvex feasible region induced by the dose–volume constraint that no more than 50% of the rectum volume exceeds 20 Gy is explicitly given by the union of regions $v_1 \leq 20$ and $v_2 \leq 20$. Contours of objective function (3.1) are plotted along with a local minimum at point A and global minimum at B.

The FMO problem with nonconvex dose–volume constraints is related to the general convex-cardinality problem

$$\underset{x \in \mathcal{C}}{\text{minimize}} \quad f(x) \tag{3.4a}$$

$$\text{subject to} \quad \|x\|_0 \leq k, \tag{3.4b}$$

where $f(x)$ is a convex function and $\mathcal{C}$ is a convex set. The ℓ_0-norm constraint defines a highly nonconvex set: the *union* of all subspaces of dimension at most k. Models with these cardinality (or sparsity) constraints have become widely used within the past three decades [166], with applications such as portfolio optimization [98], signal processing [40], and compressive sensing [41]. Due to the combinatorial nature of these constraints, the related feasibility problem is NP-complete [27], and the problem itself is known to be NP-hard in general [27, 177], except for in special cases (e.g., [60]). More generally, it is an NP-hard problem just to certify whether a point of a nonlinear function is a local minimum (let alone a global one) [118]. For a more detailed overview of cardinality problems, see [18, 38, 161, 166].

Many of the standard approaches for solving the convex-cardinality problem have been applied to the FMO problem with nonconvex dose–volume constraints, most relying upon convex relaxations or approximations. One technique is to reformulate both problems as a mixed-integer programming (MIP) problem, which can be solved using branch-and-bound or branch-and-cut methods [27, 91, 93], simplex methods using big M variables [189], or by replacing the binary constraint $x \in \{0,1\}^n$ with the linear constraint $x \in [0,1]^n$ [39, 65, 116]. Another popular technique for solving the convex-cardinality problem is to relax the nonconvex ℓ_0-norm constraint (3.4b) using the convex ℓ_1-norm as either a constraint or regularizer [34, 165, 166]. For example, convex approximations to the nonconvex dose–volume constraint $\|(Ax - d^{\text{dv}})_+\|_0 \leq k$ include the convex constraint $\|(1 + \lambda(Ax - d^{\text{dv}}))_+\|_1 \leq k$ [59] and conditional value-at-risk (CVaR) constraints [134] that limit the average dose at the

tail of the dose distribution [4, 87, 136, 137]. One common dose–volume penalty includes the convex objective term $\lambda\|(Ax - d^{\mathrm{dv}})_+\|_2^2$ [32, 56, 159, 181], though the ℓ_2-norm is less effective at promoting sparsity than the ℓ_1-norm [165]. When replacing constraints with objectives, weight factors can be applied to individual vector elements, and iterative reweighting schemes can then be used to improve convergence [1, 41, 48, 65].

The complication of nonconvexity arises from the combinatorial nature of the cardinality constraint (3.4b); once a sparsity pattern for x has been chosen, the problem becomes convex. This observation is the motivation behind the multi-stage approach of 1) solving an approximation of the original problem, 2) determining a sparsity pattern for x based on the resulting solution, and 3) *polishing* the solution from Step 1 by solving a convex problem using the sparsity pattern determined in Step 2 [15, 34, 55, 161]. In the FMO problem with nonconvex dose–volume constraints, a variety of heuristics have been used to choose the $p\%$ of organ voxels that may receive more than d^{dv} Gy in Step 3. For example, Morrill et al. [114, 115] divide organ volumes *a priori* into high-dose and low-dose regions based on tumor proximity, Saberian et al. [141, 142] and Hou et al. [74] solve the FMO problem without dose–volume constraints to determine sparsity patterns, while Fu et al. [59] and Mukherjee et al. [116] use convex relaxations of the original problem to determine which voxels may receive more than d^{dv} Gy.

The great watershed in optimization is between convex and nonconvex problems [133]. When minimizing a convex objective function over a convex set, all local minima are also global. Conversely, optimizing any function over a nonconvex set leads to a nonconvex problem with potentially many local minima. Even in the small 2-D example in Figure 3.1, the dose–volume constraint creates a nonconvex feasible set where each choice of voxel corresponds to a different local minimum. As mentioned above, early IMRT treatment planning relied on stochastic methods such as simulated annealing and genetic algorithms to avoid converging to local minima, but these methods suffered from slow convergence rates [159, 175].

Fortunately, many investigations into the properties of local minima induced by nonconvex dose–volume constraints have concluded that, given a good initialization, it is possible to find local minima whose dose distributions are close to the global minimum using fast gradient methods [30, 179, 182], so the use of stochastic methods may not be warranted [9, 45, 96]. In this case, the emphasis in IMRT is often on finding a good local minimum that satisfies clinical objectives rather than finding the global minimum [45, 190].

3.1.4 Contributions

In this chapter, we develop a novel approach to handle dose–volume constraints in the radiation treatment planning problem. In contrast to prior efforts, which focused on convex approximations, our formulation is nonconvex, yet can be efficiently solved using continuous optimization techniques. We start with the idealized problem, formulated using objectives for the tumor and constraints for the OARs, and extend it to a flexible formulation that allows re-weighting strategies to balance infeasible solutions (with respect to the OARs) against delivering a sufficient amount of radiation to the tumor. To solve this formulation, we develop a customized algorithm that combines quadratic programming with fast nonlinear operators, and show that it is guaranteed to converge to a stationary point of the proposed model. The resulting approach is scalable, handles multiple objectives on tumors and OARs, and returns actionable treatment plans on patient-scale datasets.

The chapter proceeds as follows. In Section 3.2 we formulate radiation treatment planning as a nonconvex optimization problem over the fluence map. In Section 3.3 we develop a custom optimization algorithm and provide a convergence analysis. In Section 3.4 we present multiple radiation treatment use cases with data from an anonymized cancer patient dataset. Finally, in Section 3.5 we summarize the results and discuss future avenues for research.

3.2 Problem Formulation

Given a fixed set of beam angles, the dose delivered to the patient is approximated with a linear mapping from beamlet intensity to voxel dose. The FMO step is to calculate beamlet intensities that achieve the desired dose distribution to planning target volumes (PTVs) indexed by $i \in \mathcal{I}$, while respecting constraints on OARs indexed by $j \in \mathcal{J}$. Let d_i denote target delivery doses (often in Gy) to PTV $i \in \mathcal{I}$, and let $d_j^{\max}$, d_j^{mean}, and d_j^{dv} denote the maximum, mean, and dose–volume tolerance values for OAR $j \in \mathcal{J}$, where the size of dose parameters (i.e., scalar or vector) can be determined by context. The sets of OARs that have maximum-dose, mean-dose, and dose–volume constraints are denoted by $\mathcal{J}^{\max}, \mathcal{J}^{\text{mean}}$, and $\mathcal{J}^{\text{dv}}$. Let A_i and A_j be the linear beamlet-to-voxel maps for PTV i and OAR j. The vector x contains the nonnegative beamlet intensities. The idealized FMO problem with a uniform dose to the PTVs and the three constraints described in Section 3.1 on the OARs is given by

$$(\mathcal{P}_0) \underset{x \geq 0}{\text{minimize}} \quad \sum_{i \in \mathcal{I}} \frac{\alpha_i}{2n_i} \|A_i x - d_i\|_2^2 + \frac{\lambda}{2} \|x\|_2^2 \tag{3.5a}$$

$$\text{subject to} \quad A_j x \leq d_j^{\max}, \qquad j \in \mathcal{J}^{\max}, \tag{3.5b}$$

$$1^T A_j x \leq n_j d_j^{\text{mean}}, \qquad j \in \mathcal{J}^{\text{mean}}, \tag{3.5c}$$

$$\|(A_j x - d_j^{\text{dv}})_+\|_0 \leq \frac{n_j p_j}{100}, \quad j \in \mathcal{J}^{\text{dv}}, \tag{3.5d}$$

where n_i and n_j indicate the number of voxels in the ith PTV and jth OAR, and the percentage of voxels in the jth OAR that may receive a dose exceeding d_j^{dv} is specified by p_j. The least-squares term for the uniform target dose is widely used [32, 48, 56, 159, 181]. In principle, Problem $(\mathcal{P}_0)$ can be replaced by a feasibility problem with lower and upper dose–volume constraints also placed on the PTV [44, 111, 127]. Here we focus on $(\mathcal{P}_0)$, but the techniques we propose extend to the feasibility problem.

Problem $(\mathcal{P}_0)$ is difficult for several reasons. First, it is high-dimensional (beamlets

can be on the order of $10^3 - 10^5$) [179]. Second, beamlet-to-voxel maps are generally ill-conditioned [3, 52, 152, 175]. Third, dose–volume constraints are combinatorial in nature, as constraint (3.5d) requires choosing a set of voxels allowed to exceed d_j^{dv}, so finding the global solution is NP-hard [167]. Finally, while Problem $(\mathcal{P}_0)$ is always feasible (i.e., $x = 0$ is a solution), it may not yield a clinically useful radiation distribution. Feasibility-type reformulations of $(\mathcal{P}_0)$ often fail to have any meaningful solutions because the target doses almost always compete with OAR doses, as it is common for tumors to be close to or within OARs. Practical radiation treatment planning converts constraints into objectives, and then looks for the best trade-off among competing objectives.

In the next section, we formulate the idealized problem using constraints, and then develop a relaxation that always has a feasible solution. This allows significant modeling and computational flexibility, yielding efficiently computable and clinically useful solutions that balance the need to deliver enough radiation to the tumor while approximately satisfying OAR constraints.

3.2.1 Nonconvex relaxation

To simplify the exposition, we explain the new formulation using only dose–volume constraints on the OARs:

$$(\mathcal{P}_1) \underset{x \geq 0}{\text{minimize}} \quad \sum_{i \in \mathcal{I}} \frac{\alpha_i}{2n_i} \|A_i x - d_i\|_2^2 + \frac{\lambda}{2} \|x\|_2^2 \tag{3.6a}$$

$$\text{subject to} \quad \|(A_j x - d_j^{\mathrm{dv}})_+\|_0 \leq \frac{n_j p_j}{100}, \quad j \in \mathcal{J}^{\mathrm{dv}}. \tag{3.6b}$$

Solutions to $(\mathcal{P}_1)$ may not be clinically useful if the dose–volume constraints prevent sufficient dose delivered to the tumor. We develop an extended, nonconvex formulation that relaxes the dose–volume constraints. The new formulation includes additional variables $w_j \approx A_j x - d_j^{\mathrm{dv}}$,

and penalizes deviations of the delivered OAR doses from w_j:

$$(\mathcal{P}_2)\ \underset{x \geq 0,\, w}{\text{minimize}} \quad f(x,w) := \sum_{i \in \mathcal{I}} \frac{\alpha_i}{2n_i} \|A_i x - d_i\|_2^2 + \sum_{j \in \mathcal{J}^{\text{dv}}} \frac{\alpha_j}{2n_j} \|w_j - (A_j x - d_j^{\text{dv}})\|_2^2 + \frac{\lambda}{2}\|x\|_2^2 \tag{3.7a}$$

$$\text{subject to} \quad \|(w_j)_+\|_0 \leq \frac{n_j p_j}{100}, \quad j \in \mathcal{J}^{\text{dv}}. \tag{3.7b}$$

In formulation $(\mathcal{P}_2)$, the auxiliary variables w_j are forced to satisfy the dose–volume constraints, while the residuals $A_j x - d_j^{\text{dv}}$ may not necessarily do so. The weights α_j control how closely w_j must approximate $A_j x - d_j^{\text{dv}}$, and as $\alpha_j \uparrow \infty$, problem $(\mathcal{P}_2)$ converges to the idealized problem $(\mathcal{P}_1)$ in an epigraphical sense.

The epigraph of a function is defined by

$$\text{epi}(f) = \{(x, \beta) \,:\, f(x) \leq \beta\}. \tag{3.8}$$

Epigraphs make it possible to go between sets and functions, and the notion of epigraphical convergence [135] is exactly that of set convergence for epigraphs, that is,

$$f^\alpha \to_{\text{epi}} f \quad \text{exactly when} \quad \text{epi}(f^\alpha) \to \text{epi}(f). \tag{3.9}$$

We know that our objective functions (3.7a) converge epigraphically to the original problem by [135, Theorem 7.4(d)] because of monotonicity, i.e.,

$$f^{\alpha_j} \leq f^{\alpha_{j+1}} \quad \text{when} \quad \alpha_j \leq \alpha_{j+1}, \tag{3.10}$$

which we always have when using quadratic penalties for prox-bounded functions (see discussion below Theorem 7.4(d)).

Formulation $(\mathcal{P}_2)$ has several desirable features. First, while most of the heuristic approaches described in Section 3.1.3 rely on convex approximations, $(\mathcal{P}_2)$ fully captures the nonconvex structure of the dose–volume constraints, while giving the modeler flexibility to

match them approximately using weights α_j. Specifically, the weights α_i and α_j can be tuned to balance the trade-off between covering the tumor and approximately meeting OAR constraints. Second, from a computational perspective, we can design efficient and provably convergent algorithms that aggregate information from multiple tumors and OARs using partial minimization, as described in the next section. Finally, approximate solutions can be computed efficiently, then refined using techniques such as iterative re-weighting schemes or polishing.

3.3 Algorithmic Approach

Our main strategy is to *partially minimize* formulation $(\mathcal{P}_2)$ with respect to x, viewing $(\mathcal{P}_2)$ as a *value-function optimization* problem

$$(\mathcal{P}_3)\ \underset{w}{\text{minimize}} \quad g(w) := \min_{x \geq 0} \sum_{i \in \mathcal{I}} \frac{\alpha_i}{2n_i} \|A_i x - d_i\|_2^2 + \sum_{j \in \mathcal{J}^{\text{dv}}} \frac{\alpha_j}{2n_j} \|w_j - (A_j x - d_j^{\text{dv}})\|_2^2 + \frac{\lambda}{2}\|x\|_2^2 \tag{3.11a}$$

$$\text{subject to} \quad w_j \in \Omega_j := \left\{ w_j \in \mathbb{R}^{n_j} : \|(w_j)_+\|_0 \leq \frac{n_j p_j}{100} \right\}, \quad j \in \mathcal{J}^{\text{dv}}. \tag{3.11b}$$

The algorithm is centered on optimizing the value function $g(w)$ in (3.11a) using a projected gradient approach to take care of the nonconvex constraints encoded by Ω_j. The high-level iteration is simply

$$w^{(k+1)} \in \text{proj}_\Omega \left[w^{(k)} - T \nabla g \left(w^{(k)} \right) \right], \tag{3.12}$$

where $w = (w_{j_1}, w_{j_2}, \dots)$, $T = \text{diag}(t_{j_1}, \dots, t_{j_2}, \dots)$, and $\Omega = \Omega_{j_1} \times \Omega_{j_2} \times \dots$ for $j_i \in \mathcal{J}^{\text{dv}}$. Evaluation of g and ∇g is done by computing the partial minimum over x, as detailed in Algorithm 2. Projection onto Ω_j can be implemented efficiently by ordering the elements of w_j and setting the lowest $\lceil (100 - p_j) n_j / 100 \rceil$ entries of w_j to $\min\{0, w_j\}$ [127].

The differentiability of $g(w)$, formula for the derivative, and its Lipschitz constant follow

Algorithm 2 Projected gradient descent for $g(w)$ in $(\mathcal{P}_3)$.

Input ϵ, $t_j \leq {}^{n_j}\!/_{\alpha_j}$, $x^{(0)}$. Initialize $k = 0$, $\delta > \epsilon$, $w_j^{(0)} = \mathrm{proj}_{\Omega_j}\left(A_j x^{(0)} - d_j^{\mathrm{dv}}\right)$.

while $\delta > \epsilon$ **do**

$x^{(k+1)} \leftarrow \arg\min_{x \geq 0} f\left(x, w^{(k)}\right)$

for $j \in \mathcal{J}^{\mathrm{dv}}$ **do**

$w_j^{(k+1)} \leftarrow \mathrm{proj}_{\Omega_j}\left\{w_j^{(k)} - \frac{\alpha_j t_j}{n_j}\left[w_j^{(k)} - \left(A_j x^{(k+1)} - d_j^{\mathrm{dv}}\right)\right]\right\}$

end for

$\delta \leftarrow \sum_{j \in \mathcal{J}^{\mathrm{dv}}} \frac{1}{t_j}\left\|w_j^{(k+1)} - w_j^{(k)}\right\|_2$

$k \leftarrow k + 1$

end while

return $\arg\min_{x \geq 0} f\left(x, w^{(k)}\right)$

from [192, Theorem 1]. In particular, for one PTV and one OAR, define

$$g(w) = \min_{x \geq 0} \frac{\alpha_1}{2n_1}\|A_1 x - d_1\|_2^2 + \frac{\alpha_2}{2n_2}\|w - (A_2 x - d_2^{\mathrm{dv}})\|_2^2 + \frac{\lambda}{2}\|x\|_2^2. \tag{3.13}$$

Next, taking $h(x) = \delta_+(x) + {}^1\!/_2\|A_1 x - d_1\|_2^2 + {}^\lambda\!/_2\|x\|_2^2$, where $\delta_+(x)$ is the indicator function of the nonnegative orthant, we have

$$g(w) = \min_x h(x) + \frac{\alpha_2}{2n_2}\|w - (A_2 x - d_2^{\mathrm{dv}})\|_2^2, \tag{3.14}$$

$$x(w) \in \arg\min_x h(x) + \frac{\alpha_2}{2n_2}\|w - (A_2 x - d_2^{\mathrm{dv}})\|_2^2, \tag{3.15}$$

and from [192, Theorem 1], we get

$$\nabla g(w) = \frac{\alpha_2}{n_2}\left\{w - \left[A_2 x(w) - d_2^{\mathrm{dv}}\right]\right\} \quad \text{and} \quad \mathrm{lip}(\nabla g) \leq \frac{\alpha_2}{n_2}. \tag{3.16}$$

Similarly, for multiple OARs we have ${}^{\partial g}\!/_{\partial w_j}(w) = {}^{\alpha_j}\!/_{n_j}\left\{w_j - \left[A_j x(w) - d_j^{\mathrm{dv}}\right]\right\}$ and

$\operatorname{lip}(\partial g/\partial w_j) \leq \alpha_j/n_j$ for all $j \in \mathcal{J}^{\mathrm{dv}}$, and for all $w, w' \in \mathbb{R}^{\sum_{j\in\mathcal{J}^{\mathrm{dv}}} n_j}$ we have

$$\|\nabla g(w) - \nabla g(w')\|_2^2 = \sum_{j\in\mathcal{J}^{\mathrm{dv}}} \left\| \tfrac{\partial g}{\partial w_j}(w) - \tfrac{\partial g}{\partial w_j}(w') \right\|_2^2, \tag{3.17}$$

$$\leq \sum_{j\in\mathcal{J}^{\mathrm{dv}}} \left(\frac{\alpha_j}{n_j}\right)^2 \left\|w_j - w_j'\right\|_2^2, \tag{3.18}$$

$$\leq \max_{j\in\mathcal{J}^{\mathrm{dv}}} \left(\frac{\alpha_j}{n_j}\right)^2 \sum_{j\in\mathcal{J}^{\mathrm{dv}}} \left\|w_j - w_j'\right\|_2^2, \tag{3.19}$$

$$= \max_{j\in\mathcal{J}^{\mathrm{dv}}} \left(\frac{\alpha_j}{n_j}\right)^2 \|w - w'\|_2^2, \tag{3.20}$$

so $\operatorname{lip}(\nabla g) \leq \max_{j\in\mathcal{J}^{\mathrm{dv}}} (\alpha_j/n_j)$.

Algorithm 2 is projected gradient descent for the value function (3.11a), which is equivalent to the simple iteration (3.12). Under fairy weak assumptions[1], the nonconvex gradient projection method converges for a fixed step size $t < 1/\operatorname{lip}(\nabla g)$ [14, Theorem 5.3]. On the other hand, Algorithm 2 allows the flexibility to choose different step sizes $t_j \leq n_j/\alpha_j$ for each $j \in \mathcal{J}^{\mathrm{dv}}$. We prove the convergence of this variant of projected gradient descent, starting with some definitions to quantify stationarity and optimality.

Definition 1 (Stationary Point)**.** *A vector $w \in \Omega$ is called a stationary point for $(\mathcal{P}_3)$ if*

$$0 \in \nabla g(w) + N_\Omega(w), \tag{3.21}$$

where N_Ω is the normal cone to Ω at w.

This definition follows from the basic first-order conditions for optimality [135, Theorem 6.12], where (3.21) is a necessary condition for w to be a local minimizer of g on Ω. This definition of stationarity motivates the following measure of optimality.

---

[1]Specifically, the sequence $\{w^{(k)}\}_{k\in\mathbb{N}}$ generated by (3.12) must be bounded, and $g + \delta_\Omega$ must be a Kurdyka-Łojasiewicz function (which covers a wide range of functions, including the class of real semi-algebraic functions).

Definition 2 (ϵ-Accurate Solution)**.** *Given fixed step sizes $t_j > 0$ for all $j \in \mathcal{J}^{dv}$, we call a vector $w \in \Omega$ an ϵ-accurate solution for $(\mathcal{P}_3)$ if*

$$\sum_{j \in \mathcal{J}^{dv}} \frac{1}{t_j^2} \|w_j - \bar{w}_j\|_2^2 \leq \epsilon, \tag{3.22}$$

where $\bar{w}_j \in proj_{\Omega_j}\left[w_j - t_j {}^{\partial g}/_{\partial w_j}(w)\right]$.

By the first-order optimality conditions of the problem defining the projection in (3.12), for each $j \in \mathcal{J}^{\mathrm{dv}}$ the vectors w_j and $\bar{w}_j$ satisfy

$$\frac{1}{t_j}\left(w_j - \bar{w}_j\right) \in \frac{\partial g}{\partial w_j}(w) + N_{\Omega_j}\left(\bar{w}_j\right), \tag{3.23}$$

so if $\sum_{j \in \mathcal{J}^{\mathrm{dv}}} {}^{1}/_{t_j^2} \|w_j - \bar{w}_j\|_2^2 = 0$, then $0 \in {}^{\partial g}/_{\partial w_j}(w) + N_{\Omega_j}(w_j)$ for all $j \in \mathcal{J}^{\mathrm{dv}}$. Furthermore, by [135, Proposition 6.41] we have $0 \in \nabla g(w) + N_{\Omega}(w)$, which means that w is a stationary point for problem $(\mathcal{P}_3)$.

We now establish the convergence of Algorithm 2, which follows from several auxiliary lemmas. For simplicity, we let m be the number of beamlets and $N := \sum_{j \in \mathcal{J}^{\mathrm{dv}}} n_j$ be the number of dose–volume constrained voxels, so that $x \in \mathbb{R}^m$ and $w \in \mathbb{R}^N$. We begin by proving that all fixed points of Algorithm 2 are local minimizers of g on Ω via the strong convexity of objectives f and g.

Lemma 1. *The objective $f(x, w)$ is strongly convex on $\mathbb{R}^m \times \mathbb{R}^N$.*

Proof. For simplicity, we consider the case with one PTV and one OAR, but the arguments used generalize to include additional terms. Rearranging variables, the objective $f(x, w)$ can

be expressed using a single quadratic term,

$$f(x,w) = \frac{\alpha_1}{2n_1}\left\|A_1x - d_1\right\|_2^2 + \frac{\alpha_2}{2n_2}\left\|w - \left[A_2x - d_2^{\mathrm{dv}}\right]\right\|_2^2 + \frac{\lambda}{2}\|x\|_2^2, \tag{3.24}$$

$$\begin{aligned} &= \frac{1}{2}\left\|\begin{bmatrix}\sqrt{\frac{\alpha_1}{n_1}}A_1 & 0\end{bmatrix}\begin{bmatrix}x\\ w\end{bmatrix} - \sqrt{\tfrac{\alpha_1}{n_1}}d_1\right\|_2^2 \\ &\quad + \frac{1}{2}\left\|\begin{bmatrix}\sqrt{\frac{\alpha_2}{n_2}}A_2 & -\sqrt{\frac{\alpha_2}{n_2}}I\end{bmatrix}\begin{bmatrix}x\\ w\end{bmatrix} - \sqrt{\tfrac{\alpha_2}{n_2}}d_2^{\mathrm{dv}}\right\|_2^2 \\ &\quad + \frac{1}{2}\left\|\begin{bmatrix}\sqrt{\lambda}I & 0\end{bmatrix}\begin{bmatrix}x\\ w\end{bmatrix}\right\|_2^2, \end{aligned} \tag{3.25}$$

$$= \frac{1}{2}\left\|\begin{bmatrix}\sqrt{\frac{\alpha_1}{n_1}}A_1 & 0\\ \sqrt{\frac{\alpha_2}{n_2}}A_2 & -\sqrt{\frac{\alpha_2}{n_2}}I\\ \sqrt{\lambda}I & 0\end{bmatrix}\begin{bmatrix}x\\ w\end{bmatrix} - \begin{bmatrix}\sqrt{\frac{\alpha_1}{n_1}}d_1\\ \sqrt{\frac{\alpha_2}{n_2}}d_2^{\mathrm{dv}}\\ 0\end{bmatrix}\right\|_2^2, \tag{3.26}$$

$$= \frac{1}{2}\left\|\tilde{A}\tilde{x} - \tilde{d}\right\|_2^2, \tag{3.27}$$

with symmetric Hessian matrix

$$\nabla^2 f(\tilde{x}) = \tilde{A}^T\tilde{A} = \begin{bmatrix}\frac{\alpha_1}{n_1}A_1^TA_1 + \frac{\alpha_2}{n_2}A_2^TA_2 + \lambda I & -\frac{\alpha_2}{n_2}A_2^T\\ -\frac{\alpha_2}{n_2}A_2 & \frac{\alpha_2}{n_2}I\end{bmatrix}. \tag{3.28}$$

Consider the matrix $(\alpha_2/n_2)\,I$ and the Schur complement of $\tilde{A}^T\tilde{A}$ in $(\alpha_2/n_2)\,I$,

$$\tilde{A}^T\tilde{A}\,/\left(\frac{\alpha_2}{n_2}I\right) = \frac{\alpha_1}{n_1}A_1^TA_1 + \frac{\alpha_2}{n_2}A_2^TA_2 + \lambda I - \left(-\frac{\alpha_2}{n_2}A_2^T\right)\left(\frac{\alpha_2}{n_2}I\right)\left(-\frac{\alpha_2}{n_2}A_2\right), \tag{3.29}$$

$$= \frac{\alpha_1}{n_1}A_1^TA_1 + \frac{\alpha_2}{n_2}A_2^TA_2 + \lambda I - \frac{\alpha_2}{n_2}A_2^TA_2, \tag{3.30}$$

$$= \frac{\alpha_1}{n_1}A_1^TA_1 + \lambda I. \tag{3.31}$$

Because the coefficients α_1/n_1, α_2/n_2, and λ are positive, the matrices $(\alpha_2/n_2)\,I$ and $\tilde{A}^T\tilde{A}\,/\,[(\alpha_2/n_2)\,I]$ are positive definite (note that this does not depend on the rank of matrices A_1 or A_2), which

implies that the Hessian matrix $\tilde{A}^T\tilde{A}$ is also positive definite [34]. Therefore the objective $f(x,w)$ with one PTV and one OAR is strongly convex on $\mathbb{R}^m \times \mathbb{R}^{n_2}$.

For multiple PTVs and OARs, the lower-right block of the Hessian matrix becomes a block diagonal matrix with matrices $({}^{\alpha_j}/_{n_j})\, I$ for $j \in \mathcal{J}^{\text{dv}}$ along the diagonal, and the Schur complement becomes the matrix $\sum_{i\in\mathcal{I}} A_i^T A_i + \lambda I$. Because the coefficients ${}^{\alpha_i}/_{n_i}$ for $i \in \mathcal{I}$, ${}^{\alpha_j}/_{n_j}$ for $j \in \mathcal{J}^{\text{dv}}$, and λ are positive, the arguments above can again be used to show that the objective $f(x,w)$ with multiple PTVs and OARs is strongly convex on $\mathbb{R}^m \times \mathbb{R}^N$. □

Lemma 2. *The objective $g(w)$ is strongly convex on $\mathbb{R}^N$.*

Proof. Because the objective $f(x,w)$ is strongly convex on $\mathbb{R}^m \times \mathbb{R}^N$ [Lemma 1], there exists a constant $\mu > 0$ such that for all vector pairs $(x,w),(x',w') \in \mathbb{R}^m \times \mathbb{R}^N$ and constants $t \in [0,1]$, we have

$$\begin{aligned} f\left[tx + (1-t)x', tw + (1-t)w'\right] \leq\ & tf(x,w) + (1-t)f\left(x',w'\right) \\ & - \frac{\mu}{2}t(1-t)\left(\|x-x'\|_2^2 + \|w-w'\|_2^2\right). \end{aligned} \tag{3.32}$$

Let $x(w) \in \arg\min_{x\geq 0} f(x,w)$. Since $f(x,w)$ is a strongly convex function and the nonnegative orthant is a nonempty closed convex set, the vector $x(w) \in \mathbb{R}^m_{\geq 0}$ exists and is uniquely defined for all vectors $w \in \mathbb{R}^N$ [34], and $g(w) = \min_{x\geq 0} f(x,w) = f(x(w),w)$. Furthermore, for all vectors $w, w' \in \mathbb{R}^N$ and constants $t \in [0,1]$, we have

$$g\left[tw + (1-t)w'\right] = \min_{x\geq 0} f\left[x, tw + (1-t)w'\right], \tag{3.33}$$

$$\leq f\left[tx(w) + (1-t)x\left(w'\right), tw + (1-t)w'\right], \tag{3.34}$$

$$\begin{aligned} \leq\ & tf\left[x(w),w\right] + (1-t)f\left[x\left(w'\right),w'\right] \\ & - \frac{\mu}{2}t(1-t)\left(\|x(w) - x\left(w'\right)\|_2^2 + \|w-w'\|_2^2\right), \end{aligned} \tag{3.35}$$

$$\begin{aligned} =\ & tg(w) + (1-t)g\left(w'\right) \\ & - \frac{\mu}{2}t(1-t)\left(\|x(w) - x\left(w'\right)\|_2^2 + \|w-w'\|_2^2\right), \end{aligned} \tag{3.36}$$

$$\leq tg(w) + (1-t)g\left(w'\right) - \frac{\mu}{2}t(1-t)\|w-w'\|_2^2, \tag{3.37}$$

where line (3.34) follows from the fact that the nonnegative orthant is a convex set and $x(w) \geq 0$ for all vectors $w \in \mathbb{R}^N$, so $tx(w) + (1-t)x(w') \geq 0$ for all vectors $w, w' \in \mathbb{R}^N$ and constants $t \in [0,1]$. Therefore $g(w)$ is strongly convex on $\mathbb{R}^N$. □

Theorem 1. *The fixed points of Algorithm 2 are local minimizers of g on Ω.*

Proof. Fixed points $w^* = (w^*_{j_1}, w^*_{j_2}, \dots)$ for $j_i \in \mathcal{J}^{\mathrm{dv}}$ of Algorithm 2 satisfy

$$w^*_j \in \mathrm{proj}_{\Omega_j}\left[w^*_j - t_j \frac{\partial g}{\partial w_j}(w^*)\right]. \tag{3.38}$$

Based on the form of the projection onto Ω_j, for each element i of a stationary vector w^*_j, we know that

1. if $(w^*_j)_i < 0$, then $[\partial g/\partial w_j\,(w^*)]_i = (w^*_j)_i - (A_j x(w^*) - d^{\mathrm{dv}}_j)_i = 0$, and the dose to voxel i is less than d^{dv}_j,

2. if $(w^*_j)_i > 0$, then $[\partial g/\partial w_j(w^*)]_i = 0$, $(w^*_j)_i$ is one of the $\lfloor n_j p_j/100 \rfloor$ largest positive elements of vector w^*_j, and the dose to voxel i is greater than d^{dv}_j, and

3. if $(w^*_j)_i = 0$, then either $[\partial g/\partial w_j(w^*)]_i = 0$, and the dose to voxel i is exactly d^{dv}_j, or $-1/t_j[\partial g/\partial w_j(w^*)]_i > 0$ is less than the $\lfloor n_j p_j/100 \rfloor$ largest positive elements of w^*_j, and the dose to voxel i is greater than d^{dv}_j (i.e., voxel i violates the dose–volume constraint on OAR j).

Recall that for each $j \in \mathcal{J}^{\mathrm{dv}}$, the set $\Omega_j = \{w \in \mathbb{R}^{n_j} : \|w_+\|_0 \leq n_j p_j/100\}$ is the union of finitely many convex sets, each representing a different combination of the $q_j := \lfloor n_j p_j/100 \rfloor$ elements of the vector w_j allowed to be positive. Specifically, for $\ell = 1, \dots, n_j!/(n_j - q_j)!q_j!$, let M^ℓ_j be the $n_j \times n_j$ identity matrix with q_j of its diagonal elements set to zero. Then the convex set $\Omega^\ell_j := \left\{w \in \mathbb{R}^{n_j} : M^\ell_j w \leq 0\right\}$ represents the set of vectors that satisfy element combination ℓ, and $\Omega_j = \cup_{\ell=1,\dots,n_j!/(n_j-q_j)!q_j!}\Omega^\ell_j$ is the union of all such convex sets. Consequently, the set $\Omega = \prod_{j \in \mathcal{J}^{\mathrm{dv}}} \Omega_j$, the Cartesian product of the feasible sets for vectors w_j, is

also the union of finitely many convex sets, each representing a different combination of the $Q := \sum_{j \in \mathcal{J}^{\mathrm{dv}}} q_j$ elements of w allowed to be positive. Because the objective g is strongly convex on any convex subset of its domain $\mathbb{R}^N$ [Lemma 2], there is a unique global minimum of g relative to any of these convex subsets of Ω [34], each corresponding to a local minimum of g on Ω. As a result, there are finitely many local minimizers of g on Ω.

Suppose w^* is a fixed point of Algorithm 2. If $w^* \leq 0$, then the vector satisfies all feasible combinations of the Q elements of w allowed to be positive, and thus lies within the intersection of all corresponding convex subsets of Ω. Furthermore, from observations 1 and 3 above, we have $\nabla g\left(w^*\right) = 0$, and using the strong convexity of g we can conclude that w^* is the unique global minimimizer of g on Ω [135, Theorem 6.12]. On the other hand, suppose that at least one element of w^* is positive, and let M^* be the $\mathbb{R}^N \times \mathbb{R}^N$ identity matrix where the diagonal elements corresponding to the at most Q positive elements of w^* are set to zero. Then $\Omega^* := \left\{w \in \mathbb{R}^N \,:\, M^* w \leq 0\right\}$ is a convex subset of Ω containing w^*. Again, from observations 1–3 above, we have $\left[\nabla g\left(w^*\right)\right]_i = 0$ where $\left(w^*\right)_i \neq 0$ and $\left[\nabla g\left(w^*\right)\right]_i \leq 0$ where $\left(w^*\right)_i = 0$, so for all $w \in \Omega^*$ we have

$$-\nabla g\left(w^*\right)^T \left(w - w^*\right) = -\nabla g\left(w^*\right)^T w \leq 0, \tag{3.39}$$

because positive elements of w correspond to zero elements of the gradient. Therefore $-\nabla g\left(w^*\right) \in N_{\Omega^*}\left(w^*\right)$ [135, Theorem 6.9], and by the convexity of g and Ω^*, we know that w^* is the unique global minimizer of g on Ω^* [135, Theorem 6.12], and hence a local minimizer of g on Ω. □

We now establish the convergence of Algorithm 2 to a local minimum of g on Ω.

Lemma 3. *For all vectors $w, w' \in \mathbb{R}^N$, we have*

$$g(w) \leq g\left(w'\right) + \nabla g\left(w'\right)^T \left(w - w'\right) + \sum_{j \in \mathcal{J}^{dv}} \frac{\alpha_j}{2n_j} \left\|w - w'\right\|_2^2. \tag{3.40}$$

Proof. Following the proof from [120, Lemma 1.2.3], for all vectors $w, w' \in \mathbb{R}^N$, we have

$$g(w) = g\left(w'\right) + \int_0^1 \nabla g\left[w' + \tau\left(w - w'\right)\right]^T \left(w - w'\right) d\tau, \tag{3.41}$$

$$= g\left(w'\right) + \nabla g\left(w'\right)^T \left(w - w'\right) + \int_0^1 \left\{\nabla g\left[w' + \tau\left(w - w'\right)\right] - \nabla g\left(w'\right)\right\}^T \left(w - w'\right) d\tau. \tag{3.42}$$

Recall that for OARs $j \in \mathcal{J}^{\mathrm{dv}}$, we have $\operatorname{lip}\left({}^{\partial g}/_{\partial w_j}\right) \leq \frac{\alpha_j}{n_j}$. Therefore

$$\left|g(w) - g\left(w'\right) - \nabla g\left(w'\right)^T \left(w - w'\right)\right| \tag{3.43}$$

$$= \left|\int_0^1 \left\{\nabla g\left[w' + \tau\left(w - w'\right)\right] - \nabla g\left(w'\right)\right\}^T \left(w - w'\right) d\tau\right|, \tag{3.44}$$

$$\leq \int_0^1 \left|\left\{\nabla g\left[w' + \tau\left(w - w'\right)\right] - \nabla g\left(w'\right)\right\}^T \left(w - w'\right)\right| d\tau, \tag{3.45}$$

$$= \int_0^1 \left|\sum_{j \in \mathcal{J}^{\mathrm{dv}}} \left\{\frac{\partial g}{\partial w_j}\left[w_j' + \tau\left(w_j - w_j'\right)\right] - \frac{\partial g}{\partial w_j}\left(w_j'\right)\right\}^T \left(w_j - w_j'\right)\right| d\tau, \tag{3.46}$$

$$\leq \int_0^1 \sum_{j \in \mathcal{J}^{\mathrm{dv}}} \left|\left\{\frac{\partial g}{\partial w_j}\left[w_j' + \tau\left(w_j - w_j'\right)\right] - \frac{\partial g}{\partial w_j}\left(w_j'\right)\right\}^T \left(w_j - w_j'\right)\right| d\tau, \tag{3.47}$$

$$\leq \int_0^1 \sum_{j \in \mathcal{J}^{\mathrm{dv}}} \left\|\frac{\partial g}{\partial w_j}\left[w_j' + \tau\left(w_j - w_j'\right)\right] - \frac{\partial g}{\partial w_j}\left(w_j'\right)\right\|_2 \left\|w_j - w_j'\right\|_2 d\tau, \tag{3.48}$$

$$\leq \int_0^1 \sum_{j \in \mathcal{J}^{\mathrm{dv}}} \frac{\tau \alpha_j}{n_j} \left\|w_j - w_j'\right\|_2^2 d\tau, \tag{3.49}$$

$$= \sum_{j \in \mathcal{J}^{\mathrm{dv}}} \frac{\alpha_j}{2n_j} \left\|w_j - w_j'\right\|_2^2, \tag{3.50}$$

which implies that

$$g(w) \leq g\left(w'\right) + \nabla g\left(w'\right)^T \left(w - w'\right) + \sum_{j \in \mathcal{J}^{\mathrm{dv}}} \frac{\alpha_j}{2n_j} \left\|w_j - w_j'\right\|_2^2 \tag{3.51}$$

for all $w, w' \in \mathbb{R}^N$. □

Lemma 4. *For fixed step sizes $t_j \in (0, {}^{n_j}/_{\alpha_j}]$ for all $j \in \mathcal{J}^{dv}$, the sequence of objective values $\left\{g\left(w^{(k)}\right)\right\}_{k \in \mathbb{N}}$ generated by Algorithm 2 is nonincreasing.*

Proof. For each $j \in \mathcal{J}^{\mathrm{dv}}$, let vectors $w_j^{(k+1)}, w_j^{(k)} \in \Omega_j$ satisfy $w_j^{(k+1)} \in \mathrm{proj}_{\Omega_j}\left[w_j^{(k)} - t_j {}^{\partial g}/_{\partial w_j}\left(w^{(k)}\right)\right]$. Then from the definition of the projection operator,

$$\left\| w_j^{(k+1)} - \left[w_j^{(k)} - t_j \frac{\partial g}{\partial w_j}\left(w^{(k)}\right)\right]\right\|_2^2 = \min_{w_j \in \Omega_j} \left\| w_j - \left[w_j^{(k)} - t_j \frac{\partial g}{\partial w_j} g\left(w^{(k)}\right)\right]\right\|_2^2, \tag{3.52}$$

$$\leq \left\| w_j^{(k)} - \left[w_j^{(k)} - t_j \frac{\partial g}{\partial w_j} g\left(w^{(k)}\right)\right]\right\|_2^2, \tag{3.53}$$

$$= \left\| t_j \frac{\partial g}{\partial w_j}\left(w^{(k)}\right)\right\|_2^2, \tag{3.54}$$

which simplifies to

$$\frac{\partial g}{\partial w_j}\left(w^{(k)}\right)^T \left(w_j^{(k+1)} - w_j^{(k)}\right) \leq -\frac{1}{2t_j}\left\| w_j^{(k+1)} - w_j^{(k)}\right\|_2^2. \tag{3.55}$$

Noting that $\Omega \subset \mathbb{R}^N$, we can combine inequality (3.55) with Lemma 3 to get

$$g\left(w^{(k+1)}\right) \leq g\left(w^{(k)}\right) + \nabla g\left(w^{(k)}\right)^T \left(w^{(k+1)} - w^{(k)}\right) + \sum_{j \in \mathcal{J}^{\mathrm{dv}}} \frac{\alpha_j}{2n_j}\left\| w^{(k+1)} - w^{(k)}\right\|_2^2, \tag{3.56}$$

$$= g\left(w^{(k)}\right) + \sum_{j \in \mathcal{J}^{\mathrm{dv}}} \frac{\partial g}{\partial w_j}\left(w^{(k)}\right)^T \left(w_j^{(k+1)} - w_j^{(k)}\right) + \sum_{j \in \mathcal{J}^{\mathrm{dv}}} \frac{\alpha_j}{2n_j}\left\| w^{(k+1)} - w^{(k)}\right\|_2^2, \tag{3.57}$$

$$\leq g\left(w^{(k)}\right) - \sum_{j \in \mathcal{J}^{\mathrm{dv}}} \frac{1}{2t_j}\left\| w_j^{(k+1)} - w_j^{(k)}\right\|_2^2 + \sum_{j \in \mathcal{J}^{\mathrm{dv}}} \frac{\alpha_j}{2n_j}\left\| w_j^{(k+1)} - w_j^{(k)}\right\|_2^2, \tag{3.58}$$

$$= g\left(w^{(k)}\right) + \sum_{j \in \mathcal{J}^{\mathrm{dv}}} \left(\frac{\alpha_j}{2n_j} - \frac{1}{2t_j}\right)\left\| w_j^{(k+1)} - w_j^{(k)}\right\|_2^2. \tag{3.59}$$

Therefore we have $g\left(w^{(k+1)}\right) \leq g\left(w^{(k)}\right)$ for all vectors $w_j^{(k+1)}, w_j^{(k)}$ satisfying $w_j^{(k+1)} \in \mathrm{proj}_{\Omega_j}\left[w_j^{(k)} - t_j \frac{\partial g}{\partial w_j}\left(w^{(k)}\right)\right]$ with $t \in (0, {}^{n_j}/_{\alpha_j}]$ for all $j \in \mathcal{J}^{\mathrm{dv}}$. □

Theorem 2. *For fixed step sizes $t_j \in (0, {}^{n_j}/_{\alpha_j}]$ for all $j \in \mathcal{J}^{dv}$, Algorithm 2 generates a sequence of iterates $\left\{w^{(k)}\right\}_{k \in \mathbb{N}}$ such that $g\left(w^{(k)}\right) \downarrow g^* \geq 0$; in particular, $g^* \leq g\left(w^{(k)}\right)$ for all $k \in \mathbb{N}$. Furthermore, for any convergent subsequence $w^{(p_k)} \to w_p^*$, the vector w_p^* is a local minimizer of g on Ω, and $g\left(w_p^*\right) = g^*$.*

Proof. From Lemma 4, we know that if $t_j \in (0, {}^{n_j}/_{\alpha_j}]$ for all $j \in \mathcal{J}^{\mathrm{dv}}$, then the sequence of objective values $\left\{g\left(w^{(k)}\right)\right\}_{k\in\mathbb{N}}$ is nonincreasing, and $g(w) \geq 0$ for all $w \in \mathbb{R}^N$ by definition, so the sequence converges to some constant $g^* \geq 0$ where $g^* \leq g\left(w^{(k)}\right)$ for all $k \in \mathbb{N}$. Furthermore, for all convergent subsequences $w^{(p_k)} \to w_p^*$ as $k \to \infty$, we have $g\left(w_p^*\right) = g^*$. Because w_p^* is a fixed point of Algorithm 2, it must also be a local minimizer of g on Ω [Theorem 1], so g^* is a local minimum of g on Ω. □

We note that the convergence of our sequence of objective values $\left\{g\left(w^{(k)}\right)\right\}_{k\in\mathbb{N}}$ only implies subsequence convergence for our iterates $\left\{w^{(k)}\right\}_{k\in\mathbb{N}}$. In what follows, we prove the convergence of our sequence of iterates to a local minimizer of g on Ω for the case where $t_j < {}^{n_j}/_{\alpha_j}$ for all $j \in \mathcal{J}^{\mathrm{dv}}$.

Theorem 3. *For fixed step sizes $t_j \in (0, {}^{n_j}/_{\alpha_j})$ for all $j \in \mathcal{J}^{dv}$, the sequence of iterates $\left\{w^{(k)}\right\}_{k\in\mathbb{N}}$ generated by Algorithm 2 converges to a local minimizer of g on Ω. Furthermore, we reach an ϵ-accurate solution in $\mathcal{O}({}^1/_\epsilon)$ iterations.*

Proof. From Theorem 2 we have $g\left(w^{(k)}\right) \downarrow g^*$ as $k \to \infty$, so for all $\epsilon > 0$, there exists some positive integer K such that for all $k \geq K$, we have $g\left(w^{(k)}\right) - g^* < \epsilon$. Furthermore, from Lemma 2 we know that g is strongly convex. In particular, there exists some scalar $\mu > 0$ such that

$$g(w') \geq g(w) + \nabla g(w)^T(w' - w) + \frac{\mu}{2}\|w' - w\|_2^2 \tag{3.60}$$

for all vectors $w, w' \in \Omega$. Let w^* be a fixed point of Algorithm 2 that satisfies $g(w^*) = g^*$. Then for any $k \geq K$, we have

$$\frac{\mu}{2}\left\|w^{(k)} - w^*\right\|_2^2 \leq g\left(w^{(k)}\right) - g(w^*) - \nabla g(w^*)^T\left(w^{(k)} - w^*\right), \tag{3.61}$$

$$= g\left(w^{(k)}\right) - g^* - \nabla g(w^*)^T w^{(k)}, \tag{3.62}$$

$$< \epsilon - \nabla g(w^*)^T w^{(k)}, \tag{3.63}$$

where $\nabla g(w^*)^T w^* = 0$ because nonzero elements of w^* correspond to zero elements of $\nabla g(w^*)$ (see observations 1 and 2 from Theorem 1). We consider two cases based on the gradient of g at w^*.

Case 1. Suppose $\nabla g(w^*) = 0$, and set $\delta = \sqrt{2\epsilon/\mu}$. Then from Theorem 2 and inequality (3.63), we know that for all $\delta > 0$, there exists some positive integer K such that for all $k \geq K$, we have

$$\left\| w^{(k)} - w^* \right\|_2 < \sqrt{\frac{2\epsilon}{\mu}} = \delta. \tag{3.64}$$

Therefore $w^{(k)} \to w^*$ as $k \to \infty$, and the sequence of iterates $\left\{ w^{(k)} \right\}_{k \in \mathbb{N}}$ converges to a fixed point of Algorithm 2. In fact, because g is strongly convex and $\nabla g(w^*) = 0$, the fixed point w^* is the unique global minimizer of g on Ω, and $g(w^*) = g^*$ is the global minimum.

Case 2. Suppose $[\nabla g(w^*)]_i \neq 0$ for at least one element i. From observation 3 of Theorem 1, this will occur exactly when $-1/t_j [\partial g/\partial w_j(w^*)]_i > 0$ is less than the q_j largest positive elements of w_j^*, which means that the subvector w_j^* has reached its maximum number of positive entries. In other words, all elements of $\nabla g(w^*)$ are zero, except for possibly in entries belonging to subvectors w_j^* that have reached their maximum number of positive elements; in this case, we will have $w_i^* = 0$ and $[\nabla g(w^*)]_i < 0$.

From Theorem 2, we know that all convergent subsequences of $\left\{ w^{(k)} \right\}_{k \in \mathbb{N}}$ converge to vectors w^* such that $g(w^*) = g^*$. Consider the convergent subsequence $w^{(p_k)} \to w^*$ as $k \to \infty$. For all $\delta > 0$, there exists some positive integer L such that for all $k \geq L$, we have

$$\left\| w^{(p_k)} - w^* \right\|_2 < \delta. \tag{3.65}$$

Let ω denote the smallest positive element of w^*, and let $\delta < \omega/2$. Then for all $k \geq L$, the iterates $w^{(p_k)}$ must be positive everywhere that w^* is positive, including all subvectors w_j^* that have reached their maximum number of positive elements. In particular, this means that $w^{(p_k)}$ cannot be positive where $[\nabla g(w^*)]_i < 0$, which implies that $\nabla g(w^*)^T w^{(p_k)} \geq 0$.

Suppose that for some $k \geq L$ and $n > 0$, we also have

$$\left\| w^{(p_k+n)} - w^* \right\|_2 < \delta. \tag{3.66}$$

Then the iterate $w^{(p_k+n)}$ is positive everywhere that w^* is positive, which means that $\nabla g(w^*)^T w^{(p_k+n)} \geq 0$. If we can guarantee that the next iterate $w^{(p_k+n+1)}$ also satisfies $\nabla g(w^*)^T w^{(p_k+n+1)} \geq 0$, and we let both $\sqrt{2\epsilon/\mu} < \delta$ and $p_k + n + 1 \geq K$, then from inequality (3.63) we will have

$$\left\| w^{(p_k+n+1)} - w^* \right\|_2 < \sqrt{\frac{2\epsilon}{\mu}} < \delta. \tag{3.67}$$

The only way that we could have $\nabla g(w^*)^T w^{(p_k+n+1)} < 0$ would be if there were some element i such that $w_i^{(p_k+n+1)} > 0$ where $w_i^* = 0$ and $[\nabla g(w^*)]_i < 0$, which would need to occur on a subvector w_j^* that has reached its maximum number of positive elements. Because $w^{(p_k+n)}$ is positive everywhere that w^* is positive, the subvector $w_j^{(p_k+n)}$ would also have reached its maximum number of positive elements. In order for $w_i^{(p_k+n+1)} > 0$ where $w_i^* = 0$ and $w_i^{(p_k+n)} \leq 0$, we would require $w_\ell^{(p_k+n+1)} \leq 0$ in an entry ℓ where $w_\ell^{(p_k+n)} > 0$, that is, we would need to swap the location of a positive element in the subvector. However, since we assumed that

$$\left\| w^{(p_k+n)} - w^* \right\|_2 < \delta < \frac{\omega}{2}, \tag{3.68}$$

this would require

$$\left\| w^{(p_k+n+1)} - w^{(p_k+n)} \right\|_2 > \frac{\omega}{2}. \tag{3.69}$$

Now suppose $t_j < {}^{n_j}/_{\alpha_j}$ for all $j \in \mathcal{J}^{\mathrm{dv}}$, and define $\beta := \min_{j \in \mathcal{J}^{\mathrm{dv}}} ({}^1/_{2t_j} - {}^{\alpha_j}/_{2n_j}) > 0$. Then from inequality (3.59), we have for all $k \geq K$

$$g\left(w^{(k+1)}\right) \leq g\left(w^{(k)}\right) + \sum_{j \in \mathcal{J}^{\mathrm{dv}}} \left(\frac{\alpha_j}{2n_j} - \frac{1}{2t_j} \right) \left\| w_j^{(k+1)} - w_j^{(k)} \right\|_2^2 \tag{3.70}$$

$$\leq g\left(w^{(k)}\right) - \beta \left\| w^{(k+1)} - w^{(k)} \right\|_2^2, \tag{3.71}$$

which means that

$$\left\| w^{(k+1)} - w^{(k)} \right\|_2^2 \leq \frac{1}{\beta} \left[g\left(w^{(k)}\right) - g\left(w^{(k+1)}\right) \right] \leq \frac{1}{\beta} \left[g\left(w^{(k)}\right) - g^* \right] < \frac{\epsilon}{\beta}. \tag{3.72}$$

In particular, if we add the assumption that $\sqrt{\epsilon/\beta} < \omega/2$, then

$$\left\|w^{(p_k+n+1)} - w^{(p_k+n)}\right\|_2 < \sqrt{\frac{\epsilon}{\beta}} < \frac{\omega}{2}, \tag{3.73}$$

and it is impossible to swap the location of the positive elements in the iterates $w^{(p_k+n)}$ and $w^{(p_k+n+1)}$. Therefore $\nabla g(w^*)^T w^{(p_k+n+1)} \geq 0$, and we have

$$\left\|w^{(p_k+n+1)} - w^*\right\|_2 < \delta. \tag{3.74}$$

By induction, if we choose $\delta < \omega/2$ and $\epsilon < \min\left\{\mu\delta^2/2, \beta\omega^2/4\right\}$, then for all integers $k \geq \max\left\{p_L, K\right\}$, we have

$$\left\|w^{(k)} - w^*\right\|_2 < \delta, \tag{3.75}$$

and the sequence of iterates $\left\{w^{(k)}\right\}_{k\in\mathbb{N}}$ converges to a fixed point w^* of Algorithm 2 such that $g(w^*) = g^*$, where g^* is a local minimum of g on Ω.

To establish the convergence rate of Algorithm 2, we rearrange the terms in (3.71) and sum over k to get

$$\sum_{\ell=0}^{k} \left\|w^{(\ell+1)} - w^{(\ell)}\right\|_2^2 \leq \frac{1}{\beta}\sum_{\ell=0}^{k}\left[g\left(w^{(\ell)}\right) - g\left(w^{(\ell+1)}\right)\right] \tag{3.76}$$

$$= \frac{1}{\beta}\left[g\left(w^{(0)}\right) - g\left(w^{(k+1)}\right)\right] \tag{3.77}$$

$$\leq \frac{1}{\beta}\left[g\left(w^{(0)}\right) - g^*\right]. \tag{3.78}$$

Define $\zeta := \min_{j\in\mathcal{J}^{\text{dv}}}(t_j) > 0$. For any $\epsilon > 0$, we let

$$\min_{\ell=0,\ldots,k} \sum_{j\in\mathcal{J}^{\text{dv}}} \frac{1}{t_j^2}\left\|w_j^{(\ell+1)} - w_j^{(\ell)}\right\|_2^2 \leq \sum_{\ell=0}^{k}\sum_{j\in\mathcal{J}^{\text{dv}}} \frac{\left\|w_j^{(\ell+1)} - w_j^{(\ell)}\right\|_2^2}{(k+1)t_j^2} \tag{3.79}$$

$$\leq \sum_{\ell=0}^{k} \frac{\left\|w^{(\ell+1)} - w^{(\ell)}\right\|_2^2}{(k+1)\zeta^2} \tag{3.80}$$

$$\leq \frac{g\left(w^{(0)}\right) - g^*}{(k+1)\beta\zeta^2} < \epsilon, \tag{3.81}$$

and solving for k, we find that we reach an ϵ-accurate solution after $\mathcal{O}\left(1/\epsilon\right)$ iterations. $\square$

Algorithm 3 Block coordinate descent (Algorithm 2 with $t_j = {}^{n_j}/_{\alpha_j}$).

Input ϵ, $x^{(0)}$. Initialize $k = 0$, $\delta > \epsilon$, $w_j^{(0)} = \mathrm{proj}_{\Omega_j}\left(A_j x^{(0)} - d_j^{\mathrm{dv}}\right)$.

while $\delta > \epsilon$ **do**

$x^{(k+1)} \leftarrow \arg\min_{x \geq 0} f\left(x, w^{(k)}\right)$

$w^{(k+1)} \leftarrow \arg\min_{w \in \Omega} f\left(x^{(k+1)}, w\right)$

$\delta \leftarrow \sum_{j \in \mathcal{J}^{\mathrm{dv}}} \frac{\alpha_j}{n_j} \left\| w_j^{(k+1)} - w_j^{(k)} \right\|_2$

$k \leftarrow k + 1$

end while

return $\arg\min_{x \geq 0} f\left(x, w^{(k)}\right)$

If we choose $t_j = {}^{n_j}/_{\alpha_j}$ for each $j \in \mathcal{J}^{\mathrm{dv}}$, Algorithm 2 reduces to block coordinate descent, detailed in Algorithm 3, and its convergence to a function value corresponding to a local minimizer follows immediately from Theorem 2.

3.3.1 Simple example

To demonstrate how the relaxed formulation $(\mathcal{P}_2)$ and Algorithm 3 behave compared to the idealized problem $(\mathcal{P}_1)$, we revisit our example from the introduction. For iterations $k = 0, 3, 32$ of Algorithm 3 to solve $(\mathcal{P}_2)$, we plot the contour lines of the objective function

$$f(x, w^{(k)}) = \frac{1}{2}\|A_1 x - 81\|_2^2 + \frac{5}{2}\|w^{(k)} - (A_2 x - 20)\|_2^2 + \frac{\lambda}{2}\|x\|_2^2 \tag{3.82}$$

in terms of the beamlet intensity variable x for fixed $w^{(k)}$. Here $A_1 \in \mathbb{R}^{1\times 2}$ and $A_2 \in \mathbb{R}^{2\times 2}$ are the beamlet-to-voxel maps for the prostate tumor and rectum, $x \in \mathbb{R}^2$ contains the beamlet intensities, $w^{(k)} \in \mathbb{R}^2$ approximates the deviation between the dose received by the rectum and the dose–volume constraint, and λ is set at 5×10^{-6}. In each figure, $x^{(k+1)}$ corresponds to the global minimum of the relaxed problem at iteration k, A and B correspond to the

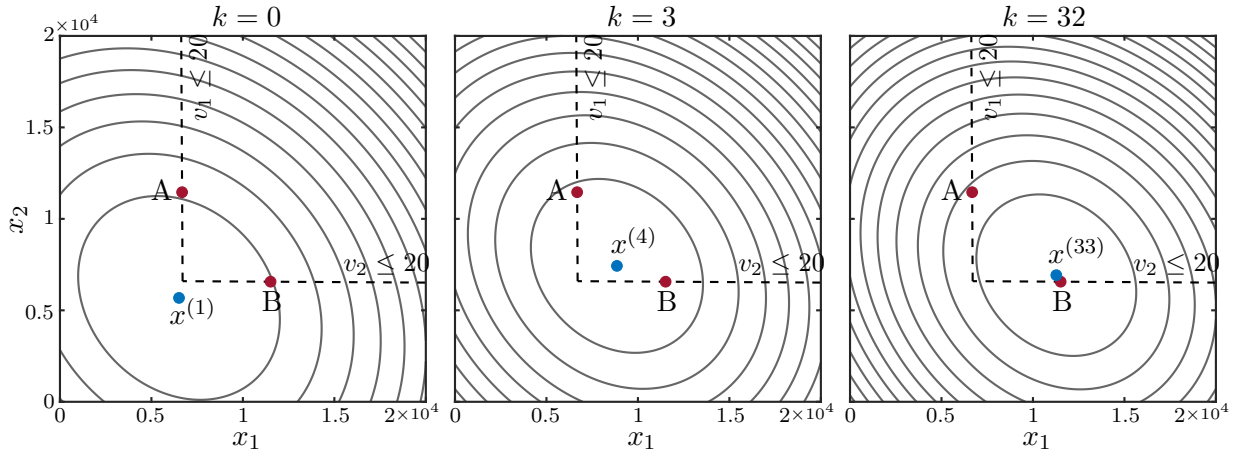

Figure 3.2: Contours of (3.82) in x for fixed iterates $w^{(k)}$ with $k = 0, 3, 32$. The points A and B correspond to the local and global minima of the idealized problem $(\mathcal{P}_1)$. Starting with the zero vector $x^{(0)}$, the global minimum of the relaxed problem $(x^{(k+1)})$ moves towards the global minimum of the idealized problem at point B.

local and global minima of the idealized problem, and the nonconvex feasible region induced by the dose–volume constraint that no more than 50% of the rectum volume exceeds 20 Gy is given by the union of regions $v_1 \leq 20$ and $v_2 \leq 20$.

In Figure 3.2 (left), we initialize the problem with $x^{(0)}$ as the zero vector. The contours of $(\mathcal{P}_2)$ are more circular than those of $(\mathcal{P}_1)$ in Figure 3.1, indicating that the relaxation has improved the problem conditioning. Additionally, as $w^{(k)}$ changes with each iteration, the global minimum of the objective function in terms of x at point $x^{(k+1)}$ moves closer to the global minimum of $(\mathcal{P}_1)$ at point B. For stopping tolerance $\epsilon = 10^{-3}$, our algorithm converges after 32 iterations, and in Figure 3.2 (right) we see that the global minimum of our relaxed problem, corresponding to our final beamlet intensities, is just shy of the global minimum of our original problem. The dose–volume constraint is met approximately, and we may increase the weight coefficient α_2 or decrease the dose level d_2^{dv} to move closer to the feasible region if necessary.

Next, we plot our iterates along with the contour lines of the objective function

$$g(w) = \min_{x \geq 0} \frac{1}{2}\|A_1 x - 81\|_2^2 + \frac{5}{2}\|w - (A_2 x - 20)\|_2^2 + \frac{\lambda}{2}\|x\|_2^2 \tag{3.83}$$

in terms of the variable w. In Figure 3.3 (left), points A and B again correspond to the local and global minima of the idealized problem, and the nonconvex feasible region $\Omega = \{w \in \mathbb{R}^2 \,:\, \|(w)_+\|_0 \leq 1\}$ is given by the union of regions $w_1 \leq 0$ and $w_2 \leq 0$. The results of the gradient step in (3.12) are plotted as squares, and the projections onto Ω are plotted as circles. Just as we saw in x-space, our algorithm in w-space converges to a point near the global minimum at B.

3.3.2 Initialization

Because our formulation is nonconvex, the initialization can affect both the solution and the time to convergence. For example, using the initialization $x^{(0)} = (0, 2 \times 10^3)$ for our simple example, we converge near the local minimum at A after 37 iterations (center panel of Figure 3.3). Fortunately, as demonstrated in other papers (see e.g., [96, 182]), local minima do not have a large effect on the solutions to these least-squares formulations, especially when the convex tumor terms have weights relative in size to the OAR term weights [138]. Given enough time, different initializations tend to converge to similar objective values, fluence maps, and dose distributions.

For the remainder of our examples, we initialize the beamlet intensity vector as the solution to $(\mathcal{P}_2)$ without any dose–volume terms, i.e.,

$$x^{(0)} = \arg\min_{x \geq 0} \sum_{i \in \mathcal{I}} \frac{\alpha_i}{2n_i}\|A_i x - d_i\|_2^2 + \frac{\lambda}{2}\|x\|_2^2. \tag{3.84}$$

In Figure 3.3 (right), we use this initialization for our simple example, converging near the global minimum at B after 26 iterations. For this initialization, we see that $A_2 x^{(0)} - 20$, the value of $w^{(0)}$ before projection onto Ω, corresponds to the global minimum of $g(w)$. This is

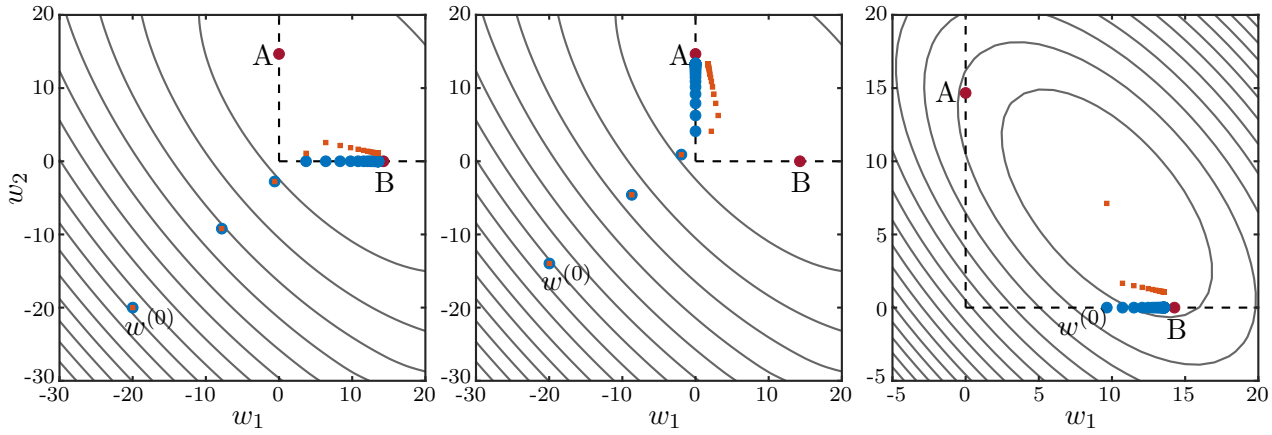

Figure 3.3: Contours of (3.83) with respect to w for different initializations. The points A and B correspond to the local and global minima of the idealized problem $(\mathcal{P}_1)$, the squares correspond to the iterates $w^{(k)}$ before projection onto the feasible set Ω, and the circles correspond to the iterates after projection. (Left) Starting with the zero vector $x^{(0)}$, we converge after 32 iterations at a point near the global minimum at B. (Center) Starting with the initialization $x^{(0)} = (0, 2 \times 10^3)$, we converge near the local minimum at A after 37 iterations. (Right) Starting with the initialization $x^{(0)} = \arg\min_{x \geq 0} {}^1\!/_2 \|A_1 x - 81\|_2^2 + {}^\lambda\!/_2 \|x\|_2^2$, we converge near the global minimum at B after 26 iterations.

true for the general problem $(\mathcal{P}_2)$ as well. Specifically, for all $j \in \mathcal{J}^{\mathrm{dv}}$ we have

$$\arg\min_{w_j} f(x, w) = A_j x - d_j^{\mathrm{dv}}, \tag{3.85}$$

which means that

$$\min_{w} f(x, w) = \sum_{i \in \mathcal{I}} \frac{\alpha_i}{2n_i} \|A_i x - d_i\|_2^2 + \frac{\lambda}{2} \|x\|_2^2. \tag{3.86}$$

Therefore letting $x^{(0)} = \arg\min_{x \geq 0} (\min_w f(x, w))$, we are guaranteed that

$$A_j x^{(0)} - d_j^{\mathrm{dv}} = \arg\min_{w_j} g(w) \tag{3.87}$$

for all $j \in \mathcal{J}^{\mathrm{dv}}$ because

$$\min_{x \geq 0} \left(\min_{w} f(x, w) \right) = \min_{w} g(w). \tag{3.88}$$

Since we are ultimately interested in solving $\min_{w \in \Omega} g(w)$, using an initialization for $x^{(0)}$ that corresponds to $w^{(0)} = \mathrm{proj}_\Omega(\arg\min_w g(w))$ is a natural choice.

3.4 Numerical Examples

Our examples are performed on data from an anonymized prostate cancer patient included in the CORT dataset [49], available in the GigaScience repository GigaDB [50]. The dataset includes beamlet-to-voxel maps for beam angles ranging from 0° to 358° in increments of 2°, PTVs in the prostate and lymph nodes, and OARs including the rectum, bladder, femoral heads, and unspecified tissue (Figure 3.4). In the CORT dataset, beamlet intensities are measured in monitor units (MU), where 100 MU delivers a dose of 1 Gy in 10 cm depth in water in the center of a 10 cm $\times$ 10 cm radiation field [49].

For a fixed set of beams, we consider the problem of treating tumors with dose–volume constraints of increasing difficulty on various combinations of critical structures. In Section 3.4.1, we look at an example with one PTV and one OAR, and show how the proposed approach works for a single dose–volume constraint on the OAR. In Section 3.4.2, we increase the complexity of the previous example by imposing multiple dose–volume constraints on the same PTV and OAR. In Section 3.4.3, we show how the approach handles multiple PTVs and OARs. Finally, in Section 3.4.4 we compare our approach to other published methods. For all of our examples, we use seven equally spaced beams with angles ranging from 0° to 312° in increments of 52°. We set our regularization parameter at $\lambda = 10^{-8}$, our stopping tolerance at $\epsilon = 10^{-3}$, and for simplicity we let $\alpha_i = \alpha_j = 1$ for all $i \in \mathcal{I}$ and all $j \in \mathcal{J}^{\mathrm{dv}}$, unless stated otherwise. Whenever there is overlap between PTVs and OARs, the voxels in common are assigned to the PTV. However, this choice is not required for our model, and a single voxel may be assigned to multiple structures if desired. All examples were run on a computer with a 2.9 GHz dual-core Intel Core i5 processor with 8 GB RAM.

We use the cumulative dose–volume histogram to evaluate the quality of the resulting treatment plans. In a dose–volume histogram, each point represents the percentage of the organ volume that receives at least a particular dose. For example, we aim to deliver a uniform dose of 81 Gy to the prostate tumor, corresponding to the dose–volume histogram

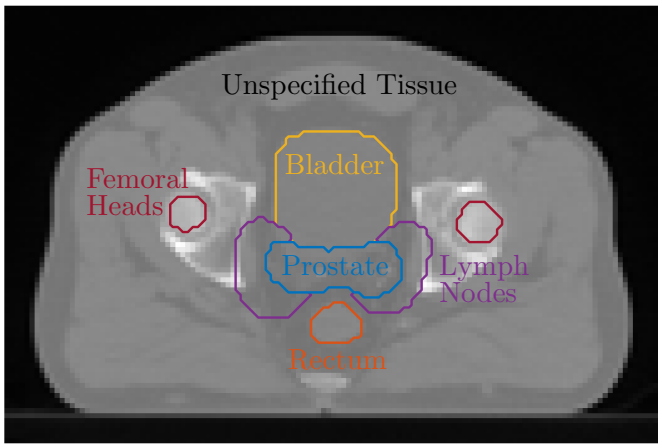

Figure 3.4: A representative axial cross-section of organs from the CORT prostate case. The prostate corresponds to planning target volume PTV_68 in the CORT dataset, while the lymph nodes correspond to planning target volume PTV_56.

on the left panel of Figure 3.5. In this case, 100% of the organ volume receives exactly 81 Gy. As we cannot deliver a perfectly uniform dose, our goal is to find a treatment plan that is as close as possible to this idealized dose distribution. For instance, in the right panel of Figure 3.5 we see the dose–volume histogram of the solution to (3.84) for one prostate tumor with regularization parameter $\lambda = 10^{-8}$. We use the beamlet intensities from this solution as the initialization for our examples in Sections 3.4.1, 3.4.2, and 3.4.4.1.

Cumulative dose–volume histograms can also be used to determine whether or not an upper dose–volume constraint on an OAR has been met by a particular treatment plan. For example, we consider the dose–volume constraint that no more than 50% of the rectum volume exceeds 50 Gy. In Figure 3.6, the rectangular region within the dotted lines contains all points where at most 50% of the rectum volume receives more than 50 Gy. Therefore any dose curve that lies outside of the box does not meet the constraint, while any curve that travels inside the box meets the constraint. In the left panel of Figure 3.6, approximately 57% of the rectum volume receives more than 50 Gy, so the dose–volume constraint is not

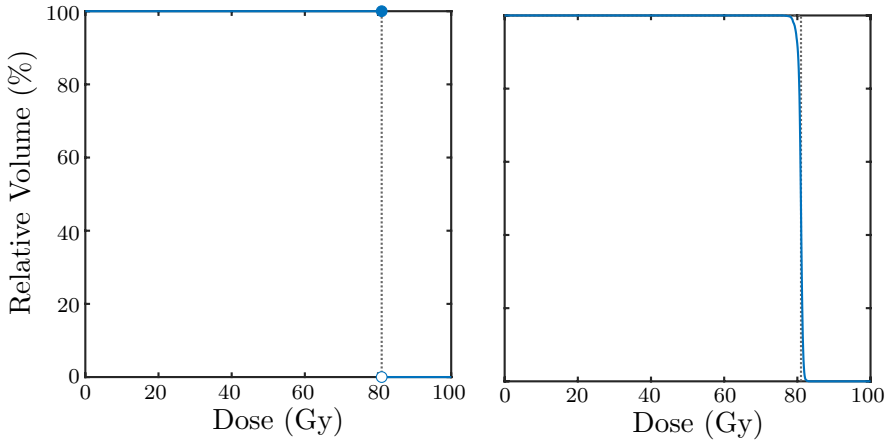

Figure 3.5: Dose–volume histograms for a uniform PTV prescription of 81 Gy. (Left) An ideal dose profile, with 100% of the target volume receiving exactly 81 Gy. (Right) A more realistic dose profile, corresponding to our initialization for Sections 3.4.1 and 3.4.2, obtained by solving (3.84) for one prostate tumor with regularization parameter $\lambda = 10^{-8}$.

met. However, in the right panel of Figure 3.6, approximately 42% of the rectum volume receives more than 50 Gy, so the dose–volume constraint is met.

3.4.1 One PTV and one OAR with one dose–volume constraint

For our first example, we consider the problem of delivering a uniform dose of 81 Gy to a prostate tumor, while satisfying the dose–volume constraint that no more than 30% of the rectum volume receives more than 30 Gy. In this case, the PTV has 6770 voxels, the OAR has 1764 voxels, and there are 986 beamlets. In Figure 3.7, we see a representative axial cross-section of the calculated dose along with the intensities of four of the beams (due to symmetry, the remaining three beams have similar intensity patterns). In this cross-section, a nearly uniform dose of 81 Gy is delivered to the tumor, while only a small percentage of the rectum volume receives more than 30 Gy.

Figure 3.8 shows the behavior of our model using Algorithm 3. The objective $g(w)$ decreases monotonically (top left), and Algorithm 3 converges after 42 iterations. The trade-

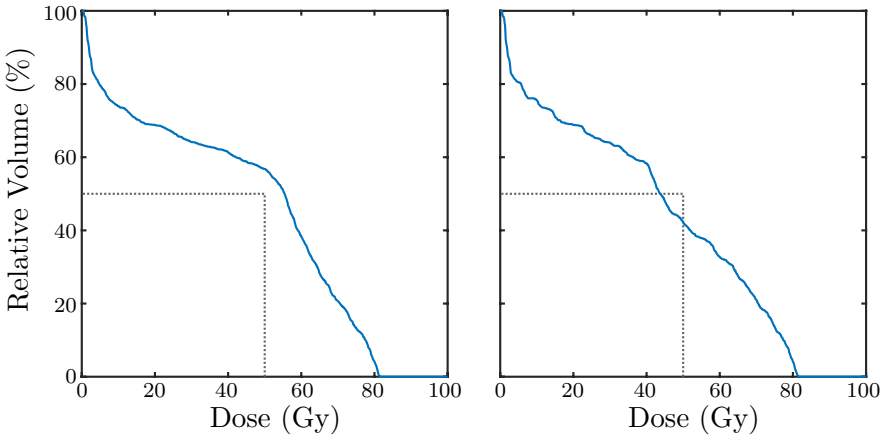

Figure 3.6: Dose–volume histograms for an OAR dose–volume constraint that no more than 50% of the rectum volume exceeds 50 Gy. The solid lines correspond to the dose that the organ receives, and the dotted lines enclose the region that satisfies the dose–volume constraint. (Left) A treatment plan that does not meet the dose–volume constraint, with approximately 57% of the organ volume exceeding 50 Gy. (Right) A treatment plan that does meet the dose–volume constraint, with approximately 42% of the organ volume exceeding 50 Gy.

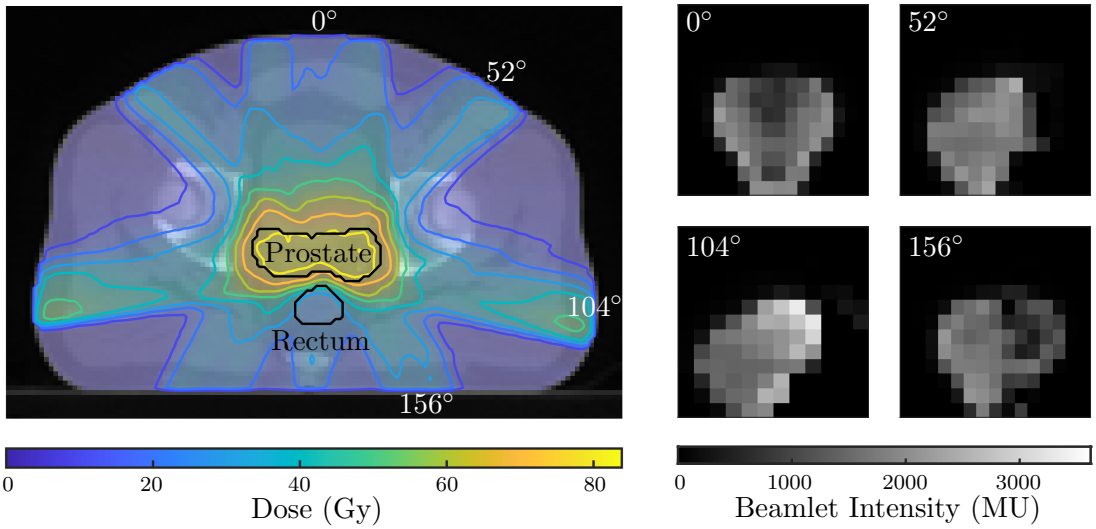

Figure 3.7: Solution for example in Section 3.4.1 using Algorithm 3. (Left) The prostate tumor is irradiated by seven equally spaced beams with a dose–volume constraint on the rectum. (Right) Beamlets intensities for four of the seven beams, calculated to deliver a nearly uniform dose of 81 Gy to the PTV while ensuring that no more than 30% of the OAR volume receives more than 30 Gy. Due to symmetry, the remaining three beams have similar intensity patterns.

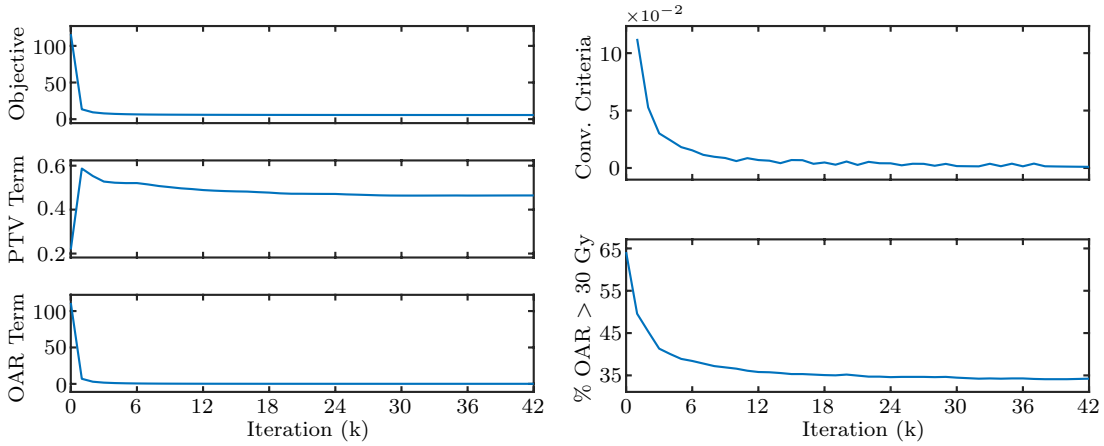

Figure 3.8: Convergence of example in Section 3.4.1 using Algorithm 3. (Left) Iterates of the objective function $g(w)$ and individual organ terms (PTV term $= {}^{1}/_{(2\times 6770)}\|A_1x - 81\|_2^2$ and OAR term $= {}^{1}/_{(2\times 1764)}\|w - (A_2x - 30)\|_2^2$). While both $g(w)$ (top) and the OAR term (bottom) decrease monotonically, the PTV term (center) adjusts to accommodate the competing OAR objective. (Right) Convergence of the auxiliary variable measured by ${}^{1}/_{t}\|w^{(k)} - w^{(k-1)}\|_2$ (top) and our approximation of the dose–volume constraint (bottom).

off between the competing PTV and OAR objectives is shown in the remaining left panels of Figure 3.8. Because our initial beamlet intensities correspond to a solution with no dose–volume constraints, the tumor term achieves its lowest value and the rectum term achieves its highest value at the initialization. Once we begin iterating, the inclusion of the OAR term causes the tumor term value to increase and the rectum term to decrease. On the right panel of Figure 3.8, we show the behavior of the auxiliary variable w (top right) and our approximation of the dose–volume constraint (bottom right). The percent of rectum voxels exceeding 30 Gy approaches the dose–volume constraint; however, because of the relaxation, we do not meet the dose–volume constraint exactly. Nevertheless, we improve upon the initialization value according to the trade-off implicitly specified by our weights α_1 and α_2.

To demonstrate how the weight factors and dose–volume parameters can influence solutions, we implement a variation of the penalty decomposition method for the convex-cardinalty problem proposed in [100]. Specifically, we use the iterative re-weighting scheme

Algorithm 4 Iterative re-weighting for Algorithm 3.

Input σ, γ, $x^{(0)}$. Initialize $k = 0$, $\alpha_j^{(0)}$, $d_j^{\mathrm{dv}(0)}$, $p_j^{(0)}$, $\epsilon^{(0)}$.

while not converged **do**

$\left(x^{(k+1)}, w^{(k+1)}\right) \leftarrow$ solution to $(\mathcal{P}_2)$ from Algorithm 3 with $\left(x^{(k)}, \alpha_j^{(k)}, d_j^{\mathrm{dv}(k)}, p_j^{(k)}, \epsilon^{(k)}\right)$

for $j \in \mathcal{J}^{\mathrm{dv}}$ **do**

if $\left\| \left(A_j x^{(k+1)} - d_j^{\mathrm{dv}(0)}\right)_+ \right\|_0 > \frac{n_j p_j^{(0)}}{100}$ **then**

$\alpha_j^{(k+1)} \leftarrow (1+\sigma)\alpha_j^{(k)}$

$d_j^{\mathrm{dv}(k+1)} \leftarrow (1-\sigma) d_j^{\mathrm{dv}(k)}$

$p_j^{(k+1)} \leftarrow (1-\sigma) p_j^{(k)}$

end if

end for

$\epsilon^{(k+1)} \leftarrow \gamma \epsilon^{(k)}$

$k \leftarrow k + 1$

end while

return $x^{(k)}$

given in Algorithm 4, where $(\mathcal{P}_2)$ is solved multiple times with increasing OAR weights, decreasing dose–volume parameters, and decreasing stopping tolerance values. In this example, we use initial OAR weight $\alpha_2^{(0)} = 1$, dose parameter $d_2^{\mathrm{dv}(0)} = 30$, volume parameter $p_2^{(0)} = 30$, and stopping tolerance $\epsilon^{(0)} = 10^{-3}$, with weight update parameter $\sigma = 0.01$ and stopping tolerance update parameter $\gamma = 0.99$. For simplicity, we use the same update parameter σ for the weight, dose, and volume, but different parameters or update schedules could be used if needed.

We compare the approximation of the dose–volume constraint in our solutions with and without re-weighting in Table 3.1 and Figure 3.9. To illustrate the trade-off between the

competing PTV and OAR objectives, we include results from Algorithm 4 using different convergence criteria: (a) stopping once the dose–volume constraint is met, and (b) stopping once the PTV D95 (i.e., the minimum dose delivered to 95% of the PTV) is less than 98% of its initial value, where weight factors and dose–volume parameters are updated regardless of whether or not the dose–volume constraint has been satisfied. There is little difference between the initial tumor dose and the final tumor dose for each solution, both in terms of the PTV D95 and the dose–volume histogram. At the same time, all solutions have reduced the dose to the rectum to approximately satisfy the dose–volume constraint, where re-weighting produces a lower dose at the expense of increased computation time. Without re-weighting, the constraint is roughly met, with only 34.16% of the volume receiving more than 30 Gy, in contrast to the initialization, where 64.14% of the volume exceeds 30 Gy. With re-weighting, the constraint is satisfied, with 29.61% and 19.84% of the volume receiving more then 30 Gy for stopping conditions (a) and (b), respectively.

In general, as the dose to the rectum decreases, the PTV D95 also decreases, while the dose to the tumor deviates more from the target uniform dose. If the dose–volume constraint makes the problem completely infeasible, the approach still returns some solution, but it need not meet the dose–volume constraint nor the target tumor dose. In this case, the planner can influence the trade-off between treatment goals by modifying weighting factors, dose–volume parameters, and convergence criteria.

3.4.2 One PTV and one OAR with multiple dose–volume constraints

Next we consider the problem of delivering a uniform dose of 81 Gy to the prostate tumor, with multiple dose–volume constraints on the tumor and rectum:

- No more than 5% of the tumor volume receives less than 81 Gy,
- no more than 0% of the tumor volume receives more than 85 Gy,
- no more than 20% of the rectum volume receives more than 60 Gy,

Table 3.1: Results for example in Section 3.4.1 using Algorithms 3 and 4.

	PTV D95 (Gy)	OAR % > 30 Gy	Time (s)
Initialization	79.65	64.14	0.18
Solution w/ Alg. 3	79.17	34.16	7.22
Solution w/ Alg. 4 (a)	79.03	29.61	33.61
Solution w/ Alg. 4 (b)	78.05	19.84	163.51

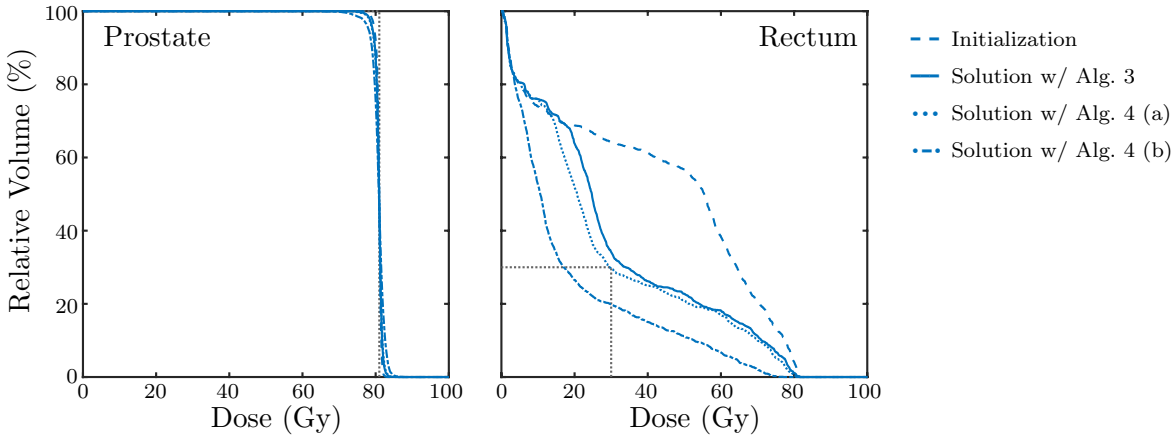

Figure 3.9: Dose–volume histograms for example in Section 3.4.1 using Algorithms 3 and 4. (Left) There is little notable difference between the initial and final dose to the PTV. (Right) Without re-weighting, the constraint that no more than 30% of the OAR volume receives more than 30 Gy is approximately met, with 34.16% of the volume exceeding 30 Gy, as opposed to the initialization where 64.14% does. With re-weighting, 29.61% and 19.84% the volume exceeds 30 Gy for stopping conditions (a) and (b), respectively.

- no more than 40% of the rectum volume receives more than 40 Gy, and
- no more than 60% of the rectum volume receives more than 20 Gy.

The first constraint, a lower dose–volume constraint on the PTV, is implemented similarly to the upper dose–volume constraint on the OAR, with the exception that the corresponding term in the objective function is changed to

$$\frac{\alpha_i}{2n_i}\|w_i - (d_i^{\mathrm{dv}} - A_i x)\|_2^2, \tag{3.89}$$

and the projection onto Ω_i is computed by ordering the elements of w_i and setting the highest $\lceil (100-p_i)n_i/100 \rceil$ entries of w_i to $\max\{0, w_i\}$. The second constraint corresponds to a maximum-dose constraint, but expressing it as an upper dose–volume constraint allows us to apply our relaxation rather than enforcing a hard constraint.

As in the previous example, the PTV has 6770 voxels, the OAR has 1764 voxels, and there are 986 beamlets. One downside to adding multiple dose–volume constraints on a particular structure is the inclusion of more terms in the objective function, increasing the problem dimension. However, we can reduce the number of objective terms by combining each of the K_j dose–volume constraint terms corresponding to organ j,

$$\frac{\alpha_{j_k}}{2n_j}\|w_{j_k} - (A_j x - d_{j_k}^{\mathrm{dv}})\|_2^2 \quad \text{s.t.} \quad \|(w_{j_k})_+\|_0 \leq \frac{n_j p_{j_k}}{100} \quad \text{for} \quad k = 1, \ldots, K_j, \tag{3.90}$$

into one term,

$$\frac{\alpha_j}{2n_j}\|y_j - A_j x\|_2^2 \quad \text{s.t.} \quad y_j \in \tilde{\Omega}_j, \tag{3.91}$$

where $\tilde{\Omega}_j = \tilde{\Omega}_{j_1} \cap \cdots \cap \tilde{\Omega}_{j_{K_j}}$ represents the set of vectors that satisfy all of the organ's K_j dose–volume constraints, with each constraint defined as

$$y_j \in \tilde{\Omega}_{j_k} := \left\{ y_j \in \mathbb{R}^{n_j} \,:\, \|(y_j - d_{j_k}^{\mathrm{dv}})_+\|_0 \leq \frac{n_j p_{j_k}}{100} \right\} \quad \text{for} \quad k = 1, \ldots, K_j. \tag{3.92}$$

Projections onto $\tilde{\Omega}_j$ can be computed efficiently by ordering the elements of y_j, setting the lowest $(100 - p_{j_k})\%$ of y_j to $\min\{d_{j_k}, y_j\}$ for upper dose–volume constraints, and setting

the highest $(100 - p_{j_k})\%$ of y_j to $\max\{d_{j_k}, y_j\}$ for lower dose–volume constraints, for all $k = 1, \ldots, K_j$. We note that by combining terms for a particular OAR, we lose the ability to assign different weights to each of the OAR's dose–volume objectives. However, one can always create different groupings of terms based on weight factor.

Using Algorithm 3 with $\alpha_1 = 10$ for the PTV dose–volume constraint term, we converge after 21 iterations. In Table 3.2 and Figure 3.10, we see that the dose–volume constraints have been approximately met. Specifically, none of the PTV volume receives more than 85 Gy, 24.47% of the PTV volume receives less than 81 Gy (53.26% initially), and the PTV D95 increases from 79.65 Gy to 80.26 Gy. Furthermore, 61.95% of the OAR volume exceeds 20 Gy (68.81% initially), 42.9% exceeds 40 Gy (61.29% initially), and 21.84% exceeds 60 Gy (38.47% initially). Importantly, it is possible to specify the shape of the PTV and OAR dose distributions in terms of multiple dose–volume constraints without any significant increase in computation.

Next, we use Algorithm 4 with weight factors and dose–volume parameters initialized with the values used in Algorithm 3, initial stopping tolerance $\epsilon^{(0)} = 10^{-3}$, weight update parameter $\sigma = 0.01$, and stopping tolerance update parameter $\gamma = 0.99$. For the lower dose–volume constraint on the PTV, the dose parameter increases each iteration,

$$d_i^{\mathrm{dv}(k+1)} = (1 + \sigma) d_i^{\mathrm{dv}(k)}, \tag{3.93}$$

rather than decreasing like the upper dose–volume constraints. Stopping once all dose–volume constraints are met, none of the PTV volume receives more than 85 Gy, 4.09% of the PTV volume receives less than 81 Gy, and the PTV D95 is 81.32 Gy. In addition, only 59.1% of the OAR volume exceeds 20 Gy, 39.26% exceeds 40 Gy, and 19.78% exceeds 60 Gy. In this case it is possible to satisfy all of the dose–volume constraints, but in cases of infeasibility due to competing PTV and OAR dose–volume constraints, additional stopping criteria for Algorithm 4 can be specified, e.g., setting a maximum number of iterations.

Table 3.2: Results for example in Section 3.4.2 using Algorithms 3 and 4.

	PTV D95 (Gy)	PTV % < 81 Gy	OAR % > 20 Gy	OAR % > 40 Gy	OAR % > 60 Gy	Time (s)
Initialization	79.65	53.26	68.81	61.29	38.47	0.18
Solution w/ Alg. 3	80.26	24.27	61.95	42.90	21.84	8.89
Solution w/ Alg. 4	81.32	4.09	59.10	39.26	19.78	55.86

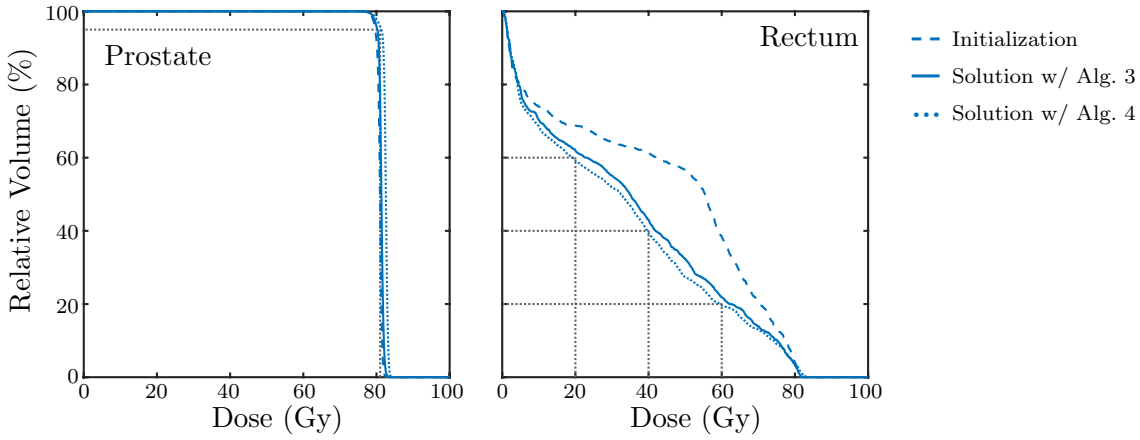

Figure 3.10: Dose–volume histograms for example in Section 3.4.2 using Algorithms 3 and 4. (Left) Without re-weighting, the dose–volume constraints on the PTV are approximately met, with 24.27% of the volume receiving less than 81 Gy (53.26% initially) and none of the volume exceeding 85 Gy. With re-weighting, only 4.09% of the volume receives less than 81 Gy. (Right) Without re-weighting, the dose–volume constraints on the OAR are approximately met, with 61.95% of the volume exceeding 20 Gy (68.61% initially), 42.9% exceeding 40 Gy (61.29% initially), and 21.84% exceeding 60 Gy (38.47% initially). With re-weighting, only 59.1% of the volume exceeds 20 Gy, 39.25% exceeds 40 Gy, and 19.78% exceeds 60 Gy.

3.4.3 Multiple PTVs and OARs

Now we consider the problem of delivering a uniform dose of 81 Gy to the prostate tumor and a uniform dose of 60 Gy to the lymph nodes, while satisfying the dose–volume constraints that no more than 50% of the rectum receives more than 50 Gy and no more than 30% of the bladder receives more than 30 Gy. In this case the first PTV has 6770 voxels, the second PTV has 8483 voxels, the first OAR has 1643 voxels, the second OAR has 9708 voxels, and there are 986 beamlets. We converge after 48 iterations of Algorithm 3.

Due to the proximity of the two PTVs and the difference between their prescription doses, it is much more difficult to deliver a uniform dose to both targets. The increase in geometric constraints and competing objectives results in more complicated dose distributions and beamlet intensity patterns, illustrated in Figure 3.11. More of the prostate volume receives less than 81 Gy than in previous examples, while more of the lymph node volume exceeds 60 Gy than we would like (Figure 3.12, top row). In this case, it may be important to add lower and upper dose–volume constraints on the PTVs to ensure that appropriate doses are delivered. The dose–volume constraints on the rectum and bladder are both approximately met, with 57.33% of the rectum exceeding 50 Gy (82.47% initially) and 38.14% of the bladder exceeding 30 Gy (91.87% initially) (Figure 3.12, bottom row).

Using Algorithm 4 with weight factors and dose–volume parameters initialized with the values used in Algorithm 3, initial stopping tolerance $\epsilon^{(0)} = 10^{-3}$, weight update parameter $\sigma = 0.01$, and stopping tolerance update parameter $\gamma = 0.99$, we are able to meet both of the dose–volume constraints on the OARs. Specifically, only 47.53% of the rectum volume receives more than 50 Gy and 29.93% of the bladder volume exceeds 30 Gy, with little difference to the PTV doses.

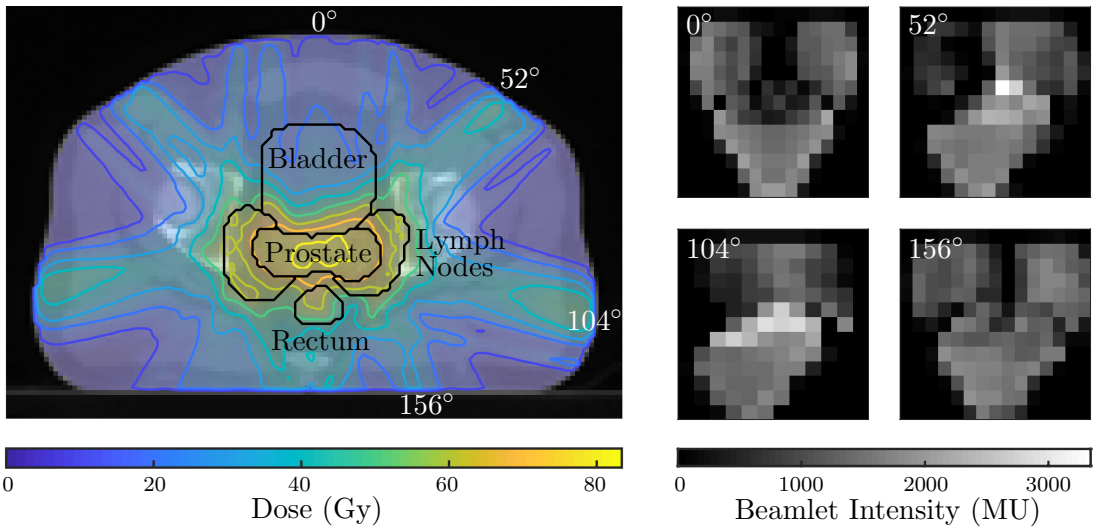

Figure 3.11: Solution for example in Section 3.4.3 using Algorithm 3. (Right) Targets in the prostate and lymph nodes are irradiated by seven equally spaced beams with dose–volume constraints on the rectum and bladder. (Left) Beamlet intensities for four of the seven beams, calculated to deliver a nearly uniform dose of 81 Gy to the prostate and 60 Gy to the lymph nodes while ensuring that no more than 50% of the rectum volume receives more than 50 Gy and no more than 30% of the bladder volume receives more than 30 Gy. Due to symmetry, the remaining three beams have similar intensity patterns.

Table 3.3: Results for example in Section 3.4.3 using Algorithms 3 and 4.

	PTV 1 D95 (Gy)	PTV 2 D95 (Gy)	OAR 1 % > 50 Gy	OAR 2 % > 30 Gy	Time (s)
Initialization	76.96	58.53	82.47	91.87	0.37
Solution w/ Alg. 3	76.70	57.51	57.33	38.14	15.52
Solution w/ Alg. 4	76.48	57.31	47.53	29.93	84.83

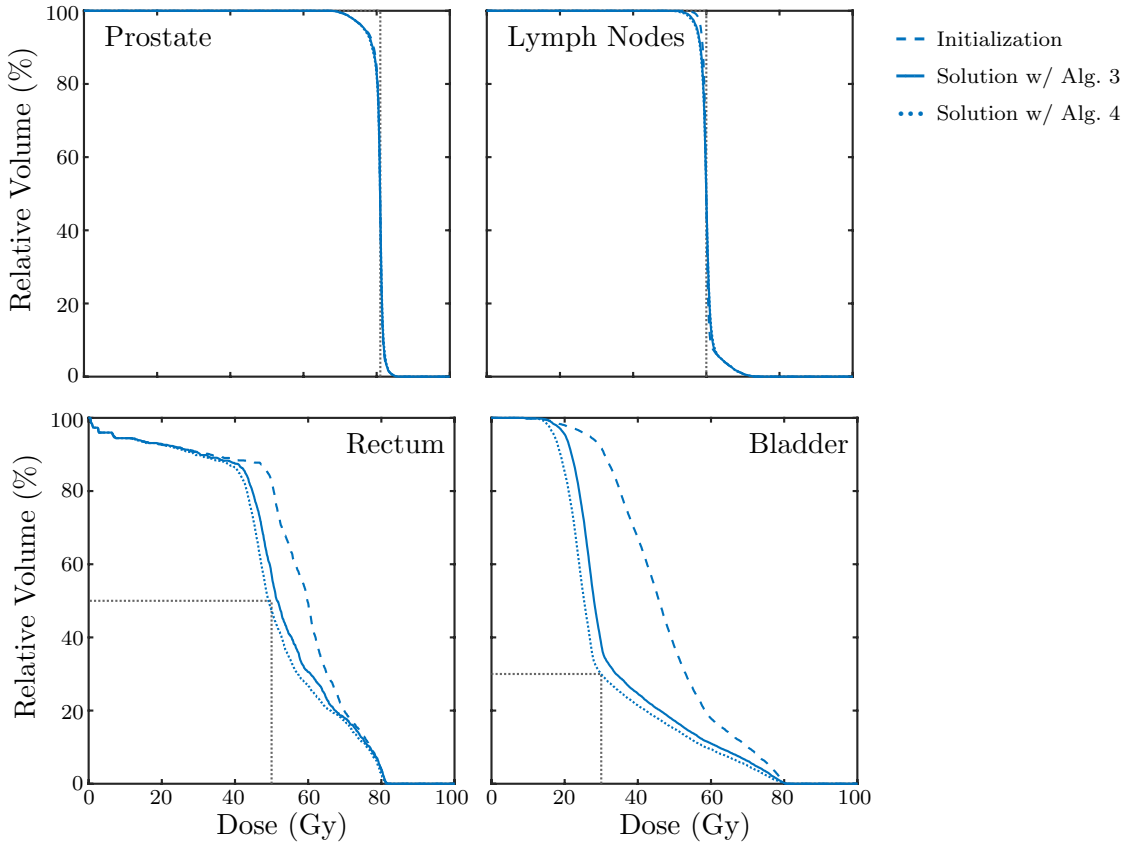

Figure 3.12: Dose–volume histograms for example in Section 3.4.3 using Algorithms 3 and 4. (Top) There is little noticeable difference between the dose to the PTVs in the prostate and lymph nodes, though their proximity causes less prostate voxels to receive their target dose and more lymph node voxels to exceed their target dose than we would like. (Bottom) Without re-weighting, the dose–volume constraints to the OARs are approximately met, with 57.33% of the rectum exceeding 50 Gy (82.47% initially) and 38.14% of the bladder exceeding 30 Gy (91.87% initially). With re-weighting, only 47.53% of the rectum volume exceeds 50 Gy and 29.93% of the bladder volume exceeds 30 Gy.

3.4.4 Comparisons with other methods

Finally, we compare results from Algorithm 3, Algorithm 4, and four other published methods for solving the FMO problem with dose–volume constraints. We start with two methods that, like our approach, approximate dose–volume constraints with objective terms, and thus provide flexibility for cases of infeasiblity. In particular, as mentioned in Section 3.1.3, many least-squares models include a penalty term for OAR voxels that exceed a specified dose (see e.g., [32, 56, 96, 97, 159, 181, 190, 191]). This penalty may apply to the entire OAR, or in the case of dose–volume objectives, to a fraction of the OAR volume. Following the approach used in [96, 97, 159], we replace our dose–volume terms in (3.7a) with

$$\sum_{j\in\mathcal{J}^{\mathrm{dv}}} \frac{\alpha_j}{2n_j}\|(A_jx-\tilde{d}_j^{\mathrm{dv}})_+\|_2^2, \tag{3.94}$$

where the vector $\tilde{d}_j^{\mathrm{dv}}$ changes each iteration. Specifically, after each iteration we sort the voxel dose vector A_jx, assigning the highest $p_j\%$ to the dose value $\tilde{d}_j^{\mathrm{dv}}=10^6$ and the remaining $(100-p_j)\%$ to $\tilde{d}_j^{\mathrm{dv}}=d_j^{\mathrm{dv}}$. The beamlet vector x is updated each iteration with a step of gradient descent, where only the the lowest $(100-p_j)\%$ of voxels that exceed $\tilde{d}_j^{\mathrm{dv}}$ influence the direction of the gradient, followed by a projection onto the nonnegative orthant. If $\tilde{d}_j^{\mathrm{dv}}$ converges to a fixed dose vector, then the problem becomes convex [96].

Alternatively, we can penalize voxels that exceed a given dose by adding nonnegative slack variables to the OAR terms. For example, following the model used in the Sensitivity-Driven Greedy Algorithm [190, 191], we replace our dose–volume terms in (3.7a) with

$$\sum_{j\in\mathcal{J}^{\mathrm{dv}}} \frac{\alpha_j}{2n_j}\|A_jx+s_j-y_j\|_2^2 \quad \text{s.t.} \quad s_j\geq 0, \quad y_j\in\tilde{\Omega}_j, \tag{3.95}$$

which is similar to our modified OAR terms in (3.91). The alternating projections approach of the Sensitivity-Driven Greedy Algorithm is equivalent to the block coordinate descent approach of Algorithm 3, with two important differences. First, the nonnegative least-squares step to compute $x^{(k+1)}$ also updates the slack variables $s_j^{(k+1)}$, so the number of

variables in this subproblem increases by $\sum_{j\in\mathcal{J}^{\mathrm{dv}}} n_j$. Second, to ensure the convergence of the dose variables, the $y^{(k+1)}$ update includes a projection onto the modified set $\tilde{\Omega}_j \cap \{y_j \in \mathbb{R}^{n_j} : y_j \geq y_j^{(k)}\}$. Using the initialization $y_j^{(0)} = d_j^{\mathrm{dv}}$, this approach can be interpreted as iteratively selecting the voxels that may exceed d_j^{dv}, and once $\lfloor {}^{n_j p_j}/_{100} \rfloor$ elements of vector $y_j^{(k)}$ surpass d_j^{dv}, the problem becomes convex.

Next we look at two methods that solve a convex problem by imposing maximum-dose constraints on a chosen subvolume of each OAR. As mentioned in Section 3.1.3, many methods take a multi-stage approach of 1) solving an approximation of the original problem, 2) determining which voxels may exceed d_j^{dv} based on the resulting solution, and 3) polishing the solution from Step 1 by solving a convex problem using the subvolumes determined in Step 2. For example, one heuristic to choose the $p_j\%$ of the voxels that may receive more than d_j^{dv} Gy is to solve the FMO problem without dose–volume constraints [141, 142, 74]. Following this approach, we first solve $(\mathcal{P}_1)$ without any dose–volume constraints (which is equivalent to our initialization in (3.84)), and then re-solve with maximum-dose constraints applied to the $(100-p_j)\%$ of the voxels that received the lowest dose in the previous solution.

Another heuristic is to solve an approximation of $(\mathcal{P}_1)$ where the dose–volume constraints in (3.6b) are replaced by the convex constraint

$$\|(\nu_j + A_j x - d_j^{\mathrm{dv}})_+\|_1 \leq \frac{\nu_j n_j p_j}{100} \quad \text{s.t.} \quad \nu_j \geq 0, \tag{3.96}$$

where ν_j is optimized along with the beamlet vector x [59]. If this problem is feasible, the solution is guaranteed to satisfy the dose–volume constraints, but it may not be optimal for $(\mathcal{P}_1)$. Therefore, the problem is polished by re-solving with maximum-dose constraints applied to the $(100 - p_j)\%$ of the voxels that received the lowest dose in the approximate solution. We compare results from these two convex methods with solutions from our re-weighting scheme in Algorithm 4 and solutions from Algorithm 3 which are then polished to satisfy the dose–volume constraints.

3.4.4.1 One PTV and one OAR.

First we consider the problem from Section 3.4.1 with stopping tolerance $\epsilon = 10^{-2}$, where an additional lower dose–volume constraint that no more than 5% of the tumor volume receives less than 81 Gy has been applied to the PTV. In Table 3.4 and Figure 3.13, we see the results of using approximate dose–volume objectives. Because none of the solutions come close to satisfying the lower dose–volume constraint on the PTV, and only one satisfies the upper dose–volume constraint on the OAR, we look for the best candidate method for re-weighting or polishing. Compared to the solution with (3.94), our solution achieves a higher PTV D95 and a lower OAR dose in less time. Because the method with (3.94) does not solve a subproblem to update the beamlet vector, it unsurprisingly requires less time per iteration than the other two methods, though it does require more iterations overall. The method with (3.95) requires significantly more time than the other two approaches due to the additional slack variables, but it produces the only solution that meets the dose–volume constraint on the OAR. However, its lower OAR dose comes at the expense of a lower PTV D95, and in less time we can use re-weighting or polishing with $(\mathcal{P}_2)$ to meet all of the dose–volume constraints (first two rows of Table 3.5).

Next we examine results computed through either re-weighting or polishing. All four solutions in Table 3.5 and Figure 3.14 satisfy the dose–volume constraints on the PTV and OAR, and are thus feasible solutions of $(\mathcal{P}_1)$. However, the lowest objective value for the idealized problem in (3.6a) is achieved by solving $(\mathcal{P}_2)$ with Algorithm 3, then re-solving with maximum-dose constraints applied to subvolumes of the PTV and OAR. The solution with (3.84) has the highest objective value, perhaps because it is the only method that does not include any approximations of the dose–volume constraints in Step 1, while the solution with (3.96) requires the most time. The solution computed with Algorithm 4, on the other hand, requires the least amount of time, though it is not as good with respect to the objective value as the solution computed with the combination of Algorithm 3 and polishing.

Table 3.4: Results for example in Section 3.4.4.1 with dose–volume objectives.

	PTV D95 (Gy)	PTV % < 81 Gy	OAR % > 30 Gy	Iter. (k)	Time (s)
Solution w/ ($\mathcal{P}_2$)	79.65	55.32	38.35	5	1.83
Solution w/ (3.94)	78.74	44.33	56.55	173	5.20
Solution w/ (3.95)	77.89	43.74	20.63	5	198.70

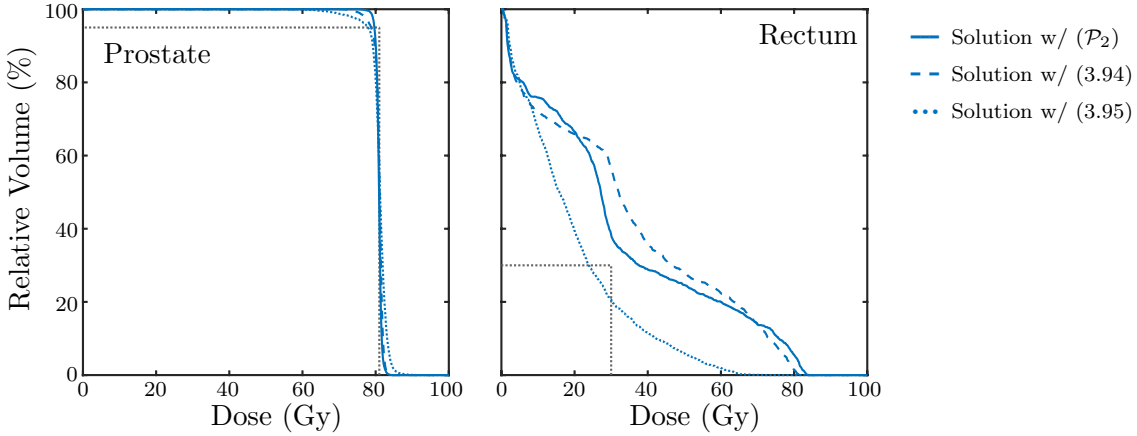

Figure 3.13: Dose–volume histograms for example in Section 3.4.4.1 with dose–volume objectives. (Left) None of the solutions come close to satisfying the lower dose–volume constraint on the PTV, though the solution with ($\mathcal{P}_2$) achieves the highest PTV D95. (Right) The solution with (3.95) satisfies the upper dose–volume constraint on the OAR, with only 20.63% of the rectum volume exceeding 30 Gy, while 38.35% and 56.55% of the rectum volume exceeds 30 Gy for the solutions with ($\mathcal{P}_2$) and (3.94), respectively.

Table 3.5: Results for example in Section 3.4.4.1 with dose–volume constraints.

	PTV D95 (Gy)	PTV % < 81 Gy	OAR % > 30 Gy	Obj.	Time (s)
Solution w/ $(\mathcal{P}_2)^a$	81.32	2.97	29.43	7.80	46.94
Solution w/ $(\mathcal{P}_2)^b$	81.00	4.37	29.43	6.95	65.56
Solution w/ (3.84)	81.00	3.26	28.40	9.12	63.60
Solution w/ (3.96)	81.00	2.20	28.16	8.94	379.46

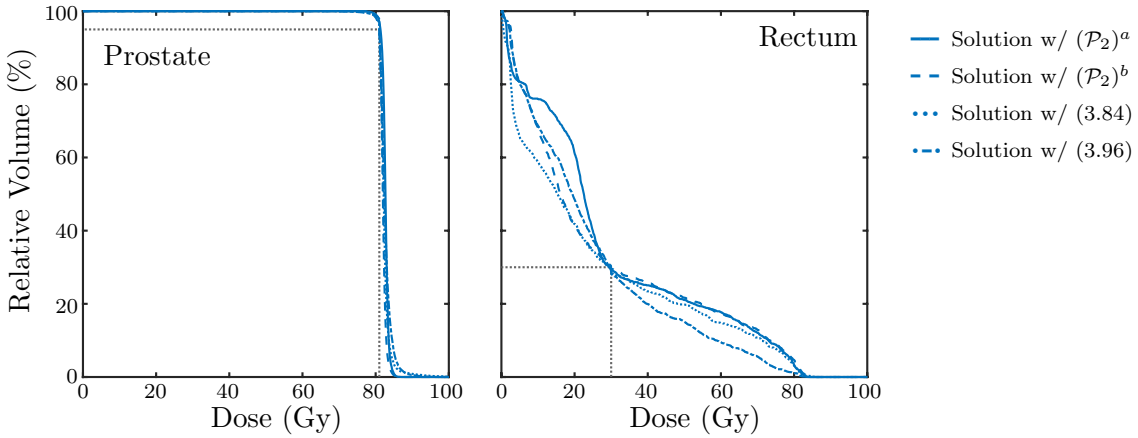

Figure 3.14: Dose–volume histograms for example in Section 3.4.4.1 with dose–volume constraints. (Left) All solutions meet the lower dose–volume constraint on the PTV. (Right) All solutions meet the upper dose–volume constraint on the OAR.

[a] Re-weighted

[b] Polished

3.4.4.2 Multiple PTVs and OARs.

Next we consider the problem from Section 3.4.3 with stopping tolerance $\epsilon = 10^{-2}$, with additional lower dose–volume constraints that no more than 5% of the tumor volumes receives less than 81 Gy and 60 Gy for the prostate and lymph nodes, respectively. In Table 3.6 and Figure 3.15, we see the results of using approximate dose–volume objectives, with patterns similar to those in Table 3.4. Specifically, none of the solutions come close to meeting the lower dose–volume constraints on the PTVs, our solution from Algorithm 3 achieves the highest PTV D95 values and requires the least amount of time, and the solution with (3.95) is the only method to satisfy the upper dose–volume constraints on the OARs, again at the expense of significantly more time. While the method with (3.94) produces a lower dose to the bladder, our method is still the best candidate for re-weighting or polishing in terms of efficiency, where it achieves solutions that meet all dose–volume constraints in less time than it takes the method with (3.95) to compute its approximate solution (first two rows of Table 3.7).

Finally, we see the results for this example computed through either re-weighting or polishing in Table 3.7 and Figure 3.16. This time, the PTV and OAR subvolumes determined through the solution computed with (3.84) does not lead to a feasible solution for $(\mathcal{P}_1)$, though the subvolumes chosen by the other three methods do. Additionally, the method with (3.96) delivers the lowest dose to the OARs, at the expense of a higher objective value and computation time, while the solution found using $(\mathcal{P}_2)$ and re-weighting produces the lowest objective value in the least amount of time. Overall, these results demonstrate that, compared to other convex and nonconvex methods, our approach can efficiently deliver similar or better solutions to the FMO problem in the presence of competing PTV and OAR dose–volume constraints.

Table 3.6: Results for example in Section 3.4.4.2 with dose–volume objectives.

	PTV 1 D95 (Gy)	PTV 1 % < 81 Gy	PTV 2 D95 (Gy)	PTV 2 % < 60 Gy	OAR 1 % > 50 Gy	OAR 2 % > 30 Gy	Iter. (k)	Time (s)
Solution w/ ($\mathcal{P}_2$)	76.95	57.86	57.51	52.95	64.58	49.47	3	2.80
Solution w/ (3.94)	76.57	42.07	56.87	41.18	77.48	44.45	195	12.98
Solution w/ (3.95)	74.18	45.73	57.34	42.25	25.68	27.55	3	2014.94

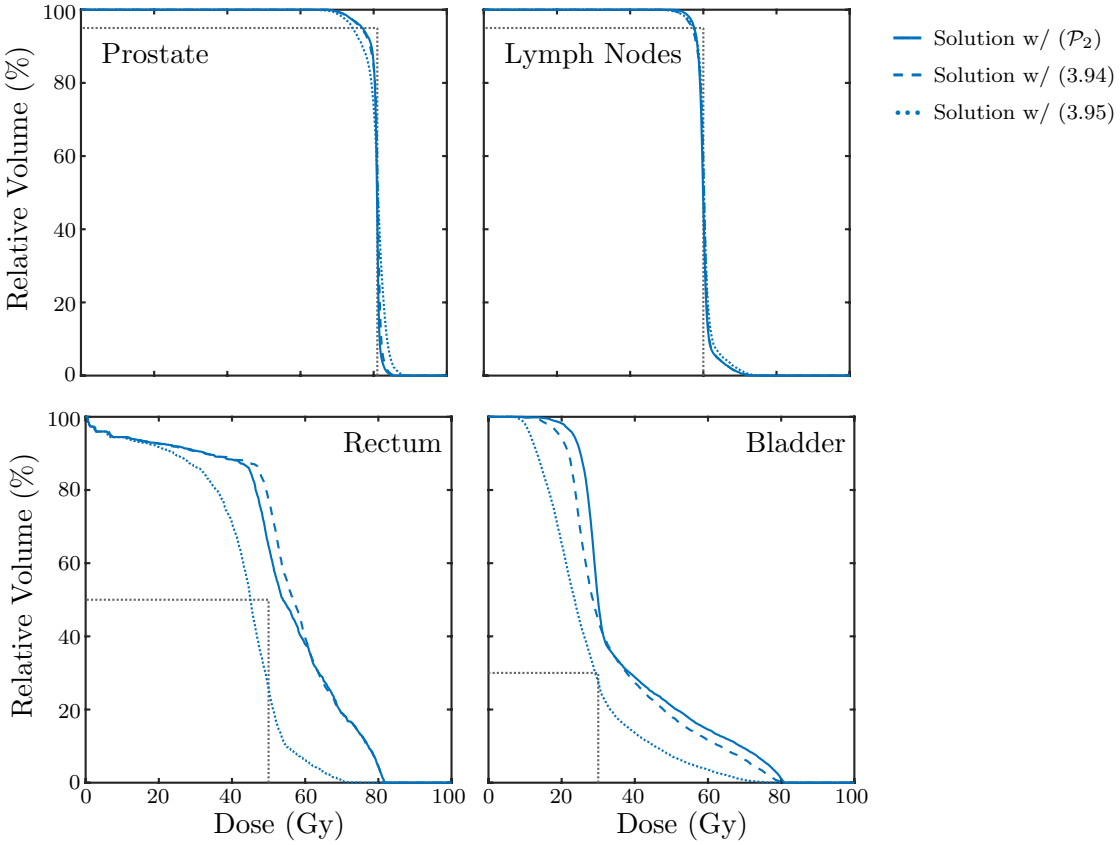

Figure 3.15: Dose–volume histograms for example in Section 3.4.4.2 with dose–volume objectives. (Top) None of the solutions come close to satisfying the lower dose–volume constraints on the PTVs, though the solution with ($\mathcal{P}_2$) achieves the highest PTV D95 values. (Bottom) The solution with (3.95) satisfies the upper dose–volume constraints on the OARs, with only 25.68% of the rectum volume exceeding 50 Gy and 27.55% of the bladder volume exceeding 30 Gy, while 64.58% and 77.48% of the rectum volume exceeds 50 Gy and 49.47% and 44.45% of the bladder volume exceeds 30 Gy for the solutions with ($\mathcal{P}_2$) and (3.94), respectively.

Table 3.7: Results for example in Section 3.4.4.2 with dose–volume constraints.

	PTV 1 D95 (Gy)	PTV 1 % < 81 Gy	PTV 2 D95 (Gy)	PTV 2 % < 60 Gy	OAR 1 % > 50 Gy	OAR 2 % > 30 Gy	Obj.	Time (s)
Solution w/ $(\mathcal{P}_2)^a$	81.17	4.71	60.32	3.90	48.39	29.84	27.71	138.36
Solution w/ $(\mathcal{P}_2)^b$	81.00	4.46	60.01	3.68	44.31	28.78	28.64	309.83
Solution w/ (3.96)	81.00	3.07	60.01	2.69	33.41	25.02	36.02	1118.13

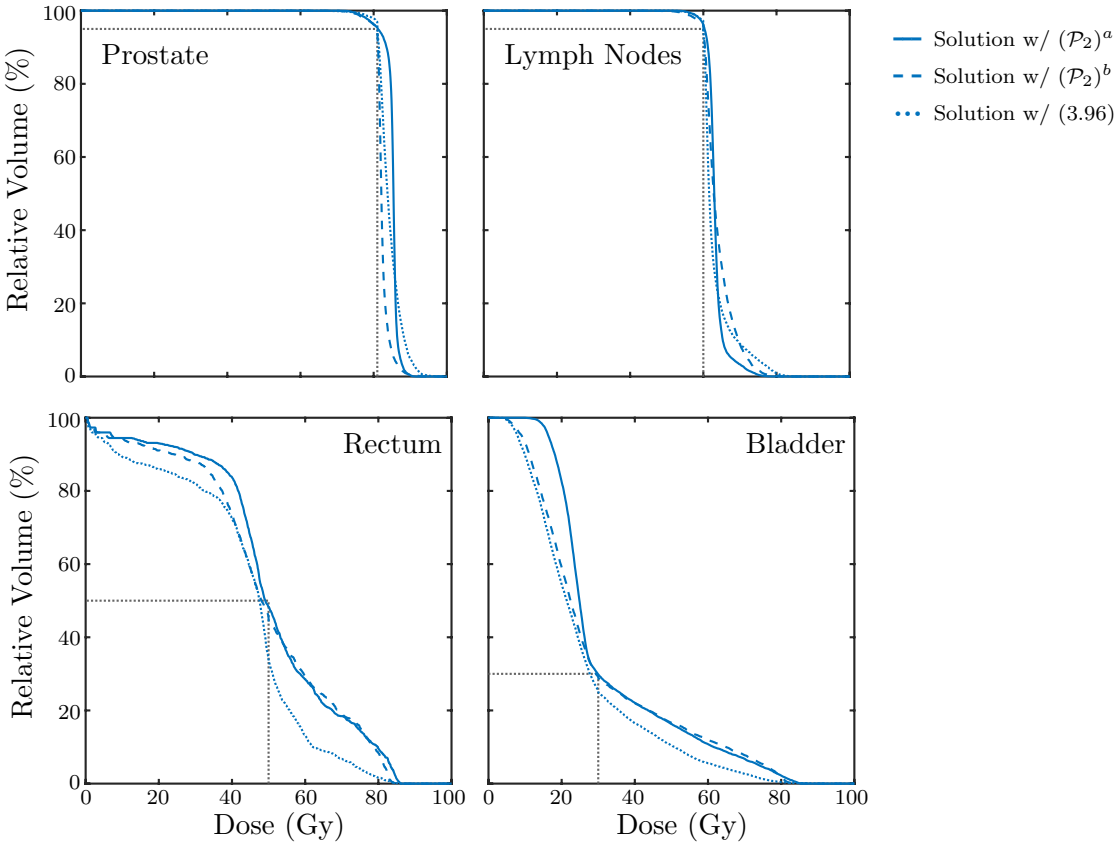

Figure 3.16: Dose–volume histograms for example in Section 3.4.4.2 with dose–volume constraints. (Top) All solutions meet the lower dose–volume constraints on the PTVs. (Bottom) All solutions meet the upper dose–volume constraints on the OARs.

[a] Re-weighted

[b] Polished

3.5 Discussion

The main challenge of radiation treatment planning is to deliver a prescribed dose to the tumor while sparing surrounding healthy tissue. The fluence map optimization problem for intensity-modulated radiation therapy can be formulated as a large-scale inverse problem with competing objectives and constraints on the planning target volumes and organs-at-risk. Unfortunately, the clinically relevant dose–volume constraints are nonconvex, so solving treatment planning problems exactly with dose–volume constraints is NP-hard. We proposed a new approach that is well adapted to handle nonconvex dose–volume constraints without relying on convex relaxations. To solve this formulation, we developed a customized algorithm that is guaranteed to converge to a stationary point of the proposed model. The overall approach is scalable, handles multiple objectives on PTVs and OARs, and returns actionable treatment plans as demonstrated with examples on the CORT dataset.

Because our approach models dose–volume constraints using relaxations, our solutions are not guaranteed to meet each dose–volume constraint exactly. This means that in cases where no feasible solution exists, we can still find treatment plans that balance competing objectives on PTVs and OARs. On the other hand, we also demonstrated that in many cases, treatment plans can be computed efficiently using our approach, then iteratively reweighted or polished to find a feasible solution. In fact, results using Algorithm 4 show that by changing the weights and dose–volume parameters in our objective function, it may be possible to find treatment plans that not only meet a dose–volume constraint, but also lower the dose to the OAR below the constraint value (e.g., Figure 3.9). This suggests that for some cases, tuning the objective function parameters may yield better treatment plans. We explore the problem of optimizing the objection function parameters in the FMO problem in the following chapter.

While this chapter focuses primarily on the methodology of our new approach, future work on this project will be directed towards practical considerations. We will test our

model on additional datasets including different tumor types, patient geometries, and prescriptions. The comparison of the results using our approach and commercial treatment planning systems is left for future work.

Chapter 4

A HYPERPARAMETER OPTIMIZATION APPROACH FOR AUTOMATED RADIOTHERAPY TREATMENT PLANNING

4.1 Introduction

As we saw in the previous chapter, a crucial step of IMRT planning is solving the FMO problem, where inverse planning is used to calculate beamlet intensities that deliver a desired dose distribution to the patient. Using a treatment planning system (TPS), the treatment planner translates a patient's clinical goals into a mathematical objective function, which is then minimized to determine the optimal treatment plan. The objective function typically contains terms corresponding to different clinical goals, such as maximum doses or dose–volume objectives, and each term is assigned a weight factor to specify the relative importance of treating or sparing particular structures. Unfortunately, the treatment plan that minimizes a given objective function may not be clinically acceptable or optimal, either because clinical goals have not been met, or further improvement (e.g., better PTV coverage or lower OAR doses) could still be made. To find a suitable plan, treatment planners often modify the objective function weights through an inefficient trial-and-error approach, where the influence of individual parameters on the resulting treatment plan is not always intuitive and can depend on a variety of factors including model formulation, optimization algorithm, and patient anatomy. This means that plan quality is often determined by the experience of an individual planner and the time available to try different parameter configurations [29, 46].

To improve efficiency and plan quality, many efforts have been made to either guide or automate this "manual planning" step. One approach is to explore the relationships between the objective function parameters and the resulting dose distributions be means of

an effect–volume histogram (EVH) [10] or an analytical sensitivity analysis [7, 88, 155]. In the first case, the EVH can identify organ subvolumes where weight adjustment can have the greatest impact. It can also help determine if goals can be met by either changing weights or including additional objective terms, or if no further improvement can be made due to conflicting objectives. In the second case, a sensitivity analysis of the objective and constraint parameters can quantify the influence of each parameter with respect to the objective function, which can then be used to predict changes to the resulting plan without re-optimization. Of course, this requires knowledge of the specific form of the objective and its Lagrangian function, which may not be available for all TPSs. In this case, an empirical sensitivity analysis can be performed by computing treatment plans for randomly sampled objective function parameters [99]. This sensitivity information can then be used to classify parameters into three sets: a fixed set (often normalized to meet clinical specifications), a sensitive set, and an insensitive set, which allows planners to reduce the dimensionality of the problem by restricting their search to the sensitive set.

Another approach, often referred to as Knowledge-based Planning (KBP), uses a library of previously optimized plans to predict the DVHs, dose objectives, and weight factors for a new patient. For example, the dose distributions and DVHs from reference plans can be used to guide the optimization of new plans via the constraint that OAR doses cannot be worse than the reference [58]. In addition, correlations between the spatial relationships amongst OARs and PTVs, such as what is captured in the overlap–volume histogram (OVH), and the dose delivered in prior plans can be leveraged to develop models to predict expected OAR DVHs from patient geometry. This can in turn be used to initialize planning goals for new patients [178] or to develop quality control tools that can automatically detect suboptimal plans [11]. Features of patient anatomy can also be incorporated into models such as k-nearest neighbors or multinomial logistic regression to predict objective function weights to warm start the manual planning process [33]. While these tools can reduce planning time by

identifying a good starting point for a particular patient, or by indicating when an acceptable plan has been found, they still require the planner to manually tune plans by trial and error. Furthermore, they need access to previously optimized plans with similar patient geometry and objectives, which may not always be available, and the resulting plan quality can be dependent upon the quality of the training set [90].

In an effort to automate the search for objective functions that result in clinically acceptable plans, others view the objective term weights as tunable hyperparameters, taking advantage of recent advances in hyperparameter optimization from the machine learning (ML) community. In the ML setting, hyperparameters can include values related to the structure of a learning model (e.g., regularization weights, number of features, or the architecture of a neural network) or the optimization approach used to train the model (e.g., learning rate, mini-batch size, or number of epochs). This approach seeks to minimize a loss function with respect to its hyperparameters, which can be challenging due the the fact that the function is typically nonconvex, expensive to evaluate (e.g., it may involve training an instance of a neural network), and does not have a closed-form expression or access to its derivatives [172]. In the same way that treatment plan quality often depends on the experience of the planner, the manual tuning of ML hyperparameters can be considered a "black art" requiring expert intuition and heuristics [23, 154].

To improve efficiency, results, and reproducibility, ML researchers have developed a variety of approaches for automated hyperparameter optimization. For a small number of hyperparameters, a naive grid search can be effective, but it suffers the curse of dimensionality and is thus infeasible for most problems [23, 70]. In some cases, when certain continuity and differentiability conditions are satisfied, gradients with respect to continuous hyperparameters can be calculated using back-propagation and automatic differentiation, which can allow for the optimization of a large number of hyperparameters [20, 106]. Two important classes of hyperparameter optimization methods include model-free and model-based approaches.

Some examples of model-free approaches include evolutionary [66, 94] and random searches [23, 95], which are amenable to parallel computing. In contrast, in Sequential Model-based Optimization (SMBO), the loss function is approximated by a surrogate that is cheaper to evaluate, and new hyperparameters are chosen by optimizing an acquisition function based on the surrogate. Examples include Bayesian optimization using Gaussian processes [22, 154] or random forests [78, 164], which we discuss in more detail in Section 4.3.

Using a treatment plan utility function or other clinical criteria to quantitatively compare different treatment plans, the hyperparameter optimization approach for radiotherapy treatment planning iteratively selects new parameters to improve plan quality using various search methods. For example, [16] applied a combination of grid search and linear interpolation to find parameter values that improved PTV coverage while satisfying OAR constraints, [99] developed a recursive random search method to optimize a plan score function based on clinical objectives, [171] used a library of optimized plans with similar patient geometry to guide weight selection and adjustment to improve the plan DVH, [90] used Bayesian optimization to update weights to meet or improve upon dose objectives predicted using KBP, and [150] used deep reinforcement learning to automate the trial-and-error process of tuning weights to lower OAR doses. Other aspects of the treatment planning workflow have also been explored, including a combination of KBP for personalized patient objective selection and an automated search for MLC jaw configurations to lower OAR doses [75].

While many of these studies have focused on automated methods to tune weight factors, it is not clear that these are always the most effective parameters of the FMO objective function to consider. For example, [76, 77] found that increasing the weight of an OAR term was more likely to worsen the PTV coverage than to lower the normal-tissue dose, and that decreases to the OAR dose only occurred for severe dose constraints. Furthermore, they noted that optimization results are often more sensitive to changes in the constraint dose, which tend to contribute quadratically to the objective, as apposed to changes to

the linear weight factors. Rather than increasing the OAR weights, they suggested that the best plans result from objective functions where the OAR weights are equal to or less than the PTV weights and the user-specified dose values for the PTVs and OARs are more stringent than the actual clinical goals. Additionally, equalizing the importance of PTV and OAR weight factors may also reduce the probability of finding local minima in the search space [138], which can be a concern for some formulations involving dose–volume objectives. Of course, the influence of weight factors and dose parameters are determined by the specific model formulation used. However, even in approaches where updating weights and dose parameters is mathematically equivalent, modifying references doses is typically more natural for planners because the weights lack an intuitive clinical interpretation [180].

In this chapter, we explore the feasibility of using hyperparameter optimization to improve treatment plan quality and efficiency for stereotactic body radiation therapy (SBRT) using the RayStation TPS [28]. Specifically, we automatically tune objective function parameters using random and Bayesian search routines implemented in the Python package Scikit-Optimize [71], focusing on objective function dose values rather than weight factors. We compare our automatically generated plans for peripherally located lung tumors with plans created by medically certified dosimetrists at the University of Washington Medical Center. Using intuitive utility functions that can be personalized to reflect individual patient goals, we demonstrate that it is often possible to decrease doses to OARs below clinical goal limits, resulting in treatment plans that are as good as or better than manually-tuned plans. Our approach can be applied to any TPS with scripting capabilities, does not require access to previously optimized plans, and can improve upon manual plans by exploring a larger parameter space than what is typically considered in clinical settings. Furthermore, Gaussian process models built within the Bayesian optimization routine can provide *a posteriori* sensitivity information in cases where access to specific optimization function information is restricted.

The rest of this chapter is organized as follows. We introduce our model formulation in Section 4.2, including information on clinical goals, TPS objective functions, and treatment plan utility functions. Next, we provide details on our solution method in Section 4.3, with a brief overview of Bayesian optimization and Gaussian processes. In Section 4.4 we demonstrate the feasibility of our approach through various numerical examples on lung tumors treated with SBRT using the RayStation TPS. Finally, we conclude the chapter and discuss directions of future work in Section 4.5.

4.2 Problem Formulation

In the IMRT treatment planning workflow, clinical goals are specified by the physician, often in terms of minimum and maximum doses or dose–volume objectives, average doses, or uniform doses. It is then the job of the treatment planner to translate these goals into a mathematical objective function, which is optimized by a TPS to determine beamlet intensities that deliver an acceptable dose distribution to the patient. The TPS often calculates beamlet intensities $x \in \mathbb{R}^m_{\geq 0}$ by solving an FMO problem of the form

$$\underset{x \geq 0}{\text{minimize}} \quad f(x; w, \theta) := \sum_{i=1}^{k} w_i f_i(x; \theta_i), \tag{4.1}$$

where constituent functions f_i penalize beamlet configurations that violate a particular clinical goal. Positive weight factors w_i are used to specify the relative importance of treating or sparing particular structures, while nonnegative dose parameters θ_i are chosen based on prescription values $\gamma_i \geq 0$. Some constituent function types, such as those corresponding to dose–volume objectives, require additional parameters related to volume limits. Examples of constituent functions formulated with quadratic penalties are given in Table 4.1, where $d_j : \mathbb{R}^m_{\geq 0} \to \mathbb{R}^{n_j}_{\geq 0}$ maps the m beamlet intensity values to the n_j voxel doses for the jth region of interest (ROI). Depending upon the TPS, the specific penalties used for each type of constituent function may not be available to the planner.

Table 4.1: Examples of constituent functions using quadratic penalties.

ROI	Clinical Goal	Constituent Function
PTV_j	Uniform dose $= \gamma_i$	$\|d_j(x) - \theta_i\|_2^2$
PTV_j	Minimum dose $\geq \gamma_i$	$\|[\theta_i - d_j(x)]_+\|_2^2$
OAR_j	Maximum dose $\leq \gamma_i$	$\|[d_j(x) - \theta_i]_+\|_2^2$

If the treatment plan that minimizes a given objective function is not clinically acceptable, the planner will update the objective by modifying weights or parameters, adding constituent functions, or even creating optimization-only ROIs (often corresponding to hot or cold spots) [76]. After re-optimizing with the new objective, the planner will again evaluate the plan based on their clinical goals and experience. Due to time constraints, this iterative process often ends once an acceptable plan is identified, even if more improvements could be made. To automate this trail-and-error process for improved efficiency and treatment plan quality, we formulate it as a hyperparameter optimization problem of maximizing a treatment plan utility function with respect to the objective function parameters.

For a fixed set of k constituent functions, let $w \in \mathbb{R}^k_{>0}$ denote the vector of weight factors and $\theta \in \mathbb{R}^k_{\geq 0}$ the vector of dose parameters. Viewing the results of the TPS as a black-box function, we let $x : \mathbb{R}^k_{>0} \times \mathbb{R}^k_{\geq 0} \to \mathbb{R}^m_{\geq 0}$ return the solution to (4.1) for a given w and θ. We define the treatment plan utility function $g(x(w, \theta); \gamma)$ to quantify treatment plan quality with respect to the clinical goal reference doses encoded in the vector $\gamma \in \mathbb{R}^k_{\geq 0}$. Our hyperparameter optimization problem can then be expressed as

$$\underset{w > 0, \theta \geq 0}{\text{maximize}} \quad g[x(w, \theta); \gamma] \tag{4.2a}$$

$$\text{subject to} \quad x(w, \theta) \in \underset{x \geq 0}{\arg\min}\, f(x; w, \theta). \tag{4.2b}$$

Depending on the patient, the planner may choose to optimize over only a subset of the

weight or dose vectors. For example, treatment plans are often normalized to meet specific prescriptions, such as the PTV D95, or the planner might know from experience that solutions are relatively insensitive to changes in a given variable. In this case, the dose parameters θ_i in these constituent functions could be set to their corresponding goal values γ_i.

4.3 Algorithmic Approach

The optimization problem given in (4.2) is challenging for a number of reasons, many of them due to the fact that the utility function depends on the results of the TPS. First, there is no closed-form expression for g or its derivatives with respect to w or θ. This means that there is no *a priori* information about the properties of the function (e.g., continuity or convexity) or the influence of the decision variables. Furthermore, each evaluation of g requires the TPS to solve an instance of (4.1), which can be expensive. Depending upon the difficulty of the case and the settings of the TPS solver (e.g., stopping tolerance or maximum number of iterations), calculating the beamlet intensities typically takes 3-5 minutes. While this is nothing compared to the time it takes to train a deep neural network, the treatment planner has a fixed time budget and multiple cases to consider. Finally, the values sampled from the utility may be noisy. This uncertainty can arise from various sources, such as the TPS solver failing to converge to a solution before reaching the maximum number of iterations, or from the dose calculation algorithm used to approximate the resulting dose distribution [43].

These challenges make (4.2) a good candidate for a Bayesian optimization approach, where the search for good parameters is guided by a surrogate model that is sequentially queried and refined, as described in Algorithm 5. Specifically, starting with a prior over our utility landscape, we iteratively sample from the utility function and refine our surrogate using a Bayesian posterior update. To determine which points to sample (i.e., which parameters to use next in the TPS objective function), we optimize an acquisition function that quantifies the benefit of sampling from different locations. There are many choices for

Algorithm 5 Bayesian optimization for $g(w, \theta)$ in (4.2).

Specify surrogate model, acquisition function, and number of iterations T.

for $t = 1, \ldots, T$ **do**

Optimize acquisition function to determine $(w^{(t)}, \theta^{(t)})$.

Sample $g[x(w^{(t)}, \theta^{(t)}); \gamma]$ by solving (4.1) with TPS.

Update surrogate model using data $\{g[x(w^{(\ell)}, \theta^{(\ell)}); \gamma]\}_{\ell=1,\ldots,t}$.

end for

return $\max_{\ell=1,\ldots,T}\{g[x(w^{(\ell)}, \theta^{(\ell)}); \gamma]\}$

acquisition functions, such as Thompson sampling, probability of improvement, expected improvement, upper confidence bounds, and entropy search [148]. In essence, these functions are designed so that their optima reflect the dual objectives of exploration (sampling where the surrogate's uncertainty is high) and exploitation (sampling where the predicted utility is high). Through this approach, we can simultaneously learn about the sensitivity of treatment plan quality with respect to the objective function parameters via the posterior surrogate model, while also identifying parameters that result in clinically superior treatment plans. For a more detailed overview of Bayesian optimization and its applications, see [2, 82, 113, 148].

In our implementation, we approximate the utility function using the nonparametric Gaussian process surrogate model. A Gaussian process, defined in terms of mean and covariance functions, is a generalization of a multivariate Gaussian probability distribution, which is specified in terms of a mean vector and covariance matrix. While a Gaussian distribution models random scalars or vectors, a Gaussian process describes a distribution over functions. For the treatment planning problem, our Gaussian process is a collection of random treatment plan utility variables $g(w, \theta)$, any finite subset of which are jointly Gaussian

distributed. In particular, for any input (w, θ), there is an associated random variable $g(w, \theta)$ with mean $\mu(w, \theta)$ and standard deviation $\sigma(w, \theta)$. For simplicity, the mean function is initially assumed to be zero, and the covariance is defined by a covariance function or kernel $k[(w, \theta), (w', \theta')]$, which specifies assumptions about the similarity between nearby points. Popular kernel choices include radial basis functions, Matérn covariance functions, rational–quadratic kernels, exp–sine–squared kernels, and dot product kernels [126]. For more details about Gaussian processes and kernel functions, see [64, 123, 131].

4.4 Numerical Examples

In this section, we demonstrate the feasibility of using hyperparameter optimization to automate the search for objective function parameters for SBRT with the RayStation TPS. Using Bayesian optimization routines from Scikit-Optimize, we automatically tune objective function dose values to calculate beamlet intensities for five lung tumor cases treated at the University of Washington Medical Center. For all cases, the objective function in (4.1) consists of eleven constituent functions corresponding to the clinical goals given in Table 4.2, chosen according to RTOG 0915 [169], with tunable weights w and dose parameters θ. While we leave the volume values of the dose–volume objectives fixed, these could also be tuned using our approach. We treat the PTV D95 dose level as a fixed parameter, with all treatment plans normalized so that exactly 95% of the PTV receives 48 Gy for the first four cases and 50 Gy for the fifth case.

Our treatment plan utility function in (4.2) is defined as a linear combination of terms corresponding to each clinical goal,

$$g[x(w, \theta); \gamma] := \sum_{i=1}^{11} g_i[x(w, \theta); \gamma_i], \tag{4.3}$$

though the term corresponding to the PTV D95 could be omitted due to normalization. We

Table 4.2: Clinical goals used in numerical examples.

i	ROI	Type	Dose (γ_i)	Volume
1	Chestwall	MaxDvh	30 Gy	30 cm^3
2	D2cm	MaxDose	24-30[1] Gy	
3	Esophagus	MaxDvh	18.8 Gy	5 cm^3
4	Lungs	MaxDvh	11.6 Gy	1500 cm^3
5	Lungs	MaxDvh	12.4 Gy	1000 cm^3
6	PTV	MinDvh	48-50[2] Gy	95%
7	PTV	MaxDose	80 Gy	
8	Ribs	MaxDvh	32 Gy	1 cm^3
9	Ribs	MaxDose	40 Gy	
10	Spinal Cord	MaxDvh	13.6 Gy	1.2 cm^3
11	Spinal Cord	MaxDose	26 Gy	

use a combination of linear and linear–quadratic utility terms to reward plans that decrease doses below (or increase doses above) the clinical goal limits. Our linear utility terms can be interpreted as percent decrease (or increase) relative to the reference dose values. For example, we use linear terms

$$g_i^{\ell}[x(w,\theta);\gamma_i] = \frac{100\{\gamma_i - d_j^{\max}[x(w,\theta)]\}}{\gamma_i} \tag{4.4}$$

for ROIs with maximum-dose objectives, depicted as the solid blue line in Figure 4.1, where $d_j^{\max}$ calculates the maximum dose delivered to the jth ROI. While the linear utility terms have an intuitive interpretation, they do not distinguish between doses that meet a clinical

[1]Determined by PTV volume.

[2]Dose level γ_6 is 48 Gy for cases 1–4 and 50 Gy for case 5.

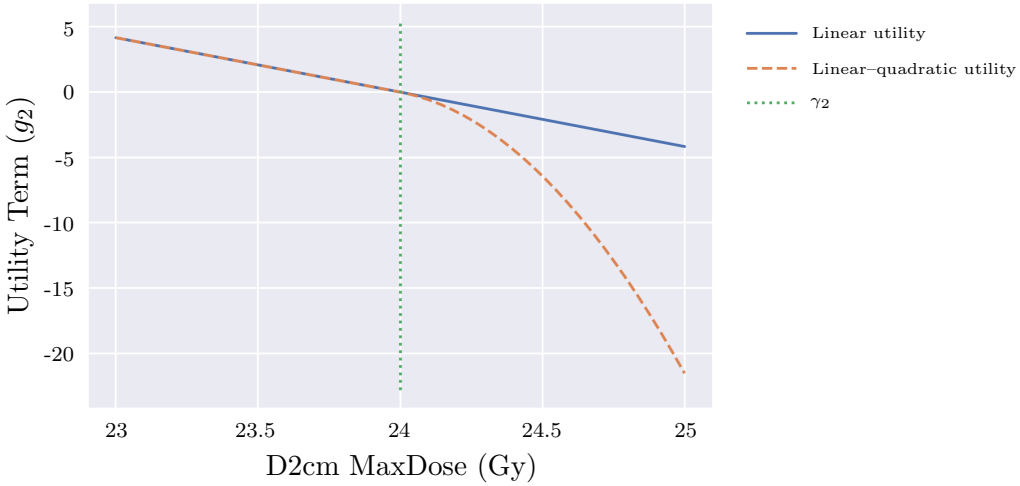

Figure 4.1: Linear and linear–quadratic utility terms corresponding to a maximum-dose clinical goal of 24 Gy for the D2cm, an optimization-only ROI that includes all voxels within 2 cm of the PTV.

goal from doses that do not. Specifically, a decrease to doses above γ_i is rewarded the same as an equal decrease to doses below γ_i (i.e., the slope of the utility is the same). For clinical goals with higher priority, we can instead use a linear–quadratic utility,

$$g_i^{\ell q}[x(w,\theta);\gamma_i] = \begin{cases} g_i^{\ell}[x(w,\theta);\gamma_i] & d_j^{\max}[x(w,\theta)] \leq \gamma_i, \\ \left\{1 - g_i^{\ell}[x(w,\theta);\gamma_i]\right\} g_i^{\ell}[x(w,\theta);\gamma_i] & d_j^{\max}[x(w,\theta)] > \gamma_i, \end{cases} \tag{4.5}$$

depicted as the dashed orange line in Figure 4.1, which prioritizes meeting these goals over decreasing the doses to ROIs that have already met their respective goals.

Our examples are implemented with RayStation 8B, which provides scripting capabilities in CPython 3.6.5. A call to our utility function updates the RayStation objective function weight factors and dose parameters, runs RayStation's optimization routine, retrieves the resulting ROI doses, and returns the calculated utility value. Our hyperparameter optimization problem is solved using the SMBO package Scikit-Optimize, which is built upon Python

packages NumPy, SciPy, and Scikit-Learn [67, 170, 126]. We compare the results of uniform random sampling to Bayesian optimization based on Gaussian processes using the functions `dummy_minimize` and `gp_minimize`, respectively, though we note that Scikit-Optimize also provides methods that utilize decision trees (e.g., random forest and extra trees regressors) and gradient-boosted trees. Except for the number of iterations and initial random points (used to initialize the surrogate model), we use the default settings for `gp_minimize`. This includes a Matérn kernel with hyperparameters (e.g., length scale, covariance amplitude, and additive Gaussian noise) automatically tuned, and the acquisition function `gp_hedge`, which randomly selects either the lower confidence bound, expected improvement, or probability of improvement acquisition function each iteration. While the default settings yield promising results, further improvements could be made by modifying the kernel and acquisition function settings.

We first explore the relative sensitivity of ROI doses with respect to the dose parameters and weight factors in Section 4.4.1, which motivates our choice to tune the dose values exclusively. Next, we illustrate the behavior of random and Bayesian search methods with reference to grid search results on a small 2-D problem in Section 4.4.2. The full 10-D problem is implemented in Section 4.4.3, with comparisons to clinical treatment plans. Finally, we demonstrate the potential benefits of dimensionality reduction in Section 4.4.4.

4.4.1 Dose sensitivity with respect to weight factors and dose parameters

For our first example, we demonstrate the relative sensitivity of ROI doses with respect to weight factors and dose parameters. This is done by first fixing all weights $w_i = 1$ and dose parameters $\theta_i = \gamma_i$. Next, we compute treatment plans by separately varying the weight factor w_2 and dose parameter θ_2 for the D2cm, an optimization-only ROI that includes all voxels within 2 cm of the PTV.

In general we would expect the maximum dose to the D2cm to decrease as we increase

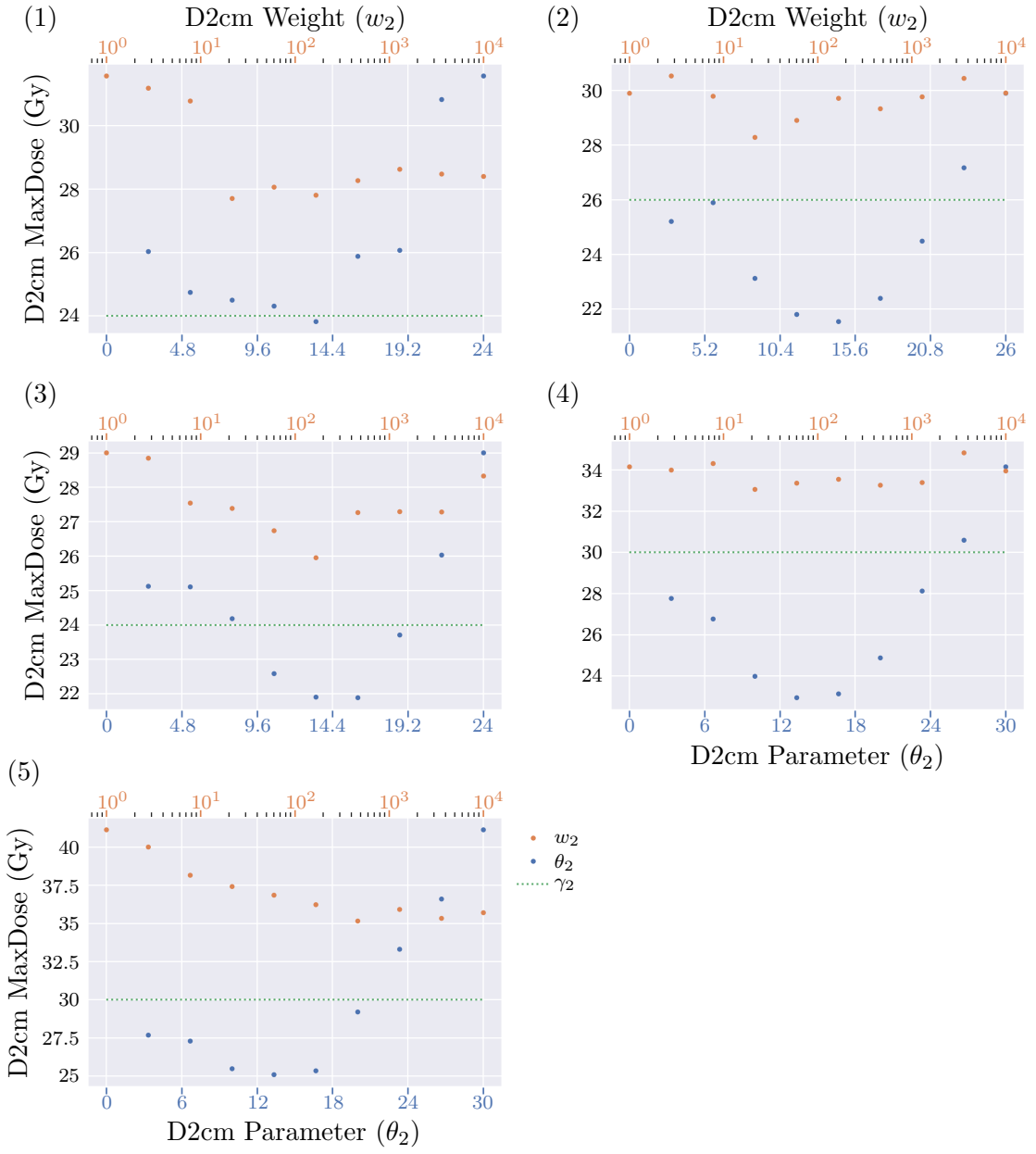

Figure 4.2: D2cm maximum-dose sensitivity with respect to objective function weight factors and dose parameters for all five cases. None of the plans computed by varying w_2 (orange dots) satisfy the maximum-dose clinical goal for the D2cm, while varying θ_2 (blue dots) leads to treatment plans that both meet the goal and decrease the dose further below γ_2 (green dotted lines).

w_2 and decrease θ_2; however, this doesn't appear to be the case. In Figure 4.2, we plot the results of varying the D2cm weight (orange dots) and dose parameter (blue dots) for all five patients. The maximum allowed dose γ_2, denoted by the green dotted line, is chosen for each patient based on the PTV volume. As we increase w_2, the resulting D2cm maximum doses behave differently for each of the five patients. For example, the dose in case 5 steadily decreases as w_2 increases, while the dose in case 2 doesn't seem to follow any particular pattern. Notably, none of the plans computed by varying w_2 satisfy the maximum-dose clinical goal for the D2cm. Alternatively, as we decrease θ_2 the maximum dose to the D2cm initially decreases, then begins to increase as the objective becomes too difficult for the TPS. This highlights the importance of identifying a good parameter range, because unrealistic parameters can lead to worse treatment plans. In all cases, varying θ_2 leads to treatment plans that meet the maximum-dose clinical goal for the D2cm, and decreases the dose further below γ_2 for cases 2–5. Based on the results of this section, and in order to keep our problem dimension low, we fix all weight factors at $w_i = 1$ and tune only the dose parameters θ_i for the remainder of our examples.

4.4.2 Comparison of search methods and utility functions

While an exhaustive grid search is simple to implement, it is infeasible for most practical problems because the number of sample points increases exponentially with the number of hyperparameters [23, 70]. In order to compare our Bayesian search method and uniform random sampling with a "ground truth" grid search, in this section we consider a toy problem with only two tunable parameters. Using our first lung case, we again vary the D2cm maximum-dose parameter θ_2, and we add an additional dimension by varying the ribs maximum-dose parameter θ_9. All weights are fixed at $w_i = 1$, and all remaining dose parameters are set to their limit values $\theta_i = \gamma_i$.

For our 2-D grid search, we calculate treatment plans for parameter pairs (θ_2, θ_9) over

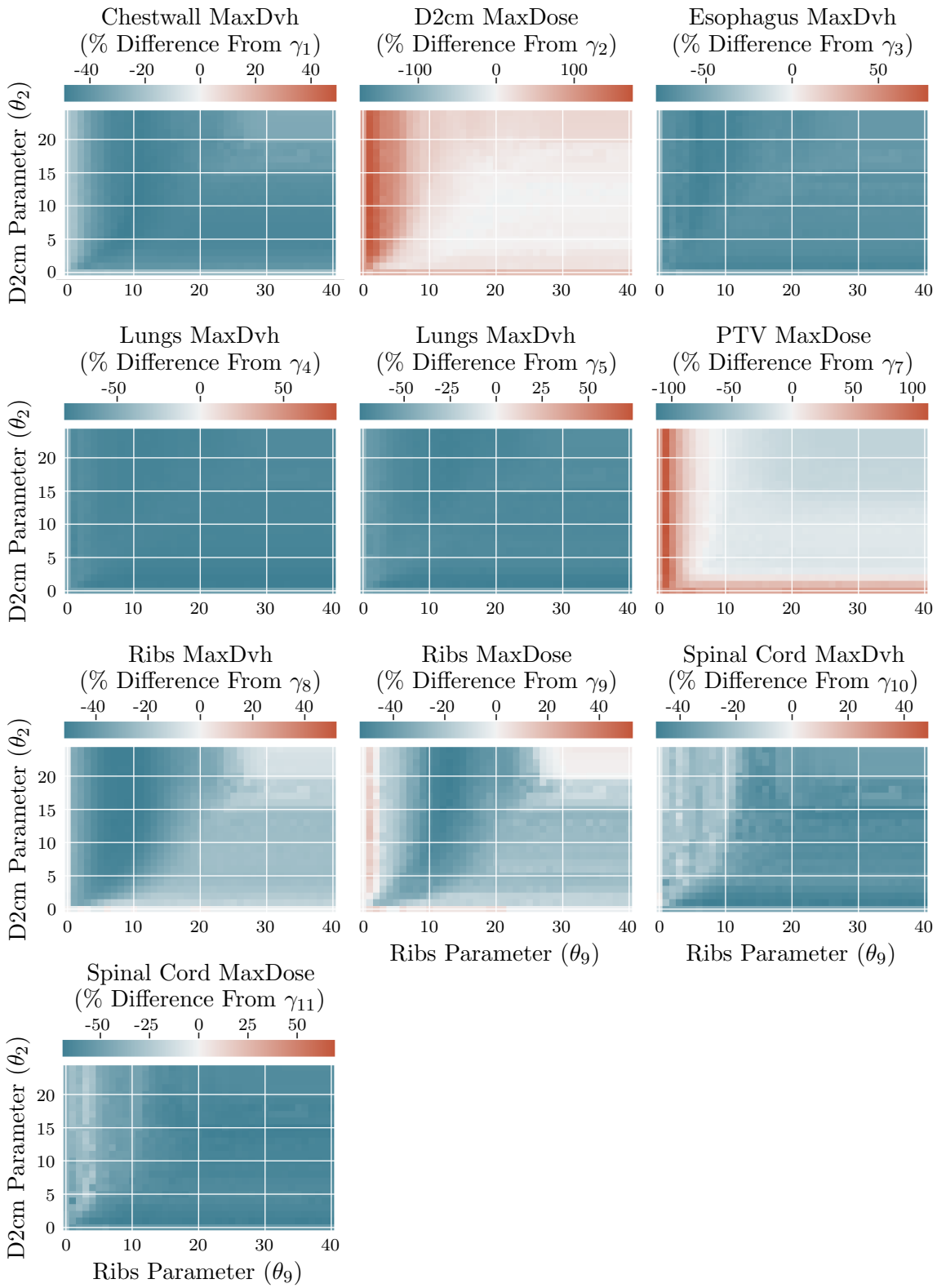

Figure 4.3: ROI doses corresponding to clinical goals computed by varying D2cm and ribs maximum-dose parameters θ_2 and θ_9. While the chestwall, esophagus, lungs, and spinal cord meet their respective clinical goals, the D2cm, PTV, and ribs have regions where their maximum-dose clinical goals are not met.

the range $[0, \gamma_2] \times [0, \gamma_9]$ with 1 Gy increments. In Figure 4.3 we see the resulting dose values for each clinical goal, excluding the PTV D95, plotted in terms of percent difference from γ_i. The clinical goals for the chestwall, esophagus, lungs, and spinal cord are met in all treatment plans , with doses far below γ_i in most cases. In contrast, the D2cm, PTV, and ribs are more affected by changes to θ_2 and θ_9, with regions where the clinical goals are not met. The trade-offs between meeting these different ROI goals will presumably impact on the shape of the utility function.

4.4.2.1 Linear utility functions

We first consider the problem of maximizing the linear utility function using random sampling and Bayesian optimization with Gaussian processes. Each method is run for 50 iterations, where the first 10 iterations of the Bayesian routine correspond to the first 10 random samples. In the first row of Figure 4.4, we plot the random and Bayesian samples over the utility landscape calculated using the grid search results, along with the convergence of the two search methods. While the Bayesian approach converges to a maximum utility of 465.94 after 28 iterations, compared to the grid search optimum of 467.58, the random search takes 37 iterations to converge to a lower utility of 451. In the second row of Figure 4.4, we plot the mean and standard deviation of the Bayes posterior surrogate function. The mean function is qualitatively similar to the utility landscape, though naturally it has more accuracy and less uncertainty near the sampled points. Finally, in Figure 4.5 we include the doses delivered in the optimal treatment plans, expressed as the percent difference from γ_i. All of the clinical goals are met with the exception of the D2cm maximum dose, where the lowest and highest doses are achieved in the Bayesian and random solutions, respectively. Based upon the location of the optima and the ROI doses, it appears that the trade-off specified by the linear utility function favors lowering the dose to the ribs over meeting the maximum-dose limit for the D2cm.

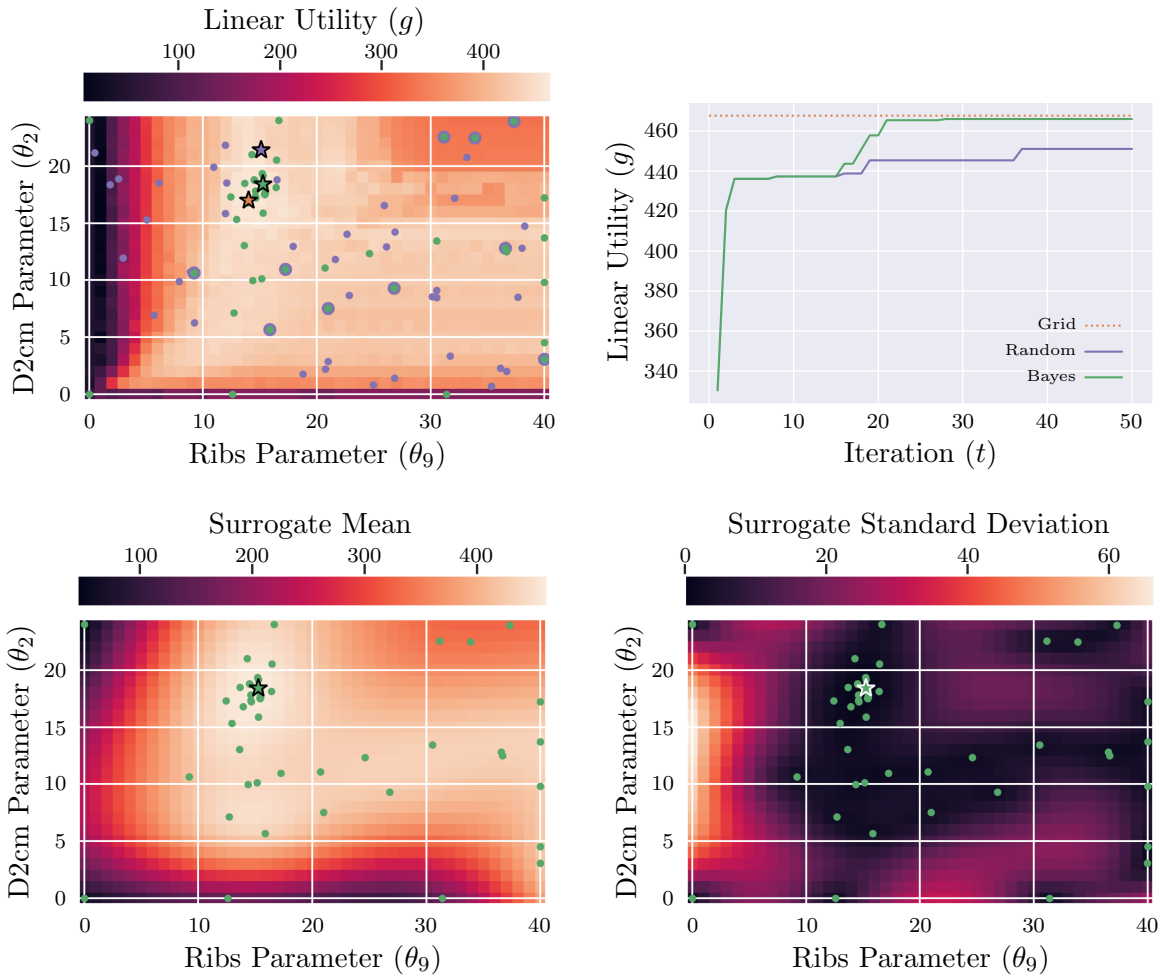

Figure 4.4: Comparison of 2-D grid search, random sampling, and Bayesian optimization using the linear utility function. (Top) Random (purple) and Bayesian (green) samples are plotted over the utility landscape calculated with the grid search plans on the left, with convergence results on the right (optimal grid search utility in orange). (Bottom) Bayesian samples are plotted over the Bayes posterior surrogate function mean and standard deviation.

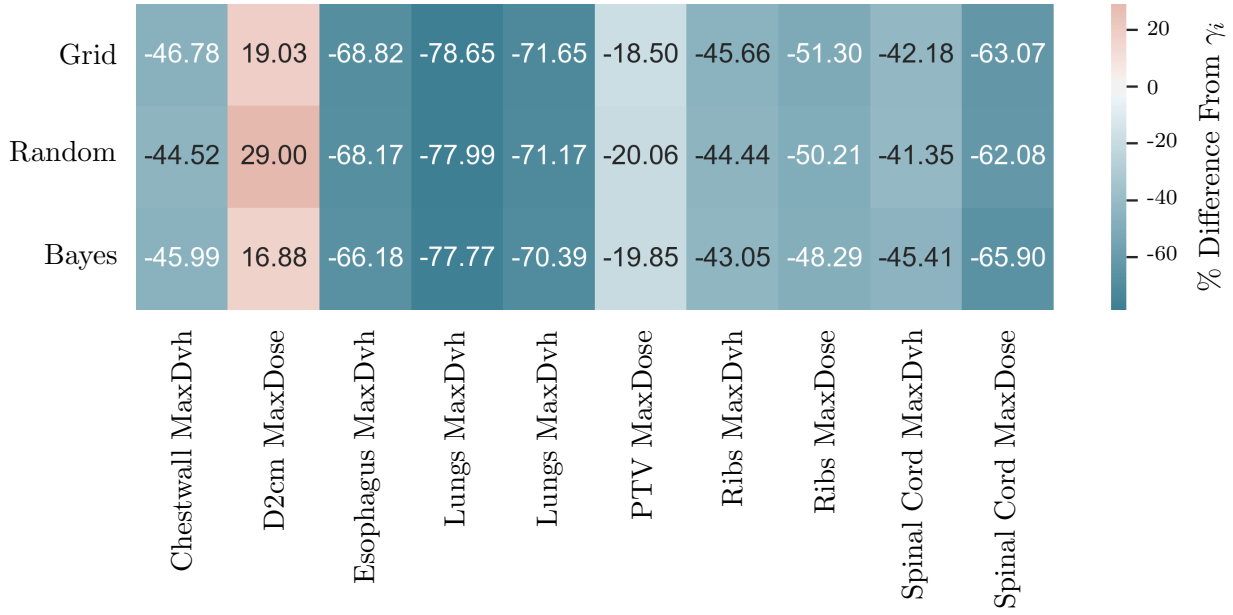

Figure 4.5: Doses delivered in the plans calculated with the linear utility function using grid search, random sampling, and Bayesian optimization, expressed as percent difference from γ_i.

4.4.2.2 Linear–quadratic utility functions

To modify our utility function so that it emphasizes meeting the most important clinical goals, we use a combination of the linear and linear–quadratic utility terms. Specifically, we use linear–quadratic utilities for all ROIs except for the chestwall and ribs, which are considered lower priority in the sense that doses may exceed their goal limits for the sake of an otherwise superior plan. For these two ROIs we use linear utility terms. In the first row of Figure 4.6, we plot the iterates of our random sampling and Bayesian optimization on top of the linear–quadratic utility landscape calculated from our grid search. By setting the lower limit of our color range to zero, we see that the region of positive utility mimics the shape of the space where the D2cm satisfies its maximum-dose clinical goal in Figure 4.3. This demonstrates the ability of the linear–quadratic utility to better prioritize meeting important clinical goals.

For this case, the best plan identified by the grid search has a utility of 445.13, while random sampling achieves an optimum of 442.32 after 28 iterations, and Bayesian optimization produces a maximum utility of 442.98 after 47 iterations. The second row of Figure 4.6 contains the mean and standard deviation functions of the Bayes posterior surrogate model. On the left, we see that the Bayesian method was able to identify a region of positive utility similar to the D2cm feasible region, which enabled it to avoid sampling bad parameter values in contrast to the random search method. We compare the doses delivered to each ROI in Figure 4.7, where the maximum dose to the D2cm is reduced as compared to the plans generated with the linear utility function. In particular, the Bayesian method produces a plan that meets all of the clinical goals. We also see that while the grid search optimum has a higher utility, it is achieved through a different trade-off than that found by the other two plans. Specifically, the treatment plan identified though the grid search exchanges a higher D2cm dose for lower rib doses, which does not correspond to our clinical intentions. This illustrates the importance of the utility function design, and the difficulty of translating clinical intuition into a mathematical objective function.

4.4.3 Comparison with clinical plans

After demonstrating how our approach works for a small 2-D problem, we now consider the 10-D problem where each of the dose parameters θ_i are tuned over the range $[\gamma_i/4, \gamma_i]$, with the exception of the fixed PTV D95 parameter $\theta_6 = \gamma_6$. The lower bound of $\gamma_i/4$ is chosen empirically based on the observation in Section 4.4.1 that unreasonably low parameters can lead to worse treatment plans. Because the PTV maximum-dose parameter θ_7 cannot be less than the PTV D95 parameter, it is instead tuned over the range $[(\gamma_7 + 3\gamma_6)/4, \gamma_7]$. As before, all weight factors are fixed at $w_i = 1$, and all terms in the utility function use the linear–quadratic utility, except for those corresponding to the chestwall and ribs, where we use the linear utility. For each case, we run random and Bayesian search for 100 iterations, where

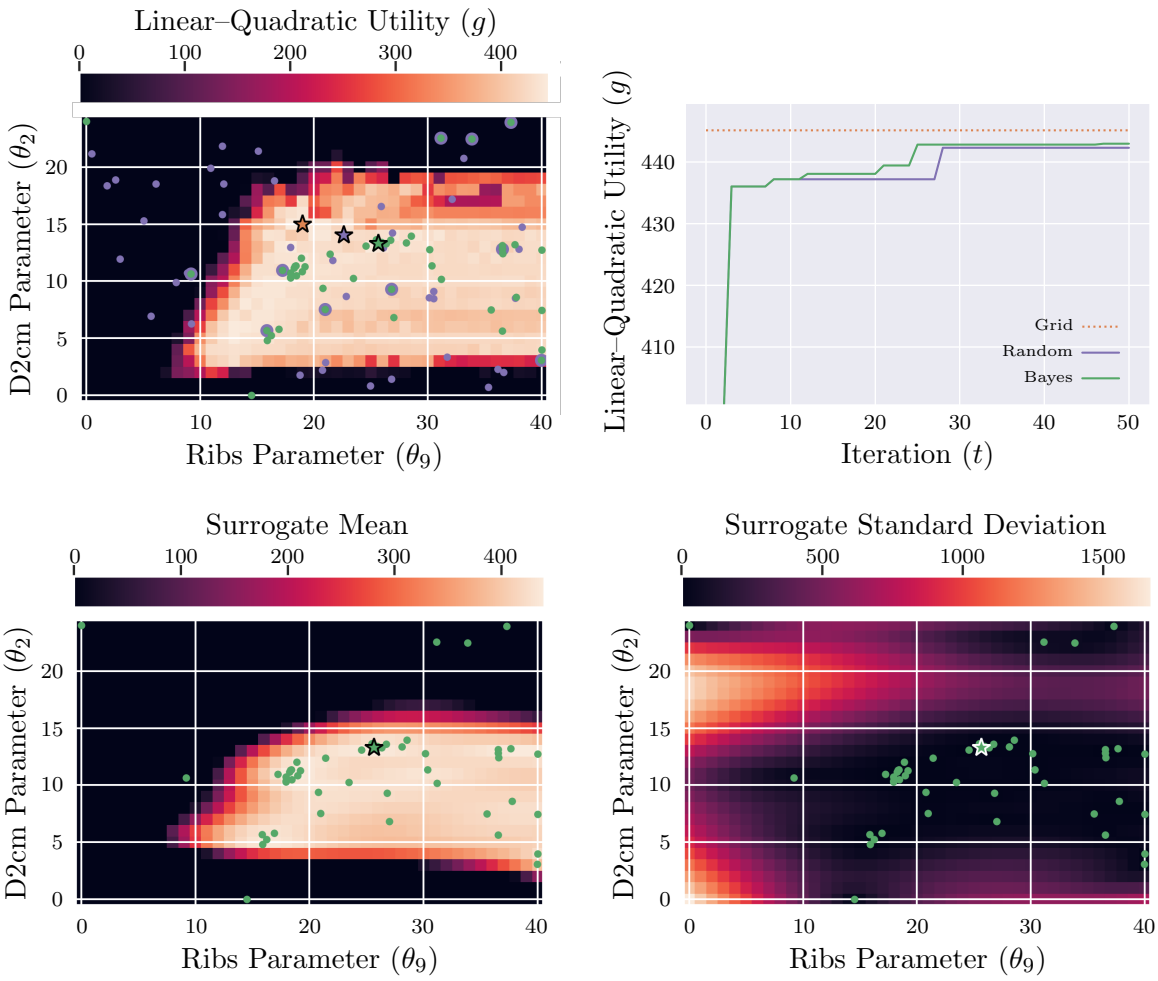

Figure 4.6: Comparison of 2-D grid search, random sampling, and Bayesian optimization using the linear–quadratic utility function. Lower limit of the color range is set to zero to demonstrate the shape of the positive utility region. (Top) Random (purple) and Bayesian (green) samples are plotted over the utility landscape calculated with the grid search plans on the left, with convergence results on the right (optimal grid search utility in orange). (Bottom) Bayesian samples are plotted over the Bayes posterior surrogate function mean and standard deviation.

Figure 4.7: Doses delivered in the optimal plans calculated with the linear–quadratic utility function using grid search, random sampling, and Bayesian optimization, expressed as percent difference from γ_i.

the first 20 iterations of the Bayesian routine correspond to the first 20 random samples. We compare the resulting optimal plans with the doses from clinical plans manually generated by a certified medical dosimetrist.

In Figure 4.8, we plot the convergence results along with the utility evaluated for the clinical and default plans, where all dose parameters are set to their clinical goal limits, i.e., $\theta_i = \gamma_i$. For most cases, the default utility is worse than the clinical utility; the utility is particularly low for the case 4 clinical plan because the spinal cord maximum dose–volume value is 33.85% above γ_{10} (see Figure 4.9), where a linear–quadratic utility term is used. After tuning the objective function dose parameters, the plans computed with Bayesian optimization achieve the highest utility, on average 4.82% higher than the plans computed with random sampling and 39.4% higher than the clinical utilities (median 11.58%). The random search method takes between 1.31–2.88 hours to complete 100 iterations, with a

mean of 1.99 hours total computation time. This is on average 29.9 minutes faster than the Bayesian routine, which takes between 1.81–3.64 hours with a mean of 2.53 hours total. This is most likely due to the fact that the latter method also computes the Bayes posterior surrogate model and optimizes the acquisition function every iteration.

Of course, a higher utility only matters if the resulting plans have improved with respect to the clinical goals. To evaluate plan quality, we include the ROI doses for each optimal plan as a percent difference from γ_i (Figure 4.9) and as a percent difference from the clinical values (Figure 4.10). For all cases, the Bayesian method meets the clinical goals for every term using the linear–quadratic utility. For cases 1–4, the automatically generated plans deliver up to a 76.75% lower dose than the clinical plans for all ROI goals that are violated in the latter. Where the ROI doses were larger in the automatically generated plans, they were still lower than the clinical goal limit γ_i by 1.64–98.87%, indicating that they were clinically acceptable. For case 5, a trade-off is observed between the spinal cord dose (lower in the Bayesian plan) and the rib dose (lower in the clinical plan), a result of using a linear–quadratic utility for the spinal cord and a linear utility for the ribs. In Figure 4.11, we plot the dose–volume histograms for each case, with dotted lines for the clinical plans, dashed lines for the plans computed through random sampling, and solid lines for the plans identified from Bayesian optimization. The most consistent trade-off made by the automatically generated plans, as compared to the clinical plans, appears to be an increased dose to the PTV in exchange for a decreased dose to the spinal cord.

Finally, in Figure 4.12 we look at the distribution of optimal dose parameters θ_i, expressed as percent difference from γ_i, found for all automatically generated plans. We first observe that many of the optimal plans are achieved with parameters on the lower boundary of their search ranges, which is often much lower than the parameters values typically considered by a manual planner. This confirms the observation made in [72] that more flexibility in the search space can lead to improved treatment plans. The optimal parameters with the widest

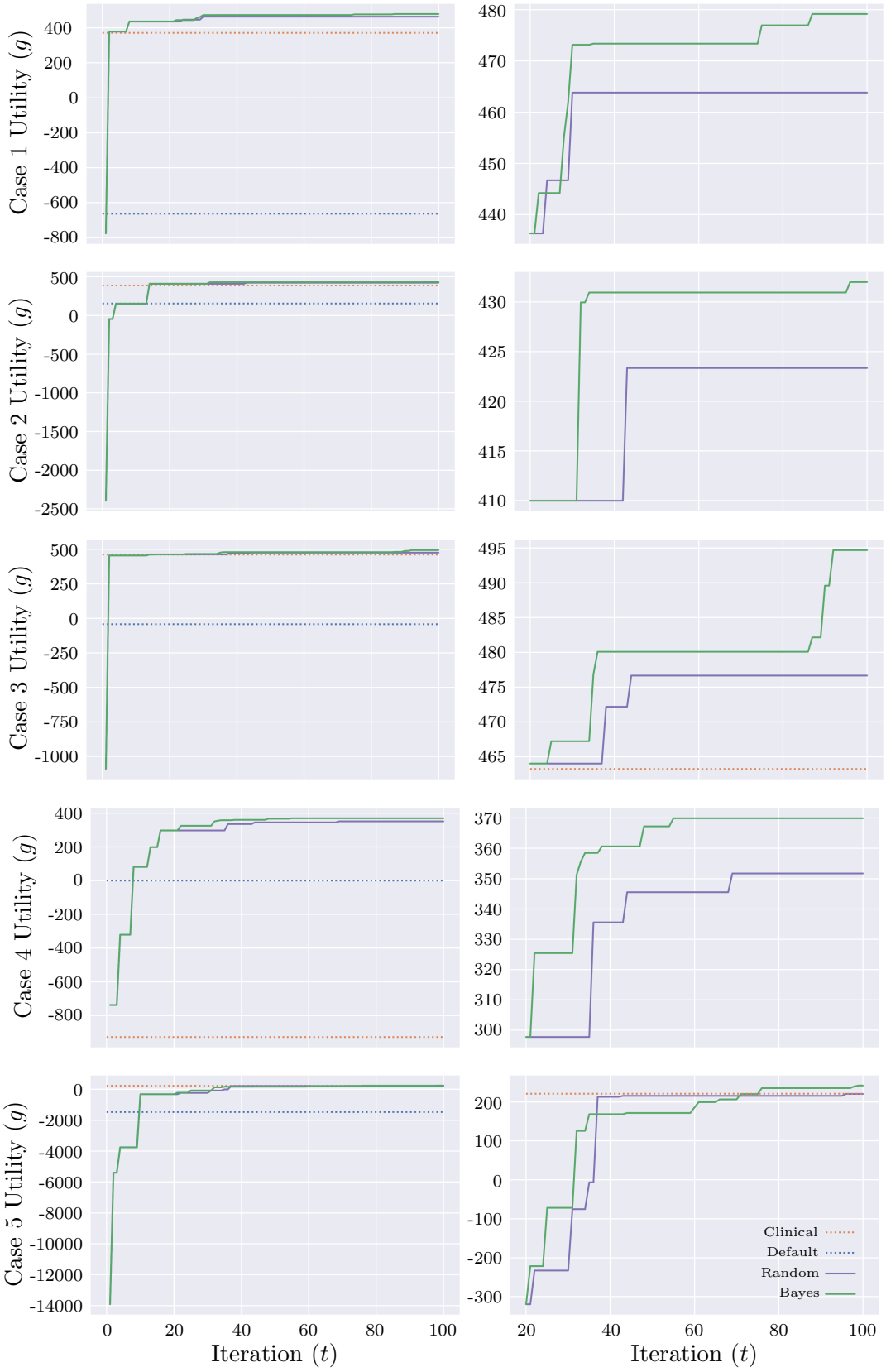

Figure 4.8: Convergence results for 10-D random (purple) and Bayesian (green) search methods, along with the clinical (orange) and default (blue) plan utilities. For all five cases, the Bayesian method achieves the highest utility.

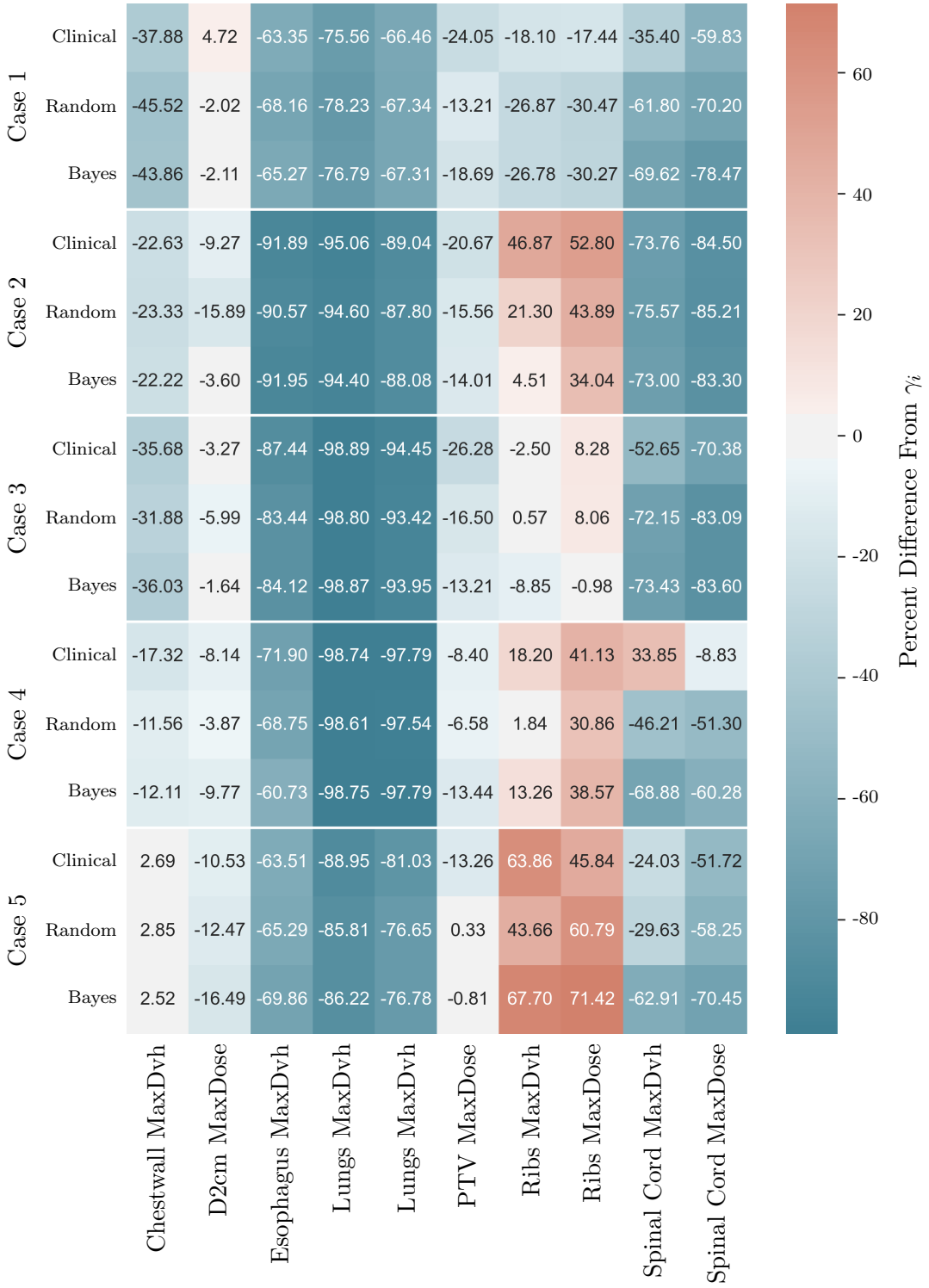

Figure 4.9: Doses delivered in the optimal plans expressed as percent difference from γ_i. For all cases, the Bayesian method meets the clinical goals for every term using the linear–quadratic utility.

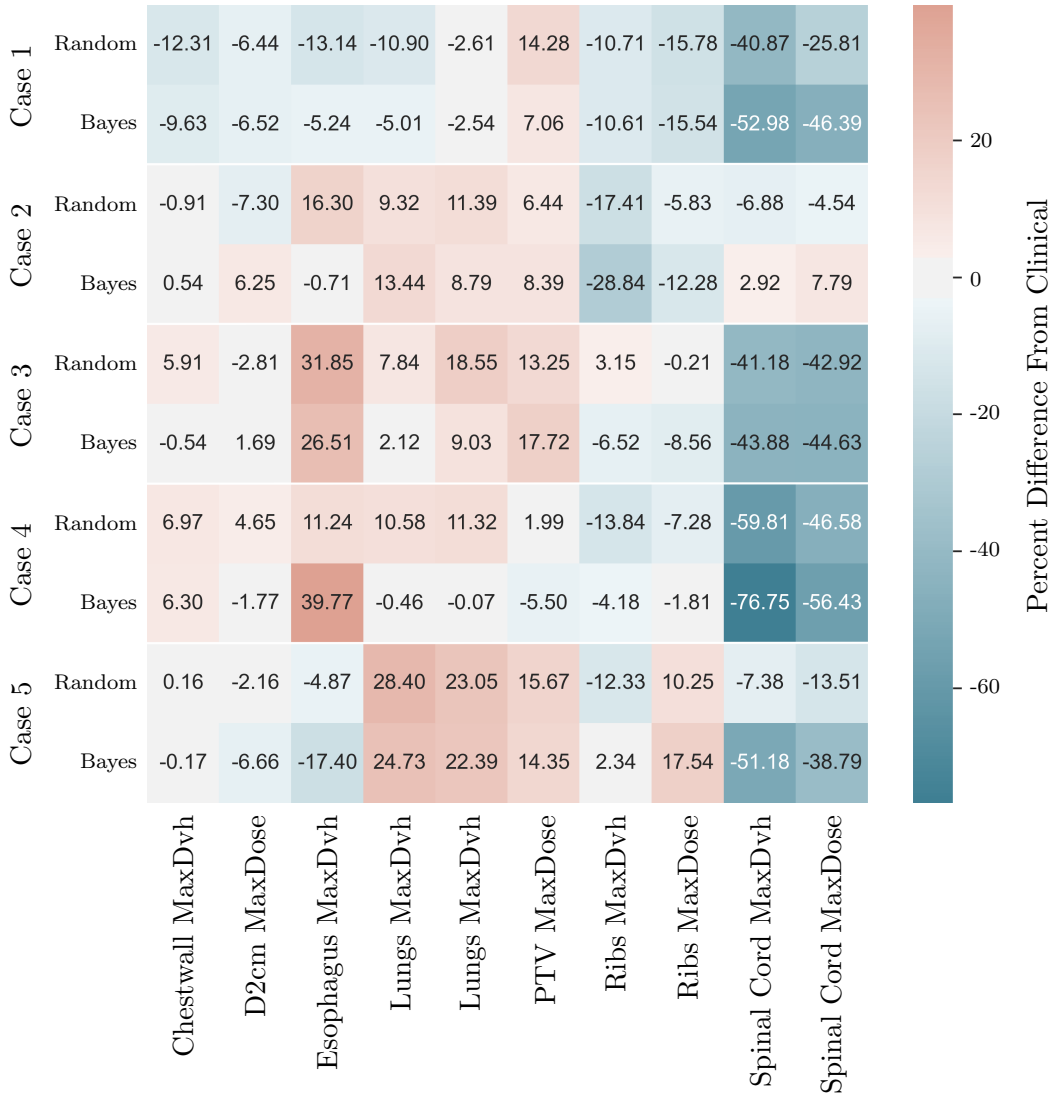

Figure 4.10: Doses delivered in the optimal plans expressed as percent difference from clinical plan doses. For cases 1–4, the Bayesian method delivers up to a 76.75% lower dose than the clinical plans for all ROI goals that are violated in the latter. For case 5, a trade-off is observed between the spinal cord dose (lower in the Bayesian plan) and the rib dose (lower in the clinical plan).

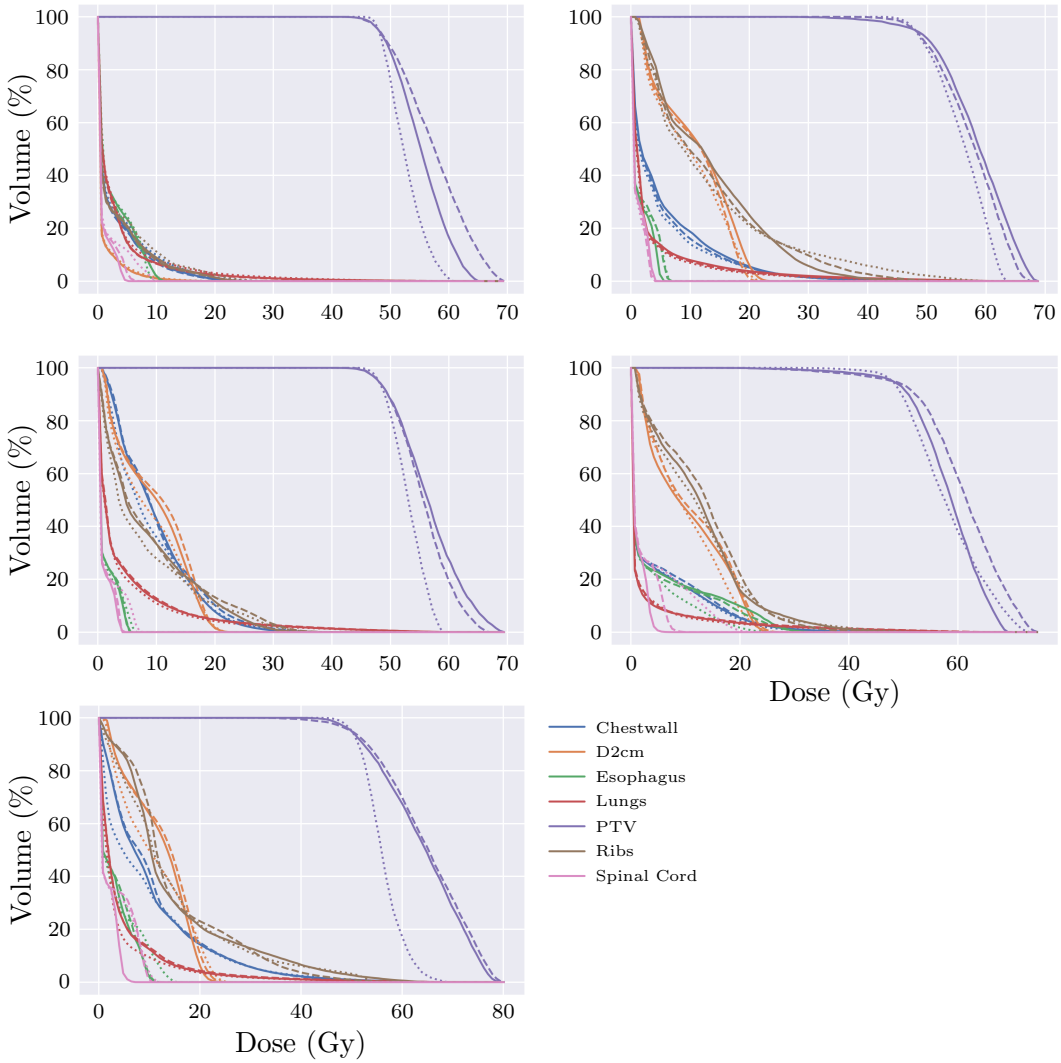

Figure 4.11: Dose–volume histograms for plans computed with random and Bayesian search compared to clinical plans. Dotted, dashed, and solid lines correspond to the clinical, random sampling, and Bayesian optimization plans, respectively. The most consistent trade-off made by the automatically generated plans, as compared to the clinical plans, appears to be an increased dose to the PTV in exchange for a decreased dose to the spinal cord.

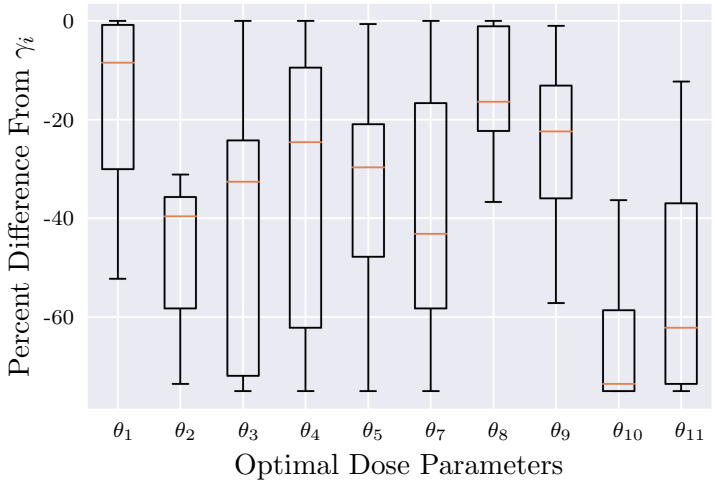

Figure 4.12: Optimal dose parameters θ_i for all automatically generated plans, expressed as percent difference from γ_i.

ranges correspond to clinical goals that appear to be the easiest to meet across all five cases, specifically those for the esophagus and the lungs. This seems to suggest that the utility may be insensitive to changes in these parameters, though other factors, such as the correlation between these parameters, the ROI doses, and the utility values must also be considered. We also note that some optimal parameters never occur in the upper portion of their respective ranges (D2cm and spinal cord), while others never occur in the lower portion (chestwall and ribs). This may in part be due to the choice of linear vs. linear–quadratic utility terms.

4.4.4 Dimensionality considerations

In practice, Bayesian optimization methods have been shown to work efficiently for problems of moderate dimension; however, their success is often limited to no more than 10–20 parameters [83, 106, 172]. Therefore, scaling these approaches to higher dimensions remains a challenge, and efforts to reduce problem dimension can lead to increased efficiency. A useful feature of many problem classes is low effective dimensionality: most dimensions don't

significantly affect the utility function. Low effective dimensionality is often present in the treatment planning problem, e.g. (4.1) and (4.2), and as mentioned in Section 4.1, there have been previous approaches to analytically and empirically identify important parameters for the treatment planning problem. In this section, we motivate future work in automatic parameter selection for our hyperparameter optimization approach by demonstrating the benefit, both in terms of efficiency and treatment plan quality, of identifying a subset of sensitive parameters and corresponding parameter ranges.

We revisit our fifth case, arguably the most challenging, to see if re-optimizing on a subset of the parameters can improve results. Specifically, we fix all parameters at $\theta_i = \gamma_i$, except for those corresponding the D2cm, PTV, and rib maximum doses and the spinal cord maximum dose–volume objective. Furthermore, these parameters are tuned over the modified ranges of $\theta_2 \in [\gamma_2/4, 3\gamma_2/4]$, $\theta_7 \in [\gamma_7/4, \gamma_7]$, $\theta_9 \in [\gamma_9/2, \gamma_9]$, and $\theta_{10} \in [\gamma_{10}/4, 3\gamma_{10}/4]$. These parameters and ranges are chosen empirically, based on the correlations between ROI doses and parameter configurations from the previous example.

In Figure 4.13, we compare the convergence results of using a subset of our parameter space against the results from Section 4.4.3. While our initial example achieved a maximum utility of 241.88 at iteration 99, the re-optimized version finds a treatment plan with a higher utility of 280.7 after only 56 iterations. This improvement in utility corresponds to a decrease in the dose to both the ribs and the spinal cord, as seen in Figures 4.14. This example demonstrates the potential benefit of choosing a good set of parameters to improve the efficiency of the Bayesian optimization routine. While in this case we manually chose our parameter subset, methods to automatically discover sensitive parameters could further improve results.

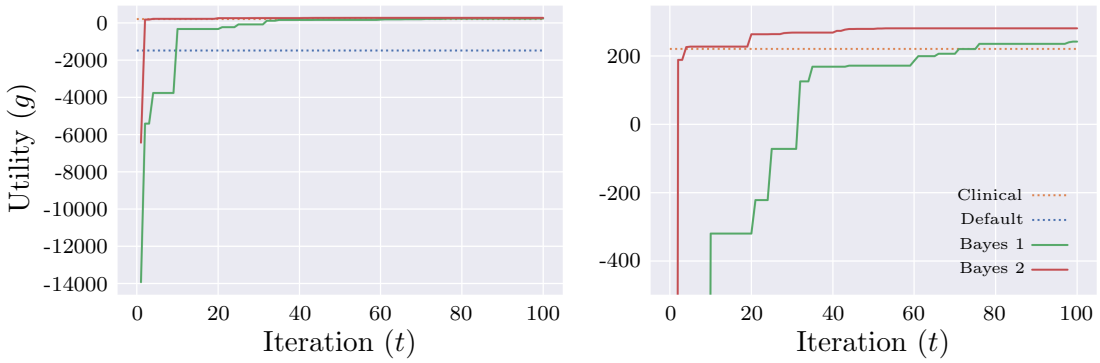

Figure 4.13: Convergence results for Bayesian search methods using different tunable parameters and ranges, along with the clinical (orange) and default (blue) plan utilities. Results using a subset of the tunable parameters (red) are compared to those from Section 4.4.3 (green).

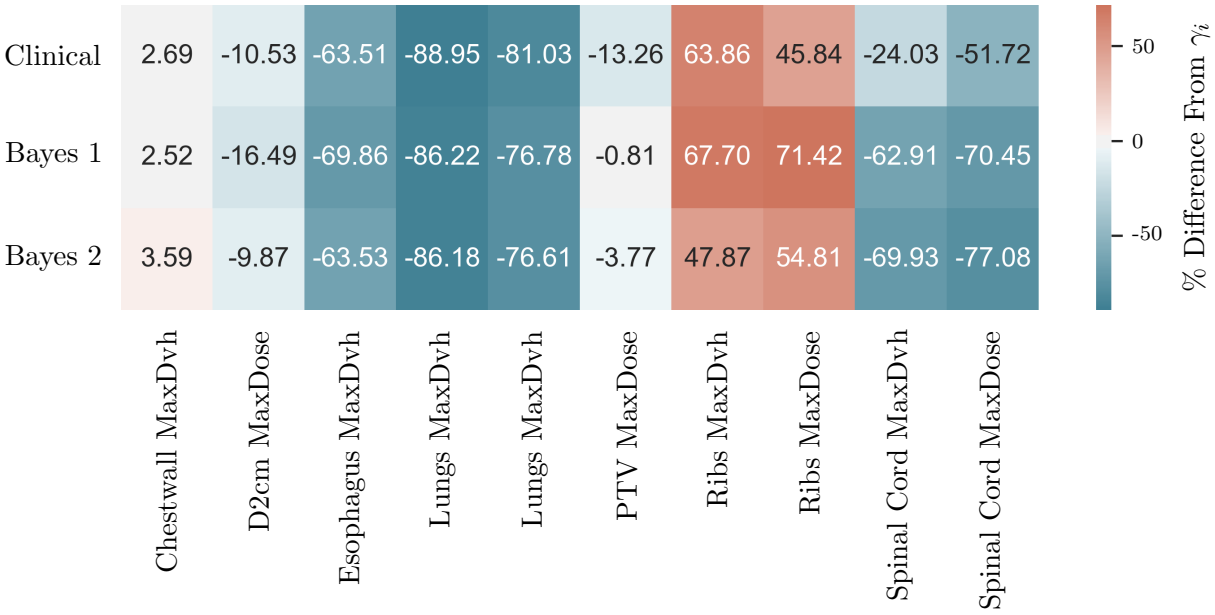

Figure 4.14: Doses delivered in the optimal plans expressed as percent difference from γ_i. The middle row contains doses from the Bayesian optimal plan for case 5 in Section 4.4.3, while the bottom row corresponds to the results computed using a subset of the tunable parameters.

4.5 Discussion

Radiotherapy inverse planning in current practice often requires treatment planners to tune multiple objective function parameters to produce clinically acceptable plans. This manual trial-and-error approach is inefficient, and the plan quality can vary widely depending on the time available and the treatment planner's skills. In this chapter, we proposed a hyperparameter optimization framework to automate the search for optimal objective function parameters, which can help maintain plan quality and reduce treatment planning time. We demonstrated the feasibility and potential benefit of our model through numerical examples for peripherally located lung tumors treated with SBRT using the RayStation TPS. Specifically, we automatically generated treatment plans for five cases using both random sampling and Bayesian optimization with Gaussian processes to optimize linear and linear–quadratic utility functions. The average treatment planning time (no planner intervention required) was 2.2 hours, and plan quality improved for four out of the five cases. Our approach can be easily integrated with any commercial TPS with a scripting interface and does not require training a model on a library of previously optimized plans as in the KBP approach.

There are many potentials directions to expand upon the work in this chapter. For example, developing methods to automatically identify sensitive parameters and effective parameter ranges could guide dimensionality reduction to improve efficiency and utility, as discussed in Section 4.4.4. When a library of previously optimized plans is available, our approach could also be used in conjunction with KBP methods. Other aspects of our implementation, including utility function design and the choice of the surrogate model's kernel and acquisition functions, could also enhance performance. Additionally, further advances in ML hyperparameter optimization can continue to motivate future work, including techniques such as early stopping [95], parallel implementations for Bayesian optimization [154], taking advantage of utility functions with additive structure [83], or collaborative approaches that can tune multiple similar problems simultaneously [17].

BIBLIOGRAPHY

[1] M. J. Abdi. Comparison of several reweighted L1-algorithms for solving cardinality minimization problems. *arXiv preprint arXiv:1304.6655*, 2013.

[2] A. Agnihotri and N. Batra. Exploring Bayesian optimization. *Distill*, 2020. https://distill.pub/2020/bayesian-optimization.

[3] S. U. Ahmad and S. W. A. Bergen. A genetic algorithm approach to the inverse problem of treatment planning for intensity-modulated radiotherapy. *Biomedical Signal Processing and Control*, 5(3):189–195, 2010.

[4] S. Ahmed, O. Gozbasi, M. Savelsbergh, I. Crocker, T. Fox, and E. Schreibmann. An automated intensity-modulated radiation therapy planning system. *INFORMS Journal on Computing*, 22(4):568–583, 2010.

[5] J.-H. Ahn and J. C. Hornberger. Involving patients in the cadaveric kidney transplant allocation process: A decision-theoretic perspective. *Management Science*, 42(5):629–641, 1996.

[6] O. Alagoz, L. M. Maillart, A. J. Schaefer, and M. S. Roberts. The optimal timing of living-donor liver transplantation. *Management Science*, 50(10):1420–1430, 2004.

[7] M. Alber, M. Birkner, and F. Nüsslin. Tools for the analysis of dose optimization: II. Sensitivity analysis. *Physics in Medicine & Biology*, 47(19):N265–N270, 2002.

[8] M. Alber, G. Meedt, F. Nüsslin, and R. Reemtsen. On the degeneracy of the IMRT optimization problem. *Medical Physics*, 29(11):2584–2589, 2002.

[9] M. Alber and F. Nüsslin. An objective function for radiation treatment optimization based on local biological measures. *Physics in Medicine & Biology*, 44(2):479–493, 1999.

[10] M. Alber and F. Nüsslin. Tools for the analysis of dose optimization: I. Effect–volume histogram. *Physics in Medicine & Biology*, 47(14):2451–2458, 2002.

[11] L. M. Appenzoller, J. M. Michalski, W. L. Thorstad, S. Mutic, and K. L. Moore. Predicting dose–volume histograms for organs-at-risk in IMRT planning. *Medical Physics*, 39(12):7446–7461, 2012.

[12] M. Artac, L. Korkmaz, B. El-Rayes, and P. A. Philip. An update on the multimodality of localized rectal cancer. *Critical Reviews in Oncology/Hematology*, 108:23–32, 2016.

[13] A. E. Attema, W. B. F. Brouwer, O. l'Haridon, and J. L. Pinto. An elicitation of utility for quality of life under prospect theory. *Journal of Health Economics*, 48:121–134, 2016.

[14] H. Attouch, J. Bolte, and B. F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: Proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods. *Mathematical Programming*, 137(1-2):91–129, 2013.

[15] G. Banjac and P. J. Goulart. A novel approach for solving convex problems with cardinality constraints. *IFAC-PapersOnLine*, 50(1):13182–13187, 2017.

[16] J. Barbiere, M. F. Chan, J. Mechalakos, D. Cann, K. Schupak, and C. Burman. A parameter optimization algorithm for intensity-modulated radiotherapy prostate treatment planning. *Journal of Applied Clinical Medical Physics*, 3(3):227–234, 2002.

[17] R. Bardenet, M. Brendel, B. Kégl, and M. Sebag. Collaborative hyperparameter tuning. In *Proceedings of the 30th International Conference on Machine Learning*, volume 28, pages 199–207. PMLR, 2013.

[18] A. Beck and Y. C. Eldar. Sparsity constrained nonlinear optimization: Optimality conditions and algorithms. *SIAM Journal on Optimization*, 23(3):1480–1509, 2013.

[19] D. R. Beil and L. M. Wein. Analysis and comparison of multimodal cancer treatments. *Mathematical Medicine & Biology*, 18(4):343–376, 2001.

[20] Y. Bengio. Gradient-based optimization of hyperparameters. *Neural Computation*, 12(8):1889–1900, 2000.

[21] C. C. Bennett and K. Hauser. Artificial intelligence framework for simulating clinical decision-making: A Markov decision process approach. *Artificial Intelligence in Medicine*, 57(1):9–19, 2013.

[22] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl. Algorithms for hyper-parameter optimization. In *Proceedings of the 24th Annual Conference on Neural Information Processing Systems*, volume 24, pages 2546–2554, 2011.

[23] J. Bergstra and Y. Bengio. Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(10):281–305, 2012.

[24] J. Bernier, E. J. Hall, and A. Giaccia. Radiation oncology: A century of achievements. *Nature Reviews Cancer*, 4(9):737–747, 2004.

[25] D. P. Bertsekas. *Nonlinear Programming: 3rd Edition.* Athena Scientific, 2016.

[26] D. Bertsimas, V. Cacchiani, D. Craft, and O. Nohadani. A hybrid approach to beam angle optimization in intensity-modulated radiation therapy. *Computers & Operations Research*, 40(9):2187–2197, 2013.

[27] D. Bienstock. Computational study of a family of mixed-integer quadratic programming problems. *Mathematical Programming*, 74(2):121–140, 1996.

[28] D. Bodensteiner. RayStation: External beam treatment planning system. *Medical Dosimetry*, 43(2):168–176, 2018.

[29] J. Bohsung, S. Gillis, R. Arrans, A. Bakai, C. De Wagter, T. Knöös, B. J. Mijnheer, M. Paiusco, B. A. Perrin, H. Welleweerd, and P. Williams. IMRT treatment planning: A comparative inter-system and inter-centre planning exercise of the ESTRO QUASIMODO group. *Radiotherapy and oncology*, 76(3):354–361, 2005.

[30] T. Bortfeld. Optimized planning using physical objectives and constraints. *Seminars in Radiation Oncology*, 9(1):20–34, 1999.

[31] T. Bortfeld. IMRT: A review and preview. *Physics in Medicine & Biology*, 51:R363–R379, 2006.

[32] T. Bortfeld, J. Bürkelbach, R. Boesecke, and W. Schlegel. Methods of image reconstruction from projections applied to conformation radiotherapy. *Physics in Medicine & Biology*, 35(10):1423–1434, 1990.

[33] J. J. Boutilier, T. Lee, T. Craig, M. B. Sharpe, and T. C. Y. Chan. Models for predicting objective function weights in prostate cancer IMRT. *Medical Physics*, 42(4):1586–1595, 2015.

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

[35] A. Brahme. Optimization of stationary and moving beam radiation therapy techniques. *Radiotherapy and Oncology*, 12(2):129–140, 1988.

[36] A. Brahme, J. E. Roos, and I. Lax. Solution of an integral equation encountered in rotation therapy. *Physics in Medicine & Biology*, 27(10):1221–1229, 1982.

[37] A. Briggs and M. Sculpher. An introduction to Markov modelling for economic evaluation. *PharmacoEconomics*, 13(4):397–409, 1998.

[38] M. Bruglieri, M. Ehrgott, H. W. Hamacher, and F. Maffioli. An annotated bibliography of combinatorial optimization problems with fixed cardinality constraints. *Discrete Applied Mathematics*, 154(9):1344–1357, 2006.

[39] O. P. Burdakov, C. Kanzow, and A. Schwartz. Mathematical programs with cardinality constraints: Reformulation by complementarity-type conditions and a regularization method. *SIAM Journal on Optimization*, 26(1):397–425, 2016.

[40] E. J. Candes, J. K. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, 59(8):1207–1223, 2006.

[41] E. J. Candes, M. B. Wakin, and S. P. Boyd. Enhancing sparsity by reweighted L1 minimization. *Journal of Fourier Analysis and Applications*, 14(5-6):877–905, 2008.

[42] D. Carradice, F. A. K. Mazari, N. Samuel, V. Allgar, J. Hatfield, and I. C. Chetter. Modelling the effect of venous disease on quality of life. *British Journal of Surgery*, 98(8):1089–1098, 2011.

[43] W.-Z. Chen, Y. Xiao, and J. Li. Impact of dose calculation algorithm on radiation therapy. *World Journal of Radiology*, 6(11):874–880, 2014.

[44] P. S. Cho, S. Lee, R. J. Marks, S. Oh, S. G. Sutlief, and M. H. Phillips. Optimization of intensity modulated beams with volume constraints using two methods: Cost function minimization and projections onto convex sets. *Medical Physics*, 25(4):435–443, 1998.

[45] C.-S. Chui and S. V. Spirou. Inverse planning algorithms for external beam radiation therapy. *Medical Dosimetry*, 26(2):189–197, 2001.

[46] H. T. Chung, B. Lee, E. Park, J. J. Lu, and P. Xia. Can all centers plan intensity-modulated radiotherapy (IMRT) effectively? An external audit of dosimetric comparisons between three-dimensional conformal radiotherapy and IMRT for adjuvant chemoradiation for gastric cancer. *International Journal of Radiation Oncology • Biology • Physics*, 71(4):1167–1174, 2008.

[47] A. M. Cormack. A problem in rotation therapy with X rays. *International Journal of Radiation Oncology • Biology • Physics*, 13(4):623–630, 1987.

[48] C. Cotrutz and L. Xing. Using voxel-dependent importance factors for interactive DVH-based dose optimization. *Physics in Medicine & Biology*, 47(10):1659–1669, 2002.

[49] D. Craft, M. Bangert, T. Long, D. Papp, and J. Unkelbach. Shared data for intensity modulated radiation therapy (IMRT) optimization research: The CORT dataset. *GigaScience*, 3(1):1–12, 2014.

[50] D. Craft, M. Bangert, T. Long, D. Papp, and J. Unkelbach. Supporting materials for: "Shared data for intensity modulated radiation therapy (IMRT) optimization research: The CORT dataset". *GigaScience*, 3(1):2047–217X, 2014.

[51] D. Craft, T. Halabi, H. A. Shih, and T. Bortfeld. An approach for practical multi-objective IMRT treatment planning. *International Journal of Radiation Oncology • Biology • Physics*, 69(5):1600–1607, 2007.

[52] S. M. Crooks and L. Xing. Linear algebraic methods applied to intensity modulated radiation therapy. *Physics in Medicine & Biology*, 46(10):2587–2606, 2001.

[53] C. J. Currie, C. Ll. Morgan, C. D. Poole, P. Sharplin, M. Lammert, and P. McEwan. Multivariate models of health-related utility and the fear of hypoglycaemia in people with diabetes. *Current Medical Research and Opinion*, 22(8):1523–1534, 2006.

[54] J. O. Deasy. Multiple local minima in radiotherapy optimization problems with dose–volume constraints. *Medical Physics*, 24(7):1157–1161, 1997.

[55] S. Diamond, R. Takapoui, and S. Boyd. A general system for heuristic solution of convex problems over nonconvex sets. *arXiv preprint arXiv:1601.07277*, 2016.

[56] M. Ehrgott, Ç. Güler, H. W. Hamacher, and L. Shao. Mathematical optimization in intensity modulated radiation therapy. *Annals of Operations Research*, 175(1):309–365, 2010.

[57] P. Franco, A. Fiorentino, F. Dionisi, M. Fiore, S. Chiesa, S. Vagge, F. Cellini, L. Caravatta, M. Tombolini, F. De Rose, et al. Combined modality therapy for thoracic and head and neck cancers: A review of updated literature based on a consensus meeting. *Tumori Journal*, 102(5):459–471, 2016.

[58] A. Fredriksson. Automated improvement of radiation therapy treatment plans by optimization under reference dose constraints. *Physics in Medicine & Biology*, 57(23):7799–7811, 2012.

[59] A. Fu, B. Ungun, L. Xing, and S. Boyd. A convex optimization approach to radiation treatment planning with dose constraints. *Optimization and Engineering*, 20(1):277–300, 2019.

[60] J. Gao and D. Li. A polynomial case of the cardinality-constrained quadratic optimization problem. *Journal of Global Optimization*, 56(4):1441–1455, 2013.

[61] P. E. Gill, W. Murray, and M. A. Saunders. SNOPT: An SQP algorithm for large-scale constrained optimization. *SIAM Review*, 47(1):99–131, 2005.

[62] M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer, 2008.

[63] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`, 2014.

[64] J. Görtler, R. Kehlbeck, and O. Deussen. A visual exploration of Gaussian processes. *Distill*, 2019. https://distill.pub/2019/visual-exploration-gaussian-processes.

[65] T. Halabi, D. Craft, and T. Bortfeld. Dose–volume objectives in multi-criteria optimization. *Physics in Medicine & Biology*, 51(15):3809–3818, 2006.

[66] N. Hansen and A. Ostermeier. Completely derandomized self-adaptation in evolution strategies. *Evolutionary Computation*, 9(2):159–195, 2001.

[67] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, 2020.

[68] L. Hathout, B. Ellingson, and W. Pope. Modeling the efficacy of the extent of surgical resection in the setting of radiation therapy for glioblastoma. *Cancer Science*, 107(8):1110–1116, 2016.

[69] M. Hauskrecht and H. Fraser. Planning treatment of ischemic heart disease with partially observable Markov decision processes. *Artificial Intelligence in Medicine*, 18(3):221–244, 2000.

[70] E. Hazan, A. Klivans, and Y. Yuan. Hyperparameter optimization: A spectral approach. *arXiv preprint arXiv:1706.00764*, 2018.

[71] T. Head, M. Kumar, H. Nahrstaedt, G. Louppe, and I. Shcherbatyi. scikit-optimize/scikit-optimize. `https://doi.org/10.5281/zenodo.4014775`, 2020.

[72] C. Holdsworth, M. Kim, J. Liao, and M. Phillips. The use of a multiobjective evolutionary algorithm to increase flexibility in the search for better IMRT plans. *Medical Physics*, 39(4):2261–2274, 2012.

[73] T. Holmes, T. R. Mackie, D. Simpkin, and P. Reckwerdt. A unified approach to the optimization of brachytherapy and external beam dosimetry. *International Journal of Radiation Oncology • Biology • Physics*, 20(4):859–873, 1991.

[74] Q. Hou, J. Wang, Y. Chen, and J. M. Galvin. An optimization algorithm for intensity modulated radiotherapy — the simulated dynamics with dose–volume constraints. *Medical Physics*, 30(1):61–68, 2003.

[75] Y. Huang, H. Yue, M. Wang, S. Li, J. Zhang, Z. Liu, and Y. Zhang. Fully automated searching for the optimal VMAT jaw settings based on Eclipse Scripting Application Programming Interface (ESAPI) and RapidPlan knowledge-based planning. *Journal of Applied Clinical Medical Physics*, 19(3):177–182, 2018.

[76] M. A. Hunt and C. M. Burman. Treatment planning considerations using IMRT. In *A practical guide to intensity-modulated radiation therapy*, chapter 6, pages 103–121. Memorial Sloan-Kettering Cancer Center, 2003.

[77] M. A. Hunt, C.-Y. Hsiung, S. V. Spirou, C.-S. Chui, H. I. Amols, and C. C. Ling. Evaluation of concave dose distributions created using an inverse planning system. *International Journal of Radiation Oncology • Biology • Physics*, 54(3):953–962, 2002.

[78] F. Hutter, H. H. Hoos, and K. Leyton-Brown. Sequential model-based optimization for general algorithm configuration. In *International Conference on Learning and Intelligent Optimization*, pages 507–523. Springer, 2011.

[79] National Cancer Institute. Cancer statistics. `https://www.cancer.gov/about-cancer/understanding/statistics`, 2018. Accessed: 2020-02-24.

[80] National Cancer Institute. External beam radiation therapy for cancer. `https://www.cancer.gov/about-cancer/treatment/types/radiation-therapy/external-beam`, 2018. Accessed: 2020-02-24.

[81] National Cancer Institute. Radiation therapy to treat cancer. `https://www.cancer.gov/about-cancer/treatment/types/radiation-therapy`, 2019. Accessed: 2020-02-15.

[82] D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13:455–492, 1998.

[83] K. Kandasamy, J. Schneider, and B. Póczos. High dimensional Bayesian optimisation and bandits via additive models. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37, pages 295–304. PMLR, 2015.

[84] J. A. Kemp, M. S. Shim, C. Y. Heo, and Y. J. Kwon. "Combo" nanomedicine: Co-delivery of multi-modal therapeutics for efficient, targeted, and safe cancer therapy. *Advanced Drug Delivery Reviews*, 98:3–18, 2016.

[85] M. Kim, A. Ghate, and M. H. Phillips. A Markov decision process approach to temporal modulation of dose fractions in radiation therapy planning. *Physics in Medicine & Biology*, 54(14):4455–4476, 2009.

[86] K. Kinjo, E. Uchibe, and K. Doya. Evaluation of linearly solvable Markov decision process with dynamic model learning in a mobile robot navigation task. *Frontiers in Neurorobotics*, 7(7):1–13, 2013.

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

[88] M. Krause, A. Scherrer, and C. Thieke. On the role of modeling parameters in IMRT plan optimization. *Physics in Medicine & Biology*, 53(18):4907–4926, 2008.

[89] M. Lahanas, E. Schreibmann, and D. Baltas. Multiobjective inverse planning for intensity modulated radiotherapy with constraint-free gradient-based optimization algorithms. *Physics in Medicine & Biology*, 48(17):2843–2871, 2003.

[90] A. Landers. *Fully Automated Radiation Therapy Treatment Planning Through Knowledge-Based Dose Predictions*. PhD thesis, UCLA, 2018.

[91] M. Langer, R. Brown, M. Urie, J. Leong, M. Stracher, and J. Shapiro. Large scale optimization of beam weights under dose–volume restrictions. *International Journal of Radiation Oncology • Biology • Physics*, 18(4):887–893, 1990.

[92] M. Langer and J. Leong. Optimization of beam weights under dose–volume restrictions. *International Journal of Radiation Oncology • Biology • Physics*, 13(8):1255–1260, 1987.

[93] E. K. Lee, T. Fox, and I. Crocker. Integer programming applied to intensity-modulated radiation therapy treatment planning. *Annals of Operations Research*, 119:165–181, 2003.

[94] A. Li, O. Spyra, S. Perel, V. Dalibard, M. Jaderberg, C. Gu, D. Budden, T. Harley, and P. Gupta. A generalized framework for population based training. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1791–1799, 2019.

[95] L. Li, K. Jamieson, G. DeSalvo, A. Rostamizadeh, and A. Talwalkar. Hyperband: A novel bandit-based approach to hyperparameter optimization. *The Journal of Machine Learning Research*, 18(185):1–52, 2018.

[96] J. Llacer, J. O. Deasy, T. R. Bortfeld, T. D. Solberg, and C. Promberger. Absence of multiple local minima effects in intensity modulated optimization with dose–volume constraints. *Physics in Medicine & Biology*, 48(2):183–210, 2003.

[97] J. Llacer, T. D. Solberg, and C. Promberger. Comparative behaviour of the dynamically penalized likelihood algorithm in inverse radiation therapy planning. *Physics in Medicine & Biology*, 46(10):2637–2663, 2001.

[98] M. S. Lobo, M. Fazel, and S. Boyd. Portfolio optimization with linear and fixed transaction costs. *Annals of Operations Research*, 152:341–365, 2007.

[99] R. Lu, R. J. Radke, L. Happersett, J. Yang, C.-S. Chui, E. Yorke, and A. Jackson. Reduced-order parameter optimization for simplifying prostate IMRT planning. *Physics in Medicine & Biology*, 52(3):849–870, 2007.

[100] Z. Lu and Y. Zhang. Penalty decomposition methods for [l]0-norm minimization. *arXiv preprint arXiv:1008.5372v1*, 2010.

[101] K. Maass. Figures and plotting scripts for "A Markov decision process approach to optimizing cancer therapy using multiple modalities". `https://doi.org/10.5281/zenodo.3595648`, 2019.

[102] K. Maass. Figures and plotting scripts for "A nonconvex optimization approach to IMRT planning with dose–volume constraints". `https://doi.org/10.5281/zenodo.4265172`, 2020.

[103] K. Maass. Figures and plotting scripts for "A hyperparameter optimization approach for automated radiotherapy treatment planning". `https://doi.org/10.5281/zenodo.4603472`, 2021.

[104] K. Maass and M. Kim. A Markov decision process approach to optimizing cancer therapy using multiple modalities. *Mathematical Medicine and Biology: A Journal of the IMA*, 37(1):22–39, 2020.

[105] K. Maass, M. Kim, and A. Aravkin. A nonconvex optimization approach to IMRT planning with dose–volume constraints. *arXiv preprint arXiv:1907.10712*, 2020.

[106] D. Maclaurin, D. Duvenaud, and R. Adams. Gradient-based hyperparameter optimization through reversible learning. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37, pages 2113–2122. PMLR, 2015.

[107] G. S. Mageras and R. Mohan. Application of fast simulated annealing to optimization of conformal radiation treatments. *Medical Physics*, 20(3):639–647, 1993.

[108] P. Magni, S. Quaglini, M. Marchetti, and G. Barosi. Deciding when to intervene: A Markov decision process approach. *International Journal of Medical Informatics*, 60(3):237–253, 2000.

[109] M. Matzenauer, D. Vrana, and B. Melichar. Treatment of brain metastases. *Biomedical Papers*, 160(4):484–490, 2016.

[110] A. Mehrez and A. Gafni. Quality-adjusted life years, utility theory, and healthy-years equivalents. *Medical Decision Making*, 9(2):142–149, 1989.

[111] D. Michalski, Y. Xiao, Y. Censor, and J. M. Galvin. The dose–volume constraint satisfaction problem for inverse treatment planning with field segments. *Physics in Medicine & Biology*, 49(4):601–616, 2004.

[112] M. Miravitlles, A. Huerta, M. Valle, P. García-Sidro, C. Forné, C. Crespo, and J. L. López-Campos. Clinical variables impacting on the estimation of utilities in chronic obstructive pulmonary disease. *International Journal of Chronic Obstructive Pulmonary Disease*, 10:367–377, 2015.

[113] J. Mockus. *Bayesian Approach to Global Optimization: Theory and Applications*, volume 37. Springer, 1989.

[114] S. M. Morrill, R. G. Lane, J. A. Wong, and I. I. Rosen. Dose–volume considerations with linear programming optimization. *Medical Physics*, 18(6):1201–1210, 1991.

[115] S. M. Morrill, I. I. Rosen, R. G. Lane, and J. A. Belli. The influence of dose constraint point placement on optimized radiation therapy treatment planning. *International Journal of Radiation Oncology • Biology • Physics*, 19(1):129–141, 1990.

[116] S. Mukherjee, L. Hong, J. O. Deasy, and M. Zarepisheh. Integrating soft and hard dose–volume constraints into hierarchical constrained IMRT optimization. *Medical Physics*, 47(2):414–421, 2019.

[117] C. J. Murray. Quantifying the burden of disease: The technical basis for disability-adjusted life years. *Bulletin of the World health Organization*, 72(3):429–445, 1994.

[118] K. G. Murthy and S. N. Kabadi. Some NP-complete problems in quadratic and linear programming. *Mathematical Programming*, 39(2):117–129, 1987.

[119] L. B. Nabors, J. Portnow, M. Ammirati, J. Baehring, H. Brem, P. Brown, N. Butowski, M. C. Chamberlain, R. A. Fenstermaker, A. Friedman, et al. Central nervous system cancers, Version 1.2015. *Journal of the National Comprehensive Cancer Network*, 13(10):1191–1202, 2015.

[120] Y. Nesterov. *Introductory lectures on convex optimization: A basic course*, volume 87. Springer, 2013.

[121] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, 2006.

[122] A. Ólafsson, R. Jeraj, and S. J. Wright. Optimization of intensity-modulated radiation therapy with biological objectives. *Physics in Medicine & Biology*, 50(22):5357–5379, 2005.

[123] M. A. Osborne, R. Garnett, and S. J. Roberts. Gaussian processes for global optimization. In *3rd International Conference on Learning and Intelligent Optimization (LION3)*, pages 1–15, 2009.

[124] M. J. O'Connor. Targeting the DNA damage response in cancer. *Molecular Cell*, 60(4):547–560, 2015.

[125] N. Parikh and S. Boyd. Proximal algorithms. *Foundations and Trends in Optimization*, 1(3):127–239, 2014.

[126] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[127] S. Penfold, R. Zalas, M. Casiraghi, M. Brooke, Y. Censor, and R. Schulte. Sparsity constrained split feasibility for dose–volume constraints in inverse planning of intensity-modulated photon or proton therapy. *Physics in Medicine & Biology*, 62(9):3599–3618, 2017.

[128] D. Pflugfelder, J. J. Wilkens, S. Nill, and U. Oelfke. A comparison of three optimization algorithms for intensity modulated radiation therapy. *Zeitschrift für Medizinische Physik*, 18(2):111–119, 2008.

[129] W. B. Powell. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Wiley Series in Probability and Statistics. John Wiley & Sons, 2nd edition, 2011.

[130] M. L. Puterman. *Markov decision processes: Discrete stochastic dynamic programming*. Wiley Series in Probability and Statistics. John Wiley & Sons, 2014.

[131] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2019.

[132] R. T. Rockafellar. *Convex Analysis*, volume 36. Princeton University Press, 1970.

[133] R. T. Rockafellar. Lagrange multipliers and optimality. *SIAM Review*, 35(2):183–238, 1993.

[134] R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2(3):21–41, 2000.

[135] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*, volume 317. Springer, 2009.

[136] 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.

[137] H. E. Romeijn, R. K. Ahuja, J. F. Dempsey, A. Kumar, and J. G. Li. A novel linear programming approach to fluence map optimization for intensity modulated radiation therapy treatment planning. *Physics in Medicine & Biology*, 48(21):3521–3542, 2003.

[138] C. G. Rowbottom and S. Webb. Configuration space analysis of common cost functions in radiotherapy beam-weight optimization algorithms. *Physics in Medicine & Biology*, 47(1):65–77, 2002.

[139] N. Roy and S. Thrun. Coastal navigation with mobile robots. *Advances in Neural Information Processing Systems*, 12:1043–1049, 1999.

[140] J. Rust. Numerical dynamic programming in economics. *Handbook of Computational Economics*, 1:619–729, 1996.

[141] F. Saberian, A. Ghate, and M. Kim. Optimal fractionation in radiotherapy with multiple normal tissues. *Mathematical Medicine and Biology: A Journal of the IMA*, 33(2):211–252, 2016.

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

[143] S. Saokaew, A. Rayanakorn, D. B.-C. Wu, and N. Chaiyakunapruk. Cost effectiveness of pneumococcal vaccination in children in low- and middle-income countries: A systematic review. *PharmacoEconomics*, 34(12):1211–1225, 2016.

[144] V. Sathyanarayanan and S. S. Neelapu. Cancer immunotherapy: Strategies for personalization and combinatorial approaches. *Molecular Oncology*, 9(10):2043–2053, 2015.

[145] A. J. Schaefer, M. D. Bailey, S. M. Shechter, and M. S. Roberts. Modeling medical treatment using Markov decision processes. In *Operations Research and Health Care*, pages 593–612. Springer, 2005.

[146] M. Schmidt. minConf: Projection methods for optimization with simple constraints in Matlab. `http://www.cs.ubc.ca/schmidtm/Software/minConf.html`, 2008.

[147] M. Schmidt, E. Berg, M. Friedlander, and K. Murphy. Optimizing costly functions with simple constraints: A limited-memory projected quasi-Newton algorithm. In *Artificial Intelligence and Statistics*, pages 456–463, 2009.

[148] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175, 2016.

[149] J. Shapiro, J. J. B. van Lanschot, M. C. C. M. Hulshof, P. van Hagen, M. I. van Berge Henegouwen, B. P. L. Wijnhoven, H. W. M. van Laarhoven, G. A. P. Nieuwenhuijzen, G. A. P. Hospers, J. J. Bonenkamp, et al. Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): Long-term results of a randomised controlled trial. *The Lancet Oncology*, 16(9):1090–1098, 2015.

[150] C. Shen, Y. Gonzalez, P. Klages, N. Qin, H. Jung, L. Chen, D. Nguyen, S. B. Jiang, and X. Jia. Intelligent inverse treatment planning via deep reinforcement learning, a proof-of-principle study in high dose-rate brachytherapy for cervical cancer. *Physics in Medicine & Biology*, 64(11):115013, 2019.

[151] 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.

[152] D. M. Shepard, G. H. Olivera, P. J. Reckwerdt, and T. R. Mackie. Iterative approaches to dose optimization in tomotherapy. *Physics in Medicine & Biology*, 45(1):69–90, 2000.

[153] R. L. Siegel, K. D. Miller, and A. Jemal. Cancer statistics, 2020. *CA: A Cancer Journal for Clinicians*, 70(1):7–30, 2020.

[154] J. Snoek, H. Larochelle, and R. P. Adams. Practical Bayesian optimization of machine learning algorithms. *arXiv preprint arXiv:1206.2944*, 2012.

[155] B. Sobotta, M. Söhn, M. Pütz, and M. Alber. Tools for the analysis of dose optimization: III. Pointwise sensitivity and perturbation analysis. *Physics in Medicine & Biology*, 53(22):6337–6343, 2008.

[156] American Cancer Society. Cancer facts & figures 2016, 2016.

[157] American Cancer Society. Global cancer facts & figures, 4th edition, 2018.

[158] American Cancer Society. Cancer facts & figures 2020, 2020.

[159] S. V. Spirou and C.-S. Chui. A gradient inverse planning algorithm with dose–volume constraints. *Medical Physics*, 25(3):321–333, 1998.

[160] M. D. Stubblefield. *Cancer Rehabilitation: Principles and Practice*. Demos Medical, 2 edition, 2018.

[161] X. Sun, X. Zheng, and D. Li. Recent advances in mathematical programming with semi-continuous variables and cardinality constraint. *Journal of the Operations Research Society of China*, 1(1):55–77, 2013.

[162] N. K. Taunk, F. Y. Moraes, F. E. Escorcia, L. C. Mendez, K. Beal, and G. N. Marta. External beam re-irradiation, combination chemoradiotherapy, and particle therapy for the treatment of recurrent glioblastoma. *Expert Review of Anticancer Therapy*, 16(3):347–358, 2016.

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

[164] C. Thornton, F. Hutter, H. H. Hoos, and K. Leyton-Brown. Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 847–855, 2013.

[165] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

[166] J. A. Tropp and S. J. Wright. Computational methods for sparse solution of linear inverse problems. *Proceedings of the IEEE*, 98(6):948–958, 2010.

[167] A. T. Tuncel, F. Preciado, R. L. Rardin, M. Langer, and J.-P. P. Richard. Strong valid inequalities for fluence map optimization problem under dose–volume restrictions. *Annals of Operations Research*, 196:819–840, 2012.

[168] B. Ungun, L. Xing, and S. Boyd. Real-time radiation treatment planning with optimality guarantees via cluster and bound methods. *INFORMS Journal on Computing*, 31(3):544–558, 2019.

[169] G. M. M. Videtic, C. Hu, A. K. Singh, J. Y. Chang, W. Parker, K. R. Olivier, S. E. Schild, R. Komaki, J. J. Urbanic, and H. Choy. NRG Oncology RTOG 0915 (NCCTG N0927): A randomized phase II study comparing 2 stereotactic body radiation therapy (SBRT) schedules for medically inoperable patients with stage I peripheral non-small cell lung cancer. *International Journal of Radiation Oncology • Biology • Physics*, 93(4):757—764, 2015.

[170] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.

[171] H. Wang, P. Dong, H. Liu, and L. Xing. Development of an autonomous treatment planning strategy for radiation therapy with effective use of population-based prior data. *Medical physics*, 44(2):389–396, 2017.

[172] Z. Wang, M. Zoghi, F. Hutter, D. Matheson, and N. de Freitas. Bayesian optimization in high dimensions via random embeddings. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence*, pages 1778–1784. IJCAI, 2013.

[173] S. Webb. Optimisation of conformal radiotherapy dose distribution by simulated annealing. *Physics in Medicine & Biology*, 34(10):1349–1370, 1989.

[174] S. Webb. Optimization by simulated annealing of three-dimensional, conformal treatment planning for radiation fields defined by a multileaf collimator: II. Inclusion of two-dimensional modulation of the X-ray intensity. *Physics in Medicine & Biology*, 37(8):1689–1704, 1992.

[175] S. Webb. The physical basis of IMRT and inverse planning. *The British Journal of Radiology*, 76(910):678–689, 2003.

[176] M. C. Weinstein, G. Torrance, and A. McGuire. QALYs: The basics. *Value in Health*, 12:S5–S9, 2009.

[177] W. J. Welch. Algorithmic complexity: Three NP-hard problems in computational statistics. *Journal of Statistical Computation and Simulation*, 15(1):17–25, 1982.

[178] B. Wu, F. Ricchetti, G. Sanguineti, M. Kazhdan, P. Simari, R. Jacques, R. Taylor, and T. McNutt. Data-driven approach to generating achievable dose–volume histogram objectives in intensity-modulated radiotherapy planning. *International Journal of Radiation Oncology • Biology • Physics*, 79(4):1241–1247, 2011.

[179] C. Wu, R. Jeraj, and T. R. Mackie. The method of intercepts in parameter space for the analysis of local minima caused by dose–volume constraints. *Physics in Medicine & Biology*, 48(11):N149–N157, 2003.

[180] C. Wu, G. H. Olivera, R. Jeraj, H. Keller, and T. R. Mackie. Treatment plan modification using voxel-based weighting factors/dose prescription. *Physics in Medicine & Biology*, 48(15):2479–2491, 2003.

[181] Q. Wu and R. Mohan. Algorithms and functionality of an intensity modulated radiotherapy optimization system. *Medical Physics*, 27(4):701–711, 2000.

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

[183] X. Wu, Y. Zhu, J. Dai, and Z. Wang. Selection and determination of beam weights based on genetic algorithms for conformal radiotherapy treatment planning. *Physics in Medicine & Biology*, 45(9):2547–2558, 2000.

[184] L. Xing and G. T. Y. Chen. Iterative methods for inverse treatment planning. *Physics in Medicine & Biology*, 41(10):2107–2123, 1996.

[185] L. Xing, Russell J. Hamilton, D. Spelbring, C. A. Pelizzari, G. T. Y. Chen, and A. L. Boyer. Fast iterative algorithms for three-dimensional inverse treatment planning. *Medical Physics*, 25(10):1845–1849, 1998.

[186] M. Zaghian, G. Lim, W. Liu, and R. Mohan. An automatic approach for satisfying dose–volume constraints in linear fluence map optimization for IMPT. *Journal of Cancer Therapy*, 5(2):198–207, 2014.

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

[188] H. H. Zhang, R. R. Meyer, J. Wu, S. A. Naqvi, L. Shi, and W. D. D'Souza. A two-stage sequential linear programming approach to IMRT dose optimization. *Physics in Medicine & Biology*, 55(3):883–902, 2010.

[189] P. Zhang, N. Fan, J. Shan, S. E. Schild, M. Bues, and W. Liu. Mixed integer programming with dose–volume constraints in intensity-modulated proton therapy. *Journal of Applied Clinical Medical Physics*, 18(5):29–35, 2017.

[190] Y. Zhang and M. Merritt. Fluence map optimization in IMRT cancer treatment planning and a geometric approach. In W. W. Hager, S. J. Huang, P. M. Pardalos, and O. A. Prokopyev, editors, *Multiscale Optimization Methods and Applications. Nonconvex Optimization and Its Applications*, volume 82, pages 205–227. Springer, 2006.

[191] Y. Zhang and M. Merritt. Dose–volume-based IMRT fluence optimization: A fast least-squares approach with differentiability. *Linear Algebra and its Applications*, 428(5-6):1365–1387, 2008.

[192] P. Zheng and A. Aravkin. Relax-and-split method for nonconvex inverse problems. *Inverse Problems*, 36(9):095013, 2020.

Appendix A

AVAILABILITY OF DATA AND MATERIALS

The figures in this dissertation, as well as the MATLAB and Python scripts necessary to reproduce them, are available openly under the CC-BY license [101, 102, 103]. All examples for Chapter 3 are performed on the CORT prostate dataset [49, 50]. Solutions to nonnegative least-squares problems are solved with the function `minConf_TMP` from the minConf package [146, 147], and solutions for (3.96) in Section 3.4.4 are solved with CVX, a package for specifying and solving convex programs [62, 63]. All examples for Chapter 4 are computed using the RayStation TPS [28] and the Python package Scikit-Optimize [71].

VITA

Kelsey Lynn Maass earned a B.A. *summa cum laude* in Mathematics from Azusa Pacific University in 2009, after which she received a 2009–2010 Fulbright English Teaching Assistantship to the Principality of Andorra. She attended the University of Washington from 2013–2021, where she obtained an M.S. (2015) and Ph.D. (2021) in Applied Mathematics. Kelsey conducted research at Pacific Northwest National Laboratory from 2019–2020 and is currently a postdoctoral scholar at the Institute for Health Metrics and Evaluation.