ROBUST OPTIMIZATION METHODS FOR
BREAST CANCER RADIATION THERAPY

by

Houra Mahmoudzadeh

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Mechanical and Industrial Engineering
University of Toronto

Copyright © 2015 by Houra Mahmoudzadeh
Abstract

Robust Optimization Methods for Breast Cancer Radiation Therapy

Hora Mahmoudzadeh
Doctor of Philosophy
Graduate Department of Mechanical and Industrial Engineering
University of Toronto
2015

The goal of radiation therapy (RT) is to eliminate cancerous cells by directing radiation beams to the hit the cancer target while sparing the surrounding healthy tissue. In breast cancer, the proximity of the heart to the radiation beams often results in the delivery of excessive dose to the heart, leading to a high risk of cardiac complications after the treatment. The heart moves in and out of the radiation beams due to breathing motion which is often irregular and unpredictable. The challenge is to remain within the clinical limits on the volume of the heart receiving a high dose in the presence of breathing motion uncertainty.

This thesis investigates robust optimization (RO) methods to minimize the radiation dose to the heart in breast cancer RT. First, we present a new optimization framework that combines robust optimization, to take into account breathing motion uncertainty, with a conditional value-at-risk (CVaR) representation of clinical dose-volume criteria. This framework is general and applicable to any problem with an underlying loss/return distribution that changes over time based on the state of some system.

We then explore a range of constraint generation solution strategies to solve the resulting large-scale RO problems and compare the computational efficiency of the proposed approach with that of traditional solution methods. Our computational experiments show that the proposed constraint generation strategies can reduce the solution time by
an order of magnitude.

Next, we apply our robust method to real data of several breast cancer patients and dosimetrically compare the results with those of the current clinical treatment planning methods. Our results demonstrate that the robust approach can substantially reduce the radiation to the heart and is less sensitive to breathing motion compared to the conventional clinical method.

Finally, we apply the concept of Pareto robust optimization (PRO) to breast cancer RT. Using clinical patient datasets, we show that RO and PRO solutions are very close for all patients. We find PRO solutions that have a potentially lower heart dose than RO solutions under non-worst-case breathing scenarios while maintaining the worst-case performance of RO solutions.
To my loving husband,

Mehrdad,

who unconditionally supported me all the way.
To the memory of my father, who was my first academic role model,

ë

To my beloved mother, whom I will forever be indebted to.
Acknowledgements

My sincerest thanks to my advisor, mentor, and friend, Timothy C. Y. Chan, for the endless guidance, support, understanding, patience, and encouragement throughout the years. You have set an example of excellence as a researcher, a mentor, and an instructor. Working with you has been an invaluable experience.

A special thanks to Thomas G. Purdie for introducing me to breast cancer radiation therapy, for patiently answering all my questions, and for the numerous meetings and discussions we had regarding the clinical aspects of this work. I also would like to thank my committee members, Dionne M. Aleman and Chi-Guhn Lee, for their valuable comments and guidance along the way. Many thanks to my internal examiner, Michael W. Carter, and to my external examiner, Marina A. Epelman, for the insightful suggestions and comments on my dissertation.

Finally, a warm thank you to all my labmates, the past and present members of AOL, who always supported me throughout my PhD and helped me in many different ways, and to all my friends in UTORG and the MIE department. Thank you all for making my time at UofT so memorable.
## Contents

1 Introduction

1.1 Breast Cancer Radiation Therapy ........................................... 2
1.2 Treatment Planning in IMRT .................................................. 6
1.3 Robust Optimization ......................................................... 7
1.4 Conditional Value-at-Risk .................................................... 9
1.5 Thesis Structure ............................................................. 11
1.6 Contributions ............................................................... 12

2 A Robust-CVaR Optimization Approach .................................. 14

2.1 Introduction ................................................................. 14
2.2 A Robust-CVaR Model ..................................................... 15
  2.2.1 A CVaR model without uncertainty ................................. 16
  2.2.2 A CVaR model with uncertainty .................................... 16
  2.2.3 Duality-based re-formulations ....................................... 21
  2.2.4 Comparison with worst-case and stochastic CVaR models ... 22
2.3 Application to Breast Cancer Radiation Therapy .................. 23
  2.3.1 Formulations ......................................................... 23
  2.3.2 Generalizing current clinical formulations ....................... 26
2.4 Results ................................................................. 27
  2.4.1 Comparing the average, ABC, and robust-CVaR methods ... 28
## List of Tables

2.1 Problem sizes of the robust counterpart ........................................... 22
2.2 Problem sizes and computation times ................................................. 28
2.3 Dose-volume limits and the planned/realized doses (in Gy) for the average and robust-CVaR methods ................................................................. 32
3.1 Strategies for adding constraints ....................................................... 46
3.2 Computation times ........................................................................... 49
3.3 Detailed computation times and the number of iterations for each strategy 49
3.4 The effect of using different values of $\epsilon$ ....................................... 53
4.1 Comparison of robust and clinical methods with different motion magnitudes for each patient ................................................................. 67
4.2 Comparison of important metrics for 200 realizations of breathing patterns 72
5.1 Comparing the average change in the mean heart dose of the RO and PRO plans ................................................................. 92
# List of Figures

1.1 Tangential breast cancer radiation therapy ........................................... 3
1.2 An example of a differential dose-volume histogram ........................... 10
2.1 The uncertainty set $P^{10}$ ................................................................. 27
2.2 Performance of the treatment planning methods ................................. 30
2.3 Comparison of planned and realized dose distributions .................... 34
3.1 The number of constraints added at each iteration .......................... 51
3.2 Comparing the average number of constraints per iteration and total time 52
3.3 The effect of $\epsilon$ on solution quality .............................................. 52
4.1 Magnitude of displacement of the voxels over all breathing phases ....... 59
4.2 The beam’s eye view and the axial view on the average and ABC scans for patient 2 ................................................................. 61
4.3 Comparison of the planned and accumulated dose-volume histograms ... 64
4.4 Comparison of $D_{99\%}$ optBreast and $D_{10cc}$ Heart for the simulated breathing patterns ................................................................. 66
4.5 Comparison of $D_{99\%}$ optBreast and $D_{10cc}$ Heart for the simulated breathing patterns for all patients ............................................. 79
5.1 Comparing results of RO (using $w^*$) and PRO (using $w^* - y^*$) for $\tilde{\pi}$ ... 90
5.2 Histogram of the difference in mean heart dose from RO to PRO ........ 97
5.3 Comparing RO and PRO results for patient 1 .................................... 101
5.4 Comparing RO and PRO results for patient 2  
5.5 Comparing RO and PRO results for patient 3  
5.6 Comparing RO and PRO results for patient 4  
5.7 Comparing RO and PRO results for patient 5  
5.8 Comparing RO and PRO results for patient 6  
5.9 Comparing RO and PRO results for patient 7
Chapter 1

Introduction

Breast cancer is the most frequently diagnosed type of cancer and the leading cause of cancer death in females worldwide (Ma and Jemal, 2013; Jemal et al., 2011). For patients receiving external beam radiation therapy (RT) as part of their treatment, the heart is subject to breathing motion and is at risk of being exposed to excessive radiation. The potential uncertainty in breathing motion motivates the use of optimization techniques to minimize the damage to the heart. In this thesis, we develop robust optimization (RO) models and solution methods for breast cancer RT treatment planning. We incorporate the concept of conditional value-at-risk within an RO framework to generate models that can ensure clinical criteria on the organs are met under breathing motion uncertainty, and we develop specialized solution methods that can efficiently solve large-scale RO problems of this type. The research in this thesis was performed in collaboration with the Princess Margaret Cancer Centre, Toronto, Canada, which is one of the top five cancer research centres in the world.

In this chapter, we first provide background on breast cancer RT and relevant optimization literature. Next, we review robust optimization and solution methods for large-scale RO problems. We then discuss the concept of conditional value-at-risk (CVaR) and its combination with RO. Finally, we present the thesis structure and contributions.
1.1 Breast Cancer Radiation Therapy

Breast cancer patients often undergo either breast-conserving surgery or mastectomy to remove the cancerous tissue. Adjuvant radiation therapy is an additional treatment after the surgery to reduce the risk of recurrence. Adjuvant RT is always performed after breast-conserving surgery and sometimes after a mastectomy (Canadian Cancer Society, 2015). Breast-conserving surgery with adjuvant radiation has been proven to provide equivalent survival and local control compared to mastectomy, and therefore has become an increasingly popular treatment option for women with early stage breast cancer (Giordano et al., 2005; Harris et al., 2006; Hooning et al., 2007; Marks et al., 2005; Lind et al., 2003).

In traditional external beam radiation therapy, high energy photon beams are delivered to the body in order to eliminate the cancerous cells. The cancerous cells are not able repair the damage as fast as the healthy tissue. Therefore, over the course of the treatment, the cancerous cells will have a larger accumulated damage than the healthy tissue. In left-sided breast cancer, parts of the left lung and the heart are usually inside the treatment field and are considered to be organs-at-risk (OAR). Studies have shown that patients receiving radiation therapy treatment for left-sided breast cancer are at a higher risk of long-term cardiac morbidity after the treatment (Giordano et al., 2005; Harris et al., 2006; Hooning et al., 2007), which is correlated to the volume of heart exposed to radiation during treatment (Marks et al., 2005; Lind et al., 2003). This fact motivates the development of specialized radiation therapy treatment methods for left-sided breast cancer to minimize damage to the heart. Furthermore, breathing motion during treatment causes the heart to move in and out of the radiation field (Jagsi et al., 2007), which adds to the challenge of effectively and consistently sparing the heart.

Intensity-modulated radiation therapy (IMRT) is a conventional cancer treatment method that aims to deliver sufficient radiation dose to a target while sparing the surrounding healthy organs. Furthermore, IMRT allows for varying the intensity of radiation
across a beam and therefore enables a more conformal dose to the target. Treatment planning for breast cancer is most commonly implemented using tangential IMRT, in which a simple radiation beam geometry with two opposed tangential beams (Figure 1.1) are used to irradiate the whole breast volume (Kestin et al., 2000; Landau et al., 2001; Vicini et al., 2002; Purdie et al., 2011).

A distinguishing feature of treating breast cancer using radiation therapy is the presence of breathing motion. First, breathing motion makes the breast a moving target. The motion pattern may be difficult to predict and can change without warning. Second, the motion induces anatomical changes that fluctuate back and forth between geometries that are more favorable and less favorable for treatment. Specifically, at full inhale, the expanded lung pushes the heart away from the breast and, therefore, the treatment field. At exhale, the heart is closer to the breast and is generally exposed to more radiation. Thus, the quality of the treatment is influenced by the amount of time the patient spends in different phases of the breathing cycle. The concave shape of the chest wall also results in unavoidable dose to parts of the heart.

To create a treatment plan, the current clinical standard is to acquire a single computed tomography (CT) scan while the patient is breathing freely. Based on this scan,
patients are classified to determine whether a free-breathing treatment or a breath-hold technique is most appropriate for imaging and treatment delivery (Wang et al., 2010; Quirk et al., 2014). Breath-hold approaches are used for patients whose heart is at risk of receiving excessive radiation, since it can effectively force patients to maintain an inhale breath-hold, pushing the heart away from the treatment beam. Breathing-adapted radiotherapy using inhale gating has also been studied (Korreman et al., 2006; Qi et al., 2012), but gating methods are generally more time-consuming (Keall et al., 2006) and less effective than breath-hold techniques for breast cancer RT (Korreman et al., 2005).

A free-breathing treatment is planned on either a helical scan or a time-averaged scan based on four-dimensional CT (4D-CT) images that contain geometrical information about the organs over the phases of the patient’s breathing cycle from inhale to exhale. The helical scan captures images at various breathing positions in a single image dataset and the average scan is an averaged image dataset using all the 4D-CT phase datasets. Neither the helical scan nor the average scan explicitly accounts for breathing motion uncertainty. Therefore, the fraction of time a patient spends at inhale or exhale, for example, will impact the volume of the heart inside the radiation treatment field. Breathing motion results in image artifacts. The artifacts can be image blurring, which makes structures (organs) appear enlarged or image under-sampling, which results in images with truncated structures. The radiation treatment is designed using this “average scan” (based on either helical or 4D-CT average scan) as input, and the treatment is delivered while the patient is breathing freely. We denote this method the “average method.”

Breath-hold treatments rely on deep inspiration breath-hold (DIBH) methods in which the expanded lung pushes the heart away from the radiation beam (Remouchamps et al., 2003; Sixel et al., 2001). Both the initial imaging and the treatment delivery are performed during inhale breath-hold which may be voluntary or involuntary. The DIBH methods that are based on voluntary breath-hold and rely on surrogates for monitoring
(external markers) may suffer from variability in patient immobilization and these may not be well-assessed during imaging (Wang et al., 2012; McIntosh et al., 2011). Involuntary breath-hold, on the other hand, uses an active breathing control (ABC) device which holds the patient’s lung at a specified and reproducible volume (Remouchamps et al., 2003; Krauss et al., 2005; Jagsi et al., 2007). The drawback of all DIBH methods is that the patients cannot always tolerate breath-hold. At the Princess Margaret Cancer Centre, when there is a large volume of the heart in the radiation field according to the average scan, the patient is re-scanned at a fixed inhale breathing phase while using an ABC device (Wang et al., 2010). The radiation treatment is then designed using the “ABC scan” as input. During each treatment session, the ABC device is again used to ensure the patient reproduces the same lung volume. We denote this method the “ABC method.”

The ABC method better spares the heart, while maintaining target coverage, compared to the average method. However, there are downsides. The ABC method is more time-consuming at imaging and treatment delivery, it is expensive because of the device, and not all patients are able to tolerate long breath holds. The average method requires less time for delivery and can be applied to all patients. However, it does not explicitly consider breathing motion, does not explicitly aim to limit heart dose, and the actual radiation dose accumulated in the various parts of the chest during free breathing may be different from what is planned due to the image artifacts described earlier.

In this thesis, we attempt to capture the benefits of both the average and ABC methods: The ability to deliver the treatment while the patient is breathing freely and improved heart sparing even in the face of uncertainty in the patient’s breathing motion. Our approach to breast cancer IMRT follows the clinical protocol at the Princess Margaret Cancer Centre (Purdie et al., 2011, 2014), where tangential IMRT with two opposed beams that are tangent to the body is used to deliver radiation to a target volume in the breast tissue. Our approach leverages 4D-CT imaging data, which provides multiple
snapshots of the anatomy over the entire breathing cycle, as opposed to a single average scan.

1.2 Treatment Planning in IMRT

In IMRT, the goal is to deliver radiation beams from different angles to a cancer patient so that the beams intersect at the cancerous target (i.e., a tumor), while sparing as much of the surrounding healthy tissue as possible. *Inverse planning* uses computerized algorithms to find a treatment plan that satisfies a set of given criteria.

In IMRT, optimization methods have been widely used for inverse planning (see surveys by Bortfeld (2006); Romeijn and Dempsey (2008); Ehrgott et al. (2010)). For instance, optimization models have been used to find optimal beam angles (e.g., Bortfeld and Schlegel, 1993; Pugachev et al., 2000; Aleman et al., 2009; Bertsimas et al., 2013) and beam shapes (e.g., Shepard et al., 2002; Cotrutz and Xing, 2003; Romeijn et al., 2005). One of the most important problems in IMRT optimization is to find the beam intensities. In IMRT, the radiation beams can be modeled as a collection of small *beamlets* whose intensities are optimized. The optimization problem is then to find the intensity of each beamlet such that sufficient dose is delivered to the target while minimizing the dose to the surrounding healthy tissue. This process is called fluence map optimization (FMO).

A wide range of optimization techniques have been used for FMO in IMRT, from linear programming (e.g., Hamacher and Küfer, 2002; Holder, 2003; Romeijn et al., 2006; Lim et al., 2008) to mixed-integer linear programming (e.g., Bednarz et al., 2002; Lee et al., 2003; Ferris et al., 2006; Lee et al., 2006) and penalty-based quadratic and nonlinear programming (e.g., Webb, 1989; Bortfeld, 1999; Wu and Mohan, 2000; Romeijn et al., 2004). The objectives on target coverage and sparing different organs at risk are often conflicting and there has been extensive research on how to prioritize the objectives.
(e.g., Wilkens et al., 2007), to find the relative objective weights (e.g., Cotrutz and Xing, 2003; Holdsworth et al., 2012; Chan et al., 2014), and to navigate the multiobjective Pareto surface (e.g., Craft et al., 2006).

Stochastic models have been used to formulate uncertainties in IMRT treatment planning (e.g., Löf et al., 1995, 1998; Unkelbach and Oelfke, 2004). Robust optimization, which is the focus of this thesis, has also been used to address uncertainties. In the next section, we review the general robust optimization background and its application to IMRT treatment planning in more detail.

### 1.3 Robust Optimization

Robust optimization (Ben-Tal et al., 2009; Bertsimas et al., 2011) is a technique that can be used to manage uncertainties in the data of an optimization problem. The uncertainty is often expressed using an uncertainty set that contains all possible scenarios of the uncertain data. In traditional robust optimization, an uncertain hard constraint is replaced by its robust counterpart, to ensure that the solution remains feasible under any realization of the uncertain data within the specified uncertainty set (Ben-Tal and Nemirovski, 1999, 2000; Bertsimas and Sim, 2004).

Robust optimization has been applied to a wide range of applications (see Gabrel et al. (2014) and Bertsimas et al. (2011) for comprehensive reviews). One of the areas in which robust optimization has been applied is RT. Robust optimization has been used to manage uncertainties in RT treatment planning problems including uncertainties in patient geometry and day-to-day positioning errors (Chu et al., 2005), dose calculation errors (Ólafsson and Wright, 2006), breathing motion during the treatment session (Chan et al., 2006; Bortfeld et al., 2008; Chan and Mišić, 2013; Mišić and Chan, 2015; Mar and Chan, 2015), and range and setup errors in proton therapy (Unkelbach et al., 2007; Liu et al., 2012; Cao et al., 2012; Fredriksson and Bokrantz, 2014).
Much effort in RO is placed on deriving tractable robust counterparts, which are finite-sized deterministic equivalents to the original RO problem. However, the resulting robust counterpart can still be quite large and computationally challenging to solve for real-world problem instances. In RT treatment planning for example, the original problem is often of a very large scale and the robust counterpart is even larger. Therefore, there is a need for specialized solution methods to solve these problems. Decomposition methods, which have a long history (Benders, 1962; Dantzig, 1965), represent a wide range of methods that can be used to solve large-scale optimization problems. Constraint generation is one type of decomposition method that has been used extensively to solve large-scale optimization problems in applications such as timetable scheduling (Odijk, 1996), network reliability (Shaio, 2001), network design (Andreas and Smith, 2009), facility location (Siddiq, 2013), and network interdiction (Brown et al., 2006, 2009). Oskoorouchi et al. (2011) developed an interior point constraint generation algorithm for semi-infinite problems that was applied to radiation therapy. In this thesis, we will develop constraint generation solution methods to solve large-scale RO problems in RT.

The concept of Pareto efficiency in robust optimization has recently been introduced by Iancu and Trichakis (2013) and has since been applied to a few areas such as hospital scheduling (Baum et al., 2014) and energy pricing (Wei et al., 2015), but has not yet been explored in radiation therapy optimization. Traditional RO models only consider the worst-case scenario of the uncertain parameters, and do not optimize for cases when non-worst-case scenarios inside the uncertainty set occur. Pareto robust optimization (PRO) finds solutions that have the same worst-case performance as RO, but potentially provide a better performance under some non-worst-case scenarios inside the uncertainty set (Iancu and Trichakis, 2013). We explore the PRO concept in our breast cancer RT application.
1.4 Conditional Value-at-Risk

The value-at-risk (VaR) of a loss distribution at the confidence level $\beta$ is the smallest loss such that the probability of exceeding such a loss is at most $\beta$. The VaR concept is useful to measure quantities associated with the tail of a loss distribution. It has found application in areas such as finance (Duffie and Pan, 1997; Linsmeier and Pearson, 2000; El Ghaoui et al., 2003) and healthcare (Bortfeld, 1999; Wu and Mohan, 2000; Erkut et al., 2008), but solving optimization problems with VaR metrics are challenging since these problems are non-convex (Artzner et al., 1999). Conditional value-at-risk (CVaR) is an alternative metric that measures the average of the tail loss values (Rockafellar and Uryasev, 2000). One of the attractive features of CVaR is that it is convex (Rockafellar and Uryasev, 2002). However, uncertainty in the underlying loss distribution results in uncertainty in the CVaR measurements.

An IMRT treatment generates a dose distribution – a distribution of dose to points inside the body. The clinical acceptability of a radiation treatment is usually assessed using dose-volume criteria (the dose to a certain percentage of a structure), which is an equivalent concept to the VaR of a distribution. For example, at Princess Margaret Cancer Centre in Toronto, Canada, one of the dose-volume criteria for breast cancer treatment is that at most 10 cubic centimeters (cc) of the volume of the heart may receive a dose higher than 50% of the prescribed dose. In the literature of IMRT treatment planning, the CVaR concept has been used to formulate constraints on the tails of a dose distribution (Romeijn et al., 2003, 2006). For example, treatments generally aim to minimize underdose (i.e., lower tail of the dose distribution) to the tumor or overdose to healthy organs. CVaR can be used to constrain the mean dose received by the subset of “voxels” (small volume elements) receiving the highest or lowest dose in a structure. The main advantage of CVaR models over other types of treatment planning models (e.g., using VaR metrics or quadratic penalty functions) is that CVaR provides a linear optimization model that is more tractable (as opposed to mixed integer or quadratic models),
especially when the problem is large-scale. Figure 1.2 shows an example distribution of the dose delivered to the clinical target volume (CTV), the primary target that we aim to treat, and the CVaR metric describes the average dose to the sub-volume of the CTV receiving the most (or least) dose. The horizontal axis shows the dose, where the prescribed dose is 42.4 Gy (Gray (Gy) is the standard unit of measurement for radiation dose).

Many researchers have considered robust optimization in a CVaR framework, particularly in financial engineering. One general approach aims to optimize the worst-case with respect to the choice of the underlying distribution of the stochastic parameter that characterizes the loss distribution (Jabbour et al., 2008; Huang et al., 2008, 2010; Zhu and Fukushima, 2009). Quaranta and Zaffaroni (2008) considered optimizing a CVaR objective, subject to standard linear constraints with uncertain coefficients, whereas Natarajan et al. (2009) considered CVaR with uncertain parameters in the objective function and formulated worst-case CVaR models for different types of uncertainty sets. In this thesis, we formulate robust-CVaR models that optimize the expected tail of a distribution that
depends on the state of the underlying system, under uncertainty in the fraction of the
time that the system spends in each state.

1.5 Thesis Structure

The rest of this thesis is structured as follows. Chapter 2 is an expanded version of the
paper “A robust-CVaR optimization approach with application to breast cancer therapy”,
chapter, a robust-CVaR optimization model that focuses on the tails of a distribution is
presented, and its application to breast cancer RT treatment planning is demonstrated.
We show that the proposed model generalizes CVaR models in the literature and is less
conservative than the worst-case CVaR models. We demonstrate the application of this
model in formulating the clinical dose limits for RT treatment planning under breathing
motion uncertainty.

Chapter 3 is an expanded version of the paper “Constraint generation methods for
2015. In this chapter, solution methods for the large-scale robust optimization models
in Chapter 2 are presented, and their computational efficiency is compared with that of
the conventional solution method for solving robust problems.

Chapter 4 is an expanded version of the paper “Robust optimization methods for
cardiac sparing in tangential breast IMRT”, *Medical Physics*, Volume 42, No. 5, 2212–
2222, 2015. In this chapter, the robust optimization methodology developed in Chapter 2
is applied to several breast cancer patient datasets, and the clinical applicability of the
proposed method is illustrated. The quality of the treatment plan is compared with that
of the current clinical methods for all patients.

In Chapter 5, Pareto robust optimization models for breast cancer radiation therapy
are developed with the aim of improving the solution under non-worst-case scenarios.
We develop models to compare the set of RO and PRO solutions and to directly find a PRO solution. Using the set of patient datasets from Chapter 4, we find a PRO solution for each patient and compare its quality with that of the RO solution.

Finally, in Chapter 6 we conclude the thesis and present directions for future research.

1.6 Contributions

While the methodologies presented in this thesis are all motivated by the breast cancer radiation therapy treatment planning, the models and solution methods are general and applicable to problems in other application areas. Specifically, Chapters 2 and 3 focus on developing new methodologies, while Chapters 4 and 5 focus on the application of previously established methods to breast cancer RT. We view the contributions of this thesis as follows:

- Methodology:

1. We develop a novel robust-CVaR framework to optimize the tails of a loss distribution that changes over time based on the states of some system, under uncertainty in the fraction of time spent in each state. Our framework generalizes the current CVaR methods in the literature and is less conservative than worst-case CVaR methods. We demonstrate how this approach can be used to formulate the breast cancer radiation therapy problem under breathing motion uncertainty.

2. We develop a constraint generation solution method to solve large-scale robust optimization models with a large number of robust constraints. We develop and compare several strategies for adding constraints at each iteration. We compare the computational efficiency of our solution approach with that of conventional solution methods.
• Application:

1. We demonstrate the clinical applicability of a robust optimization approach for cardiac sparing in breast cancer radiation therapy using real data from several patients. We compare the outcome of the robust approach with that of the current clinical treatment planning methods under a set of simulated breathing scenarios. We demonstrate that the robust approach always reduces the heart dose compared to the conventional treatment method and robust treatments are dosimetrically closer to ABC treatments, without requiring breath-hold.

2. We present the first application of Pareto robust optimization in radiation therapy treatment planning. We compare the results of PRO models with that of the RO approach for breast cancer RT and calculate the aggregate reduction in heart dose over a set of simulated breathing scenarios. We demonstrate that Pareto robust models can potentially improve the performance of our robust approach in non-worst-case breathing scenarios while maintaining the worst-case performance.
Chapter 2

A Robust-CVaR Optimization Approach

2.1 Introduction

Consider a system that can be in one of many states at any point in time. The number of states and the loss distribution corresponding to each state are known, but the fraction of time that the system spends in each state is unknown. For example, system states could correspond to states of the economy (e.g., recession, growth) and a different distribution of stock market returns would exist in each state. In this chapter, we develop an optimization approach that models the uncertain fraction of time the system spends in each state using robust optimization and aims to optimize or constrain the accumulated mean tail loss using conditional value-at-risk (CVaR) over some period of time. Our robust-CVaR model generalizes some of the existing CVaR models in the literature and is less conservative than worst-case approaches. Our approach is similar in mathematical structure to the mixture and discrete distribution models of Zhu and Fukushima (2009), who consider a likelihood distribution on the sample points and derive a min-max approach to minimize the worst-case scenario. However, we consider a different interpretation of the
uncertainty and optimize the expected tail loss as opposed to minimizing the worst-case.

The development of our robust-CVaR model is motivated by the application in breast cancer radiation therapy. In breast cancer IMRT, the breathing phases represent the state of the system in our general framework. Uncertainty in a patient’s breathing pattern leads to potential uncertainty in the delivered dose distribution and hence treatment quality. There is no literature that combines robust optimization and CVaR in IMRT. We will apply the robust-CVaR method developed in this chapter to breast cancer IMRT, where CVaR constraints are used to limit tail dose to the organs and robust optimization is used to mitigate the effects of breathing motion uncertainty, which causes the chest to move unpredictably during treatment. The robust-CVaR model aims to ensure that the clinical dose-volume criteria are met in the presence of uncertainty due to breathing.

Our specific contributions in this chapter are as follows:

1. We develop a robust-CVaR framework that models the tail of a distribution, which changes over time depending on the state of some system. Our framework is tractable through a duality-based reformulation, and generalizes the stochastic and worst-case CVaR models in the optimization literature.

2. We develop the first optimization model in IMRT treatment planning that embeds robust optimization within a CVaR framework. This model generalizes the existing clinical treatment planning methods for breast cancer.

2.2 A Robust-CVaR Model

First, we briefly review the standard (stochastic) CVaR model, closely following the conceptual framework of Rockafellar and Uryasev (2000, 2002). Then, we extend this model to formulate CVaR in systems with multiple loss distributions corresponding to different states.
2.2.1 A CVaR model without uncertainty

Let \( f(x; Y) \) be the loss function associated with a decision vector \( x \) and a vector of stochastic parameters \( Y \). Let \( \Psi(x, \alpha) \) be the probability that the amount of loss does not exceed a threshold \( \alpha \):

\[
\Psi(x, \alpha) = P_Y (f(x; Y) \leq \alpha).
\]

The upper value-at-risk at the level \( \beta \) (upper \( \beta \)-VaR) is defined as

\[
\zeta_\beta(x; Y) = \min \{ \alpha \in \mathbb{R} | \Psi(x, \alpha) \geq \beta \}.
\]

The upper conditional value-at-risk at the level \( \beta \) (upper \( \beta \)-CVaR), denoted \( \phi_\beta(x; Y) \), is defined as

\[
\phi_\beta(x; Y) = \min_{\zeta_\beta \in \mathbb{R}} \left\{ \zeta_\beta + \frac{1}{1 - \beta} E_Y [(f(x; Y) - \zeta_\beta)_+] \right\},
\]

where \((\cdot)_+\) is defined as \( \max\{\cdot, 0\} \). We can sample a collection of vectors \( y_1, \ldots, y_n \) from the distribution of \( Y \) and approximate the upper \( \beta \)-CVaR as:

\[
\phi_\beta(x; y^1, \ldots, y^n) = \min_{\zeta_\beta \in \mathbb{R}} \left\{ \zeta_\beta + \frac{1}{1 - \beta} \left( \frac{1}{n} \sum_{k=1}^n (f(x; y^k) - \zeta_\beta)_+ \right) \right\}.
\]  (2.1)

2.2.2 A CVaR model with uncertainty

Consider a system that at any time can be in one of a finite number of states \( i \in I \). For each state, there is a different loss distribution associated with it. In our model, the decision maker is optimizing over a certain time horizon in which the state of the system will fluctuate between states of \( I \) unpredictably and the cumulative loss will be uncertain, depending on the amount of time the system spends in each state. The proportion of time that the system spends in state \( i \), \( \tilde{p}(i) \), is unknown and belongs to an uncertainty
Let $p(i)$ be the nominal probability that the system is in state $i$. We denote by $\mathbf{p}$ the nominal probability mass function (PMF), and we use $\hat{\mathbf{p}}$ to denote an uncertain PMF. Let $f(x; Y_i)$ be the loss in state $i$, and assume that the system can be in state $i \in \mathcal{I}$ at any time with probability $\tilde{p}(i)$. Similar to before, we can calculate the probability $\Psi_i(x, \alpha)$ that the amount of loss does not exceed a threshold $\alpha$ when the system is in state $i$. Denote the upper $\beta$-CVaR for state $i$ as $\bar{\phi}_\beta(x; Y_i)$. This state-specific CVaR can be formulated as

$$
\bar{\phi}_\beta(x; Y_i) = \min_{\zeta_{\beta,i} \in \mathbb{R}} \left\{ \zeta_{\beta,i} + \frac{1}{1 - \beta} E_{Y_i}[(f(x; Y_i) - \zeta_{\beta,i})_+] \right\}.
$$

The index $i$ can be viewed as the realization of a random variable $I$ with (discrete) state space $\mathcal{I}$ and probability mass function $\hat{\mathbf{p}}$. We use the same overall upper $\beta$-VaR variable for all states. Next, we propose two definitions of the overall (accounting for all states) upper $\beta$-CVaR metric.

Our first definition corresponds to computing the state-specific CVaR values, weighted by the time spent in each state. Thus, the overall CVaR, denoted $\bar{\phi}_\beta^w(x; \hat{\mathbf{p}}, \mathbf{Y})$, can be defined as

$$
E_{I} \left[ \bar{\phi}_\beta(x; Y_I) \right] = \sum_{i \in \mathcal{I}} \tilde{p}(i) \bar{\phi}_\beta(x; Y_i)
$$

$$
= \sum_{i \in \mathcal{I}} \tilde{p}(i) \min_{\zeta_{\beta} \in \mathbb{R}} \left\{ \zeta_{\beta} + \frac{1}{1 - \beta} E_{Y_i}[(f(x; Y_i) - \zeta_{\beta})_+] \right\}
$$

$$
= \min_{\zeta_{\beta} \in \mathbb{R}} \left\{ \zeta_{\beta} + \frac{1}{1 - \beta} \sum_{i \in \mathcal{I}} \tilde{p}(i) E_{Y_i}[(f(x; Y_i) - \zeta_{\beta})_+] \right\}.
$$

We interchange the minimization and summation because the same overall $\beta$-VaR has to be used for all states, and the variable $\zeta_{\beta}$ does not depend on $i$. We can sample a
collection of \( n_i \) vectors \( y_i^1, \ldots, y_i^{n_i} \) for each state \( i \), and define:

\[
\phi_w^{\beta}(x; \tilde{p}, y_i^1, \ldots, y_i^{n_i} \forall i \in I) = \min_{\zeta_{\beta} \in \mathbb{R}} \left\{ \zeta_{\beta} + \frac{1}{1 - \beta} \sum_{i \in I} \left( \frac{1}{n_i} \sum_{k=1}^{n_i} (f(x; y_i^k) - \zeta_{\beta})_+ \right) \tilde{p}(i) \right\}. 
\]

\((w\text{-CVaR})\)

We denote this CVaR variant as \((w\text{-CVaR})\), referring to the weighted CVaR of state-specific distributions.

For the second definition of the overall upper \( \beta \)-CVaR, consider calculating the accumulated tail loss over all states for a particular PMF \( \tilde{p} \). Instead of weighting the state-specific CVaR values, we generate a single loss distribution using \( \tilde{p} \) and then calculate the CVaR value corresponding to this distribution. Given a particular PMF \( \tilde{p} \), we define

\[
Y = \sum_{i \in I} \tilde{p}(i) Y_i.
\]

Thus, the overall upper \( \beta \)-CVaR can now be defined as

\[
\overline{\phi}_\beta^a(x; \tilde{p}, Y) = \overline{\phi}_\beta(x; \sum_{i \in I} \tilde{p}(i) Y_i)
\]

\[
= \min_{\zeta_{\beta} \in \mathbb{R}} \left\{ \zeta_{\beta} + \frac{1}{1 - \beta} E_Y \left[ \left( f \left( x; \sum_{i \in I} \tilde{p}(i) Y_i \right) - \zeta_{\beta} \right)_+ \right] \right\}.
\]

This definition corresponds to computing the CVaR value associated with the single, accumulated loss distribution corresponding to \( \tilde{p} \), as opposed to weighting multiple CVaR values associated with multiple distributions (one per state). Again, we can sample
\( y_1^i, \ldots, y_n^i \) for each state \( i \) and write

\[
\bar{\phi}_\beta^a(x; \tilde{p}, y_1^i, \ldots, y_n^i \forall i \in \mathcal{I}) = \min_{\zeta_{\beta} \in \mathbb{R}} \left\{ \bar{\zeta}_{\beta} + \frac{1}{(1 - \beta) n} \sum_{k=1}^{n} \left( f \left( x; \sum_{i \in \mathcal{I}} \tilde{p}(i) y_k^i \right) - \bar{\zeta}_{\beta} \right) \right\}.
\]

(\text{CVaR-a})

We denote this CVaR variant as (\text{CVaR-a}), referring to the CVaR of the accumulated distribution. Note that when sampling is used to approximate the loss distributions, the (\text{CVaR-a}) definition can only be used if the same set of samples is used for all states. In other words, if we let

\[
y^k = \sum_{i \in \mathcal{I}} \tilde{p}(i) y_k^i,
\]

then \( y^k \) can be viewed as a sample from the overall accumulated distribution of \( Y \), and (\text{CVaR-a}) can be summarized as

\[
\bar{\phi}_\beta^a(x; \tilde{p}, y^1, \ldots, y^n) = \min_{\zeta_{\beta} \in \mathbb{R}} \left\{ \bar{\zeta}_{\beta} + \frac{1}{(1 - \beta) n} \sum_{k=1}^{n} \left( f(x; y^k) - \bar{\zeta}_{\beta}(x) \right) \right\}.
\]

The samples per state represent random draws from the state-specific loss distributions, and the time-weighted cumulative loss distribution can be found by accumulating the loss for each sample.

Both the (\text{w-CVaR}) and the (\text{CVaR-a}) formulation can be applied if the same set of random samples is used for all states. However, if a different set of samples is used per state, we can only calculate the state-specific loss distributions and the (\text{CVaR-a}) formulation does not apply. In Section 2.3, we will show that the issue of sampling has an important interpretation from a medical physics perspective in the IMRT application.

These robust-CVaR definitions can be used in the objective or constraints of a mathematical program. In this thesis, we focus on robust constraints that involve bounds
on CVaR. An upper robust-CVaR constraint at the level $\beta$, for example, can be written using some parameter $U_\beta$ as follows:

$$\bar{\phi}_\beta(x; \mathbf{\tilde{p}}, Y) \leq U_\beta, \quad \forall \mathbf{\tilde{p}} \in \mathcal{P}. \quad (2.2)$$

Formulating a robust-CVaR constraint using the (w-CVaR) and (CVaR-a) definitions will result in different optimization problems. Note that if the same set of samples is used in each state (thus, $n_i = n, \forall i \in \mathcal{I}$) and if $f(x, Y)$ is convex in $Y$, the upper $\beta$-CVaR derived from equation (CVaR-a) is always less than or equal to that of equation (w-CVaR):

$$\sum_{k=1}^{n} \left( f\left(x; \sum_{i \in \mathcal{I}} \tilde{p}(i) y_{i}^{k}\right) - \zeta_\beta \right)_+ \leq \sum_{k=1}^{n} \left( \sum_{i \in \mathcal{I}} f(x; y_{i}^{k}) \tilde{p}(i) - \zeta_\beta \right)_+$$

$$\leq \sum_{i \in \mathcal{I}} \sum_{k=1}^{n} \left( f(x; y_{i}^{k}) - \zeta_\beta \right)_+ \tilde{p}(i).$$

Similarly, the lower-CVaR calculated from equation (CVaR-a) is always greater than or equal to that from equation (w-CVaR). Thus, in general, when the same set of samples is used for each state, using the (CVaR-a) definition results in a less conservative constraint (2.2) than using the (w-CVaR) definition.

The choice of CVaR variant also affects the size of the optimization problem. An upper robust-CVaR constraint at the level $\beta$ using the (w-CVaR) definition is

$$\zeta_\beta + \frac{1}{1 - \beta} \sum_{i \in \mathcal{I}} \left( \frac{1}{n_i} \sum_{k=1}^{n_i} \bar{d}_{\beta, i}^{k} \right) \tilde{p}(i) \leq U_\beta, \quad \forall \mathbf{\tilde{p}} \in \mathcal{P}$$

$$\bar{d}_{\beta, i}^{k} \geq f(x; y_{i}^{k}) - \zeta_\beta, \quad \forall k \in \{1, \ldots, n_i\}, \forall i \in \mathcal{I}, \quad (2.3)$$

$$\bar{d}_{\beta, i}^{k} \geq 0, \quad \forall k \in \{1, \ldots, n_i\}, \forall i \in \mathcal{I},$$

where $\bar{d}_{\beta, i}^{k} = (f(x; y_{i}^{k}) - \zeta_\beta)_+$. Similarly, an upper robust-CVaR constraint at the level
Chapter 2. A Robust-CVaR Optimization Approach

\( \beta \) using the (CVaR-a) definition is

\[
\bar{\zeta}_\beta + \frac{1}{(1 - \beta)} n \left( \sum_{k=1}^{n} \bar{d}_\beta^k \right) \leq U_\beta, \\
\bar{d}_\beta^k \geq f \left( x; \sum_{i \in I} y_i^k \bar{p}(i) \right) - \bar{\zeta}_\beta, \quad \forall k \in \{1, \ldots, n\}, \forall \bar{p} \in P \quad (2.4)
\]

\( \bar{d}_\beta^k \geq 0, \quad \forall k \in \{1, \ldots, n\}, \)

where \( \bar{d}_\beta^k = \left( f \left( x; \sum_{i \in I} y_i^k \bar{p}(i) \right) - \bar{\zeta}_\beta \right)_+ \).

Note that when the same set of samples is used for each state, formulation (2.3) has
1 + \( n |I| \) variables (in addition to \( x \)), 2n|I| normal constraints, and one set of robust constraints \( \forall \bar{p} \in P \). Formulation (2.4), on the other hand, has 1 + \( n \) variables, and 1 + \( n \) normal constraints, and \( n \) sets of robust constraints \( \forall \bar{p} \in P \).

In the remainder of this chapter, we will use terms such as “(w-CVaR) model” and “(CVaR-a) model” to refer to optimization models that involve constraints like (2.3) and (2.4), respectively. In Section 2.4, we will also show that a (CVaR-a) model generates less conservative solutions than a (w-CVaR) model.

2.2.3 Duality-based re-formulations

The tractability of the (w-CVaR) and (CVaR-a) formulations is dependent on the structure of \( P \). If \( P \) is ellipsoidal, then the robust-CVaR formulations can be re-written as second-order conic programs, and if \( P \) is polyhedral, they can be re-formulated as linear programs (Ben-Tal et al., 2009). These methods are standard, duality-based reformulations. In the IMRT application, we focus on the polyhedral uncertainty set \( P = \{ p \in P \mid p \leq \bar{p} \leq \bar{p} \} \), where \( P \) denotes the probability simplex and \( \bar{p}, \bar{p} \) are given parameters. Table 2.1 shows the size of the robust counterpart of two variations of the problem with a polyhedral uncertainty set, where \( m \) is the number of CVaR constraints (e.g., for different \( \beta \) values), and \( l \) is the size of the decision vector \( x \). In Section 2.3, we
Table 2.1: Problem sizes of the robust counterpart. Sign constraints are not included.

<table>
<thead>
<tr>
<th>Formulation</th>
<th>Variables</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>(w-CVaR)</td>
<td>$l + m(1 + 2n +</td>
<td>I</td>
</tr>
<tr>
<td>(CVaR-a)</td>
<td>$l + m(2 +</td>
<td>I</td>
</tr>
</tbody>
</table>

will compare the size and solution time of these two formulations for the IMRT application.

### 2.2.4 Comparison with worst-case and stochastic CVaR models

Worst-case CVaR is normally defined as the CVaR value corresponding to a probability distribution that generates the highest upper CVaR (in the case of minimizing the tail loss). Let $I$ index the set of distributions. Then, the worst-case CVaR equals

$$\sup_{i \in I} \{ \phi_{\beta}(x; Y_i) \}.$$  

Suppose that the uncertainty set $P$ is confined to only one element, \{p\}, where p is a unit vector with a 1 for distribution i (assume for convenience that the set of distributions is finite). If distribution i generates the worst-case CVaR value, then the robust-CVaR formulation generates the worst-case solution over the set of distributions. If the uncertainty set is confined to one element $\hat{p}$, which represents the nominal PMF of the system, the robust formulation in this case will represent the traditional stochastic formulation. In other words, our robust-CVaR framework generalizes the worst-case and nominal (stochastic) CVaR models from the literature. Also, it easy to see that our robust-CVaR method is less conservative than the worst-case models because it takes into account the fraction of time that the system spends in the worst-case state along with the other states of the system, rather than considering the worst case only.
2.3 Application to Breast Cancer Radiation Therapy

2.3.1 Formulations

Our approach to modeling the breast cancer IMRT problem is an example of the general framework described earlier in this chapter. The patient’s breathing trajectory is discretized into phases (e.g., full inhale, full exhale, and all phases in between). The states represent phases of the patient’s breathing cycle. The quantity $\tilde{p}(i)$ represents the fraction of the time the patient is in breathing phase $i \in I$. As the patient breathes, a varying amount of heart and lung will enter the treatment field, which affects the dose delivered to those organs. Our robust-CVaR model controls the tails of the dose distributions to the organs of interest, while accounting for breathing motion uncertainty. First, we formulate the basic CVaR model without uncertainty with the IMRT context. Then, we introduce breathing motion uncertainty.

2.3.1.1 A CVaR model without motion

Let $B$ be the set of beamlets and $w_b$ be the decision variable that represents the intensity of beamlet $b \in B$. Let $S$ be the set of all structures. Each structure $s \in S$ is divided into a set of voxels, $\mathcal{V}_s$. The set of all voxels is $\mathcal{V} = \bigcup_{s \in S} \mathcal{V}_s$. The parameter $D_{v,b}$, called the “dose-influence matrix”, quantifies the dose voxel $v$ receives per unit intensity of beamlet $b$. This $D$ matrix depends on the anatomical setup depicted by a particular CT scan. Given a beamlet intensity vector $w$, the total dose delivered to voxel $v$ is $\sum_{b \in B} D_{v,b} w_b$. For structure $s$, let $\mathcal{V}_s^\alpha(w)$ be the set of all voxels in $\mathcal{V}_s$ that receive more than $\alpha$ dose under beamlet intensity vector $w$:

$$\mathcal{V}_s^\alpha(w) = \left\{ v \in \mathcal{V}_s \left| \sum_{b \in B} D_{v,b} w_b \geq \alpha \right. \right\}.$$
Let \( f^\alpha_s(w) = |\mathcal{V}^\alpha_s(w)|/|\mathcal{V}_s| \) be the fraction of \( \mathcal{V}_s \) that receives more than \( \alpha \) dose. For structure \( s \), the upper \( \beta \)-value-at-risk (upper \( \beta \)-VaR) is the minimum dose level \( \alpha \) such that no more than \((1 - \beta)\)% of \( \mathcal{V}_s \) receives more than \( \alpha \) dose:

\[
\zeta_{\beta,s}(w) = \inf \{ \alpha \mid f^\alpha_s(w) \leq 1 - \beta \}.
\]

For structure \( s \), the upper \( \beta \)-conditional value-at-risk (upper \( \beta \)-CVaR), \( \phi_{\beta,s}(w) \), is the average dose in the parts of \( \mathcal{V}_s \) that receive more than the upper \( \beta \)-VaR dose:

\[
\phi_{\beta,s}(w) = \min_{\zeta_{\beta,s} \in \mathbb{R}} \left\{ \zeta_{\beta,s} + \frac{1}{(1 - \beta)|\mathcal{V}_s|} \sum_{v \in \mathcal{V}_s} \left( \sum_{b \in \mathcal{B}} D_{v,b} w_b - \zeta_{\beta,s} \right) \right\}.
\]  

Lower CVaR constraints can be defined analogously to upper CVaR constraints. Normally, upper CVaR constraints are included for both the healthy organs and the target. Lower CVaR constraints, however, are typically required only for the target. The full treatment planning model without motion uncertainty is shown in Section 2.7.1 in the appendices.

### 2.3.1.2 A CVaR model with motion uncertainty

Next, we incorporate breathing motion into the model and develop a robust version to account for breathing motion uncertainty. Each breathing pattern is represented by an uncertain PMF. Background information on how actual breathing pattern traces (e.g., sinusoidal waves) can be converted into PMFs can be found in Lujan et al. (1999) and Chan et al. (2006). Let \( \mathcal{P} \) be an uncertainty set representing the set of possible PMFs that a treatment should be robust against. Since each phase corresponds to a different anatomical geometry, we now require more than a single \( \mathbf{D} \) matrix. Let \( \Delta_{v,i,b} \) be the dose voxel \( v \) receives per unit intensity of beamlet \( b \) when the patient is in phase \( i \).

A 4D-CT scan is used as the input data for this model. It is composed of multiple
Chapter 2. A Robust-CVaR Optimization Approach

images, one per phase. Assuming the voxel size remains constant throughout the images, each structure may be represented by a different number of voxels in each phase. The number of voxels in each phase corresponds to the number of samples to approximate the distribution in formulation (w-CVaR). Let \( V^i_s \) be the set of voxels in organ \( s \) in phase \( i \). We can define the upper \( \beta \)-CVaR as

\[
\phi^w_{\beta,s}(w; \tilde{p}) = \min_{\zeta_{\beta,s} \in \mathbb{R}} \left\{ \bar{\zeta}_{\beta,s} + \frac{1}{1 - \beta} \sum_{i \in I} \left( \frac{\sum_{b \in B} \Delta_{v,i,b} w_b - \bar{\zeta}_{\beta,s}}{|V^i_s|} \right) \tilde{p}(i) \right\}.
\] (2.6)

Equation (2.6) corresponds to computing the phase-specific CVaR values, weighted by the time spent in each phase, and is equivalent to the (w-CVaR) equation.

To be able to use the (CVaR-a) equation, we need to use the same set of voxels in each structure throughout the breathing phases. One way to do that is to create a one-to-one mapping of the voxels across phases. This task can be accomplished using deformable registration (Brock et al., 2003, 2005). Typically, each structure is represented as a mesh, which is deformed from phase to phase so that the vertices on the surface of the mesh are in one-to-one correspondence throughout the phases. With such a mapping between the images, the number of voxels remains constant in each phase, but their sizes are changing.

As a result of the deformation mapping, the \( \Delta_{v,i,b} \) matrix for each phase will correspond to the same set of voxels, and \( V^i_s = V_s \) for all \( i \). In this case, for a particular PMF \( \tilde{p} \), the accumulated dose over all phases to a particular voxel \( v \) can be written as \( \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \tilde{p}(i) w_b \). Then, the upper \( \beta \)-CVaR is

\[
\phi^a_{\beta,s}(w; \tilde{p}) = \min_{\zeta_{\beta,s} \in \mathbb{R}} \left\{ \bar{\zeta}_{\beta,s} + \frac{1}{1 - \beta} \sum_{v \in V^i_s} \left( \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \tilde{p}(i) w_b - \bar{\zeta}_{\beta,s} \right) \right\}.
\] (2.7)

which corresponds to equation (CVaR-a).

If we can track the voxels across phases, both formulations can be used. Other-
wise, only the (w-CVaR) formulation can be used. The full clinical formulation based on (CVaR-a), which includes both lower and upper CVaR constraints, is given in Section 2.7.2 along with its duality-based reformulation. Specific parameter settings and a description of the clinical data are provided in Section 2.7.3. The same parameter settings were used in both formulations.

2.3.2 Generalizing current clinical formulations

Before demonstrating computational results, we explain how our robust-CVaR formulation generalizes the two existing clinical methods previously described: the average and ABC methods. In the average method, the treatment plan is based on a single “averaged” scan of the patient’s anatomy captured during free breathing. Only the information from this scan is used to generate the treatment plan. In our CVaR-based treatment planning paradigm, the average method can be modeled using our initial CVaR model (2.5) with a dose-influence matrix $D_{v,b}^A$ representing the dose voxel $v$ receives per unit intensity of beamlet $b$ in the anatomical setup depicted by the average scan. In the ABC method, the treatment plan is based on an image of the patient anatomy at inhale. Again, the ABC method can be modeled using our initial CVaR model (2.5) and a dose-influence matrix $D_{v,b}^I$ representing the dose voxel $v$ receives per unit intensity of beamlet $b$ in the geometry induced by the patient at full inhale.

Given the above discussion, it follows immediately that the robust-CVaR models are a generalization of the two clinical methods. Take formulation (2.6) as an example and set $\mathcal{P} = \{p^A\}$ (a singleton). If the average scan can be represented as a convex combination of the scans corresponding to the different breathing phases, then setting $p^A$ accordingly and $D_{v,b}^A = \sum_{t \in \mathcal{I}} \Delta_{v,t,b}p^A(t)$ recovers the average method. If the average scan cannot be represented as a weighted average of the 4D-CT images, then simply add the average scan to the set of 4D-CT images (if there are $k$ phases, make the average scan the “$(k+1)$-th” phase), and set $p^A$ to be a unit vector with a 1 in the $(k+1)$-th position. In the case of
Figure 2.1: The uncertainty set $P^{10}$. The dots represent the nominal PMF and the errorbars show the upper and lower bounds for the probability of each phase. The probabilities of the phases should sum up to one.

the ABC method, let $P = \{p^I\}$ (a singleton) where $p^I$ is a unit vector with a 1 at the inhale phase and $D_{v,b}^I = \sum_{i \in I} \Delta_{v,i,b} p^I(i)$. The case of formulation (2.7) is similar.

2.4 Results

In this section, we demonstrate the application of the (w-CVaR) and (CVaR-a) formulations using clinical data. We divided the patient’s breathing cycle into five phases from inhale to exhale. The nominal PMF had 50% of the probability at the exhale phase, and the remaining 50% was evenly distributed across the remaining four phases. This PMF is structurally representative of regular, exhale-weighted breathing (Lujan et al., 1999; Bortfeld et al., 2002; Engelsman et al., 2005). We define the uncertainty set $P^{10}$ to be the set of all PMFs that in each phase have a probability within $\pm 0.10$ of the nominal probability (see Figure 2.1). We define $P^5$ similarly, where all PMFs are within $\pm 0.05$ of the nominal PMF in each phase.

Both optimization models were solved using the IBM ILOG CPLEX 12.3 via AMPL on a computer cluster using a single Linux node with a 3.07 GHz 12-core CPU and 32 GB of RAM. Problem sizes and computation times for the robust counterpart of
Table 2.2: Problem sizes and computation times

<table>
<thead>
<tr>
<th>Formulation</th>
<th>Variables</th>
<th>Constraints</th>
<th>CPU time (min.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(w-CVaR)</td>
<td>96,439</td>
<td>81,890</td>
<td>68.2</td>
</tr>
<tr>
<td>(CVaR-a)</td>
<td>69,155</td>
<td>68,252</td>
<td>164.0</td>
</tr>
</tbody>
</table>

the (w-CVaR) and (CVaR-a) models with $\mathcal{P}^{10}$ as the uncertainty set are given in Table 2.2.

### 2.4.1 Comparing the average, ABC, and robust-CVaR methods

We generated 300 random PMFs (100 inside $\mathcal{P}^5$, 100 inside $\mathcal{P}^{10} \setminus \mathcal{P}^5$, and 100 outside $\mathcal{P}^{10}$) and measured the difference in performance among the average, ABC, and robust-CVaR methods. The random PMFs were drawn uniformly from the probability simplex using a Dirichlet distribution with the parameter $\alpha = [1, 1, 1, 1, 1]$ (Ng et al., 2011). Then, we categorized the resulting PMF into one of the three sets mentioned above. We repeated this process until 100 PMFs were generated in each set.

The results are shown in Figure 2.2. Both axes measure VaR metrics. The x-axis measures the minimum dose to the 25 cubic centimeters (cc) of the heart that receive the most dose (the “hottest” 25cc of the heart). The y-axis measures the dose at the 95th percentile of the dose distribution of the CTV. Each point represents one realized PMF. The filled points correspond to PMFs that are inside the uncertainty set, while the unfilled points correspond to ones outside of the uncertainty set. The dashed lines indicate the clinically acceptable limits for the two metrics. If a dose distribution exceeds the limit, generally the treatment would be considered unacceptable and would need to be re-planned. The planned dose based on the average and robust methods are shown with black stars. The “planned average” point shows the dose that was planned based on the average method (using the average scan), and the “planned robust” shows the dose
that was planned based on the robust method, using a nominal breathing pattern. The ABC dose distribution is also shown. It represents the dosimetric gold standard, because the treatment is delivered in the absence of any motion. To facilitate the comparison, all dose distributions were normalized to a mean CTV dose of 42.4 Gy (the prescribed dose).

Figures 2.2(a) and 2.2(b) show the results for the (w-CVaR) and (CVaR-a) models, respectively, with the uncertainty set $P_{10}$. It can be seen that the (CVaR-a) formulation results in a less conservative dose to the CTV compared to that of the (w-CVaR) formulation; hence, it allows for a lower dose on the heart. In other words, the dose to the CTV in Figure 2.2(b) is closer to the clinical limit while meeting the limits under all realized scenarios inside the uncertainty set.

We repeated the computational experiment using the smaller uncertainty set $P^5$ to represent a decreased level of conservatism. This analysis seeks to shed light on the trade-off between reducing heart dose and maintaining target coverage. The results are shown in Figures 2.2(c) and 2.2(d). In this case, the robust dose distributions generally have lower heart dose, with a more noticeable separation between the robust and average results seen in Figure 2.2(d). This reduction in the heart dose is very small in Figure 2.2(c) since the (w-CVaR) formulation is more conservative on the CTV and does not allow for a reduction of the heart dose at the price of reducing the CTV dose. The trade-off for the improved performance on the heart is that in a few of the simulated scenarios, the robust treatment violates the minimum dose target for the CTV. These results illustrate the expected trade-off between being robust against target coverage and sparing the heart under uncertainty. It should be noted that the primary goal in breast cancer IMRT is to cover the target, which was accounted for in our models as a CVaR constraint. Therefore, being robust in target coverage under a larger uncertainty set is possible only at the price of more dose to the heart. On the other hand, if one can guarantee less variability in the patient’s breathing pattern, then it is possible to use a smaller uncertainty set in the
Chapter 2. A Robust-CVaR Optimization Approach

Figure 2.2: Performance of the treatment planning methods in 300 simulated PMF scenarios. The dashed lines represent the clinical acceptability limits.

(a) (w-CVaR), Uncertainty set $\mathcal{P}^{10}$

(b) (CVaR-a), Uncertainty set $\mathcal{P}^{10}$

(c) (w-CVaR), Uncertainty set $\mathcal{P}^5$

(d) (CVaR-a), Uncertainty set $\mathcal{P}^5$
robust-CVaR model to reduce dose to the heart.

Figure 2.2 shows that even if the planned dose distribution for the average method is entirely acceptable on the average scan, the realized dose distribution during treatment may be worse. This result has important clinical implications, as it appears that using the average scan to accumulate dose may result in a poor representation of the actual dose delivered to the patient. Furthermore, many of the dose distributions arising from the average treatment end up violating the clinical criteria for acceptability, even if the breathing pattern is inside the uncertainty set. On the other hand, the robust treatment satisfies the clinical criteria under all scenarios and generally dominates the average treatment. The performance of the robust solution pushes the efficient frontier towards the ABC solution, which is the dominant method from a dosimetric perspective.

Notice that because our uncertainty sets $\mathcal{P}^5$ and $\mathcal{P}^{10}$ surround an exhale-weighted nominal PMF, they are “conservative” uncertainty sets. In this case, “conservative” refers less to the size of the uncertainty set – traditional robust optimization produces more conservative solutions given larger uncertainty sets – and more to the nature of the solution should the uncertain data be realized outside the uncertainty set, rather than inside. Recall that the heart generally receives the most dose in the exhale phase. Therefore, an exhale-weighted uncertainty set focuses on protecting against uncertainty in the case where heart dose is generally higher. When a PMF is realized outside of the uncertainty set, it has a higher chance of being inhale-weighted, which results in lower heart dose. In other words, PMFs realized inside the uncertainty set will result in the worst heart dose, and other PMFs can only improve on what was planned.

Figure 2.2 examines only two of the relevant metrics for breast cancer IMRT. A more comprehensive list of VaR metrics used at Princess Margaret Cancer Centre is shown in Table 2.3 for the results obtained using the uncertainty set $\mathcal{P}^{10}$. Similar results were observed for $\mathcal{P}^5$. Metrics of the form “$D_{x\%} \leq y$” require that the dose to $x\%$ of the organ be at most $y$ Gy. Some are specified in terms of cubic centimeters instead of percent
Table 2.3: Dose-volume limits and the planned/realized doses (in Gy) for the average and robust-CVaR methods with $\mathcal{P}^{10}$. Ranges for the realized doses represent the minimum and maximum doses over all realized PMFs. Results that meet the criteria are bolded. Shaded cells indicate that the robust-CVaR method performed better than the average method.

<table>
<thead>
<tr>
<th>Organ</th>
<th>Metric</th>
<th>Average</th>
<th>Robust-CVaR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Planned</td>
<td>Realized</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>44.72</td>
<td>[43.95,44.30]</td>
</tr>
<tr>
<td>CTV</td>
<td>D0.5% ≤ 45.79</td>
<td>40.36</td>
<td>[38.87,39.50]</td>
</tr>
<tr>
<td></td>
<td>D95.0% ≥ 39.01</td>
<td>42.40</td>
<td>[41.68,41.82]</td>
</tr>
<tr>
<td></td>
<td>D25cc ≤ 21.20</td>
<td>2.56</td>
<td>[2.41,2.70]</td>
</tr>
</tbody>
</table>

The “planned” column shows the dose that was calculated on the planning image and the “realized” range shows the minimum and maximum realized dose over all simulated PMFs (inside and outside of $\mathcal{P}^{10}$).

First, notice that the planned dose distribution corresponding to the robust solution is always inside the range of the realized distribution for all metrics. On the other hand, the planned dose corresponding to the average solution can be quite different (both higher and lower) than the realized interval. This observation confirms that our robust-CVaR framework provides a better, more stable estimate of the realized dose. Another observation is that the realized robust dose distributions meet the clinical criteria in almost all of the simulated cases, generally outperforming the average solution. Finally, note that the (w-CVaR) formulation dominates the average method in all metrics for all PMFs, but only slightly reduces the heart dose; whereas the (CVaR-a) formulation reduces the heart dose more, by giving up a small (clinically acceptable) amount of CTV dose for some breathing patterns. The reason is that the (CVaR-a) formulation is less

volume.
conservative on the target coverage compared to the \((w-CVaR)\) formulation, and therefore can reduce the dose to the heart while maintaining clinical target coverage for breathing patterns within the uncertainty set.

Lastly, in Figure 2.3, we show dose-volume histograms associated with the average and \((CVaR-a)\) models using the simulated PMFs inside \(P^5\) and \(P^{10}\). The results for the \((w-CVaR)\) formulation are similar. The horizontal axis of a DVH is dose, and the vertical axis is the percentage of the volume of an organ receiving that amount of dose or higher. For the average model (Figures 2.3(a) and 2.3(b)), we see that the realized dose distribution on the breast is worse than what was planned. Similarly, we see a substantial difference between the planned and realized dose-volume curves for the heart. This difference indicates that the planned dose distribution on the average scan is not always a reliable indicator of what the actual delivered dose will be, highlighting a drawback of using the average method to do treatment planning. On the other hand, Figures 2.3(c) and 2.3(d) show that our robust-CVaR model can generate a planned dose distribution that is a better estimate of the actual delivered dose, even in the face of uncertainty.

2.5 Potential Impact of the Robust-CVaR Approach in IMRT

Here, we expand on the clinical implications of the results seen in Section 2.4. Recall that we set out to develop an optimization-based IMRT treatment planning method for left-sided breast cancer that combines the advantages of the average and ABC methods. Our robust-CVaR treatments can be delivered during free breathing for any patient (like the average method) and produces dose distributions that push the frontier forward and towards the ABC dose distribution. While the average method often results in dose distributions that violate the clinical criteria, the robust treatments almost always satisfy the criteria. Furthermore, the robust treatments have the benefit of improved
Chapter 2. A Robust-CVaR Optimization Approach

Figure 2.3: Comparison of planned and realized dose distributions. The dashed lines represent the planned dose, and the solid lines represent the realized dose. The dark (red) lines are for PMFs inside the uncertainty set $\mathcal{P}^5$ or $\mathcal{P}^{10}$, and the light grey lines are those outside of the uncertainty sets.
predictability between the planned and realized dose distributions.

The robust-CVaR method is also implicitly advocating for the use of 4D-CT imaging in breast cancer patients. If patients are given 4D-CT scans, better measurement of the true amount of heart in the radiation field is possible, which may allow treatment planners to refer fewer patients to ABC (as long as dose criteria are met) and possibly reduce the burden on the resources of the health system. Although 4D-CT scans result in a little more dose being delivered to the patient in comparison to a regular CT scan, it is possible that this increase is more than offset by a reduction in dose delivered to critical organs, made possible by the robust-CVaR method. This hypothesis should be tested in a more clinically oriented study on a large cohort of patient data, which would also serve to validate our findings more broadly. This idea is a topic of future study.

2.6 Conclusions

In this chapter, we develop two robust-CVaR definitions for formulating the tail of a loss distribution under uncertainty. We use a novel interpretation of a system with a loss distribution that is state dependent and the time spent in each state is uncertain. The cumulative loss depends on the uncertain fraction of time the system spends in each state. Our robust-CVaR framework generalizes existing models in the literature and is less conservative than worst-case methods.

We demonstrate an application of our framework to the IMRT treatment planning problem of breast cancer where we optimize the tails of the distribution of dose under uncertainty in the patient’s breathing pattern. Our model generalizes two existing clinical methods and has advantages of both – deliverability while the patient is breathing freely and improved cardiac sparing. Our computational results demonstrate improved dosimetric performance over the standard treatment planning method under free breathing.
2.7 Appendix

2.7.1 A CVaR model without motion

Recall from Equation (2.5) that

\[
\bar{\phi}_{\beta,s}(w) = \min_{\zeta_{\beta,s} \in \mathbb{R}} \left\{ \zeta_{\beta,s} + \frac{1}{(1 - \beta)|V_s|} \sum_{v \in V_s} \left( \sum_{b \in B} D_{v,b} w_b - \zeta_{\beta,s} \right) \right\}.
\]

Let \( U_{\beta,s} \) be an upper bound for \( \bar{\phi}_{\beta,s}(w) \). Each structure \( s \) may be prescribed multiple dose-volume criteria, each indexed by a specific \( \beta \). The weights in the objective are controlled via parameters \( c_s \) for each structure \( s \). Let \( A_s \) be the set of all upper \( \beta \) levels for structure \( s \). The initial CVaR model is

\[
\text{minimize} \quad \sum_{s \in S} \frac{c_s}{|V_s|} \sum_{v \in V_s} \sum_{b \in B} D_{v,b} w_b \tag{2.8}
\]

subject to

\[
\zeta_{\beta,s} + \frac{1}{(1 - \beta)|V_s|} \sum_{v \in V_s} \bar{d}_{v,\beta,s} \leq U_{\beta,s}, \quad \forall \beta \in \bar{A}_s, s \in S,
\]

\[
\bar{d}_{v,\beta,s} \geq \sum_{b \in B} D_{v,b} w_b - \zeta_{\beta,s}, \quad \forall v \in V_s, \beta \in \bar{A}_s, s \in S,
\]

\[
\zeta_{\beta,s} \geq 0, \quad \forall \beta \in \bar{A}_s, s \in S,
\]

\[
\bar{d}_{v,\beta,s} \geq 0, \quad \forall v \in V_s, \beta \in \bar{A}_s, s \in S,
\]

\[
w_b \geq 0, \quad \forall b \in B.
\]

2.7.2 The complete robust-CVaR treatment planning model and its robust counterpart

Formulation (2.9) is the (CVaR-a) model that is used in Section 2.4 and includes lower CVaR constraints. Sets \( \bar{A}_s \) and \( A_s \) index the upper and lower \( \beta \) levels for structure \( s \), respectively. Parameters \( U_{\beta,s} \) and \( L_{\beta,s} \) are the upper and lower CVaR limits for structure \( s \),
respectively.

\[
\begin{align*}
\text{minimize} & \quad \sum_{s \in S} \frac{c_s}{|V_s|} \sum_{v \in V_s} \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} p(i) w_b \\
\text{subject to} & \quad \bar{\zeta}_{\beta,s} + \frac{1}{(1 - \beta)|V_s|} \sum_{v \in V_s} d_{v,\beta,s} \leq U_{\beta,s}, \quad \forall \beta \in \overline{A}_s, s \in S, \\
& \quad \bar{d}_{v,\beta,s} \geq \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \bar{p}(i) w_b - \tilde{\zeta}_{\beta,s}, \quad \forall v \in V_s, \beta \in \overline{A}_s, s \in S, \bar{p} \in \mathcal{P}, \\
& \quad \zeta_{\beta,s} - \frac{1}{(1 - \beta)|V_s|} \sum_{v \in V_s} d_{v,\beta,s} \geq L_{\beta,s}, \quad \forall \beta \in \overline{A}_s, s \in S, \\
& \quad \bar{d}_{v,\beta,s} \geq \zeta_{\beta,s} - \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \bar{p}(i) w_b, \quad \forall v \in V_s, \beta \in \overline{A}_s, s \in S, \bar{p} \in \mathcal{P}, \\
& \quad \zeta_{\beta,s} \geq 0, \quad \forall \beta \in \overline{A}_s, s \in S, \\
& \quad \zeta_{\beta,s} \geq 0, \quad \forall \beta \in \overline{A}_s, s \in S, \\
& \quad \bar{d}_{v,\beta,s} \geq 0, \quad \forall v \in V_s, \beta \in \overline{A}_s, s \in S, \\
& \quad d_{v,\beta,s} \geq 0, \quad \forall v \in V_s, \beta \in \overline{A}_s, s \in S, \\
& \quad w_b \geq 0, \quad \forall b \in B.
\end{align*}
\]

Formulation (2.10) is the linear program equivalent of formulation (2.9) and is derived using LP duality.

\[
\begin{align*}
\text{minimize} & \quad \sum_{s \in S} \frac{c_s}{|V_s|} \sum_{v \in V_s} \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} p(i) w_b \\
\text{subject to} & \quad \bar{\zeta}_{\beta,s} + \frac{1}{(1 - \beta)|V_s|} \sum_{v \in V_s} d_{v,\beta,s} \leq U_{\beta,s}, \quad \forall \beta \in \overline{A}_s, s \in S, \\
& \quad \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} p(i) w_b - \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \bar{p}(i) w_b + \sum_{i \in I} p(i) q_{v,\beta,s} \\
& \quad + \sum_{i \in I} r_{v,i,\beta,s} \leq \zeta_{\beta,s} + \bar{d}_{v,\beta,s}, \quad \forall v \in V_s, \beta \in \overline{A}_s, s \in S, \\
& \quad \sum_{b \in B} \Delta_{v,i,b} (\bar{p}(i) + p(i)) w_b - r_{v,i,\beta,s} - (\bar{p}(i) + p(i)) q_{v,\beta,s} \leq 0, \\
& \quad \forall i \in \mathcal{I}, v \in V_s, \beta \in \overline{A}_s, s \in S.
\end{align*}
\]
\[ \zeta_{\beta,s} - \frac{1}{1-\beta} \sum_{v \in V_s} d_{v,\beta,s} \geq L_{\beta,s}, \quad \forall \beta \in A_s, s \in S, \]
\[ \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} p(i) w_b - \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} p(i) w_b + \sum_{i \in I} p(i) q_{v,\beta,s} \]
\[ \quad - \sum_{i \in I} r_{v,i,\beta,s} \geq \zeta_{\beta,s} - d_{v,\beta,s}, \quad \forall v \in V_s, \beta \in A_s, s \in S, \]
\[ \sum_{b \in B} \Delta_{v,i,b} (\bar{p}(i) + p(i)) w_b + r_{v,i,\beta,s} + (\bar{p}(i) + p(i)) q_{v,\beta,s} \geq 0, \]
\[ \forall i \in I, v \in V_s, \beta \in A_s, s \in S, \]
\[ \zeta_{\beta,s} \geq 0, \quad \forall \beta \in A_s, s \in S, \]
\[ \zeta_{\beta,s} \geq 0, \quad \forall \beta \in A_s, s \in S, \]
\[ d_{v,\beta,s} \geq 0, \quad \forall v \in V_s, \beta \in A_s, s \in S, \]
\[ d_{v,\beta,s} \geq 0, \quad \forall v \in V_s, \beta \in A_s, s \in S, \]
\[ q_{v,\beta,s} \text{ free}, \quad \forall \beta \in A_s \cup A_s, s \in S, \]
\[ r_{v,i,\beta,s} \geq 0, \quad \forall i \in I, \forall v \in V_s, \beta \in A_s, s \in S, \]
\[ w_b \geq 0, \quad \forall b \in B. \]

### 2.7.3 Clinical data

This section describes the clinical data and parameter value settings used in the computational experiments. A 4D-CT dataset from a left-sided breast cancer patient was obtained from Princess Margaret Cancer Centre. The heart, lung and CTV were contoured on each phase of the 4D-CT scan. To keep track of the deformation between the voxels of all phases, the contours on the inhale scan of the 4D-CT were deformed over five phases so that the fifth phase replicated the exhale scan of the 4D-CT. The average scan was also available with the actual planned dose for this patient. Since this patient was not classified as an ABC patient, the inhale scan of the 4D-CT was used to generate the ABC results. The contours were imported into MATLAB® using the
Computation Environment for Radiotherapy Research (CERR) software package (Deasy et al., 2003). CERR was also used to calculate the dose-influence matrices. The upper and lower dose-volume criteria for the CTV were used as hard constraints and the only objective was minimizing the dose to the heart. That is, $c_s = 1$ for $s = \{\text{heart}\}$, and 0 otherwise. The values of $\beta$, $U_{\beta,s}$, and $L_{\beta,s}$ used were taken from the dose-volume limits for the CTV shown in Table 2.3. For example, for $\beta = 0.5\%$ the upper limit $U_{\beta,s}$ was 45.79 Gy.

The CT scan resolution was $1 \times 1 \times 2$ mm$^3$ for the voxels. With this resolution, the (CVaR-a) formulation had 2,642,012 constraints and 6,165,593 variables. In clinical treatment planning at Princess Margaret Cancer Centre, voxels of dimension $4 \times 4 \times 4$ mm$^3$ are used. Thus, to reduce the problem size, we randomly sampled one out of every 32 voxels ($= 4 \cdot 4 \cdot 2$) in each structure and solved the optimization problems on this reduced set of voxels. We compared the results of solving an instance of the full-sized problem with those of the sampled problem, and there was negligible difference in the DVHs. Thus we used the sampled voxel set for all the computational experiments.
Chapter 3

A Constraint Generation Solution Method

3.1 Introduction

In this chapter, we develop a family of constraint generation strategies to solve large-scale robust optimization problems in radiation therapy. We focus on problems with multiple sets of robust constraints, which necessitates exploring different strategies for choosing constraints to be added at each iteration. We test several strategies for finding and adding constraints efficiently. We also compare the computational efficiency of the constraint generation methods with that of directly solving the robust counterpart. Our solution approach is motivated by the robust intensity-modulated radiation therapy (IMRT) treatment planning problem for breast cancer developed in Chapter 2, in which there exists a large number of robust constraints.

3.2 Large-scale Robust IMRT Models

In Chapter 2, model (2.9) showed a robust conditional value-at-risk (CVaR) model for breast cancer IMRT with one set of upper and lower $\beta$-CVaR constraints for limiting
overdose and undertose to the organs. In particular, constraints similar to (3.1a) and (3.1b) are used in formulating upper and lower CVaR limits on the target, respectively.

\[ \bar{d}_{v,\beta}^T \geq \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \tilde{p}(i) w_b - \xi^T \beta, \quad \forall v \in V^T, \quad \beta \in A^T, \quad \tilde{p} \in \mathcal{P}, \quad (3.1a) \]

\[ d_{v,\beta}^T \geq \xi^T \beta - \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \tilde{p}(i) w_b, \quad \forall v \in V^T, \quad \beta \in A^T, \quad \tilde{p} \in \mathcal{P}. \quad (3.1b) \]

Recall that \( w_b \) is the intensity of beamlet \( b \in B \), \( B \) is the set of all beamlets, \( V^T \) is the set of all voxels in the clinical target volume (inside the breast), and \( I \) is the set of all breathing phases. \( \Delta_{v,i,b} \) is the amount of dose that voxel \( v \) receives per unit intensity of beamlet \( b \) when the patient is in phase \( i \). The variable \( \xi \) is the value-at-risk (VaR) of the dose distribution, which captures the \( \beta \) percent of an organ that is receiving the highest amount of dose. In a robust IMRT problem, there may exist multiple CVaR constraints on an organ for different values of \( \beta \).

Notice that constraints (3.1a) and (3.1b) must hold for all voxels in the tumor. We define a type of robust constraint as one that must hold for the same set of voxels for the same \( \beta \) value and in the same inequality direction (i.e., upper or lower CVaR constraints) for all \( \tilde{p} \in \mathcal{P} \). For example, for a given \( \beta \), all the upper \( \beta \)-CVaR constraints in (3.1a) are of the same type, although they are separate constraints for each voxel on the tumor. On the other hand, the sets of robust constraints (3.1a) and (3.1b) are of different types. Similarly, constraints for different \( \beta \) values or for different organs would be of different types.

Because the uncertainty set is polyhedral, an equivalent reformulation exists by simply enumerating the vertices of \( \mathcal{P} \). In general though, enumerating all the vertices of a polyhedron is NP-hard (Khachiyan et al., 2008) and could lead to an exponential number of constraints. Alternatively, because the original problem is linear and the uncertainty set is polyhedral, the robust counterpart of this problem which is a linear program can be used. However, the large number of robust constraints make the robust counterpart
very large. Thus, we consider constraint generation as an alternative solution method.

### 3.3 A Constraint Generation Solution Method

In this section, we first formulate a general form of the previous robust problem with multiple uncertain constraints. Then, we define the steps of the constraint generation algorithm and develop several constraint addition strategies.

#### 3.3.1 A robust optimization problem with uncertain constraints

Consider a robust optimization problem with uncertain constraints of different types $k \in \mathcal{K}$ that must hold for every $v \in \mathcal{V}(k)$. Let $w$ be the decision vector. The uncertainty is in the vector $\tilde{p} \in \mathcal{P}$, which is a parameter that affects all robust constraints. Recall that in the IMRT application, the uncertain parameter $\tilde{p}$ is the patient’s breathing pattern, which is modeled as a probability mass function (PMF) that captures the fraction of time that the patient spends in each breathing phase. We construct the uncertainty set $\mathcal{P}$ by including upper and lower error bounds on a nominal breathing pattern. All robust constraints must hold for all values of $\tilde{p}$ in the uncertainty set $\mathcal{P}$. Let $c$ be the vector of objective function coefficients and $A_{v,k}$ be the constraint coefficient matrix for constraint type $k \in \mathcal{K}$ for every $v \in \mathcal{V}(k)$. Given that $|\mathcal{I}|$ and $|\mathcal{B}|$ are the sizes of the vectors $\tilde{p}$ and $w$, respectively, the size of the matrix $A_{v,k}$ is $|\mathcal{I}| \times |\mathcal{B}|$ for each $v, k$. The parameter $a_{v,k}$ is the right hand side and corresponds to the upper or lower bound on the dose to be delivered. The general robust optimization problem is

\[
\begin{align*}
\text{minimize} & \quad c'w \\
\text{subject to} & \quad \tilde{p}'A_{v,k}w \geq a_{v,k}, \quad \forall v \in \mathcal{V}(k), \quad k \in \mathcal{K}, \quad \tilde{p} \in \mathcal{P}, \\
& \quad w \geq 0.
\end{align*}
\]
For simplicity, we omit constraints that do not depend on $\mathcal{P}$, which may also be present in the problem.

### 3.3.2 Problem decomposition

We decompose model (3.2) into a master problem and a subproblem. The master problem is a linear program. An optimal solution to the master problem is passed to the subproblem, which identifies constraints to be added to the master problem or provides a certificate of optimality for the original problem.

#### 3.3.2.1 The master problem

Formulation (3.3) shows a general form of the master problem.

\[
\begin{align*}
\text{minimize} \quad & c'w \\
\text{subject to} \quad & \tilde{p}'_{v,k}^e A_{v,k} w \geq a_{v,k}, \quad \tilde{p}_{v,k} \in \mathcal{P}_{v,k}^n, \quad \forall v \in \mathcal{V}(k), \quad k \in \mathcal{K}, \\
& w \geq 0.
\end{align*}
\]

The initial set $\mathcal{P}_{v,k}^1$ consists of a single $p \in \mathcal{P}$ for all $v \in \mathcal{V}(k), k \in \mathcal{K}$. At every iteration $n$, a finite number of constraints are added to the master problem, updating the set $\mathcal{P}_{v,k}^n$. Thus, the master problem is a linear program. An optimal solution $w^*$ is passed to the subproblem.
3.3.2.2 The subproblem

The subproblem (3.4) finds a distinct vector \( p_{v,k} \in \mathcal{P} \) for each \( v, k \) for the corresponding constraint in (3.3b) to maximize the violation of the constraints.

\[
\begin{align*}
\text{maximize} & \quad \sum_{v \in \mathcal{V}(k)} \sum_{k \in \mathcal{K}} \text{viol}_{v,k} \\
\text{subject to} & \quad \text{viol}_{v,k} = \max \{ 0, a_{v,k} - p'_{v,k} A_{v,k} w^* \}, \quad \forall v \in \mathcal{V}(k), \quad k \in \mathcal{K}, \quad (3.4b) \\
& \quad p_{v,k} \in \mathcal{P}, \quad \forall v \in \mathcal{V}(k), \quad k \in \mathcal{K}. \quad (3.4c)
\end{align*}
\]

If the optimal value of the subproblem is 0, then the current solution to the master problem is optimal. The optimal solution \( p^*_{v,k} \) generates the largest violation for the constraint index \( v \) of type \( k \). In other words, \( p^*_{v,k} \) is the breathing pattern that generates the worst-case violation of dose metric type \( k \) for voxel \( v \). Note that the subproblem is separable in \( v, k \). In general, the complexity of the subproblem depends on the uncertainty set \( \mathcal{P} \). Because we consider a polyhedral uncertainty set in this thesis, the subproblem is linear.

3.3.2.3 Constraint addition strategies

Our robust model (3.2) has \(|\mathcal{K}|\) sets of robust constraints, each constituting a set of constraints for all \( v \in \mathcal{V}(k) \). An optimal solution to the subproblem may identify numerous \( p^*_{v,k} \) that result in a positive violation of its corresponding constraint. This leads to several natural questions about adding constraints to the master problem:

- How many constraints should be added to the master problem at every iteration?

- Since different \( p^*_{v,k} \) vectors may generate worst case violations for different \( v \) and \( k \), which one(s) should be added to the master problem?

- Is it efficient to add a PMF, \( p \), for all \( v \) of the same constraint type \( k \)? In other
words, is there a PMF that generates a large aggregate violation over many \( v \in \mathcal{V}(k) \)?

Let \( p_{v,k}^* \) generate the maximum violation for constraint index \( v \) in constraint type \( k \). Let \( p^* \) be the PMF that generates the highest violation among all constraints, which corresponds to a particular index \( v^* \) and constraint type \( k^* \). In other words, \( p^* = p_{v^*,k^*}^* \) where \( (v^*, k^*) = \arg\max_{v,k} (\text{viol}_{v,k}^*) \). Some of the constraint addition strategies we explore will add the same PMF, \( p^* \), for all voxels and other strategies add different PMFs, \( p_{v,k}^* \), for each voxel. First assume we have only one type of robust constraint (\(|K| = 1\)) for all \( v \in \mathcal{V} \), and we need to choose constraint index \( v \) for which we need to add constraints to the master problem. We call these index-based strategies. These strategies are as follows:

**S1.** Add \( p^* \) for all \( v \in \mathcal{V}(k) \).

**S2.** Add \( p^* \) for all \( v \in \mathcal{V}(k) \) for which the maximum violation is greater than some threshold \( \delta \geq 0 \).

**S3.** Add \( p_{v,k}^* \) for all \( v \in \mathcal{V}(k) \) for which the maximum violation is greater than some threshold \( \delta \geq 0 \).

**S4.** Add \( p^* \) for only \( v^* \).

The strategies are ordered based on the number of constraints that they add at each iteration. Strategies S2 and S3 add the same number of constraints but for different \( p \) vectors in the first iteration.

Now consider the case where we have more than one type of constraint. All the index-based strategies can be implemented either for all constraint types or only for constraint type \( k^* \). We call this second decision the type-based strategy. We can test S2 and S3 using different values for \( \delta \). These strategies are summarized in Table 3.1.
Table 3.1: Strategies for adding constraints

<table>
<thead>
<tr>
<th>Strategy name</th>
<th>Added $p$</th>
<th>Index-based</th>
<th>Type-based</th>
<th>Constraints per it.</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1-1</td>
<td>$p^*$</td>
<td>$\forall v \in \mathcal{V}(k)$</td>
<td>$k = k^*$</td>
<td>$</td>
</tr>
<tr>
<td>S1-2</td>
<td></td>
<td>$\forall k$</td>
<td>$\sum_{k \in \mathcal{K}}</td>
<td>\mathcal{V}(k)</td>
</tr>
<tr>
<td>S2-1</td>
<td>$p^*$</td>
<td>$\forall v \in \mathcal{V}(k) :</td>
<td>\text{viol}_{v,k}</td>
<td>\geq \delta$</td>
</tr>
<tr>
<td>S2-2</td>
<td></td>
<td>$\forall k$</td>
<td>$\leq \sum_{k \in \mathcal{K}}</td>
<td>\mathcal{V}(k)</td>
</tr>
<tr>
<td>S3-1</td>
<td>$p^*_{v,k}$</td>
<td>$\forall v \in \mathcal{V}(k) :</td>
<td>\text{viol}_{v,k}</td>
<td>\geq \delta$</td>
</tr>
<tr>
<td>S3-2</td>
<td></td>
<td>$\forall k$</td>
<td>$\leq \sum_{k \in \mathcal{K}}</td>
<td>\mathcal{V}(k)</td>
</tr>
<tr>
<td>S4-1</td>
<td>$p^*$</td>
<td>$v = v^*$</td>
<td>$k = k^*$</td>
<td>1</td>
</tr>
<tr>
<td>S4-2</td>
<td></td>
<td>$\forall k$</td>
<td></td>
<td>$</td>
</tr>
</tbody>
</table>

3.3.2.4 Solution update

After updating the master problem with new constraints, standard dual simplex iterations are used to find a new optimal solution. The solution update time depends on the number of constraints added at each iteration, which depends on the constraint addition strategy employed. The last column in Table 3.1 shows the number of constraints that are added in each iteration for each strategy.

3.3.2.5 Stopping criterion

The algorithm terminates when the maximum violation of all constraints is less than a given tolerance $\epsilon > 0$. In Section 3.4, we explore the sensitivity of the results to different values of $\epsilon$.

3.3.2.6 Overview of the Solution method

Algorithm 1 shows an overview of the solution method.
Algorithm 1

1: \( n = 1 \). Let \( \mathcal{P}_{v,k}^1 \) be a single element from \( \mathcal{P} \), for all \( v \in \mathcal{V}(k), k \in \mathcal{K} \).
2: Solve master problem (3.3) with \( \mathcal{P}_{v,k}^n \) and pass optimal \( w^n \) to the subproblem.
3: Solve the subproblem (3.4) and find optimal \( p^n_{v,k} \) for all \( v \in \mathcal{V}(k), k \in \mathcal{K} \).
4: if the optimal value of the subproblem is positive and the stopping criterion is not met then
5: go to 9.
6: else
7: go to 10.
8: end if
9: Add new constraints to the set \( \mathcal{P}_{v,k}^n \) to construct \( \mathcal{P}_{v,k}^{n+1} \), increment \( n \), and go to 2.
10: Output \( w^n \) and stop.

3.4 Results

We applied the constraint generation approach with all the different addition strategies on the breast cancer IMRT treatment planning model (3.1). A clinical patient dataset was provided by the Princess Margaret Cancer Centre, Toronto, Canada. There were 6824 voxels in the target volume and 13249 voxels in the heart. The beam included 901 beamlets. The problem included upper and lower robust-CVaR constraints on the target, with \( \beta = 0.5\%, U = 45.79 \) and \( \beta = 95\%, L = 39.01 \), respectively. For the parameter \( \delta \) we test three different values of 0, 0.05 and 0.1Gy. We do not consider larger values of \( \delta \) since only a dose less than 0.1Gy is considered a clinically negligible dose to the target. The objective was to minimize the mean dose to the heart. There were 13,648 \( (\sum_{k \in \mathcal{K}} |\mathcal{V}(k)|) \) robust constraints in total. The constraint generation algorithm was coded using C++ and IBM ILOG CPLEX 12.1 was used to solve the optimization problems. The robust counterpart was solved using both C++/CPLEX and AMPL/CPLEX for comparison. The default pre-solve setting of CPLEX was used in all cases. All trials were run using a single Linux node of a Dell PowerEdge R410 computer with a 3.07 GHz 12-core CPU.
3.4.1 Computational efficiency

Table 3.2 shows the total run times for all constraint generation strategies and the robust counterpart. Strategy S3 tailors the added vector $p$ to each violated constraint indexed by $v$, which results in a faster solution time than strategy S2, which adds the same vector $p$ to all violated constraints. Both strategies S2 and S3 start with adding the same number of constraints in the first iteration (for the same $\delta$), but strategy S3 takes fewer iterations than strategy S2, since the constraints it adds are more effective. Overall, tailoring the added vector $p$ to each violated constraint produces the fastest solution times. For S2, using a $\delta$ value of 0 or 0.05 resulted in a shorter computation time than $\delta = 0.1$, whereas for S3, the computation time for $\delta = 0.1$ and $\delta = 0.05$ were similar. In general, strategy S3 (with any $\delta$) had a shorter computation time than the other strategies. Strategy S4 did not find an optimal solution within the time limit provided, so we did not further examine this strategy.

The robust counterparts were an order of magnitude slower than the fastest constraint generation approach. Table 3.2 also suggests tailoring the added constraint to the particular type of robust constraint that exhibited a violation may speed up the computation. AMPL’s pre-processing capability was the reason AMPL’s robust counterpart solved more quickly than the C++ equivalent.

Table 3.3 shows the breakdown between master and subproblem computation times for each constraint generation strategy. For all strategies, the time to solve the initial master problem was the same (4.2 min). The subproblem column reports the total time spent solving the subproblem over all iterations. Similarly, the time spent for updating the master problem over all iterations is shown.

Figure 3.1 shows the exact number of constraints that were added per iteration for strategies S2 and S3 for different $\delta$ values. Strategies S1 and S4 add a fixed number
Table 3.2: Computation times (min)

<table>
<thead>
<tr>
<th>Strategy</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>AMPL</th>
<th>C++</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>δ = 0</td>
<td>δ = 0.05</td>
<td>δ = 0.1</td>
<td>δ = 0</td>
<td>δ = 0.05</td>
<td>δ = 0.1</td>
</tr>
<tr>
<td>k = k∗</td>
<td>170</td>
<td>35</td>
<td>33</td>
<td>38</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>∀k</td>
<td>195</td>
<td>30</td>
<td>29</td>
<td>38</td>
<td>19</td>
<td>13</td>
</tr>
</tbody>
</table>

Table 3.3: Detailed computation times (min) and the number of iterations for each strategy. The total time includes post-optimization processing between the iterations of the algorithm.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Subproblem</th>
<th>Master problem</th>
<th>Total time</th>
<th># of iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1-1</td>
<td>0.29</td>
<td>156</td>
<td>170</td>
<td>40</td>
</tr>
<tr>
<td>S1-2</td>
<td>0.18</td>
<td>184</td>
<td>195</td>
<td>24</td>
</tr>
<tr>
<td>S2-1</td>
<td>δ = 0</td>
<td>0.37</td>
<td>19</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>δ = 0.05</td>
<td>0.38</td>
<td>19</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>δ = 0.1</td>
<td>0.54</td>
<td>17</td>
<td>38</td>
</tr>
<tr>
<td>S2-2</td>
<td>δ = 0</td>
<td>0.27</td>
<td>17</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>δ = 0.05</td>
<td>0.30</td>
<td>16</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>δ = 0.1</td>
<td>0.50</td>
<td>18</td>
<td>38</td>
</tr>
<tr>
<td>S3-1</td>
<td>δ = 0</td>
<td>0.05</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>δ = 0.05</td>
<td>0.07</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>δ = 0.1</td>
<td>0.07</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td>S3-2</td>
<td>δ = 0</td>
<td>0.06</td>
<td>12</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>δ = 0.05</td>
<td>0.05</td>
<td>7</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>δ = 0.1</td>
<td>0.04</td>
<td>7</td>
<td>13</td>
</tr>
</tbody>
</table>
of constraints and are not shown. The strategies that add \( \mathbf{p} \) vectors for all types of constraints (\( \forall k \)) have a smoother trend compared to those that add them for only one type (\( k = k^* \)). The reason is that different types of constraints will most likely have different worst-case \( \mathbf{p} \) vectors at each iteration.

Figure 3.2 compares the solution times (in minutes) of the strategies versus the average number of constraints that are added per iteration. It shows that adding a moderate number of constraints in each iteration generally results in the shortest computation time. Strategies S1 and S4 (omitted from the figure) exhibit the largest computation time and add the largest and the smallest number of constraints in each iteration, respectively.

### 3.4.2 Solution quality vs. computation time

The results presented so far are based on \( \epsilon = 0.1 \). In this section, we compare the solution quality and computation time for different values of \( \epsilon \) using strategy S3-1 with \( \delta = 0 \), which is the fastest strategy. To quantify the solution quality, we calculate the maximum violation (in units of Gy, which is a measure of radiation dose) of all constraints in the final solution.

Table 3.4 compares the results using different values of \( \epsilon \). It shows the number of iterations, total run time and the maximum violation from any of the constraints in units of Gray (Gy). The last two columns show the percentage of voxels that have a violation for some constraint and the mean of positive violations for all voxels, for the upper bound (U) and lower bound (L) constraints. It can be seen that as \( \epsilon \) is varied, the solution time is minimally affected and the number and magnitude of constraint violations is well-controlled. Figure 3.3 shows a dose-volume histogram (DVH) which provides a more clinical view of the solution quality. A DVH shows the fraction of an organ that receives a certain dose or higher. As Figure 3.3 illustrates, the solutions corresponding to \( \epsilon = 1 \) and \( \epsilon = 10^{-10} \) result in treatments with essentially identical dosimetric properties.
Figure 3.1: The number of constraints added at each iteration
Figure 3.2: Comparing the average number of constraints per iteration and total time

Figure 3.3: The effect of $\epsilon$ on solution quality
Table 3.4: The effect of using different values of $\epsilon$. The last two columns show the percentage of voxels that have a violation for some constraint and the mean of positive violations for all voxels, for the upper bound (U) and lower bound (L) constraints.

<table>
<thead>
<tr>
<th>$\epsilon$</th>
<th># it.</th>
<th>Run time (min)</th>
<th>Maximum violation (Gy)</th>
<th>Percentage of voxels</th>
<th>Mean of viols &gt; 0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>U</td>
<td>L</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>10.07</td>
<td>0.90</td>
<td>5.29</td>
<td>4.60</td>
</tr>
<tr>
<td>0.8</td>
<td>5</td>
<td>11.05</td>
<td>0.57</td>
<td>3.53</td>
<td>3.50</td>
</tr>
<tr>
<td>0.5</td>
<td>6</td>
<td>11.50</td>
<td>0.37</td>
<td>2.29</td>
<td>3.09</td>
</tr>
<tr>
<td>0.1</td>
<td>7</td>
<td>11.83</td>
<td>0.031</td>
<td>3.33</td>
<td>3.33</td>
</tr>
<tr>
<td>0.03</td>
<td>8</td>
<td>12.14</td>
<td>0.026</td>
<td>2.62</td>
<td>2.78</td>
</tr>
<tr>
<td>0.01</td>
<td>9</td>
<td>12.39</td>
<td>$1.8 \times 10^{-11}$</td>
<td>2.21</td>
<td>3.18</td>
</tr>
<tr>
<td>$10^{-10}$</td>
<td>9</td>
<td>12.39</td>
<td>$1.8 \times 10^{-11}$</td>
<td>2.21</td>
<td>3.18</td>
</tr>
</tbody>
</table>
3.5 Discussion

The fact that strategy S3-1 was the fastest constraint addition strategy reinforces the need to tackle individual voxels separately as each may have different maximum violations for different PMFs. It seems that adding a moderate number of constraints at each iteration resulted in the best performance. The natural strategy of adding one constraint at a time (S4) turned out to be the most inefficient, due to the large number of iterations that were required to converge. At the other end of the spectrum, adding the same vector \( p \) for each \( k \) and \( v \) also resulted in long computation times.

For strategies S1 and S3, adding the new constraint(s) for the same type \( k = k^* \) results in a shorter computation time than adding them for all \( k \). This confirms that identifying different types of constraints can help reduce the computation time. In the IMRT case, for example, inhale- and exhale-weighted breathing patterns can have different effects on upper and lower CVaR constraints on the dose to each organ.

Adding a large violation threshold \( \delta = 0.1 \) (used in S2 and S3) resulted in a slightly higher computation time in most of the cases, due to an increase in the number of iterations and corresponding decrease in the number of constraints added per iteration. Our hypothesis was that some constraints added earlier in the algorithm might obviate the need for additional constraints later, because those corresponding infeasibilities would have been addressed by the earlier constraints. This turned out not to be the case. For some strategies, \( \delta = 0.05 \) was more efficient than \( \delta = 0 \), while for the fastest strategy, S3-1, \( \delta = 0 \) resulted in the shortest computation time.

Finally, we solved the robust problem by enumerating all vertices of the uncertainty set (which was possible here since it was only five dimensional) as another comparison. This version of the problem took about 30 minutes to solve, which is faster than the robust counterpart, but still much slower than the best constraint generation strategy.
3.6 Conclusions

In this chapter, we developed a decomposition-based solution method for robust optimization problems with a large number of uncertain constraints. We developed strategies for finding and adding constraints at each iteration. We defined types of constraints and demonstrated the computational benefit of categorizing the constraints based on their type. We applied our method to a large-scale robust IMRT optimization model for breast cancer and compared the computation time of our solution method with that of solving the robust counterpart. Our results showed that constraint generation can typically save one order of magnitude in computation time.

We believe that using geometric information about the voxels (i.e., which voxels are in close proximity to each other) may result in further reductions in computation time for the breast cancer IMRT problem. Also, clustering the breathing patterns based on similarity and prioritizing the addition of different breathing patterns may also help improve computational efficiency. These are topics for future study.
Chapter 4

Robust Optimization for
Cardiac Sparing in Breast Cancer Radiation Therapy

4.1 Introduction

In this chapter, we explore the benefit of applying the robust optimization (RO) approach developed in Chapter 2 for cardiac sparing in tangential breast cancer intensity-modulated radiation therapy (IMRT). We use several clinical patient datasets that represent a range of different breast cancer patients from patients treated under deep inspiration breath-hold (DIBH) to patients who have a very small volume of the heart inside the radiation field. Using this approach, we design treatment plans that can be delivered during free-breathing for all patients by accounting for breathing motion with deformable registration and motion uncertainty with RO. We compare the outcomes of the robust plan with those of the clinical free-breathing plan for both free-breathing and breath-hold patients. We test the broad application of the robust method using these patient datasets. We also explore the use of the RO method in classifying patients appropriately
for using a DIBH treatment.

4.2 Methods and Materials

In this section, we first describe the patient data and the clinical methods for treatment planning. Then, we describe the deformable registration method used for modeling the breathing motion. Lastly, we explain how the accumulated dose is calculated and briefly describe the robust optimization method.

4.2.1 Study population and clinical IMRT treatment planning

Eight patients treated with standard whole breast IMRT from the Princess Margaret Cancer Centre were included in this study, approved by our institutional research ethics board with written informed consent. These were the first eight patients who consented to be included in this study. Patients included in the study were of Stage 0, I, or II left-sided breast cancer, and they were treated under normal free-breathing conditions (six patients) or controlled breath-hold using the active breathing control (ABC, Elekta, Crawley, UK) device (two patients: patient 2 and patient 4). Patients at our institution are classified based on the V50 volume of the heart (the volume that receives 50% of the prescribed dose) on the average scan (Wang et al., 2010). If the V50 of heart on the average scan is larger than 10cc, then the patient is selected to receive treatment under inhale breath-hold. For these patients, an active breathing control (ABC) device is used to help patients hold their breath at inhale (Krauss et al., 2005; Remouchamps et al., 2003), and a scan is acquired during inhale breath-hold (ABC scan). Such a treatment is planned based on the ABC scan and is delivered under breath-hold with the same settings, and thus there is minimal uncertainty in the heart position during treatment. Using an ABC technique is time and labour intensive at both planning simulation and treatment delivery, since the patients cannot hold their breath for the duration of the
treatment and the treatment has to pause when the patient catches her breath.

Planning simulation was acquired on a CT scanner (Brilliance, Philips Medical Systems, Milpitas, CA), where patients were positioned and immobilized as per our current clinical protocol (Purdie et al., 2011, 2014). A 4D-CT image acquisition generated ten breathing phase datasets for each patient (including those who received an ABC treatment). An average CT image was also reconstructed based on a pixel-by-pixel averaging of the 4D-CT scan.

The CT datasets were then transferred to a treatment planning system (Pinnacle, Philips Healthcare, Fitchburg, WI) for standard treatment planning on the average CT. Treatment plans were generated by a radiation therapist using an automated planning method and our standard clinical objectives (Purdie et al., 2011). The target volume (optBreast) was defined as the volume irradiated by the tangential beams at the 55% isodose level, excluding the left lung and heart, and then contracting 0.5 cm in the axial plane and 1.0 cm in the superior-inferior direction. The optBreast is the same volume as the clinical target volume (CTV) in Chapter 2. The prescribed dose for all patients was 4240 cGy in 16 fractions.

4.2.2 Deformable registration

The motion due to breathing was modeled with deformable registration based on the 4D-CT datasets using RayStation (version 3.99.0.8, RaySearch Laboratories AB, Stockholm, Sweden). The registration method used required the manual delineation of the heart, left lung and left breast for each 4D-CT imaging dataset and also on the average CT dataset. These ROIs were contoured by the same person on the 4D-CT images and the average CT. The manually delineated volumes were used for dose evaluation.

A hybrid deformable registration method based on image intensity and deforming the regions of interest (ROIs) was used in this study (Weistrand and Svensson, 2015). To capture the ROI deformation, triangular surface meshes were first generated from the
contours of the heart, left lung and left breast, in order to compute displacement boundary conditions. Then, they were set as controlling ROIs to be modeled as linear elastic materials with voxel compression ratios of 0.48 to match the image pairs. Then, these ROIs were also set as focus ROIs for highlighting a specific region for higher accuracy. The average CT was the reference dataset, and each 4D-CT phase dataset was the target dataset, giving ten registration pairs. The registration grid was isotropic with a voxel size of 0.25 cm.

The deformable registration resulted in a displacement vector field from the average CT to each phase. To summarize the displacement over a particular ROI, this vector field was linearly interpolated from the registration grid onto the dose grid, as these two grids were not aligned, and the functionality to classify which voxels belong to which ROI was only available on the dose grid. For all patients, the projected displacement on the normal of the tangent plane as well as the overall displacement (in any direction) for the optBreast and the heart are shown in Figure 4.1. The particular direction of interest

![Graphs showing displacement](image)

(a) Displacement projected on the normal of the tangent plane

(b) Overall displacement

Figure 4.1: Magnitude of displacement of the voxels over all breathing phases. The upper and lower errorbars show the 90th and 10th percentile, respectively.

was normal to the tangent plane, which allows us to capture the motion in and out of
the treatment field. This direction is depicted by a dotted arrow in Figure 4.2(b). The tangent plane was formed by the edge of the opposing beams that was intersecting the isocenter of the plan. Then the displacement vector was projected onto this normal vector and the motion was quantified as the magnitude of the projected displacement as well as the original displacement. Displacement was summarized over the whole optBreast volume; for the heart, it was summarized over the region inside the field plus 2.5 cm from the tangent plane outside of the field. For the projected displacement on the normal of the tangent plane, the displacement is computed from average CT image to each phase, where positive is going into the field and negative is going out of the field.

4.2.3 Dose accumulation and robust optimization

Our optimization method utilizes a precalculated influence matrix, $D_{i,v,b}$, that captures the dose that voxel $v$ receives per unit intensity of beamlet $b$, for each beam on phase $i$, which is mapped onto the average CT image. The dose grid was isotropic with a voxel size of 0.25 cm, and the fluence grid was isotropic with a beamlet size of 0.5 cm. The $D_{i,v,b}$ matrices for each beam and each phase $i$ were extracted in our research version of RayStation. Finally, the $D_{v,b}^{i,0}$ matrix was defined as the $D_{v,b}^{i}$ mapped onto the average CT (phase 0) in MATLAB (MATLAB®, 2013) such that $D_{v,b}^{i,0} = \mathbf{S}_{i,0}^{v,b} D_{v,b}^{i}$, where $\mathbf{S}_{i,0}^{v,b}$ is the interpolation matrix (Trofimov et al., 2005) from phase $i$ to the average CT. We note that the average CT was used as the reference scan for mapping the dose, and we are not propagating contours from the average scan to the phases.

The clinical plan was also imported into RayStation for extracting the fluence of each beam. The clinical fluence grid had a beamlet size of 0.2 cm by default and was linearly interpolated onto our 0.5 cm fluence grid to be compatible with our exported $D_{v,b}^{i}$ matrices for computing and accumulating dose over the breathing cycle.

The $D_{v,b}^{i}$ matrices were used to accumulate the dose on all phases for both the clinical plan and the robust plan. For the robust plan, we set two dose volume criteria of
Figure 4.2: The beam’s eye view and the axial view on the average and ABC scans for patient 2
D95% and D0.5% on the optBreast and aimed to minimize the mean dose to the heart. These criteria are formulated as a mathematical optimization problem where the objective function is minimizing the heart dose and the constraints are the dose volume criteria that must hold for a large range of breathing patterns, as modeled by our uncertainty set presented in Chapter 2. The optimization problems were coded in C++ and solved using the IBM ILOG CPLEX 12.3 solver. In order to incorporate the uncertainty in breathing motion into the optimization, we divided the patient’s breathing cycle into 5 phases from inhale to exhale. The five phases were represented by five of the ten scans in the 4D-CT dataset that captured the phases from inhale to exhale and were representative of the range of the motion. Each breathing pattern can be interpreted as a probability mass function (PMF) where the probability of each phase is the fraction of time spent at that phase. We assume that in a nominal breathing pattern the patient spends 50% of the time at exhale and the rest of the time equally divided among the other four phases. We consider ±10% error bounds above and below the nominal breathing pattern to construct the uncertainty set, thus including 40% to 60% exhale and 2.5% to 22.5% in each of the other phases. This is a fairly large uncertainty set since it includes a range of 20% uncertainty for each of the five phases.

We accumulate the dose on the clinical and robust plans using 200 simulated breathing patterns (100 inside and 100 outside of the uncertainty set) uniformly drawn over the probability simplex. We used PMFs both inside and outside of the uncertainty set to check the sensitivity of the results to unusual breathing patterns. For the clinical case, the plan on the average scan is compared to the accumulated dose on all phases using the set of simulated breathing PMFs. For the robust case, the nominal plan (using the nominal breathing pattern to accumulate dose) is compared to the accumulated dose using the simulated breathing PMFs. To facilitate the comparison, we normalize both the robust plan and the clinical plan to the mean prescribed dose of 4240 cGy on the optBreast. Then, the normalized plans are used to calculate the accumulated dose in
both cases. The mean normalization factor was 0.99 and 0.95 for the average and robust methods, respectively.

4.3 Results

In this section, we first compare the difference between the planned dose and the accumulated dose for each method. Next, we calculate metrics on heart sparing and target coverage and compare the two methods with respect to these metrics. Lastly, we quantify the benefit of using an RO approach by calculating the absolute difference between the accumulated doses of the two methods.

4.3.1 Planned dose versus accumulated dose

First we compare the dose accumulated using the simulated PMFs with the planned dose for each method. Figure 4.3 shows the dose-volume histograms of the planned dose versus the accumulated dose for each patient and for each of the two planning methods, robust and clinical. The dashed curves represent the planned dose and the clouds of solid curves represent the accumulated dose for the set of simulated PMFs. It can be seen that in the clinical method, the planned dose differs from the actual accumulated dose, whereas the planned and accumulated doses are almost identical in the robust method. In the clinical method, the difference can be seen in both the optBreast dose and the heart dose.

A complete comparison of relevant clinical metrics on the planned and accumulated dose for each planning method are given in the Appendix in Table 4.2. For the robust method, the values of the metrics for the planned dose always lie inside the range of the metrics for the accumulated dose using the simulated breathing PMFs. For the clinical method, on the other hand, the metrics of the planned dose are typically outside the range for the accumulated dose.
Figure 4.3: Comparison of the planned and accumulated dose-volume histograms. The dashed lines represent the planned dose and the solid lines represent the accumulated dose.
4.3.2 Heart sparing and target coverage

Next, we compare the robust and the clinical method with respect to heart sparing and target coverage. Figure 4.4 shows the heart D10cc (the dose to the hottest 10cc of the heart) and the optBreast D99% (the dose to 99% of optBreast) for both the robust and the clinical methods for all patients. Each point represents one simulated breathing pattern. The planned dose is also shown for each method.

In Figure 4.4, it can be seen that the simulated cloud of the accumulated robust doses dominates that of the accumulated clinical doses both in terms of target coverage and heart sparing for all patients. Figure 4.4 also confirms that in the robust method, the planned dose is similar to the accumulated dose, whereas in the clinical method, the planned dose is very far away from the simulated cloud of accumulated doses. Table 4.2 also confirms that the robust accumulated dose dominates the clinical accumulated dose in the relevant metrics.

4.3.3 Breathing motion and the benefit of RO

In Table 4.1, we calculated the maximum motion from inhale to exhale in the direction normal to the tangent plane for optBreast and Heart. We then calculated the difference between the planned and accumulated dose distributions for both robust and clinical methods by finding the absolute difference of planned dose and the mean of accumulated doses. We also calculated the improvement gained by the robust method over the clinical method in each metric using the absolute difference between the mean of the cloud of accumulated doses of the two methods. In Table 4.1, the column “Error” shows the absolute difference between the planned dose and the accumulated dose, for each method. The column “RO Benefit” shows the absolute dosimetric benefit gained by using RO over the clinical method for each metric.

The robust method is not sensitive to the magnitude of motion and regardless of the
Figure 4.4: Comparison of D99% optBreast and D10cc Heart for the simulated breathing patterns. Note that the vertical axes are in reverse direction, and although the axes are on the same scale for all patients, their ranges vary per patient.
Table 4.1: Comparison of robust and clinical methods with different motion magnitudes for each patient. The reported values are the mean value over all simulated breathing patterns.

<table>
<thead>
<tr>
<th>Patient</th>
<th>Organ</th>
<th>Motion (cm)</th>
<th>Metric</th>
<th>Absolute Error (cGy)</th>
<th>RO Benefit (cGy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>optBreast</td>
<td>0.11</td>
<td>D0.5%</td>
<td>2</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>0.31</td>
<td>D10cc</td>
<td>8</td>
<td>182</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.6</td>
</tr>
<tr>
<td>2</td>
<td>optBreast</td>
<td>0.12</td>
<td>D0.5%</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>0.21</td>
<td>D10cc</td>
<td>39</td>
<td>471</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>140</td>
</tr>
<tr>
<td>3</td>
<td>optBreast</td>
<td>0.31</td>
<td>D0.5%</td>
<td>6</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>0.36</td>
<td>D10cc</td>
<td>119</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>199</td>
</tr>
<tr>
<td>4</td>
<td>optBreast</td>
<td>0.18</td>
<td>D0.5%</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>0.36</td>
<td>D10cc</td>
<td>6</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>182</td>
</tr>
<tr>
<td>5</td>
<td>optBreast</td>
<td>0.08</td>
<td>D0.5%</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>0.23</td>
<td>D10cc</td>
<td>11</td>
<td>154</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>522</td>
</tr>
<tr>
<td>6</td>
<td>optBreast</td>
<td>0.05</td>
<td>D0.5%</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>0.41</td>
<td>D10cc</td>
<td>30</td>
<td>762</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>322</td>
</tr>
<tr>
<td>7</td>
<td>optBreast</td>
<td>0.11</td>
<td>D0.5%</td>
<td>2</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>0.42</td>
<td>D10cc</td>
<td>115</td>
<td>558</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>420</td>
</tr>
<tr>
<td>8</td>
<td>optBreast</td>
<td>0.34</td>
<td>D0.5%</td>
<td>8</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>0.72</td>
<td>D10cc</td>
<td>0</td>
<td>275</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>
range of motion, there is almost always a benefit in using the RO approach. In particular, there is always a benefit in the heart dose and this improvement is typically substantial. The only exception is the D0.5% of optBreast for patient 4, in which there is a slight overdose to the optBreast as a result of dose normalization. If we adjust the dose to reduce this overdose, the other metrics for this patient will further improve.

The error in the robust method is almost always smaller than that of the clinical method. This robust error is in part related to the magnitude of motion. The one exception is patient 3 where a large reduction in the heart dose was possible since there was a large heart motion in the favorable direction (out of the radiation field). Therefore, there was a large difference between the planned and accumulated heart dose in the robust method. Except for this one case, the error of the robust method is always very small and less than the error of the clinical method.

For the clinical method, the error partly depends on the motion from the phases to the average scan (Figure 4.1). Therefore, it does not seem to directly relate to the magnitude of motion between phases. For patient 3, both optBreast and heart had a large displacement with respect to the average scan, and as a result, there is a large difference between the planned and accumulated dose for both of these organs.

The case of Patient 8 was special because the heart was not inside the field in any of the phases. Therefore, the accumulated D10cc on the heart was zero for both the clinical and the robust methods. On the average scan, however, the heart was contoured such that there was a small dose to the heart which resulted in the high error for the clinical method in this case.

4.4 Discussion

For all patients, the robust planned dose is very close to the cloud of simulated accumulated doses. Figures 4.3 and 4.4 and Table 4.2 confirm that for all patients, the robust plan
provides a better prediction of the accumulated dose compared to the clinical method, because the robust method uses the information of all the breathing phases on the 4D-CT and takes into account breathing uncertainty. As a result, the robust method can ensure that even under breathing motion uncertainty the delivered dose will closely approximate the planned dose. We note that the clinical heart dose on the average scan is higher than the clinical accumulated dose using the phases. The heart contour on the average scan approaches the limit of the union of the heart contours for all the phases as the heart contours for each phase will always be smaller. Therefore, it is expected that there is less low dose on the clinical accumulated DVH compared to the clinical planned DVH.

Similarly, for all patients, the robust plan provides better heart sparing and better target coverage compared to the clinical accumulated dose. In other words, the robust dose dominates the clinical plan in both metrics simultaneously. Currently, the doses are normalized so that the mean optBreast dose is set to the prescribed dose. For all patients, the D95% and D99% of optBreast are higher in the robust method compared to the clinical method. If the doses are normalized to these metrics instead of mean optBreast, the robust plan can provide an even lower dose to the heart compared to the clinical method. Overall, the robust plan provides better cardiac sparing for patients under free breathing without any degradation of target coverage.

All the robust results meet the clinical acceptability criteria on the heart except for patient 4 who was an ABC patient. In this case, the robust method reduced the heart dose compared to the clinical accumulated dose, but was still not able to meet the clinical criteria due to the substantial amount of heart in the field throughout the breathing cycle. Thus, although the robust method can result in the reduction of heart dose compared to the clinical method, ABC cases cannot be avoided entirely. In other words, the robust approach can potentially reduce the number of cases that require an ABC treatment, but cannot fully replace the ABC method. Overall, we believe our robust approach can be applied to any case that currently is treated using free-breathing and realize a positive
There was no direct relationship between the magnitude of motion and the difference between the RO and clinical methods. For example, some patients with small motion experienced a large benefit with RO (e.g., patient 5). The difference in performance between the RO and clinical methods is due to many factors including the quality of the average scan, the position of the organs (especially the heart) inside the field, and the direction and the magnitude of motion. For example, if the heart motion is large but the heart always remains outside of the field, then there is little deviation to the heart dose and perhaps little benefit of using RO.

Lastly, the RO approach may also be useful in patient classification. Currently, patients at our institution are classified based on the V50 of the heart on the average scan. Using this metric, patients 2 and 4 were classified as ABC patients. However, the V50 volume on the average scan itself depends on the breathing pattern at the time of planning. Thus, it is an uncertain value. Alternatively, patients could be classified based on the accumulated dose to the heart using the robust plan. Because the planned dose distribution using the robust method better predicts the accumulated dose, patients can be better classified based on the robust plan. The benefit of such a classification is that the robust plan will more accurately predict which patients need ABC, whereas the current clinical method might not provide a correct prediction of the accumulated dose. For example, patient 2 was classified as an ABC patient based on the average scan, but a free-breathing treatment using the robust method would have resulted in an acceptable plan and avoided ABC.

We presented a fluence-based robust optimization approach for breast IMRT to reduce the dose to the heart under free breathing. Currently, the optimization solver is separate from the treatment planning system, and the next step is to integrate them. Future work consists of placing constraints on the apertures and developing robust direct aperture optimization models.
4.5 Conclusions

In this chapter, we tested the applicability of a previously developed robust optimization method for left-sided breast IMRT using a set of clinical 4D-CT patient datasets. Clinical plans based on free-breathing were compared with robust plans when accumulated over breathing phases from inhale to exhale. Results for 200 simulated breathing patterns showed that the robust method reduces the accumulated dose to the heart for all patients. The robust plan had a much lower deviation from the accumulated dose compared to the clinical plan for all patients. Overall, we believe that the robust method can be used whenever the clinical method with free breathing is currently used. In addition, since the robust plan can meet the clinical acceptability limits under free breathing, the robust approach can potentially reduce the need for breath-hold techniques.

4.6 Appendix

In Table 4.2, we compare a set of clinical metrics for the simulated breathing patterns for all patients. The results for the robust, clinical free breathing and ABC plan (if available) are compared. The “Planned” columns show the value of the metric on the robust plan (using the nominal PMF) and the clinical plan (on the average scan), respectively. The “accumulated” columns show the ranges of the metrics for the accumulated dose using the simulated PMFs. The ABC dose is calculated using the plan on the ABC (inhale breath-hold) scan, which was only available for patients who were selected to use an ABC device for treatment. In the optBreast, the D0.5%, D1% and D2% represent the overdose where the lower values are preferred, and D95% and D99% represent the underdose and higher values are preferred. The mean optBreast dose should be as close as possible to the prescribed dose, 4240cGy. The value of the heart metrics should be as low as possible. Figure 4.5 compares the D10cc of Heart and D95% of optBreast for all patients for the simulated PMFs.
Table 4.2: Comparison of important metrics for 200 realizations of breathing patterns

<table>
<thead>
<tr>
<th>Patient</th>
<th>Organ</th>
<th>Metric</th>
<th>Robust method</th>
<th>Clinical method</th>
<th>ABC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Planned</td>
<td>Accumulated</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>optBreast</td>
<td>D0.5%</td>
<td>4432</td>
<td>[4432, 4442]</td>
<td>4427 [4459, 4469]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D1%</td>
<td>4409</td>
<td>[4409, 4420]</td>
<td>4412 [4446, 4455]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D2%</td>
<td>4388</td>
<td>[4387, 4402]</td>
<td>4393 [4428, 4436]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D95%</td>
<td>4076</td>
<td>[4073, 4081]</td>
<td>4086 [3933, 4008]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D99%</td>
<td>3873</td>
<td>[3858, 3920]</td>
<td>3892 [3605, 3704] N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>4240</td>
<td>[4239, 4246]</td>
<td>4240 [4254, 4262]</td>
</tr>
<tr>
<td>2</td>
<td>optBreast</td>
<td>D1cc</td>
<td>1595</td>
<td>[1536, 1601]</td>
<td>3486 [3016, 3050]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D10cc</td>
<td>352</td>
<td>[347, 376]</td>
<td>1343 [1146, 1183]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D25cc</td>
<td>24</td>
<td>[22, 34]</td>
<td>437 [61, 88]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>17</td>
<td>[17, 18]</td>
<td>166 [45, 47]</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>D0.5%</td>
<td>4411</td>
<td>[4410, 4421]</td>
<td>4533 [4551, 4556] 4539</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D1%</td>
<td>4391</td>
<td>[4391, 4401]</td>
<td>4516 [4530, 4534] 4522</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D2%</td>
<td>4378</td>
<td>[4377, 4384]</td>
<td>4490 [4506, 4510] 4497</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D95%</td>
<td>4060</td>
<td>[4055, 4065]</td>
<td>4041 [3870, 3903] 4044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D99%</td>
<td>3787</td>
<td>[3772, 3830]</td>
<td>3825 [3452, 3491] 3842</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>4240</td>
<td>[4239, 4243]</td>
<td>4240 [4236, 4238] 4240</td>
</tr>
<tr>
<td>2</td>
<td>Heart</td>
<td>D1cc</td>
<td>3933</td>
<td>[3902, 4002]</td>
<td>3953 [3671, 3721] 744</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D10cc</td>
<td>2169</td>
<td>[2098, 2367]</td>
<td>3144 [2618, 2780] 388</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D25cc</td>
<td>800</td>
<td>[741, 944]</td>
<td>1802 [1300, 1531] 304</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>94</td>
<td>[90, 104]</td>
<td>315 [117, 133] 130</td>
</tr>
</tbody>
</table>

Continued on the next page
Table 4.2 continued:

<table>
<thead>
<tr>
<th>Patient</th>
<th>Organ</th>
<th>Metric</th>
<th>Robust method</th>
<th>Clinical method</th>
<th>ABC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Planned</td>
<td>Accumulated</td>
<td>Planned</td>
</tr>
<tr>
<td>3</td>
<td>optBreast</td>
<td>D0.5%</td>
<td>4452 [4451, 4479]</td>
<td>4469 [4538, 4546]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>D1%</td>
<td>4437 [4436, 4454]</td>
<td>4449 [4513, 4522]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>D2%</td>
<td>4427 [4426, 4438]</td>
<td>4429 [4487, 4495]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>D95%</td>
<td>3951 [3937, 3952]</td>
<td>4020 [3628, 3670]</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>D99%</td>
<td>3825 [3780, 3829]</td>
<td>3813 [3042, 3155]</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>4240 [4236, 4241]</td>
<td>4240 [4210, 4214]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>D1cc</td>
<td>2814 [1651, 3068]</td>
<td>3181 [1933, 1975]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>D10cc</td>
<td>661 [235, 735]</td>
<td>671 [364, 967]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>D25cc</td>
<td>24 [0, 28]</td>
<td>315 [1, 39]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>26 [12, 29]</td>
<td>128 [16, 33]</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>optBreast</td>
<td>D0.5%</td>
<td>4494 [4493, 4509]</td>
<td>4460 [4450, 4465]</td>
<td>4483</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D1%</td>
<td>4476 [4475, 4489]</td>
<td>4441 [4434, 4450]</td>
<td>4444</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D2%</td>
<td>4458 [4457, 4469]</td>
<td>4417 [4418, 4432]</td>
<td>4413</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D95%</td>
<td>3910 [3902, 3931]</td>
<td>4051 [3812, 3929]</td>
<td>3974</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>D99%</td>
<td>3548 [3479, 3720]</td>
<td>3639 [2975, 3330]</td>
<td>3616</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>4240 [4237, 4255]</td>
<td>4240 [4208, 4229]</td>
<td>4240</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>D1cc</td>
<td>4221 [4196, 4266]</td>
<td>4273 [4304, 4326]</td>
<td>1180</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D10cc</td>
<td>3941 [3934, 3975]</td>
<td>4042 [4113, 4153]</td>
<td>376</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D25cc</td>
<td>3487 [3361, 3598]</td>
<td>3001 [3452, 3686]</td>
<td>266</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>244 [224, 262]</td>
<td>329 [260, 299]</td>
<td>98</td>
</tr>
<tr>
<td>Patient</td>
<td>Organ</td>
<td>Metric</td>
<td>Robust method</td>
<td>Clinical method</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>--------</td>
<td>---------------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Planned</td>
<td>Accumulated</td>
<td>Planned</td>
</tr>
<tr>
<td>5</td>
<td>optBreast</td>
<td>D0.5%</td>
<td>4383</td>
<td>[4381, 4406]</td>
<td>4527</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D1%</td>
<td>4368</td>
<td>[4367, 4384]</td>
<td>4504</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D2%</td>
<td>4359</td>
<td>[4359, 4370]</td>
<td>4474</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D95%</td>
<td>4099</td>
<td>[4093, 4100]</td>
<td>4018</td>
</tr>
<tr>
<td>5</td>
<td>Heart</td>
<td>D99%</td>
<td>3898</td>
<td>[3886, 3940]</td>
<td>3806</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>4240</td>
<td>[4239, 4243]</td>
<td>4240</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D1cc</td>
<td>3818</td>
<td>[3792, 3897]</td>
<td>4056</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D10cc</td>
<td>737</td>
<td>[614, 783]</td>
<td>1094</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D25cc</td>
<td>33</td>
<td>[20, 42]</td>
<td>358</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>39</td>
<td>[33, 42]</td>
<td>156</td>
</tr>
<tr>
<td>6</td>
<td>optBreast</td>
<td>D0.5%</td>
<td>4380</td>
<td>[4379, 4392]</td>
<td>4445</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D1%</td>
<td>4363</td>
<td>[4363, 4368]</td>
<td>4430</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D2%</td>
<td>4352</td>
<td>[4351, 4354]</td>
<td>4410</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D95%</td>
<td>4123</td>
<td>[4122, 4125]</td>
<td>4067</td>
</tr>
<tr>
<td>6</td>
<td>Heart</td>
<td>D99%</td>
<td>3915</td>
<td>[3901, 3956]</td>
<td>3851</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>4240</td>
<td>[4239, 4243]</td>
<td>4240</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D1cc</td>
<td>3039</td>
<td>[2680, 3245]</td>
<td>3678</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>D10cc</td>
<td>241</td>
<td>[216, 348]</td>
<td>1355</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D25cc</td>
<td>1</td>
<td>[0, 3]</td>
<td>419</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>17</td>
<td>[16, 21]</td>
<td>181</td>
</tr>
</tbody>
</table>

Continued on the next page
Table 4.2 continued:

<table>
<thead>
<tr>
<th>Patient</th>
<th>Organ</th>
<th>Metric</th>
<th>Robust method</th>
<th>Clinical method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Planned</td>
<td>Accumulated</td>
</tr>
<tr>
<td>optBreast</td>
<td></td>
<td>D0.5%</td>
<td>4406</td>
<td>[4405, 4415]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D1%</td>
<td>4388</td>
<td>[4387, 4394]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D2%</td>
<td>4376</td>
<td>[4375, 4381]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D95%</td>
<td>4068</td>
<td>[4066, 4072]</td>
</tr>
<tr>
<td>7</td>
<td>Heart</td>
<td>D99%</td>
<td>3857</td>
<td>[3838, 3895]</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>Mean</td>
<td>4240</td>
<td>[4239, 4244]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D1cc</td>
<td>1641</td>
<td>[1643, 1859]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D10cc</td>
<td>1425</td>
<td>[1304, 2035]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D25cc</td>
<td>269</td>
<td>[206, 415]</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>Mean</td>
<td>49</td>
<td>[46, 63]</td>
</tr>
<tr>
<td>optBreast</td>
<td></td>
<td>D0.5%</td>
<td>4410</td>
<td>[4409, 4452]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D1%</td>
<td>4393</td>
<td>[4393, 4428]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D2%</td>
<td>4383</td>
<td>[4382, 4407]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D95%</td>
<td>4018</td>
<td>[4002, 4023]</td>
</tr>
<tr>
<td>8</td>
<td>Heart</td>
<td>D99%</td>
<td>3813</td>
<td>[3780, 3820]</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>Mean</td>
<td>4240</td>
<td>[4239, 4246]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D1cc</td>
<td>0.3528</td>
<td>[0.0018, 0.8586]</td>
</tr>
<tr>
<td>Heart</td>
<td></td>
<td>D10cc</td>
<td>0</td>
<td>[0, 0]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D25cc</td>
<td>0</td>
<td>[0, 0]</td>
</tr>
<tr>
<td></td>
<td>Heart</td>
<td>Mean</td>
<td>0.0109</td>
<td>[0.0025, 0.0247]</td>
</tr>
</tbody>
</table>
Chapter 4. RO for Cardiac Sparing in Breast Cancer RT

(a) Patient 1

(b) Patient 2
Chapter 4. RO for Cardiac Sparing in Breast Cancer RT

(c) Patient 3

(d) Patient 4
(e) Patient 5

(f) Patient 6
Figure 4.5: Comparison of D99% optBreast and D10cc Heart for the simulated breathing patterns for all patients
Chapter 5

Pareto Robust Optimization in Breast Cancer Radiation Therapy

5.1 Introduction

This chapter explores Pareto robust optimization (PRO) in robust optimization (RO) models for breast cancer radiation therapy. Traditional RO models only focus on the worst-case scenarios that can be realized in the uncertainty set and do not optimize for non-worst-case scenarios. On the other hand, Pareto robust optimization can potentially improve the solution of an RO problem by improving the outcome of non-worst-case scenarios while maintaining the worst-case performance. In the breast cancer radiation therapy problems considered in the previous chapters, finding a PRO solution ensures that there are no other treatment plans that provide the same worst-case performance but strictly dominate the PRO plan in non-worst-case scenarios. In the rest of this chapter, we first briefly review the concept of Pareto robust optimization, and then formulate and solve PRO models for the breast cancer radiation therapy application.
5.2 Pareto Robust Optimization

An RO solution is called a PRO solution if there is no other solution that performs at least as well in all scenarios and strictly better in some non-worst-case scenarios over the other points in the uncertainty set (Iancu and Trichakis, 2013). In other words, it cannot be Pareto dominated by any other solution. Consider the following RO problem with a robust min-max objective function.

\[
\begin{align*}
\text{minimize} & \quad \max_{\tilde{u} \in \mathcal{U}} \tilde{u}'w, \\
\text{subject to} & \quad w \in X,
\end{align*}
\]

where \( X \) is the feasible region of the problem, \( w \) is the decision vector, and \( \tilde{u} \) is the uncertain parameter in an uncertainty set \( \mathcal{U} \). Let \( X^{RO} \) be the set of all optimal solutions to the RO problem (5.1). A solution \( w \in X^{RO} \) is PRO if

\[
\begin{align*}
\tilde{u}'x \leq \tilde{u}'w & \quad \forall \tilde{u} \in \mathcal{U}, \quad \text{and} \\
u'x < u'w & \quad \text{for some} \quad u \in \mathcal{U}.
\end{align*}
\]

That is, there is no other feasible solution \( x \) that performs at least as good as the RO solution \( w \) for all scenarios and strictly better for some scenarios. Equivalently, we can say \( w \) is a PRO solution if \( \forall x \in X \):

\[
\begin{align*}
u'x > u'w, & \quad \text{for some} \quad u \in \mathcal{U} \quad \text{or} \\
\tilde{u}'x \geq \tilde{u}'w & \quad \forall \tilde{u} \in \mathcal{U}.
\end{align*}
\]

In other words, \( w \) is a PRO solution if for all feasible solutions \( x \in X, w \) is at least as good as \( x \) in all scenarios, or \( w \) is strictly better than \( x \) in at least one scenario.

To test if the set of RO and PRO solutions are equivalent, we can use model (5.4),
which is equivalent to the optimization problem in Corollary 2 in Iancu and Trichakis (2013):

\[
\begin{align*}
\text{minimize} & \quad \hat{u}'y, \\
\text{subject to} & \quad y \in \mathcal{U}^*, \\
& \quad w - y \in X, \\
& \quad w \in X^{\text{RO}}.
\end{align*}
\]

where \(\hat{u}\) is an arbitrary point in the relative interior of the uncertainty set \(U\). We denote this by \(\hat{u} \in \text{ri}(\mathcal{U})\). Also, \(\mathcal{U}^*\) is the dual cone of the uncertainty set \(\mathcal{U}\), defined as \(\mathcal{U}^* \overset{\text{def}}{=} \{y \in \mathbb{R}^n \mid y^\top u \geq 0, \forall u \in \mathcal{U}\}\). At optimality, if \(y^*\) is non-zero, the RO solution \(w^*\) is dominated by the PRO solution \(w^* - y^*\). On the other hand, if \(y^* = 0\), the RO solution \(w^*\) is a PRO solution and the sets of RO and PRO solutions are equal. This means that all RO solutions are also PRO solutions.

To test if a given RO solution, \(w^*\), is a PRO solution, we can use model (5.5), which is equivalent to the optimization problem in Theorem 1 in Iancu and Trichakis (2013).

This formulation aims to find a solution \(w^* - y\) that Pareto dominates \(w^*\). If \(y^* = 0\), the RO solution \(w^*\) is PRO. Using a \(\hat{u} \in \text{ri}(\mathcal{U})\), this model can be formulated as follows:

\[
\begin{align*}
\text{minimize} & \quad \hat{u}'y, \\
\text{subject to} & \quad y \in \mathcal{U}^*, \\
& \quad w^* - y \in X.
\end{align*}
\]

If \(y^* \neq 0\), then \(w^* - y^*\) is a PRO solution that dominates \(w^*\). If \(y^* = 0\), then \(w^*\) is a PRO solution. The latter does not necessarily mean that the sets of RO and PRO solutions are equal, it only shows that the RO solution \(w^*\) is PRO; there might be other
RO solutions that are not PRO.

Finally, to find a PRO solution over the set of RO solutions, we can use model (5.6) which is equivalent to the optimization problem in Corollary 1 in Iancu and Trichakis (2013).

\[
\begin{align*}
\minimize \quad & \hat{u}' w, \\
\text{subject to} \quad & w \in X^{RO}.
\end{align*}
\] (5.6a) (5.6b)

Different \( \hat{u} \) vectors can lead to different PRO solutions. For any PRO solution, there exists some \( \hat{u} \) that will result in that PRO solution when used in model 5.6 (see Proposition 1 in Iancu and Trichakis (2013)).

### 5.3 Application to Breast Cancer Radiation Therapy

Recall that in our breast cancer RT problem (i.e., model (2.9)), the objective is to minimize the heart dose and the constraints are on the dose to the target. To find a PRO solution, we need to find a set of beamlets that generate the same dose as the RO solution under a worst-case breathing pattern, but potentially perform better under some non-worst-case scenarios. There are two possible approaches:

1. **Pareto-robustness on the target dose**: A PRO solution would aim to further optimize the dose to the target under non-worst-case scenarios. Because the target dose is in the constraints, a better target dose is equivalent to larger optimal slacks to the constraints on the target dose.

2. **Pareto-robustness on the heart dose**: A PRO solution would aim to minimize the heart dose under non-worst-case scenarios.

In the models presented in Chapter 2, we only considered robustness in the constraints, which focus on the target dose. In this chapter, we also consider robustness on the heart
dose (in the objective function) in order to explore the benefit of the second approach mentioned above. Clinically, if the constraints on the target dose are met for all breathing patterns, the treatment is considered acceptable, regardless of the slacks. On the other hand, reducing the heart dose by even a small amount may be vital. Therefore, the second approach, where Pareto-robustness on the heart dose is explored, is more clinically relevant and will be the focus of this chapter.

For simplicity, let us rewrite model (2.9), which is the robust formulation for breast cancer radiation therapy presented in Chapter 2 as follows:

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{|V_H|} \hat{P}' \Delta w, \\
\text{subject to} & \quad w \in X,
\end{align*}
\]

(5.7a)

(5.7b)

where the objective only includes minimizing the heart dose \((c_H = 1)\), \(\Delta\) is the dose influence matrix for the heart, \(\hat{p}\) is a nominal breathing PMF, \(w\) is the vector of beamlet intensities, and \(X\) is the set of all feasible solutions. Note that the robustness was originally in the constraints on the target only.

In order to explore Pareto-robustness on the heart dose, we modify the formulation to incorporate robustness in the objective function as follows:

\[
\begin{align*}
\text{minimize} & \quad \max_{p \in \mathcal{P}} \frac{1}{|V_H|} \hat{p}' \Delta w, \\
\text{subject to} & \quad w \in X.
\end{align*}
\]

(RO)

model (RO) is equivalent to the general model (5.6) when we set

\[
\tilde{u} = \frac{1}{|V_H|} \hat{p}' \Delta,
\]

(5.9)
and the corresponding uncertainty set $\mathcal{U} = \left\{ \tilde{\mathbf{u}} \mid \tilde{\mathbf{u}} = \frac{1}{|V_H|} \tilde{\mathbf{p}}' \Delta, \tilde{\mathbf{p}} \in \mathcal{P} \right\}$. Thus any $\mathbf{p} \in \text{ri}(\mathcal{P})$ would result in a $\mathbf{u} \in \text{ri}(\mathcal{U})$ due to the linearity of the transformation (Boyd and Vandenberghe, 2004). The formulations in Section 5.2 can thus be directly applied to this problem.

Let $\mathbf{X}^{\text{RO}}$ denote the set of all optimal RO solutions and $z^*$ the optimal objective value of model (RO). To formulate the set $\mathbf{X}^{\text{RO}}$ in the breast cancer IMRT problem, we can say $\mathbf{w} \in \mathbf{X}^{\text{RO}}$ if

$$
\begin{align*}
\mathbf{w} \in \mathbf{X}, \quad \text{and} \quad & \frac{1}{|V_H|} \tilde{\mathbf{p}}' \Delta \mathbf{w} \leq z^*, \quad \forall \tilde{\mathbf{p}} \in \mathcal{P} \\
\end{align*}

(5.10)

which ensures that $\mathbf{w}$ is feasible for the RO problem and has an objective value equal to that of the optimal solution of the RO problem.

In the rest of this section, we formulate three models to answer the following questions for the breast cancer RT problem:

1. Are the sets of RO and PRO solutions equivalent? (Section 5.3.1)
2. Is a given optimal RO solution a PRO solution? (Section 5.3.2)
3. How can we find a PRO solution that has the lowest heart dose for a specific non-worst-case breathing pattern? (Section 5.3.3)

### 5.3.1 Testing the equivalence of the sets of RO and PRO

For the breast cancer IMRT problem, to test if the set of RO and PRO solutions are equivalent, we can use model (PRO-1) which is an expanded version of model (5.4) using notation from Chapter 2:

$$
\text{minimize} \quad \frac{1}{|V_H|} \sum_{v \in V_H} \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \pi(i) y_b \\
(\text{PRO-1})
$$
subject to

\[
\frac{1}{|V_H|} \sum_{v \in V_H} \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \tilde{p}(i) y_b \geq 0, \quad \forall \tilde{p} \in \mathcal{P}, \quad (5.11a)
\]

\[
\zeta' + \frac{1}{(1 - \beta)|V_T|} \sum_{v \in V_T} \bar{d}_v \leq U, \quad (5.11b)
\]

\[
\bar{d}_v \geq \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \tilde{p}(i)(w_b - y_b) - \zeta', \quad \forall v \in V_T, \tilde{p} \in \mathcal{P}, \quad (5.11c)
\]

\[
\zeta' - \frac{1}{(1 - \beta)|V_T|} \sum_{v \in V_T} \bar{d}_v \geq L, \quad (5.11d)
\]

\[
d'_v \geq \zeta - \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \tilde{p}(i)(w_b - y_b), \quad \forall v \in V_T, \tilde{p} \in \mathcal{P}, \quad (5.11e)
\]

\[
\bar{\zeta} + \frac{1}{(1 - \beta)|V_T|} \sum_{v \in V_T} \bar{d}_v \leq U, \quad (5.11f)
\]

\[
\bar{d}_v \geq \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \tilde{p}(i) w_b - \bar{\zeta}, \quad \forall v \in V_T, \tilde{p} \in \mathcal{P}, \quad (5.11g)
\]

\[
\zeta - \frac{1}{(1 - \beta)|V_T|} \sum_{v \in V_T} d_v \geq L, \quad (5.11h)
\]

\[
d_v \geq \zeta - \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \tilde{p}(i) w_b, \quad \forall v \in V_T, \tilde{p} \in \mathcal{P}, \quad (5.11i)
\]

\[
\frac{1}{|V_H|} \sum_{v \in V_H} \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \tilde{p}(i) w_b \leq z^*, \quad \forall \tilde{p} \in \mathcal{P}, \quad (5.11j)
\]

\[
\zeta', \bar{\zeta}, \bar{\zeta}, \zeta \geq 0, \quad (5.11k)
\]

\[
\bar{d}_v, d'_v, \bar{d}_v, d_v \geq 0, \quad \forall v \in V_T, \quad (5.11l)
\]

\[
y_b, w_b \geq 0, \quad \forall b \in B, \quad (5.11m)
\]

where $\pi \in \text{ri}(\mathcal{P})$, and $V_T$ and $V_H$ are the sets of all voxels in the target and the heart, respectively. Parameter $z^*$ is the optimal objective value of model (RO). The rest of the parameters and variables are similar to those introduced in Chapter 2. Constraint (5.11a) ensures that $w^* - y^*$ Pareto dominates $w^*$. Constraints (5.11b) to (5.11e) ensure that $w^* - y^*$ is feasible for the RO problem. Constraints (5.11f) to (5.11j) ensure that $w^*$ is a robust optimal solution. Regardless of the choice of $\pi$, this model will determine whether the sets of RO and PRO solutions are identical.
5.3.2 Testing if a given RO solution is PRO

Given an optimal RO solution \( w^* \), to test if \( w^* \) is PRO, we can use model (PRO-2), which is an expanded version of model (5.5), as follows:

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{|V_H|} \sum_{v \in V_H} \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \pi(i) y_b \\
\text{subject to} & \quad \frac{1}{|V_H|} \sum_{v \in V_H} \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \tilde{p}(i) y_b \geq 0, \quad \forall \tilde{p} \in \mathcal{P}, \\
& \quad \zeta' + \frac{1}{(1 - \beta)|V_T|} \sum_{v \in V_T} \bar{d}'_v \leq U, \\
& \quad \bar{d}'_v \geq \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \tilde{p}(i) (w^*_b - y_b) - \zeta', \quad \forall v \in V_T, \tilde{p} \in \mathcal{P}, \\
& \quad \zeta' - \frac{1}{(1 - \beta)|V_T|} \sum_{v \in V_T} d'_v \geq L, \\
& \quad d'_v \geq \zeta' - \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \tilde{p}(i) (w^*_b - y_b), \quad \forall v \in V_T, \tilde{p} \in \mathcal{P}, \\
& \quad \zeta', \zeta' \geq 0, \\
& \quad \bar{d}'_v, d'_v \geq 0, \quad \forall v \in V_T \\
& \quad y_b \geq 0, \quad \forall b \in B.
\end{align*}
\]

In model (PRO-2), we again use an arbitrary parameter \( \pi \in \text{ri}(\mathcal{P}) \). Constraint (5.12a) ensures that \( w^* - y^* \) Pareto dominates \( w^* \), and constraints (5.12b) to (5.12e) ensure that \( w^* - y^* \) is feasible in the original robust problem. Similar to model (PRO-1), the choice of \( \pi \) does not play a role in determining whether the given \( w^* \) is PRO. If \( w^* \) is not PRO, the choice of \( \pi \) could affect the optimal \( y^* \) values; otherwise, \( y^* = 0 \) for all \( \pi \).

5.3.3 Finding a PRO solution

Lastly, we formulate a model that can help us find a PRO solution that optimizes the dose for a specific non-worst-case scenario. Note that the resulting solution is PRO, so
it cannot be Pareto dominated by any other solution, but it does not necessarily Pareto dominate the given RO solution.

In the IMRT problem, this model explores the set of RO solutions to find a solution that has the lowest heart dose for a specific non-worst-case breathing pattern \( \pi \in \text{ri}(P) \). This problem is equivalent to model (5.6) and can be formulated as follows:

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{|V_H|} \sum_{v \in V_H} \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \pi(i) w_b \\
& \quad \bar{\zeta} + \frac{1}{1 - \beta} \sum_{v \in V_T} \bar{d}_v \leq U, \\
& \quad \bar{d}_v \geq \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \bar{\mu}(i) w_b - \bar{\zeta}, \quad \forall v \in V_T, \bar{p} \in P, \\
& \quad \zeta - \frac{1}{1 - \beta} \sum_{v \in V_T} d_v \geq L, \\
& \quad d_v \geq \zeta - \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \bar{\mu}(i) w_b, \quad \forall v \in V_T, \bar{p} \in P, \\
& \quad \frac{1}{|V_H|} \sum_{v \in V_H} \sum_{i \in I} \sum_{b \in B} \Delta_{v,i,b} \bar{\mu}(i) w_b \leq z^*, \quad \forall \bar{p} \in P \\
& \quad \bar{\zeta}, \zeta \geq 0, \\
& \quad \bar{d}_v, d_v \geq 0, \quad \forall v \in V_T, \\
& \quad w_b \geq 0, \quad \forall b \in B.
\end{align*}
\]

Constraints (5.13a) to (5.13e) ensure that \( w \) is optimal in the original RO problem. The optimal solution to model (PRO-3) depends on \( \pi \). In Section 5.4, we test this model with different \( \pi \) vectors that represent inhale or exhale-weighted breathing patterns.

### 5.4 Results

In this section, we first solve the aforementioned models (PRO-1) and (PRO-2) for the first patient (from the set of patients in Chapter 4), to compare the sets of RO and PRO
solutions and to test if the RO solution was PRO. The main focus, however, is on finding
PRO solutions for all patients. To do so, we solve model (PRO-3) using two different
values of $\pi$ for all patients and find the difference in the mean heart dose between the
PRO and RO solutions for a set of non-worst-case breathing scenarios.

5.4.1 Testing the equivalence of the sets of RO and PRO

First, we used model (PRO-1) with $\hat{\pi} = [0.1693, 0.1552, 0.0768, 0.0918, 0.5069] \in \text{ri}(\mathcal{P})$
to test the equivalence of the sets of RO and PRO solutions (recall that any arbitrary
$\pi \in \text{ri}(\mathcal{P})$ can be used in this model). Solving model (PRO-1) using $\hat{\pi}$ yielded an optimal
$y^* \neq 0$ where only one element of the vector $y^*$ was non-zero. This means that for this
patient, the set of RO and PRO solutions are not identical. Figure 5.1 compares the RO
solution $w^*$ and the dominating PRO solution $w^* - y^*$ found by solving model (PRO-1)
using this fixed $\hat{\pi}$ for a set of 200 simulated non-worst-case breathing scenarios. Note
that this $w^*$ is an RO solution found in model (PRO-1) and is not necessarily equal to the
original RO solution from solving model (RO). Each pair of dots represents the resulting
RO and PRO dose for one simulated breathing pattern. It can be seen that the RO and
PRO doses are almost identical for all simulated non-worst-case breathing patterns.

5.4.2 Testing if a given RO solution is PRO

The next question is to test whether the original RO solution is a PRO solution. We
used the $w^*$ from solving model (RO) and then solved the PRO model (PRO-2) to check
if this solution is PRO. We used the same $\hat{\pi}$ as in Section 5.4.1. The optimal $y^*$ was
the zero vector, so the solution in the original RO problem for this test dataset was a
PRO solution. Note that for this patient, since the sets of RO and PRO solutions are
not equivalent, there exist other RO solutions that are not PRO.
5.4.3 Finding a PRO solution

Although the RO solution for the first patient was PRO, we cannot guarantee that solving model \((RO)\) will always produce PRO solutions. Even if the set of RO and PRO solutions are equivalent, when there are multiple PRO solutions, we would like to be able to choose a PRO solution that has a lower heart dose for most non-worst-case breathing patterns. In model \((PRO-3)\), the optimal PRO solution depends on \(\pi\). In this section, we find two PRO solutions per patient using an inhale-weighted and an exhale-weighted breathing pattern. The goal is to determine which type of breathing pattern (inhale or exhale-weighted) when used in model \((PRO-3)\) results in a treatment plan with a lower heart dose for a larger number of non-worst-case scenarios.

We compare the difference between the mean heart dose of the RO and PRO solutions for a set of non-worst-case scenarios, for both the inhale and exhale-weighted \(\pi\) vectors. To quantify this difference, we use a set of simulated breathing patterns inside the uncertainty set and calculate the mean heart dose for the RO and PRO solutions.
when each of these non-worst-case scenarios are realized. The following breathing PMFs inside the uncertainty set were used as $\pi$ in model (PRO-3) for comparison:

- Inhale-weighted: $\pi_1 = [0.2116, 0.1160, 0.1715, 0.0969, 0.4040] \in P$,

- Exhale-weighted: $\pi_2 = [0.0541, 0.1058, 0.1002, 0.1426, 0.5973] \in P$.

Table 5.1 compares the mean heart dose of the RO and PRO plans for the set of patients mentioned in Chapter 4. Patient 8 was excluded since the heart dose for this patient was nearly zero, as observed in Chapter 4 (see Tables 4.1 and 4.2). The values in Table 5.1 quantify the average PRO benefit (in cGy) over all simulated scenarios compared to the RO solution (negative is improvement). Note that a small change in the mean heart metric can be clinically relevant, since the majority of the heart is outside the treatment and receives no dose. In our set of patients, between 4% to 15% of the volume of the heart is inside the radiation field. For instance, if 5% of the volume of the heart is inside the radiation field, a 0.1 cGy decrease in the mean heart dose is equivalent to 2 cGy decrease for the parts inside the treatment field, on average. Thus a small gain on the mean heart dose can potentially be equivalent to a larger gain on the parts of the heart inside the treatment field.

In Table 5.1, the columns labelled “Change in mean heart dose” show the average difference between the RO and PRO solutions for a set of 200 non-worst-case breathing patterns inside the uncertainty set using the two different values of $\pi \in \text{ri}(P)$ in the objective function of model (PRO-3). The last two columns show the number of scenarios for which the mean heart dose of the PRO plan was lower than that of the RO plan. Both the inhale-weighted and exhale-weighted cases improve the dose under many scenarios for most patients. For most patients, using an exhale-weighted breathing pattern in model (PRO-3) results in a larger number of improved scenarios. It also results in a larger change in the mean heart dose metric for most patients.

Figure 5.2 shows a histogram of the change in mean heart dose over the set of 200
Table 5.1: Comparing the average change in the mean heart dose of the RO and PRO plans over a set of 200 simulated non-worst-case scenarios.

<table>
<thead>
<tr>
<th>Patient</th>
<th>Change in mean heart dose (cGy)</th>
<th># of improved scenarios (out of 200)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inhale-weighted $(\pi_1)$</td>
<td>Exhale-weighted $(\pi_2)$</td>
</tr>
<tr>
<td>1</td>
<td>-0.01</td>
<td>-0.01</td>
</tr>
<tr>
<td>2</td>
<td>0.04</td>
<td>-0.12</td>
</tr>
<tr>
<td>3</td>
<td>-0.02</td>
<td>-0.04</td>
</tr>
<tr>
<td>4</td>
<td>-0.13</td>
<td>-0.03</td>
</tr>
<tr>
<td>5</td>
<td>-0.09</td>
<td>-0.10</td>
</tr>
<tr>
<td>6</td>
<td>-0.01</td>
<td>0.08</td>
</tr>
<tr>
<td>7</td>
<td>-0.05</td>
<td>-0.15</td>
</tr>
</tbody>
</table>
simulated scenarios when the inhale-weighted or exhale-weighted $\pi$ is used for each patient. The horizontal axis shows the change in the mean heart dose and the vertical axis shows the number of scenarios (out of 200). The horizontal axis is on the same scale and has the same range for all patients. Figure 5.2 confirms that for most patients, using an exhale-weighted $\pi$ results in larger reduction in heart dose for a larger number of scenarios. For most patients, the variance in the change in the mean heart dose among the simulated scenarios is relatively smaller for the exhale-weighted breathing pattern compared to the inhale-weighted case, which further confirms that an exhale-weighted $\pi$ is more likely to reduce the overall mean heart dose over the uncertainty set, although the magnitude of reduction may not be as large as the reduction by the inhale-weighted $\pi$ for some scenarios on the left tail of the histogram.

Figures 5.3 to 5.9 in Appendix 5.7 further compare the dose metrics on the heart and the target when the inhale or exhale-weighted $\pi$ is used in model (PRO-3). These figures illustrate that the choice of $\pi$ will affect the dose metric on target coverage as well. For all scenarios, the dose to the target would change but remain inside the clinical acceptability limits since the robust constraints on the target must still be met for all scenarios in the uncertainty set (see Appendix 5.7 for further details).

5.5 Discussion

Comparing the RO and PRO solutions for inhale and exhale-weighted breathing patterns in Table 5.1 and Figure 5.2 demonstrated that for most of the patients, using an exhale-weighted breathing pattern ($\pi_2$) resulted in a larger number of improved scenarios compared to the inhale-weighted breathing pattern ($\pi_1$). The only exception was patient 6, for whom forcing the model to minimize the heart dose for the exhale-weighted breathing scenario increased the dose for most other scenarios. When an inhale-weighted breathing pattern was used, the number of improved scenarios and the aggregate reduc-
(a) Patient 1

(b) Patient 2
Chapter 5. Pareto Robust Optimization in Breast Cancer RT

(c) Patient 3

(d) Patient 4
Figure 5.2: Histogram of the difference in mean heart dose from RO to PRO over 200 non-worst-case scenarios (negative is improvement).
tion in mean heart dose were smaller for most patients. Recall from Chapters 2 and 4 that an exhale-weighted breathing pattern results in a potentially larger dose to the heart, since the heart is closest to the radiation field at exhale. The worst-case heart dose most likely happens at exhale, and if in addition to the worst case, we optimize for a non-worst-case exhale-weighted breathing PMF in the PRO approach, it may result in further reduction in the heart dose over all simulated scenarios.

The potential benefit of the PRO approach is larger when the patient is at a higher risk of having excessive radiation delivered to the heart due to breathing motion. For example, in Chapter 4, we saw that patients 4, 5 and 7 had the largest positive displacement towards the inside of the radiation field for both organs (See Figure 4.1(a)). For these patients, in Table 5.1 we observe that PRO had a larger improvement in the mean heart dose for non-worst-case scenarios. It is important to note again that since the majority of the volume of the heart is receiving no dose, a small reduction in the mean dose can potentially result in a larger reduction in the dose to a subvolume of the heart that is receiving radiation.

For all patients, the RO and PRO solutions were very close dosimetrically, and the reduction in the mean heart dose was not clinically significant in most cases. However, there is no guarantee that this is always the case. In order to ensure that we provide the best treatment plan that cannot be dominated, it is best to find a PRO solution directly instead of an RO solution which might not be PRO. This will guarantee that the solution cannot be Pareto dominated by any other solution.

The computational expense for finding a PRO solution is to solve an additional LP after solving the RO problem. The second LP (i.e., model (PRO-3)), uses the result of the first LP (i.e., model (RO)) to formulate the set of RO solutions ($X_{RO}$). The downside is that we need solve the RO model to optimality with high precision, and any method that results in even a small precision loss in the RO solution can result in a change in the feasible set of the PRO model. Solving the RO model for these patients using
constraint generation took between 15 minutes to 4 hours, whereas solving for the exact optimal solution using the robust counterpart took between 1.5 to 41 hours. Solving the PRO model for each patient took an additional 1.5 to 78 hours. One alternative is to formulate both problems in one step by taking the dual of the robust counterpart of the RO problem and enforcing strong duality of the RO problem in the constraints the PRO problem. This method would require solving a much larger LP instead of solving two smaller LPs. The constraint generation methods presented in Chapter 3 can potentially be adjusted to solve this even larger-scale LP problem. This is a topic for future research. Note that our focus in this chapter is not on the computational time, but on modelling the concept of PRO in breast cancer RT and the potential benefit of this approach.

A possible future direction is to incorporate an additional objective on the dose to the subvolume of the heart that is inside the field. For example, the D25cc metric that captures the dose to 25cc of the heart that is receiving the highest dose can be added to the objective function. This metric can be approximately formulated as a CVaR metric, as discussed in Chapter 2. The resulting model would be multi-objective with robust-CVaR and robust mean heart dose objective and robust-CVaR constraints. The benefit of using such a model is that the PRO model would directly aim to reduce the dose to the subvolume of the heart receiving the highest radiation. The challenge would be to find the appropriate weights for the multiple objectives.

Finally, we note that the size of the uncertainty set could potentially affect the difference between the RO and PRO solutions. If the size of the uncertainty set is increased, it will include a larger number of non-worst-case scenarios that are far from the worst case. Therefore, the PRO model can potentially make a greater improvement in the outcome of non-worst-case scenarios. It would be interesting to study how the size of the uncertainty set would affect the difference between the set of RO and PRO solutions.
5.6 Conclusions

Pareto robust optimization can help us find a solution that cannot be Pareto dominated by any other solution under realizations of non-worst-case scenarios in the uncertainty set. For breast cancer radiation therapy treatment planning, we can find a PRO solution by solving an additional LP after solving the RO problem. Our results showed that optimizing for a non-worst-case exhale-weighted breathing pattern in the PRO model results in a reduction in the mean heart dose over many breathing scenarios. Although the reductions were modest for these patients, using a PRO model is advisable since it guarantees that the treatment plan cannot be Pareto dominated. Future work consists of reducing the computational time of solving the PRO problem and exploring multi-objective PRO models for breast cancer RT.

5.7 Appendix

To compare the RO and PRO solutions per scenario, in addition to the mean heart dose, we calculated the difference in the D99% CTV dose for the RO and PRO solutions per scenario. Figures 5.3 to 5.9 compare the two PRO solutions with the original RO solution for each patient. Each pair of dots represents one breathing pattern. The RO dots use the optimal $w^*$ from model (RO) and show the dose when one of the simulated breathing patterns is realized. Similarly, the PRO dots use the optimal solution $w^*$ from model (PRO-3) with a simulated breathing pattern. The lines pair the RO and PRO solution for the same simulated PMF. The star and the larger dot connected to it show the result for the RO and PRO methods, respectively, when the $\pi$ used in the objective function is realized. The dots that have a lower heart dose (towards the left of the diagram) are preferred. The target dose will be affected, but it always remains within the clinical acceptability limits. Note that for each patient, the two figures are on the same scale, but the figures for different patients are not.
Figure 5.3: Comparing RO and PRO results for patient 1 using 200 simulated breathing patterns
Figure 5.4: Comparing RO and PRO results for patient 2 using 200 simulated breathing patterns.
Figure 5.5: Comparing RO and PRO results for patient 3 using 200 simulated breathing patterns
Figure 5.6: Comparing RO and PRO results for patient 4 using 200 simulated breathing patterns
Figure 5.7: Comparing RO and PRO results for patient 5 using 200 simulated breathing patterns
Figure 5.8: Comparing RO and PRO results for patient 6 using 200 simulated breathing patterns
Figure 5.9: Comparing RO and PRO results for patient 7 using 200 simulated breathing patterns
Chapter 6

Conclusions

Radiation therapy (RT) aims to deliver sufficient radiation dose to the cancerous cells while sparing the nearby healthy organs. In breast cancer radiation therapy, the heart moves in and out of the radiation beams due to breathing motion. Therefore, breast cancer patients are at a high risk of developing cardiac complications after a successful treatment. The breathing motion is often irregular and unpredictable which motivates the use of optimization techniques to minimize the dose to the heart.

This thesis presents a robust optimization (RO) framework to address these challenges in breast cancer radiation therapy. We incorporate the concept of conditional value-at-risk (CVaR) within robust optimization and formulate robust-CVaR models that ensure the clinical dose-volume criteria on the organs are met under breathing motion uncertainty. We then explore several constraint generation strategies to solve these large-scale robust problems. Our computational experiments show that the proposed constraint generation strategies can reduce the computation time by an order of magnitude compared to directly solving the robust counterpart.

Using clinical patient data, we compare our robust models with current standard clinical treatment planning methods. Compared to the conventional treatment method using free breathing, our models reduce the dose to the heart by 29% on average, and
up to 70% for some patients, without compromising target coverage. In addition to RO
models that optimize the worst-case breathing scenario, we also formulate Pareto robust
optimization (PRO) models which can potentially further reduce the heart dose under
non-worst-case breathing scenarios. For all patients, we find PRO treatment plans that
cannot be dominated by any other plan. Our results show that the RO solutions are very
close to PRO solutions for all patients.

Although the methodologies in this thesis are motivated by the breast cancer RT
problem, they have general applications beyond RT. Our robust-CVaR framework can
be adapted to other applications having an uncertain underlying probability distribu-
tion which depends on the state of a system. For instance, in financial engineering, the
return/loss distribution for an investment could depend on the state of the economy
(e.g., recession, growth). The possible states and the corresponding distributions may be
known, but the fraction of time that the economy spends in each state can be uncertain.
Our robust-CVaR models can be used to optimize the tail loss or to ensure that the tail
loss is less than a certain threshold under uncertainty. The robust-CVaR models pre-
sented in this thesis generalize the existing CVaR methods and are less conservative than
the worst-case robust methods in the literature. The constraint generation strategies can
also be used for any large-scale robust problem with a large number of robust constraints
that must hold for a set of potentially uncountable number of uncertain scenarios.

One of the possible future research directions is to adapt the uncertainty set over the
course of the treatment. In breast cancer, the RT treatment is fractionated and patients
receive radiation therapy over several treatment sessions. Often, the same treatment plan
is delivered in all fractions, regardless of the changes that may occur over the course of
the treatment. In our robust model, we have assumed that the size of the uncertainty
set is fixed throughout the course of the treatment. However, based on the observed
breathing pattern during each treatment, the uncertainty set can be updated and a re-
optimized treatment can be delivered at each fraction. Thus, over time, we may have
a tighter uncertainty set that would potentially converge to the patient’s true breathing pattern during treatment. Previously, an adaptive approach has been developed for lung cancer and has shown to improve the treatment quality over the non-adaptive approach (Chan and Mišić, 2013). We believe that adaptation of the uncertainty set in our RO approach for breast cancer RT can result in treatments that have a lower heart dose and are dosimetrically closer to an ABC treatment (which is the gold standard) compared to the non-adaptive RO approach.

In this thesis, we focused on fluence map optimization (FMO), which finds the optimal intensities of the beamlets. The FMO needs to be followed by a leaf sequencing algorithm that divides the optimized intensity map into a set of deliverable apertures (segments). These apertures must meet specific criteria in order to be deliverable. For instance, in breast cancer, there are limitations on the shape of the segments and the total number of segments per beam. There is usually a large segment that has the highest weight and covers the entire breast tissue, along with a number of smaller segments to compensate for dose requirements. Planners tend to avoid more complicated treatment plans so as to ensure homogeneity on the breast and to maintain short treatment delivery times.

Instead of the robust FMO approach, another future research direction is to develop a robust direct aperture optimization (DAO) technique to control the shape of the segments and limit the number of segments under breathing motion uncertainty. To limit the total number of segments, we can decompose the beam into different levels of intensity. Let $k$ be the maximum number of segments. Define $\mathbf{w}^1, ..., \mathbf{w}^k$ as the the vectors of different levels of beams. The intensity of beamlet $b$ in segment $j$ can be defined as $w^j_b = (0 \text{ or } \hat{w}^j)$ for all $b \in \mathcal{B}$, where $\hat{w}^j$ is the uniform intensity for segment $j$. The dose to voxel $v$ at phase $i$ is $\sum_{b \in \mathcal{B}} (\sum_{j=1}^k w^j_b) \Delta_{v,i,b}$. Let $\mathcal{B}^1 \subset \mathcal{B}$ be the given set of the beamlets in the large
segment that is a subset of beamlets inside each beam. Define the vector $\mathbf{w}^1$ as

$$w^1_b = \begin{cases} \hat{w} & \text{if } b \in \mathcal{B}^1 \\ 0 & \text{if } b \notin \mathcal{B}^1, \end{cases}$$

where $\hat{w}$ is the uniform intensity of the large segment $\mathcal{B}^1$. The resulting model would be a mixed integer linear program (MILP).

To solve the robust DAO problem, the constraint generation models presented in this thesis can potentially be improved by using geometrical information of the voxels. Since the number of segments are small and the segments have uniform intensities, if we add a constraint for a voxel, the intensity of the entire beam segment changes to meet the constraint. Therefore, the dose requirements for the neighbouring voxels would most likely be met as well. Using the information on the location of the voxels in the constraint generation method allows us to avoid adding multiple constraints for neighbouring voxels in the same iteration, which could potentially reduce the computation time. By doing so, we can provide clinically applicable treatment plans in an efficient and timely manner, which allows for a feasible clinical work flow.
Bibliography


