Reducing Conservatism in Pareto Robust Optimization

by

Fahimeh Rahimi

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Management Sciences

Waterloo, Ontario, Canada, 2022

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Ehsan Salari
Associate Professor
Department of Industrial, Systems, and Manufacturing Engineering
Wichita State University

Supervisor(s): Hossein Abouee Mehrizi
Associate Professor
Department of Management Sciences
University of Waterloo

Houra Mahmoudzadeh
Assistant Professor
Department of Management Sciences
University of Waterloo

Internal Member: Fatih Safa Erenay
Associate Professor
Department of Management Sciences
University of Waterloo

Internal Member: Jim Bookbinder
Professor
Department of Management Sciences
University of Waterloo

Internal-External Member: Chaitanya Swamy
Professor
Department of Combinatorics and Optimization
University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Robust optimization (RO) is a sub-field of optimization theory with set-based uncertainty. A criticism of this field is that it determines optimal decisions for only the worst-case realizations of uncertainty. Several methods have been introduced to reduce this conservatism. However, non of these methods can guarantee the non-existence of another solution that improves the optimal solution for all non-worse-cases.

Pareto robust optimization ensures that non-worse-case scenarios are accounted for and that the solution cannot be dominated for all scenarios. The problem with Pareto robust optimization (PRO) is that a Pareto robust optimal solution may be improved by another solution for a given subset of uncertainty. Also, Pareto robust optimal solutions are still conservative on the optimality for the worst-case scenario.

In this thesis, first, we apply the concept of PRO to the Intensity Modulated Radiation Therapy (IMRT) problem. We will present a Pareto robust optimization model for four types of IMRT problems. Using several hypothetical breast cancer data sets, we show that PRO solutions decrease the side effects of overdosing while delivering the same dose that RO solutions deliver to the organs at risk.

Next, we present methods to reduce the conservatism of PRO solutions. We present a method for generating alternative RO solutions for any linear robust optimization problem. We also demonstrate a method for determining if an RO solution is PRO. Then we determine the set of all PRO solutions using this method. We denote this set as the "Pareto robust frontier" for any linear robust optimization problem. Afterward, we present a set of uncertainty realizations for which a given PRO solution is optimal. Using this approach, we compare all PRO solutions to determine the one that is optimal for the maximum number of realizations in a given set. We denote this solution as a "superior" PRO solution for that set.

At last, we introduce a method to generate a PRO solution while slightly violating the optimality of the optimal solution for the worst-case scenario. We define these solutions as "light PRO" solutions. We illustrate the application of our approach to the IMRT problem for breast cancer. The numerical results present a significant impact of our method in reducing the side effects of radiation therapy.

Acknowledgements

My sincerest thanks to my supervisor and friend, Prof. Houra Mahmoudzadeh, for her support, understanding and patience. Without her help and wise guidance, this project would not have been the same. I also would like to thank my supervisor, Prof. Hossein Abouee Mehrizi, for his guidance and support throughout my time as his student.

I would like to thank the members of my dissertation committee, Prof. Fatih Safa Erenay, Prof. Jim Bookbinder, Prof. Chaitanya Swamy, and Prof. Ehsan Salari, for their time spent reading the thesis, insightful suggestions, and comments.

I thank all my friends in Waterloo, especially Matin, Mahsa, Nastaran and Mina, for the support and happy moments that they provided for me.

I must express my very profound gratitude to my parents. They were my first teachers in my life, and I find myself in outstanding debt to them for their love and support. I sincerely appreciate their belief in me.

Last but not least, my most heartfelt thanks to my love and supportive husband, Amin, for his continued and unfailing love, support and understanding during my pursuit of the Ph.D. degree. He was always around at times, and when I thought it was impossible to continue, he helped me keep things in perspective.

Dedication

To my dear parents, Zahra and Akbar, and the love of my life, Amin, who mean everything to me.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CT	Computed Tomography
CVar	Conditional Value at Risk
DVH	Dose-Volume Histogram
FO	Forward Optimization
GY	Gray
IMRT	Intensity-Modulated Radiation Therapy
IO	Inverse Optimization
OAR	Organs at Risk
PMF	Probability Mass Function
PRO	Pareto Robust Optimization
RO	Robust Optimization
Var	Value at Risk

Chapter 1

Introduction

Robust optimization (RO) is a sub-field of optimization theory with set-based uncertainty used to determine optimal decisions for the worst-case realizations of uncertainty. This approach was first introduced by Soyster [1973]. A robust optimal solution, denoted as an RO solution, to a robust optimization problem is sometimes criticized for being too conservative as it is only optimal for the worst-case realization of uncertainty. One of the methods introduced in the literature to reduce this conservatism is applying Pareto efficiency Iancu and Trichakis [2014] in robust optimization. Iancu and Trichakis [2014] defined Pareto robust optimal solutions, denoted as a PRO solution, that provide the same protection as RO solutions under worst-case scenario but are less conservative for non-worst-case scenarios. In this thesis, we first develop the concept of Pareto robust optimization (PRO) in an intensity-modulated radiation therapy (IMRT) problem to generate PRO solutions. We then define the concept of Superior PRO solution for a determined uncertainty set. We also define Pareto robust frontier as the set of all PRO solutions for an RO problem and present an algorithm for generating the Pareto robust frontier for any linear robust optimization problem along with an algorithm to present a set of uncertainties for which a PRO solution is optimal, using inverse optimization. Finally, we propose solutions for linear robust optimization problems that are less conservative than PRO solutions, denoted as light Pareto robust optimal solutions.

In this chapter, we first discuss the literature on robust optimization and methods for handling conservatism in this field. Then, we provide the background information in Pareto robust optimization. Next, we discuss the background in IMRT and relevant robust optimization literature for this problem for handling uncertainty. We also briefly review the literature on inverse optimization. Finally, we present the structure of the thesis and highlight the contributions of this work.

1.1 Robust Optimization

Robust optimization was introduced by Soyster [1973] and has been extended over the years. RO has been used in different applications such as health care [Meng et al., 2015], inventory management [Bienstock and ÖZbay, 2008], field of machine learning [Xu, 2009] and statistical problems [Fertis, 2009]. Ben-Tal et al. [2009] and Gabrel et al. [2014] provided a comprehensive review of robust optimization and discussed its applications. A fundamental criticism of robust optimization is that the robust optimal solutions are too conservative and only optimal for the worst-case realization of uncertainty. Recently, several methods have been presented to mitigate this conservatism. Below, we will present several approaches introduced to reduce the conservatism in robust optimization. Then, we will explain and compare our approach with these methods.

Light Robustness: A problem in RO is that the robust optimal solution must be feasible for all constraints and uncertainty realizations. The light robustness approach introduced by Fischetti and Monaci [2009] proposed a method to violate RO solutions from optimality for the worst-case scenarios. Fischetti and Monaci [2009] defined a hard upper bound for the objective value. Then, they reduced the conservatism by allowing violations for the constraints. They presented these violations using slack variables and proposed a model to minimize the sum of the slacks. This well-known method for reducing conservatism, has been discussed in several papers, such as Cacchiani and Toth [2012] and Schöbel [2014].

Decreasing Price of Robustness: Bertsimas and Sim [2004] suggested using the budget of uncertainty as a parameter that controls the deviation between an uncertain parameter and its nominal value. Thus, if this constant is zero, then the parameter is not uncertain, and it is equal to its nominal, so the optimal solution to the problem is the same as the solution to the nominal problem. The uncertainty will increase by increasing the budget of uncertainty, so the solution's conservatism will increase.

Bertsimas and Sim [2004] called this trade-off between the budget of uncertainty and the optimal solution as a price of robustness. They proposed methods to decrease the price of robustness. Their methodology has been discussed in several papers, such as, Dehghani Filabadi and Mahmoudzadeh [2022] and Gorissen et al. [2015].

Globalized Robust Optimization: Globalized robust optimization defines a range of uncertainty, called "normal range", and considers that constraints are hard for the uncertainties in this range. This approach also defines violations for the constraints regarding uncertainty realizations outside the normal range. This violation is dependent on the distance between the uncertainty scenarios and the normal range. For more information we refer the readers to Ben-Tal et al. [2006] and Ben-Tal et al. [2017]. This method has been discussed in several papers, such as Roos and den Hertog [2020] and Zhao et al. [2019].

There are other methods for reducing conservatism in RO. For more information, we refer readers to the review by Goerigk and Schöbel [2016] on methods for reducing conservatism in robust optimization.

All of these methods still focus on the worst-case scenario to the best of our knowledge. Therefore, the optimal solutions, that all these methods present, can be improved by another solution for non-worst-cases. That is, if an RO problem has multiple solutions using any method for reducing conservatism, these methods do not compare the performance of the optimal solutions for the non-worst-case scenarios.

Using the concept of Pareto efficiency, Iancu and Trichakis [2014] introduced new solutions in robust optimization. They defined Pareto robust optimal solutions for linear robust optimization problems with multiple optimal solutions as a subset of RO solutions that cannot be dominated by any other RO solutions for all uncertainty realizations, including non-worst-case scenarios. This method focuses on the behaviour of RO solutions for all realizations of uncertainty at the same time and proposes an RO solution with less conservatism.

1.2 Pareto Robust Optimization

Pareto robust optimization, introduced by Iancu and Trichakis [2014], aims to improve the non-worst-case behaviour of RO solutions. A Pareto robust optimal solution is a solution that has the same worst-case performance as an RO solution but cannot be dominated in non-worst-case scenarios. Iancu and Trichakis [2014] presented a method to generate a PRO solution for any linear robust optimization problem, along with a procedure for determining whether all RO solutions are PRO. They applied this new methodology to inventory management, portfolio optimization, and project management.

This new concept of Pareto efficiency in robust optimization has been noticed and discussed in several papers. Dunning [2016] extended this work to two-stage adaptive robust optimization. Gorissen et al. [2014], following a similar approach, developed a re-optimization procedure to improve the average performance of the robust adjustable counterpart. Botte and Schöbel [2019] generalized the concept of Pareto robust solutions for multi-objective cases.

Although Pareto robust optimization proposes a solution with better outcome for non-worst-cases than the RO solution, there is still some criticism of this approach. 1) It is not guaranteed that there is no other RO solution that is optimal for more non-worst-cases than a PRO solution. 2) A PRO solution is still an optimal solution to the traditional form of the robust optimization and is still focusing on the worst-case scenario.

To tackle these problems, we propose two methods. For the first one, we suggest generating the Pareto robust frontier and proposing a method to compare PRO solutions to find the optimal one for more uncertainty realizations in a given subset of uncertainty.

For the second problem, we will present solutions by violating the Pareto robust solutions from the optimal solution for the worst-case scenarios. We call these solutions light Pareto robust optimal solutions.

To the best of our knowledge, there has been no research on generating Pareto robust frontiers. However, several methods exist to generate such Pareto frontiers in multi-objective optimization. This thesis presents a method for generating multiple RO solutions in linear robust optimization problems. This study also proposes a method for determining whether an RO solution is PRO and presents an algorithm to generate the Pareto robust frontier for any linear robust optimization problem using these two methods. We also define the superior PRO solution for a predetermined subsets of the uncertainty set. That is, a solution optimal for the largest number of uncertainty uncertainty in that set.

Moreover, this thesis introduces a method to generate PRO solutions while slightly violating the optimality from optimal solution for the worst-case realization and reducing the conservatism of robustness, light Pareto robust optimal solutions.

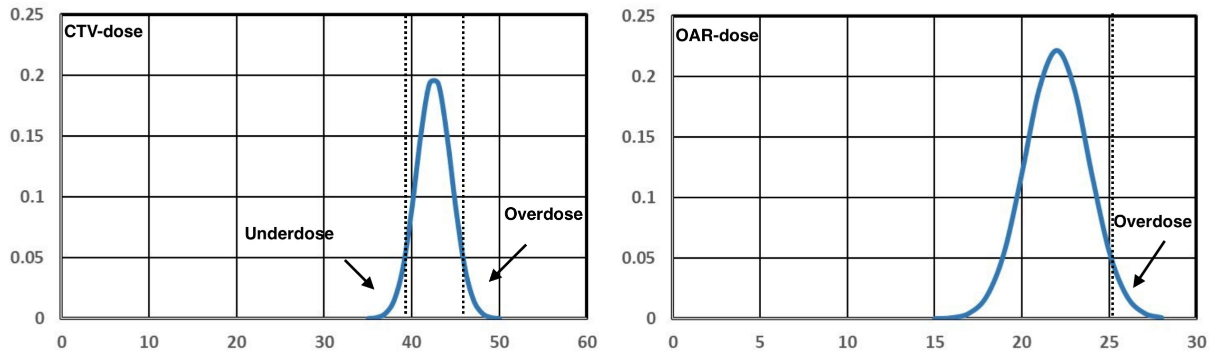
1.3 Intensity Modulated Radiation Therapy

IMRT is a well-known method in radiation therapy that was first presented by Brahme et al. [1982]. It uses a computer-controlled device called a linear accelerator to deliver a precise radiation dose to a malignant tumour or specific areas within the tumour. To do so, the linear accelerator continuously shoots a high-energy beam to the target tissue [Ahunbay and Li, 2007]. The beam, however, also damages any healthy tissue around the tumour. Therefore, radiation doses are delivered to the cells from different angles so that they all overlap on the tumour. Thus, as little dose as possible is delivered to adjacent healthy tissue. For more information about IMRT, we refer readers to reviews by Jiang [2008], Bortfeld et al. [2008], Lin et al. [2018] and Park et al. [2018].

IMRT treatment includes three steps: 1) imaging the tumour tissue; 2) delineating different organs, including the target and healthy tissues; and, 3) planning treatment. A computed Tomography (CT) scan, used for the imaging step, takes a 3D image of the organs exposed to radiation. In the second step, different organs, such as the healthy organs at risk (OAR) and the tumour in the target region, are delineated on the CT scan. Finally, the prescribed dose of radiation for the tumour and the maximum allowed dose for the healthy tissues are determined, and a treatment plan, consisting of the angles, shapes, and intensities of each beam, is prepared.

Treatment planning is an essential step in the radiation therapy procedure. To determine what the treatment will involve, an oncologist first determines a set of dose criteria for the tumour and/or the organs at risk. For example, at most 10% of the lung volume can receive a dose higher than 21.1 Gray (GY), a Gray being the standard unit of measurement for radiation doses, which is an example of a dose-volume criterion.

Suitable criteria are necessary for the evaluation and comparison of treatment plans. Quantitative evaluation of dose-volume histograms (DVH), a histogram that captures the dose to every volume of tissue, provides meaningful criteria for radiation treatment planning. Figures 1.1(a) and 1.1(b) show two examples of distributions of the dose delivered to the tissues. Figure 1.1(a) is related to the tumour, and Figure 1.1(b) corresponds to healthy tissues. It can be seen that for treating the tumour, it is essential to have both under-dose and overdose criteria. However, since we need to deliver as little radiation as possible to



(a) Dose distribution of a clinical target volume (b) Dose distribution of an organ at risk

Figure 1.1: Examples of differential dose volume histograms. The dashed line indicates the dose limits for the organs

healthy tissues, only overdose criteria are necessary for the organs at risk. The concept of dose-volume criteria is equivalent to the value-at-risk (VaR) metric of a loss distribution. The VaR metric measures the probability that a case will not exceed or break a threshold loss value. The VaR metric for a confidence level β is the lowest amount of loss such that the probability of having more than that loss is at most β Abad et al. [2014]. VaR metric measures the tail of the loss distribution, and has applications in different areas such as finance (e.g., Duffie and Pan [1997], Ghaoui et al. [2003]) and healthcare (e.g., Chan et al. [2014], Romeijn et al. [2006, 2003]).

One drawback of using VaR metrics is that optimization problems with VaR measures are non-convex. An alternative measure is conditional value-at-risk (CVaR). CVaR qualifies the average of the tail of the loss distribution and is convex (Rockafellar et al. [2000]). Romeijn et al. [2006, 2003] introduced a conditional value at risk (CVaR) formulation for IMRT to present bounds in the tails of the dose distribution for the overdose of healthy tissues and under-dose of the target. CVaR helps formulate clinical DVH criteria while keeping the model tractable.

IMRT planning is subject to uncertainties. Examples include uncertainty in dose calculation and organ motion during treatment. Robust optimization allows for incorporating uncertainties into IMRT planning (e.g., Chan et al. [2006], Lohr et al. [2009], Chu et al. [2005]). In IMRT, it is essential to know the exact location of the tumour and healthy tissues. However, the measurement of these locations is prone to errors of two types: 1) "systematic errors," which are due to the methods of observation or instruments used, for example, the image of the tumour may be not accurate; 2) "random errors," which are due to any natural variations in the process. An example for a random error is that patients may not always remain in the same position on the treatment couch. These errors cause uncertainty about the tumour's location and the organs at risk. Another example is breathing motion uncertainty, because the chest may move unpredictably during treatment. These uncertainties cause blurring of the dose distribution, which may result in excessive radiation to healthy tissues and under-dosing of the tumour. The dose blurring can be modelled statistically as a dose-distribution problem with the motion probability mass function (PMF). Bortfeld et al. [2008],Baum et al. [2006] and Unkelbach and Oelfke [2004] developed a robust optimization for the problem with motion uncertainty. Chan et al. [2014] presented a robust optimization framework for IMRT models with CVaR constraints and applied it to the IMRT for breast cancer. In their problem, the constraints limit overdosing and under-dosing of a partial volume of the healthy organs and the tumour, respectively, by considering that a tumour's position may be uncertain due to breathing motion. Several papers proposed RO models for the IMRT problem. We refer the readers to An et al. [2017], Unkelbach et al. [2018], Vrančić et al. [2009] and Barragán-Montero et al. [2019] for more information.

One drawback of robust optimization for IMRT is that it considers only the worst-case uncertainty scenario. However, in IMRT, it is also essential to consider non-worst-case scenarios. Using the concept of Pareto efficiency in robust optimization, this thesis presents PRO solutions for IMRT problems that focus on both worst-case and non-worst-case scenarios.

We also propose light Pareto robust optimal solutions for the robust form of the IMRT problem. These solutions focus on non-worst case scenarios more than PRO solutions.

1.4 Inverse Optimization

A conventional "forward" optimization problem determines the optimal value of a problem considering a given set of parameters. Inverse optimization, on the other hand, estimates the parameters of a mathematical optimization model considering a given observation as its optimal solution [Ahuja and Orlin, 2001]. The inverse approach aims to find the underlying optimization model needed to render a given solution optimal.

Inverse optimization has found applications in different areas, such as, finance [Bertsimas et al., 2012], and medical sciences [Erkin et al., 2010, Ayer, 2015], and the subject has been studied in different frameworks, such as, integer and mixed-integer (Schaefer [2009], Wang [2009], Lamperski and Schaefer [2015]) and combinatorial optimization (Ahmadian et al. [2018]).

In Chapter 3, we propose an inverse optimization model to determine if an RO solution is PRO. We also use this concept to determine an uncertainty set corresponding to a PRO solution where that solution is optimal for all members of the uncertainty set. Inverse optimization has also been applied to IMRT optimization. In chapter 5, we apply inverse optimization to the IMRT problem and present a robust inverse model for this problem. We will discuss the robust inverse IMRT problem as a potential topic for future work in Chapter 5.

1.5 Thesis Structure

The rest of this thesis is structured as follows. Chapter 2 will present Pareto robust optimization models for four types of IMRT problems. These models present solutions that other RO solutions cannot dominate. We will present an application of the Pareto optimization for an IMRT problem of breast cancer with a hypothetical data set. We will show that Pareto robust optimization significantly decreases the side effects of overdosing on the tissues.

Chapter 3 will define Pareto robust frontier as the set of all PRO solutions to a robust optimization problem. This chapter will present a method to generate alternative optimal solutions for any linear optimization problem. In Chapter 3, we will propose an algorithm to generate all RO solutions for a robust optimization problem using this method. We will also present a method to determine if an RO solution is PRO. Using this method, we will examine all RO solutions and present a procedure to determine all PRO solutions to a robust optimization problem. We will denote the set of all PRO solutions as the Pareto robust frontier for an RO problem.

Chapter 3 will present a method to generate a set of all uncertainty realizations for which a given PRO is optimal. Using this method, we can compare PRO solutions and present the one that is an optimal solution for the maximum number of elements in a given subset of uncertainty. We denote this PRO solution as a "superior" PRO solution for the given set.

Chapter 4 will define new solutions for robust optimization problems by violating the RO solutions from optimality for the worst-case scenario. We call these solutions light PRO solutions. We will prove that these solutions are less conservative than the PRO solutions.

Chapter 5 will present an idea for the direction of the future work. We will present a robust inverse optimization form for the IMRT problem. We will discuss the idea of generating PRO solutions and light Pareto robust optimal solutions for this problem as a direction for the future research.

Finally, Chapter 6 will conclude the thesis.

Chapter 2

Pareto Efficiency in Radiation Therapy

2.1 Introduction

Intensity-modulated radiation therapy is a well-known method of radiation therapy. It has been modeled as an optimization problem. Due to many sources of uncertainty in this methodology, one of the well-studied area in this subject is robust optimization. A robust optimal solution to the RT problem is optimal for the worst-case scenario. However, it is conservative and may not be the ideal solution for a non-worst-case scenario. In this chapter, to present a less conservative solution, we use Pareto robust optimization to generate a PRO solution for this problem.

Several criteria can be considered in the IMRT problem. This chapter will discuss two general forms of criteria: namely, the full volume and the partial volume criteria. For both forms, we present models that result in a PRO solution. In Section 2.2, based on the work of Iancu and Trichakis [2014], we discuss the background of using Pareto efficiency for robust optimization. We then present the general definitions of PRO solutions and basic theoretical results related to these solutions in Section 2.2. Next, the general optimization form of the IMRT problem is presented in Section 2.3. Lastly, we present PRO solutions for four categories of the IMRT problem in Section 2.4.

2.2 PRO Solutions for Robust Optimization Problems

A PRO solution for a robust optimization problem assures improvement in the objective function or slack size for a non-worst-case scenario without deteriorating the function in other scenarios. We first present the definition of PRO solutions and a method for finding a PRO solution of a robust optimization problem with uncertainty in the objective function [Iancu and Trichakis, 2014]. Then, we present the background for a setting with uncertainty in its constraints. Finally, the robust optimization problem with multiple constraints will be discussed.

2.2.1 Uncertainty in the objective function

Consider the following robust optimization problem with uncertainty in the objective function:

$$\max_{\mathbf{x}\in\mathcal{X}} \min_{\mathbf{p}\in\mathcal{U}} \mathbf{p}'\mathbf{x}, \tag{2.1}$$

where $\mathcal{X}$ is the feasible region for decision variable $\mathbf{x}$, and $\mathcal{U}$ is the uncertainty set on parameter $\mathbf{p}$ defined as:

$$\begin{aligned} \mathcal{X} &= \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}, \quad \text{and} \\ \mathcal{U} &= \{\mathbf{p} \in \mathbb{R}^n : \mathbf{D}\mathbf{p} \geq \mathbf{d}\}, \end{aligned}$$

respectively. Note that $\mathbf{A} \in \mathbb{R}^{m_x \times n}$, $\mathbf{D} \in \mathbb{R}^{m_u \times n}$, $\mathbf{b} \in \mathbb{R}^{m_x}$, and $\mathbf{d} \in \mathbb{R}^{m_u}$ for some positive integers n, m_x and m_u. Iancu and Trichakis [2014] presented the definition for a PRO solution. To solve this problem, first we reformulate $\min_{\mathbf{p}\in\mathcal{U}} \mathbf{p}'\mathbf{x}$ using the dual of this model as follows:

$$\begin{aligned} \underset{\mathbf{y}}{\text{maximize}} \quad & \mathbf{y}'\mathbf{d} \\ \text{subject to} \quad & \mathbf{D}'\mathbf{y} = \mathbf{x}, \\ & \mathbf{y} \geq \mathbf{0}, \\ & \mathbf{y} \in \mathbb{R}^{m_u}. \end{aligned} \tag{2.2}$$

Therefore, we can write Model (2.1) as

$$\begin{aligned} \max_{\mathbf{x}\in\mathcal{X}} \max_{\mathbf{y}} \quad & \mathbf{y}'\mathbf{d} \\ \text{subject to:}\quad & \mathbf{D}'\mathbf{y} = \mathbf{x}, \\ & \mathbf{y} \geq \mathbf{0}, \\ & \mathbf{y} \in \mathbb{R}^{m_u}, \end{aligned} \tag{2.3}$$

and so,

$$\begin{aligned} \max_{\mathbf{x}} \quad & \mathbf{y}'\mathbf{d} \\ \text{subject to:}\quad & \mathbf{D}'\mathbf{y} = \mathbf{x} \\ & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{y} \geq \mathbf{0} \\ & \mathbf{y} \in \mathbb{R}^{m_u}, \mathbf{x} \in \mathbb{R}^{n}. \end{aligned} \tag{2.4}$$

Let z^{RO} be equal to $\mathbf{y}^{*\prime}\mathbf{d}$ for an optimal solution $\mathbf{y}^*$ to Model (2.4). Thus, for any optimal solution $(\mathbf{x}^*, \mathbf{y}^*)$ to Model (2.4), we have $\mathbf{D}'\mathbf{y}^* = \mathbf{x}^*$, $\mathbf{A}\mathbf{x}^* \leq \mathbf{b}$ and $\mathbf{y}^{*\prime}\mathbf{d} \geq z^{RO}$. Since Model (2.4) is a reformulation of Model (2.1), we can present the set of all robust optimal solutions to Model (2.1) as

$$\mathcal{X}^{RO} = \{\mathbf{x} \in \mathcal{X} : \exists \mathbf{y} \in \mathbb{R}_+^{m_u} \text{ such that } \mathbf{D}'\mathbf{y} = \mathbf{x}, \mathbf{y}'\mathbf{d} \geq z^{RO}\}.$$

Based on Iancu and Trichakis [2014], a solution $\mathbf{x}$ is a PRO solution for problem (2.1), if it is robustly optimal and there is no other $\bar{\mathbf{x}}$ such that

$$\mathbf{p}'\bar{\mathbf{x}} \geq \mathbf{p}'\mathbf{x}, \quad \forall \mathbf{p} \in \mathcal{U}, \text{ and}$$
$$\bar{\mathbf{p}}'\bar{\mathbf{x}} > \bar{\mathbf{p}}'\mathbf{x}, \quad \text{for some } \bar{\mathbf{p}} \in \mathcal{U}.$$

For set $\mathcal{U}$, consider $ri(\mathcal{U})$ as the relative interior of $\mathcal{U}$ which is defined as $ri(\mathcal{U}) = \{\mathbf{p_1} \in \mathcal{U} | \forall \mathbf{p2} \in \mathcal{U} \quad \exists \lambda > 1 : \lambda \mathbf{p_1} + (1-\lambda)\mathbf{p_2} \in \mathcal{U}\}$. Now, we can find a PRO solution to Model (2.1) using an element of $ri(\mathcal{U})$.

For any $\bar{\mathbf{p}} \in ri(\mathcal{U})$, it is shown that all of the optimal solutions for the problem

$$\max_{\mathbf{x} \in \mathcal{X}^{RO}} \bar{\mathbf{p}}'\mathbf{x} \tag{2.5}$$

are PRO solution [Iancu and Trichakis, 2014]. The following example presents two robust optimal solutions for a network problem such that one of them is a Pareto robust optimal solution that dominates the other one which is not a PRO solution. The general form of this example with $n \geq 3$ number of links is discussed by Iancu and Trichakis [2014]. The following example discusses the problem with thirteen links.

Example 2.2.1. Consider the following network structure which has two channels, A and B. Both channels have unit capacity. Consider 13 links $l_0, l_1, ..., l_{12}$ such that l_0 and l_1 utilize channel A, l_2 utilizes both channels A and B. Links $l_3, ..., l_{12}$ utilize channel B. The transmission rate of link l_i over channel A is denoted by a_i for $i = 0, 1, 2$. The transmission rate of link l_i over channel B is denoted by b_i for $i = 2, ..., 12$. Link l_0 is for general purposes, and $l_1, ..., l_{12}$ are for emergency purposes. Let x_i be the transmission rate of the

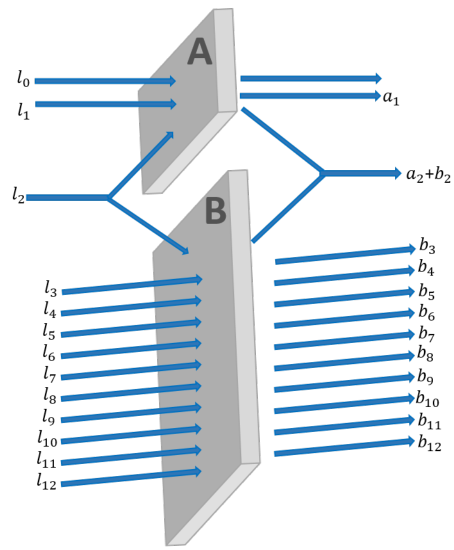


Figure 2.1: Network structure of two channels utilized by thirteen links [Iancu and Trichakis, 2014]

links l_i for $i = 1, ..., 12$. Then $x_1 = a_1$, $x_2 = a_2 + b_2$ and $x_i = b_i$ for $i = 3, ..., 12$. Consider f_i as the fraction of emergency transmission via l_i for $i = 1, ..., 12$. Therefore, the net emergency transmission rate is $\mathbf{f}^T\mathbf{x} = \sum_{i=1}^{12} f_i x_i$. Let $\mathcal{U} = \{\mathbf{f} \in \mathbb{R}_+^{12}, \mathbf{1^T f} = 1\}$ be the set of all realizations of vector $\mathbf{f}$. Therefore, we have the following robust problem.

$$
\begin{aligned}
\underset{\mathbf{x},\mathbf{a},\mathbf{b}}{\text{maximize}} \quad & \min_{\mathbf{f}\in\mathcal{U}} \mathbf{f}'\mathbf{x} \\
\text{subject to} \quad & x_1 = a_1, \\
& x_2 = a_2 + b_2, \\
& x_i = b_i \quad \text{for } i = 3, ..., 12, \\
& a_0 + a_1 + a_2 = 1, \\
& b_2 + b_3 + ... + b_{12} = 1, \\
& a_i, b_i \geq 0
\end{aligned}
\tag{2.6}
$$

The optimal objective value for this problem is $\mathbf{f}^T\mathbf{x} = \frac{1}{10}$. Let $\mathcal{X}$ be the set of all feasible solutions to (2.6). The set of all robust optimal solutions is $\mathcal{X}^{RO} = \{(\mathbf{a}, \mathbf{b}, \mathbf{x}) \in \mathcal{X}, \mathbf{x} \geq \frac{1}{10}\mathbf{1}\}$ where $\mathbf{a} = (1 - x_1 - x_2, x_1, x_2)$, $\mathbf{b} = (0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1)$ and $\mathbf{x} = (x_1, x_2, 0.1, ..., 0.1)$ where $x_1 + x_2 \leq 1$. These solutions are optimal considering worst-case realization of $\mathbf{f}$. For instance, the following solutions are robust optimal solution to Model (2.6).

$$\mathbf{a^{RO}} = (0.8, 0.1, 0.1), \mathbf{b^{RO}} = (0, 0.1, 0.1, ..., 0.1), \mathbf{x^{RO}} = (0.1, 0.1, ..., 0.1)$$

and

$$\mathbf{a^{P}} = (0, 0.9, 0.1), \mathbf{b^{P}} = (0, 0.1, 0.1, ..., 0.1).\mathbf{x^{P}} = (0.9, 0.1, ..., 0.1),$$

We will next show that $\mathbf{x^P}$ is not dominated by other RO solutions and is a PRO one,

and also $\mathbf{x^P}$ dominates $\mathbf{x^{RO}}$. Each element $\mathbf{f}$ of $\mathcal{U}$ can be written as a linear combination of $\mathbf{e_1}, \mathbf{e_2}, ..., \mathbf{e_{12}}$, such that $\mathbf{e_i}$ is a unit vector in $\mathbb{R}^{12}$, with i-th entry of $\mathbf{e_i}$ being 1, and all other entries zero. Thus, we can write $\mathbf{f}$ as $\sum_{i=1}^{12} c_i \mathbf{e_i}$ such that $c_i \geq 0$ for all $1 \leq i \leq 12$.
To prove that $\mathbf{x^P}$ is a PRO solution, suppose there exists $\mathbf{x^{P_1}} = (x_1, x_2, 0.1, ..., 0.1)$ that dominates $\mathbf{x^P}$. We have $\mathbf{e_i}'\mathbf{x^P} = \mathbf{e_i}'\mathbf{x^{P_1}}$ for all $1 \leq i \leq 12$. Without loss of generality, suppose that $\mathbf{e_1}'\mathbf{x^P} < \mathbf{e_1}'\mathbf{x^{P_1}}$. Thus, $x_1 > 0.9$. Since, $x_1 + x_2 \leq 1$, $x_2 < 0.1$, and so $\mathbf{e_2}'\mathbf{x^P} > \mathbf{e_2}'\mathbf{x^{P_1}}$. Therefore, $\mathbf{x^{P_1}}$ does not dominate $\mathbf{x^P}$, and it is a contradiction. Hence, $\mathbf{x^P}$ is not dominated by other RO solutions, and is PRO.
Now, we prove that $\mathbf{x^P}$ dominates $\mathbf{x^{RO}}$. We have $\mathbf{e_1}'\mathbf{x^P} > \mathbf{e_1}'\mathbf{x^{RO}}$ and for each $i = 1, ..., 12$, $\mathbf{e_i}'\mathbf{x^P} = \mathbf{e_i}'\mathbf{x^{RO}}$. Thus, for each $\mathbf{f} \in \mathcal{U}$, $\mathbf{f}'\mathbf{x^P} \geq \mathbf{f}'\mathbf{x^{RO}}$. Therefor, $\mathbf{x^P}$ dominates $\mathbf{x^{RO}}$. Thus, $\mathbf{x^P}$ has the same objective value as the solution $\mathbf{x}$ for the worst-case scenario, and $\mathbf{x^P}$, however, it has a better objective value for non-worst-cases.

2.2.2 Uncertainty in constraints

Now, consider a problem with uncertainty in constraints as follows:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c}'\mathbf{x} \\ \text{subject to: } & \mathbf{A}\mathbf{x} \geq \mathbf{b} \quad \forall \mathbf{A} \in \mathcal{U}_A, \end{aligned} \tag{2.7}$$

where $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$ and $\mathcal{U}_A \subset \mathbb{R}^{m \times n}$ for some positive integers n and m. A PRO solution is defined by Iancu and Trichakis [2014] as follows.

Let

$$\mathcal{X}^{RO} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{c}'\mathbf{x} \leq z^{RO}, \mathbf{A}\mathbf{x} - \mathbf{b} \geq 0, \quad \forall \mathbf{A} \in \mathcal{U}_A\}$$

denote the set of all of the RO solutions. A solution $\mathbf{x}$ is a PRO solution for problem (2.7) if it is a robust optimal solution and there is no solution $\bar{\mathbf{x}} \in \mathcal{X}^{RO}$ such that

$$\begin{aligned} &\mathbf{v}'s(\bar{\mathbf{x}}, \mathbf{A}) \geq \mathbf{v}'s(\mathbf{x}, \mathbf{A}), \quad \forall \mathbf{A} \in \mathcal{U}_A, \text{ and} \\ &\mathbf{v}'s(\bar{\mathbf{x}}, \bar{\mathbf{A}}) > \mathbf{v}'s(\mathbf{x}, \bar{\mathbf{A}}), \quad \text{ for some } \bar{\mathbf{A}} \in \mathcal{U}_A \end{aligned} \tag{2.8}$$

where $\mathbf{v}$ is a slack value vector in $\mathbb{R}^m$ and $s(\mathbf{x}, \mathbf{A}) = \mathbf{A}\mathbf{x} - \mathbf{b}$ [Iancu and Trichakis, 2014]. Assuming $\mathcal{U} = \{\mathbf{p} = \mathbf{A}'\mathbf{v} : \mathbf{A} \in \mathcal{U}_A\}$, for any $\hat{\mathbf{p}} \in ri(\mathcal{U})$, all solutions to the problem

$$\max_{\mathbf{x} \in \mathcal{X}^{RO}} \hat{\mathbf{p}}'\mathbf{x}$$

are PRO solutions for problem (2.7), [Iancu and Trichakis, 2014].

Problem (2.7) can be expanded as

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{c}'\mathbf{x} \\ \text{subject to} \quad & \mathbf{a_i}'\mathbf{x} \geq b_i \quad \forall \mathbf{a_i} \in \mathcal{U}_{a_i}, \text{ for } i = 1, ..., m_x, \\ & \mathbf{x} \in \mathbb{R}^n \end{aligned} \tag{2.9}$$

where $b_i \in \mathbb{R}$, and $\mathcal{U}_{a_i} = \{\mathbf{a_i} | \mathbf{D_i}\mathbf{a_i} \geq \mathbf{d_i}\}$ for some $\mathbf{D_i} \in \mathbb{R}^{m_u \times n}$ and $\mathbf{d_i} \in \mathbb{R}^{m_u}$. We will be using this expanded form in Chapter 3.

2.2.3 Multiple types of constraints

In this section, we extend the idea of Pareto robust optimality for problems with uncertainty in multiple types of constraints. The problem can have both types of greater-than or less-than constraints. Consider the following problem:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c}'\mathbf{x} \\ \text{Subject to: } & \mathbf{A_1}\mathbf{x} \geq \mathbf{b_1} \quad \mathbf{A_1} \in \mathcal{U}_1, \\ & \mathbf{A_2}\mathbf{x} \leq \mathbf{b_2} \quad \mathbf{A_2} \in \mathcal{U}_2. \end{aligned} \tag{2.10}$$

The vector of slacks for the constraints is as follows

$$s(\mathbf{x}, \mathbf{A_1}) = \mathbf{A_1}\mathbf{x} - \mathbf{b_1} \quad \text{for } \mathbf{A_1} \in \mathcal{U}_1,$$

$$s(\mathbf{x}, \mathbf{A_2}) = \mathbf{b_2} - \mathbf{A_2}\mathbf{x} \quad \text{for } \mathbf{A_2} \in \mathcal{U}_2.$$

Considering slack value vectors $\mathbf{v_1} \in \mathbb{R}^{m_1}$ and $\mathbf{v_2} \in \mathbb{R}^{m_2}$, we have the following definition. A solution $\mathbf{x}$ is called a Pareto robust optimal solution if it is a robust optimal solution and there is no $\bar{\mathbf{x}}$ such that

$$\begin{aligned} \mathbf{v}'s(\bar{\mathbf{x}}, \mathbf{A}) &\geq \mathbf{v}'s(\mathbf{x}, \mathbf{A}) \quad \forall \mathbf{A} \in \mathcal{U}_1 \cup \mathcal{U}_2 \\ \mathbf{v}'s(\bar{\mathbf{x}}, \bar{\mathbf{A}}) &> \mathbf{v}'s(\mathbf{x}, \bar{\mathbf{A}}) \quad \text{for some } \bar{\mathbf{A}} \in \mathcal{U}_1 \cup \mathcal{U}_2, \end{aligned}$$

where $\mathbf{v} = \mathbf{v_1}$ if $\mathbf{A} \in \mathcal{U}_1$ and $\mathbf{v} = \mathbf{v_2}$ if $\mathbf{A} \in \mathcal{U}_2$.

Let $\mathcal{V}_1 = \{\mathbf{p_1} | \mathbf{p_1} = \mathbf{A_1 v_1}$ for all $\mathbf{A_1} \in \mathcal{U}_1\}$ and $\mathcal{V}_2 = \{\mathbf{p_2} | \mathbf{p_2} = \mathbf{A_2 v_2}$ for all $\mathbf{A_2} \in \mathcal{U}_2\}$. Then, we can generate a PRO solution for problem (2.10) as follows.

Proposition 2.2.1. *If $\bar{\mathbf{p}}_1 \in ri(\mathcal{V}_1)$ and $\bar{\mathbf{p}}_2 \in ri(\mathcal{V}_2)$, the solutions of the following problem are Pareto robust optimal solutions:*

$$\max_{x \in \mathcal{X}^{RO}} (\bar{\mathbf{p}}_1 - \bar{\mathbf{p}}_2)\mathbf{x}. \tag{2.11}$$

Proof. See Appendix A.1.1. □

2.3 PRO Formulation of the IMRT Problem

In IMRT planning, each beam is assumed to be divided into many small beamlets, each of which can take a different radiation intensity. The decision variable is to find the optimal intensity of each beamlet subject to a set of constraints that capture the resulting dose to the tumour and healthy tissue.

To model the IMRT problem, we have the following parameters. The set of all beamlets is denoted by $\mathcal{B}$. The irradiated area divided to a set of small volume elements called voxel. The dose points to the voxels. Thus, we can represent each tissue by a set of voxels. We denote the tumour by the voxel set $\mathcal{T}$ and the healthy tissues by the voxel set $\mathcal{H}$. The set of all voxels, including the tumour and the adjacent healthy tissues, is denoted by $\mathcal{V} = \mathcal{T} \cup \mathcal{H}$. Consider $D_{v,b}$ as the dose that voxel v receives from beamlet b, θ_v as the prescribed dose for $v \in \mathcal{T}$, and δ_v as an upper bound for the dose delivered to the tumour voxel v. Let $\underline{\mathcal{A}}$

and $\bar{\mathcal{A}}$ sets index the upper α levels for $\mathcal{H}$ and lower β levels for $\mathcal{T}$, respectively. let U_α be an upper bound for the average dose received by the voxels in $\mathcal{H}$ and L_β be a lower bound for the average dose received by the voxels in $\mathcal{T}$.

The decision variable is the intensity of each beamlet $b \in \mathcal{B}$, denoted as w_b. Variable $\bar{\zeta}_\alpha$ is the minimum dose level such that no more than $(1-\alpha)\%$ of $\mathcal{H}$ receives more than that dose level. Variable $\underline{\zeta}_\beta$ is the maximum dose level such that no more than $(1-\beta)\%$ of $\mathcal{T}$ receives less than that dose level. Variable $\bar{d}_{v,\alpha}$ is the difference between the dose received by voxel v and $\bar{\zeta}_\alpha$. Similarly, $\underline{d}_{v,\beta}$ is the difference between the dose received by voxel v and $\underline{\zeta}_\beta$.

The general optimization form of the IMRT problem is presented in (2.12). The goal is to minimize the total dose received by all voxels while meeting the dose constraints on $\mathcal{T}$ and $\mathcal{H}$. Therefore, the objective function of the problem is the total dose delivered to all voxels:

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

We consider four types of constraints for this problem.

(1) Under-dosing of the tumour: Each voxel of the tumour should receive the prescribed dose of radiation for the tumour.

(2) Overdosing of healthy tissues: Each voxel of the organ at risk must receive lower than a certain radiation dose threshold. This constraint provides an upper bound for the dose delivered to the OAR.

(3) Dose-volume criteria for the tumour: A lower bound for the mean dose received by the subset of "voxels" in the tumour.

(4) Dose-volume criteria for healthy tissues: An upper bound for the mean dose received by the subset of "voxels" in healthy tissues.

The general IMRT problem is as follows:

$$\text{Minimize} \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} D_{v,b} w_b \tag{2.12a}$$

$$\text{Subject to: } \sum_{b \in \mathcal{B}} D_{v,b} w_b \geq \theta_v \quad \forall v \in \mathcal{T}, \tag{2.12b}$$

$$\sum_{b \in \mathcal{B}} D_{v,b} w_b \leq \delta_v \quad \forall v \in \mathcal{H}, \tag{2.12c}$$

$$\underline{\zeta}_\beta - \frac{1}{(1-\beta)|\mathcal{T}|} \sum_{v \in \mathcal{T}} \underline{d}_{v,\beta} \geq L_\beta, \quad \forall \beta \in \underline{\mathcal{A}}, \tag{2.12d}$$

$$\underline{d}_{v,\beta} \geq \underline{\zeta}_\beta - \sum_{b \in \mathcal{B}} D_{v,b} w_b, \quad \forall v \in \mathcal{T}, \forall \beta \in \underline{\mathcal{A}}, \tag{2.12e}$$

$$\bar{\zeta}_\alpha + \frac{1}{(1-\alpha)|\mathcal{H}|} \sum_{v \in \mathcal{H}} \bar{d}_{v,\alpha} \leq U_\alpha, \quad \forall \alpha \in \bar{\mathcal{A}}, \tag{2.12f}$$

$$\bar{d}_{v,\alpha} \geq \sum_{b \in \mathcal{B}} D_{v,b} w_b - \bar{\zeta}_\alpha, \quad \forall v \in \mathcal{H}, \alpha \in \bar{\mathcal{A}}, \tag{2.12g}$$

$$\bar{\zeta}_\alpha \geq 0, \quad \forall \alpha \in \bar{\mathcal{A}}, \tag{2.12h}$$

$$\underline{\zeta}_\beta \geq 0, \quad \forall \beta \in \underline{\mathcal{A}}, \tag{2.12i}$$

$$\bar{d}_{v,\alpha} \geq 0, \quad \forall v \in \mathcal{V}, \alpha \in \bar{\mathcal{A}}, \tag{2.12j}$$

$$\underline{d}_{v,\beta} \geq 0, \quad \forall v \in \mathcal{V}, \beta \in \underline{\mathcal{A}}, \tag{2.12k}$$

$$w_b \geq 0, \forall b \in \mathcal{B}. \tag{2.12l}$$

Constraint (2.12b) presents a lower bound for the dose that is received by the tumour, and constraint (2.12c) presents an upper bound for the dose that is received by healthy tissues around the tumour. Constraints (2.12d) and (2.12e) control the tail of the dose distribution to the tumour, by defining a lower bound for different lower β levels. Constraints (2.12f) and (2.12g) control the tail of the dose distribution to healthy tissues, by defining an upper bound for different lower α levels.

In this model, as mentioned in Chapter 1, there exist uncertainties in the location of the tumour or healthy tissues around it, which can be modelled as a robust optimization problem.

Let $\mathcal{X}$ be a finite set of possible phases of the motion (e.g., inhale, exhale). The probability mass function (PMF) of the breathing motion determines the proportion of time that patient spends in each of a breathing motion states $x \in \mathcal{X}$, from inhale to exhale. We define the motion PMF by a non-negative real function $f : \mathcal{X} \to \mathbb{R}$, where $\sum_{x \in \mathcal{X}} f(x) = 1$. We start with a nominal motion PMF denoted as $\mathbf{p}$, which is constructed with the data taken during treatment planning. Now, we consider a realized PMF denoted as $\tilde{\mathbf{p}}$, where

$$\mathbf{p}(x) - \underline{\mathbf{p}}(x) \leq \tilde{\mathbf{p}}(x) \leq \mathbf{p}(x) + \bar{\mathbf{p}}(x) \quad x \in \mathcal{X},$$

such that $\underline{\mathbf{p}}$ and $\bar{\mathbf{p}}$ are the bounds for lower and upper errors between the actual and nominal PMF, respectively. We assume that $\mathbf{p} - \underline{\mathbf{p}} \geq 0$ and $\mathbf{p} + \bar{\mathbf{p}} \leq 1$. Therefore, the uncertainty set can be defined as follows:

$$\mathcal{U}_p = \{\tilde{\mathbf{p}} \in \mathbb{R}^{|\mathcal{X}|} : \tilde{\mathbf{p}}(x) \in [\mathbf{p}(x) - \underline{\mathbf{p}}(x), \mathbf{p}(x) + \bar{\mathbf{p}}(x)] \quad \forall x \in \mathcal{X}; \sum_{x \in \mathcal{X}} \tilde{\mathbf{p}}(x) = 1\}. \tag{2.13}$$

The motion affects the tumour and the healthy tissues surrounding it.

In Section 2.4, we consider different classes of IMRT problems by choosing different constraints from the Model (2.12).

2.4 Application of Pareto Approach in IMRT

Consider the following four RO models for the IMRT problem.

(1) Delivering the prescribed dose to all voxels of the tumour with uncertainty in the tumour's motion; we call this model "Uncertainty in the constraint: Full volume criteria".

(2) Having dose criteria for a partial volume of the tumour with uncertainty in the tumour's motion; "Uncertainty in the constraint: Partial volume criteria ".

(3) Delivering the prescribed dose to all voxels of the tumour with uncertainty in OAR's motion; "Uncertainty in the objective function: Full volume criteria ".

(4) Having dose criteria for a partial volume of the healthy organs while delivering the prescribed dose to the tumour with uncertainty in OAR's motion; we call this model "Uncertainty in the objective function: Partial volume criteria ".

In the remainder of this chapter, we present the PRO solutions for them in the following four subsections.

2.4.1 Uncertainty in the constraint: full volume criteria

The full volume IMRT problem with uncertainty in constraint has uncertainty in the position of the tumour. This problem aims to minimize the total radiation dose received by all of the voxels, considering that the tumour receives the required dose.

We consider two types of problems. First, we only have one constraint on the lower bound for the dose received by the tumour voxels. The second one presents both the lower bound and upper bound on the dose delivered to the tumour. We discuss these problems with single and multiple constraints, in what follows. Suppose $\Delta_{v,x,b}$ is the dose of the per-unit intensity of beamlet b, delivered to voxel v when the anatomy is in the phase x.

Single Constraint: Using the approach demonstrated by Bortfeld et al. [2008], the formulation of the single constraint problem is

$$\min_{w_b} \sum_{v\in\mathcal{V}} \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b} p(x) w_b \tag{2.14}$$

$$\text{subject to: } \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b} \tilde{p}(x) w_b \geq \theta_v \quad \forall v \in \mathcal{T}, \tilde{p} \in \mathcal{U}_p, w_b \geq 0, b \in \mathcal{B}.$$

This problem is an RO problem and the RO solution is optimal for the worst case. The worst-case occurs when the slack is as low as possible for all constraints. Here, we assume that the goal is to deliver as much radiation as possible to the tumour. In other words, the left side of the constraint should be as high as possible. Therefore, larger slacks are preferable. The PRO solution for this problem is a robust solution such that no slacks can be improved without sacrificing other slacks. Using the approach given by Iancu and Trichakis [2014], we present a PRO solution for this problem in Theorem 2.4.1. Let

$\mathbf{t} = (t_1, ..., t_{|\mathcal{T}|})$ be the value vector of constraints of Model (2.14); that is, a vector that determines the importance of the voxel of the tumour. Consider Z^{RO} as the optimal objective value of the problem (2.14). The following theorem presents an RO solution to Model (2.14) such that there is not another solution that delivers more dose to the tumour in all tumour motion's PMFs.

Theorem 2.4.1. *For $\hat{p} \in ri(\mathcal{U}_p)$, the solution of following problem is a PRO solution to Model (2.14).*

$$\max_{(w_1,...,w_{|\mathcal{B}|})} \sum_{v\in\mathcal{T}} \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} t_i \Delta_{v,x,b} \hat{p}(x) w_b \tag{2.15}$$

subject to:

$$\begin{aligned}
&\sum_{v\in\mathcal{V}} \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b} p(x) w_b \leq Z^{RO},\\
&\sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b} p(x) w_b - \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b} \underline{p}(x) w_b +\\
&\sum_{x\in\mathcal{X}} \underline{p}(x) q_v - \sum_{x\in\mathcal{X}} r_{v,x} \geq \theta_v \quad \forall v \in \mathcal{T},\\
&(\bar{p}(x) + \underline{p}(x)) q_v - r_{v,x} \leq \sum_{b\in\mathcal{B}} \Delta_{v,x,b} (\bar{p}(x) + \underline{p}(x)) w_b \quad \forall v \in \mathcal{T}, \forall x \in \mathcal{X},\\
&q_v \geq 0 \quad \forall v \in \mathcal{T},\\
&r_{v,x} \geq 0 \quad \forall v \in \mathcal{T}, x \in \mathcal{X}.
\end{aligned}$$

Proof. See Appendix A.1.2. □

Sometimes overdosing the tumour can cause side effects such as skin burns. The goal in such cases is to minimize the side effect by minimizing the total dose delivered to the tumour while still delivering the prescribed dose to each tumour voxel. To find such a PRO solution, we have the same model as problem (2.15) except that its objective function, is changed to

$$\min_{(w_1,\dots,w_{|\mathcal{B}|})} \sum_{v\in\mathcal{T}} \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} t_i \Delta_{v,x,b}\hat{p}(x) w_b.$$

To avoid overdosing to the tumour, we define an upper bound for the dose delivering to each tumour voxel. Thus, in this case, the IMRT problem has two types of constraints. We will discuss this problem and present a PRO solution for it.

Multiple Constraints: Now consider the case where both upper bound and lower bound are imposed the tumour dose. The mathematical formulation of this problem can be written as follows,

$$\min \sum_{v\in\mathcal{V}} \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b} p(x) w_b \tag{2.16}$$

$$\text{subject to:} \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b}\tilde{p}(x) w_b \geq \theta_v \quad \forall v \in \mathcal{T}, \tilde{p} \in \mathcal{U}_p, w_b \geq 0, b \in \mathcal{B}, \tag{2.17}$$

$$\sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b}\tilde{p}(x) w_b \leq \delta_v \quad \forall v \in \mathcal{T}, \tilde{p} \in \mathcal{U}_p, w_b \geq 0, b \in \mathcal{B}. \tag{2.18}$$

Let $\mathbf{t} = (t_1, ..., t_{|\mathcal{T}|})$ and $\mathbf{s} = (s_1, ..., s_{|\mathcal{T}|})$ be the value vectors of the tumour voxels corresponding to the first and second constraints respectively. Vectors $\mathbf{t}$ and $\mathbf{s}$ determine the importance of the tumour voxels for the constraints (2.17) and (2.18), respectively. Consider $\mathcal{W}^{RO}$ as the set of all RO solutions for Model (2.16). Theorem 2.4.1 can be extended as follows.

Theorem 2.4.2. *Pick $\hat{p} \in ri(\mathcal{U}_p)$, then the solution of the following problem is a PRO* *solution to Model* (2.16).

$$\max_{(w_1,\dots,w_{|\mathcal{B}|})\in\mathcal{W}^{RO}} \sum_{v\in\mathcal{T}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}}(t_i - s_i)\Delta_{v,x,b}\hat{p}(x)w_b. \tag{2.19}$$

Proof. See Appendix A.1.3. □

2.4.2 Uncertainty in the constraint: partial volume criteria

In most cases, clinicians consider a percentage of a structure (either the tumour or the adjacent healthy tissues) and determine dose-volume criteria for that volume. We develop this PRO formulation base on the RO formulation of Chan et al. [2014] that includes DVH criteria. The CVaR modelling for IMRT with uncertainty in the constraints can be written as follows [Chan et al., 2014],

$$\begin{aligned}
\min \quad & \sum_{v\in\mathcal{H}}\sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}}\Delta_{v,x,b}p(x)w_b && (2.20)\\
\text{subject to: } & \underline{\zeta}_\beta - \frac{1}{(1-\beta)|\mathcal{T}|}\sum_{v\in\mathcal{T}}\underline{d}_{v,\beta} \ge L_\beta, \quad \forall\beta\in\underline{\mathcal{A}},\\
& \underline{d}_{v,\beta} \ge \underline{\zeta}_\beta - \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}}\Delta_{v,x,b}\tilde{p}(x)w_b, \quad \forall v\in\mathcal{T}, \beta\in\underline{\mathcal{A}}, \tilde{p}\in\mathcal{U}_p,\\
& \underline{\zeta}_\beta \ge 0, \quad \forall\beta\in\underline{\mathcal{A}},\\
& \underline{d}_{v,\beta} \ge 0, \quad \forall v\in\mathcal{T}, \beta\in\underline{\mathcal{A}},\\
& w_b \ge 0, \forall b\in\mathcal{B}.
\end{aligned}$$

Consider Z^{RO} as the optimal objective value of Model (2.20). Theorem 2.4.3 presents PRO solutions for the case in which there is only one DVH constraint ($|\underline{\mathcal{A}}| = 1$). We then present PRO solutions for problem (2.20) in Proposition 2.4.1.

Theorem 2.4.3. *For $\hat{p} \in ri(\mathcal{U}_p)$, the solution of the following problem is a PRO solution for the problem with $\underline{\mathcal{A}} = \{\beta\}$,*

$$
\begin{aligned}
\max_{(\mathcal{W},\underline{\mathcal{D}},\underline{\zeta}_\beta)} \quad & \sum_{v\in\mathcal{T}} t_v(\underline{d}_{v,\beta} - \underline{\zeta}_\beta + \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}\hat{p}(x)w_b) \qquad (2.21)\\
\textit{subject to:} \quad & \sum_{v\in\mathcal{H}}\sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}p(x)w_b \le Z^{RO},\\
& \underline{\zeta}_\beta - \frac{1}{(1-\beta)|\mathcal{T}|}\sum_{v\in\mathcal{T}} \underline{d}_{v,\beta} \ge L_\beta,\\
& \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}p(x)w_b - \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}\underline{p}(x)w_b + \sum_{x\in\mathcal{X}} \underline{p}(x)q_v\\
& - \sum_{x\in\mathcal{X}} r_{v,x} \ge \underline{\zeta}_\beta - \underline{d}_{v,\beta} \quad \forall v \in \mathcal{T},\\
& \sum_{b\in\mathcal{B}} \Delta_{v,x,b}(\bar{p}(x) + \underline{p}(x))w_b + r_{v,x} + (\bar{p}(x) + \underline{p}(x))q_v \ge 0 \quad \forall x \in \mathcal{X}, v \in \mathcal{T}.
\end{aligned}
$$

Proof. See Appendix A.1.4. □

Theorem 2.4.3 discusses the case in which $\underline{\mathcal{A}}$ has one member; the following proposition presents the PRO solutions for the general case. Let Z^{RO} be the optimal objective value of problem (2.20).

Proposition 2.4.1. *For $\hat{p} \in ri(\mathcal{U}_p)$, the solution of the following problem is a PRO solution* *for problem* (2.20),

$$\max_{(\mathcal{W},\underline{\mathcal{D}},\underline{\zeta})} \quad \sum_{\beta\in\underline{\mathcal{A}}}\sum_{v\in\mathcal{T}} t_v(\underline{d}_{v,\beta} + \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}\hat{p}(x)w_b - \underline{\zeta}_\beta) \tag{2.22}$$

$$\begin{aligned}
\textit{subject to:}\quad & \sum_{v\in\mathcal{H}}\sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}p(x)w_b \le Z^{RO},\\
& \underline{\zeta}_\beta - \frac{1}{(1-\beta)|\mathcal{T}|}\sum_{v\in\mathcal{T}} \underline{d}_{v,\beta} \ge L_\beta \quad \forall \beta\in\underline{\mathcal{A}},\\
& \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}p(x)w_b - \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}\underline{p}(x)w_b + \sum_{x\in\mathcal{X}} \underline{p}(x)q_{v,\beta} - \sum_{x\in\mathcal{X}} r_{v,x,\beta} \ge \underline{\zeta}_\beta - \underline{d}_{v,\beta},\\
& \forall v\in\mathcal{T},\ \beta\in\underline{\mathcal{A}},\\
& \sum_{b\in\mathcal{B}} \Delta_{v,x,b}(\bar{p}(x)+\underline{p}(x))w_b + r_{v,x,\beta} + (\bar{p}(x)+\underline{p}(x))q_{v,\beta} \ge 0 \quad \forall x\in\mathcal{X}, v\in\mathcal{T}, \beta\in\underline{\mathcal{A}},\\
& q_{v,\beta}\ge 0 \quad \forall v\in\mathcal{T},\beta\in\underline{\mathcal{A}},\\
& ,r_{v,x,\beta}\ge 0 \quad \forall v\in\mathcal{T}, x\in\mathcal{X},\beta\in\underline{\mathcal{A}}.
\end{aligned}$$

Proof. See Appendix A.1.5. □

So far, we considered uncertainty in constraints. We next, consider uncertainty in the objective function.

2.4.3 Uncertainty in the objective function: full volume criteria

This section presents a method to find a PRO solution for the IMRT problem with uncertainty in the objective parameters. This problem aims to minimize the worst-case dose to all voxels while delivering the prescribed dose to the tumour. This robust optimization problem is as follows

$$\begin{aligned}
&\min_{w}\max_{\tilde{p}} \quad \sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}}\Delta_{v,x,b}\tilde{p}(x)w_b && (2.23)\\
&\text{Subject to:} \quad \sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}}\Delta_{v,x,b}p(x)w_b \geq \theta_v \quad \forall v\in\tau\\
&\qquad\qquad\quad w_b \geq 0
\end{aligned}$$

To find the PRO solutions, we first formulate this problem similar to the formulation of the robust optimization problem given by Iancu and Trichakis [2014] which was a "max min" problem. To do so, we multiple $\Delta_{v,x,b}$ by (-1) and change "min max" function to a "max min" one. Thus, the problem is as follows,

$$\begin{aligned}
&\max_{w}\min_{\tilde{p}} \quad \sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}}-\Delta_{v,x,b}\tilde{p}(x)w_b && (2.24)\\
&\text{Subject to:} \quad \sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}}\Delta_{v,x,b}p(x)w_b \geq \theta_v \quad \forall v\in\tau\\
&\qquad\qquad\quad w_b \geq 0.
\end{aligned}$$

Let Z^{RO} be the optimal value of problem (2.24). We present the Pareto robust optimal solutions of this problem in Theorem 2.4.4.

Theorem 2.4.4. *For any* $\hat{\mathbf{p}} \in ri(\mathcal{U}_p)$ *an optimal solution to the following problem is a*

PRO solution:

$$
\begin{aligned}
\max_{(w_1,..,w_{|\mathcal{B}|})} \quad & \sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} -\Delta_{v,x,b}\hat{p}(x)w_b \\
\textit{subject to:} \quad & \sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} -\Delta_{v,x,b}\bar{p}(x)w_b + \sum_{x\in\mathcal{X}}\bar{p}(x)\bar{q}_x - \sum_{x\in\mathcal{X}}\underline{p}(x)\underline{q}_x \geq Z^{RO}, \\
& \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}\bar{p}(x)w_b \geq \theta_v \;\; \forall v \in \mathcal{T}, \\
& y + \bar{q}_x + \underline{q}_x = \sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}} -\Delta_{v,x,b}w_b \;\; \forall x \in \mathcal{X}.
\end{aligned}
$$

Proof. See Appendix A.1.6. □

In this problem, the goal of the treatment was to minimize the dose delivered to all voxels. Next, we discuss PRO solutions for partial volume criteria.

2.4.4 Uncertainty in the objective function: partial volume criteria

In IMRT, the goal is often to minimize the total dose received by a percentage of the voxels. Chan et al. [2014] presented CVaR modelling for this problem in which the uncertainty is in the objective function, and they minimize the mean dose received by a subset of the voxels. The following problem is a robust optimization.

$$
\begin{aligned}
\min_{w,\bar{\zeta}_\alpha} \quad & \bar{\zeta}_\alpha + \frac{1}{(1-\alpha)|\mathcal{H}|}\sum_{v\in\mathcal{H}}\Big(\sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}}\Delta_{v,x,b}\tilde{p}(x)w_b - \bar{\zeta}_\beta\Big)^+ && (2.25)\\
\text{Subject to:} \quad & \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}}\Delta_{v,x,b}w_b p(x) \geq \theta_v \quad \forall v\in\mathcal{T},\\
& \bar{\zeta}_\alpha \geq 0\\
& w_b \geq 0\\
& \tilde{p}\in\mathcal{U}_p.
\end{aligned}
$$

An RO solution is optimal for the worst-case, and it may not be optimal for any other scenarios, as we mentioned. The PRO solution to this problem is a solution that no other feasible robust solution exists that has a better objective value for at least one scenario, and for all other scenarios, the objective value is not worst. Let Z^{RO} be the optimal objective value to Model (2.25). The following theorem presents a procedure to find a PRO solution for problem (2.25).

Theorem 2.4.5. *For any $\hat{p}\in\mathcal{U}_p$, the solution of the following problem is a PRO solution* *if problem* (2.25).

$$
\begin{aligned}
\max_{(W,Z,\bar{\zeta}_\alpha)} \quad & \sum_{v\in\mathcal{H}}(z_v + \bar{\zeta}_\alpha - \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}}\Delta_{v,x,b}\hat{p}(x)w_b) \qquad (2.26)\\
\textit{subject to: } & \bar{\zeta}_\alpha + \frac{1}{(1-\alpha)|\mathcal{H}|}\sum_{v\in\mathcal{H}} z_v \le Z^{RO},\\
& \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}}\Delta_{v,x,b}p(x)w_b \ge \theta_v \quad \forall v\in\mathcal{T},\\
& z_v + \bar{\zeta}_\alpha - \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}}\Delta_{v,x,b}p(x)w_b + \sum_{x\in\mathcal{X}}\bar{p}(x)r_{v,x} + \sum_{x\in\mathcal{X}}\underline{p}(x)s_{v,x} \ge 0 \quad \forall v\in\mathcal{V}\\
& q_v + r_{v,x} + s_{v,x} = \sum_{b\in\mathcal{B}}\Delta_{v,x,b}w_b \quad \forall v\in\mathcal{H},\ x\in\mathcal{X}\\
& r_{v,x}\le 0, s_{v,x}\ge 0 \quad \forall v\in\mathcal{H},\ x\in\mathcal{X}.
\end{aligned}
$$

Proof. See Appendix A.1.7. □

2.5 Case Study

This section tests the proposed methodology using a hypothetical breast cancer data set. This data set includes 2108 voxels, including 1224 tumour voxels and 884 heart voxels, which is the organ at risk. We focus on uncertainty in the constraints and full volume criteria, as discussed in Section 2.4.1.

2.5.1 IMRT formulation

We consider the following RO problem for this data set,

$$\begin{aligned} \min_{w_b} \quad & \sum_{v\in\mathcal{H}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} \Delta_{v,x,b}p(x)w_b \\ \text{subject to:} \quad & \sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} \Delta_{v,x,b}\tilde{p}(x)w_b \geq \theta_v \quad \forall v \in \mathcal{T}, \tilde{p} \in \mathcal{U}_p, w_b \geq 0, b \in \mathcal{B}, \end{aligned} \tag{2.27}$$

which is the same as Model (2.14) with the objective of minimizing the total dose delivered to the heart under a nominal breathing pattern $\mathbf{p}$. The prescribed dose, to each tumour voxel is $\theta_v = 42.5$ for all tumour voxels. The RO solution delivers about a total dose of 14×10^6 GY to the tumour for a random PMF $\hat{\mathbf{p}}$, that is, $\sum_{v\in\mathcal{T}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} \Delta_{v,x,b}\hat{\mathbf{p}}(x)w_b^{RO}$ for a robust solution w^{RO}. This solution is clinically unacceptable. This causes severe side effects, such as skin burns.

To tackle this problem, we found a PRO solution that delivers a significantly lower dose to the tumour voxels and the same dose to the heart. We found this solution using the following model, which is based on Model (2.15) in Section 2.4.1.

$$
\begin{aligned}
\min_{(w_1,\ldots,w_{|\mathcal{B}|})} \quad & \sum_{v\in\mathcal{T}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} \Delta_{v,x,b}\hat{\mathbf{p}}(x)w_b \\
\text{subject to:} \quad & \sum_{v\in\mathcal{H}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} \Delta_{v,x,b}p(x)w_b \le Z^{RO}, \\
& \sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} \Delta_{v,x,b}p(x)w_b - \sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} \Delta_{v,x,b}\underline{p}(x)w_b + \\
& \sum_{x\in\mathcal{X}} \underline{p}(x)q_v - \sum_{x\in\mathcal{X}} r_{v,x} \ge \theta_v \quad \forall v\in\mathcal{T}, \\
& (\bar{p}(x)+\underline{p}(x))q_v - r_{v,x} \le \sum_{b\in\mathcal{B}} \Delta_{v,x,b}(\bar{p}(x)+\underline{p}(x))w_b \quad \forall v\in\mathcal{T}, \forall x\in\mathcal{X}, \\
& q_v \ge 0 \quad \forall v\in\mathcal{T}, \\
& r_{v,x} \ge 0 \quad \forall v\in\mathcal{T}, x\in\mathcal{X}
\end{aligned}
\tag{2.28}
$$

In the following section, we present the results for an RO solution and a PRO one and compare them with each other.

2.5.2 IMRT results

Using Model (2.28) we could find a PRO solution to Model (2.27). We computed the total dose delivered to each voxel for both RO and PRO solutions for the realization of uncertainty $\hat{\mathbf{p}}$. The following table compares the RO and PRO solutions for heart and tumour voxels.

Table 2.1: Comparing RO and PRO solutions for both heart and tumour voxels for a specific PMF

	RO solution		PRO solution	
Dose(GY)	tumour	Heart	tumour	Heart
Min	42.6	0	42.6	0
Average	12001	2.91	3494	2.91
Max	425079	49	165481	59
95% CI	(9411,14591)	(2.26,3.57)	(2747,4241)	(2.26,3.57)

The performance of both solutions is the same for the heart voxels. However, the PRO one delivers significantly less dose to the tumour voxels for the specific $\hat{\mathbf{p}}$. We also compared these solutions regarding tumour voxels. The average and maximum dose delivered to each tumour voxel by the PRO solution is much lower than the RO solution.

We also presented the box plot regarding the dose delivering to the tumour voxels. That is, for the amount of $\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}}\Delta_{v,x,b}\hat{\mathbf{p}}(x)w_b$ for all voxels. The blue box corresponds to the PRO solution, w^{PRO}, and the pink one corresponds to the RO solution, w^{RO}.

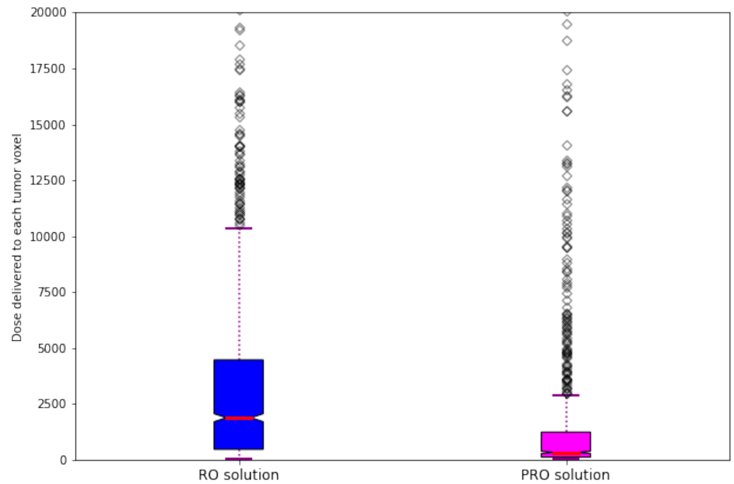


Figure 2.2: Comparison of box plot regarding dose delivered to the tumour voxels for a specific PMF

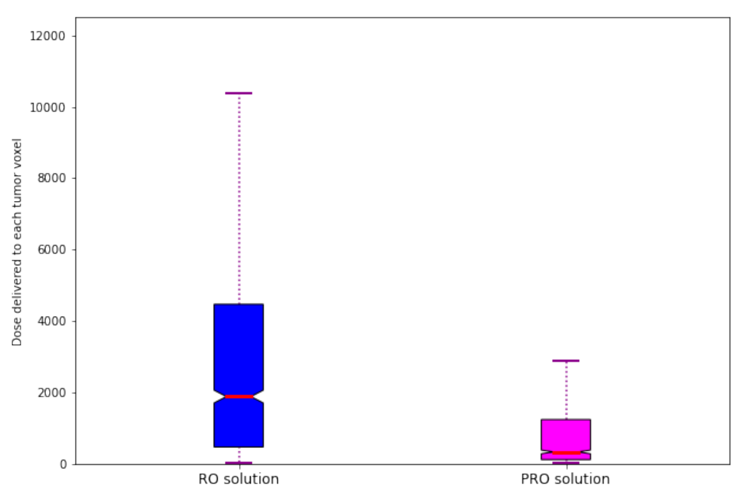


Figure 2.3: Comparison of box plot regarding dose delivered to the tumour voxels for a specific PMF without outlier

Figure 2.2 indicates that the PRO one delivers significantly less dose to the tumour voxels. Even the median of the box related to the PRO one is less than the box's minimum for the RO solution, and the maximum of the PRO box is less than the median of the RO one. To make it more clear, we presented Figure 2.3, that is, the boxes without outliers.

We also compared the solutions using DVH histograms. Figure 2.4 presents the DVH histogram for tumour voxels comparing PRO solution and RO solution for a specific PMF $\hat{\mathbf{p}}$. The red line indicates the prescribed dose 42.5.

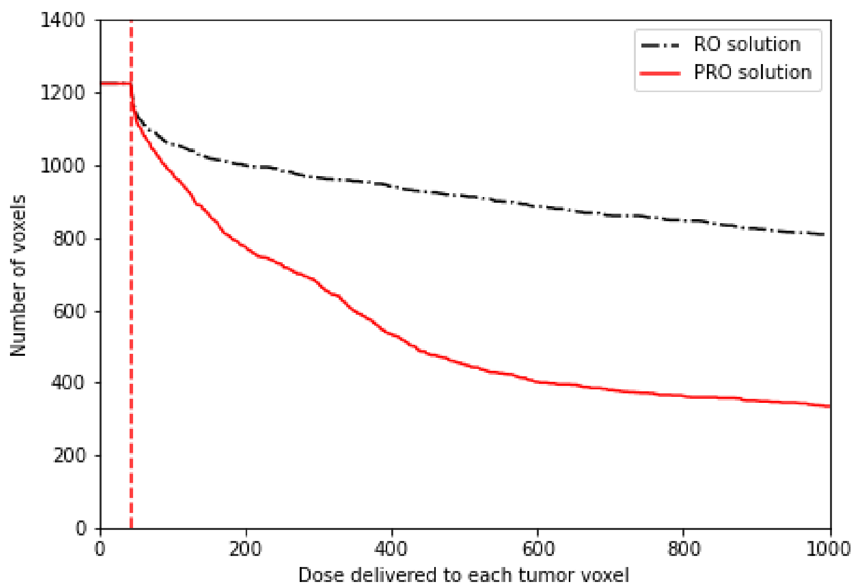


Figure 2.4: Comparison of the tumour dose volume histogram for RO and PRO solution for a specific PMF

In Figure 2.4, the x-axis is related to the dose delivered to the voxels, and the y-axis presents the number of voxels. It indicates that while both solutions deliver at least the prescribed dose to all tumour voxels, these voxels receive significantly more doses for the robust solution than the PRO solution. The number of voxels that receive at least a determined dose by the PRO solution is about half of the number of voxels that receive this dose by the RO solution.

We propose a DVH histogram for the tumour and heart voxels and at most 100 GY dose to clear the comparison.

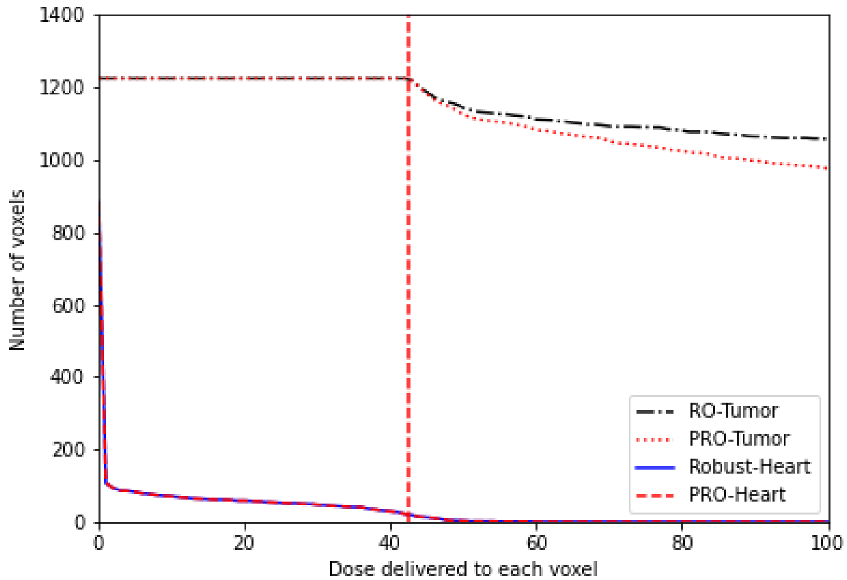


Figure 2.5: Comparison of the dose volume histogram for an RO and a PRO solution for a specific PMF

Figure 2.5 presents DVH for both tumour and heart voxels. The heart DVH histogram for both solutions is the same, and both solutions deliver the same dose to the heart voxels. However, the solutions' dose delivers to the tumour voxels is significantly different. Thus, while the PRO solution has the same performance regarding the worst-case scenario, it has a better performance for the non-worst-case $\hat{\mathbf{p}}$.

2.5.3 Extending the results for thirty random scenarios

So far, we only considered a specific PMF. Next, we examine the performance of the PRO solution for other realizations of uncertainty. We generated 30 random PMFs and compared the RO and PRO solutions for all these scenarios. We labelled these PMFs as $\{\mathbf{p_1}, \mathbf{p_2}, ..., \mathbf{p_{30}}\}$, computed the total dose delivered to each voxel for each $\mathbf{p_i}$ for $i \in \{1, ..., 30\}$, and calculated the average dose delivered to each voxel. That is, for each voxel, we computed $\frac{\sum_{i=1}^{30} \sum\limits_{b \in \mathcal{B}} \sum\limits_{x \in \mathcal{X}} \Delta_{v,x,b} \mathbf{p_i}(x) w_b}{30}$ for the RO and PRO solutions, where w_b is the intensity of beamlet b in each solution.

Figure 2.6 presents the DVH plot for average tumour voxels across all scenarios comparing RO and the PRO one. The red dash line presents the prescribed dose, that is 42.5. This figure indicates that, on average, the PRO solution delivers much lower total doses to the tumour.

We also compared the DVH histogram for the heart voxels for both solutions on average in Plot 2.7. This histogram indicates that the performances of the solutions are the same as each other.

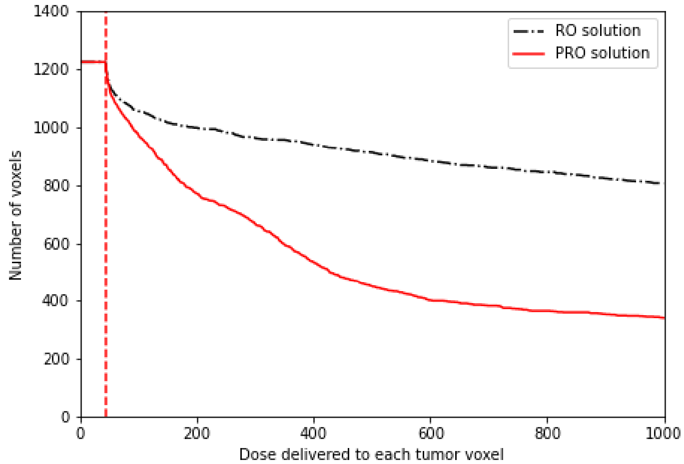


Figure 2.6: Comparison of the tumour dose volume histogram for a PRO and an RO solution for thirty random PMFs

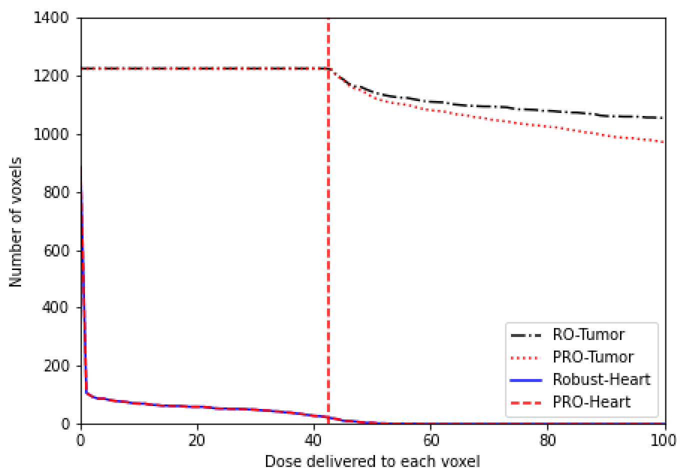


Figure 2.7: Comparison of the dose volume histogram for a PRO and an RO solution for thirty random PMFs

We compared the RO and PRO solutions for all thirty random **p**s. The following table compares the RO and PRO solutions in minimum, average and maximum dose delivered to each tumour voxel for all random uncertainty realizations. Each row presents the results for each **p**.

Table 2.2: Comparing the dose delivered to a tumour voxel for thirty random PMFs

Minimum dose (GY)		Average dose (GY)			Maximum dose (GY)	
RO	PRO	RO	PRO	Gap percentage	RO	PRO
42.6	42.6	12125	**3469**	71	426798	**164650**
42.6	42.6	11984	**3523**	70	424358	**156709**
42.6	42.6	11908	**3521**	70	423656	**155572**
42.6	42.6	11793	**3484**	70	422124	**163961**
42.6	42.6	11999	**3513**	71	425082	**155752**
42.5	42.5	11947	**3552**	70	424438	**156111**
42.5	42.5	11861	**3539**	70	423824	**141350**
42.5	42.5	12013	**3551**	70	424541	**162097**
42.5	42.5	11882	**3531**	70	423345	**157891**
42.6	42.6	11985	**3482**	71	425284	**164483**
42.5	42.5	11927	**3512**	70	424519	**160226**
42.6	42.6	12079	**3516**	71	425106	**165286**
42.5	42.5	11893	**3515**	70	423847	**160303**
42.5	42.5	11918	**3555**	70	423620	**153964**
42.5	42.5	12021	**3518**	71	424984	**161931**
42.5	42.5	11886	**3532**	70	424323	**150384**
42.5	42.5	11751	**3542**	69	421583	**152160**
42.6	42.6	11776	**3519**	70	421836	**156790**
42.5	42.5	11912	**3550**	70	423702	**151951**
42.5	42.5	12118	**3498**	71	425911	**166908**
42.6	42.6	11994	**3516**	71	424771	**154500**
42.5	42.5	11886	**3514**	70	424259	**155373**
42.5	42.5	11916	**3465**	71	424402	**166463**
42.5	42.5	11842	**3531**	70	423324	**154930**
42.6	42.6	12069	**3465**	69	425992	**170162**
42.5	42.5	11894	**3539**	70	423620	**158297**
42.6	42.6	12013	**3526**	71	424764	**156628**
42.6	42.6	11974	**3489**	71	424380	**163500**
42.6	42.6	11927	**3482**	71	424080	**166203**
42.5	42.5	11841	**3462**	71	423367	**167724**

This table indicates that the PRO solution delivers a much lower dose to the tumour voxels on average for all scenarios. The PRO one delivers about 70% GY lower than the RO solution on average to the tumour voxels. Also, the maximum dose the PRO solution delivers to tumour voxels is much lower than the RO solution. Therefore, the PRO one dominates the RO solution on average for all thirty PMFs.

We also solved this problem for four other samples. All samples have the same size. We found a PRO solution using the same $\hat{\mathbf{p}}$ for all samples. For each sample, we compared the dose delivered to a tumour voxel for the RO and PRO solutions and uncertainty realization $\hat{\mathbf{p}}$.

Table 2.3: Comparing the dose delivered to a tumour voxel for five samples and thirty random PMFs

	Minimum dose (GY)		Average dose (GY)			Maximum dose (GY)	
Sample	RO	PRO	RO	PRO	Gap percentage	RO	PRO
1	42.6	42.6	12001	**3494**	71	425079	**165481**
2	42.5	42.5	8316	**4370**	48	406751	**203626**
3	42.5	42.5	10015	**4248**	58	414023	**178268**
4	42.5	42.5	9826	**4379**	55	435941	**215239**
5	42.6	42.6	13786	**4315**	69	430374	**220647**

Table 2.3 indicates that for all samples the PRO solutions dominate the RO ones in average. The gap percentage determines that the PRO solutions deliver a much lower dose to the tumour voxels.

2.6 Conclusion

Pareto robust optimization presents an RO solution with less conservatism. A Pareto robust optimal solution is a robust optimal solution that any other RO solution cannot dominate.

It is important for the robust IMRT problem to find an optimal solution for all uncertainty realizations. We discussed four types of robust IMRT problems. For each type, we presented a method to find a PRO solution. We also demonstrated an application of our method for a hypothetical breast cancer data set. We presented a PRO solution for the IMRT problem with uncertainty in the tumour motion.

The PRO one delivered a much lower dose to the tumour voxels while delivering the prescribed dose to those voxels.

Chapter 3

Finding the Pareto Robust Frontier

3.1 Introduction

In this chapter, we propose an algorithm to generate the Pareto Robust frontier for a linear robust optimization problem. To do so, we first present an algorithm to generate alternative optimal solutions for a linear optimization problem with multiple optimal solutions. We then provide a method to determine if an RO solution is PRO. We also, define a "superior" PRO solution for a subset of uncertainty when the solution is optimal for a larger proportion realizations in that subset. Lastly, we present an algorithm to generate the corresponding subset of uncertainty for which a given PRO solution is superior. The rest of this chapter is organized as follows. In Section 3.2, we present the definition of a superior PRO solution for a given subset of uncertainty. Next, we present an algorithm to generate the Pareto robust frontier and demonstrate a procedure to find a superior PRO solution for a given subset of uncertainty in Section 3.3. In Section 3.4, we generalize the results for the robust problem with uncertainty in the constraint. At last, we present an application of our approach for a network structure and present a superior PRO solution for a given subset of uncertainty in Section 3.5.

3.2 Problem Definition

Robust optimization is used to determine optimal decisions for the worst-case realization of a parameter in an uncertainty set. This chapter considers the following robust problem.

$$\underset{\mathbf{x} \in \mathcal{X}}{\text{maximize}} \quad \min_{\mathbf{p} \in \mathcal{U}} \mathbf{p}'\mathbf{x}, \tag{3.1}$$

where $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ is the feasible region, $\mathcal{U} = \{\mathbf{p} \in \mathbb{R}^n : \mathbf{D}\mathbf{p} \geq \mathbf{d}\}$ is the uncertainty set of parameter $\mathbf{p}$, and parameters $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{D} \in \mathbb{R}^{l \times n}$, $\mathbf{b} \in \mathbb{R}^m$ and $\mathbf{d} \in \mathbb{R}^l$ are given. The solution to this problem is optimal for the worst-case realization of $\mathbf{p}$. However, non-worst-case realizations of $\mathbf{p}$ are not taken into account. If Model (3.1) has multiple optimal solutions, the behaviour of these solutions may differ for non-worst-case realizations of uncertainty. Pareto robust optimization [Iancu and Trichakis, 2014] aims to improve the non-worst-case behaviour of RO solutions.

Let $\mathcal{X}^{RO}$ be the set of all robust optimal solutions to Model (3.1) and $ri(\mathcal{U})$ be the relative interior of uncertainty set $\mathcal{U}$. It has been shown that each PRO solution is an optimal solution to the following problem for $\hat{\mathbf{p}} \in ri(\mathcal{U})$

$$\underset{\mathbf{x} \in \mathcal{X}^{RO}}{\text{maximize}} \quad \hat{\mathbf{p}}'\mathbf{x}. \tag{3.2}$$

A PRO solution is an RO solution that cannot be dominated in non-worst-case scenarios. Multiple PRO solutions can often be found by varying $\hat{\mathbf{p}}$ in formulation (3.2). Now, assume that a subset of the uncertain realizations are considered more important than the rest. The problem is to find a PRO solution that works best for the non-worst-case scenarios within this specific uncertainty subset. Consider a given discrete set with finite cardinality $\mathcal{V} \subset \mathcal{U}$. We denote a PRO solution as "superior" for $\mathcal{V}$ when it is optimal to Model (3.2) for the largest number of cases in the set $\mathcal{V}$, as defined below.

Definition 3.2.1. *A PRO solution is "superior" for discrete subset of uncertainty $\mathcal{V} \subset \mathcal{U}$ if it is the optimal solution of Model (3.2) for the largest number of elements in $\mathcal{V}$.*

For a given PRO solution $\hat{\mathbf{x}}$ define

$$\mathcal{V}^{\hat{\mathbf{x}}} = \{\mathbf{p} \in \mathcal{V} | \hat{\mathbf{x}} \in \arg\max_{\mathbf{x} \in \mathcal{X}^{RO}} \mathbf{p}'\mathbf{x}\}.$$

The cardinality of the set $\mathcal{V}^{\hat{\mathbf{x}}}$, denoted by $|\mathcal{V}^{\hat{\mathbf{x}}}|$, is the number of scenarios in $\mathcal{V}^{\hat{\mathbf{x}}}$. Since $|\mathcal{V}|$ is finite, $|\mathcal{V}^{\hat{\mathbf{x}}}|$ is also finite for all $\hat{\mathbf{x}} \in \mathcal{X}^{RO}$. That is, $|\mathcal{V}^{\hat{\mathbf{x}}}|$ is the number of $\mathbf{p} \in \mathcal{V}$ for which $\hat{\mathbf{x}}$ is an optimal solution to Model (3.2).

Consider two PRO solutions $\mathbf{x^1}$ and $\mathbf{x^2}$. We say $\mathbf{x^1}$ performs better than $\mathbf{x^2}$ for $\mathcal{V}$ if $|\mathcal{V}^{\mathbf{x^1}}| > |\mathcal{V}^{\mathbf{x^2}}|$. That is, $\mathbf{x^1}$ performs better than $\mathbf{x^2}$ if the number of $\mathbf{p} \in \mathcal{V}$ for which $\mathbf{x^1}$ is an optimal solution to Model (3.2) is larger than that of $\mathbf{p} \in \mathcal{V}$ for which $\mathbf{x^2}$ is an optimal solution to Model (3.2).

Our goal is to find a superior PRO solution $\hat{\mathbf{x}}$ for each subset $\mathcal{V}$. To do so, we suggest to generate the Pareto robust frontier, that is defined below, and compare Pareto solutions based on the number of elements $\mathbf{p}$ in $\mathcal{V}$ for which the solution $\hat{\mathbf{x}}$ is optimal.

Definition 3.2.2. *For any linear robust optimization problem, the Pareto robust frontier is the set of all PRO solutions.*

To find a superior PRO solution for a given subset of uncertainty, we need to generate the Pareto robust frontier and compare Pareto solutions based on the number of elements $\mathbf{p}$ in $\mathcal{V}$ for which the solution $\hat{\mathbf{x}}$ is optimal.

We next discuss how PRO solutions can be compared. Consider a Pareto robust optimal solution $\hat{\mathbf{x}}$. Let $\mathcal{U}_{\hat{\mathbf{x}}}$ be the set of all $\hat{\mathbf{p}} \in \mathcal{U}$ for which $\hat{\mathbf{x}}$ is an optimal solution for Model (3.2); that is,

$$\mathcal{U}_{\hat{\mathbf{x}}} = \{\mathbf{p} \in \mathcal{U} | \hat{\mathbf{x}} \in \arg \max_{\mathbf{x} \in \mathcal{X}^{RO}} \mathbf{p}'\mathbf{x}\}.$$

Therefore, for a given $\mathcal{V} \subseteq \mathcal{U}$, $\mathcal{V}^{\hat{\mathbf{x}}} = \mathcal{U}_{\hat{\mathbf{x}}} \cap \mathcal{V}$. To find $\mathbf{x}$ which is superior for set $\mathcal{V}$ among all Pareto robust optimal solutions, after generating the Pareto robust frontier, we compute $|\mathcal{U}_{\hat{\mathbf{x}}} \cap \mathcal{V}|$ for each PRO solution. Since $\mathcal{V}$ is a set with finite cardinality, the cardinality of $\mathcal{U}_{\hat{\mathbf{x}}} \cap \mathcal{V}$ is also finite. Then, we pick $\hat{\mathbf{x}}$, which is an optimal solution to

$$\max_{\mathbf{x} \in \mathcal{X}^{PRO}} |\mathcal{U}_{\mathbf{x}} \cap \mathcal{V}|,$$

where $\mathcal{X}^{PRO}$ is the set of all PRO solutions to Model (3.1). Therefore, we need to find $\mathcal{U}_{\hat{x}}$ for each PRO solution $\hat{\mathbf{x}}$.

In the next sections, we propose an approach to generate the Pareto robust frontier and also develop a method for finding the corresponding $\mathcal{U}_{\hat{\mathbf{x}}}$ for each Pareto robust optimal solution $\hat{\mathbf{x}}$.

3.3 Determining the Superior RPO Solution

This section presents an algorithm to generate the Pareto robust frontier for any robust optimization problem with uncertainty in the objective function. A robust problem has either multiple PRO solutions, or only one PRO solution. In this chapter, we consider the first case.

The idea behind our algorithm for generating the Pareto robust frontier is to search for robust optimal solutions one by one until there are no more unique PRO. First, we generate a PRO solution $\mathbf{x^0}$ using the method by Iancu and Trichakis [2014] given in equation (2.5). Then, we find a new RO solution that is not equal to the RO solutions found so far, at each iteration. Therefore, we need a search method to generate multiple optimal solutions. We present a method that results in alternative optimal solutions for any linear optimization problem in Section 3.3.1. Using this method, we will generate alternative RO solutions for Model (3.1). After finding a new RO solution at each iteration, we check if that solution is PRO. For this purpose, we present a method to determine if an RO solution is PRO in Section 3.3.2. Lastly, we construct an algorithm to generate the Pareto robust frontier for any robust optimization problem in Section 3.3.3.

After finding each Pareto solution $\hat{\mathbf{x}}$, we present another algorithm to generate the set $\mathcal{U}_{\hat{\mathbf{x}}}$ in Section 3.3.4.

3.3.1 Generating alternative optimal solutions for a linear problem

Consider the following linear optimization problem with a finite feasible region and multiple optimal solutions

$$\min_{\mathbf{x}} \quad \mathbf{c}'\mathbf{x} \tag{3.3}$$
$$\text{Subject to } \mathbf{Ax} \geq \mathbf{b}$$

where $\mathbf{x} \in \mathbb{R}^n$. Let $\mathcal{X}^{OPT}$ be the set of all optimal solutions to the Model (3.3), that is, $\mathcal{X}^{OPT} = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{Ax} \geq \mathbf{b}, \mathbf{c}'\mathbf{x} \leq w^{OPT}\}$ where w^{OPT} is the optimal objective value to Model (3.3). Suppose we find an optimal solution $\mathbf{x^0}$ to this problem. Let $\mathbf{x^1}$ be another element in $\mathcal{X}^{OPT}$. We define the distance between $\mathbf{x^0}$ and $\mathbf{x^1}$ using L_1-norm. That is, $||\mathbf{x^0} - \mathbf{x^1}||_1 = \sum_{i=1}^{n} |\mathbf{x^0}_i - \mathbf{x^1}_i|$. The idea behind our method is to find an optimal solution that has a maximum distance with a given set of optimal solutions. Specifically, if we have a given set $\mathcal{G} = \{\mathbf{x^0}, \mathbf{x^1}, ..., \mathbf{x^k}\} \subset \mathcal{X}^{OPT}$, the goal is to find the solution to

$$\max_{x \in \mathcal{X}^{OPT}} \min_{\mathbf{x^i} \in \mathcal{G}} ||\mathbf{x} - \mathbf{x^i}||_1.$$

First, we present a method to find an optimal solution $\mathbf{x^1}$ to Model (3.3) that has a maximum distance with a given optimal solution $\mathbf{x^0}$ in Theorem 3.3.1

Theorem 3.3.1. *Suppose $\mathbf{x^0}$ is an optimal solution to Model (3.3). The solution to the following problem is the farthest optimal solution from $\mathbf{x^0}$, where w^{OPT} is the optimal* *objective value for problem (3.3).*

$$\max_{\mathbf{x},\mathbf{l},\mathbf{z}} \sum_{i=1}^{n} \mathbf{l}_i \tag{3.4}$$

$$\textit{Subject to } \mathbf{A}\mathbf{x} \geq \mathbf{b} \tag{3.5}$$

$$\mathbf{c}'\mathbf{x} \leq w^{OPT} \tag{3.6}$$

$$\mathbf{x^0} - \mathbf{x} + \mathbf{l} \leq M\mathbf{z}$$

$$\mathbf{x} - \mathbf{x^0} + \mathbf{l} \leq M(1 - \mathbf{z})$$

$$\mathbf{l} \in \mathbb{R}^n_+, \mathbf{z} \in \{0,1\}^n, \mathbf{x} \in \mathbb{R}^n.$$

Proof. See Appendix A.2.1. □

Now, let $\mathbf{x^0}, \mathbf{x^1}, ..., \mathbf{x^{k-1}}$ be k optimal solutions to problem (3.3). We seek to find an optimal solution $\mathbf{x^k}$ that is far from previous solutions $\mathbf{x^0}, \mathbf{x^1}, ..., \mathbf{x^{k-1}}$:

$$\mathbf{x^k} \in \arg\max\{\min_{\mathbf{x}} ||\mathbf{x} - \mathbf{x^j}||_1, \quad \forall j \in \{0, ..., k-1\}\}$$

The following corollary presents such a solution.

Corollary 3.3.1. *An optimal solution of the following model has the maximum distance* *from all elements of* $\mathcal{X}' = \{\mathbf{x^0}, ..., \mathbf{x^{k-1}}\} \subset \mathcal{X}^{OPT}$.

$$\begin{aligned}
\max_{\mathbf{x},\mathbf{l},\mathbf{z},m} \quad & m && (3.7)\\
\textit{subject to:} \quad & \mathbf{Ax} \geq \mathbf{b} \\
& \mathbf{c}'\mathbf{x} \leq w^{OPT} \\
& \mathbf{x}^j - \mathbf{x} + \mathbf{l}^j \leq M\mathbf{z}^j \quad \textit{for } j = 0, ..., k-1 \\
& \mathbf{x} - \mathbf{x}^j + \mathbf{l}^j \leq M(1 - \mathbf{z}^j) \quad \textit{for } j = 0, ..., k-1 \\
& m \leq \sum_{i=1}^{n} \mathbf{l}_i^j \quad \textit{for } j = 0, ..., k-1 && (3.8)\\
& \mathbf{x} \in, \mathbb{R}^n, \mathbf{l^j} \in \mathbb{R}_+^n, \mathbf{z^j} \in \{0,1\}^n, m \in \mathbb{R}_+.
\end{aligned}$$

Proof. See Appendix A.2.2. □

As we mentioned in Section 2.2.1, we can write Model (3.1) as

$$\begin{aligned}
\max_{\mathbf{x}} \quad & \mathbf{y}'\mathbf{d} && (3.9)\\
\text{subject to:} \quad & \mathbf{D}'\mathbf{y} = \mathbf{x} \\
& \mathbf{Ax} \leq \mathbf{b} \\
& \mathbf{y} \in \mathbb{R}_+^{m_u}, \mathbf{x} \in \mathbb{R}^n.
\end{aligned}$$

Therefore, we can use Corollary 3.9 to find alternative optimal solutions for Model (3.1) at each iteration in our algorithm.

We next present a method to determine if an RO solution is PRO using inverse optimization. To do so, we first briefly discuss inverse optimization.

We know that an RO solution $\hat{\mathbf{x}}$ is a PRO solution to Model (3.1) if and only if there exists some $\mathbf{p}$ in $ri(\mathcal{U})$ where $\hat{\mathbf{x}}$ is an optimal solution to

$$\max_{\mathbf{x}\in\mathcal{X}^{RO}} \mathbf{p}'\mathbf{x}. \tag{3.10}$$

To check if $\hat{\mathbf{x}}$ is PRO, we can find $\mathbf{p}$ for which $\hat{\mathbf{x}}$ is an optimal solution to (3.10), and then evaluate if $\mathbf{p}$ can be in $ri(\mathcal{U})$ or not. We use inverse optimization to find $\mathbf{p}$.

Inverse optimization for linear programming was introduced by Ahuja and Orlin [2001]. Consider the following problem as a forward linear optimization (FO) problem for a specific $\mathbf{c} \in \mathbb{R}^n$,

$$\begin{aligned} \text{FO}(\mathbf{c}): \quad & \min_{\mathbf{x}} \mathbf{c}'\mathbf{x} \\ & \text{subject to: } \mathbf{A}\mathbf{x} \geq \mathbf{b}, \end{aligned} \tag{3.11}$$

where $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{b} \in \mathbb{R}^m$. Suppose that $\mathcal{X}$ is the set of all feasible solutions to Model (3.11), and $\mathcal{X}^{OPT}(\mathbf{c})$ is the set of all optimal solutions to model $FO(\mathbf{c})$. Let $\mathcal{X}^{OPT} = \bigcup_{\mathbf{c}} \mathcal{X}^{OPT}(\mathbf{c})$ for all $\mathbf{c} \neq 0$ is a non-empty set.

Now, consider $\mathbf{x}^* \in \mathcal{X}^{OPT}$ be a given optimal solution. The goal in inverse optimization is to find $\mathbf{c}$ for which $\mathbf{x}^*$ is an optimal solution to $FO(\mathbf{c})$. A basic form of inverse optimization

(IO) can be written as follows:

$$IO(\mathbf{x}^*): \quad \min_{\mathbf{c},\mathbf{y}} 0 \tag{3.12a}$$

$$\text{Subject to: } \mathbf{A}'\mathbf{y} = \mathbf{c} \tag{3.12b}$$

$$\mathbf{c}'\mathbf{x}^* = \mathbf{b}'\mathbf{y} \tag{3.12c}$$

$$||\mathbf{c}||_1 = 1 \tag{3.12d}$$

$$\mathbf{y} \geq 0. \tag{3.12e}$$

Constraints (3.12b) and (3.12e) apply dual feasibility and constraint (3.12c) applies strong duality. Constraint (3.12d) imposes a norm on $\mathbf{c}$ to avoid trivial solution.

3.3.2 Determining if an RO solution is PRO

Using inverse optimization, we can find $\hat{\mathbf{p}}$ for which $\mathbf{x}^*$ is a solution to Model (3.10). The following theorem presents a method to determine if an RO solution is PRO.

Theorem 3.3.2. *An RO solution $\mathbf{x}^*$ is PRO if and only if $m > 0$ at optimality in the* *following problem:*

$$\begin{aligned} \max_{\mathbf{w},\mathbf{v},\mathbf{s},\mathbf{p}} \quad & m && (3.13) \\ \textit{subject to:} \quad & \mathbf{bs} + z^{RO}w - \mathbf{px}^* = 0 \\ & \mathbf{s}'\mathbf{A} - \mathbf{v} - \mathbf{p} = 0 \\ & \mathbf{Dv} + \mathbf{wd} \geq \mathbf{0} \\ & \mathbf{Dp} - \mathbf{d} \geq m\mathbf{1} \\ & \mathbf{s} \geq \mathbf{0}, w \leq 0, \mathbf{s} \in \mathbb{R}^{m_x}, \mathbf{v} \in \mathbb{R}^n, \mathbf{p} \in \mathbb{R}^n, w \in \mathbb{R}, m \in \mathbb{R}^+, \end{aligned}$$

where vector $\mathbf{1}$ *is an all-ones vector.*

Proof. See Appendix A.2.3. □

In the following section, we use the elements discussed so far to develop an algorithm for finding the Pareto robust frontier.

3.3.3 Pareto robust frontier

The idea behind our algorithm is to search across the set $\mathcal{X}^{RO}$ and find distinct RO solutions. First, we find a PRO solution by solving Model (3.10) for some $\hat{\mathbf{p}} \in ri(\mathcal{U})$. Then, Corollary 3.3.1 presents an RO solution that has maximum distance from the RO solutions that have been found so far, at each iteration. After finding each RO solution, we check if that RO solution is PRO using Theorem 3.3.2.

The following theorem provides a stopping criteria for the algorithm, and Corollary provides proof of convergence.

Theorem 3.3.3. *If the solution m for Model (3.7) at any iteration is zero, then all RO solutions to Model (3.1) have been generated.*

Proof. See Appendix A.2.4. □

Corollary 3.3.2. *The proposed algorithm will converge to the set of all RO solutions to Model (3.1).*

Proof. See Appendix A.2.5. □

Therefore, the proposed algorithm can generate all PRO solutions. For computational purposes, we define a threshold ϵ_x for the distance between the new solution and the solutions that we already have and stop the algorithm when this distance is less than ϵ_x. That means, we stop the algorithm if

$$\max_{\mathbf{x} \in \mathcal{X}^{RO}} \min_{0 \leq j \leq k-1} ||\mathbf{x^j} - \mathbf{x}||_1 < \epsilon_x.$$

Algorithm 1 returns the Pareto Robust frontier for Model (3.1).

Algorithm 1 Generating Pareto Robust Frontier for Model (3.1)

Pick $\mathbf{p}' \in ri(\mathcal{U})$

Solve $\max_{\mathbf{x} \in \mathcal{X}^{RO}} \mathbf{p}'\mathbf{x}$ to find a PRO solution $\mathbf{x^0}$

Define $\mathcal{R} = \{\mathbf{x^0}\}$

Solve Model (3.4) for $\mathbf{x^0}$ and the robust problem (3.1)

Denote the solution by $\mathbf{x^1}$

Define $\mathcal{X}' = \{\mathbf{x^0}, \mathbf{x^1}\}$

Solve model (3.13) for $\mathbf{x^1}$

Define m^* equal to the optimal objective value of Model (3.13)

if m^* is positive **then**

  $\mathcal{R} = \mathcal{R} \cup \{\mathbf{x^1}\}$

end if

Set $d = ||\mathbf{x^0} - \mathbf{x^1}||_1$

while $d \geq \epsilon_x$ **do**

  Solve Model (3.7) for $\mathcal{X}'$ and the robust problem (3.1)

  Denote the solution by $\mathbf{x}^*$

  Set $d = \min_{\mathbf{x} \in \mathcal{X}'} ||\mathbf{x}^* - \mathbf{x}||_1$

  Set $\mathcal{X}' = \mathcal{X}' \cup \{\mathbf{x}^*\}$

  Solve (3.13) for $\mathbf{x}^*$

  Set m^* equal to the optimal objective value

  if m^* is positive **then**

    $\mathcal{R} = \mathcal{R} \cup \{\mathbf{x}^*\}$

  end if

end while

Return the set $\mathcal{R}$

3.3.4 Corresponding set of uncertainty for each PRO solution

For each PRO solution $\mathbf{x}^*$ we need to find $\mathcal{U}_{\mathbf{x}^*}$ which denotes the corresponding subset of $\mathcal{U}$ for which $\mathbf{x}^*$ is preferred. In this section, we present an algorithm that finds such subsets of $\mathcal{U}$. First, we present a model that presents the inverse optimization model for problem (3.10).

Proposition 3.3.1. *The following problem is an inverse optimization model for a specific RO solution* $\mathbf{x}^*$ *of Model* (3.10)*,*

$$\begin{aligned}
\min_{\mathbf{w},\mathbf{v},\mathbf{s},\mathbf{p}} \quad & 0 && (3.14)\\
& \mathbf{bs} + z^{RO}w - \mathbf{px}^* = 0 \\
& \mathbf{s}'\mathbf{A} - \mathbf{v} - \mathbf{p} = 0 \\
& \mathbf{Dv} + \mathbf{wd} \geq \mathbf{0} \\
& \mathbf{Dp} \geq \mathbf{d} \\
& \mathbf{s} \geq \mathbf{0}, w \leq 0, \mathbf{s} \in \mathbb{R}^{m_x}, \mathbf{v} \in \mathbb{R}^n, \mathbf{p} \in \mathbb{R}^n, w \in \mathbb{R}.
\end{aligned}$$

Proof. The proof of this proposition is provided in the proof of Theorem 3.3.2, in Appendix A.2.3. □

Using Model (3.14), we first find $\mathbf{p^1} \in \mathcal{U}$ such that $\mathbf{x}^*$ is an optimal solution to $\max_{\mathbf{x}\in\mathcal{X}^{RO}} \mathbf{p^1x}$. Next, we look for $\mathbf{p^2}$ that is the farthest $\mathbf{p}$ from $\mathbf{p^1}$ such that $\mathbf{x}^*$ is an optimal solution for $\max_{\mathbf{x}\in\mathcal{X}^{RO}} \mathbf{p^2x}$. That is, we seek for a new solution $\mathbf{p^2}$ that has a maximum distance $||\mathbf{p^2} - \mathbf{p^1}||_1$ from $\mathbf{p^1}$ using Theorem 3.3.1. Using Corollary 3.3.1, we can find different $\mathbf{p} \in \mathcal{U}_{\mathbf{x}^*}$ at each iteration.

The following theorem provides a stopping criteria for the algorithm, and Corollary provides proof of convergence.

Theorem 3.3.4. *If the algorithm returns zero as minimum distance between new* $\mathbf{p}$ *and the* $\mathbf{p}$*s that have already been found, then all* $\mathbf{p}$*s for which* $\mathbf{x}^*$ *is an optimal solution to Model* (3.10) *have been determined.*

Proof. The proof of this theorem is similar to the proof of Theorem 3.3.3, presented in Appendix A.2.4. □

Corollary 3.3.3. *The proposed algorithm will converge to the set of all* $\mathbf{p} \in \mathcal{U}$ *for which* $\mathbf{x}^*$ *is an optimal solution to model* $\max_{\mathbf{x} \in \mathcal{X}^{RO}} \mathbf{p}'\mathbf{x}$.

Proof. The proof of this corollary is similar to the proof of Corollary 3.3.2, presented in Appendix A.2.5. □

We define a threshold ϵ_u for the minimum distance between new $\mathbf{p}$ and the set of $\mathbf{p}$s that we already have, and stop the algorithm when this distance is less than ϵ_u. That is,

$$min_{\hat{\mathbf{p}} \in \mathcal{U}_x} ||\mathbf{p} - \hat{\mathbf{p}}||_1 < \epsilon_u.$$

Algorithm 2 returns $\mathcal{U}_{\mathbf{x}^*}$ for any PRO solution $\mathbf{x}^*$.

Algorithm 2 Corresponding Uncertainty Set for a PRO Solution to Model (3.1)

Solve model (3.14) to find $\mathbf{p^0}$ for the PRO solution $\mathbf{x}^*$

Solve Model (3.4) for $\mathbf{p^0}$ and linear problem (3.14)

Denote the solution by $\mathbf{p^1}$

Define $\mathcal{U}_{\mathbf{x}^*} = \{\mathbf{p^0}, \mathbf{p^1}\}$

Define $d = ||\mathbf{p^0} - \mathbf{p^1}||_1$

while $d \geq \epsilon_u$ **do**

Solve Model (3.7) for linear problem (3.14) and the set of optimal solutions $\mathcal{U}_{\mathbf{x}^*}$

Denote the solution by $\mathbf{p}^*$

Set $d = \min_{\mathbf{p} \in \mathcal{U}_{\mathbf{x}^*}} ||\mathbf{p}^* - \mathbf{p}||_1$

$\mathcal{U}_{\mathbf{x}^*} = \mathcal{U}_{\mathbf{x}^*} \cup \{\mathbf{p}^*\}$

end while

Return $\mathcal{U}_{\mathbf{x}^*}$

3.4 Generalization

The main problem that we discussed in this chapter was the robust optimization problem with uncertainty in the objective function, as shown in (3.1). In this section, we consider robust optimization problems with uncertainty in the constraints as follows

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c}'\mathbf{x} \\ \text{subject to: } & \mathbf{A}\mathbf{x} \geq \mathbf{b} \quad \forall \mathbf{A} \in \mathcal{U}_A, \end{aligned} \tag{3.15}$$

where $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$ and $\mathcal{U}_A \subset \mathbb{R}^{m \times n}$. Note that PRO solutions for this type of problems were discussed in Section 2.2.2. In this section, we present an algorithm to generate the Pareto robust frontier for this problem. Consider the expanded form of Model (3.15) as follows.

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{c}'\mathbf{x} \\ \text{subject to} \quad & \mathbf{a_i'x} \geq b_i \quad \forall \mathbf{a_i} \in \mathcal{U}_{a_i}, \ \text{for } i = 1, ..., m_x, \\ & \mathbf{x} \in \mathbb{R}^n \end{aligned} \tag{3.16}$$

where $b_i \in \mathbb{R}$, $\mathcal{U}_{a_i} = \{\mathbf{a_i} | \mathbf{D_i a_i} \geq \mathbf{d_i}\}$ for some $\mathbf{D_i} \in \mathbb{R}^{m_u \times n}$, $\mathbf{d_i} \in \mathbb{R}^{m_u}$, and each vector $\mathbf{a_i}$ is the $i-$th row of matrix $\mathbf{A}$. Let $f = \min_{\mathbf{p} \in \mathcal{U}} \mathbf{p}'\mathbf{x}$. Then, we can write Model (3.1) as follows

$$\begin{aligned} \underset{\mathbf{x} \in \mathcal{X}}{\text{maximize}} \quad & f \\ \text{subject to} \quad & \mathbf{p}'\mathbf{x} \geq f \quad \forall \mathbf{p} \in \mathcal{U} \end{aligned} \tag{3.17}$$

It can be seen that Model (3.16) is a general form of (3.17), and in turn, (3.1). Iancu and Trichakis [2014] defined Pareto robust optimal solutions for this type of problems. For each constraint $\mathbf{a_i x} \geq b_i$ and each robust optimal solution $\mathbf{x}$, define

$$s_i(\mathbf{x}, \mathbf{a_i}) = \mathbf{a_i'x} - b_i.$$

It can be seen that an RO solution $\hat{\mathbf{x}}$ is PRO if there does not exist any other RO solution $\mathbf{x}$ such that

$$\sum_{i=1}^{m_x} \mathbf{v}_i s_i(\mathbf{x}, \mathbf{a_i}) \geq \sum_{i=1}^{m_x} \mathbf{v}_i s_i(\hat{\mathbf{x}}, \mathbf{a_i}) \quad \forall(\mathbf{a_1}, ..., \mathbf{a_{m_x}}) \text{ s.t. } \mathbf{a_i} \in \mathcal{U}_{a_i}, \text{ for } 1 \leq i \leq m_x, \tag{3.18}$$
$$\sum_{i=1}^{m_x} \mathbf{v}_i s_i(\mathbf{x}, \bar{\mathbf{a}}_\mathbf{i}) > \sum_{i=1}^{m_x} \mathbf{v}_i s_i(\hat{\mathbf{x}}, \bar{\mathbf{a}}_\mathbf{i}) \quad \text{for some } (\bar{\mathbf{a}}_\mathbf{1}, ..., \bar{\mathbf{a}}_\mathbf{m_x}) \text{ s.t. } \bar{\mathbf{a}}_\mathbf{i} \in \mathcal{U}_{a_i}, \text{ for } 1 \leq i \leq m_x$$

for some slack value vector $\mathbf{v}$ that determines the relative value of the constraints. Define vector $\mathbf{p}$ such that

$$\mathbf{p}_j = \sum_{i=1}^{m_x} \mathbf{a_i}_j \mathbf{v}_i$$

and

$$\mathcal{U}^\mathbf{p} = \{\mathbf{p} \in \mathbb{R}^n | \mathbf{p}_j = \sum_{i=1}^{m_x} \mathbf{a_i}_j \mathbf{v}_i \text{ and } \mathbf{a_i} \in \mathcal{U}_{a_i} \quad \forall 1 \leq i \leq m_x\}.$$

Iancu and Trichakis [2014] also proved that an RO solution $\hat{\mathbf{x}}$ is a PRO solution if and only if there exists an element $\hat{\mathbf{p}} \in ri(\mathcal{U}_\mathbf{p})$ such that $\hat{\mathbf{x}}$ is an optimal solution to

$$\underset{\mathbf{x} \in \mathcal{X}^{RO}}{\text{maximize}} \quad \hat{\mathbf{p}}'\mathbf{x}, \tag{3.19}$$

where $\mathcal{X}^{RO}$ is the set of all robust optimal solutions to Model (3.16).

Consider a given sequence of sets $(\mathcal{W}_1, ..., \mathcal{W}_{m_x})$ where $\mathcal{W}_i$ is a discrete finite subset of $\mathcal{U}_{a_i}$. Consider the following set regarding this sequence:

$$\mathcal{W} = \{\mathbf{p} \in \mathbb{R}^n | \mathbf{p}_j = \sum_{i=1}^{m_x} \mathbf{a_i}_j \mathbf{v}_i \text{ and } \mathbf{a_i} \in \mathcal{W}_i \quad \forall 1 \leq i \leq m_x\}.$$

For each PRO solution $\hat{\mathbf{x}}$ define

$$\mathcal{W}^{\hat{\mathbf{x}}} = \{\mathbf{p} \in \mathcal{W} | \hat{\mathbf{x}} \in \arg\max_{\mathbf{x} \in \mathcal{X}^{RO}} \mathbf{p}'\mathbf{x}\}.$$

Consider two PRO solutions $\mathbf{x^1}$ and $\mathbf{x^2}$ to problem (3.16). We say $\mathbf{x^1}$ performs better than $\mathbf{x^2}$ for $\mathcal{W}$ if $|\mathcal{W}^{\mathbf{x^1}}| > |\mathcal{W}^{\mathbf{x^2}}|$. Our goal is to find a PRO solution which is superior for $\mathcal{W}$ among all PRO solutions. That is a PRO solution which is a solution to $\max_{\mathbf{x} \in \mathcal{X}^{PRO}} |\mathcal{W}^{\mathbf{x}}|$ where $\mathcal{X}^{PRO}$ is the set of all PRO solutions to Model (3.16). Therefore, we need to find all PRO solutions and for each PRO solution $\hat{\mathbf{x}}$ we need to find the set of all $\hat{\mathbf{p}} \in \mathcal{U}_{\mathbf{p}}$ for which $\hat{\mathbf{x}}$ is an optimal solution to Model (3.19). We define this set as

$$\mathcal{U}_{\hat{\mathbf{x}}} = \{\mathbf{p} \in \mathcal{U}^{\mathbf{p}} | \hat{\mathbf{x}} \in \arg\max_{\mathbf{x} \in \mathcal{X}^{RO}} \mathbf{p}'\mathbf{x}\}.$$

Suppose that we have the Pareto robust frontier of Model (3.16), that is the set of all PRO solutions, and $\mathcal{U}_{\hat{\mathbf{x}}}$ for each PRO solution $\hat{\mathbf{x}}$. For a given $\mathcal{W}$, the PRO solution $\hat{\mathbf{x}}$ is a superior PRO if $\hat{\mathbf{x}}$ has maximum $|\mathcal{U}_{\hat{\mathbf{x}}} \cup \mathcal{W}|$ among all $\mathbf{x} \in \mathcal{X}^{PRO}$. We present an algorithm to generate the Pareto robust frontier for Model (3.16). We also present an algorithm to generate the set $\mathcal{U}_{\hat{\mathbf{x}}}$ for each PRO solution $\hat{\mathbf{x}}$.

For both algorithms, we first need to write Model (3.16) as a linear programming model.

Consider the problem (3.16). We can write this model as

$$\min_{\mathbf{x}} \mathbf{c}'\mathbf{x} \tag{3.20}$$

$$s.t. \begin{pmatrix} \min_{\mathbf{a_i}} \mathbf{a_i'x} \\ \mathbf{D_i a_i} \geq \mathbf{d_i} \end{pmatrix} \geq b_i. \tag{3.21}$$

Using duality for each $1 \leq i \leq m_x$ we can write constraint (3.21) as

$$\begin{pmatrix} \max_{\mathbf{y_i}} \mathbf{y_i'd_i} \\ \mathbf{D_i'y_i} = \mathbf{x} \\ \mathbf{y_i} \geq \mathbf{0} \end{pmatrix} \geq b_i \tag{3.22}$$

Therefore, we can write Model (3.16) as follows

$$\begin{aligned} \underset{\mathbf{x}, \mathbf{y_1}, ..., \mathbf{y_{m_x}}}{\text{minimize}} \quad & \mathbf{c}'\mathbf{x} \\ \text{subject to} \quad & \mathbf{y_i'd_i} \geq b_i \quad \text{for } i = 1, ..., m_x, \\ & \mathbf{D_i'y_i} = \mathbf{x} \quad \text{for } i = 1, ..., m_x, \\ & \mathbf{y_i} \geq \mathbf{0}, \\ & \mathbf{x} \in \mathbb{R}^n, \\ & \mathbf{y_i} \in \mathbb{R}^{m_u} \end{aligned} \tag{3.23}$$

Thus, there exists a linear form of problem (3.16), and the set of all robust optimal solutions is

$$\mathcal{X}^{RO} = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{c}'\mathbf{x} \leq z^{RO} \quad \exists\ \mathbf{y_i} \geq \mathbf{0}, \quad \mathbf{y_i'd_i} \geq b_i, \quad \mathbf{D_i'y_i} = \mathbf{x}\}$$

where z^{RO} is the optimal objective value to Model (3.16). Now, we can use the methodology in Section 3.3.1 to search for multiple robust optimal solutions to Model (3.16). We showed before in this section, that an RO solution is PRO, if and only if it is an optimal solution to Model (3.19). Therefore, using a similar approach to the one discussed in Section 3.3.2, we can determine if an RO solution is PRO using inverse optimization. Consider Model (3.19). Using the definition of $\mathcal{X}^{RO}$, we can write this model as follows.

$$
\begin{aligned}
\underset{\mathbf{x}, \mathbf{y_1}, ..., \mathbf{y_{m_x}}}{\text{maximize}} \quad & \mathbf{p}'\mathbf{x} \\
\text{subject to} \quad & \mathbf{c}'\mathbf{x} \leq z^{RO}, \\
& \mathbf{y_i}'\mathbf{d_i} \geq b_i \quad \text{for } i = 1, ..., m_x, \\
& \mathbf{D_i}'\mathbf{y_i} = \mathbf{x} \quad \text{for } i = 1, ..., m_x, \\
& \mathbf{y_i} \geq \mathbf{0}, \\
& \mathbf{x} \in \mathbb{R}^n, \\
& \mathbf{y_i} \in \mathbb{R}^{m_u}
\end{aligned}
\tag{3.24}
$$

To write the inverse model of this problem we need its dual. The dual of this model is

$$
\begin{aligned}
\underset{g, q_1, ..., q_{m_x} \mathbf{f_1}, ..., \mathbf{f_{m_x}}}{\text{minimize}} \quad & z^{RO} g + \sum_{i=1}^{m_x} b_i q_i \\
\text{subject to} \quad & \mathbf{c}' g \sum_{i=1}^{m_x} \mathbf{f_i} = \mathbf{p}, \\
& q_i \mathbf{d_i} + \mathbf{D_i} \mathbf{f_i} \geq \mathbf{0} \quad \text{for } i = 1, ..., m_x, \\
& g \geq 0, \\
& q_i \leq 0 \quad \text{for } i = 1, ..., m_x, \\
& \mathbf{f_i} \in \mathbb{R}^n
\end{aligned}
\tag{3.25}
$$

Suppose that we have an optimal solution $\mathbf{x}^*$ to Model (3.24) for some $\mathbf{p} \in \mathcal{U}_{\mathbf{p}}$. Using the primal Model (3.24) and the dual Model (3.25), we can enforce optimality conditions on $\mathbf{x}^*$ and write the following inverse optimization problem that provides us with the intended

$\mathbf{p}$, which makes this specific $\mathbf{x}^*$ optimal for (3.24),

$$\max_{g, q_1, \ldots, q_{m_x} \mathbf{f_1}, \ldots, \mathbf{f_{m_x}}, \mathbf{p}} \quad 0 \tag{3.26}$$

$$\text{subject to: } z^{RO} g + \sum_{i=1}^{m_x} b_i q_i - \mathbf{p}\mathbf{x}^* = 0 \tag{3.27}$$

$$\mathbf{c}' g - \sum_{i=1}^{m_x} \mathbf{f_i} = \mathbf{p} \tag{3.28}$$

$$q_i \mathbf{d_i} + \mathbf{D_i} \mathbf{f_i} \geq \mathbf{0} \quad \text{ for } i = 1, \ldots, m_x \tag{3.29}$$

$$\mathbf{p} \in \mathcal{U}_{\mathbf{p}} \tag{3.30}$$

$$g \geq 0$$

$$q_i \leq 0 \quad \text{ for } i = 1, \ldots, m_x$$

$$\mathbf{f_i} \in \mathbb{R}^n.$$

Constraint (3.27) applies strong duality. Constraints (3.28) and (3.29) guarantee dual feasibility and constraint (3.30) guarantees that $\mathbf{p}$ is in $\mathcal{U}_{\mathbf{p}}$.

The goal is to determine if a robust optimal solution $\mathbf{x}^*$ is PRO. Therefore, we have to change constraint (3.30) to $\mathbf{p} \in ri(\mathcal{U}_\mathbf{p})$ as follows

$$\max_{g,q_1,\dots,q_{m_x}\mathbf{f_1},\dots,\mathbf{f_{m_x}},\mathbf{p}} \quad 0 \tag{3.31}$$

$$\text{Subject to: } z^{RO}g + \sum_{i=1}^{m_x} b_i q_i - \mathbf{p}\mathbf{x}^* = 0 \tag{3.32}$$

$$\mathbf{c}'g - \sum_{i=1}^{m_x} \mathbf{f_i} = \mathbf{p} \tag{3.33}$$

$$q_i\mathbf{d_i} + \mathbf{D_i}\mathbf{f_i} \geq \mathbf{0} \quad \text{for } i = 1, ..., m_x \tag{3.34}$$

$$\mathbf{p} \in ri(\mathcal{U}_\mathbf{p}) \tag{3.35}$$

$$g \geq 0$$

$$q_i \leq 0 \quad \text{for } i = 1, ..., m_x$$

$$\mathbf{f_i} \in \mathbb{R}^n.$$

If this model has a feasible solution, then $\mathbf{x}^*$ is PRO. Otherwise, it is not a PRO solution. Therefore, using a similar approach to the one discussed in Section 3.3.1, we can generate a new robust optimal solution to Model (3.16). Also, at each iteration, we can determine if the new solution is PRO using Model (3.31). Algorithm 3 returns the Pareto robust frontier for Model (3.16).

Algorithm 3 Generating Pareto Robust Frontier for Model (3.16)

Pick $\mathbf{p'} \in ri(\mathcal{U}_{\mathbf{p}})$

Solve $\max_{\mathbf{x}\in\mathcal{X}^{RO}} \mathbf{p'x}$ to find a PRO solution $\mathbf{x^0}$

Define $\mathcal{R} = \{\mathbf{x^0}\}$

Solve Model (3.4) for $\mathbf{x^0}$ and the robust problem (3.23)

Denote the solution by $\mathbf{x^1}$

Define $\mathcal{X}' = \{\mathbf{x^0}, \mathbf{x^1}\}$

Solve (3.31) for $\mathbf{x^1}$

if Model (3.31) has a solution **then**

    $\mathcal{R} = \mathcal{R} \cup \{\mathbf{x^1}\}$

end if

Set $d = ||\mathbf{x^0} - \mathbf{x^1}||_1$

while $d \geq \epsilon_x$ **do**

    Solve Model (3.7) for $\mathcal{X}'$ and the robust problem (3.23)

    Denote the solution by $\mathbf{x}^*$

    Set $d = \min_{\mathbf{x}\in\mathcal{X}'} ||\mathbf{x}^* - \mathbf{x}||_1$

    Set $\mathcal{X}' = \mathcal{X}' \cup \{\mathbf{x}^*\}$

    Solve (3.31) for $\mathbf{x}^*$

    if Model (3.31) has a solution **then**

        $\mathcal{R} = \mathcal{R} \cup \{\mathbf{x}^*\}$

    end if

end while

Return the set $\mathcal{R}$

Now, for each PRO solution $\mathbf{x}^*$ to Model (3.16) we have to find the set $\mathcal{U}_{\mathbf{x}^*}$. For this purpose we have the same idea that is presented in Section 3.3.4. The difference is in the inverse model which gives us the corresponding $\hat{\mathbf{p}}$ for any PRO solution $\mathbf{x}^*$ which is an optimal solution to $\max_{\mathbf{x}\in\mathcal{X}^{RO}} \hat{\mathbf{p}}'\mathbf{x}$. In this algorithm Model (3.26) returns such a $\hat{\mathbf{p}}$. Algorithm 4 returns the corresponding set $\mathcal{U}_{\mathbf{x}^*}$ for each PRO solution $\mathbf{x}^*$.

Algorithm 4 Finding the Subset of Uncertainty for a Specific PRO Solution to Model (3.16)

Solve Model (3.26) to find $\mathbf{p^0}$ for the PRO solution $\mathbf{x}^*$

Solve Model (3.4) for $\mathbf{p^0}$ and linear problem (3.26)

Denote the solution by $\mathbf{p^1}$

Define $\mathcal{U}_{\mathbf{x}^*} = \{\mathbf{p^0}, \mathbf{p^1}\}$

Define $d = ||\mathbf{p^0} - \mathbf{p^1}||_1$

while $d \geq \epsilon_u$ **do**

  Solve Model (3.7) for linear problem (3.26) and the set of optimal solutions $\mathcal{U}_{\mathbf{x}^*}$

  Denote the solution by $\mathbf{p}^*$

  Set $d = \min_{\mathbf{p}\in\mathcal{U}_{\mathbf{x}^*}} ||\mathbf{p}^* - \mathbf{p}||_1$

  $\mathcal{U}_{\mathbf{x}^*} = \mathcal{U}_{\mathbf{x}^*} \cup \{\mathbf{p}^*\}$

end while

Return $\mathcal{U}_{\mathbf{x}^*}$

3.5 Numerical Result for a Network Structure

The following network structure is the one that we discussed in Example 2.2.1. We present the Pareto robust frontier for this problem and find a superior PRO solution for a determined subset of uncertainty.

Example 3.5.1. Consider the following network structure which has been discussed in Example 2.2.1.

$$\begin{aligned}
\underset{\mathbf{x},\mathbf{a},\mathbf{b}}{\text{maximize}} \quad & \min_{\mathbf{f}\in\mathcal{U}} \mathbf{f}'\mathbf{x} \\
\text{subject to} \quad & x_1 = a_1, \\
& x_2 = a_2 + b_2, \\
& x_i = b_i \qquad \text{for } i = 3, ..., 12, \\
& a_0 + a_1 + a_2 = 1, \\
& b_2 + b_3 + ... + b_{12} = 1, \\
& a_i, b_i \geq 0.
\end{aligned} \tag{3.36}$$

As we mentioned, the optimal objective value for this problem is $\mathbf{f}^T\mathbf{x} = \frac{1}{10}$. We considered the set of all robust optimal solutions as $\mathcal{X}^{RO} = \{(\mathbf{a},\mathbf{b},\mathbf{x}) \in \mathcal{X}, \mathbf{x} \geq \frac{1}{10}\mathbf{1}\}$ where $\mathbf{a} = (a_0, a_1, a_2)$, $\mathbf{b} = (b_2, b_3, ..., b_{12})$ and $\mathbf{x} = (x_1, x_2, ..., x_{12})$. Given a randomly generated subset $\mathcal{V} \subset \mathcal{U}$ where $|\mathcal{V}| = 50$, the goal is to find a superior PRO solution for $\mathcal{V}$ Such that a robust optimal solution is PRO one and maximizes the net emergency transmission rate for the elements in $\mathcal{V}$.

Using Algorithm 1, we were able to generate 200 robust optimal solutions to problem (3.36) and 31 of them were PRO solutions captured in the set $\mathcal{X}^{PRO} = \{\mathbf{x^1}, ..., \mathbf{x^{31}}\}$.

We computed $|\mathcal{V}^{\mathbf{x^i}}|$, the cardinality of the set $\mathcal{V}^{\mathbf{x^i}}$ for each PRO solution $\mathbf{x^i} \in \mathcal{X}^{PRO}$ where

$$\mathcal{V}^{\mathbf{x^i}} = \{\mathbf{f} \in \mathcal{V} | \mathbf{x^i} \in \arg \max_{\mathbf{x} \in \mathcal{X}^{RO}} \mathbf{f}'\mathbf{x}\}.$$

In Figure 3.1, we present each PRO solution using a star and each element of $\mathcal{V}$ using a dot. $\mathbf{x^1}$, denoted by the red star, is an optimal solution to model

$$\begin{array}{ll} \text{maximize} & \mathbf{f}'\mathbf{x} \\ \mathbf{x} \in \mathcal{X}^{RO} & \end{array} \tag{3.37}$$

for 29 dots, including red and blue dots. $\mathbf{x^2}$, denoted by the green star, is an optimal solution to Model (3.37) for 23 dots, including green and blue dots. All other PRO solutions, denoted by blue stars, are optimal for two blue dots.

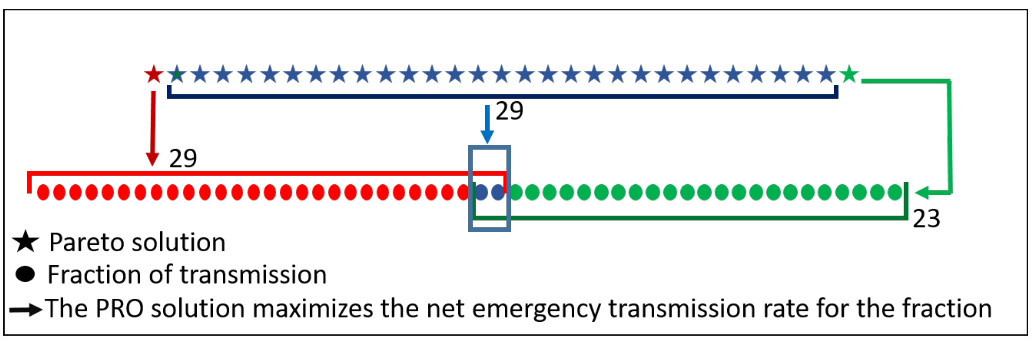


Figure 3.1: Illustration of the relationship between PRO solutions and elements of $\mathcal{V}$

Therefore,

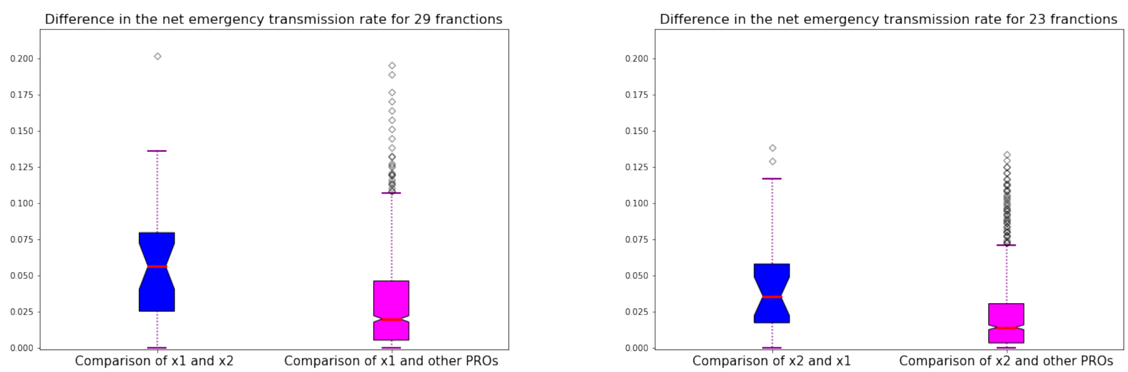
$$|\mathcal{V}^{\mathbf{x^1}}| = 29, \quad |\mathcal{V}^{\mathbf{x^2}}| = 23, \quad |\mathcal{V}^{\mathbf{x^i}}| = 2 \quad \forall 3 \leq i \leq 31.$$

Figure 3.2(a) presents the box plots for the difference in outcomes of $\mathbf{x^1}$ and other PRO solutions for 29 $\mathbf{f}s$ for which $\mathbf{x^1}$ is superior. The first (blue) box represents the data set

$\mathbf{f'x^1} - \mathbf{f'x^2}$ for all 29 $\mathbf{f}$s for which $\mathbf{x^1}$ is superior.

The second (pink) box represents the data set $\mathbf{f'x^1} - \mathbf{f'x^i}$ for all $3 \leq i \leq 31$ and 29 $\mathbf{f}$s. Both blue and pink boxes are non-negative, and it indicates that $\mathbf{x^1}$ has a better performance than other PRO solutions for these 29 elements in $\mathcal{V}$.

Figure 3.2(b) presents the the box plots for the difference in the outcomes of $\mathbf{x^2}$ and other PRO solutions for 23 elements in $\mathcal{V}$ that $\mathbf{x^2}$ is superior for. The blue box compares $\mathbf{x^2}$ and $\mathbf{x^1}$ that is about the data set $\mathbf{f'x^2} - \mathbf{f'x^1}$ for 23 $\mathbf{f}$s and the pink box compares $\mathbf{x^2}$ and all $\mathbf{x^i}$ for $3 \leq i \leq 31$ that is about the data set $\mathbf{f'x^2} - \mathbf{f'x^i}$ for all $3 \leq i \leq 31$. Both boxes are non-negative, and so the performance of $\mathbf{x^2}$ is better than the performance of other PRO solutions. Therefore, $\mathbf{x^1}$ is superior for $\mathcal{V}$.



(a) $\mathbf{f'x^1} - \mathbf{f'x^i}$ for $i = 2, 3, ..., 31$

(b) $\mathbf{f'x^2} - \mathbf{f'x^i}$ for $i = 1, 3, 4..., 31$

Figure 3.2: Comparison of performance of PROs for elements in $\mathcal{V}$

3.6 Conclusion

In this chapter, we proposed a method to generate alternative optimal solutions for any linear optimization problem with multiple solutions. Next, we presented a method to determine if an RO solution is PRO. Then, we discussed a method to generate the set of all PRO solutions, denoted as the Pareto robust frontier, for linear robust optimization problems. We also presented a method to generate a set of uncertainty realizations for which a given solution is optimal. Using this method, we proposed a procedure to compare PRO solutions and present a PRO solution, denoted as a superior PRO solution, optimal for the maximum number of elements in a given set of uncertainty realizations.

Chapter 4

Light Pareto Robust Optimal Solutions

4.1 Introduction

This chapter proposes "light PRO solutions" that accept a small optimality gap for the worst-case realization to improve non-worst-case solutions. We demonstrate the proposed method using the IMRT case study, introduces in Chapter 2. We observed empirically, that: (i) the worst-case scenario is often very unlikely to occur and (ii) giving up on the optimality of the worst-case solution may provide an opportunity to perform considerably better under non-worst-case scenarios. With this observation in mind, in this Chapter, we formally introduce light PRO solutions and discuss their theoretical properties.

In Section 4.2, we define light Pareto robust optimal solutions mathematically. Next, we discuss the theoretical properties of the improvement that light PRO solutions make for the non-worst-case scenarios, in Section 4.3. We define light PRO solutions for robust optimization problems with uncertainty in the constraints and analyze these solutions in Section 4.4. Lastly, we present an application of these solutions for an IMRT case study.

4.2 Definition

The robust optimal solution to the following problem is a solution that is optimal for the worst-case scenario of uncertainty.

$$\max_{\mathbf{x}\in\mathcal{X}} \min_{\mathbf{p}\in\mathcal{U}} \mathbf{p}'\mathbf{x}, \tag{4.1}$$

where $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ and $\mathcal{U} = \{\mathbf{p} \in \mathbb{R}^n : \mathbf{D}\mathbf{p} \geq \mathbf{d}\}$ for a given $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{D} \in \mathbb{R}^{l\times n}$, $\mathbf{b} \in \mathbb{R}^m$ and $\mathbf{d} \in \mathbb{R}^l$. To reduce the conservatism of RO solutions, PRO solutions can be used to improve the objective of the non-worst-case performance without deteriorating the worst-case performance.

As we mentioned in Section 2.2.1, the set of all robust optimal solutions to Model (4.1) is

$$\mathcal{X}^{RO} = \{\mathbf{x} \in \mathcal{X} : \exists \mathbf{y} \in \mathbb{R}_+^{m_u} \text{ such that } \mathbf{D}'\mathbf{y} = \mathbf{x}, \mathbf{y}'\mathbf{d} \geq z^{RO}\},$$

where z^{RO} is an optimal objective value to Model (4.1). To reduce the conservatism of the robust optimal solutions we let the z^{RO} be slightly worsened to $z^{RO} - \epsilon$. We refer to this ϵ parameter as the "worst-case loss".

Definition 4.2.1. *A solution* $\mathbf{x}$ *is a light RO solution if we worsen the value of the optimal solution for the worst-case scenario by* ϵ*. The set of all light RO solutions is,*

$$\mathcal{X}_{\epsilon}^{RE} = \{\mathbf{x} \in \mathcal{X} : \exists \mathbf{y} \in \mathbb{R}_{+}^{m_u} \text{ such that } \mathbf{D}'\mathbf{y} = \mathbf{x}, \mathbf{y}'\mathbf{d} \geq z^{RO} - \epsilon\}.$$

where $\epsilon \in \mathbb{R}_{+}$.

Definition 4.2.2. *A solution* $\hat{\mathbf{x}}$ *in* $\mathcal{X}_{\epsilon}^{RE}$ *is a light PRO solution if there is not any* $\mathbf{x} \in \mathcal{X}_{\epsilon}^{RE}$ *that dominates* $\hat{\mathbf{x}}$*, that is, there is not any* $\mathbf{x} \in \mathcal{X}_{\epsilon}^{RE}$ *such that*

$$\begin{aligned} &\mathbf{p}'\bar{\mathbf{x}} \geq \mathbf{p}'\mathbf{x}, \quad \forall \mathbf{p} \in \mathcal{U}, \text{ and} \\ &\bar{\mathbf{p}}'\bar{\mathbf{x}} > \bar{\mathbf{p}}'\mathbf{x}, \quad \text{for some } \bar{\mathbf{p}} \in \mathcal{U}. \end{aligned}$$

Using the same approach as Iancu and Trichakis [2014], we can find a light PRO solution, when considering $\mathcal{X}_{\epsilon}^{RE}$ instead of $\mathcal{X}^{RO}$.

Theorem 4.2.1. *For some* $\hat{\mathbf{p}} \in ri(\mathcal{U})$*, the solution to the following problem is a light PRO solution.*

$$\underset{\mathbf{x} \in \mathcal{X}_{\epsilon}^{RE}}{\text{maximize}} \quad \hat{\mathbf{p}}'\mathbf{x}. \tag{4.2}$$

Proof. See Appendix A.3.1. □

We will discuss the theoretical properties of the improvement that light PRO solutions make for non-worst-case scenarios in the following section.

4.3 Non-Worst-Case Gain

So far, we have defined the concept of light PRO solutions and hypothesized that slightly worsening the RO objective for worst-case may improve the objective for the non-worst-case scenarios. This section formalizes this gain as a function of ϵ and analyzes and discusses its theoretical properties.

Definition 4.3.1. *For a worst-case loss of ϵ, the gain function for non-worst-case scenario $\hat{\mathbf{p}}$ is defined as*

$$f(\epsilon; \hat{\mathbf{p}}) = \max_{\mathbf{x} \in \mathcal{X}_\epsilon^{RE}} \hat{\mathbf{p}}'\mathbf{x} - \max_{\mathbf{x} \in \mathcal{X}^{RO}} \hat{\mathbf{p}}'\mathbf{x}.$$

The gain function $f(\epsilon; \hat{\mathbf{p}})$ determines the improvement of the objective function of the RO model, for the non-worst-case $\hat{\mathbf{p}}$ where we deteriorate $\mathbf{p}'\mathbf{x}$ for the worst-case scenario by ϵ.

Proposition 4.3.1. *$f(\epsilon; \hat{\mathbf{p}})$ is an increasing function of ϵ.*

Proof. See Appendix A.3.2. □

Using the concept of global dependence of the optimal cost on the right hand side of a constraint, presented by Bertsimas and Tsitsiklis [1997], $f(\epsilon; \hat{\mathbf{p}})$ is a concave function. Therefore, either this function is strictly increasing or at some point it becomes a constant function, that is, there exists $\epsilon^* > 0$ such that for any $\epsilon_1 < \epsilon_2 < \epsilon^*$, $f(\epsilon_1; \hat{\mathbf{p}}) < f(\epsilon_2; \hat{\mathbf{p}}) < f(\epsilon^*; \hat{\mathbf{p}})$ and for any $\epsilon > \epsilon^*$, $f(\epsilon; \hat{\mathbf{p}}) = f(\epsilon^*; \hat{\mathbf{p}})$. We choose ϵ^* such that for any $\epsilon > \epsilon^*$, the fraction of $\frac{f(\epsilon;\hat{\mathbf{p}}) - f(\epsilon^*;\hat{\mathbf{p}})}{\epsilon - \epsilon^*}$ is less than a determined threshold. The target ϵ can be the lowest such ϵ^*.

4.4 Generalization

Now, consider the following form of a robust optimization problem with uncertainty in constraints, as discussed in Section 3.4,

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{c}'\mathbf{x} \\ \text{subject to} \quad & \mathbf{a_i}'\mathbf{x} \geq b_i \quad \forall \mathbf{a_i} \in \mathcal{U}_{a_i}, \text{ for } i = 1, ..., m_x, \\ & \mathbf{x} \in \mathbb{R}^n, \end{aligned} \tag{4.3}$$

where $b_i \in \mathbb{R}$ and $\mathcal{U}_{a_i} = \{\mathbf{a_i} | \mathbf{D_i a_i} \geq \mathbf{d_i}\}$ for some $\mathbf{D_i} \in \mathbb{R}^{m_u \times n}$ and $\mathbf{d_i} \in \mathbb{R}^{m_u}$. Recall that the set of robust optimal solutions to this problem is

$$\mathcal{X}^{RO} = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{c}'\mathbf{x} \leq z^{RO} \quad \exists\ \mathbf{y_i} \geq \mathbf{0}, \quad \mathbf{y_i}'\mathbf{d_i} \geq b_i, \quad \mathbf{D_i}'\mathbf{y_i} = \mathbf{x}\}.$$

A light PRO solution would then belong to:

$$\mathcal{X}^{RE} = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{c}'\mathbf{x} \leq z^{RO} + \epsilon \quad \exists\ \mathbf{y_i} \geq \mathbf{0}, \quad \mathbf{y_i}'\mathbf{d_i} \geq b_i, \quad \mathbf{D_i}'\mathbf{y_i} = \mathbf{x}\}.$$

Because $\mathcal{X}^{RO}$ is a subset of the set $\mathcal{X}^{RE}$, for any objective function, a problem with feasible region $\mathcal{X}^{RE}$ has the same or better optimal solution for the problem with feasible region $\mathcal{X}^{RO}$.

Now, we consider the elements of this set as our feasible solutions for the set of all robust solutions. We define a PRO solution in $\mathcal{X}^{RE}$, for the new feasible region as a solution that cannot be dominated by any other solution in $\mathcal{X}^{RE}$. That is, a solution $\hat{\mathbf{x}} \in \mathcal{X}^{RE}$ is a light

PRO solution if there is not any $\mathbf{x} \in \mathcal{X}^{RE}$ where,

$$\sum_{i=1}^{m_x} \mathbf{v}_i s_i(\mathbf{x}, \mathbf{a_i}) \leq \sum_{i=1}^{m_x} \mathbf{v}_i s_i(\hat{\mathbf{x}}, \mathbf{a_i}) \quad \forall(\mathbf{a_1}, ..., \mathbf{a_{m_x}}) \text{ s.t. } \mathbf{a_i} \in \mathcal{U}_{a_i}, \text{ for } 1 \leq i \leq m_x,$$
$$\sum_{i=1}^{m_x} \mathbf{v}_i s_i(\mathbf{x}, \bar{\mathbf{a}}_\mathbf{i}) < \sum_{i=1}^{m_x} \mathbf{v}_i s_i(\hat{\mathbf{x}}, \bar{\mathbf{a}}_\mathbf{i}) \quad \text{for some } (\bar{\mathbf{a}}_\mathbf{1}, ..., \bar{\mathbf{a}}_\mathbf{m_x}) \text{ s.t. } \bar{\mathbf{a}}_\mathbf{i} \in \mathcal{U}_{a_i}, \text{ for } 1 \leq i \leq m_x.$$

for $s_i(\mathbf{x}, \mathbf{a_i}) = \mathbf{a_i'x} - b_i$, and a value vector $\mathbf{v}$.

Theorem 4.4.1. *A solution in $\mathcal{X}^{RE}$ is a light Pareto robust one if it is an optimal solution to the following problem for some $\hat{\mathbf{p}} \in ri(\mathcal{U}_\mathbf{p})$*

$$\underset{\mathbf{x} \in \mathcal{X}^{RE}}{\text{minimize}} \quad \hat{\mathbf{p}}'\mathbf{x}. \tag{4.4}$$

Now, the question is about the improvement that right PRO solutions make for non-worst-case scenarios. The following definition formalize this improvement.

Definition 4.4.1. *For a worst-case loss of ϵ, the gain function for non-wosrt-case scenario $\hat{\mathbf{p}}$ is defined as*

$$f(\epsilon; \hat{\mathbf{p}}) = \min_{\mathbf{x} \in \mathcal{X}^{RO}} \hat{\mathbf{p}}'\mathbf{x} - \min_{\mathbf{x} \in \mathcal{X}^{RE}} \hat{\mathbf{p}}'\mathbf{x}.$$

Using same method in Section 4.3, we have the following proposition.

Proposition 4.4.1. *$f(\epsilon; \hat{\mathbf{p}})$ is an increasing function of ϵ.*

Proof. The proof of this theorem is similar to the proof of Proposition 4.3.1. □

Also, the function $f(\epsilon; \hat{\mathbf{p}})$ is concave. Therefore, it is enough to pick the lowest ϵ^* such that for any $\epsilon > \epsilon^*$, the function of $\frac{f(\epsilon;\hat{\mathbf{p}}) - f(\epsilon^*;\hat{\mathbf{p}})}{\epsilon - \epsilon^*}$ is less than a given threshold.

4.5 Case Study

Consider the problem of IMRT treatment planning for the breast cancer with the heart as the organ at risk. Recall the following problem parameters. The set of all beamlets is denoted by $\mathcal{B}$. Also, we denote the tumour by the voxel set $\mathcal{T}$ and the healthy tissues by the voxel set $\mathcal{H}$. We denote the intensity of beamlet b by w_b. Consider $D_{v,b}$ as the dose that voxel v receives from beamlet b, and θ_v as the prescribed dose for $v \in \mathcal{T}$.

Let the goal of IMRT be minimizing the total dose received by the heart voxels while meeting dose constraints on the tumour voxels. The robust IMRT problem with uncertainty in the tumour position can be written as follows [Bortfeld et al., 2008].

$$\begin{aligned} \min_{w_b} \quad & \sum_{v\in\mathcal{H}} \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b} p(x) w_b \\ \text{subject to:} & \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b} \tilde{p}(x) w_b \geq \theta_v \quad \forall v \in \mathcal{T}, \tilde{p} \in \mathcal{U}_p, w_b \geq 0, b \in \mathcal{B}, \end{aligned} \tag{4.5}$$

where θ_v is the prescribed dose to the tumour voxel, v. The uncertainty set

$$\mathcal{U}_p = \{\tilde{\mathbf{p}} \in \mathbb{R}^{|\mathcal{X}|} : \tilde{\mathbf{p}}(x) \in [\mathbf{p}(x) - \underline{\mathbf{p}}(x), \mathbf{p}(x) + \bar{\mathbf{p}}(x)] \quad \forall x \in \mathcal{X}; \sum_{x\in\mathcal{X}} \tilde{\mathbf{p}}(x) = 1\} \tag{4.6}$$

is the set of all motion PMFs where $\underline{\mathbf{p}}$ and $\bar{\mathbf{p}}$ are the bounds for the difference between the actual and nominal PMF $\mathbf{p}$ during treatment. In our case study $\theta_v = 42.5$ for all $v \in \mathcal{T}$.

A PRO solution $\hat{\mathbf{w}}$ is a solution for which there is not any RO solution $\mathbf{w}$ such that

$$\sum_{v\in\mathcal{T}}(\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}}\Delta_{v,x,b}\mathbf{p}(x)w_b-\theta_v)\leq\sum_{v\in\mathcal{T}}(\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}}\Delta_{v,x,b}\mathbf{p}(x)\hat{w}_b-\theta_v) \qquad \forall\mathbf{p}\in\mathcal{U}_p,$$

$$\sum_{v\in\mathcal{T}}(\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}}\Delta_{v,x,b}\hat{\mathbf{p}}(x)w_b-\theta_v)<\sum_{v\in\mathcal{T}}(\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}}\Delta_{v,x,b}\hat{\mathbf{p}}(x)\hat{w}_b-\theta_v) \qquad \exists\hat{\mathbf{p}}\in\mathcal{U}_p,$$

Using Model (2.28), we can find a PRO solution for a given $\hat{\mathbf{p}}$. It is evident that the PRO solutions reported in Chapter 2 are overly conservative and provide a very high dose to the tumour, as seen in Figures 2.4 and 2.6. Therefore, we generated light PRO solutions for this problem. Using the approach discussed in this chapter, we relaxed the worst-case dose by a loss parameter ϵ, and calculated the gain for different non-worst-case scenario. That is, we allowed an increase in the total dose delivered to the heart voxels for the worst-case breathing realization to decrease the total dose delivered to the tumour for non-worst-case scenarios. We solve the following model for different non-worst-case loss values of ϵ.

$$\min_{(w_1,\ldots,w_{|\mathcal{B}|})} \sum_{v\in\mathcal{T}} \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b}\hat{\mathbf{p}}(x) w_b \tag{4.7}$$

subject to:

$$\sum_{v\in\mathcal{H}} \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b} p(x) w_b \leq Z^{RO} + \epsilon,$$

$$\sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b} p(x) w_b - \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b} \underline{p}(x) w_b +$$

$$\sum_{x\in\mathcal{X}} \underline{p}(x) q_v - \sum_{x\in\mathcal{X}} r_{v,x} \geq \theta_v \quad \forall v \in \mathcal{T},$$

$$(\bar{p}(x) + \underline{p}(x)) q_v - r_{v,x} \leq \sum_{b\in\mathcal{B}} \Delta_{v,x,b} (\bar{p}(x) + \underline{p}(x)) w_b \quad \forall v \in \mathcal{T}, \forall x \in \mathcal{X},$$

$$q_v \geq 0 \quad \forall v \in \mathcal{T},$$

$$r_{v,x} \geq 0 \quad \forall v \in \mathcal{T}, x \in \mathcal{X}.$$

We ran this model for $\epsilon = 0.1, 0.2, 0.3, 0.4$ and 0.5 for a random $\hat{\mathbf{p}} \in ri(\mathcal{U}_p)$, and denoted the solutions by $\mathbf{w^{0.1}}$, $\mathbf{w^{0.2}}$, $\mathbf{w^{0.3}}$, $\mathbf{w^{0.4}}$ and $\mathbf{w^{0.5}}$ respectively. The results were considerably better than the PRO solutions with $\epsilon = 0$, denoted as $\mathbf{w^{Pr}}$.

Let $\mathcal{W}^s = \{\mathbf{w^{Pr}}, \mathbf{w^{0.1}}, \mathbf{w^{0.2}}, \mathbf{w^{0.3}}, \mathbf{w^{0.4}}, \mathbf{w^{0.5}}\}$. Figure 4.1 presents the box plots for the total dose delivered to each tumour voxel for the specific uncertainty realization $\hat{\mathbf{p}}$, that is, $\sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} \Delta_{v,x,b}\hat{\mathbf{p}}(x) w_b$ for $v \in \mathcal{T}$ and $w \in \mathcal{W}^s$. We magnified the boxes for $\epsilon = 0.1, 0.2, 0.3, 0.4$ and 0.5 to clarify the differences. It can be seen that the dose delivered to the tumour voxels decreases by increasing ϵ.

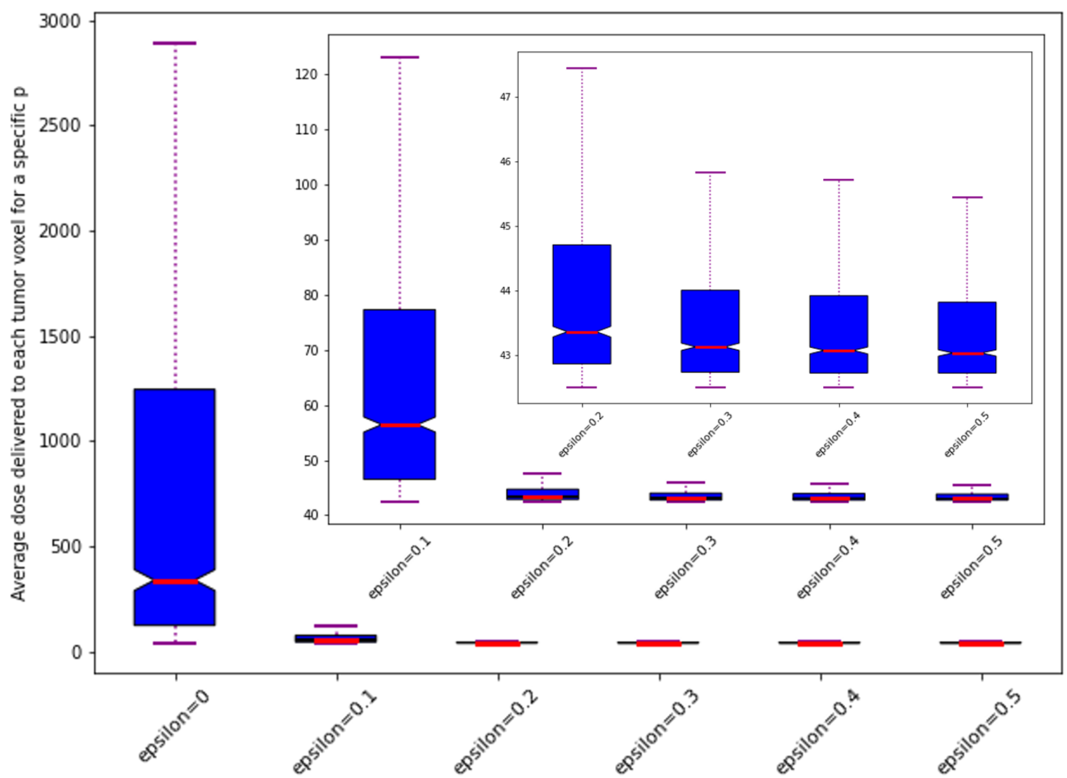


Figure 4.1: Box plots for the dose delivered to each tumour voxel for a specific uncertainty realization

We also extend our results for all realizations of the uncertainty. For this purpose, we generated a sample of $\mathcal{U}_p$ by randomly picking thirty $\mathbf{p}$s in $\mathcal{U}_p$ and denote the set of all these $\mathbf{p}$s by $\mathcal{Q}$. For each solution in $\mathcal{W}^s$ the average dose delivered to each tumour voxel for elements of $\mathcal{Q}$ is computed. That is, for each $\mathbf{w} \in \mathcal{W}^s$ and $v \in \mathcal{T}$ we computed

$$\frac{\sum\limits_{\mathbf{p}\in\mathcal{Q}} (\sum\limits_{b\in\mathcal{B}} \sum\limits_{x\in\mathcal{X}} \Delta_{v,x,b}\mathbf{p}(x)\mathbf{w}_b)}{30}. \tag{4.8}$$

Figure 4.2 presents the box plot of the data set generated by (4.8) for each $\mathbf{w} \in \mathcal{W}^s$ and $v \in \mathcal{T}$. It is clear that when we add 0.1 dose on average to each heart voxel, we significantly deliver less dose to the tumour voxels. We see this result for $\mathbf{w^{0.2}}$ and $\mathbf{w^{0.3}}$. However, $\mathbf{w^{0.4}}$ and $\mathbf{w^{0.5}}$ do not have a significant improvement. That is, the boxes related to this solutions are almost similar to the box plot related to $\mathbf{w^{0.3}}$, and the best result is for $\mathbf{w^{0.3}}$. We presented the boxes for all six element of $\mathcal{W}^s$, and magnified the boxes related to the solutions for $\epsilon = 0.1, 0.2, 0.3, 0.4, 0.5$ to clarify the differences.

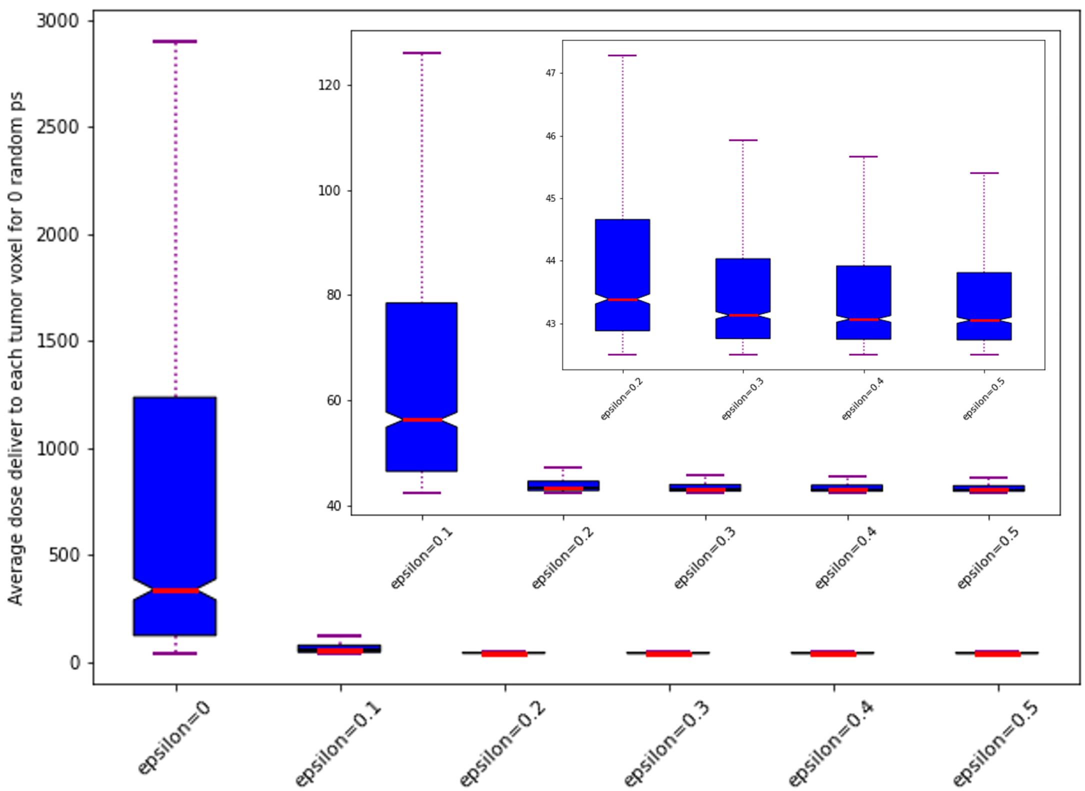


Figure 4.2: Box plots for the average dose delivered to each tumour voxel for 30 uncertainty realizations

We also compared the DVH plot for the average dose delivered to the tumour voxels for the elements in $\mathcal{W}^s$ in Figure 4.3. The red dash line presents the prescribed dose. All voxels receive the prescribed dose. This figure indicates that the average dose delivered to the tumour voxels decreases remarkably when we increase the average dose delivered to each heart voxel. Figure 4.4 compares the DVH plot for the average dose delivered to the heart voxels. This figure implies that the solutions in $\mathcal{W}^s$ do not have significant differences in delivering dose to the heart voxels, and all histograms are almost the same as others.

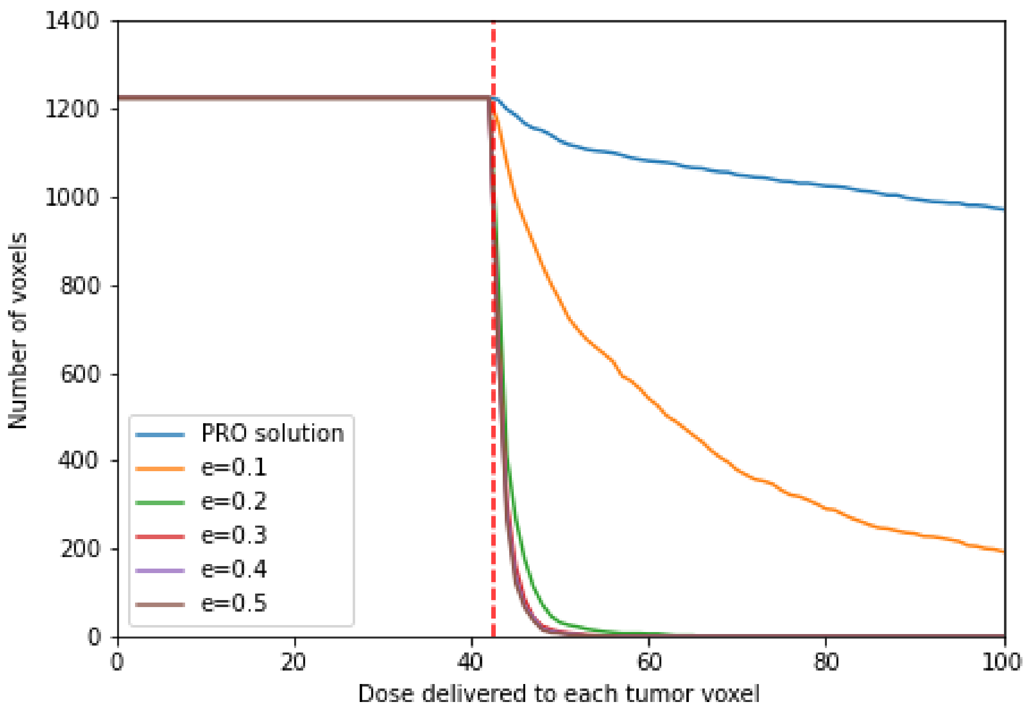


Figure 4.3: Comparison of dose volume histogram for the tumour voxels for solutions in $\mathcal{W}^s$

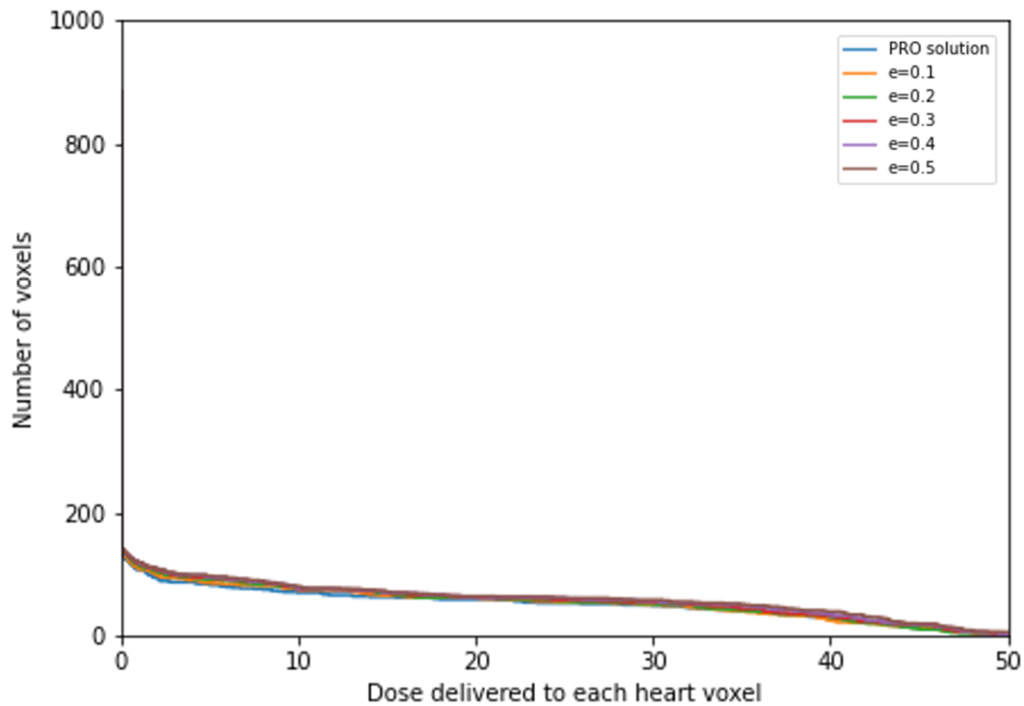


Figure 4.4: Comparison of dose volume histogram for the heart voxels for solutions in $\mathcal{W}^s$

We also generated fifty hypothetical data sets for the IMRT problem for breast cancer. The size of the data sets is different, and the total number of voxels for each sample is in the range (2108, 5058). For each data set, we generated a PRO solution. We also found light PRO solutions by adding $\epsilon = 0.1, 0.2, 0.3, 0.4, 0.5$ to the average dose delivered to each heart voxel for each sample. These solutions delivered less dose to the tumour voxels than the original PRO solution on average for all the fifty samples.

We added the results of all fifty samples together and presented the average, min, max, and the confidence interval for 95 percent of the dose delivered to each tumour voxels for different PRO solutions corresponding to different ϵ in the following table.

Epsilon added to average heart voxel	Min	Average	Max	95% CI
$\epsilon = 0$	42.5	3554	265002	(3452, 3655)
$\epsilon = 0.1$	42.5	58	21588	(57.12, 58.73)
$\epsilon = 0.2$	42.5	45	132	(44.75, 44.79)
$\epsilon = 0.3$	42.5	44.2	80.8	(44.19, 44.22)
$\epsilon = 0.4$	42.5	44	84.3	(44.04, 44.07)
$\epsilon = 0.5$	42.5	44	82.6	(43.98, 44)

Table 4.1: Comparing the dose delivered to each tumour voxel for fifty samples

Comparing the solutions, the difference between the original PRO solution with the one by adding 0.1 GY dose to the average heart voxels is considerable. Also, the solution will improved by adding 0.2 and 0.3 GY to the heart's average dose. However, the difference between the solutions for $\epsilon = 0.3, 0.4$, and 0.5 is negligible, and we can consider the light PRO solutions by adding 0.3 GY to the average heart dose as the best solutions for these fifty samples of the IMRT problem.

4.6 Conclusion

In this chapter, we defined new solutions by violating PRO solutions from optimal solutions for the worst-case scenario, denoted as light PRO solutions. We discussed and analyzed the improvement of the optimal solution for the non-worst-cases, that light PRO solutions make. We illustrated an application of light PRO solutions to breast cancer IMRT treatment planning. The results indicate that the light PRO solutions considerably decrease the side effects of overdosing.

Chapter 5

Future Research: Pareto Robust Inverse Optimization

In this chapter, we explore a possible extension of this work to the field of robust inverse optimization for the IMRT problem. In Section 5.1, we briefly introduce the problem and literature. We discuss the results from literature in robust inverse optimization in Section 5.2. We will present a robust inverse optimization model for the IMRT problem in Section 5.3. Lastly, we will discuss Pareto robust optimization for the robust inverse IMRT problem as the future research direction in Section 5.4.

5.1 Introduction

Considering that there are multiple or noisy observations for an optimization problem, the goal is to find an objective function that would make all observations "near-optimal"

Bertsimas et al. [2015],Keshavarz et al. [2011]. Recently, Ghobadi et al. [2018] combined the inverse optimization approach with robust optimization and considered uncertainty around observations. The robust inverse optimization then finds parameters that would render the worst-case scenario of the optimal observation. This method has not been applied to IMRT treatment planning to the best of our knowledge.

One of the ideas that the planners use for radiation therapy is to consider a given set of clinically “good” treatment plans and find the objective function that makes these solutions optimal by using inverse optimization. We use the approach presented by Ghobadi et al. [2018] to model the IMRT problem as a robust inverse optimization problem.

5.2 Robust Inverse Optimization

Ghobadi et al. [2018] used a robust approach to present an inverse optimization problem with multiple observations as a robust inverse optimization problem. Consider the following optimization problem,

$$\begin{aligned} \mathbf{FO}(\mathbf{c}) : \min_{\mathbf{x}} \mathbf{c}'\mathbf{x} \\ \text{subject to: } \mathbf{A}\mathbf{x} \geq \mathbf{b}, \end{aligned} \tag{5.1}$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{m\times n}$, and $\mathbf{b} \in \mathbb{R}^m$. They proved that the inverse optimization of this problem is as follows,

$$\begin{aligned} \min_{\mathbf{c},\mathbf{y},\mathbf{x}} \quad & d(\mathbf{x}, \mathbf{x^0}) && (5.2) \\ \text{subject to:} \quad & \mathbf{A}'\mathbf{y} = \mathbf{c}, \\ & \mathbf{c}'\mathbf{x} = \mathbf{b}'\mathbf{y}, \\ & \mathbf{A}\mathbf{x} \geq \mathbf{b}, \\ & \mathbf{y} \geq 0, \end{aligned}$$

where $d(\mathbf{x}, \mathbf{x^0})$ is the distance function between $\mathbf{x}$ and a given solution $\mathbf{x^0}$, and $\mathbf{y} \in \mathbb{R}^m$. They proved that we could consider the multiple data points as an uncertainty set that encapsulates all possible realizations of the input data, and they present Problem (5.2) with multiple observations as a robust inverse optimization problem as follows.

$$\begin{aligned} \min_{\mathbf{c},\mathbf{y},\mathbf{x}} \max_{\mathbf{x^0}\in\mathcal{U}} \quad & d(\mathbf{x}, \mathbf{x^0}) && (5.3) \\ \text{subject to:} \quad & \mathbf{A}'\mathbf{y} = \mathbf{c}, \\ & \mathbf{c}'\mathbf{x} = \mathbf{b}'\mathbf{y}, \\ & \mathbf{A}\mathbf{x} \geq \mathbf{b}, \\ & \mathbf{y} \geq 0. \end{aligned}$$

We can present the inverse IMRT problem with multiple "good" observations in the following section using this approach.

5.3 Robust Inverse Optimization Form of the IMRT Problem

To follow the concept of Section 5.2, we first model the IMRT problem as an inverse optimization problem. Suppose that a "good" treatment exists that is considered an optimal solution to an IMRT problem. The goal is to define an objective function to have an optimal solution equal to the observation. Consider the general formulation of the IMRT problem as

$$\begin{aligned} \min_{w_b} \quad & \sum_{v \in \mathcal{V}} t_v \sum_{x \in \mathcal{X}} \sum_{b \in \mathcal{B}} \Delta_{v,x,b} p(x) w_b \\ \text{subject to:} \quad & \sum_{x \in \mathcal{X}} \sum_{b \in \mathcal{B}} \Delta_{v,x,b} p(x) w_b \geq \theta_v \quad \forall v \in \mathcal{T}, \end{aligned} \tag{5.4}$$

where t is the value vector of the voxels. Now, define

$$z_v = \sum_{x \in \mathcal{X}} \sum_{b \in \mathcal{B}} \Delta_{v,x,b} p(x) w_b,$$

the problem can be written as

$$\begin{aligned} \min_{z_v} \quad & \sum_{v \in \mathcal{V}} t_v z_v \\ \text{subject to:} \quad & z_v \geq \theta_v \quad \forall v \in \mathcal{T}. \end{aligned} \tag{5.5}$$

Suppose that there exists a good observation w^0. Then, we define an inverse optimization problem that give us a value vector such that the result of model (5.4) be w^0. First, define

$z^0 \in \mathbb{R}^{|\mathcal{V}|}$ such that $z_v^0 = \sum_{x\in\mathcal{X}} \sum_{b\in\mathcal{B}} \Delta_{v,x,b} p(x) w_b^0$ for each $v \in \mathcal{V}$.

Let the first $|\mathcal{T}|$ elements of $\mathcal{V}$ be the voxels of the tumour $\mathcal{T}$, and the last $|\mathcal{H}|$ elements are the healthy voxels from $\mathcal{H}$. Define the vector α as

$$\alpha_{\mathbf{i}} = \begin{cases} \theta_i & \text{if } i \in \mathcal{T} \\ 0 & \text{otherwise.} \end{cases} \tag{5.6}$$

By using Formulation (5.2), we can write the following inverse optimization model for this problem.

$$\begin{aligned} \min_{\mathbf{z},\mathbf{t},\mathbf{q}} \quad & d(\mathbf{z}, \mathbf{z^0}) \\ \text{subject to:} \quad & \mathbf{A}'\mathbf{q} = \mathbf{t}, \\ & \mathbf{t}'\mathbf{z} = \alpha'\mathbf{q}, \\ & \mathbf{A}'\mathbf{z} \geq \alpha, \\ & \mathbf{q} \geq 0, \end{aligned} \tag{5.7}$$

where the matrix $\mathbf{A}$ is

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}_{|\mathcal{T}|\times|\mathcal{V}|} \\ \mathbf{0}_{|\mathcal{H}|\times|\mathcal{V}|} \end{pmatrix} \tag{5.8}$$

such that $\mathbf{I}$ is the identity matrix.

There are some cases in radiation therapy in which the health care researchers have multiple “good” treatments, and their goal is to find an optimization problem for which all of the observations are “near-optimal”.

To present a model which results in good observations, we can use robust inverse optimization. Using this approach, we can present the inverse IMRT problem with multiple “good” observations as follows:

$$\begin{aligned} \min_{\mathbf{z},\mathbf{t},\mathbf{q}} \max_{\hat{z}\in\mathcal{U}} \quad & d(\mathbf{z},\hat{\mathbf{z}}) \\ \text{subject to:} \quad & \mathbf{A}'\mathbf{q} = \mathbf{t}, \\ & \mathbf{t}'\mathbf{z} = \alpha\mathbf{q}, \\ & \mathbf{A}\mathbf{z} \geq \alpha, \\ & \mathbf{q} \geq 0. \end{aligned} \tag{5.9}$$

5.4 PRO Solutions of the Robust Inverse IMRT Problem

Since most voxels confirm the prescription dose while a small fraction may deviate markedly from prescription, the good treatment $\mathbf{z}$ is sparse. Therefore, the robust problem (5.9) would have lots of RO solutions. A question is which solution we should choose. We suggest applying Pareto robust optimization for this problem.

Since Model (5.9) is a nonlinear optimization problem, the idea for our future work is to reformulate this model as a linear robust optimization problem. Then, we can find Pareto solutions for this model. We also can generate the Pareto robust frontier for the new linear form of the robust inverse IMRT problem.

Chapter 6

Conclusions

Robust Optimization (RO) is a sub-field of optimization with deterministic uncertainty. This field is sometimes criticized for producing overly conservative solutions. A robust solution is optimal for worst-case scenarios and may not be optimal for non-worst-cases. Several methods have been introduced to reduce this conservatism, such as light robustness and globalized robust optimization. All of these methods focus on the worst-case scenario, however, the optimal solution may be dominated by another solution for non-worst-case scenarios. Pareto robust optimization (PRO) presents a solution that is optimal for the worst-case scenario and cannot be improved by other solutions for a non-worse-case scenario without deteriorating the outcome for another non-worst-case scenario.

In this thesis, we applied the concept of PRO to the intensity-modulated radiation therapy (IMRT) problem. We focused on two types of delivering radiation dose to the tissues, full volume or partial volume criteria. For each type, we considered uncertainty in either the tumour's motion or the organs at risk's motion. Therefore, we discussed four types of the

IMRT problem. We applied PRO to these four IMRT problems and presented a method for finding a PRO solution for each of these problems.

Next, we applied our method to several hypothetical data sets for the breast cancer IMRT problem with the goal of decreasing the side effects of delivering overdose to the tumour. We demonstrated a PRO and an RO solution. We compared the dose that the PRO and RO solutions deliver to the tumour and OAR tissues for thirty random uncertainty realizations. Both PRO and RO solutions delivered the same dose to the OAR tissues while delivering the prescribed dose to the tumour. However, the PRO solution decreased the overdosing to the tumour (for about 70%).

Although Pareto robust optimization presents a solution that another RO solution cannot dominate, the existence of another RO solution that is optimal for more non-worst-case scenarios cannot be guaranteed. Also, a PRO solution is still conservative in the worst-case scenario. To tackle these problems, we proposed two methods.

In the first method, we defined the superior PRO solutions for a given subset of uncertainty. A PRO solution has a better performance than another one for an uncertainty set if it is an optimal solution for more uncertainty realizations in that set. A PRO solution is superior for a given uncertainty set if no PRO solution has a better performance. We presented a method to find a superior PRO solution for a predetermined subset of the uncertainty set with finite cardinality.

To find such a solution, we suggested generating the set of all PRO solutions, denoted as the Pareto robust frontier, and then comparing the performance of the PRO solution for the given subset of uncertainty. We presented an algorithm to generate the Pareto robust frontier for any linear robust optimization problem. For this purpose, we presented a method for generating alternative optimal solutions for any linear optimization problem with multiple solutions. We also presented a method to determine if an RO solution is PRO.

To compare PRO solutions, we need to know the set of uncertainty realizations for which a given PRO solution is optimal. We demonstrated an algorithm to generate such a set of uncertainty for a given PRO solution.

The second method that we suggested to reduce the conservatism of PRO solutions is violating the PRO solution from optimal solutions for the worst-case scenario. We denoted these new solutions as light Pareto robust optimal solutions. We demonstrated that these solutions improve the optimality for the non-worse-case realizations and cannot be dominated by another solution for all uncertainty realizations. We discussed the practical result of this approach in the breast cancer IMRT problem for fifty hypothetical data sets. We showed that these solutions significantly decrease the side effects of overdosing. Light PRO solutions deliver at least the prescribed dose to the tumour tissues. However, they do not deliver a high dose volume to these tissues. It can decrease the side effects of overdosing, such as skin burns. Note that the increased dose that these solutions deliver to the OAR is negligible.

We believe that our approach for presenting superior PRO solutions for a subset of uncertainty can be applied to more real-world robust optimization problems. One of these problems is the robust inverse IMRT problem. We discussed this problem in Chapter 5 and presented a non-linear model for it. We suggest reformulating this model as a linear RO problem and presenting a method to find a superior PRO solution for a given subset of uncertainty. We also suggest presenting light PRO solutions for this problem.

Another future research direction is to apply our approach for generating light PRO solutions to more real-world robust optimization problems, such as inventory management and communication problems.

References

Pilar Abad, Sonia Benito, and Carmen López. A comprehensive review of value at risk methodologies. *The Spanish Review of Financial Economics*, 12(1):15–32, 2014.

Sara Ahmadian, Umang Bhaskar, Laura Sanità, and Chaitanya Swamy. Algorithms for inverse optimization problems. In *26th Annual European Symposium on Algorithms (ESA 2018)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

Ravindra K Ahuja and James B Orlin. Inverse optimization. *Operations Research*, 49(5): 771–783, 2001.

Ergun Ahunbay and X Allen Li. Investigation of the reliability, accuracy, and efficiency of gated imrt delivery with a commercial linear accelerator. *Medical Physics*, 34(7): 2928–2938, 2007.

Yu An, Jianming Liang, Steven E Schild, Martin Bues, and Wei Liu. Robust treatment planning with conditional value at risk chance constraints in intensity-modulated proton therapy. *Medical Physics*, 44(1):28–36, 2017.

Turgay Ayer. Inverse optimization for assessing emerging technologies in breast cancer screening. *Annals of Operations Research*, 230(1):57–85, 2015.

Ana María Barragán-Montero, Dan Nguyen, Weiguo Lu, Mu-Han Lin, Roya Norouzi-Kandalan, Xavier Geets, Edmond Sterpin, and Steve Jiang. Three-dimensional dose prediction for lung imrt patients with deep neural networks: robust learning from heterogeneous beam configurations. *Medical Physics*, 46(8):3679–3691, 2019.

Christoph Baum, Markus Alber, Mattias Birkner, and Fridtjof Nüsslin. Robust treatment planning for intensity modulated radiotherapy of prostate cancer based on coverage probabilities. *Radiotherapy and Oncology*, 78(1):27–35, 2006.

Aharon Ben-Tal, Stephen Boyd, and Arkadi Nemirovski. Extending scope of robust optimization: Comprehensive robust counterparts of uncertain problems. *Mathematical Programming*, 107(1):63–89, 2006.

Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*, volume 28. Princeton University Press, 2009.

Aharon Ben-Tal, Ruud Brekelmans, Dick Den Hertog, and Jean-Philippe Vial. Globalized robust optimization for nonlinear uncertain inequalities. *INFORMS Journal on Computing*, 29(2):350–366, 2017.

Dimitris Bertsimas and Melvyn Sim. The price of robustness. *Operations research*, 52(1):35–53, 2004.

Dimitris Bertsimas and John N Tsitsiklis. *Introduction to linear optimization*, volume 6. Athena Scientific Belmont, MA, 1997.

Dimitris Bertsimas, Vishal Gupta, and Ioannis Ch Paschalidis. Inverse optimization: A new perspective on the black-litterman model. *Operations research*, 60(6):1389–1403, 2012.

Dimitris Bertsimas, Vishal Gupta, and Ioannis Ch Paschalidis. Data-driven estimation in equilibrium using inverse optimization. *Mathematical Programming*, 153(2):595–633, 2015.

Daniel Bienstock and Nuri ÖZbay. Computing robust basestock levels. *Discrete Optimization*, 5(2):389–414, 2008.

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

Marco Botte and Anita Schöbel. Dominance for multi-objective robust optimization concepts. *European Journal of Operational Research*, 273(2):430–440, 2019.

Anders Brahme, J-E Roos, and Ingemar Lax. Solution of an integral equation encountered in rotation therapy. *Physics in Medicine & Biology*, 27(10):1221, 1982.

Valentina Cacchiani and Paolo Toth. Nominal and robust train timetabling problems. *European Journal of Operational Research*, 219(3):727–737, 2012.

Timothy CY Chan, Thomas Bortfeld, and John N Tsitsiklis. A robust approach to imrt optimization. *Physics in Medicine & Biology*, 51(10):2567, 2006.

Timothy CY Chan, Houra Mahmoudzadeh, and Thomas G Purdie. A robust-cvar optimization approach with application to breast cancer therapy. *European Journal of Operational Research*, 238(3):876–885, 2014.

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

Milad Dehghani Filabadi and Houra Mahmoudzadeh. Effective budget of uncertainty for classes of robust optimization. *INFORMS Journal on Optimization*, 2022.

Darrell Duffie and Jun Pan. An overview of value at risk. *Journal of Derivatives*, 4(3): 7–49, 1997.

Iain Robert Dunning. *Advances in robust and adaptive optimization: algorithms, software, and insights*. PhD thesis, Massachusetts Institute of Technology, 2016.

Zeynep Erkin, Matthew D Bailey, Lisa M Maillart, Andrew J Schaefer, and Mark S Roberts. Eliciting patients' revealed preferences: an inverse markov decision process approach. *Decision Analysis*, 7(4):358–365, 2010.

Apostolos Fertis. *A robust optimization approach to statistical estimation problems by Apostolos G. Fertis*. PhD thesis, Massachusetts Institute of Technology, 2009.

Matteo Fischetti and Michele Monaci. Light robustness. In *Robust and online large-scale optimization*, pages 61–84. Springer, 2009.

Virginie Gabrel, Cécile Murat, and Aurélie Thiele. Recent advances in robust optimization: An overview. *European Journal of Operational Research*, 235(3):471–483, 2014.

Laurent El Ghaoui, Maksim Oks, and Francois Oustry. Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research*, 51(4):543–556, 2003.

Kimia Ghobadi, Taewoo Lee, Houra Mahmoudzadeh, and Daria Terekhov. Robust inverse optimization. *Operations Research Letters*, 46(3):339–344, 2018.

Marc Goerigk and Anita Schöbel. Algorithm engineering in robust optimization. In *Algorithm engineering*, pages 245–279. Springer, 2016.

Bram L Gorissen, İhsan Yanıkoğlu, and Dick den Hertog. A practical guide to robust optimization. *Omega*, 53:124–137, 2015.

Bram Leendert Gorissen et al. *Practical robust optimization techniques and improved inverse planning of HDR brachytherapy.* CentER, Tilburg University, 2014.

Dan A Iancu and Nikolaos Trichakis. Pareto efficiency in robust optimization. *Management Science*, 60(1):130–147, 2014.

Runqing Jiang. Impact of geometric uncertainties on dose calculations for intensity modulated radiation therapy of prostate cancer. 2008.

Arezou Keshavarz, Yang Wang, and Stephen Boyd. Imputing a convex objective function. In *2011 IEEE International Symposium on Intelligent Control*, pages 613–619. IEEE, 2011.

Jourdain B Lamperski and Andrew J Schaefer. A polyhedral characterization of the inverse-feasible region of a mixed-integer program. *Operations Research Letters*, 43(6):575–578, 2015.

Yanzhu Lin, Kai Chen, Zhiyuan Lu, Lei Zhao, Yalan Tao, Yi Ouyang, and Xinping Cao. Intensity-modulated radiation therapy for definitive treatment of cervical cancer: a meta-analysis. *Radiation Oncology*, 13(1):177, 2018.

Frank Lohr, Mostafa El-Haddad, Barbara Dobler, Roland Grau, Hans-Joerg Wertz, Uta Kraus-Tiefenbacher, Volker Steil, Yasser Abo Madyan, and Frederik Wenz. Potential effect of robust and simple imrt approach for left-sided breast cancer on cardiac mortality. *International Journal of Radiation Oncology* Biology* Physics*, 74(1):73–80, 2009.

Fanwen Meng, Jin Qi, Meilin Zhang, James Ang, Singfat Chu, and Melvyn Sim. A robust optimization model for managing elective admission in a public hospital. *Operations research*, 63(6):1452–1467, 2015.

So-Yeon Park, Jung-in Kim, Minsoo Chun, Hyunjun Ahn, and Jong Min Park. Assessment of the modulation degrees of intensity-modulated radiation therapy plans. *Radiation Oncology*, 13(1):244, 2018.

R Tyrrell Rockafellar, Stanislav Uryasev, et al. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–42, 2000.

H Edwin Romeijn, Ravindra K Ahuja, James F Dempsey, Arvind Kumar, and Jonathan G Li. A novel linear programming approach to fluence map optimization for intensity modulated radiation therapy treatment planning. *Physics in Medicine & Biology*, 48 (21):3521, 2003.

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

Ernst Roos and Dick den Hertog. Reducing conservatism in robust optimization. *INFORMS Journal on Computing*, 32(4):1109–1127, 2020.

Andrew J Schaefer. Inverse integer programming. *Optimization Letters*, 3(4):483–489, 2009.

Anita Schöbel. Generalized light robustness and the trade-off between robustness and nominal quality. *Mathematical Methods of Operations Research*, 80(2):161–191, 2014.

Allen L Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations research*, 21(5):1154–1157, 1973.

J Unkelbach and U Oelfke. Inclusion of organ movements in imrt treatment planning via inverse planning based on probability distributions. *Physics in Medicine & Biology*, 49 (17):4005, 2004.

Jan Unkelbach, Markus Alber, Mark Bangert, Rasmus Bokrantz, Timothy CY Chan, Joseph O Deasy, Albin Fredriksson, Bram L Gorissen, Marcel Van Herk, Wei Liu, et al. Robust radiotherapy planning. *Physics in Medicine & Biology*, 63(22):22TR02, 2018.

Christian Vrančić, Alexei Trofimov, Timothy CY Chan, Gregory C Sharp, and Thomas Bortfeld. Experimental evaluation of a robust optimization method for imrt of moving targets. *Physics in Medicine & Biology*, 54(9):2901, 2009.

Lizhi Wang. Cutting plane algorithms for the inverse mixed integer linear programming problem. *Operations Research Letters*, 37(2):114–116, 2009.

Huan Xu. *Robust decision making and its applications in machine learning*. Citeseer, 2009.

Fangwei Zhao, Hua Sun, Fangxia Zhao, Hong Zhang, and Tongfei Li. A globalized robust optimization approach of dynamic network design problem with demand uncertainty. *IEEE Access*, 7:115734–115748, 2019.

APPENDICES

Appendix A

A.1 Appendix 1: Proofs of Chapter 2

We present the proofs of the theorems and propositions herein. Each section will provide the proof of one theorem discussed in this chapter.

A.1.1 Proof of Proposition 2.2.1

The method to prove this proposition is similar to the proof of the theorem that presented a PRO solution for a problem with one type of constraint [Iancu and Trichakis, 2014]. Let $\mathbf{x}$ be a solution to model (2.11). Suppose that there exists an RO solution $\mathbf{x}^*$ for model (2.10) that dominated $\mathbf{x}$. Thus,

$$\mathbf{v}'s(\mathbf{x}^*, \mathbf{A}) \geq \mathbf{v}'s(\mathbf{x}, \mathbf{A}) \quad \forall \mathbf{A} \in \mathcal{U}_1 \cup \mathcal{U}_2$$
$$\mathbf{v}'s(\mathbf{x}^*, \bar{\mathbf{A}}) > \mathbf{v}'s(\mathbf{x}, \bar{\mathbf{A}}) \quad \text{for some } \bar{\mathbf{A}} \in \mathcal{U}_1 \cup \mathcal{U}_2$$

That is $\mathbf{p_1 x^*} \geq \mathbf{p_1 x}$ for all $\mathbf{p_1} \in \mathcal{V}_1$, $\mathbf{p_2 x^*} \leq \mathbf{p_2 x}$ for all $\mathbf{p_2} \in \mathcal{V}_2$ and either $\mathbf{\hat{p}_1 x^*} > \mathbf{\hat{p}_1 x}$ for some $\mathbf{\hat{p}_1} \in \mathcal{V}_1$ or $\mathbf{\hat{p}_2 x^*} < \mathbf{\hat{p}_2 x}$ for some $\mathbf{\hat{p}_2} \in \mathcal{V}_2$. Without loss of generality, we consider $\exists \mathbf{\hat{p}}_1 \in ext(\mathcal{V}_1)$ such that $\mathbf{\hat{p}_1 x^*} > \mathbf{\hat{p}_1 x}$, where $\mathbf{\hat{p}_1}$ is a solution to $\max_{\mathbf{p} \in \mathcal{V}_1} \mathbf{p}(\mathbf{x}^* - \mathbf{x})$.

Since $\mathbf{\bar{p}_1} \in ri(\mathcal{V}_1)$, we can write $\mathbf{\bar{p}_1}$ as a convex combination of all extreme points of $\mathcal{V}_1$. That is, $\exists \lambda^1 \in \mathbb{R}^{|ext(\mathcal{V}_1)|}$ such that $\lambda^1 > \mathbf{0}$ and $\mathbf{1}'\lambda^1 = 1$ and $\mathbf{\bar{p}_1} = \sum_{i \in \mathcal{E}_1} \mathbf{p_1}^i \lambda_i^1$ where $ext(\mathcal{V}_1) = \{\mathbf{p_1}^i | i \in \mathcal{E}_1\}$. With the same method, there exists $\lambda^2 \in \mathbb{R}^{|ext(\mathcal{V}_2)|}$ such that $\lambda^2 > \mathbf{0}$ and $\mathbf{1}'\lambda^2 = 1$ and $\mathbf{\bar{p}_2} = \sum_{i \in \mathcal{E}_2} \mathbf{p_2}^i \lambda_i^2$ where $ext(\mathcal{V}_2) = \{\mathbf{p_2}^i | i \in \mathcal{E}_2\}$. Thus,

$$
\begin{aligned}
(\mathbf{\bar{p}_1} - \mathbf{\bar{p}_2})(\mathbf{x}^* - \mathbf{x}) &= \sum_{i \in \mathcal{E}_1} \mathbf{p_1}^i (\mathbf{x}^* - \mathbf{x}) - \sum_{i \in \mathcal{E}_2} \mathbf{p_2}^i (\mathbf{x}^* - \mathbf{x}) \\
&= \hat{\mathbf{p_1}}(\mathbf{x}^* - \mathbf{x}) + \sum_{i \in \mathcal{E}_1, \mathbf{p_1}^i \neq \hat{\mathbf{p_1}}} \mathbf{p_1}^i (\mathbf{x}^* - \mathbf{x}) - \sum_{i \in \mathcal{E}_2} \mathbf{p_2}^i (\mathbf{x}^* - \mathbf{x}).
\end{aligned}
$$

Therefore, $(\mathbf{\bar{p}_1} - \mathbf{\bar{p}_2})(\mathbf{x}^* - \mathbf{x}) > 0$ and so $(\mathbf{\bar{p}_1} - \mathbf{\bar{p}_2})\mathbf{x}^* > (\mathbf{\bar{p}_1} - \mathbf{\bar{p}_2})\mathbf{x}$ and $\mathbf{x}$ is not an optimal solution to model (2.11). It is a contradiction, and so $\mathbf{x}$ is not dominated by $\mathbf{x}^*$. Thus, $\mathbf{x}$ is a PRO solution.

A.1.2 Proof of Theorem 2.4.1

To prove this theorem, we have to reformulate model (2.14) as the following model,

$$
\begin{aligned}
&\min_{\mathbf{x}} \mathbf{c}'\mathbf{x} \qquad \text{(A.1)} \\
&\text{subject to: } \mathbf{Ax} \geq \mathbf{b} \quad \forall \mathbf{A} \in \mathcal{U}_A.
\end{aligned}
$$

Consider vector $\mathbf{c} \in \mathbb{R}^{|\mathcal{B}|}$ as $\mathbf{c}_b = \sum\limits_{v \in \mathcal{V}} \sum\limits_{x \in \mathcal{X}} \Delta_{v,x,b} p(x)$. Also, let $\tilde{\mathbf{A}}_{v,b} = \sum_{x \in \mathcal{X}} \Delta_{v,x,b} \tilde{p}(x)$ for all $v \in \mathcal{T}$, $b \in \mathcal{B}$, and $\tilde{p} \in \mathcal{U}_p$, and let $\mathcal{U}_A$ be the set of all such $\mathbf{A}$. Thus, we can write model (2.14) as follows,

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{c}'\mathbf{w} \\ \text{subject to: } & \tilde{\mathbf{A}}\mathbf{w} \geq \theta \quad \forall \mathbf{A} \in \mathcal{U}_A. \end{aligned} \tag{A.2}$$

Now, using the result in section 2.2.2, a solution $\mathbf{w}$ is a PRO solution if it is a solution to the following problem,

$$\min_{\mathbf{w} \in \mathcal{W}^{RO}} \mathbf{t}'\hat{\mathbf{A}}\mathbf{w}, \tag{A.3}$$

where $\mathbf{t}$ is a value vector for the tumour voxels, and $\hat{\mathbf{A}} \in ri(\mathcal{U}_A)$.

Now, we have the following lemma.

Lemma A.1.1. *Let $\tilde{\mathbf{A}}_{v,b} = \sum_{x \in \mathcal{X}} \Delta_{v,x,b} \tilde{p}(x)$ be an element of the uncertain matrix $\tilde{\mathbf{A}}_{|\mathcal{V}| \times |\mathcal{B}|} \in \mathcal{U}_A$. We have $\tilde{\mathbf{A}} \in ri(\mathcal{U}_A)$ if and only if $\tilde{p}(x) \in ri(\mathcal{U}_p)$.*

Proof. We prove each direction of the statement separately.

[$\tilde{\mathbf{A}} \in ri(\mathcal{U}_A) \Longrightarrow \tilde{\mathbf{p}} \in ri(\mathcal{U}_p)$]: Let $\mathbf{M}_v$ be a matrix such that the entry $(\mathbf{M}_v)_{b,x}$ is the dose delivered from beamlet b to voxel v in phase x. Suppose $\tilde{\mathbf{A}} \in ri(\mathcal{U}_A)$, it is sufficient to prove that $\tilde{\mathbf{p}} \in \mathcal{U}_p$ and for every $\mathbf{p}_1 \in \mathcal{U}_p$ there exists $\lambda > 1$ such that $\lambda \tilde{\mathbf{p}} + (1 - \lambda)\mathbf{p}_1 \in \mathcal{U}$.

For each $\mathbf{p}_1$ we an rewrite $\mathbf{A_1}$ as

$$\mathbf{A_1} = \begin{pmatrix} \mathbf{p}_1'\mathbf{M}_1' \\ \vdots \\ \mathbf{p}_1'\mathbf{M}_{|\mathcal{T}|}' \end{pmatrix}.$$

We know that $\tilde{\mathbf{A}}$ is in $ri(\mathcal{U}_A)$, thus there exists $\lambda > 1$ such that $\lambda\tilde{\mathbf{A}} + (1-\lambda)\mathbf{A_1} \in \mathcal{U}_A$. We have

$$\lambda\tilde{\mathbf{A}} + (1-\lambda)\mathbf{A_1} = \lambda \begin{pmatrix} \tilde{\mathbf{p}}'\mathbf{M}_1' \\ \vdots \\ \tilde{\mathbf{p}}'\mathbf{M}_{|\mathcal{T}|}' \end{pmatrix} + (1-\lambda) \begin{pmatrix} \mathbf{p}_1'\mathbf{M}_1' \\ \vdots \\ \mathbf{p}_1'\mathbf{M}_{|\mathcal{T}|}' \end{pmatrix} = \begin{pmatrix} (\lambda\tilde{\mathbf{p}} + (1-\lambda)\mathbf{p}_1)'\mathbf{M}_1' \\ \vdots \\ (\lambda\tilde{\mathbf{p}} + (1-\lambda)\mathbf{p}_1)'\mathbf{M}_{|\mathcal{T}|}' \end{pmatrix}.$$

Therefore, $\lambda\tilde{\mathbf{P}} + (1-\lambda)\mathbf{p}_1 \in \mathcal{U}_p$ and so $\tilde{\mathbf{p}} \in ri(\mathcal{U}_p)$.

[$\tilde{\mathbf{A}} \in ri(\mathcal{U}_A) \Longleftarrow \tilde{\mathbf{p}} \in ri(\mathcal{U}_p)$]: We know $\tilde{\mathbf{A}} \in \mathcal{U}_A$, and need to prove that for each $\mathbf{A_1} \in \mathcal{U}_A$, there exists $\lambda > 1$ such that $\lambda\tilde{\mathbf{A}} + (1-\lambda)\mathbf{A_1} \in \mathcal{U}_A$. Since $\tilde{\mathbf{p}} \in ri(\mathcal{U}_p)$, for each $\mathbf{p}_1 \in \mathcal{U}_p$ we have $\lambda\tilde{\mathbf{p}} + (1-\lambda)\mathbf{p}_1 \in \mathcal{U}_p$ for some $\lambda > 1$. We know that $\mathbf{A_1} \in \mathcal{U}_A$, thus there exists $\mathbf{p}_1 \in \mathcal{U}_p$ such that

$$\mathbf{A_1} = \begin{pmatrix} \mathbf{p}_1'\mathbf{M}_1' \\ \vdots \\ \mathbf{p}_1'\mathbf{M}_{|\mathcal{T}|}' \end{pmatrix}.$$

We have $\lambda\tilde{\mathbf{p}} + (1-\lambda)\mathbf{p_1} \in \mathcal{U}_p$ for some $\lambda > 1$, thus

$$A = \begin{pmatrix} (\lambda\tilde{\mathbf{p}} + (1-\lambda)\mathbf{p}_1)'\mathbf{M}_1' \\ \vdots \\ (\lambda\tilde{\mathbf{p}} + (1-\lambda)\mathbf{p}_1)'\mathbf{M}_{|\mathcal{T}|}' \end{pmatrix} \in \mathcal{U}_A.$$

Hence, we can write

$$\mathbf{A} = \lambda \begin{pmatrix} \tilde{\mathbf{p}}'\mathbf{M}_1' \\ \vdots \\ \tilde{\mathbf{p}}'\mathbf{M}_{|\mathcal{T}|}' \end{pmatrix} + (1-\lambda) \begin{pmatrix} \mathbf{p}_1'\mathbf{M}_1' \\ \vdots \\ \mathbf{p}_1'\mathbf{M}_{|\mathcal{T}|}' \end{pmatrix} = \lambda\tilde{\mathbf{A}} + (1-\lambda)\mathbf{A_1} \in \mathcal{U}_A.$$

Therefore, $\tilde{\mathbf{A}} \in ri(\mathcal{U}_A)$.

□

Now, we reformulate model (A.3), as

$$\max_{(w_1,\ldots,w_{|\mathcal{B}|})\in\mathcal{W}^{RO}} \sum_{v\in\mathcal{T}} \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} t_i \Delta_{v,x,b} \hat{p}(x) w_b \tag{A.4}$$

Now, the question is what is $\mathcal{W}^{RO}$. To find the PRO solutions, we need to determine the set of all solutions which are feasible for all uncertainty scenario and have the optimal value. We denote this set by $\mathcal{W}^{RO}$. Lemma 2, presents this set.

Lemma A.1.2. *The set of robust optimal solutions of formulation* (2.14) *is*

$$
\begin{aligned}
\mathcal{W}^{RO} = \{(w_1, ..., w_{|\mathcal{B}|}) \in \mathbb{R}^{+|\mathcal{B}|} | & \sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} \Delta_{v,x,b} p(x) w_b \leq Z^{RO}, \\
& \sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} \Delta_{v,x,b} p(x) w_b - \sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} \Delta_{v,x,b} \underline{p}(x) w_b + \\
& \sum_{x\in\mathcal{X}} \underline{p}(x) q_v - \sum_{x\in\mathcal{X}} r_{v,x} \geq \theta_v \quad \forall v \in \mathcal{T}, \\
& (\bar{p}(x) + \underline{p}(x)) q_v - r_{v,x} \leq \sum_{b\in\mathcal{B}} \Delta_{v,x,b} (\bar{p}(x) + \underline{p}(x)) w_b \quad \forall v \in \mathcal{T}, \forall x \in \mathcal{X}, \\
& q_v \geq 0 \quad \forall v \in \mathcal{T}, \\
& r_{v,x} \geq 0 \quad v \in \mathcal{T}, x \in \mathcal{X}\}.
\end{aligned}
$$

where Z^{RO} is the optimal value of the problem.

Proof. consider the following linear programming model for the radiotion therapy problem which is derived from Theorem 1 in Bortfeld et al. [2008],

$$
\begin{aligned}
& \min_{w,q,r} \sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} \Delta_{v,x,b} p(x) w_b \\
\text{subject to: } & \sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} \Delta_{v,x,b} p(x) w_b - \sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} \Delta_{v,x,b} \underline{p}(x) w_b + \\
& \sum_{x\in\mathcal{X}} \underline{p}(x) q_v - \sum_{x\in\mathcal{X}} r_{v,x} \geq \theta_v \quad \forall v \in \mathcal{T}
\end{aligned}
\tag{A.5}
$$

$$(\bar{p}(x)+\underline{p}(x))q_v - r_{v,x} \le \sum_{b\in\mathcal{B}} \Delta_{v,x,b}(\bar{p}(x)+\underline{p}(x))w_b \quad \forall v\in\mathcal{T}, \forall x\in\mathcal{X}$$

$$q_v\text{free} \quad \forall v\in\mathcal{T}$$

$$r_{v,x}\ge 0 \quad \forall v\in\mathcal{T}, \forall x\in\mathcal{X},$$

$$w_b\ge 0 \quad \forall b\in\mathcal{B}.$$

Thus,

$$\begin{aligned}
\mathcal{W}^{RO} = \{(w_1,...,w_{|\mathcal{B}|})\in\mathbb{R}^{+|\mathcal{B}|} | &\sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}}\Delta_{v,x,b}p(x)w_b \le Z^{RO}, \\
&\text{and} \quad \exists q_v \text{ for } v\in\mathcal{T}, \quad \exists r_{v,x} \quad \text{for } v\in\mathcal{T} \text{ and } x\in\mathcal{X}, \\
&\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}}\Delta_{v,x,b}p(x)w_b - \sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}}\Delta_{v,x,b}\underline{p}(x)w_b + \\
&\sum_{x\in\mathcal{X}}\underline{p}(x)q_v - \sum_{x\in\mathcal{X}}r_{v,x} \ge \theta_v \quad \forall v\in\mathcal{T} \quad \text{and} \\
&(\bar{p}(x)+\underline{p}(x))q_v - r_{v,x} \le \sum_{b\in\mathcal{B}}\Delta_{v,x,b}(\bar{p}(x)+\underline{p}(x))w_b \quad \forall v\in\mathcal{T}, \forall x\in\mathcal{X}\}.
\end{aligned} \tag{A.6}$$

□

Using Lemma A.1.1, it is sufficient to find the solution of

$$\max_{(w_1,...,w_{|\mathcal{B}|})\in\mathcal{W}^{RO}} \sum_{v\in\mathcal{T}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} t_i\Delta_{v,x,b}\hat{p}(x)w_b.$$

Using Lemma A.1.2, we can find $\mathcal{W}^{RO}$. □

A.1.3 Proof of Theorem 2.4.2

The idea to proof this theorem is similar to the proof of Theorem 2.4.1. Add the constraint $\tilde{\mathbf{A}}\mathbf{w} \geq \theta \quad \forall \mathbf{A} \in \mathcal{U}_A$ to model (A.2). Now, using Proposition 2.2.1 and Lemmas A.1.1 and A.1.2 any solution to problem

$$\max_{(w_1,\ldots,w_{|\mathcal{B}|})\in\mathcal{W}^{RO}} \sum_{v\in\mathcal{T}} \sum_{b\in\mathcal{B}} \sum_{x\in\mathcal{X}} (t_i - s_i)\Delta_{v,x,b}\hat{p}(x)w_b. \tag{A.7}$$

is a PRO solution. □

A.1.4 Proof of Theorem 2.4.3

To prove this theorem, first we reformulate problem (2.20) as model

$$\begin{aligned} &\min_{\mathbf{x}} \mathbf{c}'\mathbf{x} \\ &\text{subject to: } \mathbf{A}\mathbf{x} \geq \mathbf{b} \quad \forall \mathbf{A} \in \mathcal{U}_A. \end{aligned} \tag{A.8}$$

to present the PRO solutions.

Consider the decision variable $\mathbf{f} \in \mathbb{R}^{|\mathcal{B}|+1+|\mathcal{T}|}$ such that

$$(\mathbf{f})_i = \begin{cases} w_i & \text{for} \quad i = 1, ..., |\mathcal{B}| \\ \underline{\zeta}_\beta & \text{for} \quad i = |\mathcal{B}| + 1 \\ \underline{d}_{i-|\mathcal{B}|-1,\beta} & \text{for} \quad i = |\mathcal{B}| + 2, ..., |\mathcal{B}| + 1 + |\mathcal{T}| \end{cases} ; \tag{A.9}$$

Let $\mathbf{y}_1 \in \mathbb{R}^{\mathcal{B}|+1+|\mathcal{T}|}$ be $\mathbf{y}_1$ is a vector in $\mathbb{R}^o$ such that

$$(\mathbf{y}_1)_i = \begin{cases} 1 & \text{if} \quad i = |\mathcal{B}| + 1 \\ 0 & \text{otherwise} \end{cases}; \tag{A.10}$$

Consider$\mathbf{y}_2$ as a vector in $\mathbb{R}^{|\mathcal{B}|+1+|\mathcal{T}|}$ with the following entries

$$(\mathbf{y}_2)_i = \begin{cases} \frac{-1}{(1-\beta)|\mathcal{T}|} & \text{if} \quad |\mathcal{B}| + 2 \leq i \leq |\mathcal{B}| + |\mathcal{T}| + 1 \\ 0 & \text{otherwise} \end{cases}; \tag{A.11}$$

; and matrix $\underline{\mathbf{D}}$ as

$$(\underline{\mathbf{D}})_{k,l} = \begin{cases} 1 & \text{if} \quad l = |\mathcal{B}| + 1 + k \\ 0 & \text{otherwise} \end{cases} \tag{A.12}$$

The following matrix is denoted by $\mathbf{A}$ which is a $|\mathcal{T}| \times (|\mathcal{B}| + 1 + |\mathcal{T}|)$ matrix:

$$(\mathbf{A})_{k,l} = \begin{cases} -1 & \text{if} \quad l = |\mathcal{B}| + 1 \\ 0 & \text{otherwise} \end{cases} \tag{A.13}$$

First suppose that for a specific voxel j, $\mathbf{\Delta_j}$ is a matrix such that $(\mathbf{\Delta_j})_{i,b} = \Delta_{j,i,b}$. Considering that for each voxel $k \in \mathcal{T}$, we have $\tilde{\mathbf{a}}_\mathbf{k} = \tilde{\mathbf{p}}'(\mathbf{\Delta_k})'$ for some $\tilde{\mathbf{p}} \in \mathcal{U}_P$.

Let matrix $\tilde{\mathbf{G}}$,be

$$(\tilde{\mathbf{G}})_{k,l} = \begin{cases} (\tilde{\mathbf{a}}_{\mathbf{k}})_l & \text{if} \quad 1 \leq l \leq |\mathcal{B}| \\ 0 & \text{otherwise} \end{cases} \tag{A.14}$$

. Now we can model the problem (2.20) as

$$\begin{aligned} \min_{\mathbf{f}} \quad & j\gamma_H \mathbf{f} \\ \text{Subject to:} \quad & \mathbf{y_1 f} + \mathbf{y_2 f} \geq L_\beta, \\ & \underline{\mathbf{D}}\mathbf{f} + \mathbf{A}\mathbf{f} + \tilde{\mathbf{G}}\mathbf{f} \geq \mathbf{0}, \end{aligned} \tag{A.15}$$

Now, if $\mathcal{U}_1 = \{\tilde{\mathbf{M}} = \bar{\mathbf{D}} + \mathbf{A} + \tilde{\mathbf{G}};\text{ for all } \tilde{\mathbf{G}}\}$. Using the result in Section 2.2.2, solutions to the following problem are PRO solutions for some $\hat{\mathbf{M}} \in ri(\mathcal{U}_1)$,

$$\max_{\mathbf{f} \in \mathcal{F}^{RO}} \mathbf{t}\hat{\mathbf{M}}\mathbf{f}$$

where $\mathbf{t}$ is a value vector of the tumour voxels, and $\mathcal{F}^{RO} = \{\mathbf{f} |\ \mathbf{f}$ is constructed by$(\mathcal{W}, \underline{\mathcal{D}}, \underline{\zeta}_\beta) \in \mathcal{S}^{RO}\}$ such that $\mathcal{W} = \{w_1, ..., w_{|\mathcal{B}|}\}$,$\underline{\mathcal{D}} = \{\underline{d}_{v,\beta} | v \in \mathcal{T}\}$ and $\mathcal{S}^{RO}$ is the set of all robust optimal solutions to problem (2.20) with a single β.

Using the same idea for proving Lemma A.1.1, we prove that if $\mathbf{M}$ is generated by $\mathbf{p}$, then $\mathbf{M} \in ri(\mathcal{U}_1)$ if and only if $\mathbf{p} \in ri(\mathcal{U}_p)$. Therefor, the solutions to the following problem are

PRO solutions for some $\hat{\mathbf{p}} \in ri(\mathcal{U}_p)$,

$$\max_{(\mathcal{W},\underline{\mathcal{D}},\underline{\zeta}_\beta)\in\mathcal{S}^{RO}} \sum_{v\in\mathcal{T}} t_v(\underline{d}_{v,\beta} - \underline{\zeta}_\beta + \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}\hat{p}(x)w_b) \tag{A.16}$$

Now, the question is about the set $\mathcal{S}^{RO}$. Using equation $(B.2)$ in Chan et al. [2014], model (2.20) can be reformulated as the following linear problem

$$\begin{aligned}
\min \quad & \sum_{v\in\mathcal{H}}\sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}p(x)w_b && \text{(A.17)}\\
\text{Subject to: } & \underline{\zeta}_\beta - \frac{1}{(1-\beta)|\mathcal{T}|}\sum_{v\in\mathcal{T}} \bar{d}_{v,\beta} \geq L_\beta, \\
& \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}p(x)w_b - \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}\underline{p}(x)w_b + \sum_{x\in\mathcal{X}} \underline{p}(x)q_v - \\
& \sum_{x\in\mathcal{X}} r_{v,x} \geq \underline{\zeta}_\beta - \underline{d}_{v,\beta}, \quad \forall v \in \mathcal{T}, \\
& \sum_{b\in\mathcal{B}} \Delta_{v,x,b}(\bar{p}(x) + \underline{p}(x))w_b + r_{v,x} + (\bar{p}(x) + \underline{p}(x))q_v \geq 0 \ \forall x \in \mathcal{X}, v \in \mathcal{T}.
\end{aligned}$$

Therefore, we can write the set of all robust optimal solutions of problem (2.20) with a single β as follows:

$$
\begin{aligned}
&\mathcal{S}^{RO} = \{(\mathcal{W} = \{w_1, ..., w_{|\mathcal{B}|}\}, \underline{\mathcal{D}} = \{\underline{d}_{v,\beta} | v \in \mathcal{T}\}, \zeta_\beta), \\
&\sum_{v\in\mathcal{H}} \sum_{x\in\mathcal{X}} \sum_{b\in\mathcal{B}} \Delta_{v,x,b} p(x) w_b \leq Z^{RO}, \\
&\underline{\zeta}_\beta - \frac{1}{(1-\beta)|\mathcal{T}|} \sum_{v\in\mathcal{T}} \underline{d}_{v,\beta} \geq L_\beta \quad , \text{ and} \\
&\sum_{x\in\mathcal{X}} \sum_{b\in\mathcal{B}} \Delta_{v,x,b} p(x) w_b - \sum_{x\in\mathcal{X}} \sum_{b\in\mathcal{B}} \Delta_{v,x,b} \underline{p}(x) w_b + \sum_{x\in\mathcal{X}} \underline{p}(x) q_v - \sum_{x\in\mathcal{X}} r_{v,x} \geq \underline{\zeta}_\beta - \underline{d}_{v,\beta}, \quad \forall v \in \mathcal{T}, \\
&\sum_{b\in\mathcal{B}} \Delta_{v,x,b} (\bar{p}(x) + \underline{p}(x)) w_b + r_{v,x} + (\bar{p}(x) + \underline{p}(x)) q_v \geq 0 \\
&\forall x \in \mathcal{X}, v \in \mathcal{T}, \\
&\bar{q}_v \geq 0 \quad \forall v \in \mathcal{T}, \\
&r_{v,x} \geq 0 \quad \forall v \in \mathcal{T}, x \in \mathcal{X}\}.
\end{aligned}
\tag{A.18}
$$

Thus, all solutions to the following problem are PRO solutions for some $\hat{\mathbf{p}} \in ri(\mathcal{U}_p)$,

$$\max_{(\mathcal{W},\underline{\mathcal{D}},\underline{\zeta}_\beta)} \sum_{v\in\mathcal{T}} t_v(\underline{d}_{v,\beta} - \underline{\zeta}_\beta + \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}\hat{p}(x)w_b) \tag{A.19}$$

$$\begin{aligned}
\text{subject to:} & \sum_{v\in\mathcal{H}}\sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}p(x)w_b \le Z^{RO}, \\
& \underline{\zeta}_\beta - \frac{1}{(1-\beta)|\mathcal{T}|}\sum_{v\in\mathcal{T}} \underline{d}_{v,\beta} \ge L_\beta, \\
& \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}p(x)w_b - \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}\underline{p}(x)w_b + \sum_{x\in\mathcal{X}} \underline{p}(x)q_v \\
& -\sum_{x\in\mathcal{X}} r_{v,x} \ge \underline{\zeta}_\beta - \underline{d}_{v,\beta} \quad \forall v \in \mathcal{T}, \\
& \sum_{b\in\mathcal{B}} \Delta_{v,x,b}(\bar{p}(x) + \underline{p}(x))w_b + r_{v,x} + (\bar{p}(x) + \underline{p}(x))q_v \ge 0 \quad \forall x \in \mathcal{X}, v \in \mathcal{T}.
\end{aligned}$$

□

A.1.5 Proof of Proposition 2.4.1

The idea to prove this proposition is similar to the proof of Theorem 2.4.3 by considering a similar set of constraints for each $\beta \in \underline{\mathcal{A}}$. Thus, we just need to add the result of the problems related to each $\beta \in \underline{\mathcal{A}}$. □

A.1.6 Proof of Theorem 2.4.4

To prove this theorem, first we reformulate problem (2.24) to the following model

$$\max_{\mathbf{x}\in\mathcal{X}} \min_{\mathbf{p}\in\mathcal{U}} \mathbf{p}'\mathbf{x}, \tag{A.20}$$

that is the form of problem (2.1). Consider matrix $\mathbf{G_v} = (-\Delta_{v,x,b})_v$; that is, the matrix $\mathbf{G_v}$ is the negative of $\Delta_{v,x,b}$ for the specific v. Let, $\tilde{\mathbf{a}}_v = \mathbf{p}'\mathbf{G_v}$ for all $v \in \mathcal{V}$. Denote the matrix $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_{|\mathcal{T}|} \end{pmatrix}$ for a constant $\mathbf{p}$ such that $\mathbf{a}_i = \mathbf{p}'\mathbf{G_i}$.

Now, let

$$\tilde{\mathbf{s}} = \mathbf{j} \begin{pmatrix} \tilde{\mathbf{a}}_1 \\ \tilde{\mathbf{a}}_2 \\ \vdots \\ \tilde{\mathbf{a}}_{|\mathcal{V}|} \end{pmatrix}$$

and vector $\delta = -\theta$; also, denote the set of all $\mathbf{s}$ with $\mathcal{U}_s$. Then, the model can be written as

$$\begin{aligned} \max_{w} \min_{\tilde{\mathbf{s}}} \quad & \tilde{\mathbf{s}}\mathbf{w} \\ \text{subject to:} \quad & \mathbf{A}\mathbf{w} \leq \delta, \\ & \mathbf{w} \geq 0. \end{aligned} \tag{A.21}$$

Now, using the results in Section 2.2.1 to find a PRO solution, if there exists $\hat{\mathbf{s}} \in ri(\mathcal{U}_s)$, then all of the optimal solutions to the problem

$$\max_{w \in \mathcal{W}^{RO2}} \hat{\mathbf{s}}\mathbf{w}$$

are PRO solutions, where $\mathcal{W}^{RO2}$ is the set of RO solutions.

Using the same idea for proving Lemma A.1.1, we prove that if $\mathbf{s}$ is generated by $\mathbf{p}$, then $\mathbf{s} \in ri(\mathcal{U}_s)$ if and only if $\mathbf{p} \in ri(\mathcal{U}_p)$. Therefore, a solution to the following problem is a PRO one for some $\hat{\mathbf{p}} \in ri(\mathcal{U}_p)$.

$$\max_{(w_1,..,w_{|\mathcal{B}|}) \in \mathcal{W}^{RO2}} \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in \mathcal{X}} -\Delta_{v,x,b} \hat{p}(x) w_b$$

are Pareto optimal solutions.

Now, the question is about the set W^{RO2}. To answer this question, we find the dual of the following problem for a constant w:

$$\begin{aligned} \min_{\hat{p}} & \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in \mathcal{X}} -\Delta_{v,x,b} \hat{p}(x) w_b \\ \text{subject to:} \quad & \sum_{x \in \mathcal{X}} \hat{p}(x) = 0 \\ & -\underline{p}(x) \le \hat{p}(x) \le \bar{p}(x) \quad \forall x \in \mathcal{X} \end{aligned} \tag{A.22}$$

The dual of this problem is

$$\max_{q_x, q'_x, y} \sum_{x \in \mathcal{X}} \bar{p}(x) q_x - \sum_{x \in \mathcal{X}} \underline{p}(x) q'(x) \tag{A.23}$$

$$\begin{aligned} \text{subject to:} \quad & y + q_x + q'_x = \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} -\Delta_{v,x,b} w_b \quad \forall x \in \mathcal{X}, \\ & y \text{ is free,} \\ & q_x \leq 0, \\ & q'_x \geq 0. \end{aligned}$$

Therefore, we can write the original problem as follows.

$$\max_{q_x, q'_x, y, w} \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} \sum_{x \in \mathcal{X}} -\Delta_{v,x,b} p(x) w_b + \sum_{x \in \mathcal{X}} \bar{p}(x) q_x - \sum_{x \in \mathcal{X}} \underline{p}(x) q'(x) \tag{A.24}$$

$$\begin{aligned} \text{subject to:} \quad & \sum_{b \in \mathcal{B}} \sum_{x \in \mathcal{X}} \Delta_{v,x,b} p(x) w_b \geq \theta_v \quad \forall v \in \tau \\ & \quad y + q_x + q'_x = \sum_{v \in \mathcal{V}} \sum_{b \in \mathcal{B}} -\Delta_{v,x,b} w_b \quad \forall x \in \mathcal{X}, \\ & w_b \geq 0 \\ & y \text{ is free,} \\ & q_x \leq 0, \\ & q'_x \geq 0. \end{aligned}$$

Hence, we can write the set of robust optimal solutions of problem (2.24) as the set of all optimal solutions of the above set. Thus,

$$\begin{aligned}
\mathcal{W}^{RO2} = \{&(w_1, ..., w_{|\mathcal{B}|}) \in \mathbb{R}^{+|\mathcal{B}|} | \\
&\sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} -\Delta_{v,x,b}\bar{p}(x)w_b + \sum_{x\in\mathcal{X}} \bar{p}(x)\bar{q}_x - \sum_{x\in\mathcal{X}} \underline{p}(x)\underline{q}_x \geq Z^{RO} \text{ and} \\
&\sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}\bar{p}(x)w_b \geq \theta_v \ \forall v \in \mathcal{T}, \\
&y + \bar{q}_x + \underline{q}_x = \sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}} -\Delta_{v,x,b}w_b \ \forall x \in \mathcal{X}, \\
&\bar{q}_x \leq 0 \ \forall x \in \mathcal{X}, \\
&\underline{q}_x \geq 0 \ \forall x \in \mathcal{X}\}.
\end{aligned}$$

Therefor, the solution to the following problem is a PRO solution,

$$\begin{aligned}
\max_{(w_1,..,w_{|\mathcal{B}|})} \ & \sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} -\Delta_{v,x,b}\hat{p}(x)w_b \\
\text{subject to:} \ & \sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}} -\Delta_{v,x,b}\bar{p}(x)w_b + \sum_{x\in\mathcal{X}} \bar{p}(x)\bar{q}_x - \sum_{x\in\mathcal{X}} \underline{p}(x)\underline{q}_x \geq Z^{RO}, \\
& \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}\bar{p}(x)w_b \geq \theta_v \ \forall v \in \mathcal{T}, \\
& y + \bar{q}_x + \underline{q}_x = \sum_{v\in\mathcal{V}}\sum_{b\in\mathcal{B}} -\Delta_{v,x,b}w_b \ \forall x \in \mathcal{X}.
\end{aligned}$$

□

A.1.7 Proof of Theorem 2.4.5

First, we reformulate model (2.25) as follows

$$\min_{\mathbf{x}} \mathbf{c}'\mathbf{x} \tag{A.25}$$
$$\text{subject to: } \mathbf{A}\mathbf{x} \geq \mathbf{b} \quad \forall \mathbf{A} \in \mathcal{U}_A.$$

that is the form of problem (2.1). For this purpose, let $\mathbf{f} \in \mathbb{R}^{|\mathcal{B}|+|\mathcal{H}|}$ be

$$\mathbf{f}_i = \begin{cases} w_i & \text{for} \quad i = 1, ..., |\mathcal{B}| \\ z_{i-|\mathcal{B}|} & \text{for} \quad i = |\mathcal{B}| + 1, ..., |\mathcal{B}| + |\mathcal{H}| \\ \bar{\zeta}_\beta & \text{for} \quad i = 1 + |\mathcal{B}| + |\mathcal{H}| \end{cases}, \tag{A.26}$$

Consider $\mathbf{G_1}$ as a $(|\mathcal{T}| \times q)$ matrix where

$$(\mathbf{G_1})_{k,l} = \begin{cases} (\mathbf{a_k})_l & \text{if} \quad 1 \leq l \leq |\mathcal{B}| \\ 0 & \text{otherwise} \end{cases}, \tag{A.27}$$

such that $\mathbf{a_j} = \mathbf{p}'(\mathbf{\Delta_j})'$ for each $1 \leq j \leq |\mathcal{T}|$ for a specific p. Denote the following matrix by $\tilde{\mathbf{G}}_\mathbf{2}$; that is, a $|\mathcal{H}| \times q$ matrix,

$$(\tilde{\mathbf{G}}_{\mathbf{2}})_{k,l} = \begin{cases} -(\tilde{\mathbf{a}}_{\mathbf{k}})_l & \text{if} \quad 1 \le l \le |\mathcal{B}| \\ 1 & \text{if} \quad l = |\mathcal{B}| + k \\ 1 & \text{if} \quad l = |\mathcal{B}| + |\mathcal{H}| + 1 \\ 0 & \text{otherwise} \end{cases} \tag{A.28}$$

such that $\tilde{\mathbf{a}}_{\mathbf{j}} = \tilde{\mathbf{p}}'(\mathbf{\Delta_j})'$ for each $1 \le j \le |\mathcal{H}|$. Since we have uncertainty for $\tilde{\mathbf{p}}$, then $\tilde{\mathbf{G}}_{\mathbf{2}}$ is also uncertain. Now, model (2.25) can be reformulated as follows,

$$\min_{\mathbf{f}} \mathbf{e}_{|\mathcal{B}|+|\mathcal{H}|+\mathbf{1}}\mathbf{f} + \frac{1}{(1-\beta)|\mathcal{H}|}\mathbf{y}\mathbf{f} \tag{A.29}$$

$$\text{Subject to:} \qquad \mathbf{G_1 f} \ge \theta, \tag{A.30}$$

$$\tilde{\mathbf{G}}_{\mathbf{2}}\mathbf{f} \ge 0, \tag{A.31}$$

where $\mathbf{y}$ is a $|\mathcal{B}| + |\mathcal{H}| + 1$ vector such that $\mathbf{y}_j = 1$ for $|\mathcal{B}| \le j \le |\mathcal{B}| + |\mathcal{H}|$ and $\mathbf{y}_j = 0$ in other places.

Therefore, to find a Pareto robust solutions pick $\hat{\mathbf{G}}_{\mathbf{2}} \in ri(\mathcal{U}_G)$, the solution of the following problem is a Pareto robust optimal solution,

$$\max_{f \in \mathcal{F}^{RO}} \hat{\mathbf{G}}_{\mathbf{2}}\mathbf{f} \tag{A.32}$$

Using the same method as Lemma A.1.1, we prove that if $\mathbf{G_2}$ is generated by $\mathbf{p}$, then $\mathbf{G_2} \in ri(\mathcal{U}_G)$ if and only if $\mathbf{p} \in ri(\mathcal{U}_p)$.

We have $\mathbf{G}_2\mathbf{f} = \sum_{v\in\mathcal{H}}(z_v + \bar{\zeta}_\beta - \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}}\Delta_{v,x,b}p(x)w_b)$.

Hence, for any $\hat{\mathbf{p}} \in ri(\mathcal{U}_p)$, the solution of the following problem is a PRO solution,

$$\max_{(W,Z,\bar{\zeta}_\beta)\in\mathcal{S}^{RO}} \sum_{v\in\mathcal{H}}(z_v + \bar{\zeta}_\beta - \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}}\Delta_{v,x,b}\hat{p}(x)w_b). \tag{A.33}$$

Now, the question is about the set $\mathcal{S}^{RO}$. To answer this question, we demonstrate the following linear program model of the problem (2.25):

$$\begin{aligned}
&\min_{z,\bar{\zeta}_\beta,w_b} \bar{\zeta}_\beta + \frac{1}{(1-\beta)m}\sum_{v\in\mathcal{H}} z_v && \text{(A.34)}\\
&\text{Subject to: } \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}}\Delta_{v,x,b}p(x)w_b \geq \theta_v \quad \forall v\in\mathcal{T},\\
&z_v + \bar{\zeta}_\beta - \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}}\Delta_{v,x,b}p(x)w_b + C_v(w) \geq 0 \quad \forall v\in\mathcal{H},\\
&z_v \geq 0,\\
&\beta \in \bar{\mathcal{A}}_s,\\
&w \in \mathcal{W},
\end{aligned}$$

where

$$\begin{aligned}
&C_v(w) = \min_{p^\cdot}\sum_{b\in\mathcal{B}}\sum_{x\in\mathcal{X}}\Delta_{v,x,b}p^\cdot(x)w_b && \text{(A.35)}\\
&\text{Subject to: } \sum_{x\in\mathcal{X}}\hat{p}(x) = 0,\\
&-\underline{p}(x) \leq p^\cdot(x) \leq \bar{p}(x) \quad \forall x\in\mathcal{X}.
\end{aligned}$$

Using the dual of problem (A.35), we have the following linear programming.

$$\min_{z,\bar{\zeta}_\beta,w_b} \bar{\zeta}_\beta + \frac{1}{(1-\beta)|\mathcal{H}|}\sum_{v\in\mathcal{H}} z_v \tag{A.36}$$

$$\begin{aligned}
\text{Subject to: } & \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}p(x)w_b \geq \theta_v \quad \forall v \in \mathcal{T}, \\
& z_v + \bar{\zeta}_\beta - \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}p(x)w_b + \sum_{x\in\mathcal{X}} \bar{p}(x)r_{v,x} + \sum_{x\in\mathcal{X}} \underline{p}(x)s_{v,x} \geq 0 \quad \forall v \in \mathcal{V} \\
& q_v + r_{v,x} + s_{v,x} = \sum_{b\in\mathcal{B}} \Delta_{v,x,b}w_b \quad \forall v \in \mathcal{H},\ x \in \mathcal{X} \\
& z_v \geq 0, \quad \forall v \in \mathcal{H} \\
& \beta \in \bar{\mathcal{A}}_s, \\
& w \in \mathcal{W}, \\
& r_{v,x} \leq 0, \quad \forall v \in \mathcal{H} \\
& s_{v,x} \geq 0, \quad \forall v \in \mathcal{H},\ x \in \mathcal{X} \\
& q_v \text{ free}, \quad \forall v \in \mathcal{H}.
\end{aligned}$$

Therefore, $\mathcal{S}^{RO}$ is the set of all optimal solutions to model (A.36), and so all solutions to the following problem are PRO solutions.

$$\max_{(W,Z,\bar{\zeta}_\beta)} \sum_{v\in\mathcal{H}} t_i(z_v + \bar{\zeta}_\beta - \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}\hat{p}(x)w_b) \tag{A.37}$$

$$\text{subject to: } \bar{\zeta}_\beta + \frac{1}{(1-\beta)|\mathcal{H}|}\sum_{v\in\mathcal{H}} z_v \leq Z^{RO},$$

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

$$z_v + \bar{\zeta}_\beta - \sum_{x\in\mathcal{X}}\sum_{b\in\mathcal{B}} \Delta_{v,x,b}p(x)w_b + \sum_{x\in\mathcal{X}} \bar{p}(x)r_{v,x} + \sum_{x\in\mathcal{X}} \underline{p}(x)s_{v,x} \geq 0 \quad \forall v \in \mathcal{V}$$

$$q_v + r_{v,x} + s_{v,x} = \sum_{b\in\mathcal{B}} \Delta_{v,x,b}w_b \quad \forall v \in \mathcal{H},\ x \in \mathcal{X}$$

$$r_{v,x} \leq 0, s_{v,x} \geq 0 \quad \forall v \in \mathcal{H},\ x \in \mathcal{X}.$$

□

A.2 Appendix 2: Proofs of Chapter 3

A.2.1 Proof of Theorem 3.3.1

In model (3.4), constraints (3.5) and (3.6) present an optimal solution to model (3.3). We are looking for a solution that has the maximum distance from $\mathbf{x^0}$. That is, a solution that has maximum $||\mathbf{x^0} - \mathbf{x}||_1 = \sum_{i=1}^{n} |\mathbf{x^0}_i - \mathbf{x}_i|$. Consider $\mathbf{l} = (\mathbf{l}_1, ..., \mathbf{l}_n)$ such that

$$|\mathbf{x^0}_i - \mathbf{x}_i| \geq \mathbf{l}_i \quad \text{for } i = 1, ..., n,$$

that is, either $\mathbf{x^0}_i - \mathbf{x}_i \geq \mathbf{l}_i$ or $\mathbf{x}_i - \mathbf{x^0}_i \geq \mathbf{l}_i$ for $i = 1, ..., n$. Define $\mathbf{z} = (\mathbf{z}_1, ..., \mathbf{z}_n)$ where $\mathbf{z}_i \in \{0, 1\}$ for $i = 1, ..., n$. Therefore, it is easy to see that

$$\mathbf{x^0}_i - \mathbf{x}_i + \mathbf{l}_i \leq M\mathbf{z}_i$$
$$\mathbf{x}_i - \mathbf{x^0}_i + \mathbf{l}_i \leq M(1 - \mathbf{z}_i)$$

for a large enough M. Now, it is sufficient to maximize $\sum_{i=1}^{n} \mathbf{l}_i$ to find $\mathbf{x}$ with maximum distance from $\mathbf{x^0}$, that is, $||\mathbf{x} - \mathbf{x^0}||_1$.

A challenge in model (3.4) is selecting a proper value for M. Suppose that we have an upper bound u_x for each entry of any feasible solution $\mathbf{x}$ to model (3.3); that is, $\mathbf{x}_i \leq u_x$ for any entry $1 \leq i \leq n$ of feasible solution $\mathbf{x}$ to model (3.3). Since a feasible solution to model (3.4) is a feasible solution to model (3.3), we have $\mathbf{x}_i \leq u_x$ for any feasible solution $\mathbf{x}$ to (3.4) and $1 \leq i \leq n$. Also since $\mathbf{x^0}$ is an optimal solution to model (3.3), we have $\mathbf{x^0}_i \leq u_x$ for any $1 \leq i \leq n$. Therefore, $\mathbf{l}_i \leq |\mathbf{x}_i - \mathbf{x^0}_i| \leq u_x$ for $1 \leq i \leq n$, and $\mathbf{x}_i - \mathbf{x^0}_i + \mathbf{l}_i \leq 2u_x$. Thus, it is sufficient to have $M \geq 2u_x$ □

A.2.2 Proof of Corollary 3.3.1

Model (3.7) is an extension of model (3.4). We are looking for a solution that maximizes $m = \min\{||\mathbf{x^k} - \mathbf{x^j}||_1, 0 \leq j \leq k-1\}$. With the same method as model (3.4), we define $\mathbf{l^j}$ and $\mathbf{z^j}$ for each solution $\mathbf{x^j}$, $0 \leq j \leq k-1$. That is, define $\mathbf{l^j} = (\mathbf{l^j}_1, ..., \mathbf{l^j}_n)$ for $0 \leq j \leq k-1$ such that

$$|\mathbf{x^j}_i - \mathbf{x}_i| \geq \mathbf{l^j}_i \quad \text{for } i = 1, ..., n,$$

that is, either $\mathbf{x^j}_i - \mathbf{x}_i \geq \mathbf{l^j}_i$ or $\mathbf{x}_i - \mathbf{x^j}_i \geq \mathbf{l^j}_i$ for $i = 1, ..., n$. Define $\mathbf{z^j} = (\mathbf{z^j}_1, ..., \mathbf{z^j}_n) \in \{0,1\}^n$ for $0 \leq j \leq k-1$ to have

$$\mathbf{x^j}_i - \mathbf{x}_i + \mathbf{l^j}_i \leq M\mathbf{z^j}_i$$

$$\mathbf{x}_i - \mathbf{x^j}_i + \mathbf{l^j}_i \leq M(1 - \mathbf{z^j}_i)$$

It is sufficient to maximize $m = \min\{||\mathbf{x^k} - \mathbf{x^j}||_1, 0 \leq j \leq k-1\}$. Constraint (3.8) and the objective function address this goal. $\square$

A.2.3 Proof of Theorem 3.3.2

First, we demonstrate how we can use inverse optimization to find $\hat{\mathbf{p}} \in \mathcal{U}$. Therefore, we need the inverse model of problem (3.10). First, we have to write model (3.10) as a linear programming. Since

$$\mathcal{X}^{RO} = \{\mathbf{x} \in \mathbb{R}^n | \exists \mathbf{y} \in \mathbb{R}_+^{m_u} \text{ such that } \mathbf{Ax} \leq \mathbf{b}, \mathbf{D'y} - \mathbf{x} = 0, \mathbf{y'd} \geq z^{RO}\},$$

We can wirte model (3.10) as

$$
\begin{aligned}
\underset{\mathbf{x},\mathbf{y}}{\text{maximize}} \quad & \hat{\mathbf{p}}'\mathbf{x} \\
\text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b}, \\
& \mathbf{D}'\mathbf{y} - \mathbf{x} = 0, \\
& \mathbf{y}'\mathbf{d} \geq z^{RO}, \\
& \mathbf{x} \in \mathbb{R}^{n}, \mathbf{y} \in \mathbb{R}_{+}^{m_u}, \mathbf{x} \geq 0
\end{aligned}
\tag{A.38}
$$

Therefore, model (3.10) is a linear programming. As it was mentioned in Section **??**, to present its inverse model, we have to combine both primal and dual model of this problem. The dual of model (3.10) is,

$$
\begin{aligned}
\underset{\mathbf{w},\mathbf{v},\mathbf{s}}{\text{minimize}} \quad & \mathbf{b}\mathbf{s} + z^{RO} w \\
\text{subject to} \quad & \mathbf{s}'\mathbf{A} - \mathbf{v} = \mathbf{p}, \\
& \mathbf{D}\mathbf{v} + \mathbf{w}\mathbf{d} \geq \mathbf{0}, \\
& \mathbf{s} \geq \mathbf{0}, w \leq 0, \mathbf{s} \in \mathbb{R}^{m_x}, \mathbf{v} \in \mathbb{R}^{n}, w \in \mathbb{R}
\end{aligned}
\tag{A.39}
$$

Suppose we have a robust optimal solution $\mathbf{x}^*$, combining the primal model (A.38) and dual problem (A.39), the inverse optimization model for problem (3.10) is as follows for $\mathbf{x}^*$

$$\min_{\mathbf{w},\mathbf{v},\mathbf{s},\mathbf{p}} 0 \tag{A.40a}$$

$$\mathbf{bs} + z^{RO}w - \mathbf{px}^* = 0 \tag{A.40b}$$

$$\mathbf{s}'\mathbf{A} - \mathbf{v} - \mathbf{p} = 0 \tag{A.40c}$$

$$\mathbf{Dv} + \mathbf{wd} \geq \mathbf{0} \tag{A.40d}$$

$$\mathbf{Dp} \geq \mathbf{d} \tag{A.40e}$$

$$\mathbf{s} \geq \mathbf{0}, w \leq 0, \mathbf{s} \in \mathbb{R}^{m_x}, \mathbf{v} \in \mathbb{R}^n, \mathbf{p} \in \mathbb{R}^n, w \in \mathbb{R}.$$

Where constraint (A.40b) is presenting strong duality. Constraints (A.40c) and (A.40d) apply dual feasibility and constraint (A.40e) guarantees that $\hat{\mathbf{p}} \in \mathcal{U}$.

Therefore, using model (A.40), for each RO solution $\mathbf{x}^* \in \mathcal{X}^{RO}$ we can find $\mathbf{p} \in \mathcal{U}$ for which $\mathbf{x}^*$ is an optimal solution to model (3.10). Now, the problem is how we can find a $\mathbf{p}$ as a solution to model (A.40) where $\mathbf{p} \in ri(\mathcal{U})$ for a specific RO solution $\mathbf{x}^*$.

We know $\mathbf{p} \in \mathcal{U}$ if $\mathbf{D}'\mathbf{p} \geq \mathbf{d}$, so $\mathbf{p}$ is in $ri(\mathcal{U})$ if and only if $\mathbf{D}'\mathbf{p} > \mathbf{d}$. Thus, it is enough and sufficient to have some $m > 0$ where

$$\mathbf{Dp} - \mathbf{d} \geq m\mathbf{j}$$

and $\mathbf{j}$ be the vector of all ones. If the optimal objective value to model (3.13) is not positive, then it means that such a positive m does not exist, and so such a $\mathbf{p} \in ri(\mathcal{U})$ does not exist and $\mathbf{x}^*$ is not a PRO one. If the result is positive, then there exists $\mathbf{p} \in ri(\mathcal{U})$ corresponding to $\mathbf{x}^*$ and so $\mathbf{x}^*$ is a PRO solution. $\square$

A.2.4 Proof of Theorem 3.3.3

Assume that there exists an element in $\mathcal{X}^{RO}$ that has not been observed in the algorithm. Suppose we have found the subset $\mathcal{S} = \{\mathbf{x^0}, \mathbf{x^1}, ..., \mathbf{x^q}\} \subset \mathcal{X}^{RO}$ at some iteration in the algorithm and the optimal objective value to model (3.1) is zero. Let $\mathbf{x}^* \in \mathcal{X}^{RO} \setminus \mathcal{S}$.

Consider model (3.7). One of the feasible solutions to this model for the set of optimal solutions $\mathcal{S}$ is

$$\mathbf{x} = \mathbf{x}^*, \quad \mathbf{l^j}_i = |\mathbf{x^j}_i - \mathbf{x}^*{}_i|,$$

and

$$\mathbf{z^j}_i = \begin{cases} 1 & \text{if } \mathbf{x^j}_i - \mathbf{x}^*{}_i > 0 \\ 0 & \text{O.W.} \end{cases}$$

for $0 \leq j \leq q$. Since $\mathbf{x}^*$ is not in $\mathcal{S}$, for all $\mathbf{x^j} \in \mathcal{S}$, we have $||\mathbf{x^j} - \mathbf{x}^*||_1 > 0$. Therefore, the optimal value for m for model (3.7) is greater than zero. It is a contradiction, and so all of the elements in $\mathcal{X}^{RO}$ must be met if the optimal objective value to model (3.7) is zero at some iteration. $\square$

A.2.5 Proof of Corollary 3.3.2

We prove this corollary with contradiction. Suppose that there exists $\tilde{\mathbf{x}} \in \mathcal{X}^{RO}$ that is not generate with the algorithm. Thus, there must be a loop in this algorithm, and we only meet the elements in a set $\bar{\mathcal{X}} \in \mathcal{X}^{RO}$, and $\bar{\mathcal{X}} \neq \mathcal{X}^{RO}$. Since it is a loop in the algorithm, after generating all elements in $\bar{\mathcal{X}}$, we must generate an element in $\bar{\mathcal{X}}$ again. Therefor, the solution m for model (3.7) will be zero. Using Theorem 3.3.3, the algorithm must generate all elements in $\mathcal{X}^{RO}$ if it returns zero for the value of m. Thus, it is a contradiction and the algorithm generates all elements in $\mathcal{X}^{RO}$. $\square$

A.3 Appendix 3: Proofs of Chapter 4

A.3.1 Proof of Theorem 4.2.1

The method to prove this proposition is similar to the proof of the theorem that presented a PRO solution for this problem [Iancu and Trichakis, 2014]. Let $\mathbf{x}$ be a solution to problem (4.2). Suppose that there exists a light RO solution $\mathbf{x}^*$ for model (4.1) that dominated $\mathbf{x}$. Thus,

$$\mathbf{p}'\mathbf{x}^* \geq \mathbf{p}'\mathbf{x}, \quad \forall \mathbf{p} \in \mathcal{U}, \text{ and}$$
$$\bar{\mathbf{p}}'\mathbf{x}^* > \bar{\mathbf{p}}'\mathbf{x}, \quad \text{ for some } \bar{\mathbf{p}} \in \mathcal{U}.$$

Suppose, $\exists \bar{\mathbf{p}} \in ext(\mathcal{U})$ such that $\bar{\mathbf{p}}\mathbf{x}^* > \bar{\mathbf{p}}\mathbf{x}$, where $\bar{\mathbf{p}}$ is a solution to $\max_{\mathbf{p}\in\mathcal{U}} \mathbf{p}(\mathbf{x}^* - \mathbf{x})$. Since $\hat{\mathbf{p}} \in ri(\mathcal{V}_1)$, we can write $\hat{\mathbf{p}}$ as a convex combination of all extreme points of $\mathcal{U}$. That is, $\exists \lambda \in \mathbb{R}^{|ext(\mathcal{U})|}$ such that $\lambda > \mathbf{0}$ and $\mathbf{1}'\lambda = 1$ and $\hat{\mathbf{p}}_\mathbf{1} = \sum_{i\in\mathcal{E}} \mathbf{p}^i\lambda_i$ where $ext(\mathcal{U}) = \{\mathbf{p}^i | i \in \mathcal{E}\}$. Thus,

$$\hat{\mathbf{p}}(\mathbf{x}^* - \mathbf{x}) = \sum_{i\in\mathcal{E}} \mathbf{p}^i(\mathbf{x}^* - \mathbf{x}) = \bar{\mathbf{p}}(\mathbf{x}^* - \mathbf{x}) + \sum_{i\in\mathcal{E}, \mathbf{p}^i\neq\bar{\mathbf{p}}} \mathbf{p}^i(\mathbf{x}^* - \mathbf{x}).$$

Therefore, $\hat{\mathbf{p}}(\mathbf{x}^* - \mathbf{x}) > 0$ and so $\hat{\mathbf{p}}\mathbf{x}^* > \hat{\mathbf{p}}\mathbf{x}$ and $\mathbf{x}$ is not an optimal solution to model (4.2). It is a contradiction, and so $\mathbf{x}$ is not dominated by another light RO solution. Thus, $\mathbf{x}$ is a light PRO solution. $\square$

A.3.2 Proof of Proposition 4.3.1

Consider model

$$\max_{\mathbf{x}\in\mathcal{X}^{RO}} \hat{\mathbf{p}}'\mathbf{x}. \tag{A.41}$$

Since z^{RO} is the optimal objective value to model (4.1), then we have $\mathbf{y}'\mathbf{d} = z^{RO}$ for a solution to model (A.41). Thus, the constraint $\mathbf{y}'\mathbf{d} \geq z^{RO}$ is binding, and the solution will change by deducting ϵ from the raight hand side of this constraint. Depending on the shadow price of this constraint, the value of $\hat{\mathbf{p}}'\mathbf{x}$ might not change. Assuming the shadow price is non-zero, the optimal objective value to model (A.41) will increase by increasing ϵ. Let the shadow price be δ, the cost will improve by $\delta(\epsilon)$. Thus, $\max_{\mathbf{x}\in\mathcal{X}^{RE}_{\epsilon_1}} \hat{\mathbf{p}}'\mathbf{x} \geq \max_{\mathbf{x}\in\mathcal{X}^{RE}_{\epsilon_2}} \hat{\mathbf{p}}'\mathbf{x}$ if $\epsilon_1 > \epsilon_2$. If we consider the function $f_{\hat{\mathbf{p}}}(\epsilon)$ for different ϵs, we have

$$f(\epsilon_1; \hat{\mathbf{p}}) \geq f(\epsilon_2; \hat{\mathbf{p}}) \quad if \quad \epsilon_1 > \epsilon_2,$$

and so the function $f(\epsilon; \hat{\mathbf{p}})$ is an increasing function.

$\square$