

Intensity Modulated Proton Therapy Optimization Under
Uncertainty: Field Misalignment and Internal Organ Motion

A Dissertation

Presented to

the Faculty of the Department of Industrial Engineering

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

in Industrial Engineering

by

Li Liao

December 2016

Intensity Modulated Proton Therapy Optimization Under Uncertainty: Field Misalignment and Internal Organ Motion

Li Liao

Approved:

Chair of the Committee
Gino Lim, Professor,
Industrial Engineering

Committee Members:

Qianmei Feng, Associate Professor,
Industrial Engineering

Jiming Peng, Associate Professor,
Industrial Engineering

Xiaodong Zhang, Associate Professor,
Radiation Physics, The University of
Texas MD Anderson Cancer Center

X.Ronald Zhu, Professor,
Radiation Physics, The University of
Texas MD Anderson Cancer Center

Suresh K. Khator, Associate Dean,
Cullen College of Engineering

Gino Lim, Professor and
Chair, Industrial Engineering

Acknowledgements

First and foremost I would like to express my deepest appreciation to my advisor Professor Dr. Gino Lim. His advice on both research as well as on my career has been priceless. Without his guidance and persistent help, this dissertation would not have been possible. I especially would like to thank Dr. Xiaidong Zhang (Department of Radiation Physics, The University of Texas MD Anderson Cancer Center) for providing me the opportunity to work in the greatest groups at the best cancer hospital. His knowledge and insight inspired and motivated me.

Many thanks to Dr. Qianmei Feng, Dr. Jiming Peng and Dr. X.Ronald Zhu for serving as my committee members. Thanks for their time, insightful suggestions and providing valuable comments that improved the contents of the work.

I would also like to thank all the faculties, staff and friends who have contributed immensely to my work and life at the University of Houston.

A special thanks to my family. Words cannot express how grateful I am to my mother and parents-in-law for all of the sacrifices that they've made on my behalf. Lastly, I would like express appreciation to my beloved wife Ting who spent sleepless nights with and was always my support. Their support made me devote my time to completing my Ph.D. study. This dissertation is dedicated to them.

Intensity Modulated Proton Therapy Optimization Under
Uncertainty: Field Misalignment and Internal Organ Motion

An Abstract
of a
Dissertation
Presented to
the Faculty of the Department of Industrial Engineering
University of Houston

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
in Industrial Engineering

by
Li Liao
December 2016

Abstract

Intensity modulated proton therapy (IMPT) is one of the most advanced forms of radiation therapy, which can deliver a highly conformal dose to the tumor while sparing the dose in healthy tissues. Compared to conventional photon-based radiation therapy, IMPT is more flexible in delivering radiation dose according to different tumor shapes. However, this flexibility also makes the optimization problems in IMPT harder to solve, e.g., it requires larger memory to store data and longer computational time. Furthermore, proton beams are very sensitive to different uncertainties, such as setup uncertainty, range uncertainty and internal organ motion. These uncertainties can greatly impact the quality of clinical treatment. Therefore, this dissertation aims to investigate different optimization methods for treatment planning and to handle a variety of uncertainties in IMPT.

First, to solve the fluence map optimization (FMO) problem in IMPT, we propose a method to formulate the FMO problem into a molecular dynamics model. So that, the FMO problem can be optimized according classical dynamics system. This method combines the advantages of gradient-based algorithms and heuristic search algorithms.

Next, we develop and validate a robust optimization method for IMPT treatment plans with multi-isocenter large fields to overcome the dose inhomogeneity problem caused by the setup misalignment in field junctions. Numerical results show that the robust optimized IMPT plans create a low gradient field radiation dose in the junction regions, which can minimize the impact from misalignment uncertainty. Compare to conventional techniques, the robust optimization method leads the whole treatment much more efficient.

Lastly, we focus on a two-stage method to solve the beam angle optimization (BAO) problem in IMPT with internal organ motion uncertainty. In the first stage, a p -median algorithm is developed for beam angle clustering. In the second stage, a

bi-level search algorithm is used to find the final beam angle set for the treatment. Furthermore, Support vector machine (SVM) is used for beam angle classification to reduce the search space and the 4D-CT information is incorporated to handle the internal organ motion uncertainty. Results show that the two-stage BAO method consistently finds a high-quality solution in a short time.

Table of Contents

Acknowledgements	iv
Abstract	vi
Table of Contents	viii
List of Figures	xi
List of Tables	xv
Chapter 1 Introduction	1
1.1 Background	1
1.2 Problem Statement	8
1.3 Objectives & Contributions	12
1.4 List of Publications	14
1.5 Organization	16
Chapter 2 Literature Review	18
2.1 Fluence Map Optimization Problem	18
2.2 Robust Optimization for Radiation Therapy	20
2.3 Beam Angle Optimization Problem	22
Chapter 3 A Molecular Dynamics Method for Fluence Map Opti- mization in Intensity Modulated Proton Therapy	26
3.1 Introduction	27

3.2	Optimization Model and Solution Algorithms	29
3.2.1	FMO Problem Formulation	29
3.2.2	Solution Algorithms	32
3.3	Numerical Experiments	42
3.3.1	Patient Data	42
3.3.2	IMPT Starting Conditions	44
3.4	Results	45
3.4.1	Solution Quality	45
3.4.2	Computational Performance	48
3.5	Conclusion	50
Chapter 4 Robust Optimization for Intensity Modulated Proton Ther-		
apy Plans with Multi-Isocenter Large Fields		52
4.1	Introduction	53
4.2	Material and Methods	54
4.2.1	Field Setup and Spot Arrangement	55
4.2.2	Robust Optimization and Uncertainty Setup	57
4.2.3	Dosimetric Evaluation	58
4.2.4	Plan Robustness Evaluation	58
4.3	Results	58
4.4	Discussion	63
4.5	Conclusion	66
Chapter 5 Two-stage Method for IMPT Beam Angle Optimization		
Incorporating Internal Organ Motion		67
5.1	Introduction	68

5.2	Material and Methods	71
5.2.1	4D Robust Fluence Map Optimization Model	72
5.2.2	Beam Angle Clustering	74
5.2.3	Bi-level Local Neighborhood Search (Bi-LNS) for BAO	79
5.2.4	Patient Studies and Setup	83
5.3	Results	86
5.4	Discussion	91
5.5	Conclusion	96
	Chapter 6 Summary and Future Work	97
6.1	Current Findings	97
6.2	Future Work	99
	References	101

List of Figures

Figure 1.1	Depth-dose curves of a photon beam (red), a proton spread-out Bragg peak (blue, thick), and the proton pencil beams constituting the spread-out Bragg peak (blue, thin).	5
Figure 1.2	Proton pencil beam scanning (a) and passive scattering proton therapy (b)	6
Figure 1.3	Example of DVHs	7
Figure 1.4	The distortion of proton dose distribution in the lung with respiratory motion. (a) planned dose distribution. (b) dose distribution of the same spots with respiratory motion.	11
Figure 3.1	The three clinical cancer cases selected for the study and their corresponding field directions. (A) Prostate cancer case, (B) head-and-neck cancer case, (C) lung cancer case; the tumors are contoured in red.	42
Figure 3.2	The initial beamlet weights for a single intensity modulated proton therapy field from three initial conditions: (A) Forward wedge, (B) Inverse wedge, (C) Spread-out Bragg peak (SOBP).	45
Figure 3.3	The dose-volume results of the prostate case using quasi-Newton, L-BFGS, L-BFGS-B and the MD method starting from three initial conditions: forward wedge (FW), inverse wedge (IW) and spread-out Bragg peak (SOBP).	47

Figure 3.4	The objective function value as a function of time for the optimization processes for the prostate case starting from FW, IW and SOBP using quasi-Newton, L-BFGS, L-BFGS-B and the MD method.	49
Figure 4.1	Field arrangement for the craniospinal irradiation patient (A-F) and mesothelioma irradiation patient (G-J).	56
Figure 4.2	Dose color wash and corresponding dose profiles of the robust and conventional IMPT plans for the craniospinal irradiation patient (A, B) and the mesothelioma patient (C, D).	59
Figure 4.3	Dose profiles in junctions for the CSI IMPT plans with junction sizes of 8, 12, 16 and 26 cm and a longitudinal misalignment error of 3 mm per field (total, 6 mm).	60
Figure 4.4	Robustness comparison between a robust IMPT plan with a large dose junction (18 cm) and a robust IMPT plan with a small dose junction (7 cm) and junction shifting for the CSI patient.	61
Figure 4.5	Dose volume histograms of robust and non-robust IMPT plans for craniospinal irradiation patient (A) and mesothelioma patient (B). Solid lines: robust IMPT plan; Dashed lines: non-robust IMPT plan.	62
Figure 4.6	Dose volume histograms of robust IMPT plan with fixed brain fields uncertainty setup and non-robust IMPT plan for craniospinal irradiation patient. Solid lines: robust IMPT plan; Dashed lines: non-robust IMPT plan.	65
Figure 5.1	Normalized p-median objective function value (blue bar) and slope (orange dotted line) as a function of cluster number M.	78

Figure 5.2	Beam angle score of 36 candidate beam angles (dotted line), the centroid beam angles (triangle) and the solution found by cluster level LNS (black dot) for lung case I.	86
Figure 5.3	Beam angle appearance probability in the best 200 IMPT plans (A), and the histogram of average beam angle appearance probability (B) for lung case I.	87
Figure 5.4	Centroid distributions and final solutions obtained by TSBAO method for 4 clinical cases. Centroids are denoted by dashed lines and TSBAO solutions are solid lines.	89
Figure 5.5	Converge comparison for TSBAO (solid lines), standalone LNS (dashed lines), SA (cross solid lines), SA-LNS (triangle solid lines), GA (asterisk solid lines), and GA-LNS (circle solid lines). Global optima are shown by dotted lines.	90
Figure 5.6	Beam angle configurations selected for the Lung case I by dif- ferent algorithms. Equal spaced beam angles are denoted by dashed lines, the final solutions by green solid lines and the optimal solution by red solid lines.	93
Figure 5.7	Beam angle configurations found by cluster method and global optimal for esophagus case II.	94
Figure 1	Dose-volume histogram comparison for the head-and-neck case.	118
Figure 2	Dose-volume histogram comparison for the lung case.	119
Figure 3	The OFVs as a function of time for the optimization processes for the head-and-neck case starting from FW, IW and SOBP using quasi-Newton, L-BFGS, L-BFGS-B and the MD method.	120

Figure 4 The OFVs as a function of time for the optimization processes
for the lung case starting from FW, IW and SOBPs using quasi-Newton,
L-BFGS, L-BFGS-B and the MD method. 120

List of Tables

Table 3.1	The intensity modulated proton therapy beam angles, number of beamlets in each beam, VOIs and number of voxels of each VOI for the three cancer cases	43
Table 3.2	Dose-based objective function parameters used for optimizing the intensity modulated proton therapy plans	44
Table 3.3	Objective function value comparison using quasi-Newton, L-BFGS, L-BFGS-B and the MD methods for three tested cases starting from forward wedge (FW), inverse wedge (IW) and spread-out Bragg peak (SOBP).	46
Table 3.4	Time comparison using quasi-Newton, L-BFGS, L-BFGS-B and the MD methods for three tested cases starting from forward wedge (FW), inverse wedge (IW) and spread-out Bragg peak (SOBP).	48
Table 5.1	Prescriptions and penalty weights used for IMPT plan optimization.	84
Table 5.2	Number of voxels within each structures for different cases	84
Table 5.3	Stopping criteria for different BAO algorithms	85
Table 5.4	Dose volume data of 3 beam IMPT plans from different BAO methods for four clinical cancer cases.	92

Chapter 1

Introduction

1.1 Background

Cancer is a fatal disease that accounts for nearly one fourth of total deaths in the United States. A total of 1,658,210 new cancer cases are estimated to be diagnosed and 595,690 cancer deaths in the United States in 2016 [1]. There are different types of cancer treatment, such as surgery, chemotherapy and radiation therapy.

Radiation therapy: Radiation therapy is one of the most common treatments for many types of cancer. About 60% of cancer cases receive radiation therapy during their treatment [2]. Radiation therapy uses controlled high-energy radiation to damage cancer cells' DNA and destroy their ability to divide and grow. Abnormal cancer cells are more sensitive to radiation because they divide more quickly than normal cells. Over time, the abnormal cells die and the tumor shrinks. Since radiation can damage both cancer cells and healthy cells, the goal of radiation therapy is to deliver a prescribed dose of radiation to the tumor in order to kill or control the growth of cancerous cells, while avoiding the delivery of excessive doses of radiation to surrounding critical organs and healthy tissues. According to the type and stage of cancer, radiation therapy is used both as a stand-alone treatment and in combination with other cancer treatments such as surgery and chemotherapy.

There are two ways to deliver radiation: external beam radiation and internal radiation (brachytherapy). External beam radiation is delivered from outside the body by using a machine to aim high energy rays at the tumor. Different types of radiation are used for external beam therapy, such as x-ray, gamma rays, photons and proton beams. External beam therapy is the radiation therapy treatment option used for most cancer patients. Internal radiation (brachytherapy) is delivered from inside the body by placing radiation sources close to or inside the tumor. The radioactive sources or isotopes are in the form of wires, seeds (or molds), or rods.

Treatment planning: To ensure the patients can get the full benefit from radiation therapy while minimizing the impact on healthy organs, a careful planning is critical before treatment. The main steps of radiation treatment planning can be described as follows. In the beginning, the patients will be positioned carefully to keep the same position during every treatment. To stabilize the position, a variety of immobilization devices may be used, e.g., for a specific patient, a foam box shaped to patient form will be used to keep the body position; a thermoplastic mask may be designed to hold the patient's head in place. Then, the computed tomography (CT) images of the area of treatment will be taken to identify the internal structures of interest. Sometimes, the other image techniques such as magnetic resonance imaging (MRI) and positron emission tomography (PET) are also used to get more precise images to identify the structures. After images are taken, a physician will delineate the target tumor and the surrounding critical organs, also referred to as organs-at-risk (OARs), which desire to spare. Along with the contour of these structures, the physician will

also prescribe objective doses to the target volumes and constraints to OARs. The treatment planner will base on all of these information to generate a treatment plan to the patient.

Typically, the whole radiation therapy treatment is divided into many treatment sessions, known as "fractions", with daily intervals or breaks for the weekend. One small portion of total prescribed radiation dose is to be delivered in each fraction. The whole treatment may last for 4 to 6 weeks. The reason of fractionation is based on the radiobiological effects of the dose on the healthy and cancerous cells. Compare to healthy cells, the cancer cells generally have much less ability to repair DNA damages caused by radiation exposure. Therefore, by splitting the dose to many treatment fractions, the healthy cells can repair the damage and recover between treatment fractions, but the accumulated radiation dose can lead to lethal damages to tumor cells. This mechanism enables patients to tolerant a higher total radiation dose to expect a better treatment outcome.

Generally, there are two types of radiation treatment planning process: forward planning and inverse planning. Forward planning is often applied for three-dimensional conformal radiation therapy (3DCRT) and passive scattering proton therapy (PSPT). In which, treatment planner specifies the directions, shapes and intensity of the beams then calculate and evaluate the plan quality. If the plan is not meet the clinical requirements, the planner repeats the process until the requirements are satisfied. The whole process of forward planning is very tedious and time consuming, and the quality of treatment plan highly depends on the experience of

the treatment planner. Due to the new technology development, radiation therapy delivery modalities become more flexible and precise, such as intensity modulated radiation therapy (IMRT) and intensity modulated proton therapy (IMPT). In these radiation treatment plans, there are much more parameters need to consider during the planning. The forward planning process is not applicable anymore. Instead, a variety of optimization models were developed to solve the problems for different radiation treatment delivery modalities. Planner specifies the desired requirement, such as a prescribed dose of the tumor, max tolerance of OARs into the model, the optimization algorithms will try to determine all the parameters to achieve the requirements as good as possible. This type of treatment planning is called inverse planning. In this thesis, we focus on inverse treatment planning.

Intensity Modulated Proton Therapy: After decades of development, the proton therapy is widely adopted to treat cancer patients over the world. Compare to conventional photon-based radiotherapy, proton therapy uses charged proton beams that have a very attractive physical characteristics for radiation therapy (see Figure 1.1). First, the deposited dose of a proton beam starts from a low entrance level, its energy increases gradually while increasing depth, then suddenly jumps to a sharp peak known as the Bragg peak. Once the dose deposition reaches a few millimeters beyond this peak, it falls sharply to zero. Therefore, proton beams may deliver nearly no dose to regions beyond the target, which is typically not feasible for photon beams. Second, the depth of the Bragg peak can be controlled by alteration of the energy of the incident protons. This amounts to an additional degree of freedom as compared

to photon therapy. The superposition of pencil beams of different energies allows for spread-out Bragg peaks that cover the full target volume in depth. So, protons therapy can achieve a more conformal high dose to the tumor regions and a better dose sparing to the normal tissue region than photon-based radiation therapy modalities.

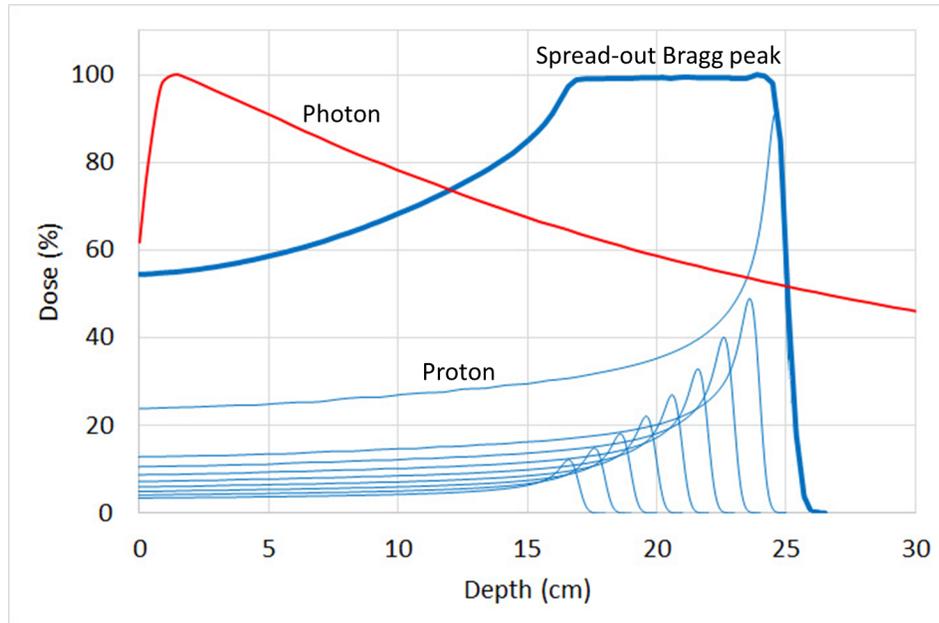


Figure 1.1: Depth-dose curves of a photon beam (red), a proton spread-out Bragg peak (blue, thick), and the proton pencil beams constituting the spread-out Bragg peak (blue, thin).

In proton therapy, a particle accelerator is used to deliver a beam of protons to the tumor. Currently, the passive scattering proton therapy (PSPT) and intensity modulated proton therapy (IMPT) or pencil beam scanning are two available proton beam delivery modalities. In PSPT, the proton beam is spread by placing scattering material into the beam path and shaped by aperture and compensator (Figure 1.2b). In IMPT, the proton beams are delivered as narrow scanning pencil beams, also called beamlets. The tumor target volume is divided into multiple scanning spots (Figure

1.2a). These three-dimensional (3D) arrangement of scanning spots is achieved by controlling the energy level of the proton beamlet (determining the depth of a spot from patient surface) and the intensity scanning magnets (determining the lateral position of a spot). Beamlets with the same energy level (or depth) are located in one layer, often called the energy layer. The intensity of a beamlet, i.e., beamlet weight, is controlled by the exposure time. A higher intensity of a beamlet, i.e., longer beam on time, results in a higher radiation dose deposited on a specific spot. So, compare to PSPT, IMPT is more flexible to generate dose distribution according to the different shape of tumors.

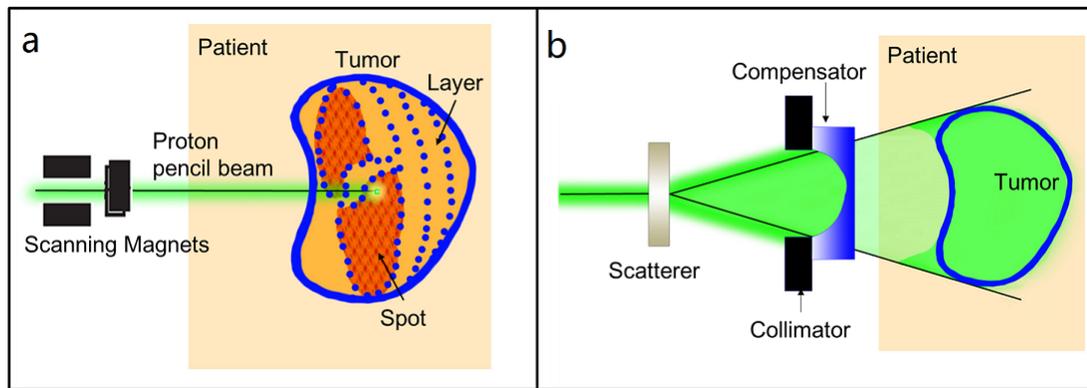


Figure 1.2: Proton pencil beam scanning (a) and passive scattering proton therapy (b)

Evaluation of treatment plan quality: The primary method to evaluate the quality of a radiation treatment plan is to analyze the resulting dose distribution associated with anatomical images. The test can be performed by checking 2D or 3D dose distribution to evaluate if coverage of target is actually adequate; or by verifying the mean dose or point dose of critical organs to assure the healthy tissue are well protected.

In treatment planning, dose volume histogram (DVH) is an important tool to evaluate the plan quality [3]. Many important dose distribution indices of a region of interest (ROI) can be easily evaluated by inspection of its cumulative DVH. DVH is also a valuable tool for treatment plan comparison for a specific patient. Examples of DVHs are shown in Figure 1.3.

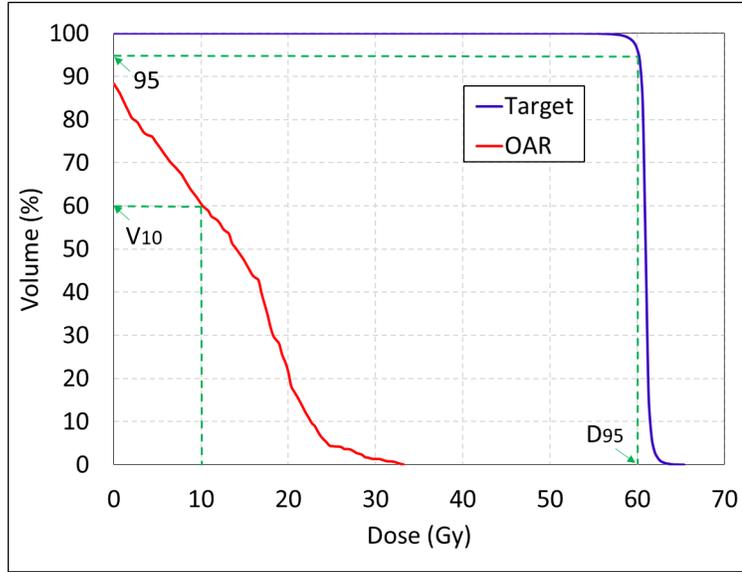


Figure 1.3: Example of DVHs

In DVH, the D_v is represent the dose level d , such that a given $v\%$ volume of an ROI receives d Gy or higher dose and V_d is represent the percent of the volume of an ROI $v\%$, such that $v\%$ volume of an ROI receives a given d Gy or higher dose. There are some DVH indices that commonly used to evaluate plan quality, e.g., D_{100} and D_0 are used to present minimum and maximum dose; D_{50} is a median dose of ROI. Based on DVH, people also derived some useful indices, such as homogeneity index which equals $(D_5 - D_{95}) / D_{mean}$, and the ratio between the patient volume that receives 95%

of the prescription and the target volume used as conformity index [4]. DVH values also used as input data to predict the biological outcome from radiation therapy. For example, the tumor control probability (TCP) and normal tissue complication probability (NTCP) [5, 6]. TCP stands the probability that a given radiation dose will kill or control the tumor cells and NTCP stands the probability that an organ or structure to have a complication with a given dose of radiation. Note that many of the parameters of TCP and NTCP models, and in fact, the models themselves, are still under investigation, and may be the subject of significant controversy. In addition, DVH clusters and band graph which consist of a group of DVHs in a variety of are always used to evaluate the impact of uncertainties to the radiation treatment plans.

1.2 Problem Statement

In radiation therapy, the fundamental goal is to deliver a prescribed dose to cover the target while sparing radiation on the surrounding OARs. These two goals are inherently contradictory if the targets and critical structures are near each other or overlapping. So, the whole treatment planning can be considered as an optimization problem of balancing these two objectives. In IMPT, different procedures of treatment planning can be formulated as different sub-problems. Generally, the beam angle optimization (BAO) and the fluence map optimization (FMO) are two major optimization problems in IMPT treatment planning.

Since the radiation would damage healthy cells in OARs which located along

the path of the external beam. To avoid the radiation dose deposition in the OARs exceeds its tolerance, treatment plans are always designed to deliver radiation from a number of different angles around the patient. Multiple beams also can lead a more uniform dose coverage on tumor than a single beam. So, selection of suitable beams is critical to making a high quality treatment plan. The problem of choosing beam angle is called beam angle optimization problem. However, the beam angle selection is typically a large-scale combinatorial optimization problem and the whole process can be very time consuming. In clinical practice, the beam angle selection is still based on the treatment planner's experience and intuition. So, efficient methods for BAO is an important topic and still under investigation.

In IMPT, a beam of radiation consists of thousands of beamlets, each one has an independently adjustable intensity. The intensity map of all beamlets from all beam directions is also called fluence map. The problem of adjusting beamlets intensity profiles to ensure the treatment plan can deliver the radiation to the tumor and avoid neighboring critical structures is fluence map optimization (FMO) problem. In this problem, the beamlets intensities are the decision variables. Due to a large number of decision variables, the FMO problem is always solved by computers. So, the final result quality directly relies on the mathematic models and the optimization approaches.

Uncertainty issue in radiation therapy: During radiation treatment process, many uncertainties can come from different procedures of radiation treatment process. These uncertainties can be classified as follows:

- Setup error: this error is due to the misalignment of incident beams and the patient anatomy during day to day treatment.
- Range error: range error arises from multiple sources, such as CT image error, CT number to stopping power conversion error, patient gain or loss weight and tumor shrinkage.
- Intrafractional organ motion: This error is caused by the internal motion of organs and tissues in a human body during a treatment session. For example, in the proton therapy of lung cancer and esophageal cancer, the respiratory motion may cause significant changes in patient geometry.

Since treatment plans are designed based on the planning CT images and assumed to be identical with the patient geometry during the treatment. So, all the uncertainties mentioned above may cause the delivered dose to seriously deviate from the planned dose distribution and lead to some unforeseen results. To reduce the uncertainties in radiation therapy, different strategies have been proposed and applied in clinical treatment.

The combination of imaging and immobilization devices have been commonly used to detect and mitigate setup errors [7, 8, 9, 10, 11]. Gating and breath holding techniques are used to reduce the impact of respiratory motion [12, 13, 14, 15]. However, all these methods require extra devices and usually technologically demanding and may extend the treatment time.

In conventional photon-based radiotherapy (e.g., IMRT and 3DCRT), the uncertainties can be accounted by adding margins, i.e., a margin added around to the

clinical target volume (CTV) to form a planning target volume (PTV) to ensure CTV can receive the prescribed dose in the presence of uncertainties. The margin approach works well in conventional photon-based therapy, due to the static dose cloud approximation, i.e., photon dose is robust to the anatomy change on the beam path [16]. However, compare to photon-based radiation therapy, proton therapy is even more vulnerable to the uncertainties. Because of the physical character of the proton beam, the position of the Bragg peak is highly sensitive to the traversed medium on proton beam path. Combined with proton pencil beams can shape very accurate dose with a sharp gradient. The geometric changes may cause significant deformation of the proton dose distribution (Figure 1.4), especially for the case with large internal motion such as lung cancer and esophageal cancer. To handle the complex

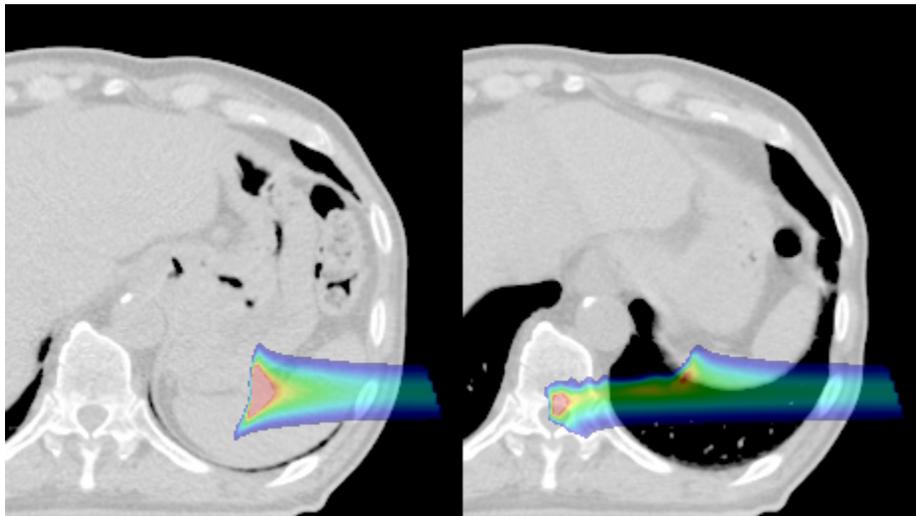


Figure 1.4: The distortion of proton dose distribution in the lung with respiratory motion. (a) planned dose distribution. (b) dose distribution of the same spots with respiratory motion.

uncertainties in radiation therapy, robust optimization is introduced to incorporate

different errors into the optimization process to improve the robustness of treatment plans. Many robust optimization models for photon-based radiation treatment planning were developed by researchers. Worst case based and probabilistic based robust optimization methods are the two major groups used to describe the uncertainties in radiation therapy. Either linear programming (LP) or non-linear programming (NLP) model can be used to handle this problem. Different researchers also reported using robust optimization to handle different uncertainties [17, 18, 19, 20, 21]. Due to the proton beam is sensitive to uncertainties, the robust optimization for IMPT treatment planning is even more critical.

1.3 Objectives & Contributions

This dissertation aims to develop new methodologies to handle the complex problems (fluence map optimization, uncertainty issue and beam angle optimization) in intensity modulated proton therapy treatment planning. The contributions of this dissertation research are listed as follows:

- **Objective 1:** We develop a molecular dynamics method for solving the FMO problem in intensity modulated proton therapy. This method combines the advantages of global and local search algorithms to overcome the local entrapment issue observed in many gradient-based algorithms that are extensively used in radiotherapy planning systems in clinics. This approach is a good alternative method of gradient-based algorithms to solve the FMO problem and consistently produces high-quality treatment solution in a clinically required time frame.

- **Objective 2:** Conventional treatment planning process for multi-isocenter large field patient case is a complex process. To improve the efficiency of treatment planning for these cases, we propose a new robust optimization approach to handle the uncertainty issue in IMPT treatment planning of the multi-isocenter large field patient case. The robust optimized IMPT plans can incorporate filed misalignment uncertainty in the treatment process. As a result, it can easily generate low gradient field dose in the junction region to minimize the dose deviation of uncertainties. This approach can greatly reduce the complexity of treatment planning for the multi-isocenter large field patient case.
- **Objective 3:** Requiring excessive amount of time and easily trapping in local minimum are the main drawbacks of conventional beam angle optimization algorithms. To overcome these problems, we investigate a two-stage robust beam angle optimization method in IMPT treatment planning for thoracic cancer. We explore the prior knowledge from dose deposition information to develop a beam angle score function to evaluate the merit of beam angles to guide the beam angle selection. Clustering technique is utilized to group beam angles and shrink beam angle search space. We designed a bi-level local neighborhood search algorithm to search the final beam angle set for the treatment. Support vector machine is also used in to reduce the search space. This algorithm provides consistent high quality solutions and outperforms other methods in computational time.

1.4 List of Publications

- Journal Publications

- Li Liao, Gino J Lim, Yupeng Li, Juan Yu, Narayan Sahoo, Heng Li, Michael Gillin, X Ronald Zhu, Anita Mahajan, Steven J Frank, David R Grosshans, Quynh-Nhu Nguyen, Daniel Gomez, Xiaodong Zhang, "Robust Optimization for Intensity Modulated Proton Therapy Plans with Multi-Isocenter Large Fields", International Journal of Particle Therapy, 2016
- Li Liao, Gino J. Lim, Xiaodong Zhang, "A Molecular Dynamics Method for Optimizing Beam Intensities in IMPT Treatment Planning", Physics in Medicine and Biology, under revision.
- Li Liao, Gino J. Lim, Nasrin Nouri, Xiaodong Zhang, "A Two-stage Method for IMPT Robust Beam Angle Optimization for Thoracic Cancer Incorporating Internal Organ Motion", Operations Research for Healthcare, under revision.
- Jen Yu, Xiaodong Zhang, Li Liao, Heng Li, Ronald Zhu, Peter C Park, Narayan Sahoo, Michael Gillin, Yupeng Li, Joe Y Chang, Ritsuko Komaki, Steven H Lin, Motion-robust Intensity Modulated Proton Therapy for Distal Esophageal Cancer, Medical physics, 2016
- Shengpeng Jiang, Jingqian Wang, Heng Li, Li Liao, Yupeng Li, Xiaochun Wang, Yining Yang, Ronald X Zhu, Narayan Sahoo, Michael T Gillin, Yoshifumi Hojo, Jian Sun, Joe Y Chang, Zhongxing Liao, David Grosshans, Steven J Frank, Xiaodong Zhang, "Novel Hybrid Scattering-and Scanning-Beam Proton Therapy Approach", International Journal of Particle Therapy, 2016

- Yupeng Li, Perttu Niemela, Li Liao, Shengpeng Jiang, Heng Li, Falk Poenisch, X Ronald Zhu, Sami Siljamaki, Reynald Vanderstraeten, Narayan Sahoo, Michael Gillin, Xiaodong Zhang, "Selective robust optimization: A new intensity modulated proton therapy optimization strategy", *Medical physics*, 2015
- Yupeng Li, Laleh Kardar, Xiaoqiang Li, Heng Li, Wenhua Cao, Joe Y. Chang, Li Liao, Ronald X. Zhu, Narayan Sahoo, Gillin Michael, Gino Lim, and Xiaodong Zhang, "On the interplay effects with proton scanning beams in stage III lung cancer", *Medical Physics*, 41(2), <http://dx.doi.org/10.1118/1.4862076>, February, 2014.
- Laleh Kardar, Yupeng Li, Xiaoqiang Li, Heng Li, Wenhua Cao, Joe Y. Chang, Li Liao, Ronald X. Zhu, Narayan Sahoo, Michael Gillin, Zhongxing Liao, Ritsuko Komaki, James D. Cox, Gino J. Lim, and Xiaodong Zhang, "Evaluation and Mitigation of the Interplay Effects for Intensity Modulated Proton Therapy for Lung Cancer in a Clinical Setting", *Practical Radiation Oncology*, 2014.
- Wenhua Cao, Gino. Lim, Li Liao, Yupeng. Li, Shengpeng Jiang, Xiaoqiang Li, Kazumichi Suzuki, X. Ronald Zhu, Daniel Gomez, Xiaodong Zhang. "Proton energy optimization and reduction for intensity modulated proton therapy", *Physics in Medicine and Biology*, 2014.

- Conference Presentation

- Li Liao, Shengpeng Jiang, Yupeng Li, Xiaochun Wang, Heng Li, X.Ronald Zhu, Narayan Sahoo, Michael Gillin, Anita Mahajan, David Grosshans, Gino Lim, Xiaodong Zhang, "Robust Intensity Modulated Proton Therapy Plan Can

Eliminate Junction Shifts for Craniospinal Irradiation", AAPM, Austin, TX, June 2014

- Li Liao, Gino J. Lim, and Wenhua Cao, "A molecular dynamics approach for optimizing beam intensities in IMPT treatment planning", INFORMS Annual Meeting, Minneapolis, MN, October 2013. - **Posters**
- Li Liao, Juan Yu, Yupeng Li, Ronald Zhu, Heng Li, Gino J. Lim and Xiaodong Zhang, "4DCT robust optimization for esophageal cancer using intensity modulated proton therapy", AAPM, Anaheim,CA, July 2015.
- Juan Yu, Li Liao, Narayan Sahoo, X.Ronald Zhu, Michael Gillin, Xiaodong Zhang, "Experimental Verification of Robustness Optimized Intensity Modulated Proton Therapy Plans for Craniospinal Irradiation", AAPM, Anaheim,CA, July 2015.
- Shengpeng Jiang, Yining Yang, Li Liao, Xiaochun Wang, Heng Li, Xiaorong Ronald Zhu, Xiaodong, "A Methodology for Quality Control of IMPT Treatment Plan based on VMAT Plan", AAPM, Anaheim,CA, July 2015.

1.5 Organization

This dissertation is organized as follows. Chapter 2 provides a comprehensive literature review for the related research on 1) the optimization methods for the fluence map optimization problem, 2) uncertainty problem of intensity modulated proton therapy and the related robust optimization methods 3) beam angle optimization of radiation therapy. In Chapter 3, we present a molecular dynamics method for

solving the fluence map optimization problem in intensity modulated proton therapy. Compare to conventional gradient based method which may yield different results when starting with different initial points, the molecular dynamics method can consistently produce solutions that are the same or within a negligible margin of error regardless of the initial conditions used. In Chapter 4, we propose and validate a robust optimization approach for multi-isocenter large field treatment plan using intensity modulated proton therapy to overcome the dose inhomogeneity caused by field misalignment in the junction regions. Results show that the robust IMPT can deliver a low gradient field dose in the junction which can minimize the dose deviation caused by misalignment. We also present the relationship between dose deviation, uncertainty and junction size. In Chapter 5, we focus on beam angle optimization problem in IMPT treatment planning. A two-stage method is developed to solve the BAO problem incorporating internal organ motion for thoracic cancer using IMPT. The first stage of the method is beam angle clustering and the second stage is final solution searching. Support vector machine is used for beam angle classification and 4D-CT is integrated to handle internal organ motion. Finally, Chapter 6 is devoted to making a summary and discuss some potential research directions following this dissertation.

Chapter 2

Literature Review

2.1 Fluence Map Optimization Problem

The fluence map optimization problem is a basic topic for radiation treatment planning optimization. The goal is to select an optimal intensity map of beamlets to deliver a uniform prescribed dose to the target while minimize the radiation dose on critical organs. To accomplish this, an objective function is used to describe the difference between the desired dose distribution and the realized dose distribution. Different formulations have been proposed in previous studies.

The most commonly used objective functions are dose based and dose-volume based objective functions [22, 23, 24, 25]. The advantage of these objective functions is they are straightforward for the treatment plan evaluation. [26] proposed an objective function based on the equivalent uniform dose (EUD) for radiation therapy optimization. This objective function uses the biologically equivalent dose to evaluate the plan quality. Linear programming models also have been used to formulate the FMO problem [27].

In order to solve the fluence map optimization problem a great number of optimization algorithms have been proposed to find the optimal beamlets intensity profiles. These strategies can be grossly classified into two groups: global optimization (GO) and local optimization (LO). GO approaches include linear programming [27, 28, 29], mixed integer programming [30, 31], simulated annealing [32] and genetic algorithms [33, 28]. These approaches are designed to reach a global optimal solution. However, they all require an excessive amount of time for optimization, which is not practical in clinical treatment planning. In addition, the performance of these approaches depends heavily on the choice of parameters [29]. For example, simulated annealing and genetic algorithms, have the advantages of avoiding getting trapped in local minima in principle, they are slow and may also get trapped in local minima if the thermal cooling process is too fast in the case of simulated annealing, or if the population evolution is not realistic in the case of genetic algorithms. Linear programming methods can incorporate constraints and guarantee to have an optimal solution. However, they are limited to linear objective functions, which are poor indicators of the response of tumors and healthy tissue to radiation. On the other hand, LO approaches include gradient-based algorithms [34, 35, 36, 37], local neighborhood search [38] and iterative methods [39]. These algorithms are designed to find a local minimum solution in a relatively short time. So, LO approaches have been commonly used for clinical treatment planning optimization to yield a clinical acceptable solution within a clinical acceptable time frame. Especially, the gradient-based algorithms approaches have been chosen for commercial treatment planning systems such

as Eclipse [Varian Associates, Palo Alto, CA] using quasi-Newton method [40] and sequential quadratic programming (SQP) employed in Pinnacle [Philips, Milpitas, CA].

2.2 Robust Optimization for Radiation Therapy

In standard radiation treatment regime, the spatial and temporal dose distribution is optimized assuming the patient geometry is static over the course of treatment and a fixed dose of radiation is delivered in every treatment fraction. However, during the course of treatment, the patient geometry may deviate from the one observed in the image on which a treatment plan is based. These uncertainties add complexity to the inherent trade-off between minimizing the healthy tissue dose (or probability of side effects) and ensuring that the tumor receives a sufficient dosage of radiation.

To date, robust optimization is widely used to incorporate different uncertainties into the optimization process to improve the robustness of treatment plans. Many robust optimization models for radiation treatment planning were developed by researchers. [18] proposed a robust optimization approach accounted for patient interfraction motion and setup uncertainties for IMRT. The results demonstrated that robust solution achieved better healthy tissue sparing than a clinical margin solution without compromising tumor coverage and robustness. [19] considered dose matrices calculation error and interfraction position uncertainties into an IMRT treatment planning problem formulation, and showed that a robust solution outperforms nominal solution (one which assumes a dose matrix in known with certainty) in terms of

tumor coverage and improved healthy tissue sparing when compared with margin solution. [17] use a three-dimensional Gaussian distribution function to simulate random organ motion for IMRT planning. [21] introduced a robust methodology for IMRT treatment planning under uncertainty and considered the specific case of intrafractional uncertainty induced by breathing motion. They incorporated the uncertainty in the probability mass function of breathing motion into the inverse planning optimization and ensured that all target voxels received sufficient expected dose for all probability distributions within a polyhedral set. [41] generalized robust optimization framework for IMRT planning without considering probability distribution of uncertainties.

Worst case robust optimization is another main approach to consider uncertainties. [37] proposed a "worst case" optimization method for IMPT by considering both setup uncertainty and range uncertainty. In this approach, the worst case dose in each voxel was calculated to evaluate the objective function. [42] use minimax robust optimization method to handle setup and range uncertainties in IMPT planning. The worst scenario among the nominal and uncertainty cases was punished by the optimization algorithm. Both of these approaches can work with a linear programming (LP) model [43] and a nonlinear programming (NLP) model [44]. The results of all these papers show that the robustness of IMPT plan can be significantly improved by robust optimization, while without loss nominal case target coverage and OAR sparing.

2.3 Beam Angle Optimization Problem

In radiation therapy, to ensure a uniform target dose coverage and to avoid the radiation dose deposition in the OARs exceed its tolerance, treatment plans are always designed to deliver radiation from a number of different angles around the patient. So, the beam angles selection is critical to making a high-quality treatment plan. However, the beam angle optimization (BAO) problem in radiation treatment planning is typically a large-scale combinatorial optimization problem. Due to the computational difficulty, BAO is not implemented in commercial treatment planning systems (TPS). In the current clinical practice, the number and angles of treatment beams are decided heavily based on the knowledge and experience of planners. To achieve the automated selection of the orientations of treatment beams for external radiation therapy, different studies have extensively investigated optimization algorithms for solving the BAO problem. These strategies can be mainly classified into two groups.

The first group algorithms combine beam angle selection and fluence map optimization to formulate the BAO problem as a mixed integer programming (MIP) problem. [45] first proposed the MIP model for beam angle selection for the conventional conformal radiation therapy. [46] introduced a MIP model to solve the BAO for IMRT, which incorporated the FMO in IMRT to guide the beam selection. [47] proposed a mixed integer linear programming technique for BAO for conventional 3D conformal radiation therapy and later [48] extended that work to IMRT optimization. However, solving the BAO problem is computational intensive because it

is typically a large-scale combinatorial optimization problem. Moreover, the BAO is highly nonconvex and may have many local optima. To solve this problem within a clinical acceptable computational time, different algorithms have been proposed. [49] utilized a generic algorithm to optimize the beam angles for conventional conformal radiation therapy. [50] extended the research using GA for the IMRT beam angle optimization. [51] proposed a particle swarm algorithm for BAO. Simulated annealing and fast simulated annealing algorithms have also been used for beam angle optimization [52, 53]. [54] also introduced an artificial neural network algorithm. [55] developed a nested partition method and [56] introduced a neighborhood search algorithm for BAO. [57] introduced a sampling strategy to reduce the size of the problem to shorten the solution time. [38] proposed a two-phase method, which using a Branch and Prune (B&P) algorithm combine with a local neighborhood search method to find solutions close to global optimal within a short time. Based on the advantages of different algorithms, [58] introduced a hybrid framework to improve the efficiency of BAO in IMRT. Although these algorithms may increase the speed of solving BAO problem, they still require a large number of iterations and the results are also influenced by initial points and the parameters choose for the algorithms.

The second class of solutions to BAO uses prior knowledge about the problem to guide the beam angles selection to reduce the search space. [54] attempted an intelligent search using an artificial neural network technique to evaluate geometric data. [59] introduced scalar scoring functions to rand candidate beam directions by using beam's-eye-view projections technique. These pre-optimization lead to the set

of treatment beams is assigned to the most favorable beam directions. [60] used a measure of angle suitability based upon the beam's-eye-view to select beams for both coplanar (i.e., all beams with the same iso-center) and non-coplanar (i.e., beams with different iso-centers) 3D conformal radiotherapy. [61, 62] neglect the beam selection from a large combination MIP problem and choose to add beams to a radiation therapy plan iteratively. [63] attempted a beam angle selection method based on target equivalent uniform dose (EUD). [64] ranked the beam orientation based on dose-volume information for IMRT beam selection. [65] reported a beam score method for BAO. The score is determined by the maximum PTV dose delivery of each beamlet and the overall score of the gantry angle was calculated as a sum of the scores of all beamlets. [66] facilitated a clustering algorithm in the context of beam angle selection by applying a Euclidean metric in a space of characteristic vectors for a set of candidate beam directions. [67] suggested a spherical K-means clustering algorithm for beam angle selection for IMRT.

Although the previously published algorithms showed the benefits of using BAO to improve the treatment plan quality, while most previously published studies on BAO were designated for conventional photon-based radiotherapy, algorithms introduced might be difficult to hold their quality and efficiency in implementation for IMPT. First, it is much more expensive to compute a score for an incident beam configuration in IMPT planning than that in IMRT or 3DCRT because of larger data size. The dose influence data for IMPT must contain information of scanning spot depth which is an additional dimension beyond typical IMRT data. Another critical

distinction is the plan robustness over delivery uncertainties in IMPT planning can be critical which is usually absent in photon-based radiotherapy. As protons deliver most of their dose at their Bragg peaks, misplaced Bragg peaks can easily cause dose inhomogeneity in the target volume and overdosing in normal tissues. Therefore, IMPT plan can be very sensitive to treatment delivery uncertainties. The final delivered dose distribution will greatly deviate from the prescribed dose if uncertainties are not considered in the treatment planning process. The beam angle can be an important fact for plan robustness, especially for the cancer cases with large organ motion such as lung and esophageal cancer. The previous algorithms haven't considered this information to perform the beam angle selection, so they may not be able to guarantee the plan quality and robustness to meet the clinical criteria at the same time.

[43] introduced a method to incorporate the setup and range uncertainties in a local neighborhood search algorithm to solve robust BAO problem. Although this local search algorithm may increase the speed of solving BAO problem, it still requires a large number of iterations and the results are influenced by initial points. Especially, for uncertainty incorporated BAO, different dose scenarios need evaluation in each iteration, an intensive computer time is still needed. To increase the effectiveness of solving robust BAO is important and still an open question in both the practice and the research domain.

Chapter 3

A Molecular Dynamics Method for Fluence Map Optimization in Intensity Modulated Proton Therapy

The fluence map optimization (FMO) problem of radiation therapy is commonly formulated as a quadratic programming (QP) model with non-negativity bounds on variables. Because the resulting QP model is very large scale, many researchers in the medical community have proposed to convert the FMO problem into an unconstrained minimization model, and then used gradient-based optimization methods such as quasi-Newton to solve the problem faster. However, there is a major issue concerning the convergence of such approaches, claimed by many researchers, that the model has multiple local optimal solutions; hence the quality of the solution varies widely depending on an initial solution to the problem. This is contradictory to the theory of a convex model. We believe that the actual issue is due to the removal of non-negativity constraint in the model coupled with a poor implementation of the algorithm. To shed the light on this problem and to make an initial attempt to overcome such shortcomings, we propose a molecular dynamics (MD) method as a new alternative for solving the QP model. A dose-based objective function is used

to compare the performance of the MD method with those of the gradient based methods using three clinical cancer cases: prostate, head-and-neck, and lung cancer. Overall, the MD method consistently converged to a solution regardless of the initial conditions as used by many researchers. Furthermore, MD converged faster than L-BFGS-B that is more reliable algorithm than L-BFGS.

3.1 Introduction

The fluence map optimization (FMO) problem of intensity modulated radiation therapy (IMRT) planning has been extensively studied and has been addressed by various solution strategies, such as linear programming, simulated annealing, genetic algorithms and gradient-based local search algorithms. In these algorithms, the gradient-based local optimization approaches are normally adopted to solve the FMO problem in clinical practice because they can yield a clinical acceptable solution within a short frame.

Based on the physics of radiation particle transport, a feasible radiation treatment plan must contain non-negative beamlet intensity. Nevertheless, researchers in the medical community have often used unconstrained gradient-based algorithms to optimize treatment plans [68, 69], as it is implemented in a leading commercial treatment planning system such as Eclipse [Varian Associates, Palo Alto, CA]. This has resulted in sub-optimality due to exclusion of the non-negative constraint, and difficulty of convergence to a global optimal solution due to a poor implementation of

the algorithm. To tackle this issue, these researchers relied on an unconstrained QP model and commonly solved it using a gradient-based method with different starting point generation strategies, which really did not resolve the fundamental issue of the problem as we point out in this paper.

Despite the fact that many solution approaches have been proposed to solve the FMO problem, a method of avoiding local minima while converging to a solution within a practical time limit has rarely been reported. [70, 71] proposed a method to formulate the IMRT FMO problem into a molecular dynamics problem which motivated the study in this chapter. Molecular dynamics is a powerful computational technique that is often used to simulate the physical movement of atoms and molecules in a many-body system. In Hou’s paper, the beamlets in IMRT were considered as virtual atoms. The weight of the beamlets were formulated as the positions of the virtual atoms and the objective function value (OFV) of the FMO problem in IMRT was formulated as the potential energy of the dynamic system. In classical molecular dynamics, because the movement of atoms follows Newton’s Law of Motion, the dynamic system will relax to an equilibrium state with the lowest free energy. In this process, the position and velocity of the atoms will change with time. Thus, following the MD formulation, the beamlets weight and virtual velocity will update over time and the OFV of the FMO problem will be minimized. The MD method’s feature of virtual velocity differs from traditional gradient algorithms in that it only updates the weight. Furthermore, within the FMO problem, virtual velocity can help atoms to keep in the bounded area, e.g., an atom goes out of the range but can move

back by changing the direction of its velocity. In addition, the search direction in MD follows dynamic equations which enable MD to converge faster than many global optimization methods. To show the performance of MD, three well cited gradient methods are selected and implemented for the IMPT FMO problem, and all four of these methods are tested using three clinical cancer patient cases and the typical three different starting point generation approaches used by many researchers in the medical community. The primary objective of this study is to demonstrate that the MD method can be a viable alternative for solving the IMPT FMO problem to overcome the major issue of traditional gradient-based methods: a premature termination to a feasible solution and sensitivity to the starting point to the algorithm.

The remainder of the paper is organized as followed: Section 3.2 describes the optimization model and the solution algorithms for the FMO problem in IMPT, which include gradient-based methods (an existing quasi-Newton method, L-BFGS and L-BFGS-B) and the MD method. The data used in the experiment and the initial configuration setups are listed in Section 3.3. Section 3.4 provides the results and discussion regarding the convergence properties of the MD method and other gradient algorithms.

3.2 Optimization Model and Solution Algorithms

3.2.1 FMO Problem Formulation

The main purpose of IMPT is to deliver the prescribed conformal radiation dose to the targeted tumor while sparing normal tissues. To achieve this goal, we define a quadratic objective function to quantify the difference between the prescribed dose

(D^n , where n is the index of organ of interest) and the actual dose (D_i , where i is the voxel index) delivered to the patient. Although different types of objective functions are reported in the literature [72, 73], dose-based quadratic objective functions are commonly used in the medical physics community [34, 35, 36, 44] for optimizing beam intensities in radiation therapy. Hence, it is used to develop the optimization model for this paper.

A dose-based objective function F is composed of two parts: F^T for the target, and F^{OAR} for the OAR. Because the primary goal of treatment planning is to obtain an actual radiation dose profile that is identical or nearly identical to the prescribed dose level on the target, F^T can be defined as the deviation of the resulting actual dose D_i on voxel i from the target prescription dose D^T :

$$F^T = \frac{1}{N^T} \sum_{i=1}^{N^T} (D_i - D^T)^2, \quad (3.1)$$

where N^T is the total numbers of voxels in the target. Similarly, F^{OAR} can be defined as

$$F^{OAR} = \frac{1}{N^{OAR}} \sum_{i=1}^{N^{OAR}} (D_i - D^{OAR})_+^2, \quad (3.2)$$

where N^{OAR} is the total number of voxels in the OAR, D^{OAR} is the specified tolerance dose for the organ, and $(\delta)_+$ is defined as $(\delta)_+ = \max(\delta, 0)$. Note, we introduced the step function F^{OAR} because healthy organs are often allowed to receive radiation doses up to a certain amount. However, once the amount is over a tolerance value, a

penalty will be imposed on the voxel according to the degree of deviation from the tolerance.

The dose D_i in voxel i can be calculated as

$$D_i = \sum_j^N k_{ij} \omega_j, \quad (3.3)$$

where ω_j is the weight or intensity of beamlet j , which is the decision variable of our IMPT optimization model. Notation N denotes the total number of beamlets and k_{ij} is the unit dose contribution of the j^{th} beamlet to the i^{th} voxel; k_{ij} is also known as the dose deposition coefficient. Here the values of k_{ij} are calculated using an in-house dose calculation engine for proton beamlets [74].

Using the notation described above, the dose-based objective function for our optimization model is

$$F = \sum_n p^T F^T + \sum_m p^{OAR} F^{OAR}, \quad (3.4)$$

where p^T and p^{OAR} denote the penalty weights of the tumor and OAR, respectively. These weights are often obtained by trial and error by planners (dosimetrists, physicists, etc.), to find a balance between tumor dose coverage and OAR dose sparing to satisfy the clinical criteria. In this study, the model follows the common practice in the medical community of containing a physical constraint: the beamlet weight cannot be negative, i.e., $\omega_j \geq 0, j = 1, 2 \dots N$. Hence, our optimization model for the

IMPT FMO problem is

$$\begin{aligned}
 & \min F \\
 & \text{s.t. } w_j \geq 0 \quad , \quad j = 1, 2, \dots, N \quad .
 \end{aligned}
 \tag{3.5}$$

3.2.2 Solution Algorithms

The general nonlinear optimization algorithms updating function can be described as

$$x_{k+1} = x_k + \alpha_k d_k, \tag{3.6}$$

where the x_k is the decision variables at the k th iteration, d_k and α_k are the corresponding direction and step size. The solution can be found using the following iterative process:

- (1) Calculate search direction d_k and step size α_k .
- (2) Update the decision variables according to the updating function.
- (3) Check whether the stopping criterion is satisfied. If it is not satisfied go to step 1; otherwise, output the final solution.

Different algorithms are classified according to the way they choose the search direction and the step size. We have selected three well-cited gradient based methods (a quasi-Newton method, L-BFGS and L-BFGS-B algorithms) to compare the performance of the proposed MD method. Note that the quasi-Newton and L-BFGS are unconstrained optimization algorithms and L-BFGS-B is designed to handle the

problem with bound constraints.

A quasi-Newton method

For the standard Newton's method, search direction d_k is generated according to the gradient and the Hessian matrix, $d_k = [\nabla^2 F(x_k)]^{-1} \nabla F(x_k)$. Accordingly, the updating function for the FMO problem can be described as

$$\omega(k+1) = \omega(k) - \alpha_k \frac{\nabla F(\omega(k))}{\nabla^2 F(\omega(k))}, \quad (3.7)$$

where the first derivative of the objective function is

$$\frac{\partial F}{\partial \omega_j} = 2 \frac{p^T}{N^T} \sum_{i=1}^{N^T} (D_i - D^T) k_{ij} + 2 \frac{p^{OAR}}{N^{OAR}} \sum_{i=1}^{N^{OAR}} (D_i - D^{OAR}) k_{ij}, \quad (3.8)$$

and second derivative is

$$\frac{\partial^2 F}{\partial \omega_j \partial \omega_k} = 2 \frac{p^T}{N^T} \sum_{i=1}^{N^T} k_{ij} k_{ik} + 2 \frac{p^{OAR}}{N^{OAR}} \sum_{i=1}^{N^{OAR}} k_{ij} k_{ik}. \quad (3.9)$$

The Newton's method can be used for solving the FMO problem, but there are inherent disadvantages: (1) the computation of the inverse of the Hessian matrix may require extensive time because of the large size of the beamlet vector $\omega(k)$, and (2) the Newton's method is an unconstrained algorithm, which cannot guarantee the feasibility of the solution to the treatment. To overcome these shortcomings, the problem can be computed in a clinical acceptable time, we approximate the Hessian matrix by its diagonal [34]. And we applied the damping factor which introduced by

Lomax(1999) as step size α_k to deal with the non-negativity constraints, which in the form of:

$$a_{ij}(k) = \frac{\omega_j(k) k_{ij}}{D_i(k)}. \quad (3.10)$$

Therefore, the updating function of the beamlet weight ω_j becomes

$$\begin{aligned} \omega_j(k+1) &= \omega_j(k) - \alpha_k \left(\frac{\frac{\partial F}{\partial \omega_j}}{\frac{\partial^2 F}{\partial \omega_j^2}} \right) \\ &= \omega_j(k) \frac{\frac{p^T}{N^T} \sum_{i=1}^{N^T} k_{ij}^2 \frac{D^T}{D_i} + \frac{p^{OAR}}{N^{OAR}} \sum_{i=1}^{N^{OAR}} k_{ij}^2 \frac{D^{OAR}}{D_i}}{\frac{p^T}{N^T} \sum_{i=1}^{N^T} k_{ij}^2 + \frac{p^{OAR}}{N^{OAR}} \sum_{i=1}^{N^{OAR}} k_{ij}^2}. \end{aligned} \quad (3.11)$$

As a result, the $\omega_j(k+1)$ equals to $\omega_j(k)$ times a positive coefficient, which guarantees the feasibility of non-negative beamlet intensities.

L-BFGS: Limited-memory Broyden-Fletcher-Goldfarb-Shanno algorithm

The major difference of an L-BFGS algorithm [75] compared to the quasi-Newton method is in the updating function that can be described as

$$x_{k+1} = x_k - \alpha_k H_k \nabla f(x_k), \quad (3.12)$$

where

$$\begin{aligned}
H_{k+1} &= (V_k^T \cdots V_{k-\hat{m}}^T) H_0 (V_{k-\hat{m}}^T \cdots V_k^T) \\
&+ \rho_{k-\hat{m}} (V_k^T \cdots V_{k-\hat{m}+1}^T) s_{k-\hat{m}} s_{k-\hat{m}}^T (V_{k-\hat{m}+1}^T \cdots V_k^T) \\
&+ \rho_{k-\hat{m}+1} (V_k^T \cdots V_{k-\hat{m}+2}^T) s_{k-\hat{m}+1} s_{k-\hat{m}+1}^T (V_{k-\hat{m}+2}^T \cdots V_k^T) , \\
&\vdots \\
&+ \rho_k s_k s_k^T
\end{aligned} \tag{3.13}$$

$$y_k = \nabla f(x_{k+1}) - \nabla f(x_k), \tag{3.14}$$

$$\rho_k = \frac{1}{y_k^T s_k}, \text{ and} \tag{3.15}$$

$$V_k = I - \rho_k y_k s_k^T. \tag{3.16}$$

In H_k , the \hat{m} is the number of stored Hessian approximation correction steps and the step size α_k is selected to satisfy the Wolfe conditions. However, L-BFGS is also an unconstrained algorithm. To handle the non-negativity issue, the beamlet intensity ω_j is often replaced by a non-negative quantity $\omega_j'^2$ [44, 36]. So, the dose in voxel j is calculated as $D_i = \sum_j^N k_{ij} \omega_j'^2$. Thus, the constrained optimization problem with respect to weights is approximated by an unconstrained one in which the square

root of the beamlet weights is used to optimize the plan rather than optimizing the beamlet weights directly in the model. As a result, the first derivative of the revised objective function becomes:

$$\frac{\partial F}{\partial \omega'_j} = 4 \frac{p^T}{N^T} \sum_{i=1}^{N^T} (D_i - D^T) k_{ij} \omega'_j + 4 \frac{p^{OAR}}{N^{OAR}} \sum_{i=1}^{N^{OAR}} (D_i - D^{OAR}) k_{ij} \omega'_j. \quad (3.17)$$

L-BFGS-B: L-BFGS algorithm with box constraints

A major drawback of using L-BFGS for the IMPT FMO problem is its inability of adding the non-negativity constraint. A better alternative is the L-BFGS-B algorithm that is designed for solving nonlinear optimization problems with simple box constraints on variables [76] as

$$\begin{aligned} \min \quad & F(x) \\ \text{s.t.} \quad & l \leq x \leq u, \end{aligned} \quad (3.18)$$

where l and u represent lower and upper bounds on the variables. So, this algorithm is capable of handling the non-negativity bounds on intensity of beamlets. Theoretically, it can solve the FMO problem to optimality and find the global optimal solution.

Molecular Dynamics

MD is a computational technique for many-body system simulation that has been widely applied in the material sciences community. In a classical MD model, the physical movements of particles in the system follow Newton's Laws of Motion [77]. Let x_j be the position and v_j be the velocity of a particle j . Then, force f_j is

the product of mass m_j and acceleration a_j of the particle:

$$\begin{aligned} f_j &= m_j \cdot a_j = m_j \frac{dv_j}{dt} = m_j \frac{d^2x_j}{dt^2}, \\ f_j &= -\nabla_j E \end{aligned} \quad (3.19)$$

where t is the time of the system and E is the potential energy of the system. The force f_j is related to the acceleration and can be expressed as the gradient of the potential energy of the particle.

Based on Newton's Laws of Motion, the position, velocity and acceleration of the particle can be described as functions of time t ;

$$v_j = \frac{dx(t)_j}{dt}, \quad a_j = \frac{dv(t)_j}{dt} = \frac{d^2x(t)_j}{dt^2}. \quad (3.20)$$

Therefore, the continuous motion configuration of the system can be calculated by integrating Newton's Laws of Motion. When the system is under the influence of continuous potential energy, the positions and velocities can be approximated using a Taylor series expansion for a small time step Δt , $\Delta t > 0$:

$$\begin{aligned} x(t + \Delta t) &= x(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 + \frac{1}{6}b(t)\Delta t^3 + \dots, \\ v(t + \Delta t) &= v(t) + a(t)\Delta t + \frac{1}{2}b(t)\Delta t^2 + \frac{1}{6}c(t)\Delta t^3 + \dots, \end{aligned} \quad (3.21)$$

where a , b and c are the second, third and fourth time derivatives of the coordinates.

This Taylor expansion serves as the basis for the most common integrators used

in MD calculations. Many classical methods for integrating equations require information from both current and previous steps to update the system. This means the information from these steps must be stored in memory and the system cannot self-start at the beginning [78]. To resolve this problem, [79] introduced the Velocity Verlet method which requires information from the previous step only:

$$\begin{aligned} x(t + \Delta t) &= x(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2, \\ v(t + \Delta t) &= v(t) + \frac{1}{2}\Delta t[a(t) + a(t + \Delta t)]. \end{aligned} \tag{3.22}$$

Therefore, this approach is selected in our algorithm to update the MD system to solve the IMPT FMO problem.

In IMPT, the optimization problem can be formulated as a dynamic system with N virtual atoms [70]. Each beamlet weight (ω_j is assumed to be the position (x_j) of a virtual atom j in 1-D dimension. The objective function F can be considered as the potential energy E of the system. As a result, the dynamic equations for virtual atom j can be expressed as

$$v_j = \frac{d\omega_j}{dt}, \quad a_j = \frac{dv_j}{dt} = \frac{d^2\omega_j}{dt^2}, \quad f_j = m_j \frac{dv_j}{dt} = \frac{\partial F}{\partial \omega_j}. \tag{3.23}$$

We followed the approach of [70], in which the mass of the virtual atom j equals the summation of the unit dose contribution of all voxels influenced by the j^{th} beamlet, written as $m_j = \sum_{i=1}^{N_j} k_{ij}$, where N_j is the total number of voxels influenced by beamlet j . Using the velocity Verlet method, the dynamic updating equations for the IMPT

FMO problem is written as

$$\begin{aligned}\omega_j(t + \Delta t) &= \omega_j(t) + v_j(t) \Delta t + \frac{1}{2m_j} \Delta t^2 f_j(t) \\ v_j(t + \Delta t) &= v_j(t) + \frac{1}{2m_j} \Delta t [f_j(t) + f_j(t + \Delta t)]\end{aligned}\tag{3.24}$$

Hence, we calculate the updating beamlet weights by Eq. 3.24 and we can combine Eq. 3.24 with Eq. 3.4 to calculate the trajectory of the OFVs.

In physics, temperature is used to specify the thermodynamic state of a system. In the MD system, temperature T is related to the kinetic energy of the system and can be calculated as

$$T = \frac{1}{3Nk} \sum_{j=1}^N m_j v_j^2,\tag{3.25}$$

where k is the Boltzman constant and N is the total number of particles in the system.

The MD system will converge to an equilibrium state with the lowest free energy. Note that free energy consists of kinetic energy and potential energy. Therefore, the potential energy equals the free energy only when kinetic energy is zero, i.e., the temperature of such system is zero. Thus, the objective function (potential energy) of our FMO model is minimized when the system reaches an equilibrium state with zero system temperature. However, in physics, the kinetic energy and potential energy of a dynamic system follow the law of energy conservation. Although kinetic and potential energy will interchange continuously, the total energy will remain unchanged. This can create an issue of convergence to a specific point because the atoms can still carry

significant speed when they reach that point. As a result, atoms may move away from an optimal point with the lowest potential energy. In order to solve this problem, a "friction" to the system (i.e., a damping factor to the MD system) is added to slow the movements of atoms as they get closer to an optimal point. The following damping function is applied in our algorithm,

$$v_j(t) = \begin{cases} \lambda v_j(t), & \text{if } v_j(t) f_j(t) < 0 \\ v_j(t), & \text{otherwise.} \end{cases} \quad (3.26)$$

where $0 < \lambda < 1$.

As we mentioned above, their speed may cause the atoms to pass the optimal point and create an issue of convergence. On the other hand, when the virtual atoms become trapped in local minima, a proper speed may help them continue to move and get out of those local minimum points. Using this feature, we employ temperature scaling to adjust the velocities to help the atoms move out from the local minimum points. From the updating function, Eq. 3.24, we define the scaling function as

$$v_j(t + \Delta t) = \sqrt{\frac{T_d}{T_0}} v_j(t) + \frac{1}{2m_j} \Delta t [f_j(t) + f_j(t + \Delta t)], \quad (3.27)$$

where T_0 is the initial temperature and T_d is the desired temperature.

An important physical constraint of IMPT planning is that the beamlet weight cannot be negative. [70] suggested a barrier potential with an infinite height at $\omega = 0$ to impose this constraint. The virtual atoms are reflected by changing the sign of

their velocity each time they try to pass the barrier: $v_j(t) = -v_j(t)$, when $\omega_j(t) < 0$.

The proposed MD method to optimize the IMPT FMO problem is outlined below. The algorithm stops when either of these following conditions are met: (a) a certain number of iterations is reached or (b) there is no change (smaller than ε , i.e., 10^{-4} in the objective value for a certain number of consecutive iterations.

Algorithm 1: Molecular dynamics method

```

1 Initialization:  $\delta t := t_0; \bar{v} := 0; \bar{\omega}^* := \bar{\omega}_0; F^* := F_0;$ 
2 while Stopping criteria are not met do
3   | Calculating force  $\bar{f}$  and updating  $\bar{\omega}$  and  $\bar{v}$ , Eq.3.24; calculating  $F$ ,
   | Eq.3.1-3.4;
4   | if  $F < F^* - \varepsilon$  then
5   |   | s
6   |   | end
7   |   |  $F^* := F; \bar{\omega}^* := \bar{\omega};$  else if  $F < F^*$  then
8   |   |   |  $F^* := F; \bar{\omega}^* := \bar{\omega};$ 
9   |   |   | Heating the system by setting the desired temperature  $T_d = \beta T_0;$ 
10  |   | end
11  |   | else
12  |   |   | Heating the system by setting the desired temperature  $T_d = \beta T_0;$ 
13  |   | end
14  |   | Damping the system by setting the atom  $j$ , if  $v_j(t) f_j(t) < 0;$ 
   |   |   |  $v_j(t) = \lambda v_j(t);$ 
15  |   |   | Checking the physical constraint, set  $v_j(t) = -v_j(t)$ , when  $\omega_j(t) < 0.$ 
16  | end
17 return  $\bar{\omega}^*$  as optimal solution.

```

In our implementation, if a new solution increases the OFV to a value larger than ε , it replaces the old one. Otherwise, if the new solution increases the OFV but the value remains smaller than ε , it replaces the old solution and the system is heated by rescaling the temperature by $T_d = \beta T_0$, where β is the heating rate and typically lies between 1.1 and 2. If there is no OFV improvement, we keep the old solution and also heat the system. When the stopping criteria is satisfied, we stop the algorithm

and return the final solution.

3.3 Numerical Experiments

All algorithms were implemented in C++ and all experiments were performed on a 64-bit Linux workstation with 128 GB of memory and quad Intel Xeon E5649 2.53GHz processor. The stopping criteria were either: (a) 10,000 iteration limit or (b) no change in the OFV for 10 consecutive iterations with $\varepsilon = 10e - 5$.

3.3.1 Patient Data

The three clinical cancer cases from The University of Texas MD Anderson Cancer Center selected for this study was a prostate cancer, head-and-neck cancer and a lung cancer case (Figure 3.1). The corresponding beam angles for each case are marked by arrows F1, F2 and F3, respectively.

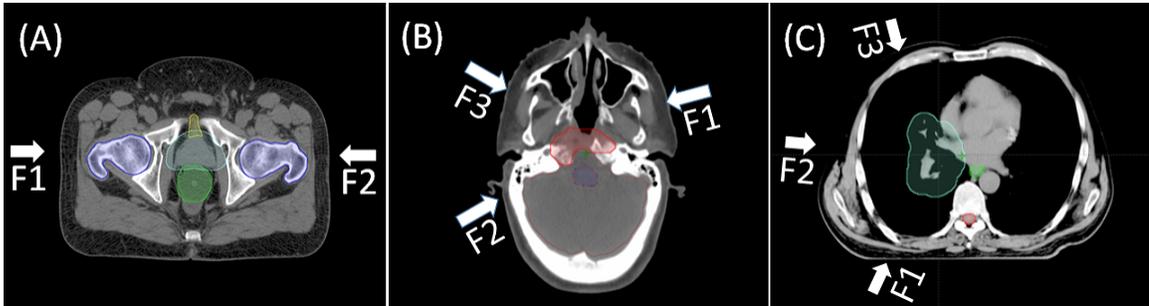


Figure 3.1: The three clinical cancer cases selected for the study and their corresponding field directions. (A) Prostate cancer case, (B) head-and-neck cancer case, (C) lung cancer case; the tumors are contoured in red.

The beam angles, number of beamlets in each beam, volumes of interest and number of voxels of each volume for each case are listed in Table 3.1. The prostate case involved a medium-sized tumor that required only a simple treatment plan in

which parallel-opposed fields were used. The head-and-neck case had a small target size but some very subtle OARs such as the optical chiasm. The lung case represents a large sized tumor case. It contains twice as much target volume than the prostate case and twice as large in total volume compared to the head-and-neck case. Because of this, more beamlets were required to cover the tumor for the lung case.

Table 3.1: The intensity modulated proton therapy beam angles, number of beamlets in each beam, VOIs and number of voxels of each VOI for the three cancer cases

Cancer Type	Beam Angle	Number of beamlets	VOI	Number of voxels
Prostate	90°	599	STV	4916
	270°	605	Bladder	15189
			Femoral heads	23908
			Rectum	8570
Head-and-neck	75°	374	CTV	2603
	240°	356	Brain	96536
	300°	365	Brainstem	2506
			Optic chiasm	110
Lung	205°	1539	PTV	11161
	275°	1218	Esophagus	1435
	345°	1042	Spinal cord	2030
			Total lung	159188
			Heart	18148

Abbreviations: VOI: volume of interest, STV: scanning target volume, CTV: clinical target volume and PTV: planning target volume.

The planned doses and penalty weights for the corresponding VOIs in a dose-based objective function for the three IMPT cases are listed in Table 3.2. The same penalty was applied to different initial conditions in each case. Because the highest priority was to satisfy the tumor coverage and dose uniformity requirements, the penalty for the target was high. Meanwhile, the target dose to the OARs is set to 0 Gy, which means we wish to minimize the dose on OARs as low as possible.

The parameter values of each algorithm were assigned the same values for all

cancer cases. The stopping criteria were either: (a) the number of iterations reached 10,000 or (b) there was no change in the OFV for 20 consecutive iterations, where $\varepsilon = 10e - 4$.

Table 3.2: Dose-based objective function parameters used for optimizing the intensity modulated proton therapy plans

Cancer Type	VOI	Dose (Gy)	Weight
Prostate	STV	78	200
	Bladder	0	1
	Femoral heads	0	1
	Rectum	0	1
Head-and-neck	CTV	74	200
	Brain	0	1
	Brainstem	0	1
	Optic chiasm	0	1
Lung	PTV	74	200
	Esophagus	0	1
	Spinal cord	0	1
	Total lung	0	1
	Heart	0	1

3.3.2 IMPT Starting Conditions

In this study, the IMPT plans of each of three cases were obtained from three different initial conditions (Figure 3.2). These initial conditions have been described by [68] and described briefly below:

(a) Forward wedge (FW). All beamlet weights are set the same creating a wedge-shaped dose that has a high dose at the proximal edge and a low dose at the distal edge (Figure 3.2A).

(b) Inverse wedge (IW). The beamlet weights are set to distal tracking creating an inverse wedge-shaped dose that has a very low dose to the proximal edge and a

high dose to the distal edge (Figure 3.2B).

(c) Spread-out Bragg peak (SOBP). The beamlet weights are arranged to deliver a flat dose on the targeted area (Figure 3.2C).

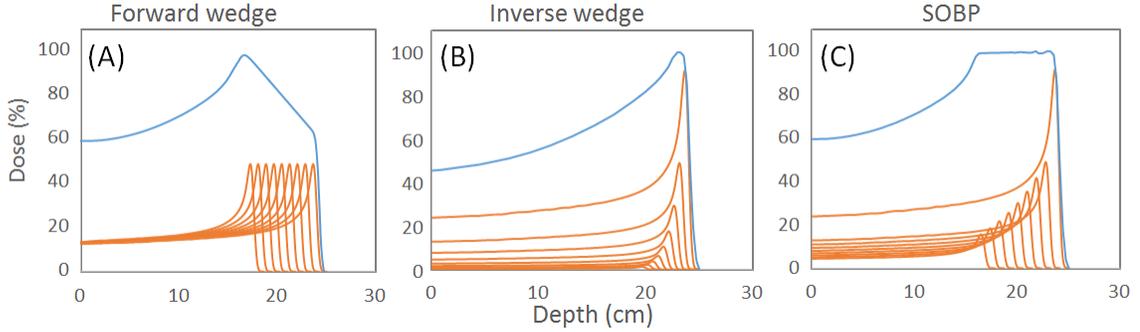


Figure 3.2: The initial beamlet weights for a single intensity modulated proton therapy field from three initial conditions: (A) Forward wedge, (B) Inverse wedge, (C) Spread-out Bragg peak (SOBP).

3.4 Results

3.4.1 Solution Quality

Table 3.3 compares the objective function values obtained from three different initial points for three patient cases using all algorithms discussed in this paper. In all patient cases, the L-BFGS-B converged consistently to the same OFV regardless of the starting points used. Furthermore, it produced lowest OFVs in all cases: 2265.1, 528.7 and 1992.6 for prostate, head-and-neck and lung cancer case, respectively. In terms of OFV, MD was the second best with the gap from L-BFGS-B within 0.1% while the MD method was not sensitive to the starting points. However, the OFVs obtained by the quasi-Newton and L-BFGS varied significantly when different starting

conditions were utilized as often mentioned in the literature. Especially, the quasi-Newton method was most sensitive to the starting point that made big differences in the final results. For example, the OFVs of head-and-neck plans are 574.1, 1150.4 and 644.3 when starting from solutions based on FW, IW and SOBP, respectively. Among these three initial points, IW yielded the worst OFVs for the quasi-Newton and L-BFGS methods. The dose-volume histograms (DVHs) of the scanning target volume

Table 3.3: Objective function value comparison using quasi-Newton, L-BFGS, L-BFGS-B and the MD methods for three tested cases starting from forward wedge (FW), inverse wedge (IW) and spread-out Bragg peak (SOBP).

Algorithm	Prostate			Head-and-neck			Lung		
	FW	IW	SOBP	FW	IW	SOBP	FW	IW	SOBP
quasi-Newton	2371.4	2657.2	2339.5	574.1	1150.4	644.3	2251.1	2262.9	2135.4
L-BFGS	2311.7	2573.2	2301.8	532.6	596.2	581.9	2087.6	2094.7	2027.6
L-BFGS-B	2265.1	2265.1	2265.1	528.7	528.7	528.7	1992.6	1992.6	1992.6
MD	2265.5	2265.4	2265.5	529.3	529.1	529.3	1992.8	1993.0	1993.0

(STV) and femoral heads for the prostate case are shown in Figure 3.3. Notice that DVHs of L-BFGS-B and the MD method were identical for all three starting points; all three lines were not distinguishable. In contrast, the lines were different when optimized using the quasi-Newton method and L-BFGS. The higher OFVs in Table 3.3 may reflect worse target coverage or OARs sparing in DVHs. In fact, the tiny differences of objective function values between the MD method and L-BFGS-B is clinically negligible. We have observed a similar result for the head-and-neck and lung cancer cases (see Figure 1,2 in Appendix).

So far, we have demonstrated that the unconstrained optimization methods are sensitive to the initial conditions and that both the L-BFGS-B and the MD method can overcome the issue. We believe that such poor results of L-BFGS and the quasi

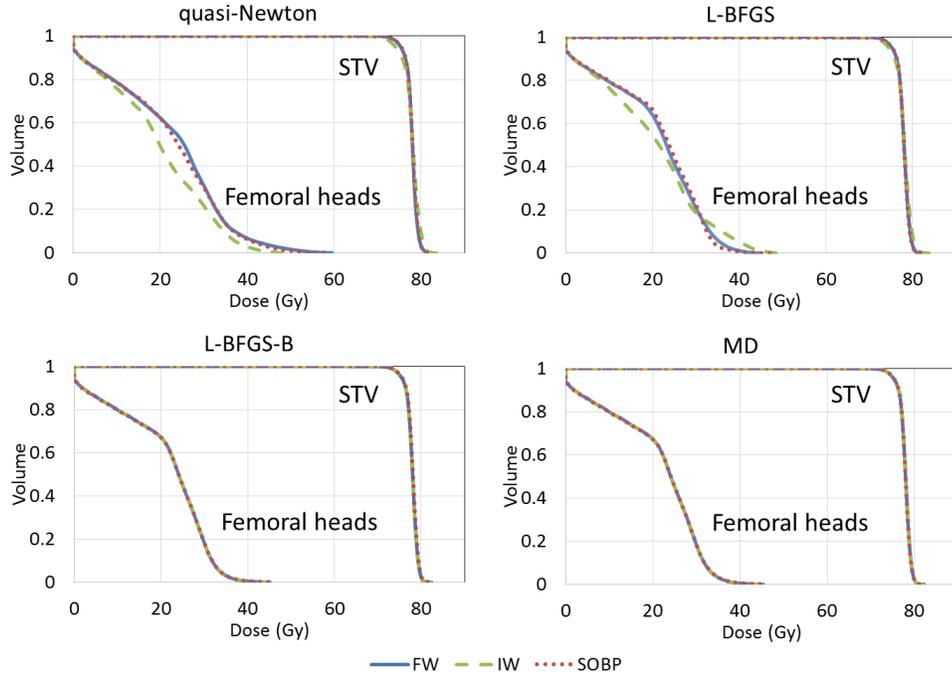


Figure 3.3: The dose-volume results of the prostate case using quasi-Newton, L-BFGS, L-BFGS-B and the MD method starting from three initial conditions: forward wedge (FW), inverse wedge (IW) and spread-out Bragg peak (SOBP).

Newton method may be caused by the modification of the objective function to convert the FMO model into an unconstrained one. From Eq. 3.11 and 3.17, the ω_j or ω'_j is a multiplier of updating function or gradient of the objective function. When ω_j or ω'_j is equal to 0, it will not be updated during the optimization process. This may happen in the initial setting or during the optimization loop. This shows the inappropriateness of solving the FMO problem as an unconstrained optimization problem. For the MD method, changing the velocity direction of virtual atoms can handle the bounds on variables, which resulted in achieving the near global optimal.

3.4.2 Computational Performance

Computational performance of the three algorithms is discussed in this section. Table 3.4 shows the CPU times in minutes of the four algorithms on three different sizes of cancer patient cases. For each pair of algorithm and a cancer case, CPU time is recorded for each starting point specified. Overall, the quasi-Newton method was faster than all the rest of methods, but the compute time was clearly influenced by the starting points. Although L-BFGS-B was the best performer in solution quality, it took considerably longer time to converge when compared with the rest of the three algorithms. The CPU times of the MD method were comparable to those of the two unconstrained algorithms and it was significantly faster than L-BFGS-B in all cases tested. For prostate, head-and-neck and lung case, the average CPU times of MD method were 80%, 76% and 58% percent faster than the results of L-BFGS-B for each of the three starting points, respectively. For prostate and head-and-neck cases, the MD method was faster than L-BFGS.

Table 3.4: Time comparison using quasi-Newton, L-BFGS, L-BFGS-B and the MD methods for three tested cases starting from forward wedge (FW), inverse wedge (IW) and spread-out Bragg peak (SOBP).

Algorithm	Prostate (min)			Head-and-neck (min)			Lung (min)		
	FW	IW	SOBP	FW	IW	SOBP	FW	IW	SOBP
quasi-Newton	16.2	23.5	10.7	27.4	14.8	18.9	121.7	103.1	92.3
L-BFGS	25.5	33.6	32.4	39.5	27.7	20.8	112.4	145.6	132.3
L-BFGS-B	112.6	106.4	103.8	99.6	78.8	82.0	315.3	362.1	354.6
MD	21.5	20.6	19.9	22.1	18.5	21.6	156.6	135.1	140.3

We further analyzed the convergence of the algorithms. Figure 3.4 shows the plots of OFVs as a function of CPU run time for the prostate cancer case. Each sub-figure shows convergence of the four algorithms for each starting point. We

observed that the MD method quickly converged to near optimal within five minutes of computation in all cases. Running MD beyond five minutes did not show much improvement in OFV. Similarly, L-BFGS-B reached a 'flat region' after 30 minutes for all three starting points. Influence of the starting points for convergence of an algorithm seems to be clear from these figures. Especially, the inverse wedge (IW) starting point seems to be much worse than FW or SOBP. The reason is that the proximal beamlets weights are more likely to be set to zero for the inverse wedge shape of initial dose. Similar results were also observed in the head-and-neck and lung cancer cases (see Figure 3,4 in Appendix).

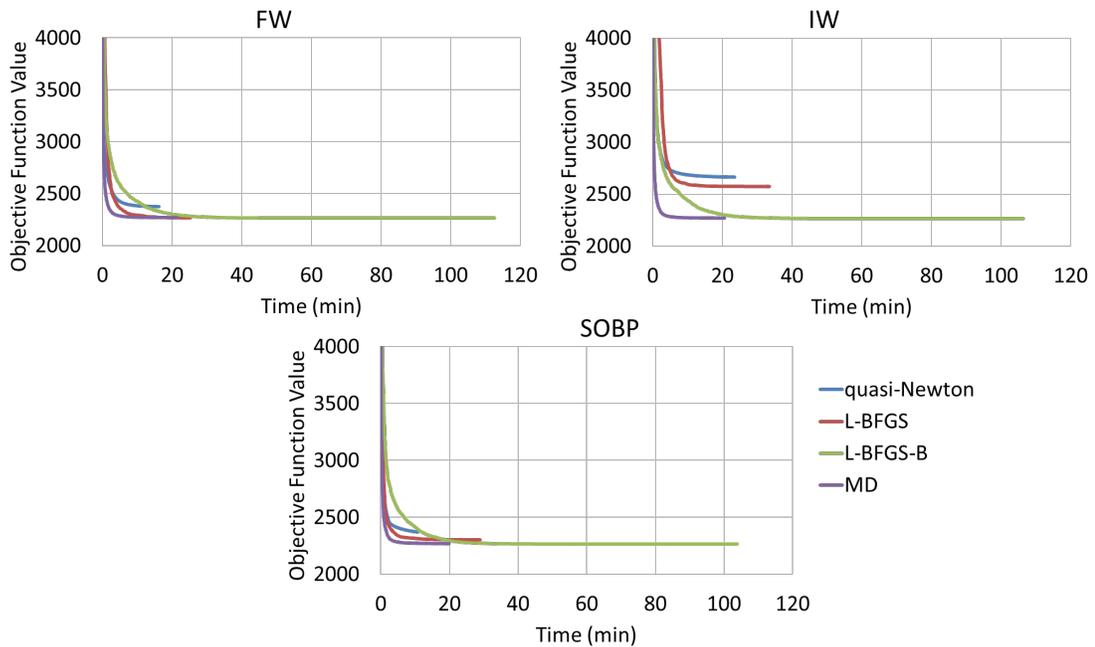


Figure 3.4: The objective function value as a function of time for the optimization processes for the prostate case starting from FW, IW and SOBP using quasi-Newton, L-BFGS, L-BFGS-B and the MD method.

Overall, the MD method stands out as the best approach when minimizing both

the objective function and the CPU time are important in obtaining a radiation treatment plan.

3.5 Conclusion

We have demonstrated that the commonly used QP model for IMPT FMO problem has a global optimal solution, which is contradictory to over a decade-old claim by researchers in the medical physics community that the model has many local optima. The literature also claims that popular gradient based solution algorithms for solving the QP model are sensitive to the starting condition. We found that there might be two compelling reasons for such shortcomings claimed by these researchers; unconstrained QP models are frequently used for solving the problem that requires non-negativity constraints on variables, and solution algorithms are incorrectly modified to address the non-negativity constraints, which results in premature termination. We then provided two remedies to fix these issues and achieve global optimal solutions: the use of a constrained QP model and a fast solution algorithm to solve the constrained optimization model. Specifically, the MD method was developed to optimize the IMPT treatment plans. The performance of the MD method was compared against a well cited quasi-Newton method and L-BFGS as well as L-BFGS-B using three clinical cancer cases of different size. For the comparison purpose, three suggested initial conditions were used to test each of these methods. By the computational results, we have confirmed that both the quasi-Newton method and L-BFGS were sensitive to the initial conditions. But more importantly, we have shown that

the MD method consistently produced solutions that were the same or within a negligible margin of error to the global optimal solutions found by L-BFGS-B regardless of the initial conditions used. Although L-BFGS-B can guarantee global optimal, it took considerably longer time to converge. The MD method converged to a 'flat area' of objective function value in five minutes, while L-BFGS-B took 30 minutes to reach a similar point.

Chapter 4

Robust Optimization for Intensity Modulated Proton Therapy Plans with Multi-Isocenter Large Fields

Conventional proton therapy for the patients with a large size tumor usually requires multiple fields combining together to create a multi-isocenter large field to cover the whole target. It is a complex treatment planning procedures, and plans can be subject to dose inhomogeneity caused by field mismatches in junction areas. In this chapter, we propose and validate a robust optimization approach for intensity modulated proton therapy treatment plans with multi-isocenter large fields to overcome these limitations and potentially improve treatment planning efficiency and patient safety. The field alignment uncertainties are incorporated into treatment planning optimization process. The results demonstrated that the robust optimized IMPT treatment plan creates a low-gradient field dose in the junction regions to mitigate the impact caused by misalignment errors and is more efficient than the conventional planning technique.

4.1 Introduction

Proton therapy is being used for an increasing range of disease sites as a result of the development of patient-specific planning and delivery techniques that improve the therapeutic ratio by taking advantage of finite proton ranges in patients [80, 81, 82, 83, 84, 85]. For large and irregular-shape tumors, such as craniospinal irradiation (CSI) [83, 84, 86] and mesothelioma irradiation [82], techniques are being developed for patient treatment. In those cases, the size of the target volume normally exceeds the mechanical limitations of the treatment field size, and multiple fields with different isocenters are required to be matched together to cover the target [87, 88]. Normally, the field dose in the junction area has a steep gradient, which makes the treatment plan sensitive to misalignment errors, and even small uncertainties can significantly affect dose uniformity. Traditionally, preventing the risk of dose deviation in junction regions usually requires a manual shift of the field junctions, which can be technically challenging.

In conjunction with the development of applying intensity modulated proton therapy (IMPT) to more disease sites, there is a major progress in the robust optimization techniques [44, 89, 42, 37]. Robust optimization methods have been developed for mitigating the effects of proton range, setup and anatomical motion uncertainties on dose delivered to a patient. However, none of the robust optimization methods reported in literature are dealing with the junction mismatch which is special for the large and irregular targets.

In this chapter, we introduce a general robust optimization approach for IMPT

plans with multi-isocenter large fields. This approach incorporates field misalignment uncertainties during the optimization process and generates a low-gradient field dose in junction regions.

4.2 Material and Methods

We selected one CSI case and one mesothelioma case to demonstrate the use of the proposed approach. Both patients underwent the simulation in the supine position. Images were obtained from patients in the treatment position with a multi-slice CT scanner at a 2.5-mm slice thickness. Target structures and organs at risk were outlined by experienced dosimetrists or radiation oncologists. The clinical target volume (CTV) in the CSI patient comprised the brain and spinal canal and was extended caudally to just beyond the thecal sac. In the mesothelioma patient, the gross tumor volume (GTV) encompassed gross disease on the postsurgical positron emission CT scan, the CTV was contoured by radiation oncologist, and the planning target volume (PTV) was consist with a 0.5-cm margin expansion around the GTV plus a 6-mm internal margin and a 1-cm external margin expansion on CTV.

For the CSI patient, a radiobiological equivalent dose of 36 Gy in 1.8-Gy fractions was prescribed for CTV. For the mesothelioma patient, the prescription dose was 45 Gy in 1.8-Gy fractions to PTV. For contouring, spot arrangement and dose we used the Eclipse version 13.0 system (Varian Medical Systems, Palo Alto, CA). The robust optimization was performed using an in-house proton treatment planning system [44]. All plans were normalized to 95% of target volume (i.e., CTV for CSI case, PTV for

mesothelioma case) received 100% prescribed dose. The homogeneity index (HI = D5/D95) was used to evaluate the target dose uniformity. The beam has a spot size with a diameter of approximately 1.6-2.2 cm (full width half maximum).

4.2.1 Field Setup and Spot Arrangement

Figure 4.1A-C show representative axial, sagittal and coronal views with marked field projections for the CSI patient. Two brain fields with the same isocenter are typically angled 15° posteriorly from the horizontal plane to reduce the dose to the lens (Figure 4.1A). For each field, the corresponding CTV included the brain contour and a portion of the upper spine contour that extended approximately 1 to 2 cm superior to the shoulders (Figure 4.1D). The spinal fields were equally spaced along the spine axis, and the isocenters were designed to minimize the total number of spinal fields and maximize the field overlap region for junctions (Figure 4.1B, C). The target covered by the spinal field immediately inferior to the brain fields may include the upper spine as well as portions of the brain target (Figure 4.1E). The maximum field size of our system is 30 cm × 30 cm; to maximize junction size, we applied a 45° couch rotation for spinal fields (Figure 4.1C, E and F). Figure 4.1G and H show representative axial and sagittal views with marked field projections for the mesothelioma patient. The PTV was covered by four fields (Figure 4.1G): two upper fields with one isocenter matched with two lower fields with another isocenter (Figure 4.1H). The corresponding targets for the upper and lower fields are shown in Figure 4.1I and J. For both patients, the spot arrangement volume of each field was

expanded by 8 mm uniformly in all directions from the corresponding target contour.

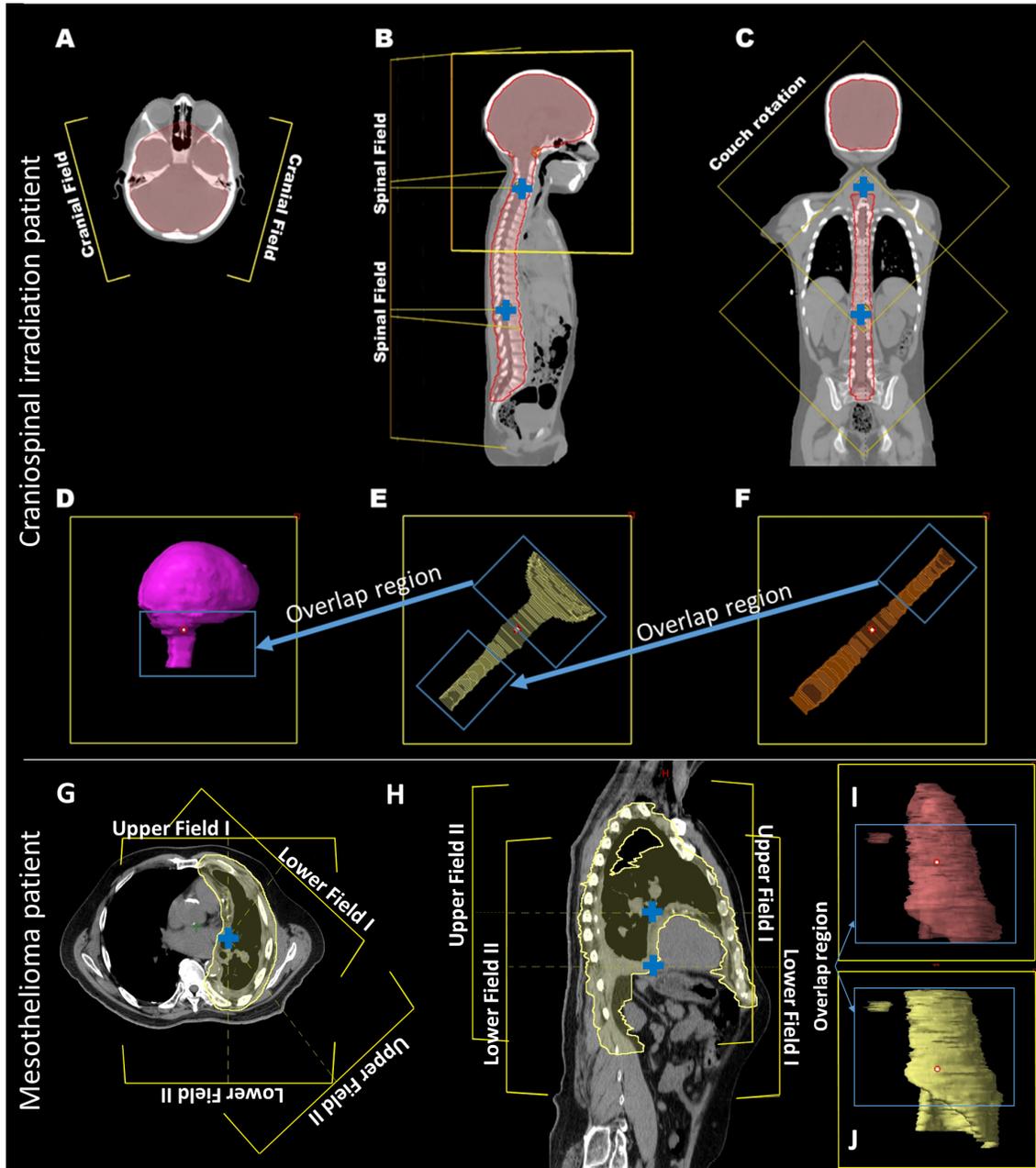


Figure 4.1: Field arrangement for the craniospinal irradiation patient (A-F) and mesothelioma irradiation patient (G-J).

4.2.2 Robust Optimization and Uncertainty Setup

In our in-house proton treatment planning system, the worst-case dose algorithm is adopted for robust optimization. In this algorithm, the dose distributions from different scenarios, including the nominal dose (i.e., without uncertainties) and different uncertainty setups, are computed. The worst-case dose distribution is represented by the maximum (for overdosage) or minimum (for underdosage) dose from all computed dose distributions in each voxel corresponding to specific structures. The formulation can be described as

$$\begin{aligned}
 \min \quad F_{Robust} = & \sum_{i \in T} \omega_{T,\min} (D_{i,\min} - D_{p,T})^2 \\
 & + \sum_{i \in T} \omega_{T,\max} (D_{i,\max} - D_{p,T})^2 \quad , \\
 & + \sum_{i \in OAR} \omega_{OAR} (D_{i,\max} - D_{p,OAR})_+^2
 \end{aligned} \tag{4.1}$$

where D_i is the worst-case (minimum or maximum) dose on voxel i , D_p is the prescription dose of target or OARs, ω is the penalty weight of the specific structure, $(\delta)_+$ is defined as $(\delta)_+ = \max(\delta, 0)$. Different dose distributions need to be computed and the worst-case dose are penalized during the optimization iterations.

We designed two uncertainty scenarios for robust optimization to simulate misalignment errors that may occur at all field junctions. In these scenarios, field isocenters shift ± 3 mm in the superior-inferior direction alternately. For example, for CSI patient, two brain fields are shifted by -3 mm, and the first and second spinal fields are shifted by +3 and -3 mm in scenario I, respectively. In scenario II, the fields are shifted by 3 mm in the opposite direction with respect to scenario I.

4.2.3 Dosimetric Evaluation

Dose uniformity in targets was evaluated with the heterogeneity index [HI = D_5/D_{95}] and inhomogeneity coefficient [IC = $(D_5 - D_{95})/D_{mean}$]. Selected dose indices for the targets and OARs were also evaluated. All plans were normalized to 95% of the CTV receiving 100% of the prescription dose. Paired t tests were used to assess potential differences (Excel, Microsoft Corp.). A dose-volume histogram (DVH) "band" was used to illustrate the robustness of the IMPT plans to uncertainties (a narrower band indicates greater robustness).

4.2.4 Plan Robustness Evaluation

Robust optimized and conventional, nonrobust IMPT plans were generated for both patients. Alternating isocenter shifts of 3 mm per field (6-mm total error) were performed to simulate the longitudinal mismatching error for robustness analysis. The dose profiles in the junction regions were used to demonstrate the deviation caused by misalignment uncertainty. For the CSI patient, robust IMPT plans with different junction sizes (8, 12, 16 and 26 cm) were generated to illustrate the relationship between junction size and dose deviation, and the robustness of a robust optimized IMPT plan with a large junction size was compared with that of a robust optimized treatment plan with a small junction and conventional junction shifting.

4.3 Results

First, we evaluated the robustness of the dose distribution in field junctions for the robust and conventional IMPT plans. The Figure 4.2 demonstrate the dose

color wash and the corresponding dose profiles from the robust and non-robust plans for two tested case. The dose color wash and dashed lines represent the dosimetric deviations resulting from a 3-mm alternating misalignment error. As shown in Figure 4.2A, the field dose in the junction region has a low smooth gradient in the robust IMPT plan but is irregular (non-smooth) in the conventional IMPT plan (Figure 4.2B). The hot and cold doses were evenly distributed in the junction region in the robust plan, and the deviation for the simulated error was around 5% (Figure 4.2A), which is significantly smaller than the 20% deviation in the conventional plan (Figure 4.2B). Similar results were observed for the mesothelioma patient (Figure 4.2C, D).

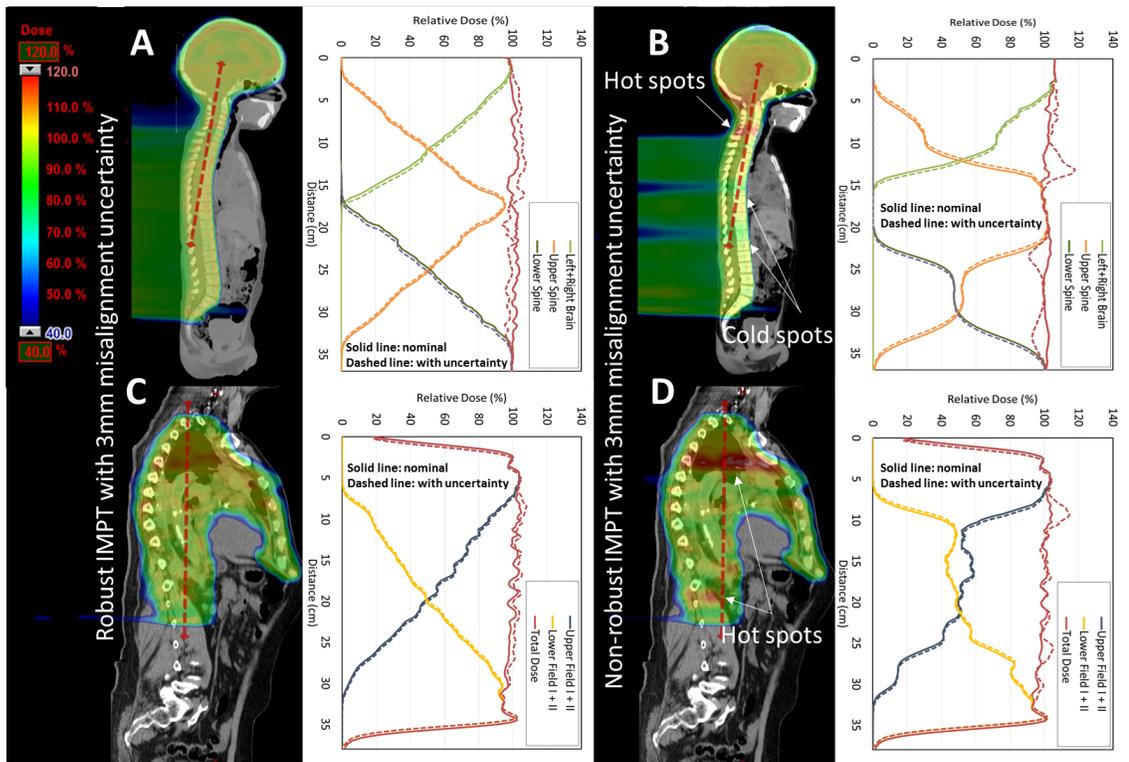


Figure 4.2: Dose color wash and corresponding dose profiles of the robust and conventional IMPT plans for the craniospinal irradiation patient (A, B) and the mesothelioma patient (C, D).

Figure 4.3 shows the dose profiles in robust IMPT plans for the CSI case with different junction sizes (8, 12, 16 and 26 cm) and a 3-mm misalignment error. And the uncertainty yield 9.9%, 5.4%, 4.5% and 2.6% dose deviation in the junction region for the IMPT plans with 8, 12, 16 and 26 cm junction size respectively. For a given uncertainty level, the dose deviation decreased as junction size increased. This result is also consistent with the results reported in previous study [83, 84]. The relationship between dose deviation, uncertainty and junction size can be rough simplified as

$$\text{Dose deviation}(\%) = \frac{\text{Uncertainty}}{\text{Junction Size}} \times 100\% \cdot \quad (4.2)$$

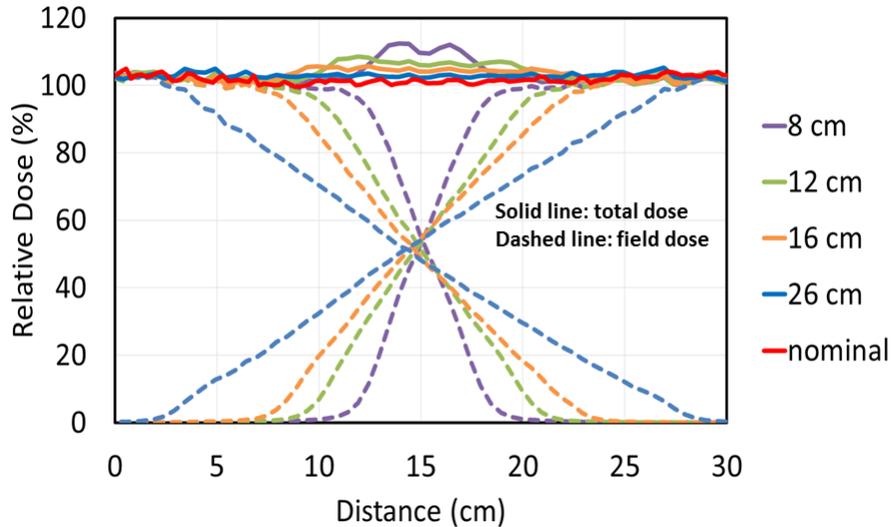


Figure 4.3: Dose profiles in junctions for the CSI IMPT plans with junction sizes of 8, 12, 16 and 26 cm and a longitudinal misalignment error of 3 mm per field (total, 6 mm).

Figure 4.4 demonstrates two strategies to increase the robustness of the misalignment errors: a robust IMPT plan with an 18-cm dose junction and a robust IMPT plan with a 7-cm dose junction and junction shifting. The second plan includes three subplans, each delivering 1/3 of the total dose. The total lateral dose profiles for the two plans are quite similar. Each subplan in the second plan has large dose deviations, but shifting the junction helps to spread the uncertainty. Thus, in general, the dose deviations of the two plans are similar. This result suggests that if the overlapping region is sufficiently enlarged, the shifting of junctions will not be necessary for the robust IMPT plan.

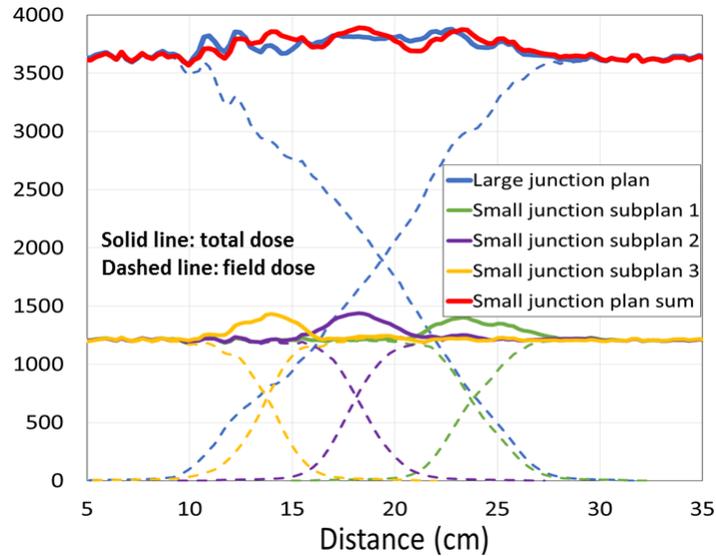


Figure 4.4: Robustness comparison between a robust IMPT plan with a large dose junction (18 cm) and a robust IMPT plan with a small dose junction (7 cm) and junction shifting for the CSI patient.

The dose volume histograms (DVHs) of robust and non-robust IMPT plans were illustrated in Figure 4.5. The tradeoff between target uniformity and robustness

between robust and non-robust IMPT plans was within 1.5% for two patient cases. For CSI patient, the HI of spinal cord, brain and cribriform plan were 1.041, 1.051 and 1.030 in robust IMPT plan compare to 1.036, 1.045 and 1.025 in non-robust IMPT plan. And the mean doses of left lens and right lens were increased from 8.9 Gy and 8.7 Gy to 10.3 Gy to 10.1 Gy from non-robust plan to robust plan. For the mesothelioma case, robust IMPT plan achieved similar plan quality of non-robust plan in nominal scenario. for two patient cases. For CSI patient, the HI of spinal cord, brain and cribriform plan were 1.041, 1.051 and 1.030 in robust IMPT plan compare to 1.036, 1.045 and 1.025 in non-robust IMPT plan. And the mean doses of left lens and right lens were increased from 8.9 Gy and 8.7 Gy to 10.3 Gy to 10.1 Gy from non-robust plan to robust plan. For the mesothelioma case, robust IMPT plan achieved similar plan quality of non-robust plan in nominal scenario.

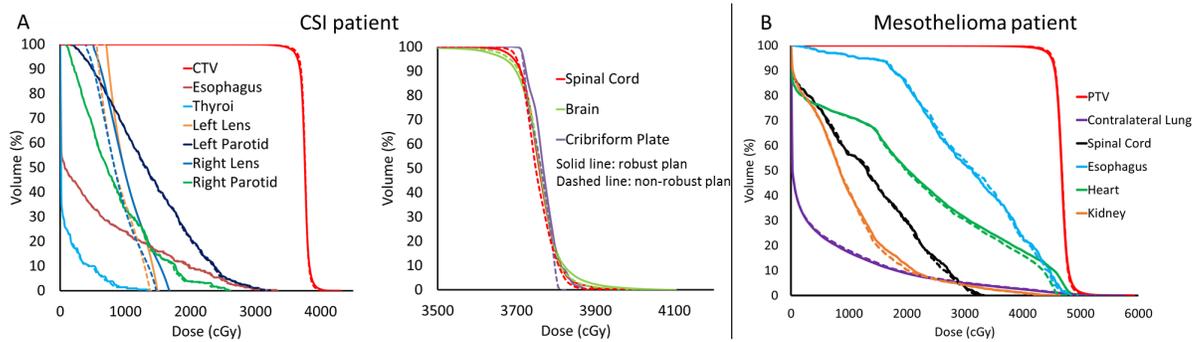


Figure 4.5: Dose volume histograms of robust and non-robust IMPT plans for craniospinal irradiation patient (A) and mesothelioma patient (B). Solid lines: robust IMPT plan; Dashed lines: non-robust IMPT plan.

4.4 Discussion

Robust optimization is aimed at reducing uncertainty in IMPT. Whereas previous studies only investigated setup errors in single-isocenter treatment plans [44], the current study provides, to our knowledge, the first demonstration of efficient integration of intrafractional setup errors for multi-isocenter fields into a general robust planning algorithm. Such robust optimization is especially important for treatment planning for large, complex and irregular-shape targets.

Many strategies have been proposed to handle field misalignment errors during treatment. For CSI treatment planning, a volumetric gradient dose optimization (GDO) methodology [90] was recently introduced for IMPT technology [83, 84]. The GDO method, which was initially introduced for volume modulated arc therapy (VMAT) planning [90], is a two-step manual planning approach. In this method, gradient volumes are generated in the overlap regions as four equally spaced sections. The first step is to optimize the first volume field so that the four gradient volumes receive 80%, 60%, 40% and 20% of the prescribed volume. The second step is to optimize the second field separately so that the four gradient volumes receive 20%, 40%, 60% and 80% of the prescribed volume. This method, which produces a tapered dose distribution in the junction regions, has several limitations. (i) In both VMAT and IMPT planning, the GDO method increases the optimization time significantly, since the manual GDO requires delineation of structures for optimizing the dose in the junction and running extra optimizations. So, an automatic process is desired. (ii) In GDO method, the assigned field dose in gradient volumes was not continuous,

so it is hard to produce a more general tapered dose distribution for large junction sizes. (iii) The GDO method applies single field optimization. This process cannot be used for mesothelioma cases since it often requires at least two fields for each isocenter. A non-optimal GDO solution for a large overlap region has been described for a VMAT optimization [90].

An important finding of the current study is that dose gradients that are low and tapered in field junctions can be achieved through a robust optimization that is much more general and simple than manual single-field optimization [83, 84]. Our approach overcomes the limitations of the GDO method in that it (i) is automated, (ii) can be used for any junction size and (iii) use multi-field optimization and can be used for large and complex targets. In addition, as the use of scanning beam proton therapy is increasing, the robust optimization planning method is being implemented in commercially available treatment planning systems, such as Eclipse V13.7 system (Varian Medical Systems, Palo Alto, CA). So, our general robust optimization method for multi-isocenter large field treatment plan can be easily applied in other proton therapy center. Our work is the first time to report the utilization of this automatic process for two distinct disease sites.

As shown above, robust IMPT greatly improves the efficiency of treatment over conventional IMPT. For the CSI treatment, one of the important results is that junction shifting was not necessary. For the mesothelioma treatment, the second isocenter was setup simply by shifting the couch during the treatment, since the plan is robust to intrafractional junction shifting. Currently, our center uses the robust

optimization planning approach for complex-target treatments and can perform CSI or mesothelioma IMPT in 45-min sessions.

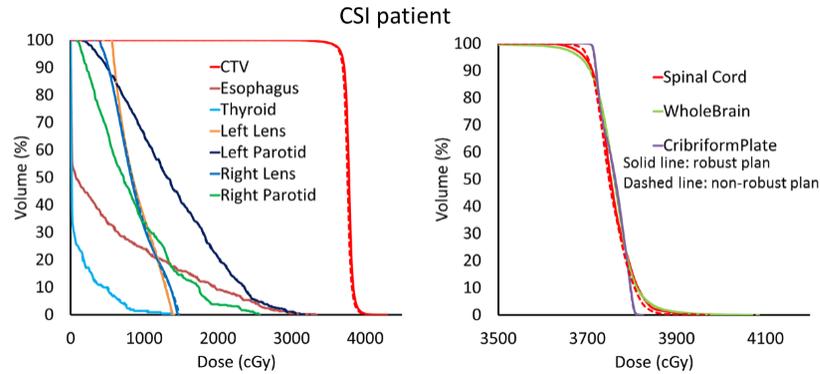


Figure 4.6: Dose volume histograms of robust IMPT plan with fixed brain fields uncertainty setup and non-robust IMPT plan for craniospinal irradiation patient. Solid lines: robust IMPT plan; Dashed lines: non-robust IMPT plan.

Although the robust optimization tools have been well developed, planners are still lack of experience in clinical application. The setup of uncertainty scenarios is crucial for the use of robust optimization in clinical practice. The inclusion of too many scenarios will increase the computation burden and thereby prevent optimization in an acceptable time frame, whereas the inclusion of too few scenarios may not guarantee robustness. How to balance the plan robustness and quality in nominal scenarios also need more experience. For example, in CSI case to increase the dose conformality in brain target and keep taped dose in junction. The uncertainty scenarios can change to two brain fields are kept still and the first and second spinal fields are shifted by $\hat{\Delta}3$ mm. The DVHs of this uncertainty setting are demonstrated in Figure 4.6. It shows that in brain target robust IMPT plan achieved the same plan quality of non-robust IMPT plan in nominal case. The selective robust optimization

strategies [91] also can apply to increase the dose uniformity in nominal case. In this study, we only discussed uncertainty scenarios to generate a robust field junction. The conventional interfractional patient setup uncertainties and system range uncertainties can also be integrated into treatment plan optimization.

4.5 Conclusion

A robust optimization approach for multi-isocenter large field IMPT has been developed. Several types of uncertainty during the CSI can be incorporated into optimization process. As a result, this approach can easily generate low gradient doses in field junctions and minimize dose deviations introduced by uncertainties.

Chapter 5

Two-stage Method for IMPT Beam Angle Optimization Incorporating Internal Organ Motion

Intensity modulated proton therapy (IMPT) is highly sensitive to uncertainties such as internal organ motion. Considering the motion uncertainty in beam angle selection is a complex optimization problem in IMPT treatment planning. To solve this problem efficiently, we developed a two-stage robust beam angle optimization (TSBAO) method for IMPT treatment planning. The goal of the first stage is to determine an appropriate initial number of beam angle clusters and assign angles to each of these clusters based on prior knowledge. A p-median algorithm is developed for beam angle clustering using two different measures: score function and similarity measure. The merit of an individual beam angle is associated with a score function incorporating the internal organ motion uncertainty. Another measure is to evaluate the similarity between two angles. For the second stage, a bi-level local neighborhood search (bi-LNS) algorithm is used to determine the final beam angle set for the treatment. Support vector machine (SVM) is used in bi-LNS to reduce the search space. Our methods were tested on four thoracic clinical cancer cases. TSBAO was

tested against several well-known heuristic methods found in literature, including a standalone LNS, simulated annealing (SA), genetic algorithms (GA), hybrid SA-LNS and hybrid GA-LNS. Results show that TSBAO consistently outperformed other methods in both objective value and CPU time. Furthermore, all TSBAO optimized treatment plans achieved more uniform and robust target dose distributions than the competitors did.

5.1 Introduction

As proton therapy has improved the therapeutic ratio by taking advantage of finite proton ranges in patients, it has been adopted for treating more and more disease sites by developing patient specific planning and delivery techniques [80, 81, 82, 83]. Intensity modulated proton therapy (IMPT) is an advanced proton delivery technique that uses modulated proton beams to produce three dimensional conformal dose distribution to cover the target [92, 93]. Similar to conventional photon based radiation therapy, the treatment planning procedure of IMPT also consists of two sequential optimization problems: beam angle optimization (BAO) which is to find an optimal beam angle set and fluence map optimization (FMO) which is to determine the optimal weights or intensities of beamlets. Often, the FMO problem is embedded into the beam angle selection problem.

The purpose of BAO of IMPT treatment planning is to select the most suitable beam angles to achieve uniform target dose coverage with a minimum radiation exposure to organs-at-risk (OAR). The BAO problem is a combinatorial optimization

problem which is known to be NP-hard [94]. In current clinical practice, treatment beam angles are manually selected by planners based on their clinical experience and/or a trial-and-error process.

Many researchers have reported methods to automatically select optimal beam angles in the literature. The majority of the algorithms combine BAO and FMO together to formulate the BAO problem as a mixed integer programming (MIP) problem. Sonderman D and Abrahamson PG [45] first proposed the MIP model for beam angle selection for the conventional conformal radiation therapy. Stain et al. [46] introduced an MIP model to solve the BAO for IMRT (intensity modulated radiation therapy), which incorporated the FMO in IMRT to guide the beam selection. In order to solve this MIP problem more efficiently, different algorithms have been proposed. Generic algorithm (GA), simulated annealing (SA) and particle swarm algorithm were designed to find the global optimal solution for BAO [49, 52, 53, 51]. Aleman et al. [56] introduced a neighborhood search algorithm for BAO to find a local optimal solution in a fast manner. Lim GJ and Cao W [38] proposed a two-phase method, which used a Branch and Prune (B&P) algorithm combined with a local neighborhood search method focusing on reducing CPU time. Based on the advantages of different algorithms, Lim GJ, Kardar L and Cao W [58] introduced a hybrid framework to improve the efficiency of BAO in IMRT. Although these algorithms can solve the beam angle selection problem, they still require a large number of iterations and the performance may be sensitive to the choice of initial solution and the algorithmic parameters [58].

To avoid excessive time for solving the BAO problem, prior knowledge was employed to guide the beam angle selection and to reduce the search space. Pugachev A and Xing L. [59] introduced scalar scoring functions to rank candidate beam directions by using beam's-eye-view (BEV) projections technique. Schreibmann E and Xing L [64] ranked the beam orientation based on dose-volume information for IMRT beam selection. Lim GJ, Holder A and Reese J [66] facilitated a Euclidean metric in a space of characteristic vectors for scoring the candidate beam directions. Based on these beam angle information, different algorithms have been introduced to solve the BAO problem [95, 73, 96]. These algorithms typically select or eliminate beam angles iteratively according to individual beam information. However, they are shown to be sensitive to initial parameter values [58].

In general, proton-based treatment plans are more sensitive to treatment uncertainties than the photon-based treatment plans such as IMRT. In IMPT treatment planning, proton beams deliver majority amount of dose at Bragg peaks that spans a small region. In consequence, even a small uncertainty during treatment can result in displacement of Bragg peaks to lead a significant dose deviation between actual dose and planning dose. It is especially for thoracic cancer patients that the respiratory motion during the treatment can induce severe target dose degradation [97, 37, 89]. In order to reduce the impact of uncertainties, various robust optimization models of FMO have been reported to generate a steady dose distribution under different uncertainty settings [37, 20, 44]. However, all these methods require extra calculation for uncertainty scenarios that can severely increase the computational efforts [98].

Solving the optimization models can be even harder if a robust FMO problem is integrated into beam angle optimization in IMPT. So, a prior knowledge-guided BAO algorithm incorporating uncertainty for IMPT treatment planning is desired because it can considerably reduce the computational time.

Therefore, we propose a two-stage beam angle optimization framework for solving the BAO problem in IMPT planning considering (1) uncertainty of internal organ motion and (2) prior knowledge of beam information. Within this framework, a p-median algorithm is used to cluster the candidate beam angles for simplify beam angle search. A score function is introduced as prior knowledge to evaluate the merit of a beam angle under uncertain internal organ motion to help expedite the beam angle selection process. To overcome the drawbacks of iteratively selecting beam angles, a bi-level local neighborhood search (bi-LNS) algorithm is developed. Bi-LNS incorporates a support vector machine (SVM) for candidate beam angle classification in order to reduce the search space.

The remainder of this paper is organized as follows. Section 5.2 describes the optimization models, solution methods, and the experiments setup. The results are shown in Section 5.3. Section 5.4 discusses the insights of results and conclusions are presented in Section 5.5.

5.2 Material and Methods

This section describes the proposed two-stage beam angle optimization method to select an optimal beam angle set for IMPT treatment. A given angle set is evaluated

by an imbedded 4D robust fluence map optimization model that determines optimal beamlet weights, while incorporating internal motion uncertainty. It is assumed that the number of total candidate beams ($N = |A|$) is finite and the number of final beam angles for the treatment (k) is given as an input.

5.2.1 4D Robust Fluence Map Optimization Model

For thoracic cancer, the tumor inside the chest usually moves significantly due to respiratory motion during the radiation treatment. The magnitude of internal organ motion can be measured by four-dimensional computed tomography (4DCT) data. The whole 4DCT includes 10 CT datasets representing temporal phases of a respiratory cycle. Phase CT_0 is the maximum inhale phase, while Phase CT_{50} is the maximum exhale phase. All other phases are between these two extremes. The average CT (CT_{ave}) is calculated using the mean CT numbers of the 10 temporal CT phases of 4DCT at each pixel location. In conventional IMPT, only CT_{ave} is used for treatment planning, that is not sufficient to guarantee a robust dose distribution in target area when accounting internal organ motion [99]. Hence, we optimize the beamlet weights based on more 4DCT data (i.e., including CT_{ave} and two extreme conditions CT_0 and CT_{50}) to yield a robust dose distribution under the internal organ motion situation.

The robust FMO model is formulated based on a worst case dose model [44], which was introduced by Lomax et al. [97]. In order to calculate the worst-case dose distribution, different dose scenarios are calculated independently on different CT

phases (i.e., CT_{ave} , CT_0 and CT_{50}). In the target volume, the worst-case dose of this voxel is set as the minimum of all dose scenarios if the actual dose is lower than the prescription dose; if the actual dose is higher than the prescription, the worst-case dose is the maximum of all dose scenarios. If the voxel is located in the surrounding OARs and normal tissues, the dose of this voxel is set as the maximum of all dose scenarios.

For illustration, suppose that there are R dose scenarios based on different CT phases. The dose deposit on voxel i under scenario r is calculated as

$$D_i^r = \sum_{a \in \bar{A}} \sum_{j \in B_a} d_{i,a,j}^r \omega_{a,j}, \quad \forall i \in V_\Omega, \Omega \in T \cup O, r \in \{R\}, \quad (5.1)$$

where $\omega_{a,j}$ is weight or intensity of beamlet j in beam angle a , $a \in \bar{A}$, $j \in B_a$ and it is the primary decision variable of the FMO model. Set \bar{A} contains the beam angles used in a treatment plan; B_a is the beamlet set for beam angle a ; $d_{i,a,j}^r$ represents the dose contribution to voxel i from beam angle a and beamlet j ; T and O are set of target structure and OAR, respectively. The voxel set within structure Ω is denote as V_Ω .

So, the worst-case dose can be calculated as

$$D_{i \in V_\Omega} = \begin{cases} \max_r D_i^r, & D_i^r \geq P_\Omega \\ \min_r D_i^r, & D_i^r < P_\Omega \end{cases} \quad \Omega \in T, \quad r \in R, \quad (5.2)$$

$$D_{i \in V_\Omega} = \max_r D_i^r, \quad D_i^r \geq P_\Omega, \quad \Omega \in O, \quad r \in R, \quad (5.3)$$

where, P_Ω denotes the control parameter for the structure Ω , i.e., dose prescription of the target and dose limitation of OARs. Thus, using this dose formulae (1-3), the robust FMO model can be expressed as

$$\begin{aligned} \min_{\omega} \quad & F(D) = \sum_{\Omega \in T} \left(\frac{1}{|V_\Omega|} \left(\sum_{i \in V_\Omega} \lambda_\Omega^+ (\max_r D_i^r - P_\Omega)_+^2 + \sum_{i \in V_\Omega} \lambda_\Omega^- (P_\Omega - \min_r D_i^r)_+^2 \right) \right) \\ & + \sum_{\Omega \in O} \left(\frac{1}{|V_\Omega|} \left(\sum_{i \in V_\Omega} \lambda_\Omega^+ (\max_r D_i^r - P_\Omega)_+^2 \right) \right), \\ \text{s.t.} \quad & \omega_{aj} \geq 0, \end{aligned} \quad (5.4)$$

where $(\cdot)_+$ represents $\max\{\cdot, 0\}$. Parameters λ_Ω^+ and λ_Ω^- are the penalty coefficients of structure Ω on hot and cold spot, respectively; and $|V_\Omega|$ is the number of voxels in structure Ω . The first and second terms penalize the maximum violations between different scenarios on the target for both hot and cold spots to achieve a uniform dose distribution. The third term penalizes overdose on OARs. The penalty coefficient of each objective function can be adjusted according to the clinical requirements.

5.2.2 Beam Angle Clustering

The first stage of the TSBAO method is to cluster the candidate beam angles according to the clinical merits of each angle. Hence, similar beam angles are grouped into the same cluster, and an angle is selected as a representative beam angle for each cluster, which is called a centroid. The p -median method is commonly used

for solving the clustering problem [100]. The beam angle clustering problem using p -median method can be formulated as

$$\begin{aligned}
 x_a &= \begin{cases} 1 & \text{if beam angle } a \text{ is selected for treatment plan} \\ 0 & \text{otherwise,} \end{cases} \\
 y_{aa'} &= \begin{cases} 1 & \text{if beam angle } a' \text{ is allocated with } a \\ 0 & \text{otherwise,} \end{cases}
 \end{aligned} \tag{5.5}$$

$$\begin{aligned}
 \min \quad & W_M = \sum_{a \in A} \sum_{a' \in A} S(a) \theta(a, a') y_{aa'} \\
 \text{s.t.} \quad & \sum_{a \in A} y_{aa'} = 1, \\
 & \sum_{a \in A} x_a = M, \\
 & x_a \geq y_{aa'}, \\
 & x_a \in \{0, 1\}, y_{aa'} \in \{0, 1\}, \\
 & a, a' \in \{A\},
 \end{aligned} \tag{5.6}$$

where, A denotes the total candidate beam angle set; M is the total number of beam angle clusters; $S(a)$ is beam angle score function to describe the potential of beam angle a will be selected in treatment plan. Likewise, $\theta(a_a, a_{a'})$ is a function that measures the similarity between beam angles a and a' . The details of these two functions are described in the following sections.

5.2.2.1 Beam Angle Scoring

To measure the potential of a beam angle to be selected in the final treatment plan, we introduce a revised beam angle score function to evaluate each candidate beam angle. This score function is calculated based on dose contribution information from each beam angle. It consists of OAR sparing score and target robustness score. Both of these scores include multiple sub-scores which measure the ratio of dosimetric deposition in a structure over the total dose contribution from the beam angle under uncertainty, and they are expressed as

$$S(a)_\Omega^r = \frac{1}{|V_\Omega|} \left(\frac{\sum_{i \in V_\Omega} \sum_{j \in B_a} d_{i,j}^r}{\sum_{i \in V_T \cup V_O} \sum_{j \in B_a} d_{i,j}^r} \right) \quad a \in A, \quad \Omega \in \{T, O\}, \quad r \in \{R\}, \quad (5.7)$$

where, $\sum_{i \in V_T \cup V_O} \sum_{j \in B_a} d_{i,j}^r$ is the total dose contribution from angle a in scenario r , and $\sum_{i \in V_\Omega} \sum_{j \in B_a} d_{i,j}^r$ is the dose deposition in structure Ω from angle a in scenario r . A smaller value of the ratio is preferred.

For the OAR sparing, the score is defined as a linear combination of sub-scores of OAR structures in the nominal scenario (CT_{ave})

$$S(a)_{\text{OAR}} = \sum_{\Omega \in O} \lambda_\Omega S(a)_\Omega^{\text{nominal}}. \quad (5.8)$$

Coefficient λ_Ω is an avoidance factor for different structures and it is consistent with the penalty coefficient in FMO model (Eq. 5.4).

For the target robustness score, the function calculates the maximum different

between sub-scores of target structures in different dose scenarios

$$S(a)_{\text{Target}} = \sum_{\Omega \in \mathbf{T}} \lambda_{\Omega} \cdot \max \left| S(a)_{\Omega}^r - S(a)_{\Omega}^{r'} \right| \quad r, r' \in R. \quad (5.9)$$

Therefore, the total beam angle score is the summation of the OAR sparing score and the target robustness score,

$$S(a) = S(a)_{\text{OAR}} + S(a)_{\text{Target}}. \quad (5.10)$$

In this approach, a beam angle with a lower score is preferred for the treatment plan.

5.2.2.2 Similarity Measure of Beam Angles

The similarity between two beam angles is calculated by measuring the Euclidean distance between sub-scores of two beam angles in the nominal scenario. The similarity function is defined as

$$\theta(a, a') = \sqrt{\sum_{\Omega \in \{\mathbf{T}, \mathbf{O}\}} (S(a)_{\Omega}^{\text{nominal}} - S(a')_{\Omega}^{\text{nominal}})^2} \quad a, a' \in A. \quad (5.11)$$

5.2.2.3 Estimating the Number of Beam Angle Clusters

In the p -median problem, the number of clusters M is an input parameter to the optimization model and the value of M has a direct impact on the objective function value (W_M^*). Therefore, choosing a right value of M is critical for the p -median problem. Different methods for selecting an appropriate number of clusters have been reported in the literature [101, 102, 103].

Figure 5.1 depicts the normalized objective value (W_M^*/W_1^*) as a function of M . Note that, the value of W_M^*/W_1^* decreases as M increases and the decreasing rate gradually slowed down. According to Lim et al. [103], The slope is defined as

$$\Psi_M = \frac{180}{\pi} \arctan\left(\frac{W_{M-1}^* - W_M^*}{W_1^*} N\right), \quad 1 < M < N, \quad (5.12)$$

where, $N = |A|$ is the total number of candidate beam angles. The slope (Ψ_M) as a function of M is demonstrated in Figure 5.1 (the line with dots). The transition point of a slope can be defined as the point where Ψ_M changes from the value $\geq 45^\circ$ to the value $< 45^\circ$. For example, the number of clusters is 6 for the data showed in Figure 5.1. In our BAO problem, because we need to find k beam angles for the treatment plan, so the number of clusters is set to be at least k (i.e., $M \geq k$).

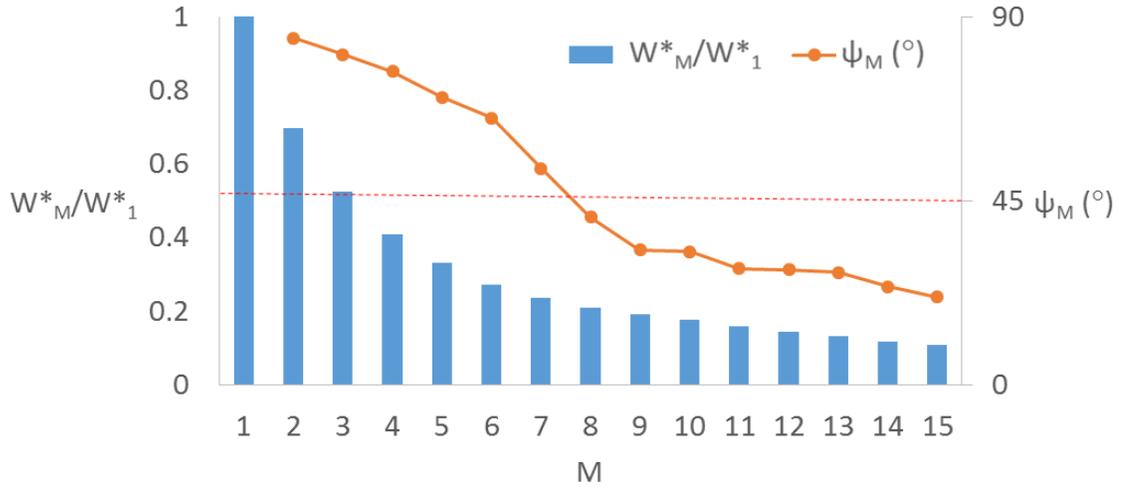


Figure 5.1: Normalized p-median objective function value (blue bar) and slope (orange dotted line) as a function of cluster number M .

5.2.3 Bi-level Local Neighborhood Search (Bi-LNS) for BAO

In this section, the goal is to search the final k beam angles for the IMPT treatment planning based on the clustering results from Section 5.2.2.

5.2.3.1 Local Neighborhood Search

The local neighborhood search (LNS) algorithm has been widely used to solve the beam angle optimization problem in radiation treatment planning [58, 98]. In general, LNS finds a local optimal solution fast subject to the neighborhood definition for a given starting feasible solution. The general procedure of LNS is described in Algorithm 2. Let \bar{A} be a starting solution, where $\bar{A} \subset A$. We solve the FMO and obtain the objective value $z(\bar{A})$. Then we search the neighborhood of $\bar{A}(N(\bar{A}))$ and solve the FMO for each neighbor. If a new solution (\bar{A}') yields a better objective value, such that $z(\bar{A}') < z(\bar{A})$, both the solution and the objective value are updated accordingly then repeat the search process. Otherwise the algorithm stops, and the current solution \bar{A} is the final solution.

Algorithm 2: Local neighborhood search

- 1 Initialization: Generate \bar{A}_0 , solve FMO and find $z(\bar{A}_0)$, $i = 0$, $\bar{A}_i^* = \bar{A}_0$,
 $z_i^* = z(\bar{A}_0)$;
 - 2 **do**
 - 3 $i = i + 1$;
 - 4 Generate $N(\bar{A}_{i-1}^*)$ the neighborhood of \bar{A}_{i-1}^* ;
 - 5 Enumerate all beam angle sets in $N(\bar{A}_{i-1}^*)$, and save the lowest objective
 value as z_i^* , $\bar{A}_i^* \leftarrow \arg \min\{z_i^*\}$;
 - 6 **while** *Stopping criteria are not met*;
 - 7 **Stop.** $(z_{i-1}^*, \bar{A}_{i-1}^*)$ is the final local optimal solution.
-

The neighborhood of beam angle set \bar{A} is achieved by exchanging one or more

beams between \bar{A} and the rest of candidate beam angles ($A \setminus \bar{A}$). The neighborhood size can increase exponentially if two or more beam changes are allowed to form a new neighborhood. In order to avoid this exponential growth of the neighborhood size, we adopt a one-angle-exchange algorithm [104] in our approach. The beam angle set in a neighborhood can have only one different beam compared to the central beam set \bar{A} . The neighborhood of \bar{A} is defined as

$$N(\bar{A}) = \{\bar{A}' : \bar{A}' = (\bar{A} \cup \{a'\}) \setminus \{a\}\}, \quad \text{for } a' \in \delta(a), \quad a \in \bar{A}, \quad (5.13)$$

where $\delta(a)$ is a neighborhood of given beam angle a . The neighbor angles are defined in the following sections according to specific requirements in different search levels.

5.2.3.2 Cluster Level LNS

As the result of Section 5.2.2, the total candidate beam set (A) has been grouped into M clusters and found a centroid beam angle for each cluster. We denote this centroid beam angle set as A_C , $|A_C| = M$. Because the centroid beam angles are representative angles of the total candidate beam angle set and M is typically much smaller than N , our strategy is to determine the optimal solution \bar{A}_C^* with k beam angles in the centroid beam angle set. This solution is used as a starting point for local adjustment to search for the final beam angle solution in the candidate beam set.

In this step, we use the LNS algorithm described in Section 5.2.3.1 to search for the solution \bar{A}_C^* . To find the optimal solution, we define the neighbor angle set $\delta(a)$ is

the whole centroid angle set A_C ,

$$\delta(a) = \{a : a \in A_C\} . \quad (5.14)$$

Because the $M \ll N$, the cluster level LNS can still be solved in an acceptable time even if the neighborhood includes all centroid angles. For example, a problem with $N = 36$, $M = 6$ and $k = 3$. In worst case, cluster level enumerates all the 3 beam angle combinations in set A_C , the maximum iteration is $C_3^6 = 20$. Although the LNS algorithm cannot guarantee optimal, this neighborhood definition ensures the algorithm can search a sufficient space to find the optimal solution or a solution very close to the optimal in the centroid beam angle set.

5.2.3.3 SVM for Reducing Feasible Region

In the cluster level LNS, the solution is based solely on the selected centroid beam angles. To explore a better solution, we need to perform another search in the total candidate beams. However, commonly the number of candidate beams is much larger than centroid beams. In this case, a preprocessing step is added to reduce the search space based on the results of the cluster level LNS.

The SVM is a type of learning algorithm that is commonly used for classification [105, 106]. Given a set of training data, labeled for belonging to one of two categories, the algorithm outputs an optimal hyperplane that assigns new input data into appropriate category.

In our study, the centroid beam angle set is used as the training data. For each

centroid angle, the sub-scores of each structure in nominal case are used as an input example, it can be denoted as $I_a = \{S(a)_{\Omega}^{\text{nominal}} \mid \Omega \in \{T, O\}\}$, $a \in C$. And the input points are labeled by variable L_a , where the value was decided by to the solution \bar{A}_C^* found in cluster level LNS. If the centroid angles belonging to the optimal solution in the cluster level LNS (i.e., $a \in \bar{A}_C^*$), we set $L_a = +1$, otherwise, the beam angle was labeled as $L_a = -1$. The decision function implemented by the SVM can be written as

$$f(\vec{I}) = \text{sgn} \left(\sum_a L_a \alpha_a \cdot K(\vec{I}, \vec{I}_a) + \beta \right), \quad (5.15)$$

where K is a kernel function and the coefficients α_a and β are determined by maximizing the following problem

$$\begin{aligned} \max \quad & \sum_a \alpha_a - \frac{1}{2} \sum_a \sum_{a'} \alpha_a \alpha_{a'} L_a L_{a'} K(\vec{I}_a, \vec{I}_{a'}) \\ \text{s.t.} \quad & \alpha_a \geq 0 \quad \text{and} \quad \sum_a \alpha_a L_a = 0 \end{aligned} \quad (5.16)$$

The Radial Basic Function (RBF) kernel [107] is applied in our algorithm.

After SVN training, the rest of the non-centroid beam angles in candidate beam angle set are classified. If a beam angle is classified as -1, it would be removed from the final candidate beam angle set A_f .

5.2.3.4 Candidate Beam Level LNS

We obtain a feasible solution \bar{A}_C^* and the final candidate beam angle set A_f from previous steps. In the last step, we perform another local neighborhood search

to find the final solution. In this step, the initial beam angle set for LNS is the optimal solution \bar{A}_C^* found in section 5.2.3.2. The local neighborhood $\delta(a)$ in candidate beams is the angles geometrically located within a 's adjacent area. It is defined as

$$\delta(a) = \{a' : a' = [a - \varepsilon, a + \varepsilon], \text{ for } a' \in A_f \text{ and } a' \notin \bar{A}\} \text{ mod } 360^\circ, \quad (5.17)$$

where ε is a parameter that determines the size of $\delta(a)$, a larger ε represents a larger neighborhood of the center angle. For example, for a beam angle with parameters $a = 0$ and $\varepsilon = 3$, the neighborhood is $\{330, 340, 350, 10, 20, 30\}$.

5.2.4 Patient Studies and Setup

The solution methods were tested on two lung cancer cases and two esophagus cases. For all cases, 36 equispaced coplanar beam angles were considered as candidate beam angles for selecting 3 beam angles, which are normally used in clinical treatment planning. Table 5.1 shows the prescribed dose to the internal clinical target volume (ICTV) for lung cases was 66 Gy and for esophageal cancer was set at 50.4 Gy. To achieve a uniform dose, the penalty weight for target was set much higher than the penalty of OARs. The numbers of voxels in major volumes of interest are listed in Table 5.2. 4DCT was acquired for each case. The nominal dose was evaluated on the average phase of the 4DCT (CT_{ave}) and the doses on the maximum inhale and exhale phases (CT_0 and CT_{50}) were considered as uncertainty scenarios.

Table 5.1: Prescriptions and penalty weights used for IMPT plan optimization.

Cancer Type	Prescription (Gy)	Structure	Penalty
Lung	66	ICTV	100
		Cord	1
		Esophagus	1
		Heart	1
		Lung	1
Esophagus	50.4	ICTV	100
		Cord	1
		Esophagus	1
		Heart	1
		Lung	1

Table 5.2: Number of voxels within each structures for different cases

Case	Number of voxels				
	ICTV	cord	Esophagus	Heart	Lung
Lung I	9974	773	741	10874	65073
Lung II	4537	466	990	13939	89165
Esophageal I	2247	1143	352	5850	18472
Esophageal II	1343	510	392	7374	38538

In order to analyze the convergence properties, we implemented a standalone LNS, a simulated annealing (SA) algorithm, a genetic algorithm (GA) and a hybrid SA-LNS and a GA-LNS to compare with TSBAO method. Table 5.3 lists the stopping criteria for all tested approaches. The neighborhood size $\varepsilon = 3$ was adopted for local neighborhood search algorithm. For result analysis, we exhaustive enumerated all 3-beam combinations out of 36 candidate beams.

The proposed TSBAO method and all comparison algorithms were implemented in C++ and GAMS. All computations in this study were performed on a 64-bit Linux server with dual ten-core Intel Xeon 3 GHz processor and 364 GB memory.

In all plans the dose was normalized to 95% of ICTV receiving 100% of the

Table 5.3: Stopping criteria for different BAO algorithms

Algorithm	Stopping criteria
SA	Number of iterations exceeds 500; or These is no better solution found in 20 successive iterations
SA-LNS	SA: Number of iterations exceeds 100; or These is no better solution found in 10 successive iterations LNS: These is no better solution found in local neighborhood
GA	Number of generations exceeds 50; or These is no better solution found in 20 successive iterations
GA-LNS	GA: Number of generations exceeds 20; or These is no better solution found in 10 successive iterations LNS: These is no better solution found in local neighborhood
LNS	These is no better solution found in local neighborhood
TSBAO	These is no better solution found in local neighborhood

prescribed dose. To evaluate the plan quality and robustness, we calculated the plan on CT_{ave} , CT_0 and CT_{50} to yield the corresponding doses $D_{CT_{ave}}$, D_{CT_0} and $D_{CT_{50}}$. The ΔD_{95} of ICTV, calculated based on max differences in D95 of ICTV between dose scenarios, was used to quantify the overall plan robustness. A heterogeneity index (HI) was computed for the ICTV to measure the dose uniformity. The HI was defined as follows

$$HI = D_5 / D_{95}, \quad (5.18)$$

where D5 and D95 correspond to the doses delivered to 5% and 95% of ICTV, respectively. A larger HI indicates a greater degree of dose heterogeneity. For normal structures, cord maximum dose, esophagus V40 (i.e., the volume of esophagus received at least 40 Gy dose) and V50, heart V30 and V40, lung V5, V20 and mean dose were evaluated.

5.3 Results

Figure 5.2 shows the beam angle score as a function of beam angle for lung case I. The centroids selected by p -median clustering for this case are indicated with triangle. It demonstrates that the clustering algorithm attempted to find the beam angles with local minimum scores as centroids. And all the centroids are kept away from each other. Total 6 beam clusters were grouped for this case. The beam angle with a local minimum beam score is prone to be selected as the centroid of the cluster.

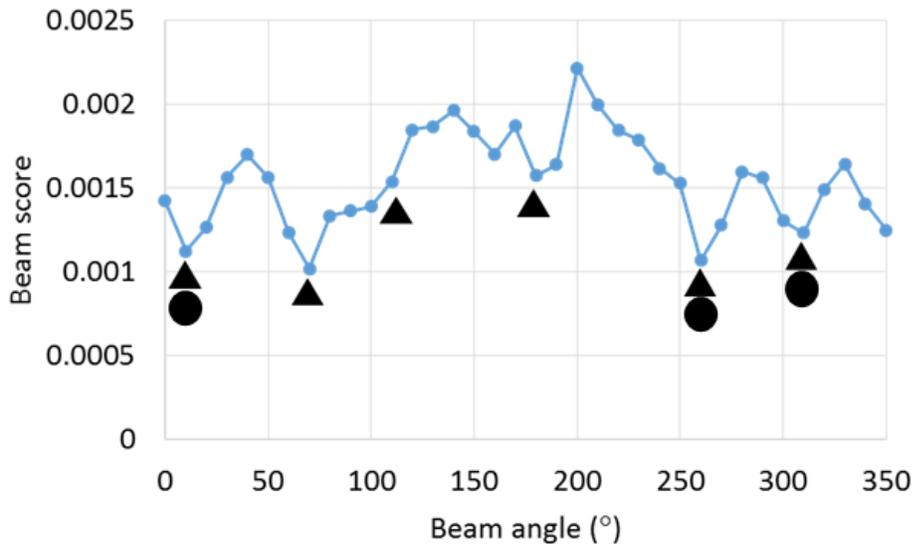


Figure 5.2: Beam angle score of 36 candidate beam angles (dotted line), the centroid beam angles (triangle) and the solution found by cluster level LNS (black dot) for lung case I.

To investigate which beam angle is more preferred to be selected in high quality IMPT plans, we optimized IMPT plans for all possible combinations of 3 beams out of a pool of 36 equispaced coplanar beams. Based on the objective function value, the best 200 (i.e., about 3% of total $C_3^{36} = 7140$ combinations) IMPT plans were selected

for the analysis. Figure 5.3A shows the probability of beam angle appearance in the best 200 IMPT plans for the lung case I. The beam angles 20 and 310 are two peaks and their adjacent angles are most probable regions to be selected in the top quality plans, while the beams from 90 to 200 are not preferred.

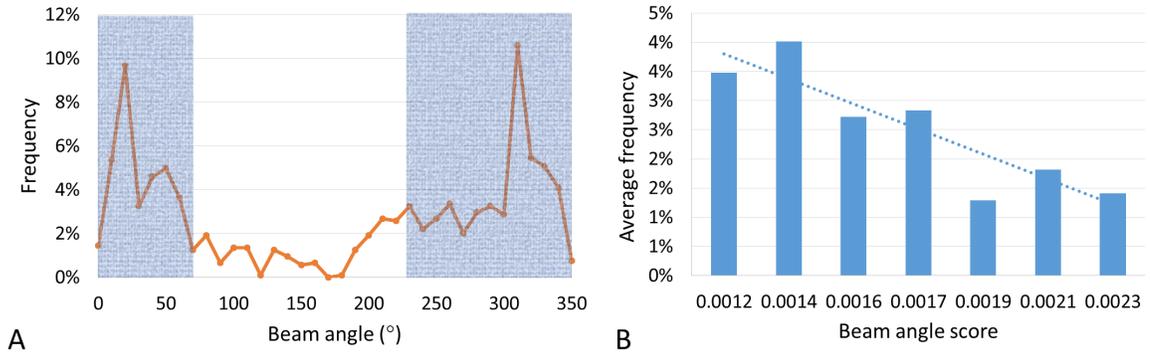


Figure 5.3: Beam angle appearance probability in the best 200 IMPT plans (A), and the histogram of average beam angle appearance probability (B) for lung case I.

The beam angles in the solution of the cluster level LNS are 10, 260 and 310, which is indicated in Figure 5.3 by black dot. And after SVM classification, the final candidate beam angle set for the candidate beam level LNS is indicated by shadow-covered area in Figure 5.3A, which includes angles 0-50 and 230-350. Around 81.5% beams appeared in the best 200 plans are included in A_f . All the high frequency beams such as angle 20 and angle 310 are included in the final candidate beam set.

To analyze the relationship between the beam angle score and the appearance probability, we plot the average beam angle appearance probability versus the beam angle score for lung case I in Figure 5.3B. The beam angles are divided into several groups according to their beam angle scores. The number of beam score interval is

calculated based on Sturges' rule ($I = 1 + \log_2 N$). In our case $N = 36$, so the number of interval is 7. The average appearance probability for SI , $\bar{P}(SI)$, is calculated as

$$\bar{P}(SI) = \frac{\sum_{S(a) \in SI} P(a)}{N_{SI}}, \quad (5.19)$$

where, $P(a)$ is the appearance probability of beam angle a ; N_{SI} is the number of beam angles with the beam angle score within a beam angle score interval. Empirically, beam angles with a lower score are associated with a higher appearance probability in the final treatment plan. It shows the effectiveness of beam angle score function. However, due to our beam score is still based on individual beam, it cannot reflect the potential of the beam be selected in good plans very precise. For example, the appearance probability of a beam with score 0.0012 is lower than the beam with score 0.0013. Overall, there is a negative correlation between the beam angle score and the appearance probability (blue dot trend line).

We applied the TSBAO to all test cases, the centroids of clusters and the final solutions found by this method for all 4 tested cases are demonstrated in Figure 5.4. For all cases, 6 or 7 clusters were estimated as a proper number by our algorithm. The centroids distribution were $\{10,70,110,180,260,310\}$, $\{0,30,60,130,160,200,290\}$ for lung case I and II, and $\{0,90,150,190,240,300,330\}$, $\{10,40,90,190,230,290\}$ for esophageal case I and II, accordingly. The corresponding solutions found by TSBAO were $\{20,260,320\}$, $\{200,270,350\}$, $\{90,200,290\}$ and $\{20,180,300\}$, respectively. All the beam angles in the final solutions were located on the centroid angle or very close

to the nearby centroid.

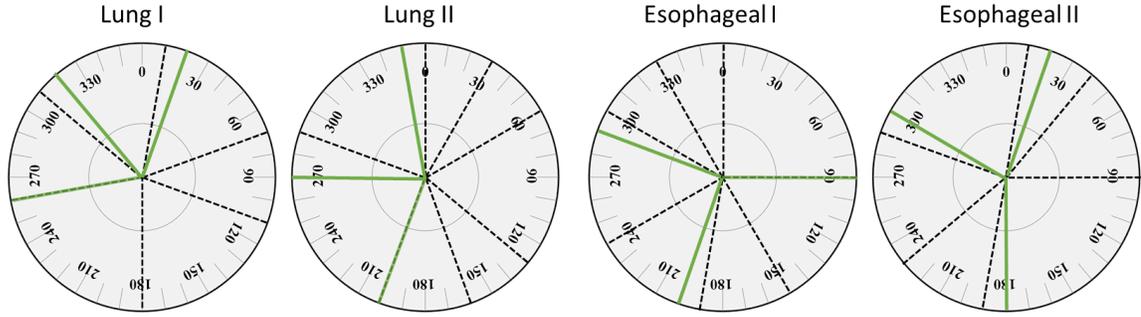


Figure 5.4: Centroid distributions and final solutions obtained by TSBAO method for 4 clinical cases. Centroids are denoted by dashed lines and TSBAO solutions are solid lines.

In order to analyze the convergence rate of the TSBAO algorithm, we tracked objective function values improving at each iteration until the algorithm terminates. Figure 5.5 shows the progression of objective convergence for the TSBAO, standalone LNS, SA, SA-LNS, GA and GA-LNS algorithms implemented on four tested patient cases. Note that, globally optimal solutions were found by exhaustive search and also shown by dashed-dotted lines in Figure 5.5. The figure demonstrated that TSBAO method found global optimal solutions for Lung I, II and esophagus case I. Compared to other algorithms, TSBAO method also converged closer to the global optimal for the esophagus case II. From computational time aspect, TSBAO was also faster than all heuristic and hybrid BAO algorithms and similar with single LNS algorithm. For all cases, TSBAO spend 50-70 minutes consistently for all tested cases which was around half of the time cost by SA, GA, SA-LNS and GA-LNS. For SA and GA, the results varied case by case. And in general, these algorithms were hard to find solutions close to the global optimal before they meet the stopping criteria. The

hybrid algorithms, i.e., SA-LNS and GA-LNS, were better than standalone SA and GA from both objective value and time spending point of view. The hybrid algorithms also can find the solution close to the global optima. In tested cases, SA-LNS found the global optimal for two lung cases. GA-LNS also found the global optimal for Esophagus case I, and the result of Lung case II was very close to the global optima.

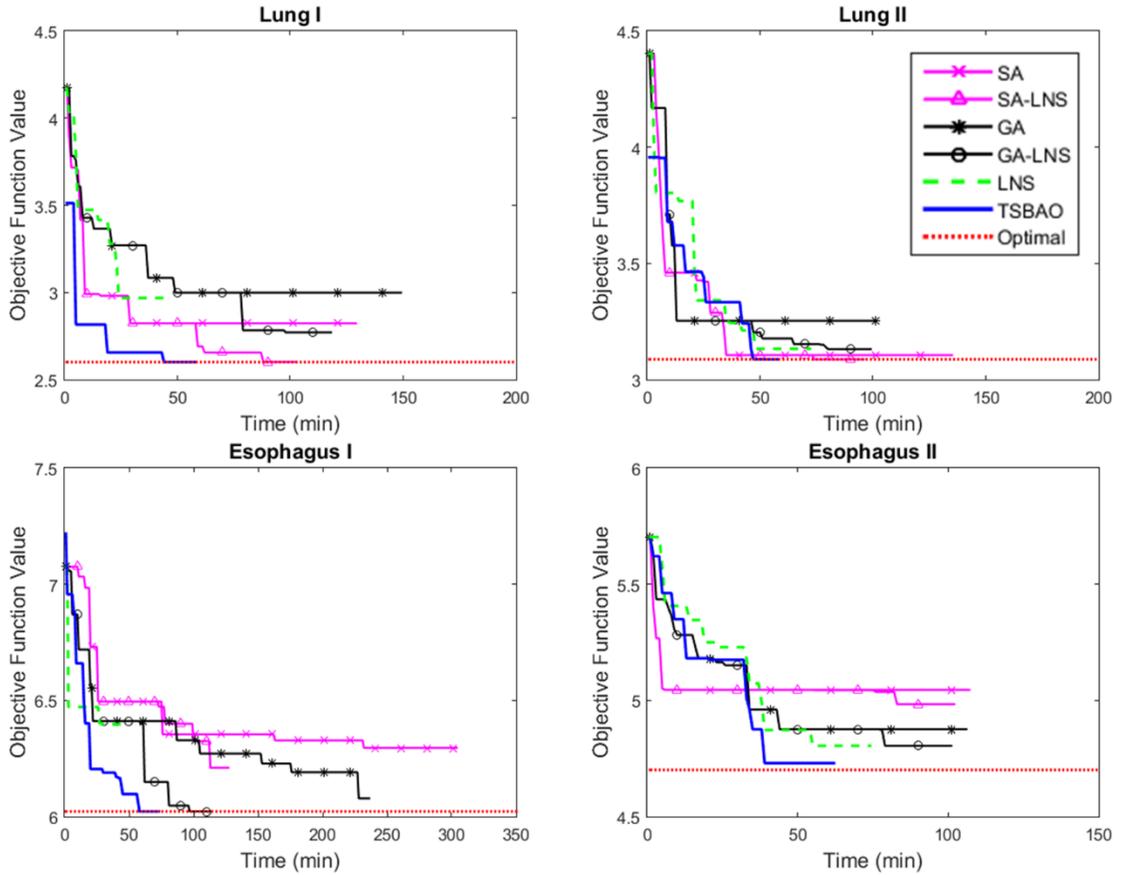


Figure 5.5: Converge comparison for TSBao (solid lines), standalone LNS (dashed lines), SA (cross solid lines), SA-LNS (triangle solid lines), GA (asterisk solid lines), and GA-LNS (circle solid lines). Global optima are shown by dotted lines.

Table 5.4 lists the dose volume results obtained from the BAO solution methods for four clinical cases. We can see from Table 5.4 that better target dose uniformity

and robustness were achieved by TSBAO for all tested cases. However, for OAR sparing the results varied case by case. TSBAO achieved best OAR sparing for majority of indices of lung II and esophagus I. For lung case I and esophagus II, the plans found by STBAO only achieve best lung sparing between all tested algorithms.

Figure 5.6 compare the beam angles selected under different solution methods for the IMPT lung case I with 36-angle configuration. The global optimal solution for this case was $\{20,260,320\}$. The TSBAO method found the global optimal in the end. All other tested algorithms were starting at the equally spaced beam angles $\{0,120,24\}$ and the final results were significantly different between each other. Specifically, SA-LNS also found the global optimal solution. The optimized beam angle configurations showed that LNS algorithm was significantly influenced by its initial point. For example, angles found by single LNS are $\{20,120,220\}$, which are close to the equal spaced starting point. And for all hybrid algorithms (i.e., SA-LNS and GA-LNS), the final solutions were also close to the results got from pure SA and GA. For example, the solution of SA is $\{30,260,310\}$ and the solution of SA-LNS is $\{20,260,320\}$.

5.4 Discussion

Different algorithms have been reported to solve the BAO problem in radiation therapy, but these algorithms need a long time to find the solution or the performance is influenced by starting point or input parameters. This paper, we focuses on a two-stage BAO method which considering prior knowledge of beam information and

Table 5.4: Dose volume data of 3 beam IMPT plans from different BAO methods for four clinical cancer cases.

	Lung I										Lung II			
	STBAO†	SA	SA-LNS†	GA	GA-LNS	LNS	STBAO†	SA	SA-LNS†	GA	GA-LNS	LNS		
ICTV	$\Delta D95$ (Gy)	3.05*	3.25	3.05*	3.54	3.36	3.35	2.24	2.04*	2.71	2.06	2.24		
	HI	1.0730*	1.0785	1.0730*	1.0825	1.0815	1.0819	1.0696*	1.0698	1.0777	1.0711	1.0744		
Cord	Max (Gy)	20.11	20.31	20.11	17.29	16.80*	20.31	17.41	15.99*	17.44	16.67	18.37		
	V40 (%)	2.28	1.93*	2.28	2.46	2.46	2.11	22.96	22.96	22.96	22.96	22.70*		
Esophagus	V50 (%)	1.41	1.58	1.41	1.23*	1.58	1.58	21.43*	21.68	21.43*	21.43*	21.94		
	V30 (%)	30.19	28.15*	30.19	30.31	26.87	25.95	19.88*	20.78	19.88*	23.15	21.20		
Heart	V40 (%)	23.46	20.35	23.46	21.16	18.90*	19.15	12.08*	12.46	18.33	15.72	12.73		
	V5 (%)	20.10	26.24	20.10	24.82	20.27	17.68*	51.05*	58.16	56.80	56.00	58.38		
Lung	V20 (%)	8.32*	10.69	8.32*	9.19	8.61	11.15	28.06	28.76	26.14	29.75	26.07*		
	Mean (Gy)	5.34*	6.51	5.34*	6.14	5.61	5.96	12.90*	13.83	13.73	13.31	13.13		

	Esophagus I										Esophagus			
	STBAO†	SA	SA-LNS	GA	GA-LNS†	LNS	STBAO†	SA	SA-LNS	GA	GA-LNS	LNS		
ICTV	$\Delta D95$ (Gy)	0.79*	1.49	1.06	0.99	0.79*	1.79	0.52*	0.53	0.66	0.53	0.53		
	HI	1.0872*	1.0921	1.0925	1.0885	1.0872*	1.0953	1.0830*	1.0868	1.0909	1.0896	1.0913		
Cord	Max (Gy)	36.68	38.79	32.98*	35.96	36.68	39.85	14.99	15.05	18.21	16.62	14.78*		
	V40 (%)	21.31	20.74	20.17*	20.74	21.31	21.31	48.68*	49.10	49.37	48.94	49.10		
Esophagus	V50 (%)	2.56	4.55	2.41*	3.41	2.56	3.13	41.01	38.52	37.62*	40.00	40.48		
	V30 (%)	10.72*	11.18	11.44	12.74	10.72*	10.74	8.26	9.99	9.43	8.75	8.14*		
Heart	V40 (%)	7.70*	8.79	8.17	7.94	7.70*	7.80	2.37	2.22	2.92	1.96	1.94*		
	V5 (%)	12.74*	14.10	14.16	14.91	12.74*	14.37	23.24	23.27	22.83*	24.23	25.09		
Lung	V20 (%)	4.60*	4.81	4.89	5.03	4.60*	5.34	6.07*	6.19	7.34	6.30	7.02		
	Mean (Gy)	3.13*	3.31	3.25	3.29	3.13*	3.39	4.29*	4.31	4.68	4.47	4.67		

Abbreviations: D95, dose received by 95% (of ICTV); $\Delta D95$, the maximum difference of (ICTV) D95 between 3 dose scenarios (D_{CTave} , D_{CT0} and D_{CT50}). V40, volume of structure exposed to at least 40 Gy; Max is the maximum dose of a structure; Mean is the mean dose of a structure. † indicates which BAO algorithms achieved best objective value for a clinical case. * is the best achieved dose volume value.

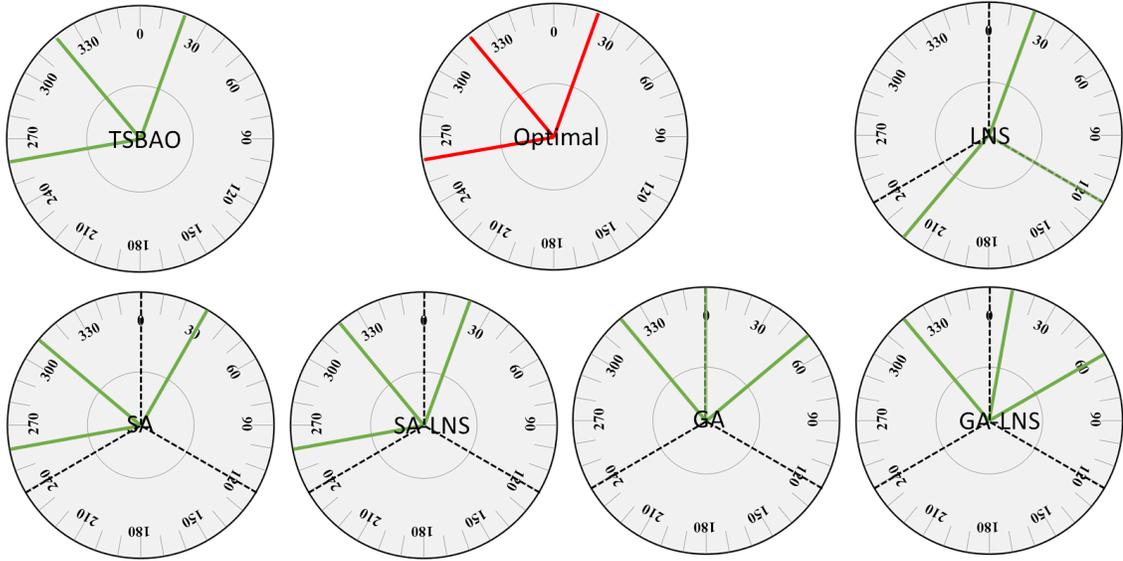


Figure 5.6: Beam angle configurations selected for the Lung case I by different algorithms. Equal spaced beam angles are denoted by dashed lines, the final solutions by green solid lines and the optimal solution by red solid lines.

incorporating uncertainty of internal organ motion to solve the BAO problem in IMPT treatment planning.

Uncertainty issue is a critical problem for IMPT treatment. In this work, to tackle the internal motion during treatment, a 4D robust fluence model is used to optimize the beamlet intensities. Furthermore, a beam angle score considering uncertainty information is introduced for beam angle clustering to help the algorithm to find the robust beam angle solutions. To our knowledge, it is the first work that incorporates uncertainty information as beam angle prior knowledge to guide beam angle optimization for IMPT planning. A bi-level deterministic local neighborhood search algorithm is used to improve the searching accuracy and reduce the total convergence time. Comparing to stochastic global algorithms such as simulated annealing and

genetic algorithm, we showed that the proposed clustering algorithm could greatly reduce the searching time. And compared to simple local neighborhood search algorithm, our algorithm can converge consistently to a good solution regardless of starting point or parameter settings.

Although the TSBAO algorithm demonstrated good convergence properties, it still does not guarantee global optimal. For example, in esophagus case II the objective value of global optimal is 4.73 while the solution found by clustering method is 4.80. Figure 5.7 shows the beam configurations of global optimal and solution found by TSBAO method. The global optimal is 20,170,290 and the TSBAO solution is 20,180,300. Although the solution is very close, two of the beams different. Using the one beam exchange local neighborhood search algorithm we cannot find the optimal solution. So, how to define the neighborhood to balance the time and result quality is still a problem for local neighborhood search algorithm for IMPT BAO.

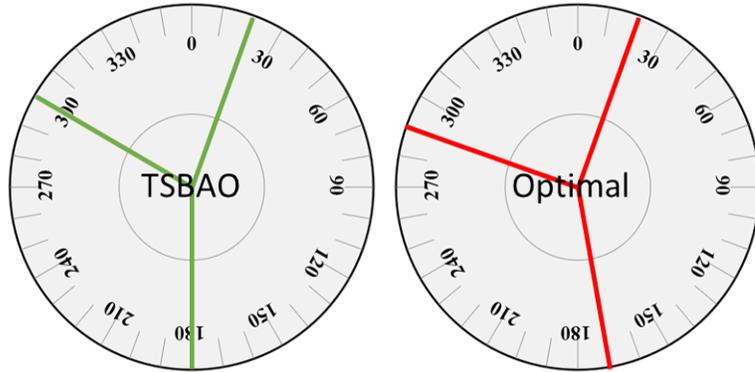


Figure 5.7: Beam angle configurations found by cluster method and global optimal for esophagus case II.

In this study, we focus on the target dose coverage and robustness, the penalty

weight of ICTV is much higher than that of OARs (i.e., 100:1). TSBAO achieved best objective values for tested cases and the plans can yield the best dose uniformity and robustness compare to other tested algorithms. However, these plans cannot guarantee the best dose sparing for all ORAs (Figure 5.7). Because OAR sparing is very sensitive to the beam angle incident direction, when there are multiple critical organs surrounding the target, it is hard to find a beam angle set can yield best dose sparing for all OARs at the same time, especially for IMPT treatment plan with much fewer beam angles than IMRT plan. So, how to balance the weight between different OARs is still depends on physicians or planners' preference.

Up to now, in all prior knowledge guided BAO algorithm, the feasibility scores are defined based on each standalone beam angle. This score definition cannot reflect the fitness of treatment plan with multiple beams directly. So, in this work, we choose using a cluster level LNS to find an optimal solution in centroid beam angles, so that we can have a good starting point in global scale. This step helps the algorithm avoid trapping in local minima. However, this step requires extra time for searching the solution. If the cluster number is large, the time for the cluster level LNS may increase rapidly. Therefore, how to find the correlation between beam angles in prior knowledge, so that it can reflect the merit of beam in a treatment plan is still under investigation.

5.5 Conclusion

This study focuses on the BAO problem in IMPT treatment planning considering internal organ motion and using prior knowledge guidance. We introduced a TSBAO method to beam angle optimization incorporating internal motion uncertainty for IMPT treatment planning. First, a p -median method was used for beam angle clustering. A beam angle score function incorporated uncertainty information was introduced to measure the merit of a beam angle for beam angle clustering. Then, a bi-LNS algorithm was designed to search the final solution for the treatment plan. We have demonstrated the algorithm can consistently find the solution of global optimal or close to it on four thoracic cancer cases. The efficiency of this algorithm was illustrated by comparing it with alternative BAO algorithms.

Chapter 6

Summary and Future Work

6.1 Current Findings

Radiation therapy is one of the most common treatments for many types of cancer. Different procedures of treatment planning can be formulated as different optimization problems. In this dissertation, we investigated two major optimization problems in intensity modulated proton therapy: fluence map optimization and beam angle optimization. Furthermore, we proposed a new robust optimization method to handle the misalignment error in multi-isocenter large field treatment planning.

Chapter 3 of the dissertation concentrated on solving the fluence map optimization problem. We found the conventional optimization algorithms require considerable long time or the performance is influenced by initial starting points. To overcome such shortcomings, we proposed a molecular dynamics method as a new alternative for solving the FMO problem in IMPT. We applied this method on three clinical cancer cases and compared the performance with three literature reported algorithms. We demonstrated that the MD method consistently performs better than other well-accepted methods, such as quasi-Newton, LBFGS and LBFGS-B, in both objective value and computational time.

Chapter 4 addressed the misalignment uncertainty issue in treating patients with large tumor size and require multi-isocenter large field treatment plan. Conventional, it requires complex treatment planning procedures for this type of patient. We proposed and validated a robust optimization approach to incorporate misalignment uncertainty into intensity modulated proton therapy treatment planning. The results demonstrated that the robust optimized IMPT treatment plan yields a low-gradient field dose in the junction regions to minimize the impact caused by misalignment errors. This method can greatly improve treatment planning efficiency and patient safety.

In Chapter 5, we focus on solving the beam angle selection problem in intensity modulated proton therapy. Beam angle optimization problem is NP hard. Conventional the global BAO methods require a long time to solve the problem and the performance is influenced by parameters used. The performance of local search methods highly depends on the starting point. Furthermore, the proton therapy is very sensitive to uncertainties, especially the large internal organ motion during the treatment. If considering the uncertainties into the BAO problem, it will increase a considerable computational time to solve. In order to find reliable solutions in a relatively short time, we developed a two-stage robust beam angle optimization method for the IMPT beam angle optimization. In this approach, we introduced a beam angle score function using the prior knowledge to incorporating the internal organ motion to measure the merit of beam angles. A p -median algorithm is developed for beam angle clustering to guide the beam angle searching. A bi-level local neighborhood

search algorithm is used to determine the final beam angle set for the treatment. Furthermore, Support vector machine (SVM) is used in bi-level LNS to reduce the search space. Our methods were tested on four thoracic cancer cases. It demonstrates that the two-stage method outperformed five widely used beam angle optimization methods, including a standalone LNS, simulated annealing (SA), genetic algorithms (GA), hybrid SA-LNS and hybrid GA-LNS in both objective value and CPU time.

6.2 Future Work

The molecular dynamics method is well accepted to solve the problem in material science. This method is a good candidate to be parallelized for running on the multi-core workstation. Based on the current progress, develop a parallel MD method for fluence map optimization can greatly improve the speed of the solving the problem. In addition, different parallel strategies can be investigated. Incorporating the uncertainty into the MD method can be another direction for the future work.

In our study, the two-stage beam angle optimization method has two hypothesis, one is the candidate beam angle set is finite and another one is the number of beam angles for treatment is given by planner. Now, the beam angle number is still decided by planner based on their experience. So, To estimate the optimum number of beams for IMPT treatment plan is an important topic for beam angle selection problem.

Up to now, the beam angle scoring is only based on the information of one individual beam angle, this score can only measure the merit of a single beam. However, a whole treatment plan consists of multiple beams, the beam angle score is hard to

correctly reflect whether this beam can collaborate with other beams to generate a high-quality plan. So, To consider the correlation between beam angles for beam angle scoring is the next step of our research.

Along with the patient case increasing, the proved high-quality treatment plan can be used as a database. Patient geometric information, beam angle setup and optimization parameters can be fully investigated. Various solution techniques, such as data mining or statistical inference, can be applied to identify the relationships between different patients. The stored beam angle sets and parameter sets can be suggested for the new patient case. At least, it can offer a good starting point for treatment plan optimization to improve the efficiency.

References

- [1] AmericanCancerSociety. Cancer facts & figures. 2016.
- [2] E. Halperin, L. Brady, D. Wazer, and C. Perez. *Perez & Brady's Principles and Practice of Radiation Oncology*. Lippincott Williams & Wilkins, 2013.
- [3] R.E. Drzymala, R. Mohan, L. Brewster, J. Chu, M. Goitein, W. Harms, and M. Urie. Dose-volume histograms. *International Journal of Radiation Oncology* Biology* Physics*, 21(1):71–78, 1991.
- [4] T. Kataria, K. Sharma, V. Subramani, K.P. Karrthick, and S.S. Bisht. Homogeneity index: An objective tool for assessment of conformal radiation treatments. *Journal of medical physics/Association of Medical Physicists of India*, 37(4):207, 2012.
- [5] J. Lyman and A. Wolbarst. Optimization of radiation therapy, iii: A method of assessing complication probabilities from dose-volume histograms. *International Journal of Radiation Oncology* Biology* Physics*, 13(1):103–109, 1987.
- [6] G. Kutcher and C. Burman. Calculation of complication probability factors for non-uniform normal tissue irradiation: The effective volume method gerald. *International Journal of Radiation Oncology* Biology* Physics*, 16(6):1623–1630, 1989.

- [7] J. Stroom, H. de Boer, H. Huizenga, and A. Visser. Inclusion of geometrical uncertainties in radiotherapy treatment planning by means of coverage probability. *International Journal of Radiation Oncology* Biology* Physics*, 43(4):905–919, 1999.
- [8] A. Bel, O. Petrascu, I. Van de Vondel, L. Coppens, N. Linthout, D. Verellen, and G. Storme. A computerized remote table control for fast on-line patient repositioning: implementation and clinical feasibility. *Medical physics*, 27(2):354–358, 2000.
- [9] H. de Boer, J. de Koste, C. Creutzberg, A. Visser, P. Levendag, and B. Heijmen. Electronic portal image assisted reduction of systematic set-up errors in head and neck irradiation. *Radiotherapy and Oncology*, 61(3):299–308, 2001.
- [10] L. Gilbeau, M. Octave-Prignot, T. Loncol, L. Renard, P. Scalliet, and V. Grégoire. Comparison of setup accuracy of three different thermoplastic masks for the treatment of brain and head and neck tumors. *Radiotherapy and Oncology*, 58(2):155–162, 2001.
- [11] E. van Lin, L. van der Vight, H. Huizenga, J. Kaanders, and A. Visser. Set-up improvement in head and neck radiotherapy using a 3d off-line epid-based correction protocol and a customised head and neck support. *Radiotherapy and oncology*, 68(2):137–148, 2003.
- [12] K. Ohara, T. Okumura, M. Akisada, T. Inada, T. Mori, H. Yokota, and

- M. Calaguas. Irradiation synchronized with respiration gate. *International Journal of Radiation Oncology* Biology* Physics*, 17(4):853–857, 1989.
- [13] T. Inada, H. Tsuji, Y. Hayakawa, A. Maruhashi, and H. Tsujii. Proton irradiation synchronized with respiratory cycle. *Nihon Igaku Hoshasen Gakkai zasshi. Nippon acta radiologica*, 52(8):1161–1167, 1992.
- [14] J. Wong, M. Sharpe, D. Jaffray, V. Kini, J. Robertson, J. Stromberg, and A. Martinez. The use of active breathing control (abc) to reduce margin for breathing motion. *International Journal of Radiation Oncology* Biology* Physics*, 44(4):911–919, 1999.
- [15] D. Kim, B. Murray, R. Halperin, and W. Roa. Held-breath self-gating technique for radiotherapy of non–small-cell lung cancer: A feasibility study. *International Journal of Radiation Oncology* Biology* Physics*, 49(1):43–49, 2001.
- [16] D. Maleike, J. Unkelbach, and U. Oelfke. Simulation and visualization of dose uncertainties due to interfractional organ motion. *Physics in medicine and biology*, 51(9):2237, 2006.
- [17] J.G. Li and L. Xing. Inverse planning incorporating organ motion. *Medical physics*, 27(7):1573–1578, 2000.
- [18] M. Chu, Y. Zinchenko, S. Henderson, and M. Sharpe. Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty. *Physics in Medicine and Biology*, 50(23):5463, 2005.

- [19] A. Olafsson and S. Wright. Efficient schemes for robust imrt treatment planning. *Physics in medicine and biology*, 51(21):5621, 2006.
- [20] J. Unkelbach, T. Chan, and T. Bortfeld. Accounting for range uncertainties in the optimization of intensity modulated proton therapy. *Physics in medicine and biology*, 52(10):2755, 2007.
- [21] T. Bortfeld, T.C.Y. Chan, A. Trofimov, and J.N. Tsitsiklis. Robust management of motion uncertainty in intensity-modulated radiation therapy. *Operations Research*, 56(6):1461–1473, 2008.
- [22] M. Langer and J. Leong. Optimization of beam weights under dose-volume restrictions. *International Journal of Radiation Oncology* Biology* Physics*, 13(8):1255–1260, 1987.
- [23] M. Langer, R. Brown, M. Urie, J. Leong, M. Stracher, and J. Shapiro. Large scale optimization of beam weights under dose-volume restrictions. *International Journal of Radiation Oncology* Biology* Physics*, 18(4):887–893, 1990.
- [24] A. Niemierko. Random search algorithm (ronsc) for optimization of radiation therapy with both physical and biological end points and constraints. *International Journal of Radiation Oncology* Biology* Physics*, 23(1):89–98, 1992.
- [25] S. Spirou and C. Chui. A gradient inverse planning algorithm with dose-volume constraints. *Medical Physics*, 25(3):321–333, 1998.
- [26] Q. Wu, R. Mohan, A. Niemierko, and R. Schmidt-Ullrich. Optimization of

- intensity-modulated radiotherapy plans based on the equivalent uniform dose. *International Journal of Radiation Oncology* Biology* Physics*, 52(1):224–235, 2002.
- [27] I. Rosen, R. Lane, S. Morrill, and J. Belli. Treatment plan optimization using linear programming. *Medical Physics*, 18(2):141–152, 1991.
- [28] X. Wu and Y. Zhu. A mixed-encoding genetic algorithm with beam constraint for conformal radiotherapy treatment planning. *Medical Physics*, 27(11):2508–2516, 2000.
- [29] W. Cao and G. Lim. Optimization models for cancer treatment planning. *Wiley Encyclopedia of Operations Research and Management Science*, 2011.
- [30] M. Langer, S. Morrill, R. Brown, O. Lee, and R. Lane. A comparison of mixed integer programming and fast simulated annealing for optimizing beam weights in radiation therapy. *Medical Physics*, 23(6):957–964, 1996.
- [31] E. Lee, T. Fox, and I. Crocker. Optimization of radiosurgery treatment planning via mixed integer programming. *Medical physics*, 27(5):995–1004, 2000.
- [32] S.M. Morrill, R.G. Lane, G. Jacobson, and I.I. Rosen. Treatment planning optimization using constrained simulated annealing. *Physics in medicine and biology*, 36(10):1341, 1991.
- [33] G. Ezzell. Genetic and geometric optimization of three-dimensional radiation therapy treatment planning. *Medical Physics*, 23(3):293–305, 1996.

- [34] T. Bortfeld, J. Bürkelbach, R. Boesecke, and W. Schlegel. Methods of image reconstruction from projections applied to conformation radiotherapy. *Physics in Medicine and Biology*, 35(10):1423, 1990.
- [35] Q. Wu and R. Mohan. Algorithms and functionality of an intensity modulated radiotherapy optimization system. *Medical physics*, 27(4):701–711, 2000.
- [36] X. Zhang, H. Liu, X. Wang, L. Dong, Q. Wu, and R. Mohan. Speed and convergence properties of gradient algorithms for optimization of imrt. *Medical physics*, 31(5):1141–1152, 2004.
- [37] D. Pflugfelder, J. Wilkens, S. Nill, and U. Oelfke. A comparison of three optimization algorithms for intensity modulated radiation therapy. *Zeitschrift für Medizinische Physik*, 18(2):111–119, 2008.
- [38] G. Lim and W. Cao. A two-phase method for selecting imrt treatment beam angles: Branch-and-prune and local neighborhood search. *European Journal of Operational Research*, 217(3):609–618, 2012.
- [39] L. Xing, J.G. Li, S. Donaldson, Q.T. Le, and A.L. Boyer. Optimization of importance factors in inverse planning. *Physics in Medicine and Biology*, 44(10):2525, 1999.
- [40] Varian Medical System. Proton algorithm reference guide, chapter 6: Eclipse proton optimizer for modulated scanning. 2007.

- [41] T. Chan, T. Bortfeld, and J. Tsitsiklis. A robust approach to imrt optimization. *Physics in medicine and biology*, 51(10):2567, 2006.
- [42] A. Fredriksson, A. Forsgren, and B. Hårdemark. Minimax optimization for handling range and setup uncertainties in proton therapy. *Medical physics*, 38(3):1672–1684, 2011.
- [43] W. Cao, G. Lim, A. Lee, Y. Li, W. Liu, X.R. Zhu, and X. Zhang. Uncertainty incorporated beam angle optimization for impt treatment planning. *Medical physics*, 39(8):5248–5256, 2012.
- [44] W. Liu, X. Zhang, Y. Li, and R. Mohan. Robust optimization of intensity modulated proton therapy. *Medical physics*, 39(2):1079–1091, 2012.
- [45] D. Sonderman and P. Abrahamson. Radiotherapy treatment design using mathematical programming models. *Operations Research*, 33(4):705–725, 1985.
- [46] J. Stein, R. Mohan, X. Wang, T. Bortfeld, Q. Wu, K. Preiser, C. Ling, and W. Schlegel. Number and orientations of beams in intensity-modulated radiation treatments. *Medical Physics*, 24(2):149–160, 1997.
- [47] C. Wang, J. Dai, and Y. Hu. Optimization of beam orientations and beam weights for conformal radiotherapy using mixed integer programming. *Physics in Medicine and Biology*, 48(24):4065, 2003.
- [48] R. Yang, J. Dai, Y. Yang, and Y. Hu. Beam orientation optimization

- for intensity-modulated radiation therapy using mixed integer programming. *Physics in medicine and biology*, 51(15):3653, 2006.
- [49] X. Wu, Y. Zhu, J. Dai, and Z. Wang. Selection and determination of beam weights based on genetic algorithms for conformal radiotherapy treatment planning. *Physics in medicine and biology*, 45(9):2547, 2000.
- [50] Y. Li, J. Yao, and D. Yao. Automatic beam angle selection in imrt planning using genetic algorithm. *Physics in Medicine and Biology*, 49(10):1915, 2004.
- [51] Y. Li, D. Yao, J. Yao, and W. Chen. A particle swarm optimization algorithm for beam angle selection in intensity-modulated radiotherapy planning. *Physics in medicine and biology*, 50(15):3491, 2005.
- [52] C. Rowbottom, V. Khoo, and S. Webb. Simultaneous optimization of beam orientations and beam weights in conformal radiotherapy. *Medical Physics*, 28(8):1696–1702, 2001.
- [53] D. Djajaputra, Q. Wu, Y. Wu, and R. Mohan. Algorithm and performance of a clinical imrt beam-angle optimization system. *Physics in Medicine and Biology*, 48(19):3191, 2003.
- [54] C.G. Rowbottom, S. Webb, and M. Oldham. Beam-orientation customization using an artificial neural network. *Physics in medicine and biology*, 44(9):2251, 1999.

- [55] W. D D'Souza, H. Zhang, D. Nazareth, L. Shi, and R. Meyer. A nested partitions framework for beam angle optimization in intensity-modulated radiation therapy. *Physics in medicine and biology*, 53(12):3293, 2008.
- [56] Dionne M Aleman, Arvind Kumar, Ravindra K Ahuja, H Edwin Romeijn, and James F Dempsey. Neighborhood search approaches to beam orientation optimization in intensity modulated radiation therapy treatment planning. *Journal of Global Optimization*, 42(4):587–607, 2008.
- [57] G. Lim, M. Ferris, S. Wright, D. Shepard, and M. Earl. An optimization framework for conformal radiation treatment planning. *INFORMS Journal on Computing*, 19(3):366–380, 2007.
- [58] G. Lim, L. Kardar, and W. Cao. A hybrid framework for optimizing beam angles in radiation therapy planning. *Annals of Operations Research*, 217(1):357–383, 2014.
- [59] A. Pugachev and L. Xing. Pseudo beam's-eye-view as applied to beam orientation selection in intensity-modulated radiation therapy. *International Journal of Radiation Oncology* Biology* Physics*, 51(5):1361–1370, 2001.
- [60] J. Meyer, S.M. Hummel, P.S. Cho, M.M. Austin-Seymour, and M.H. Phillips. Automatic selection of non-coplanar beam directions for three-dimensional conformal radiotherapy. *The British journal of radiology*, 2005.

- [61] G. Meedt, M. Alber, and F. Nüsslin. Non-coplanar beam direction optimization for intensity-modulated radiotherapy. *Physics in Medicine and Biology*, 48(18):2999, 2003.
- [62] E. Woudstra, B. Heijmen, and P. Storchi. Automated selection of beam orientations and segmented intensity-modulated radiotherapy (imrt) for treatment of oesophagus tumors. *Radiotherapy and oncology*, 77(3):254–261, 2005.
- [63] S. Das, T. Cullip, G. Tracton, S. Chang, L. Marks, M. Anscher, and J. Rosenman. Beam orientation selection for intensity-modulated radiation therapy based on target equivalent uniform dose maximization. *International Journal of Radiation Oncology* Biology* Physics*, 55(1):215–224, 2003.
- [64] E. Schreibmann and L. Xing. Dose–volume based ranking of incident beam direction and its utility in facilitating imrt beam placement. *International Journal of Radiation Oncology* Biology* Physics*, 63(2):584–593, 2005.
- [65] A. Pugachev and L. Xing. Computer-assisted selection of coplanar beam orientations in intensity-modulated radiation therapy. *Physics in Medicine and Biology*, 46(9):2467, 2001.
- [66] G. Lim, A. Holder, and J. Reese. A clustering approach for optimizing beam angles in imrt planning. 2009.

- [67] M. Bangert and U. Oelfke. Spherical cluster analysis for beam angle optimization in intensity-modulated radiation therapy treatment planning. *Physics in medicine and biology*, 55(19):6023, 2010.
- [68] F. Albertini, E.B. Hug, and A.J. Lomax. The influence of the optimization starting conditions on the robustness of intensity-modulated proton therapy plans. *Physics in medicine and biology*, 55(10):2863, 2010.
- [69] J. Llacer, N. Agazaryan, T. Solberg, and C. Promberger. Degeneracy, frequency response and filtering in imrt optimization. *Physics in medicine and biology*, 49(13):2853, 2004.
- [70] Q. Hou and Y. Wang. Molecular dynamics used in radiation therapy. *Physical review letters*, 87(16):168101, 2001.
- [71] Q. Hou, J. Wang, Y. Chen, and J. Galvin. An optimization algorithm for intensity modulated radiotherapy; the simulated dynamics with dose–volume constraints. *Medical physics*, 30(1):61–68, 2003.
- [72] E. Romeijn, J. Dempsey, and J. Li. A unifying framework for multi-criteria fluence map optimization models. *Physics in Medicine and Biology*, 49(10):1991, 2004.
- [73] G. Lim, J. Choi, and R. Mohan. Iterative solution methods for beam angle and fluence map optimization in intensity modulated radiation therapy planning. *OR Spectrum*, 30(2):289–309, 2008.

- [74] Y. Li, X. Zhang, and R. Mohan. An efficient dose calculation strategy for intensity modulated proton therapy. *Physics in medicine and biology*, 56(4):N71, 2011.
- [75] D. Liu and J. Nocedal. On the limited memory bfgs method for large scale optimization. *Mathematical programming*, 45(1-3):503–528, 1989.
- [76] R.H. Byrd, P. Lu, J. Nocedal, and C. Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208, 1995.
- [77] F. Ercolessi. A molecular dynamics primer. *Spring college in computational physics, ICTP, Trieste*, 19, 1997.
- [78] L. Verlet. Computer" experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Physical review*, 159(1):98, 1967.
- [79] W. Swope, H. Andersen, P. Berens, and K. Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics*, 76(1):637–649, 1982.
- [80] J. Chang, H. Li, R. Zhu, Z. Liao, L. Zhao, A. Liu, Y. Li, N. Sahoo, F. Poenisch, and D. Gomez. Clinical implementation of intensity modulated proton therapy for thoracic malignancies. *International Journal of Radiation Oncology* Biology* Physics*, 90(4):809–818, 2014.

- [81] S. Frank, J. Cox, M. Gillin, R. Mohan, A. Garden, D. Rosenthal, B. Gunn, R. Weber, M. Kies, and J. Lewin. Multifield optimization intensity modulated proton therapy for head and neck tumors: a translation to practice. *International Journal of Radiation Oncology* Biology* Physics*, 89(4):846–853, 2014.
- [82] H. Pan, S. Jiang, J. Sutton, Z. Liao, W. Chance, S. Frank, R. Zhu, H. Li, H. and Fontanilla, X. Zhang, and Gomez D. Early experience with intensity modulated proton therapy for lung-intact mesothelioma: a case series. *Practical radiation oncology*, 5(4):e345–e353, 2015.
- [83] J. Stoker, J. Grant, R. Zhu, R. Pidikiti, A. Mahajan, and D. Grosshans. Intensity modulated proton therapy for craniospinal irradiation: organ-at-risk exposure and a low-gradient junctioning technique. *International Journal of Radiation Oncology* Biology* Physics*, 90(3):637–644, 2014.
- [84] H. Lin, X. Ding, M. Kirk, H. Liu, H. Zhai, C. Hill-Kayser, R. Lustig, Z. Tochner, S. Both, and J. McDonough. Supine craniospinal irradiation using a proton pencil beam scanning technique without match line changes for field junctions. *International Journal of Radiation Oncology* Biology* Physics*, 90(1):71–78, 2014.
- [85] W.H. Clair, J.A. Adams, M. Bues, B.C. Fullerton, S. La Shell, H.M. Kooy, J.S. Loeffler, and N.J. Tarbell. Advantage of protons compared to conventional x-ray or imrt in the treatment of a pediatric patient with medulloblastoma.

- International Journal of Radiation Oncology* Biology* Physics*, 58(3):727–734, 2004.
- [86] L. Liao, S. Jiang, Y. Li, X. Wang, H. Li, X. Zhu, N. Sahoo, M. Gillin, A. Mahajan, and D. Grosshans. Th-c-brd-12: Robust intensity modulated proton therapy plan can eliminate junction shifts for craniospinal irradiation. *Medical Physics*, 41(6):553–553, 2014.
- [87] A. Giebeler, W. Newhauser, R. Amos, A. Mahajan, K. Homann, and R. Howell. Standardized treatment planning methodology for passively scattered proton craniospinal irradiation. *Radiation oncology*, 8(1):1, 2013.
- [88] J. Michalski, E. Klein, and R. Gerber. Method to plan, administer, and verify supine craniospinal irradiation. *Journal of applied clinical medical physics*, 3(4):310–316, 2002.
- [89] J. Unkelbach, T. Bortfeld, B. Martin, and M. Soukup. Reducing the sensitivity of impt treatment plans to setup errors and range uncertainties via probabilistic treatment planning. *Medical physics*, 36(1):149–163, 2009.
- [90] P. Myers, S. Stathakis, P. Mavroidis, C. Esquivel, and N. Papanikolaou. Evaluation of localization errors for craniospinal axis irradiation delivery using volume modulated arc therapy and proposal of a technique to minimize such errors. *Radiotherapy and Oncology*, 108(1):107–113, 2013.
- [91] Y. Li, P. Niemela, L. Liao, S. Jiang, H. Li, F. Poenisch, R. Zhu, S. Siljamaki,

- R. Vanderstraeten, N. Sahoo, M. Gillin, and X. Zhang. Selective robust optimization: A new intensity-modulated proton therapy optimization strategy. *Medical physics*, 42(8):4840–4847, 2015.
- [92] S. Register, X. Zhang, R. Mohan, and J. Chang. Proton stereotactic body radiation therapy for clinically challenging cases of centrally and superiorly located stage i non-small-cell lung cancer. *International Journal of Radiation Oncology* Biology* Physics*, 80(4):1015–1022, 2011.
- [93] X. Zhang, Y. Li, X. Pan, X Li, R. Mohan, R. Komaki, J. Cox, and J. Chang. Intensity-modulated proton therapy reduces the dose to normal tissue compared with intensity-modulated radiation therapy or passive scattering proton therapy and enables individualized radical radiotherapy for extensive stage iiib non-small-cell lung cancer: a virtual clinical study. *International Journal of Radiation Oncology* Biology* Physics*, 77(2):357–366, 2010.
- [94] A. Sultan and A. Saher. Optimization of beam orientations in intensity modulated radiation therapy planning. 2006.
- [95] H. Rocha, J. Dias, B. Ferreira, and M. Lopes. Beam angle optimization for intensity-modulated radiation therapy using a guided pattern search method. *Physics in medicine and biology*, 58(9):2939, 2013.
- [96] S. Lin, G. Lim, and J. Bard. Benders decomposition and an ip-based heuristic for selecting imrt treatment beam angles. *European Journal of Operational Research*, 251(3):715–726, 2016.

- [97] A. Lomax, E. Pedroni, H. Rutz, and G. Goitein. The clinical potential of intensity modulated proton therapy. *Zeitschrift für Medizinische Physik*, 14(3):147–152, 2004.
- [98] W. Cao, G. Lim, X. Li, Y. Li, R. Zhu, and X. Zhang. Incorporating deliverable monitor unit constraints into spot intensity optimization in intensity-modulated proton therapy treatment planning. *Physics in medicine and biology*, 58(15):5113, 2013.
- [99] X. Zhang, T. Guerrero, S. McGuire, B. Yaremko, R. Komaki, J. Cox, Z. Hui, Y. Li, W. Newhauser, R. Mohan, and Z. Liao. Four-dimensional computed tomography-based treatment planning for intensity-modulated radiation therapy and proton therapy for distal esophageal cancer. *International Journal of Radiation Oncology* Biology* Physics*, 72(1):278–287, 2008.
- [100] C. ReVelle and R. Swain. Central facilities location. *Geographical analysis*, 2(1):30–42, 1970.
- [101] G. Milligan and M. Cooper. An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, 50(2):159–179, 1985.
- [102] R. Tibshirani, G. Walther, and T. Hastie. Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):411–423, 2001.
- [103] G. Lim, S. Zangeneh, and S. Kim. Clustering approach for defining hurricane

evacuation zones. *Journal of Urban Planning and Development*, page 04016008, 2016.

[104] L. Wolsey. *Integer programming*, 1998.

[105] B.E. Boser, I.M. Guyon, and V.N. Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152. ACM, 1992.

[106] V. Vapnik. *Statistical learning theory*, volume 1. Wiley New York, 1998.

[107] M. Musavi, W. Ahmed, K. Chan, K. Faris, and D. Hummels. On the training of radial basis function classifiers. *Neural networks*, 5(4):595–603, 1992.

Appendix

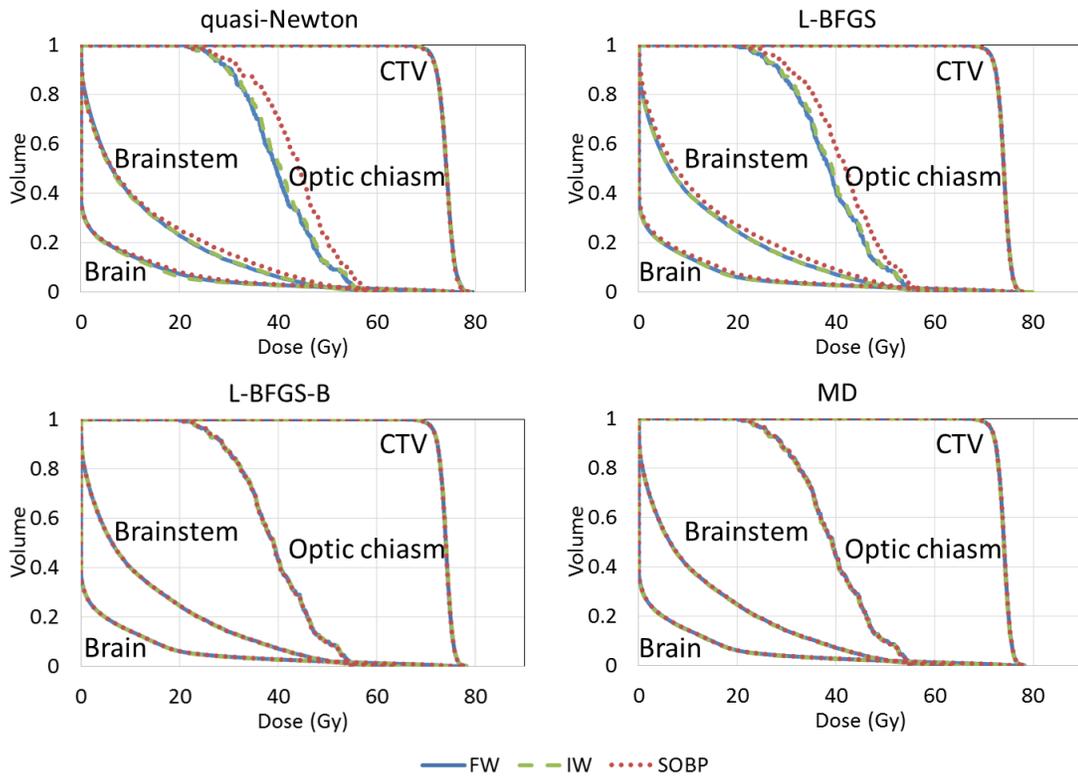


Figure 1: Dose-volume histogram comparison for the head-and-neck case.

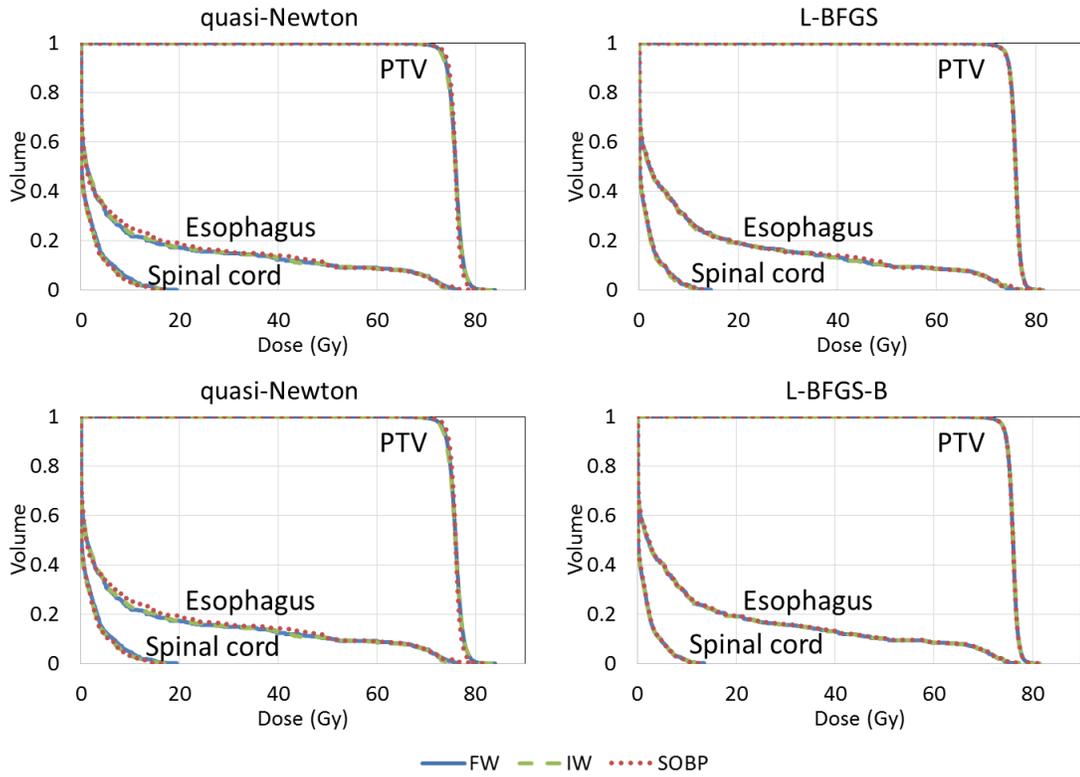


Figure 2: Dose-volume histogram comparison for the lung case.

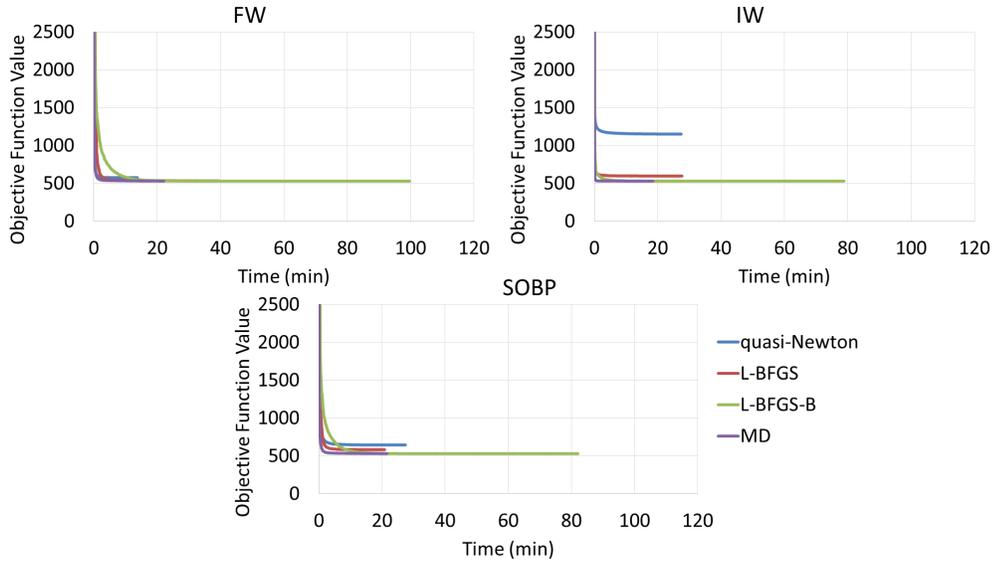


Figure 3: The OFVs as a function of time for the optimization processes for the head-and-neck case starting from FW, IW and SOBP using quasi-Newton, L-BFGS, L-BFGS-B and the MD method.

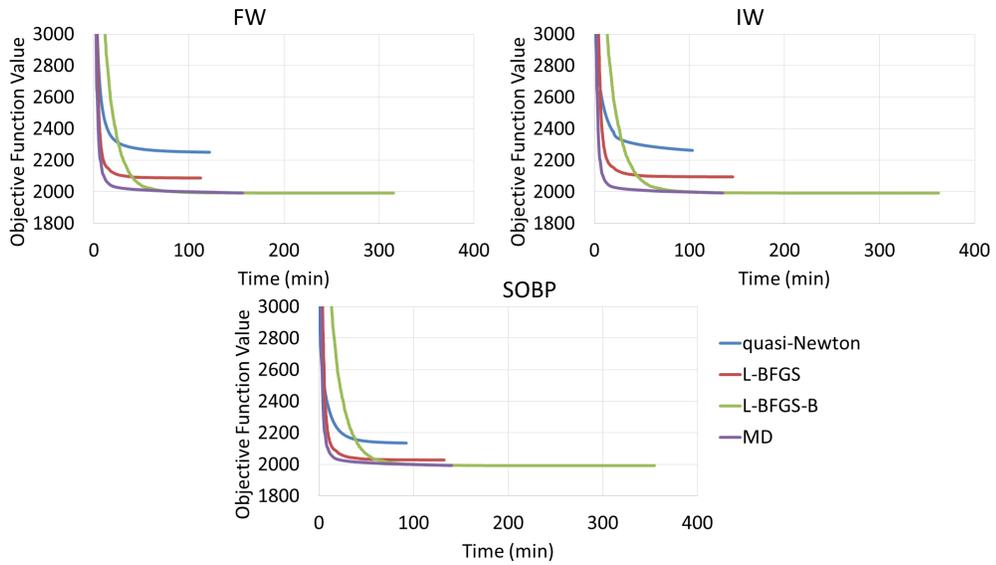


Figure 4: The OFVs as a function of time for the optimization processes for the lung case starting from FW, IW and SOBP using quasi-Newton, L-BFGS, L-BFGS-B and the MD method.