Radiation Dose Optimization For Critical Organs

by

Yasaman Khodadadegan

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2012 by the
Graduate Supervisory Committee:

Muhong Zhang, Co-Chair
William Pavlicek, Co-Chair
John Fowler
Tong Wu

ARIZONA STATE UNIVERSITY

May 2013

ABSTRACT

Ionizing radiation used in the patient diagnosis or therapy has negative effects on the patient body in short term and long term depending on the amount of exposure. More than 700,000 examinations are everyday performed on Interventional Radiology modalities [1], however; there is no patient-centric information available to the patient or the Quality Assurance for the amount of organ dose received. In this study, we are exploring the methodologies to systematically reduce the absorbed radiation dose in the Fluoroscopically Guided Interventional Radiology procedures.

In the first part of this study, we developed a mathematical model which determines a set of geometry settings for the equipment and a level for the energy during a patient exam. The goal is to minimize the amount of absorbed dose in the critical organs while maintaining image quality required for the diagnosis. The model is a large-scale mixed-integer program. We performed polyhedral analysis and derived several sets of strong inequalities to improve the computational speed and quality of the solution. Results present the amount of absorbed dose in the critical organ can be reduced up to 99% for a specific set of angles.

In the second part, we apply an approximate gradient method to simultaneously optimize angle and table location while minimizing dose in the critical organs with respect to the image quality. In each iteration, we solve a sub-problem as a MIP to determine the radiation field size and corresponding X-ray tube energy. In the computational experiments, results show further reduction (up to 80%) of the absorbed dose in compare with previous method.

Last, there are uncertainties in the medical procedures resulting imprecision of the absorbed dose. We propose a robust formulation to hedge from the worst case absorbed

dose while ensuring feasibility. In this part, we investigate a robust approach for the organ motions within a radialogy procedure. We minimize the absorbed dose for the critical organs across all input data scenarios which are corresponding to the positioning and size of the organs. The computational results indicate up to 26% increase in the absorbed dose calculated for the robust approach which ensures the feasibility across scenarios.

DEDICATION

I would like to dedicate this Doctoral dissertation to my husband for his continues help and support, to my dear father and mother for their support, patience, and continues encouragement, to my sister and my brother for their assistance and support. I also would like to dedicate it to my grand parents who always encouraged me and supported me during my education, my friends here and overseas who have always been supportive. I believe it was not possible to achieve my goal without their support and encouragements. I am so pleased that I have such thoughtful family and friends. I was inspired to

ACKNOWLEDGEMENTS

I acknowledge my advisor, Dr Muhong Zhang who helped me throughout my research and tought me how to do a high quality research. I acknowledge Dr Ronald Askin for his support to enter this PhD program and throughout the program. I acknowledge my co-advisor, Dr William Pavlicek who provided me with opportunities to work in Mayo Clinic, practice and learn radiology procedures while he encouraged me during my research, taught me a lot about the radiology procedures and spent a lot of time on revising my papers in Radiology. I acknowledge Dr Teresa Wu for her support and good advices during my PhD studies. I also acknowledge all of them for the time they spent for my publications. I acknowledge Robert G. Paden from Mayo Clinic who helped me to learn radiology procedures and answered my questions patiently. I acknowledge Mary B. Peter from Mayo Clinic for her support during our work in Mayo Clinic. I acknowledge Dr John Fowler for his time to be in my PhD committee.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 X-ray Modalities in Radiology

Radiation operative modalities are classified into two categories of the imaging and therapy modalities. Interventional Radiology (IR) is recognized as an image-guided specialty for diagnosis ([5]) and radiation therapy (radiation oncology) is usually used for the cancer treatment. Computed Tomography (CT) and Fluoroscopically Guided Interventional (FGI) modalities are instances of imaging modalities. Intensity Modulated Radiation Therapy (IMRT) is one of the most common forms of the radiation therapy. These instances use ionizing radiation which is composed of the particles, and the energy of individual particles determines the ionization degree. Ionizing radiation has dangerous effects for the patient and the interventional radiologist depending upon the amount of radiation. In this study, we focus on the radiation dose resulted from X-ray exposures and consider only modalities emitting X-ray, however, this study can be extended to gamma rays as well.

FGI and CT are fast in capturing the images of the inside of the human body using X-rays. They are composed of an X-ray tube producing the X-ray, an image receptor, and a table for the patient. The X-ray passes through the patient while attenuating by varying amounts as they interact with the different internal structures of the body, and a shadow of the structures on the image screen constructs the digital image. In the current digital radiology, a matrix represents an image which is a square or rectangular area divided into rows and columns. The smallest element of this matrix is called “pixel”. Each pixel of the matrix represents the individual grey level of an image, and the value depends upon the amount of absorbed dose in the image receptor [6].

IMRT utilizes computer-controlled linear accelerators to precisely deliver radiation doses to the tumor. By controlling and modulating the intensity of the radiation beam, it allows conforming to the three-dimensional (3D) shape of the tumor. It allows exposing higher radiation doses to the tumor while minimizing the dose to surrounding normal and critical structures. It is used extensively for a variety of cancer disease treatments [7].

1.2 Radiation Effects and Risks

Ionizing radiation interacts with atoms of the cells. From the lowest level it may affect the atoms, molecules, and cells to the higher levels of tissues and organs[8]. Radiation effects of X-ray exposure to the body are in two major types of deterministic and stochastic effects which may lead to skin injuries or cancer induction respectively.

Here we define some terminologies which are widely used in the organ dose studies [9, 10],

- Absorbed dose: radiation energy imparted per unit mass of an irradiated body. It is measured in joule per kilogram, and this unit is also called Gray (Gy).

$$D = \frac{dE(J)}{dm(kg)} \tag{1.1}$$

  where dE denotes the energy absorbed in the mass and dm is the mass of the organ.

- Equivalent dose: multiplying the absorbed dose by appropriate weighting factors depending on the type of radiation creates the equivalent dose in the relevant organ or tissue.

$$H_T = W_R \, D_{T,R} \tag{1.2}$$

  where W_R is the radiation weighting factor for the radiation type R and $D_{T,R}$ denotes the absorbed dose received by radiation type R in the organ or tissue T. The unit for H_T is Sivert (Sv).

- Effective dose: sum of the equivalent doses in the principal tissues and organs in the body, each weighted by a tissue weighting factor. This weighting factor takes account of the probability of fatal cancer, the probability of nonfatal cancer, weighted for severity, and the average length of life lost due to an induced cancer.

$$E' = \sum W_T \, H_T \tag{1.3}$$

where W_T is tissue weighting factor based on the sensitivity to stochastic radiation damage (risk factor). Table 1.1 represents Tissue Weighting Factors W_T based on ICRP [11]. E' is measured in Sv units. [a] The value 0.30 is applied to the average dose to the five remaining tissues receiving the highest dose, excluding the skin, lens of the eye, and the extremities. [b] Remainder is composed of the following tissues: adrenals, brain, extrathoracic airways, small intestine, kidneys, muscle, pancreas, spleen, thymus, and uterus. [c] The value 0.05 is applied to the mass-weighted average dose to the Remainder tissue group, except when the following splitting rule applies: If a tissue of Remainder receives a dose in excess of that received by any of the 12 tissues for which weighting factors are specified, a weighting factor of 0.025 (half of Remainder) is applied to that tissue and 0.025 to the mass-averaged committed equivalent dose in the rest of the Remainder tissues.

- Effective risk: Weighted organ doses with cancer risks estimates would perform the same comparative role as effective dose.

$$R = \sum_T r_T \, H_T \tag{1.4}$$

where H_T is the equivalent dose to the organ/tissue T and r_T is the sex-, age-, and tissue-specific risk coefficient (cases per 100000 exposed to 0.1 Gy) for lifetime attributable risk of cancer incidence. Table 1.2 presents the r_T coefficients for children and adults. Another source for indicating the risk factors is [4].

	Tissue weighting factor W_T	
Organ or Tissue	ICRP Pub. 26	ICRP Pub. 60
Gonads	0.25	0.20
Bone marrow	0.12	0.12
Colon	–	0.12
Lung	0.12	0.12
Stomach	–	0.12
Bladder	–	0.05
Breast	0.15	0.05
Liver	–	0.05
Esophagus	–	0.05
Thyroid	0.03	0.05
Skin	–	0.01
Bone surface	0.03	0.01
Remainder	0.30^{a}	$0.05^{b,c}$

Table 1.1: Tissue weighting factors W_T given in ICRP Publication 26 (1977) [2] and ICRP Publication 60 (1991) [3].

In this study, our focus is on the minimization of the absorbed dose in the critical organs of the body which may result in cancer induction. Since we only focus on a single organ, the cancer risk is proportional to the absorbed dose.

1.2.1 Deterministic Risks

Deterministic risk is pertinent to the dose at the irradiated area of skin. Skin effects include erythema (reddening like sunburn), dry desquamation (peeling), and moist desquamation (blistering). Skin effects usually occur in the relatively low exposures of X-ray. Erythema may occur when skin dose exceeds 3 Gy (unit of dose, J/kg), and blistering requires a skin dose exceeding 12 Gy. Hair loss (epilation) is another skin effect which can occur after receiving a skin dose higher than 5 Gy [8].

Several episodes of severe skin damage and harm to patients have occurred due to high X-ray exposures during FGI procedures [12, 13]. Since 1994 the FDA and more recently (2006) the Joint Commission has specified that skin exposures from fluoroscopy

	r_T		
	Children	Adults	All ages
Lung	373	166	208
Stomach	66	30	37
Colon	203	96	118
Liver	32	14	18
Bladder	153	75	91
Uterus	37	14	19
Ovary	76	28	37
Prostate	67	34	41
Breast	865	160	299
Thyroid	200	18	54
Leukemia	133	68	82

Table 1.2: Tissue weighting factors W_T given in ICRP Publication 26 (1977) [2] and ICRP Publication 60 (1991) [3].

be routinely monitored as a Quality Assurance (QA) metric. High skin doses that cause erythema must be identified and the patient informed. Appropriate medical care must be provided should this occur. An X-ray exposure to a localized region of the skin (i.e. peak skin dose) that exceeds 15 Gy is now included as a sentinel event and it is mandated to be reported to the Joint Commission [14].

Maximum entrance skin dose (ESD) for CT is estimated by a theoretical method within 0.1-0.42 Gy for one procedure [15]. However, skin dose may reach 1.2 Gy in some CT procedures [16]. IMRT avoids the high skin doses by modulating the intensity through various angles.

1.2.2 Stochastic Risks

Stochastic risk belongs to the mean organ dose which correlates with a risk of cancer induction by affecting on the cell growing and dividing mechanisms [17]. Total organ dose is indicated by total absorbed dose which is total energy absorbed in the organ/tissue per unit mass of the organ/tissue [9]. When X-ray beam passes through the patient body,

depending on the attenuation coefficient of the tissue, energy is attenuated and absorbed at each volume unit along the path of the beam.

Not all the living cells have the same sensitivity to radiation. They can be classified based on their rate of reproduction indicating their relative sensitivity to radiation. Cells producing blood are continuously regenerating, so are the most sensitive type, bone marrow cells. Reproductive and gastrointestinal cells are not regenerating as quickly as the blood forming cells, so they are less sensitive. The least sensitive cells are nerve and muscle cells which are the slowest to regenerate [8]. Therefore, radiation biological effects on living cells may have three consequences: (1) no residual damage when injured or damaged cells repair themselves; (2) cells being replaced through normal biological processes after they die; or (3) cells overproduce to repair themselves leading to a biophysical change [18].

High radiation doses in a short period of the time may kill cells following in tissue and organ damage, while low doses may damage or alter the DNA of the irradiated cells. Effects of the high radiation appears early after exposure and may result in death; however, effects of the low doses may not appear for several years. A study on women patients undergone fluoroscopic procedures determine the rate of breast cancer after 10 years [19].

People do not show the same reaction for the same amount of radiation dose, so it is not possible to indicate the exact exposure that may result in death. However, statistics show that 50% of a population may die within 30 days by receiving a dose in the range of 3.5-5 Gy to the whole body. It also depends on each individual health. The whole body receives the exposure in a short period of time (minutes to hours). Exposures more than 0.5 Gy are recognized as decision doses and over 1.5 Gy cause severe blood changes and affect many cells resulting in tissue or organ damage [8].

Low doses do not lead to the immediate problems in the organs. The effect may occur in the cell level and changes may not be recognized for many years after the exposure. However, currently there is no data to indicate the occurrence of cancer following exposure to low doses less than 0.10 Gy in the long time [18].

The correlations between radiation exposure and cancer induction are mostly according to populations exposed to relatively high levels of ionizing radiation (e.g., patients who undergone selected diagnostic or therapeutic medical procedures). A study determines the risk factors based on Japanese atomic bomb survivors [4]. Figure 1.1 presents the lifetime attributable risk of cancer incidence in terms of number of cases per 100,000 persons exposed to a single dose of 0.1 Gy for different ages and genders. The lifetime attributable risk is defined in [4] as: "The estimated rate of a disease (such as lung cancer) that could, in theory, be prevented if all exposures to a particular causative agent (such as radon) were eliminated."

Figure 1.2 presents the lifetime attributable risk of cancer mortality in terms of number of cases per 100,000 persons exposed to a single dose of 0.1 Gy for different ages and genders. Cancers induced from high dose exposure (greater than 0.5 Gy) include leukemia, breast, bladder, colon, liver, lung, esophagus, ovarian, multiple myeloma, stomach, and possibly prostate, nasal cavity/sinuses, pharyngeal and laryngeal, and pancreatic cancer [18]. Figure 1.3 shows the estimated lifetime attributable risk for the patient exposure to 1 mGy per year throughout his/her life and to 10 mGy per year from ages 18 to 65. Figures 1.1, 1.2, and 1.3 show the highest lifetime attributable risk factors belong to the female lungs and breasts in the young ages.

The radiation protection community conservatively considers some risk associated with the amount of radiation causing cancer and hereditary effect. A linear, no-threshold (LNT) model illustrates the dose and cancer risk relationship 1.4. This figure depicts a

Cancer Site	Age at Exposure (years)										
	0	5	10	15	20	30	40	50	60	70	80
Males											
Stomach	76	65	55	46	40	28	27	25	20	14	7
Colon	336	285	241	204	173	125	122	113	94	65	30
Liver	61	50	43	36	30	22	21	19	14	8	3
Lung	314	261	216	180	149	105	104	101	89	65	34
Prostate	93	80	67	57	48	35	35	33	26	14	5
Bladder	209	177	150	127	108	79	79	76	66	47	23
Other	1123	672	503	394	312	198	172	140	98	57	23
Thyroid	115	76	50	33	21	9	3	1	0.3	0.1	0.0
All solid	2326	1667	1325	1076	881	602	564	507	407	270	126
Leukemia	237	149	120	105	96	84	84	84	82	73	48
All cancers	2563	1816	1445	1182	977	686	648	591	489	343	174
Females											
Stomach	101	85	72	61	52	36	35	32	27	19	11
Colon	220	187	158	134	114	82	79	73	62	45	23
Liver	28	23	20	16	14	10	10	9	7	5	2
Lung	733	608	504	417	346	242	240	230	201	147	77
Breast	1171	914	712	553	429	253	141	70	31	12	4
Uterus	50	42	36	30	26	18	16	13	9	5	2
Ovary	104	87	73	60	50	34	31	25	18	11	5
Bladder	212	180	152	129	109	79	78	74	64	47	24
Other	1339	719	523	409	323	207	181	148	109	68	30
Thyroid	634	419	275	178	113	41	14	4	1	0.3	0.0
All solid	4592	3265	2525	1988	1575	1002	824	678	529	358	177
Leukemia	185	112	86	76	71	63	62	62	57	51	37
All cancers	4777	3377	2611	2064	1646	1065	886	740	586	409	214

NOTE: Number of cases per 100,000 persons exposed to a single dose of 0.1 Gy.

[a]These estimates are obtained as combined estimates based on relative and absolute risk transport and have been adjusted by a DDREF of 1.5, except for leukemia, which is based on a linear-quadratic model.

Figure 1.1: Lifetime Attributable Risk of Cancer Incidence [4], (Table 12D-1)

risk value even when there is no dose and that is associated with the risk of cancer even without receiving any radiation. Although this LNT model may overestimate the radiation dose, it is accepted by the Nuclear Regulatory Commission (NRC) as a conservative model for determining radiation dose standards [8].

Cancer Site	Age at Exposure (years)										
	0	5	10	15	20	30	40	50	60	70	80
Males											
Stomach	41	34	30	25	21	16	15	13	11	8	4
Colon	163	139	117	99	84	61	60	57	49	36	21
Liver	44	37	31	27	23	16	16	14	12	8	4
Lung	318	264	219	182	151	107	107	104	93	71	42
Prostate	17	15	12	10	9	7	6	7	7	7	5
Bladder	45	38	32	27	23	17	17	17	17	15	10
Other	400	255	200	162	134	94	88	77	58	36	17
All solid	1028	781	641	533	444	317	310	289	246	181	102
Leukemia	71	71	71	70	67	64	67	71	73	69	51
All cancers	1099	852	712	603	511	381	377	360	319	250	153
Females											
Stomach	57	48	41	34	29	21	20	19	16	13	8
Colon	102	86	73	62	53	38	37	35	31	25	15
Liver	24	20	17	14	12	9	8	8	7	5	3
Lung	643	534	442	367	305	213	212	204	183	140	81
Breast	274	214	167	130	101	61	35	19	9	5	2
Uterus	11	10	8	7	6	4	4	3	3	2	1
Ovary	55	47	39	34	28	20	20	18	15	10	5
Bladder	59	51	43	36	31	23	23	22	22	19	13
Other	491	287	220	179	147	103	97	86	69	47	24
All solid	1717	1295	1051	862	711	491	455	415	354	265	152
Leukemia	53	52	53	52	51	51	52	54	55	52	38
All cancers	1770	1347	1104	914	762	542	507	469	409	317	190

NOTE: Number of deaths per 100,000 persons exposed to a single dose of 0.1 Gy.

[a]These estimates are obtained as combined estimates based on relative and absolute risk transport and have been adjusted by a DDREF of 1.5, except for leukemia, which is based on a linear-quadratic model.

Figure 1.2: Lifetime Attributable Risk of Cancer Mortality [4], (Table 12D-2)

"The likelihood of cancer occurring after radiation exposure is about five times greater than a genetic effect (e.g., increased still births, congenital abnormalities, infant mortality, childhood mortality, and decreased birth weight)." A mutation created in the reproducible cells of the patient after exposure may result in genetic effects. These effects may appear in the offspring of the exposed person or several generations later.

	Incidence: Exposure Scenario		Mortality: Exposure Scenario	
Cancer site	1 mGy per Year throughout Life	10 mGy per Year from Ages 18 to 65	1 mGy per Year throughout Life	10 mGy per Year from Ages 18 to 65
Males				
Stomach	24	123	13	66
Colon	107	551	53	273
Liver	18	93	14	72
Lung	96	581	99	492
Prostate	32	164	6.3	32
Bladder	69	358	16	80
Other	194	801	85	395
Thyroid	14	28		
All solid	554	2699	285	1410
Leukemia	67	360	47	290
All cancers	621	3059	332	1700
Females				
Stomach	32	163	19	94
Colon	72	368	34	174
Liver	8.7	44	8	40
Lung	229	1131	204	1002
Breast	223	795	53	193
Uterus	14	19	3.5	18
Ovary	29	140	18	91
Bladder	71	364	21	108
Other	213	861	98	449
Thyroid	75	139		
All solid	968	4025	459	2169
Leukemia	51	270	38	220
All cancers	1019	4295	497	2389

NOTE: Number of cases or deaths per 100,000 persons exposed to 1 mGy per year throughout life or to 10 mGy per year from ages 18 to 64.

[a]These estimates are obtained as combined estimates based on relative and absolute risk transport and have been adjusted by a DDREF of 1.5, except for leukemia, which is based on a linear-quadratic model.

Figure 1.3: Lifetime Attributable Risk of Solid Cancer Incidence and Mortality [4], (Table 12D-3)

In this study, we pick female breasts which is a vulnerable organ for young females when exposed to the X-ray radiation.

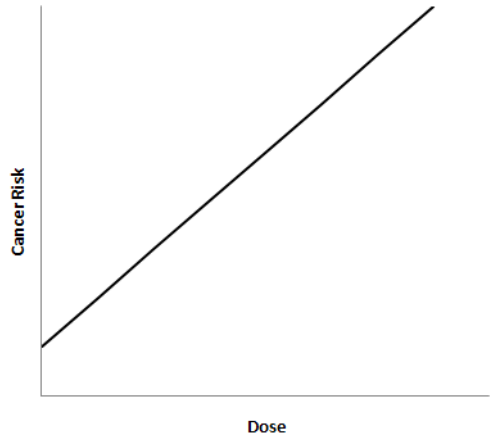


Figure 1.4: Linear No-Threshold Risk Model

1.3 Interventional Radiology

Interventional Radiology is an invasive diagnostic subspecialty that comprises a wide range of image-guided therapeutic procedures[5]. Fluoroscopic-Guided Interventional Radiology (FGI) modalities use ionizing radiation which produces X-rays. Number of radiology procedures is increasing due to having less cost and pain in compare with surgeries [20, 21]. Certain procedures can be potentially harmful because of high X-ray penetration and require to be accurately monitored to avoid exposing critical organs of the body. Among them, cardiac imaging frequently exposes patients to ionizing radiation, but its contribution to the breast cancer is unknown. A study [19] on the 31,710 records of female patients having fluoroscopic procedures after 10 years shows the mortality rates resulted from breast cancer while there is an accumulated dose of 0.1 Gy absorbed in the breasts.

International Commission on Radiological Protection (ICRP) report of “Avoidance of Radiation Injuries from Medical Procedures” addresses lack of radiologists training regarding possible radiation induced injuries resulted in FGI procedures [22]. A study on

women who received chest X-ray shows that young women who receive repeated X-rays with breast tissue included in the beam interception region with body have a higher potential risk of breast cancer (about 9 times more for the ages less than 40 in compare with ages more than 40) [23]. Nonetheless, patients are not aware of the radiation risks nor are followed up when radiation doses from these sophisticated procedures may cause any harm [24]. ICRP expect that all patients be informed of the radiation risks and effects as part of the process of informed consent and that an appropriate process for patient follow-up should be established [24]. Thus, monitoring organ doses is required for the quality assurance purposes.

In this regard, each procedure should be setup in terms of geometry to avoid exposing X-ray to the critical tissue as low as reasonably achievable (ALARA) while providing adequate image quality for the diagnosis. Geometry of each procedure includes patient position, table position, gantry angles, and collimator size. Energy of the x-ray source is determined by the potential, current and time of exposure. Equipment parameters and patient data should be investigated in order to identify the best setup for each procedure. Currently in FGI procedures, there is some instruction for each protocol which determines the setting of the equipment. However, this setting does not take into account the patient size and geometry. There is not sufficient data for the geometry setting to avoid a specific organ of the patient. An automatic control system adjusts the amount of the exposure in the further radiations. This system measures the number of the photons hitting the image receptor and energy of each individual photon. Based on this information, it adjusts tube potential and current by itself. Real time optimization of the dose for the critical organs may not be practical; however, a set of suggestions for the geometry and energy setup can result in smaller amount of absorbed dose in a critical organ. Although the current equipment has the automatic exposure control during the

procedure, changing the geometry setting would reduce the absorbed radiation in the body.

This study has been encouraged by the report of cardiovascular examinations in Mayo Clinic in Arizona where a 19 years old girl had more than 8 examinations. Such examinations for young females might increase the risk of cancer induction for the critical organs which are breasts for young females in such irradiation procedures.

One factor to quantify the image quality is Signal-to-Noise (SNR) ratio [25]. It can be defined in terms of the ratio of the average intensity of the Region of Interest(ROI) to the standard deviation of the background of the image [26]. Therefore, there is a tradeoff between energy required for the reasonable image quality and absorbed dose in the organ. By specifying a range for reasonable SNR values a constraint for image quality is added to the model. Another factor for image quality is defined as contrast. Contrast means how much distinguishable objects are from each other and in the other word, what is the smallest size of the objects distinguishable in the image. Methods to improve the image contrast are discussed in chapter 3.

We discretized the phantom volume into small volumes called voxels. We use Computed Tomography (CT) images which are usually available prior to FGI procedures to calculate linear attenuation coefficients based on the CT numbers captured from the image.

The purpose of this paper is to present an optimization model to minimize total absorbed dose in patient by determining required geometry setup of the FGI equipment subject to image quality. A mixed integer programming (MIP) model is presented to decide on the amount of energy to be generated at the X-ray tube, table position and coverage of collimator jaws.

1.3.1 Solution Approach

Mathematical programming is a method to systematically model the problems and determine the best decisions based on the problem objective and constraints. A well studied group of mathematical models are called Linear programming (LP) where there exist known polynomial algorithms to solve these problems. Variables are continues and objective function and constraints are a linear function of variables. A general linear program formulation is stated as follows.

$$\begin{aligned} Max \quad & cx \\ \text{subject to} \quad & Ax \leq b \\ & x \geq 0 \end{aligned}$$

Another class of mathematical programming models is integer programming (IP) models where all the variables are integer. No polynomial time algorithm is known to solve these problems. Mixed-integer linear and mixed-integer nonlinear models are another group where some of the variables are integer and some are continuous. There is no polynomial algorithm known to solve the mixed-integer problems in polynomial time and they are usually harder than pure integer programs. A class of problems which there is no known algorithm better than pure enumeration is called NP-Hard (non-deterministic polynomial-time hard) problems. Mixed-integer nonlinear programs have nonlinear objective function and/or constraints. It is more complicated to analyze the nonlinear region for the mixed-integer problems. Thus, we have defined some variables to linearize the nonlinear functions. We formulate radiation dose optimization model in terms of 0-1 mixed integer program. The set of variables include continuous variables and binary variables. We explain the theoretical approach for solving this problem. Due to the large size of the problem in terms of number of binary variables and constraints, we decompose

the region of the problem into smaller substructures called polyhedrons and analyze each polyhedron. A definition for the polyhedron is stated in [27], "A subset of $\mathbb{R}^n$ described by a finite set of linear constraints $P = x \in \mathbb{R}^n : Ax \leq b$ is a polyhedron." Figure 1.5 represents the feasible region corresponding to the polyhedron $P = \{(x_1, x_2) \in \mathbb{R}^2 : -18x_1 + 25x_2 \leq 55, 40x_1 + 17x_2 \leq 168, x_1, x_2 \geq 0\}$ where O, A, B, and C are corresponding vertices of this polyhedron. If variables x_1 and x_2 only take integer values, then we define a set of $S = P \cap \mathbb{Z}^2$ which is depicted by the solid black points in this figure. The convex hull of S denoted by $conv(S)$ is defined as:

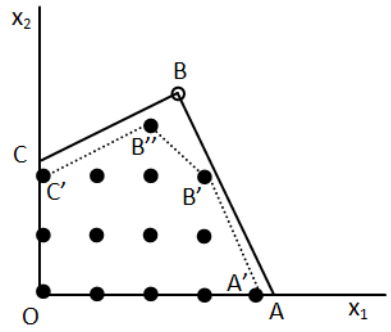


Figure 1.5: Demonstration of a convex hull of a pure IP and corresponding LP relaxation problem; Polyhedron restricted to vertices O, A, B, and C is called LP relaxation and O, A', B', B", C' is corresponding to the convex hull of IP

$conv(S) = \{x \in \mathbb{R}^2 : x = \sum_{i=1}^{2} \lambda_i x_i,\ \sum_{i=1}^{2} = 1,\ \lambda_i \geq 0 \text{ for } i = 1,2\}$ [27] where $x = (x_1, x_2)$ and vertices O, A', B', B", and C' determine the corners (extreme points) of the feasible region. $conv(S)$ is a polyhedron as well. The boundary of this convex hull is shown by dotted lines and axes x_1 and x_2. All the vertices and edges of this convex hull are called a face for this convex hull. Any of these edges is called facet defining inequality for the $conv(S)$ since they are not redundant. Facet has one dimension less than the original polyhedron dimension.

A branch and cut algorithm is applied to solve the problem. Branch and bound algorithm is a common deterministic method for solving mixed-integer models. First, it solves the problem assuming all the variables are continues (LP relaxation) and then if that solution satisfies the integrality for the integer variables, it is optimal solution and the branch and bound algorithm will stop. Otherwise, it starts branching on the integer variables having fractional value in the solution and each branch fixes corresponding integer value to an integer number and solves the problem. This continues until all the integer variables get integer value or problem is infeasible. In the branch and cut algorithm, at each step of branching a set of cuts are added to the problem to tighten the space of relaxed problem.

A cut applied to the IP or MIP for cutting off a fractional solution from the LP relaxation is depicted in figure 1.6. Two cuts are depicted by dashed lines. Cut1 separates

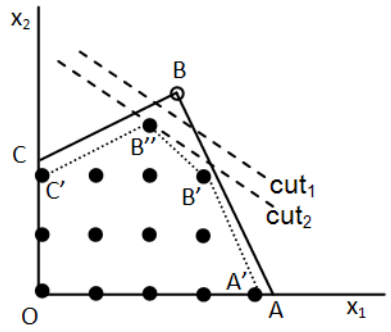


Figure 1.6: Cut1 and cut2 separates the point of B which has a fractional solution from the LP relaxation. Cut2 is a facet defining inequality.

the point B of the polyhedron restricted to O, A, B, and C which is LP relaxation, but this cut is not intersecting with convex hull of IP. Thus, it is neither called a facet defining inequality nor a strong valid inequality. However, cut2 is intersecting with B" which is a face of the convex hull, so this cut is a strong valid inequality. A cut passing B' and B" is called facet defining inequality. Mathematical methods are applied to find the cuts for the

problem. We derived strong valid inequalities, and added them in the pool of cuts in the branch and cut algorithm. These cuts are developed for the MIP problem of radiation dose reduction based on the structure of the problem and existing mathematical methods for deriving cuts.

An example of a MIP problem is given in figure 1.7 where X is an integer and Y is a continuous variables. Dash lines indicate the feasible region of the mixed-integer program. The solid lines show the feasible region of the relaxation problem. It depicts a line cutting off point 'A' from the polyhedron. This line is a strong valid inequality. Deriving facets of a MIP problem improves the quality of the solution and computational

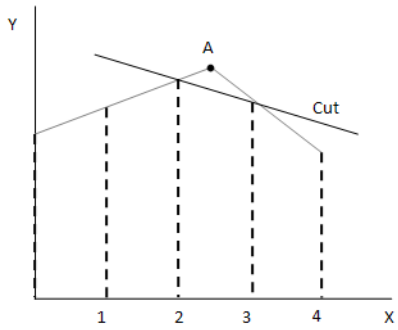


Figure 1.7: Mixed integer programming polyhedron

time.

Two sets of cuts can be added to the algorithm built in current commercial solvers. One is called lazy cuts and the other is user defined cuts. Lazy cuts are used via the methodology the commercial solver uses to pick a cut or a set of cuts from a pool of cuts. The set of user defined cuts are a set of cuts that are applied based on a methodology the user defines for the branch and cut algorithm. In the current study the method used for adding the cuts is based on the lazy cuts; however, developing efficient algorithm to customize the application of the cuts can improve the efficiency of the algorithm.

Results compare the quality of the solution before and after adding these cuts to the problem. Results show how much dose is absorbed in the critical organ for each geometry setting.

In the second part of our study, we apply an approximate gradient method to provide a method to search the continuous geometry and angles while minimizing the absorbed dose in the critical organs and remaining a reasonable image quality. This method reduces the absorbed dose by slight change in the geometry. It can be used to solve the model with a refined size of the voxels where voxels are smaller and we have more variables and constraints. The original model is very large and refining the size of the voxels cannot be handled by commercial solvers for the branch-and-cut method.

The third part of this research, investigates a robust optimization methodology to handle the uncertainty of the patient motion within a procedure. This motion can be resulted from breathing processes and other sources and changes the relative position and size of the organs. We introduce scenarios based on the breathing motion and compare the results for the absorbed dose using the robust model with each scenario solution.

As a future research this model can be applied in IMRT and a small modification of the formulation is required. However, the computation complexity needs to be investigated and compared with current methods and studies.

Chapter 2

BACKGROUND

2.1 Monte Carlo Simulation

Many studies employ Monte Carlo techniques to simulate particle tracks and interactions with matter with respect to energy spectra to estimate absorbed dose in different organs of the body. We address some of the organ dose studies using Monte Carlo in interventional radiology.

A Monte Carlo code developed by [28], simulates X-ray photon interactions in diagnostic X-ray procedures. This simulation includes the image receptor and estimates image quality in terms of contrast and signal-to-noise ratio.

Results given by Monte Carlo method for lens of the eyes in diagnostic X-ray procedures are compared with converted thermoluminescence dosimetry (TLD) ([29]) measurements [30].

A Monte Carlo code GEANT4 is an object oriented code developed in C++ which requires user code to identify the geometry, particles, physical process to execute the simulation [31]. Although, GEANT4 provides simulation of the particles, it is not specifically designed for the medical modalities. Some modifications to this code and existing algorithms enables CT simulation. A Linux cluster including 38 CPUs is used to run several pieces of the simulation program [32].

A Monte Carlo simulation of CT using GEANT4 toolkit has been implemented which includes random generation of the photons based on the spectra distribution and detection of the photons at the image receptor [33]. The energy absorbed in each pixel of the image is used for image reconstruction. An interface is designed to import data from the Digital Imaging and Communications in Medicine (DICOM) into simulation package

GEANT4 [34]. CT data are captured from DICOM and geometry parameters are fed into the simulation as geometry.

Monte Carlo code MCNP [35] used in [36] typically takes about six hours with 20 million starting particles to attain less than 0.5% uncertainty in the calculated effective dose. To compute the effective dose resulted from scattered radiation for the cardiologist, it takes four times as much processing time as required for the patient. MCNP uses mathematical phantoms to calculate the organ doses, however, authors in [36] discuss in fact that PCXMC ([37]) code may have considerably different results than MCNP for the organ doses, but it is much faster and provide relatively acceptable results in pediatric patients from catheterizations in cases of (suspected) congenital heart defects.

The Monte Carlo code MCNP has been used for dosimetry in medical irradiations within user defined 3-dimensional geometry. The image simulation has been added from the version MCNPX ([38]). Study [39] uses MCNPX code and a FAX model of the patient ([40]) to achieve the tube potential value in the computed radiography of the chest which results in the low patient dose and reasonable image quality.

Authors in [41] study pediatric patient effective doses undergone interventional cardiac procedures. Effective doses are calculated using Monte Carlo simulation and the exact radiation geometry of the X-ray tube and energy settings are used for a particular projection in a patient. An anthropomorphic phantom is used for the simulation of the patient anatomy based on the gender and length of the patient.

Bozkurt and Bor [42] employ Monte Carlo code MCNP to report the organ equivalent doses and compute the effective doses for the patient and primary physician during interventional cardiological procedures. Two voxel-based VIP-Man models are used for the patient and physician. Five most common projections with equal field sizes

were tested and seven tube potentials were studied for each projection. They indicated that effective doses for the patient determined in their study was lower than literature data and effective dose for the physician agreed with the literature data.

Organ doses for different ages and genders are indicated using simulation results by a software called CALDose [43]. The inputs include gender, potential (kV), Focus to skin distance (FSD) in cm, charge (mA), time (s), filtration Al (mm), and type of the procedure which specifies body part. The output is the total absorbed dose for each organ in addition to the effective dose and cancer risk.

Monte Carlo method is time consuming and needs to be patient-specific to have more accurate results, so it is not reasonable to run simulation for every procedure. The total time required depends upon the projection and is varies based on the phantom and codes.

2.2 Optimization in Computed Tomography (CT)

International Electrotechnical Commission [44] requires every CT system displays the average absorbed dose regarding the exposure parameters. A definition for the image quality takes into account noise and special resolution. Having a constant special resolution, image noise is proportional to the inverse of the square root of radiation dose [45]. Some dosimetry studies in CT address the current modulation of X-ray tube [45, 46, 47, 48, 49, 50] using methods other than Monte Carlo.

Authors in [45] develop a method to determine each scan parameters regarding radiation dose reduction and obtaining the same level of the image noise in different sizes of the patients. Using the exponential attenuation results in unscattered photons on the image receptor and variance of the linear attenuation, determined by Poisson statistics at the center of the homogeneous circular cylinder, establishes the relationship between

standard deviation of the pixel value, image noise, special resolution, attenuation coefficient and radiation dose. However, this method is valid for the center of the rotation and central region of the image. This derivation assumes parallel beam while the real CT emits beams more like fan shape beams.

The value of Computed Tomography Dose Index (CTDI) is based on measuring the absorbed dose in a cylindrical acrylic phantom and used to estimate the absorbed dose resulted from CT radiation. The dependence of this value on the size of the phantom, the selected kV and the scan mode are addressed in [46]. Current acrylic phantoms are in two sizes which does not represent real size of the patients. The relationship for the various scan parameters is formulated, and enables calculating CTDI for a continuous range of selectable scan variables (e.g., kVp, mmAs, scanner type, scan mode, patient size).

Automatic Tube Current Modulation is addressed as an embedded feature of the recent CT technologies [47]. This technique includes angular (x-y) modulation and z-axis modulation which maintains constant image quality. Measurements are performed for X-ray beam absorption in 100 central channels from lateral and anteroposterior views. Sinusoidal modulation for the tube current is achieved by preprogramming of 360 degrees of rotation. In z-axis modulation, a localizer radiograph is used accompanied by a scanner which computes the required tube current to attain the images within a selected noise level.

Traditional methods for dose calculation are based on the parametrizing dose distributions measured using water phantoms and employing correction factors to the beam for the nonuniform contour of the patient and tissue heterogeneities. However, convolution kernels are most often achieved by using Monte Carlo methods for transporting charged particles [48].

A developed program captures scan protocols from DICOM standard and performs the automatic computation for the CT patient doses [49]. This study use organ conversion factors for the average size and weight of adults, a 7-year-old child and 8-year-old babies for the both sexes.

CT dose metrics including CTDI, organ dose, and effective dose are evaluated based on different tube voltages and contrast-to-noise ratio for a Rando phantom [50]. Two criteria mentioned as optimization criteria are: having a constant contrast-to noise ratio at each tube voltage to specify the one minimizing the patient dose, having a constant dose for each tube voltage and determine the value maximizing the contrast-to-noise ratio. The result of the experiments indicates that the CT optimization with respect to tube voltage needs to consider patient effective dose rather than CTDI in the air or body.

2.3 Optimization in Intensity-Modulated Radiation Therapy (IMRT)

An advance method in radiation therapy called intensity-modulated radiation therapy (IMRT) delivers precise radiation to a tumor utilizing computer-controlled linear accelerators. It has the capability to be conformed to a three-dimensional (3D) shape of the tumor by modulating the intensity of radiation beams. It is desired to radiate tumor required dose and avoid surrounding organs spatially organs at risk [7]. Pencil beams used for treating cancer tumor are referred as beamlets. A fluence map is a matrix of beamlet weights [51]. Major problems discussed in IMRT are beam angle optimization, fluence map optimization and beam segmentation.

A number of linear, quadratic and integer programming models have been proposed for finding the optimal set of beam angles and corresponding fluence maps. A review of formulations can be found in [51, 52, 53]. Ehrgott et al [52] survey the use of optimization models discussed in IMRT for the selection of the beam angles, intensity map optimization, and sequence of the multileaf collimator reconfigurations including

linear and nonlinear models. The linear model minimizes the average dose on the critical structures while maximizing target dose[51, 54, 55]. A greedy search technique can be used to determine the parameters of coverage and conformity as well as beam angle selection [54]. Two sets of full-volume constraints and partial-volume constraints determine the limits that have to be satisfied by all voxels of a structure or by a specified fraction of the voxels respectively. The proportion of the voxels receiving a dose greater or equal than prescribed dose is called coverage of target structure and the ratio of the number of voxels in the structure and surrounding tissue to the number of the voxels only in the structure receiving a dose greater than or equal to the prescribed dose is called conformity of the structure[51, 54].

The nonlinear quadratic formulations are in terms of weighted least square deviation from prescribed dose and received dose in tumor, and the least square deviation from the received dose and threshold dose in critical organs and normal tissues [56, 57, 58]. Carlsson [57] reduces the problem dimension of a fluence map optimization by decomposition of the Hessian matrix corresponding to the objective function. Using a column generation approach and optimization of the segment shapes and weights, dynamically generates the multileaf collimator segments. Also using the same approach, the author finds the near optimal solutions with fewer columns comparing to the original problem dimension. A primal-dual interior point algorithm guarantees to find the optimal solution to the convex formulation of the fluence map optimization problem for a given set of beams in a clinical viable period of the time. This formulation is a quadratic programming with linear constraints [56].

Mixed integer programming is applied in relaxing dose-volume constraints [59, 60, 61], beam angle and intensity selection [51, 62, 63, 64, 65, 66]. The objective function is either minimizing doses to critical organs or maximizing doses to the targets.

If the problem is infeasible, the dose-volume constraints for each voxel are relaxed by converting to the fractional constraints. Beam angle selection can be implemented by having an indicator variable to decide selection of a known angle [51] or optimizing over several set of angles [63]. A mixed integer model optimizes beamlet weights by using an indicator variable to decide whether a beam is selected or not. An iterative method eliminates insignificant beam angles with less contribution to a better treatment plan and reduces the search space[51].

The optimality of the IMRT plans can be studied as a function of beam angles with an exhaustive search approach. Computational speed is improved by parallelizing the beam angle optimization problem [67]. Decomposition of an integral fluence map into rectangular apertures and corresponding intensities has been studied through mixed integer models as well as minimization of the total treatment time. In a bi-level optimization problem, Bender's decomposition generates columns and rows necessary within a branch-and-price-and-cut algorithm which is not computationally efficient. Thus, several valid inequalities are generated based on the structure of the problem which improve the gap between lower bound and upper bound and are able to provide the optimal solution in some instances. A partitioning strategy on input matrix also improves the gap.

Other studies minimize the total treatment time which is the amount of time required to switch from one jaw setup to another and duration of the radiation. Total set up time is assumed to be proportional to the total number of apertures in one study [62]. Some solution techniques to a large-scale problem are also discussed in [64], such as; branching on a set of binary variables rather than one variable at a time, using a heuristic which forces some of the binary variables to one, having a master problem with smaller number of body voxels and adding columns.

Chapter 3

RADIATION DOSE REDUCTION IN FGI

In this chapter, we formulate the problem of radiation dose reduction for critical organs in Fluoroscopically Guided Interventional Radiology (FGI). We consider a cardiovascular procedure where the patient breast is a critical organ and heart of the patient is the region of the interest.

3.1 FGI System Features

FGI modalities are comprised of an X-ray tube where X-ray is generated, a receptor where image is constructed, and a C-arm to connect these two. A table is located close to the equipment for patient positioning. The table can be moved in any three directions of the space. Moving collimator jaws above the X-ray tube, changes the size of the irradiated region on the patient. The goal of this study is covering the radiation field in a way that Region of Interest (ROI) retains with an acceptable resolution in the image and avoids over exposing critical organs. ROI is referred to the organ of the body which is of interest to physician.

The amount of energy produced in X-ray tube depends on the tube potential (kV), current (mA) and time (s). Lowering energy leads to smaller values of absorbed dose; however the amount of energy passing through the body and hitting the receptor will be smaller and may not have acceptable image quality for diagnosis. This system is equipped with an automatic exposure control which adjusts the exposure based on the number of photons and energy of the photon received at the image receptor. This system starts with an initial setup independent of the patient size and geometry. It does not minimize dose for the critical organs.

3.2 Assumptions

Before each procedure, there is data available prior to FGI procedure such as number of CT slices, thickness of each slice, size of field of view (at the image receptor), and source to image receptor distance, which can be used for constructing mathematical models by applying preliminary calculations. Usually a CT scan is provided to physicians prior to FGI procedures. A series of CT images can be used to construct a real patient phantom. We can discretize each CT image into equal squares, and build up a patient phantom from small volumes called voxels. We combine a set of image pixels and refer to each set as one pseudo pixel. Each voxel has a width and height equal to the size of the image pixel and length of CT slice thickness. Combined set of voxels is used to overcome the large number of variables in the model. However, when the computational capability is improved, the pseudo voxel size can be further refined. The CT number for each pixel of the CT image can be converted to the linear attenuation coefficient corresponding to the tissue [9, 68]. CT number in Hounsfield units is defined as follows,

$$\text{CT number} = \frac{\mu - \mu_{water}}{\mu_{water}} \tag{3.1}$$

where μ is the linear attenuation coefficient of the matter corresponding to a pixel of the CT image where CT number is measured. μ_{water} denotes the linear attenuation coefficient of the water. Here, we assume that for each pseudo voxel there is one attenuation coefficient μ. CT numbers also depend upon the values of the kV.

We use the same discrete size of the pseudo pixel for the image at the receptor. The energy generated by the X-ray tube is assumed to be uniformly distributed into a number of pencil beams equal to the number of the pseudo pixels considered for the image. We assign one single beam going from source to each pseudo pixel of the image at the receptor. Figure 3.1 presents a schematic view of the discretized phantom, image and

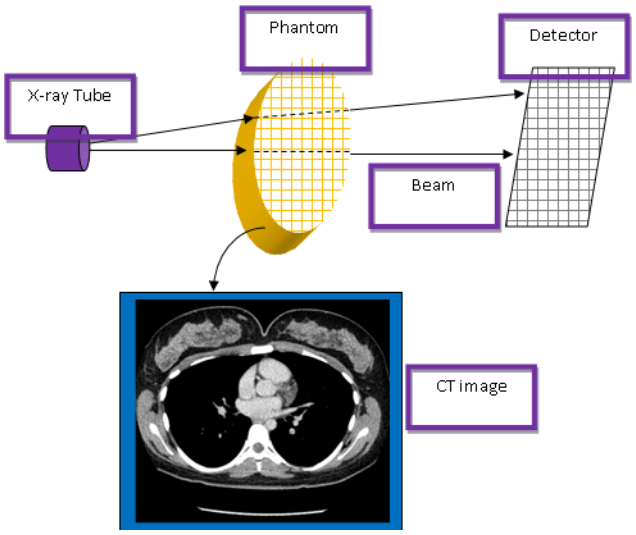


Figure 3.1: Beams going from X-ray source to image receptor through phantom pseudo voxels and are attenuated based on attenuation coefficients computed from CT slices

corresponding CT slice. The phantom for a real patient is built using the CT slices and the FGI image is constructed after each pencil beam traverses the phantom and strikes the image receptor.

By moving the collimator jaws, a set of pencil beams going to a row or column of the image are covered. Type of the procedure indicates the body part undergoing the radiology procedure. In this study, we consider a cardiovascular procedure for a female where heart is the organ which should be irradiated and is considered as Region of Interest (ROI). In this procedure the patient breast is considered as a critical organ based on the statistics given in the chapter 1 which indicates high risk of cancer for the female range age of ≤ 40.

For each pencil beam intercepting body pseudo voxels, we calculate absorbed energy in each pseudo voxel. Since attenuation values depend on the tube potential, we fix kV so energy varies by changing current (mA) and time(s). Problem can be solved for a set of kV potentials and the best kV will be determined. Since a setting for the kV alters

image contrast, it can be evaluated in the methods for improving the image quality.

It is assumed that each beam is attenuated relative to the exponential attenuation of a patient pseudo voxel. This is an approximation and can be refined in the future studies if necessary. Energy attenuated by the air before the beam enters patient and after exiting from the patient to the image receptor is small and negligible because of the small value of the air attenuation coefficient (1/1000 of μ_{water}).

The distance between the X-ray tube and receptor is assumed to be known. However, since it has effect on the image contrast, a general framework to address this problem is to solve a set of problems in different setups and select the best configuration. The table height, lateral and longitudinal increments, radiation field size and the amount of the source energy are to be determined for each scan while minimizing the total absorbed dose in the critical organ. Each axis in the space is discretized by the same size of the pseudo voxels in the patient for the corresponding direction. We assign a set of binary variables to each direction. These binary variables are x_i, y_j, and z_l in directions x, y, and z respectively. We assume that for the default patient location, all pseudo voxels of the patient are completely matching space pseudo voxels. Any change in the table position relative to X-ray source can be determined by these three variables which indicate the location of the patient. The angles of the equipment are relative to the patient lateral and longitudinal directions and assumed known for each setting. The selection of the best angle configuration can be achieved by solving the problem for a set of the most commonly used angles. Changing the angles alters the position of the X-ray tube and image receptor in a polar system.

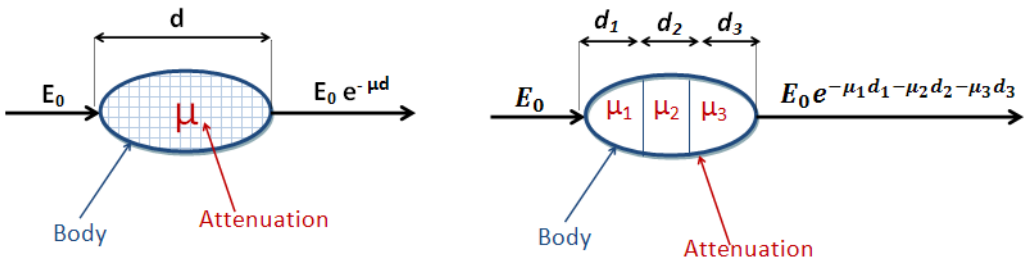
Mass of each voxel is determined based on the table of organ or tissue masses presented in [69]. For a pseudo voxel, mass is determined by mass of one voxel times the number of voxels combined in a pseudo voxel. The amount of absorbed dose for each

voxel is defined as the energy (ΔE) absorbed by the ionizing radiation per unit mass of material (Δm) [9] and calculated based on the 1.1.

$$\Delta E = E_{entrance} - E_{exit} = E_k - E_k\, e^{-\mu_k d} \tag{3.2}$$

where E_k is the entrance energy at voxel k th, μ_k is the linear attenuation coefficient of voxel k th and d is the length of the voxel the pencil beam traverses which is assumed to be equal for all the voxels.

Figure 3.2(a) presents the amount of energy exiting a homogenous material. Human body consists of several different tissues and organs with different attenuations. Figure 3.2(b) presents the value of the energy when leaving a material. It is assumed that



(a) Homogenous (b) Nonhomogenous

Figure 3.2: Attenuation of the energy, (a) pencil beam traversing a homogenous material (b) pencil beam traversing a nonhomogenous material

photons after striking the image receptor are stopped and their energy is completely absorbed in the receptor. $\mu_{i'j'}$ indicates the attenuation of the receptor for each pseudo pixel $i'j'$ and d is the distance traversed in the receptor. The value of $\mu_{i'j'}$ is considered to be fix for all the pseudo pixels of the image. Field of view is assumed a known parameter for this problem.

3.3 Mathematical model

We present a mixed-integer formulation for minimizing radiation dose in the critical organs in FGI modalities. This formulation is a general formulation in terms of the procedure type and corresponding critical organ.

3.3.1 Notations

A list of the notations used for the model is provided to facilitate understanding of the formulation.

Indices

o = Organ index

$i \in \{l_x, \ldots, u_x\}$: index of increment in x direction

$j \in \{l_y, \ldots, u_y\}$: index of increment in y direction

$l \in \{l_z, \ldots, u_z\}$: index of increment in z direction

i': index of rows of the image on the image receptor

j': index of columns of the image on the image receptor

n: number of the image rows/columns

l_x: min index in x direction

u_x: max index in x direction

l_y: min index in y direction

u_y: max index in y direction

l_z: min index in z direction

u_z: max index in z direction

l_{max}: max phantom index in z direction

$(\hat{i}, \hat{j}, \hat{l})$: index of a pseudo voxel in space

(i_p, j_p, l_p): index of a pseudo voxel in the body

Sets

VS: set of space pseudo voxels

VP: set of phantom pseudo voxels

C_o: set of pseudo voxels belonging to the critical organ, $C_o \subseteq VP$

$S_{i'j'}$: set of the space pseudo voxels intercepted with beam $i'j'$ for a fix geometry setting

Parameters

d: distance traversed in each pseudo voxel (mm)

$V_{(i_p,j_p,l_p)}$: phantom pseudo voxel (i_p, j_p, l_p)

$dose_{V_{(i_p,j_p,l_p)}}$: Absorbed dose in pseudo voxel (i_p, j_p, l_p) (Gy)

d': distance traversed in the image receptor (mm)

$dose_{V_{detector_{i'j'}}}$: dose for a pseudo pixel $i'j'$ of the image (Gy)

$m_{V_{(i_p,j_p,l_p)}}$: mass of the pseudo voxel (i_p, j_p, l_p) (kg)

$\mu_{(i_p,j_p,l_p)}$: linear attenuation coefficient of a pseudo voxel (i_p, j_p, l_p) (1/mm)

$\mu_{i'j'}$: linear attenuation coefficient corresponding to the $i'j'$ pseudo pixel of the image receptor with a known height (1/mm)

$\overline{Int}_{bg}$: average intensity of the background of the ROI (J)

$\overline{Int}_{ROI}$: average intensity over the ROI (J)

$m_{detector}$: mass of detector (kg)

ROI: region of interest of the phantom

K_1: lower bound of SNR

K_2: upper bound of SNR

$$\xi_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} = \begin{cases} 1 & \text{if there exist } i,\ j,\ l \text{ where } \hat{i} = i_p + i,\ \hat{j} = j_p + j,\ \hat{l} = l_p + l \\ 0 & \text{otherwise} \end{cases}$$

3.3.2 *Variables*

We introduce notations used for declaring decision variables including continuous and binary variables,

E_0 = Initial energy of each beam at source

$$\eta_{i'} = \begin{cases} 1 & \text{if row } i' \text{ of the image is not covered by collimator/wedges} \\ 0 & \text{otherwise} \end{cases}$$

$$\delta_{j'} = \begin{cases} 1 & \text{if column } j' \text{ of the image is not covered by collimator/wedges} \\ 0 & \text{otherwise} \end{cases}$$

$$\gamma_{i'j'} = \begin{cases} 1 & \text{if } \eta_{i'} = 1 \text{ and } \delta_{j'} = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$x_i = \begin{cases} 1 & \text{if longitudinal table increment relative to the origin is } i \text{ units of discrete size} \\ 0 & \text{otherwise} \end{cases}$$

$$y_j = \begin{cases} 1 & \text{if lateral table increment is } j \text{ units of the discrete size} \\ 0 & \text{otherwise} \end{cases}$$

$$z_l = \begin{cases} 1 & \text{if table height relative to the origin is } l \text{ units of the discrete size} \\ 0 & \text{otherwise} \end{cases}$$

$$\psi_{(i_p,j_p,l_p),i'j'} = \begin{cases} 1 & \text{if beam } i'j' \text{ intercepts pseudo voxel } (i_p, j_p, l_p) \text{ of the phantom :} \\ & \gamma_{i'j'} = 1 \text{ and } x_i = 1,\, y_j = 1,\, z_l = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} = \begin{cases} 1 & \text{if voxel } (i_p,j_p,l_p) \text{ matches with pseudo voxel } \hat{i}\hat{j}\hat{l} \\ & \text{after increments } i,\ j,\ l, \text{ where } i=\hat{i}-i_p,\ j=\hat{j}-j_p,\ l=\hat{l}-l_p \\ & \text{and beam } i'j' \text{ intercepts with this pseudo voxel} \\ 0 & \text{otherwise} \end{cases}$$

$E_{l_p,i'j'} =$ Energy at layer l_p of the phantom when beam $i'j'$ is passing

$E_{SD,i'j'} =$ Energy at the $i'j'$ pseudo pixel of the image receptor $i'j'$

$$\tau_{i'j'} = \begin{cases} E_{SD,i'j'} & \text{if } \psi_{(i_p,j_p,l_p),i'j'} = 1\ , (i_p,j_p,l_p) \in ROI \\ 0 & \text{otherwise} \end{cases}$$

$$\alpha_{i'j'} = \begin{cases} 1 & \text{if } \psi_{(i_p,j_p,l_p),i'j'} = 1\ , (i_p,j_p,l_p) \in ROI \\ 0 & \text{otherwise} \end{cases}$$

$$\Gamma_{(i_p,j_p,l_p-1),i'j'} = \begin{cases} E_{l_p,i'j'} & \text{if } \psi_{(i_p,j_p,l_p-1),i'j'} = 1 \\ 0 & \text{otherwise} \end{cases}$$

To determine whether variable $\psi_{(i_p,j_p,l_p),i'j'}$ is zero or one, we need to have binary variable $\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'}$ be equal to one to specify the destination location in the space which phantom can be moved to and for corresponding pseudo voxel in that destination, a specific beam intercepts the pseudo voxel which should exists in the radiation field and that is $\gamma_{i'j'}$. Since we defined a set of space pseudo voxels which are intercepted by a specific pencil beam, we have $\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'}$ equal to one if for corresponding pencil beam, the coordinate of a phantom pseudo voxel matches with one of the space pseudo voxels in the set and this location be selected as the destination to move the phantom to. Parameter $\xi_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'}$ is known for a specific angle configuration and indicates a

pseudo voxel of the space intercepting with each pencil beam. We set indices of the corner of the room to be (0, 0, 0). Figure 3.3 demonstrates the geometry of the mathematical model. This figure depicts two pencil beams going from the X-ray source to the image

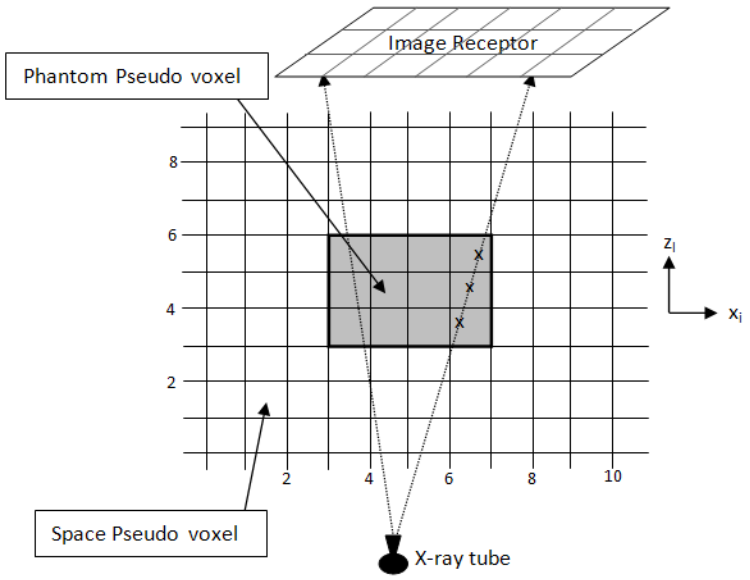


Figure 3.3: Demonstration of geometry used in the mathematical model

receptor while traversing through the phantom. Cross symbols on the phantom are the interception points of the pencil beam with patient pseudo voxels. We assign one binary variable for every x_i location and one for every z_l location in the space. This figure depicts geometry corresponding to binary decision variables $x_0 \ldots x_1 0$ and $z_0, \ldots, z_9$. This figure is in two dimensions; however, the problem is in three dimensions. Parameter $\xi_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'}$ is computed for each geometry setting that angles are known and determines all sets of $S_{i'j'}$ associated with beam $i'j'$. This figure is used for the illustration of the geometry and patient real phantom is constructed using CT slices.

3.3.3 Optimization Model

We represent the formulation of the mixed-integer program where the objective function minimizes the dose received in the critical organ with respect to the constraints which determine the table geometry, location for the collimator jaws, and range of the SNR.

$$\textbf{Min} \sum_{V_{(i_p,j_p,l_p)} \in C_o} dose_{V_{(i_p,j_p,l_p)}} = \sum_{(i_p,j_p,l_p): V_{(i_p,j_p,l_p)} \in C_o} \sum_{i'=1}^{n} \sum_{j'=1}^{n} \frac{\Gamma_{(i_p,j_p,l_p),i'j'}\left(1 - e^{-\mu_{(i_p,j_p,l_p)}d}\right)}{m_{V_{(i_p,j_p,l_p)}}} \tag{3.3}$$

Subject to

- Constraints to identify which pseudo pixels of the detector are in the field.

$$\gamma_{i'j'} \leq \eta_{i'} \qquad i', j' = 1, \ldots, n \tag{3.4}$$

$$\gamma_{i'j'} \leq \delta_{j'} \qquad i', j' = 1, \ldots, n \tag{3.5}$$

$$\gamma_{i'j'} \geq \eta_{i'} + \delta_{j'} - 1 \qquad i', j' = 1, \ldots, n \tag{3.6}$$

$$\delta_{j'} \geq \delta_p + \delta_t - 1 \qquad p \leq j' \leq t,\ \ p = 1, \ldots, n-2,\ t = p+2, \ldots, n,\ j' = p+1, \ldots, t-1 \tag{3.7}$$

$$\eta_{i'} \geq \eta_k + \eta_h - 1 \qquad k \leq i' \leq h,\ \ k = 1, \ldots, n-2,\ h = k+2, \ldots, n,\ i' = k+1, \ldots, h-1 \tag{3.8}$$

- Constraint to relate $\psi_{(i_p,j_p,l_p),i'j'}$ to $\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'}$

$$\psi_{(i_p,j_p,l_p),i'j'} \geq \beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} + \gamma_{i'j'} - 1 \quad \forall\,(i_p,j_p,l_p) \in VP,\ (\hat{i},\hat{j},\hat{l}) \in S_{i'j'},\ \hat{i} = i + i_p,\ \hat{j} = j + j_p,\ \hat{l} = l + l_p,\ i',j' = 1,..,n \tag{3.9}$$

$$\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \leq \gamma_{i'j'} \quad \forall\,(i_p,j_p,l_p) \in VP,\ i',\ j' = 1,..,n \tag{3.10}$$

$$3\,\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \leq x_i + y_j + z_l \quad \forall (i_p,j_p,l_p) \in VP, \hat{i} - i_p = i,\ \hat{j} - j_p = j,\ \hat{l} - l_p = l,\ i',j' = 1,..,n \tag{3.11}$$

$$\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \leq \xi_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \quad \forall (i_p,j_p,l_p) \in VP, \hat{i} - i_p = i,\ \hat{j} - j_p = j,\ \hat{l} - l_p = l,\ i',j' = 1,..,n \tag{3.12}$$

$$\psi_{(i_p,j_p,l_p),i'j'} \leq \sum_{(\hat{i},\hat{j},\hat{l}) \in S_{i'j'}} \beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \quad \forall (i_p,j_p,l_p),\ \in VP,\ \forall i'j' = 1,..,n \tag{3.13}$$

$$x_i + y_j + z_l + \gamma_{i'j'} + \xi_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} - 4 \leq \beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \quad \forall (i_p,j_p,l_p) \in VP,\ \hat{i} - i_p = i,\ \hat{j} - j_p = j,\ \hat{l} - l_p = l,\ \forall\, i'j' \tag{3.14}$$

- Constraint to linearize objective function

$$\Gamma_{(i_p,j_p,l_p),i'j'} \geq E_{l_p,i'j'} - M(1 - \psi_{(i_p,j_p,l_p),i'j'}) \quad \forall\, i'j',\ (i_p,j_p,l_p) \in C_o \tag{3.15}$$

- Constraints to ensure that one variable in each direction determines the location of the table

$$\sum_{i=l_x}^{u_x} x_i = 1 \tag{3.16}$$

$$\sum_{j=l_y}^{u_y} y_j = 1 \tag{3.17}$$

$$\sum_{l=l_z}^{u_z} z_l = 1 \tag{3.18}$$

- Constraints to connect different levels of energy of a pencil beam

$$E_{l_p-1,i'j'} \leq E_{l_p,i'j'} + M \sum_{i_p \in VP} \sum_{j_p \in VP} \psi_{(i_p,j_p,l_p-1),i'j'} \qquad \forall\, i', j' = 1,\ldots,n,$$

$$(i_p, j_p, l_p) \in VP \tag{3.19}$$

$$E_{l_p,i'j'} \geq E_{l_p-1,i'j'} e^{-\mu_{(i_p,j_p,l_p-1)}d} - M(1 - \psi_{(i_p,j_p,l_p-1),i'j'}) \quad \forall\, i', j' = 1,\ldots,n,$$

$$(i_p, j_p, l_p) \in VP \tag{3.20}$$

$$E_{l_p,i'j'} \leq E_{l_p-1,i'j'} e^{-\mu_{(i_p,j_p,l_p-1)}d} + M(1 - \psi_{(i_p,j_p,l_p-1),i'j'}) \quad \forall\, i', j' = 1,\ldots,n,$$

$$(i_p, j_p, l_p) \in VP \tag{3.21}$$

$$\sum_{i_p \in VP} \sum_{j_p \in VP} \psi_{(i_p,j_p,l_p),i'j'} \leq 1 \qquad \forall\, i', j' = 1,\ldots,n,$$

$$(i_p, j_p, l_p) \in VP \tag{3.22}$$

- Constraints to specify the range for energy at each level

$$E_{l_p,i'j'} \leq E_{l_p-1,i'j'} \qquad \forall\, i', j' = 1,\ldots,n,\ l_p \in VP \tag{3.23}$$

$$n^2\, E_{0,i'j'} - E_0 = 0 \qquad \forall\, i', j' = 1,\ldots,n \tag{3.24}$$

$$E_{SD,i'j'} = E_{l_p,i'j'} \qquad \forall\, i', j' = 1,\ldots,n,\ l_p \in VP \tag{3.25}$$

- Constraints to ensure the ROI will not be covered by the collimator jaws

$$\sum_{i'=1}^{n} \sum_{j'=1}^{n} \psi_{(i_p,j_p,l_p),i'j'} \geq 1 \qquad \forall\, (i_p, j_p, l_p) \in ROI \tag{3.26}$$

- Image quality constraint

$$K_1'' \sum_{i'=1}^{n} \sum_{j'=1}^{n} \alpha_{i'j'} \leq \sum_{i'=1}^{n} \sum_{j'=1}^{n} \tau_{i'j'} \tag{3.27}$$

$$\sum_{i'=1}^{n} \sum_{j'=1}^{n} \tau_{i'j'} \leq K_2'' \sum_{i'=1}^{n} \sum_{j'=1}^{n} \alpha_{i'j'} \tag{3.28}$$

- Image quality complementary constraint for linearizing and avoiding division by zero

$$\tau_{i'j'} \leq E_{SD,i'j'} \qquad \forall\, i', j' = 1,\ldots,n \tag{3.29}$$

$$\tau_{i'j'} \leq M\alpha_{i'j'} \qquad \forall i', j' = 1,\ldots,n \tag{3.30}$$

$$\tau_{i'j'} \geq E_{SD,i'j'} - M\,(1-\alpha_{i'j'}) \qquad \forall i', j' = 1,\ldots,n,\ M = E_{max} \tag{3.31}$$

$$\alpha_{i'j'} \geq \psi_{(i_p,j_p,l_p),i'j'} \qquad \forall i', j' = 1,\ldots,n,\ (i_p,j_p,l_p) \in ROI \tag{3.32}$$

$$\alpha_{i'j'} \leq \sum_{i_p,j_p,l_p \in ROI} \psi_{(i_p,j_p,l_p),i'j'} \qquad \forall i', j' = 1,\ldots,n \tag{3.33}$$

- Variable range

$$\psi_{(i_p,j_p,l_p),i'j'} \in \{0,1\} \tag{3.34}$$

$$\Gamma_{l_p,i'j'} \geq 0 \tag{3.35}$$

$$\eta_{i'}, \delta_{j'} \in \{0,1\} \tag{3.36}$$

$$\gamma_{i'j'} \in \{0,1\} \tag{3.37}$$

$$\alpha_{i'j'} \in \{0,1\} \tag{3.38}$$

$$\tau_{i'j'} \geq 0 \tag{3.39}$$

The objective function minimizes the accumulative dose which is absorbed in the pseudo voxels of the critical organ.

Constraints (3.4), (3.5) and (3.6) determines whether a beam from source to a pseudo pixel $i'j'$ of the image exists or not. This is determined by the variable $\gamma_{i'j'} = \eta_{i'}\delta_{j'}$. Variables $\eta_{i'}$ are corresponding to the rows of the image and variables $\delta_{j'}$ are corresponding to the columns of the image. These three constraints linearizes the variable $\gamma_{i'j'} = \eta_{i'}\delta_{j'}$ which indicates the existence of the beam corresponding to pseudo pixel $i'j'$. Two other set of constraints (3.7) and (3.8) ensure the connectivity of the image pseudo

pixels intercepted by the beams. These two constraints include $\frac{n(n-1)(n-2)}{6}$ members each, and enforce the continuity of the image regarding beam existence.

The other set of constraints are corresponding to $\psi_{(i_p,j_p,l_p),i'j'}$ which is defined to indicate interception of a beam with a pseudo voxel of the body. For that, two binary variables of $\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'}$ and $\gamma_{i'j'}$ should be equal to one. Having $\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} = 1$ means there exists a pseudo voxel of the body matching a pseudo voxel of the space with applied table increments where beam $i'j'$ traverses the patient pseudo voxel. Inequalities (5.7), (5.8), and (5.11) are defined to lineariz the formula $\psi_{(i_p,j_p,l_p),i'j'} = \lfloor \frac{\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} + \gamma_{i'j'}}{2} \rfloor$. Constraints (5.9) and (5.10) determine the matching of increments and variable $\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'}$. Constraint (5.12) enforces the varaible $\psi_{(i_p,j_p,l_p),i'j'}$ be equal to one if a pencil beam $\gamma[i][j] = 1$ and corrsponding x_i, y_j, z_l, and $\xi_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'}$ are equal to one.

Constraint (5.13) is defined to linearize the definition of variable $\Gamma_{(i_p,j_p,l_p),i'j'}$ which is equal to $\psi_{(i_p,j_p,l_p),i'j'}\ E_{l_p,i'j'}$.

Three constraints (3.16), (3.17), and (3.18) ensure selection of one variable indicating the increment in one direction of x, y, or z.

Different layers of the energy have to be determined when a beam is traversing the organs. Constraint (5.17) enforces the beam interception with one pseudo voxel at each layer. Two constraints (5.15) and (5.16) enforces the pencil beam energy be attenuated when passing a pseudo voxel of the body.

Constraint (3.26) activates all the beams passing through the ROI to ensure corresponding pixels of the image are visible.

Image quality constraints (5.21) and (5.22) determine the range of the values that $\frac{\sum_{i'j'} \tau_{i'j'}}{\sum_{l'j'} \alpha_{i'j'}}$ can take to have a reasonable image quality. Complementary constraints linearize

the variable $\tau_{i'j'} = \alpha_{i'j'} E_{SD,i'j'}$. The variable of $\alpha_{i'j'}$ is corresponding to the image pixels for a beam hitting ROI.

3.3.4 Embedded Computation

Having E_0 as initial X-ray source energy, patient entrance energy of beam i', j' at the first level is calculated as follows,

$$E_{1,i'j'} = E_0 \tag{3.40}$$

In the next level, there are two possibilities; intercepting with a pseudo voxel of the phantom or traversing the air. In case of traversing the air, there is no decay of energy and the next level of this beam has the same amount of energy E_0 (inequalities 5.14 and 5.18). In case of intercepting with a voxel of the phantom, energy entering in level one is calculated as follows,

$$E_{1,i'j'} = E_0 \, e^{-\mu_{(i_p,j_p,l_p)} d} \tag{3.41}$$

Thus, for any layer of the space in l direction, we calculate corresponding energy as follows,

$$E_{l+1,i'j'} = E_{l,i'j'} \; e^{-\mu_{(i_p,j_p,l_p)} d} \tag{3.42}$$

and by assuming zero attenuation for the air, the amount of energy received at the receptor is equal to the amount of energy exiting the last layer of the discretized patient. We denote the amount of intensity received at a image pixel $i'j'$ by $E_{SD,i'j'}$.

Here we prove a lemma which will be used in SNR computation and deriving cuts.

Lemma 1. $\sum_{i'j'} \alpha_{i'j'} \geq 1$

Proof. According to definition $\alpha_{i'j'}$, it takes value 1 if $\psi_{(i_p,j_p,l_p),i'j'} = 1$, $(i_p, j_p, l_p) \in ROI$. Constraint (3.26) ensures that $\exists \; \psi_{(i_p,j_p,l_p),i'j'} = 1$ and $(i_p, j_p, l_p) \in ROI$, and by constraint (5.26),it implies corresponding $\alpha_{i'j'} = 1$, therefore $\sum_{i'j'} \alpha_{i'j'} \geq 1$. □

3.3.5 Calculation of the image quality quantifier

To quantify the image quality, there are two factors to exploit the contrast of the image which is how distinguishable objects are in an image and noise of the image which is standard deviation of the signal. Signal-to-Noise Ratio (SNR) is a combined factor and better criteria for evaluation of the image quality in current imaging systems. It can be defined as the ratio for the average intensity of the ROI to the standard deviation of the background [25]. However, in a high-contrast scene, many imaging systems including digital X-ray imaging systems represent the background with a uniform black color forcing standard deviation to be zero. In this case, a better definition of SNR can be stated as the ratio for the average intensity of ROI to standard deviation of ROI. Based on this definition, SNR can be calculated from the following equations,

$$SNR = \frac{\overline{Int}_{ROI}}{\sigma_{ROI}} \tag{3.43}$$

Where $\overline{Int}_{ROI}$ is computed from the following formula,

$$\overline{Int}_{ROI} = \frac{\sum_{i'j'} \tau_{i'j'}}{\sum_{i'j'} \alpha_{i'j'}} \tag{3.44}$$

Since noise of the image is proportional to the square root of the average dose, and we are only concerned for the ROI, average dose is calculated for the ROI [45]. Without loss of generality, we assume that all the energy received at the receptor will be absorbed in the receptor.

$$\sigma \propto \frac{1}{\sqrt{\overline{dose}_{ROI}}} \Rightarrow \sigma \propto \frac{1}{\sqrt{\frac{\sum_{i'j'} \tau_{i'j'}}{m_{detector} \sum_{i'j'} \alpha_{i'j'}}}} \tag{3.45}$$

Now, SNR can be written in the form of,

$$SNR = \frac{\sum_{i'j'} \tau_{i'j'}}{\sum_{i'j'} \alpha_{i'j'}} \sqrt{\frac{\sum_{i'j'} \tau_{i'j'}}{m_{detector} \sum_{i'j'} \alpha_{i'j'}}} \tag{3.46}$$

Range of the acceptable quality is mentioned in [9]. It can be stated as,

$$K_1 \leq SNR \leq K_2 \tag{3.47}$$

By substituting the value of equation (3.46) in inequality (3.47), we have

$$K_1 \leq \frac{\sum_{i'j'} \tau_{i'j'}}{\sum_{i'j'} \alpha_{i'j'}} \sqrt{\frac{\sum_{i'j'} \tau_{i'j'}}{m_{detector} \sum_{i'j'} \alpha_{i'j'}}} \leq K_2 \tag{3.48}$$

and we can write as follows,

$$K_1 \leq (\frac{\sum_{i'j'} \tau_{i'j'}}{m_{detector} \sum_{i'j'} \alpha_{i'j'}})^{3/2} \leq K_2 \Rightarrow \; K_1^{2/3} m_{detector} \leq \frac{\sum_{i'j'} \tau_{i'j'}}{\sum_{i'j'} \alpha_{i'j'}} \leq K_2^{2/3} m_{detector} \tag{3.49}$$

By Lemma 1, we can multiply each side of by $\sum_{i'j'} \alpha_{i'j'}$, and following linear is attained for the image quality.

$$K_1'' \sum_{i'j'} \alpha_{i'j'} \leq \sum_{i'j'} \tau_{i'j'} \leq K_2'' \sum_{i'j'} \alpha_{i'j'} \tag{3.50}$$

where $K_1'' = K_1^{2/3} m_{detector}$ and $K_2'' = K_2^{2/3} m_{detector}$.

3.4 Polyhedral Analysis

The problem formulated in section 3.3.3, is a large-scale mixed-integer program with a large number of binary variables and constraints. The computation using commercial branch and cut solvers is time-consuming (more than 15 hours) and the quality of the solution is low in terms of the optimality gap attained by the solver. To increase the speed of the algorithm, the feasible region of the problem is decomposed to smaller substructures including a subset of constraints. We investigate each substructure of problem constraints to derive strong valid inequalities hoping to solve the problem faster with the minimum optimality gap.

In this problem, the feasible region is not bounded, but it has bounded solution since it is minimization problem, and there is no negative direction to improve the objective function. By deriving strong valid inequalities, we try to provide tighter bounds

closer to the convex hall of the MIP feasible region. We call polyhedron P an integral polyhedron if all the extreme points are integral.

3.4.1 Beam Inequalities

We first consider the set of inequalities (3.4)-(3.8) and we define polyhedron P_1 as follows,

$P_1 = \{(\boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\gamma}) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n^2} : \gamma_{i'j'} \leq \eta_{i'}, \gamma_{i'j'} \leq \delta_{j'}, \eta_{i'} + \delta_{j'} - \gamma_{i'j'} \leq 1, \delta_p + \delta_t - \delta_{j'} \leq 1, \eta_k + \eta_h - \eta_{i'} \leq 1,\ p \leq j' \leq t,\ k \leq i' \leq h,\ 0 \leq \gamma_{i'j'} \leq 1,\ 0 \leq \eta_{i'} \leq 1, 0 \leq \delta_{j'} \leq 1,\ \forall i', j' = 1, \ldots, n\}$ and the set of $S_1 = P_1 \cap \mathbb{B}^{2n+n^2}$. We will first show that $conv(S_1)$ is full-dimensional. Let denote the unit vector of $(0, \ldots, 1, \ldots, 0)$ by e_i, where the ith element is equal to one and the rest of the variables are zero. We assign vector $e_i,\ i = 1, \ldots, n$ to present unit vectors for set of variables $\eta_{i'}$, vector $e_j,\ j' = 1, \ldots, n$ for set of variables $\delta_{j'}$ and vector $e_{i,j},\ \forall\, i = 1, \ldots, n,\ , j = 1, \ldots, n$ corresponding to $\gamma_{i'j'}$. We show that there are $n^2 + 2n + 1$ affinely independent points for the $conv(S_1)$ as follows,

$(\mathbf{0}, \mathbf{0}, \mathbf{0})$

$(e_i, \mathbf{0}, \mathbf{0}),\ \forall\, i = 1, \ldots, n$

$(\mathbf{0}, e_j, \mathbf{0},),\ \forall\, j = 1, \ldots, n$

$(e_i, e_j, e_{i,j}),\ \forall\, i = 1, \ldots, n,\ , j = 1, \ldots, n$

We show that P_1 is not an integral polyhedron ($P_1 \neq conv(S_1)$). We find a fractional point which is an extreme point of P_1.

We can find an extreme point of $\eta_1 = 1/2,\ \eta_2 = 0,\ \eta_3 = 1,\ \eta_{i'} = 0,\ \forall\, i' \in 4, \ldots, n,$ $\delta_1 = 1/2,\ \delta_2 = 0,\ \delta_3 = 1/2,\ \delta_{j'} = 0,\ \forall\, j' \in 4, \ldots, n,\ \gamma_{11} = 1/2,\ \gamma_{12} = 0,\ \gamma_{13} = 1/2,$ $\gamma_{i'j'} = 0,\ \forall\, i', j' \in 4, \ldots, n,\ \gamma_{2j'} = 0,\ \forall\, j' \in 1, \ldots, n,\ \gamma_{31} = 1/2,\ \gamma_{32} = 0,\ \gamma_{33} = 1/2,$ $\gamma_{3j'} = 0,\ \forall\, j' \in 1, \ldots, n,\ \gamma_{i'j'} = 0,\ \forall\, i' \in 4, \ldots, n, j' \in 1, \ldots, n$ which has fractional values for η, γ and δ. It is extreme point since we can find $n^2 + 2n$ linearly independent constraints. Linearly independent constraints for $n \geq 4$ can be selected from $n^2 - n$

constraints of 3.5 where $i' \in 1,3,\ldots,n,\ j' \in 1,\ldots,n$, $n-2$ constraints of 3.5 where $i' = 2,\ j' \in 2,4,\ldots,n$, two constraints of 3.6 for the pairs of $(i',j') = (2,1)$ and $(i',j') = (2,3)$, $2n$ constraints from 3.7 and 3.8 with n constraints from each such that diagonal elements of the n dimensional matrix are nonzeros. For $n = 3$, linearly independent constraints are those mentioned above plus four constraints of $\eta_3 + \delta_{j'} - \gamma_{3j'} \leq 1,\ \forall\ j' \in 1,2,3$ and $\gamma_{11} \leq \eta_1$.

Since we presented fractional extreme points for P_1, it is not an integral polyhedron.

Therefore, we will identify a set of strong inequalities for $conv(S_1)$.

For any $1 \leq i' \leq n-2,\ 1 \leq j' \leq n-2$

$$\gamma_{i'j'} + \eta_{i'+2} - \gamma_{i'+1,j'} \leq 1 \tag{3.51}$$

$$\gamma_{i',j'+2} + \eta_{i'+2} - \gamma_{i'+1,j'+2} \leq 1 \tag{3.52}$$

$$\gamma_{i'+2,j'} + \delta_{j'+2} - \gamma_{i'+2,j'+1} \leq 1 \tag{3.53}$$

$$\gamma_{i'j'} + \delta_{j'+2} - \gamma_{i',j'+1} \leq 1 \tag{3.54}$$

$$\gamma_{i',j'+2} + \delta_{j'} - \gamma_{i',j'+1} \leq 1 \tag{3.55}$$

$$\gamma_{i'+2,j'+2} + \eta_{i'} - \gamma_{i'+1,j'+2} \leq 1 \tag{3.56}$$

$$\gamma_{i'+2,j'+2} + \delta_{j'} - \gamma_{i'+2,j'+1} \leq 1 \tag{3.57}$$

$$\gamma_{i'+2,j'} + \eta_{i'} - \gamma_{i'+1,j'} \leq 1 \tag{3.58}$$

First, we show how the strong inequality such as $\gamma_{i'j'} + \eta_{i'+2} - \gamma_{i'+1,j'} \leq 1$ is driven. Consider following inequlities:

$$\eta_{i'+1} + \delta_{j'} - \gamma_{i'+1,j'} \leq 1 \tag{3.59}$$

$$\gamma_{i'j'} - \delta_{j'} \leq 1 \tag{3.60}$$

$$\gamma_{i'j'} - \eta_{i'} \leq 1 \tag{3.61}$$

$$\eta_{i'} + \eta_{i'+2} - \eta_{i'+1} \leq 1 \tag{3.62}$$

$$\eta_{i'+2} \leq 1 \tag{3.63}$$

We sum up inequalities 3.59, 3.60, 3.61, 3.62, and 3.63,

$$2\gamma_{i'j'} + 2\eta_{i'+2} - \gamma_{i'+1,j'} \leq 3 \tag{3.64}$$

We multiply the inequality (3.64) by $1/2$, and we use C-G cut [70] to round down the coefficients. We get the following facet definfing inequality,

$$\gamma_{i'j'} + \eta_{i'+2} - \gamma_{i'+1,j'} \leq 1 \tag{3.65}$$

Lemma 2. *The set of inequalities (3.51)-(3.58) are facet defining inequalities for* $conv(S_1)$.

Proof. First, we will show that inequalities (3.51)-(3.58) are valid inequalities for $conv(S_1)$. For inequality (3.51), if $\gamma_{i'j'} = 1$ and $\eta_{i'+2}$, it follows that $\eta_{i'+1} = 1$ and $\delta_{j'} = 1$, thus $\gamma_{i'+1,j'} = 1$. If any of $\gamma_{i'j'} = 1$ or $\eta_{i'+2}$, or both are zero, then the inequality is valid independent of the value of $\gamma_{i'+1,j'}$. Similar argument applies for the rest of inequalities.

We show that there are $n^2 + 2n$ affinely independent points on the face defined by each inequality (3.51)-(3.58) which proves the inequality is a facet defining. For a given i', j', we use i, j as common indices. We list $n^2 + 2n$ affinely independent points in table 3.1.

Thus the total affinely independent points are $n^2 + 2n$ and it is easy to show they are affinely independent by subtracting all $n^2 + 2n - 1$ points from one of the points. The

Table 3.1: The set of Affinely independent points

Point	Range	Number of points on the face
$(\sum_{i=k}^{i'+2} e_i, \mathbf{0}, \mathbf{0})$ $(\sum_{i=i'+2}^{k} e_i, \mathbf{0}, \mathbf{0})$	$k = 1, \ldots, n$	n
$(e_i, e_j, e_{i,j})$	$i = i'+2,$ $j = 1, \ldots, n$	n
$(e_i, \sum_{j=k}^{j'} e_j, \sum_{j=k}^{j'} e_{i,j})$ $(e_i, \sum_{j=j'}^{k} e_j, \sum_{j=j'}^{k} e_{i,j})$	$i = i',$ $k = 1, \ldots, n$	n
$(\sum_{i=i'}^{i'+2} e_i, \sum_{j=k}^{j'} e_j, \sum_{i=i'}^{i'+2} \sum_{j=k}^{j'} e_{i,j})$ $(\sum_{i=i'}^{i'+2} e_i, \sum_{j=j'}^{k} e_j, \sum_{i=i'}^{i'+2} \sum_{j=j'}^{k} e_{i,j})$	$k = 1, \ldots, n$	n
$(\sum_{i=i'+1}^{i'+2} e_i, \sum_{j=k}^{j'} e_j, \sum_{i=i'+1}^{i'+2} \sum_{j=k}^{j'} e_{i,j})$ $(\sum_{i=i'+1}^{i'+2} e_i, \sum_{j=j'}^{k} e_j, \sum_{i=i'+1}^{i'+2} \sum_{j=k}^{j'} e_{i,j})$ $(\sum_{i=i'+2}^{l} e_i, \sum_{j=k}^{j'} e_j, \sum_{i=i'+1}^{i'+2} \sum_{j=k}^{j'} e_{i,j})$ $(\sum_{i=i'+2}^{l} e_i, \sum_{j=j'}^{k} e_j, \sum_{i=i'+1}^{i'+2} \sum_{j=k}^{j'} e_{i,j})$	$k = 1, \ldots, n$ $l = i'+3, \ldots, n$	$n(n-2)$

dimension of the valid inequality is $n^2 + 2n - 1$ which indicates it is a facet defining inequality. □

We add these inequalities to the lazy cut pool for the computation.

3.4.2 Linearized ψ Inequalities

The polyhedron restricted to the constraints (5.7 - 5.11) is denoted by Q_6.

$$
\begin{aligned}
&Q_6 = conv\{(\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\psi},\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) \in \mathbb{B}^k \times \mathbb{B}^n \times \mathbb{B}^{pn^2} \times \mathbb{B}^{a_1} \times \mathbb{B}^{a_2} \times \mathbb{B}^{a_3} : \qquad (3.66)\\
&\psi_{(i_p,j_p,l_p),i'j'} \geq \beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} + \gamma_{i'j'} - 1,\ \beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \leq \gamma_{i'j'},\\
&x_i + y_j + z_l \geq 3\ \beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'},\ \beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \leq \xi_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'},\\
&\psi_{(i_p,j_p,l_p),i'j'} \leq \sum_{(\hat{i},\hat{j},\hat{l}) \in S_{i'j'}} \beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'},\\
&\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} - (x_i + y_j + z_l + \gamma_{i'j'}) \geq \xi_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} - 4,\\
&(i_p,j_p,l_p) \in VP,\ (\hat{i},\hat{j},\hat{l}) \in S_{i'j'}, \hat{i} = i + i_p,\ \hat{j} = j + j_p,\ \hat{l} = l + l_p,\ \ i',j' = 1,..,n\}
\end{aligned}
$$

We denote number of the phantom pseudo voxels by p and number of pseudo voxels in ROI by r.

We denote number of indices in directions x, y and z by a_1, a_2, and a_3 respectively. We define a parameter $\xi_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'}$ which has value one with possible configuration of i, j and l, otherwise zero. We count the number of $\xi_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} = 1$ and denote by m where $m < pn^2 \sum_{i'j'} k_{i'j'}$. We list $m + n^2 + pn^2 + a_1 + a_2 + a_3 + 1$ affinely independent points for Q_6 and thus, $dim(Q_6) = m + n^2 + pn^2 + a_1 + a_2 + a_3$.

$(\boldsymbol{e_f},\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{0}),$ $\quad f = 1,\ldots,m$

$(\mathbf{0},\mathbf{0},\boldsymbol{e_s h},\mathbf{0},\mathbf{0},\mathbf{0}),$ $\quad s = 1,\ldots,n^2$

$(\boldsymbol{e_f},\boldsymbol{e_s},\boldsymbol{e_{sh}},\mathbf{0},\mathbf{0},\mathbf{0}),$ $\quad f = 1,\ldots,m,\ s = 1,\ldots,n^2,\ h = 1,\ldots,p$

$(\mathbf{1},\mathbf{0},\mathbf{0},\boldsymbol{e_i},\mathbf{0},\mathbf{0}),$ $\quad i = 1,\ldots,a_1$

$(\mathbf{0},\mathbf{0},\mathbf{1},\mathbf{0},\boldsymbol{e_j},\mathbf{0}),$ $\quad j = 1,\ldots,a_2$

$(\mathbf{0},\mathbf{1},\mathbf{0},,,\boldsymbol{e_l}),$ $\quad l = 1,\ldots,a_3$

$(\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{0},)$

The set of $\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} + \gamma_{i'j'} - \psi_{(i_p,j_p,l_p),i'j'} \leq 1$ inequalities are facet defining inequalities. Since we can find $m+n^2+pn^2+a_1+a_2+a_3$ affinely independent points satisfying constraint (5.7), as follows

$(\boldsymbol{e_f},\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{0})$, $f = 1,\ldots,m$

$(\mathbf{0},\mathbf{0},\boldsymbol{e_{sh}},\mathbf{0},\mathbf{0},\mathbf{0})$, $s = 1,\ldots,n^2$

$(\boldsymbol{e_f},\boldsymbol{e_s},\boldsymbol{e_{sh}},\mathbf{0},\mathbf{0},\mathbf{0})$, $f = 1,\ldots,m,\ s = 1,\ldots,n^2,\ h = 1,\ldots,p$

$(\mathbf{0},\mathbf{1},\mathbf{0},\boldsymbol{e_i},\mathbf{0},\mathbf{0})$, $i = 1,\ldots,a_1$

$(\mathbf{0},\mathbf{1},\mathbf{0},\mathbf{0},\boldsymbol{e_j},\mathbf{0})$, $j = 1,\ldots,a_2$

$(\mathbf{0},\mathbf{1},\mathbf{0},\mathbf{0},\mathbf{0},\boldsymbol{e_l})$, $l = 1,\ldots,a_3$

where a_1, a_2 and a_3 denote the number of possible increments in directions x, y, and z respectively.

Lemma 3. *The set of inequalities $\psi_{(i_p,j_p,l_p),i'j'} \leq \gamma_{i'j'}$ are facet defining inequalities.*

Proof. Affinely independent points for the set of inequalities $\psi_{(i_p,j_p,l_p),i'j'} \leq \gamma_{i'j'}$ are as follows,

$(\mathbf{0},\mathbf{0},\mathbf{0},\boldsymbol{e_i},\mathbf{0},\mathbf{0})$, $i = 1,\ldots,a_1$

$(\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{0},\boldsymbol{e_j},\mathbf{0})$, $j = 1,\ldots,a_2$

$(\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{0},\boldsymbol{e_l})$, $l = 1,\ldots,a_3$

$(\boldsymbol{e_f},\mathbf{1},\mathbf{1},\mathbf{0},\mathbf{0},\mathbf{0})$, $f = 1,\ldots,m$

$(\mathbf{0},\boldsymbol{e_s},\boldsymbol{e_{sh}},\mathbf{0},\mathbf{0},\mathbf{0})$, $s = 1,\ldots,n^2,\ h = 1,\ldots,p$

□

Lemma 4. *The set of inequalities $\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \leq x_i$, $\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \leq y_j$, and $\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \leq z_l$ are facet defining inequalities for Q_6.*

Proof. We show that there exist $m+n^2+pn^2+a_1+a_2+a_3$ affinely independent points satisfying each of these inequalities, so they are facet defining inequalities. For $\beta_{(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \leq x_i$, these affinely independent points are as follows,

$(\boldsymbol{e_f},\mathbf{0},\mathbf{0},\mathbf{1},\mathbf{1},\mathbf{1}),$ $f=1,\ldots,m$

$(\mathbf{0},\boldsymbol{e_s},\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{0}),$ $s=1,\ldots,n^2$

$(\mathbf{0},\boldsymbol{e_s},\boldsymbol{e_{sh}},\mathbf{0},\mathbf{0},\mathbf{0}),$ $s=1,\ldots,n^2,\ h=1,\ldots,p$

$(\mathbf{0},\mathbf{0},\mathbf{0},\boldsymbol{e_i},\mathbf{0},\mathbf{0}),$ $i=1,\ldots,a_1$

$(\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{0},\boldsymbol{e_j},\mathbf{0}),$ $j=1,\ldots,a_2$

$(\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{0},\mathbf{0},\boldsymbol{e_l}),$ $l=1,\ldots,a_3$

□

3.4.3 Energy Level Inequalities

Consider the polyhedron Q_8 where

$$
\begin{aligned}
Q_8 = conv\{&(E_0,\ldots,E_{l_{max}},\psi_{1,i'j'},\ldots,\psi_{p,i'j'}) \in \mathbb{R}^{l_{max}n^2} \times \mathbb{B}^{pn^2}\text{:}\\
&E_{l_p-1,i'j'} \leq E_{l_p,i'j'} + M \sum_{i_p \in VP} \sum_{j_p \in VP} \psi_{(i_p,j_p,l_p-1),i'j'},\\
&E_{l_p,i'j'} \geq E_{l_p-1,i'j'} e^{-\mu_{(i_p,j_p,l_p-1)}d} - M(1-\psi_{(i_p,j_p,l_p-1),i'j'}),\\
&E_{l_p,i'j'} \leq E_{l_p-1,i'j'} e^{-\mu_{(i_p,j_p,l_p-1)}d} + M(1-\psi_{(i_p,j_p,l_p-1),i'j'}),\\
&\sum_{i_p \in VP} \sum_{j_p \in VP} \psi_{(i_p,j_p,l_p),i'j'} \leq 1,\ \forall\, i',j' = 1,\ldots,n,\ (i_p,j_p,l_p) \in VP\}
\end{aligned}
$$

where $dim(Q_8) \leq l_{max}n^2 + pn^2$. We can find a total of $p_l n^2 + pn^2 + 1$ affinely independent points for the vector of variables $(E_0,\ldots,E_{l_{max}},\psi_{1,i'j'},\ldots,\psi_{p,i'j'})$ as follows,

$(\boldsymbol{E'},\ldots,\boldsymbol{E'},\mathbf{0}),$

$(\boldsymbol{E'},\ldots,\boldsymbol{E'e^{-\mu_{(i_p,j_p,l_p-1)}d}},\ldots,\boldsymbol{E'e^{-\mu_{(i_p,j_p,l_p-1)}d}},\mathbf{0},\ldots,e_j,\ldots,\mathbf{0}),$ $j=1,\ldots,p$

$(\mathbf{0},\ldots,\boldsymbol{E'}e_i,\ldots,\mathbf{0},\mathbf{0},\ldots,\mathbf{0}),$ $i=1,\ldots,l_{max}$

where E' can be any real value within the range $E_{l_{max}} \leq E' \leq E_0$.

Lemma 5.

$$\frac{Me^{\mu_{(i_p,j_p,l_p-1)^d}}}{e^{\mu_{(i_p,j_p,l_p-1)^d}}-1}\psi_{(i_p,j_p,l_p-1),i'j'} \leq \lfloor \frac{Me^{\mu_{(i_p,j_p,l_p-1)^d}}}{e^{\mu_{(i_p,j_p,l_p-1)^d}}-1} \rfloor + \frac{E_{l_p,i'j'}}{1-\frac{Me^{\mu_{(i_p,j_p,l_p-1)^d}}}{e^{\mu_{(i_p,j_p,l_p-1)^d}}-1}+\lfloor \frac{Me^{\mu_{(i_p,j_p,l_p-1)^d}}}{e^{\mu_{(i_p,j_p,l_p-1)^d}}-1} \rfloor} \tag{3.67}$$

is a valid cut for Q_8.

Proof. A mixed-integer rounding cut suggested by reference [27] derives a cut including integer and continuous variables as follows,

If $X = \{(x,y) \in \mathbb{R}^1_+ \times \mathbb{Z}^1 : y \leq b + x\}$, and $f = b - \lfloor b \rfloor \geq 0$, the

$$y \leq \lfloor b \rfloor + \frac{x}{1-f} \tag{3.68}$$

is valid for X.

If we multiply both sides of the inequality (5.15) by $e^{\mu_{(i_p,j_p,l_p-1)^d}}$, and sum up the resulted inequality and inequality (5.18), we have

$$e^{\mu_{(i_p,j_p,l_p)^d}} E_{l_p-1,i'j'} \geq E_{l_p,i'j'} - Me^{\mu_{(i_p,j_p,l_p)^d}} + Me^{\mu_{(i_p,j_p,l_p)^d}} \psi_{(i_p,j_p,l_p-1),i'j'} \tag{3.69}$$

We divide both sides of the inequality (3.69) by $(e^{\mu_{(i_p,j_p,l_p)^d}} - 1)$, then the value of f for cut (3.68) is

$$f = b - \lfloor b \rfloor = \frac{Me^{\mu_{(i_p,j_p,l_p)^d}}}{Me^{\mu_{(i_p,j_p,l_p)^d}}-1} - \lfloor \frac{Me^{\mu_{(i_p,j_p,l_p)^d}}}{Me^{\mu_{(i_p,j_p,l_p)^d}}-1} \rfloor \tag{3.70}$$

Thus the mixed-integer cut is

$$\frac{Me^{\mu_{(i_p,j_p,l_p)^d}}}{Me^{\mu_{(i_p,j_p,l_p)^d}}-1}\psi_{(i_p,j_p,l_p-1),i'j'} \leq \lfloor \frac{Me^{\mu_{(i_p,j_p,l_p)^d}}}{Me^{\mu_{(i_p,j_p,l_p)^d}}-1} \rfloor + \frac{E_{l_p-1,i'j'}}{1-f} \tag{3.71}$$

and f is calculated in (3.70), and the result follows. □

3.4.4 Image Quality

We define polyhedron Q_9 as follows,

$Q_9 = conv\{(\boldsymbol{\tau}, \boldsymbol{\alpha}, \boldsymbol{E_{SD}}, \boldsymbol{\psi}, \boldsymbol{\gamma}) \in \mathbb{R}^{n^2} \times \mathbb{R}^{n^2} \times \mathbb{R}^{n^2} \times \mathbb{B}^{pn^2} \times \mathbb{B}^{n^2}$:

$K_1'' \sum_{i'=1}^{n}\sum_{j'=1}^{n} \alpha_{i'j'} \leq \sum_{i'=1}^{n}\sum_{j'=1}^{n} \tau_{i'j'},\ \sum_{i'=1}^{n}\sum_{j'=1}^{n} \tau_{i'j'} \leq K_2'' \sum_{i'=1}^{n}\sum_{j'=1}^{n} \alpha_{i'j'},$
$\tau_{i'j'} \geq E_{SD,i'j'} - M\,(1-\alpha_{i'j'}),\ \tau_{i'j'} \leq E_{SD,i'j'},\ \tau_{i'j'} \leq M\alpha_{i'j'},$
$\alpha_{i'j'} \geq \psi_{(i_p,j_p,l_p),i'j'}, \alpha_{i'j'} \leq \sum_{i_p,j_p,l_p \in ROI} \psi_{(i_p,j_p,l_p),i'j'},$
$\psi_{(i_p,j_p,l_p),i'j'} \leq \alpha_{i'j'},\ \forall\, i', j' = 1,\ldots,n,$
$(i_p,j_p,l_p) \in ROI\}.$

We consider constraint (5.8) invovled in this polyhedron to use for deriving strong valid inequalities. $dim(Q_9) = 4n^2 + pn^2$ and the vector of the variables is $(\boldsymbol{\tau}, \boldsymbol{\alpha}, \boldsymbol{E_{SD}}, \boldsymbol{\psi}, \boldsymbol{\gamma})$.

Lemma 6. $\sum_{i'=1}^{n}\sum_{j'=1}^{n} \alpha_{i'j'} \leq \left\lfloor \frac{Mn^2}{M-K_2''} \right\rfloor$ *is a valid inequality for* Q_9.

Proof. Constraint (5.22) ensures $\sum_{i'}\sum_{j'} \tau_{i'j'} \leq K_2'' \alpha_{i'j'}$, so by substituting in inequality (5.25), we have,

$$K_2'' \sum_{i'}\sum_{j'} \alpha_{i'j'} \geq \sum_{i'}\sum_{j'} E_{SD,i'j'} - M(n^2 - \sum_{i'}\sum_{j'} \alpha_{i'j'}) \tag{3.72}$$

applying C-G cut, results the following ,

$$\sum_{i'}\sum_{j'} \alpha_{i'j'} \leq \left\lfloor \frac{Mn^2}{M-K_2''} \right\rfloor \tag{3.73}$$

□

Lemma 7. $\sum_{i'=1}^{n}\sum_{j'=1}^{n} \alpha_{i'j'} \leq \left\lfloor \frac{Mn^2}{M-K_2''+K_1''} \right\rfloor$ *is a valid inequality for* Q_9.

Proof. If we sum up inequalities (5.23) $\forall\, i',\ j'$, we have $\sum_{i'}\sum_{j'} \tau_{i'j'} - \sum_{i'}\sum_{j'} E_{SD,i'j'} \leq 0$.
Constraints (5.21) and (5.22) ensures $K_1'' \sum_{i'}\sum_{j'} \alpha_{i'j'} - \sum_{i'}\sum_{j'} \tau_{i'j'} \leq 0$ and $\sum_{i'}\sum_{j'} \tau_{i'j'} - K_2'' \alpha_{i'j'} \leq 0$ respectively. Also if we sum inequalities (5.25) $\forall\, i',\ j'$, we have following ,

$$\sum_{i'}\sum_{j'} \tau_{i'j'} \geq \sum_{i'}\sum_{j'} E_{SD,i'j'} - M(n^2 - \sum_{i'}\sum_{j'} \alpha_{i'j'}) \tag{3.74}$$

By summing up all four inequalities mentioned above, and using C-G cut,

$$\sum_{i'}\sum_{j'} \alpha_{i'j'} \leq \left\lfloor \frac{Mn^2}{M-K_2''+K_1''} \right\rfloor \tag{3.75}$$

is attained where by comparing the denominators of (3.75) and (3.73), it is observed that (3.75) has a larger denominator than (3.73), and therefore (3.75) is stronger. □

A tighter lower bound for $\sum \alpha_{i'j'}$ can be given by solving the original problem regarding $min\{\sum \alpha_{i'j'}\}$, and similar to that we can achieve a strong upper bound applying $max\{\sum \alpha_{i'j'}\}$ as the objective function.

3.4.5 Big M Analysis

The big M used in the MIP formulation presented in section 3.3.3 needs to be determined since a very large value for the M increases the space between the feasible region of the MIP and LP relaxation. This will result in the larger optimality gaps and slows down the algorithm performance.

From inequality (5.13), we know that M is only effective in this formula when $\psi_{(i_p,j_p,l_p),i'j'} = 0$, and we can write,

$$M \geq \frac{E_{l_p,i'j'} - \Gamma_{(i_p,j_p,l_p),i'j'}}{1 - \psi_{(i_p,j_p,l_p),i'j'}}, \qquad \forall (i_p,j_p,l_p) \in C_o,\ i',\ j' = 1,\ldots,n \tag{3.76}$$

Since inequality (3.76) is only valid when $\psi_{(i_p,j_p,l_p),i'j'} = 0$, we can write it as follows,

$$M \geq E_{l_p,i'j'} - \Gamma_{(i_p,j_p,l_p),i'j'}, \qquad \forall (i_p,j_p,l_p) \in C_o,\ i',\ j' = 1,\ldots,n \tag{3.77}$$

having $\psi_{(i_p,j_p,l_p),i'j'} = 0$ implies that $\Gamma_{(i_p,j_p,l_p),i'j'} = 0$, and we have,

$$M \geq E_{l_p,i'j'}, \qquad \forall (i_p,j_p,l_p) \in C_o,\ i',\ j' = 1,\ldots,n \tag{3.78}$$

From the set of inequalities (3.78), it implies that $M \geq max_{\forall\, l_p \in VP, i', j' = 1,\ldots,n}\{E_{l_p,i'j'}\}$.

Since $E_{l_p,i'j'} \leq E_0,\ \forall\ l_p \in C_o,\ i',\ j' = 1,\ldots,n$, therefore, $M \geq E_0$. By selecting E_0 as objective function for the original problem and maximizing the objective function, the optimal solution is a bound for big M. However, it can be time consuming a comlicated optimization problem by itself. Thus, we imply a practical bound for the cases that theoretical bound is a better bound but not an easy problem to solve in a reasonable time.

3.5 Computational Results

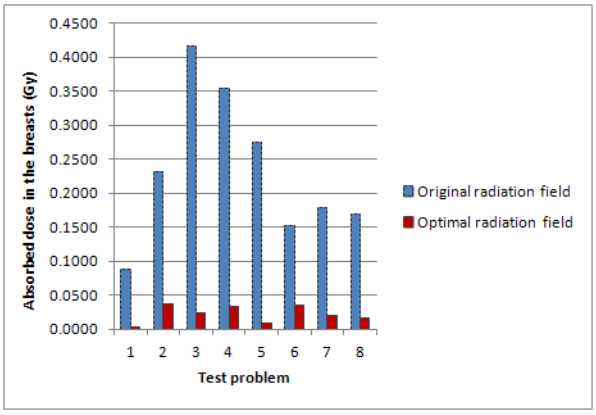
The problem formulation in section 3.3.3 is a large-scale MIP. Even in small instances with only 3 feasible geometry increments, there are about 80,000 variables and 150,000 constraints. The IBM ILOG CPLEX11.0 and Concert25 library packages are used to solve the problem. The algorithm is implemented by a desktop computer having Intel(R) Xeon(R) CPU and 4GB of RAM. We solve instances of this problem to compare the computational results before and after adding cuts derived in section 3.4. We added these cuts in terms of lazy cuts to the branch and cut algorithm. Lazy cuts are used by the methodologies built in the CPLEX and Concert solver.

The CT image and receptor image are discretized into 16×16 pseudo pixels. Each CT slice is 2mm and we merged 7 slices to have the same size of voxels ($14\text{mm} \times 14.08 \times 14.08$). Current setup enables handling the large number of binary variables (89,000) and constraints (187,000). Total number of the lazy cuts is 1878. Simple addition of the cuts to the original problem cannot be handled by the solver because of the memory limitation.

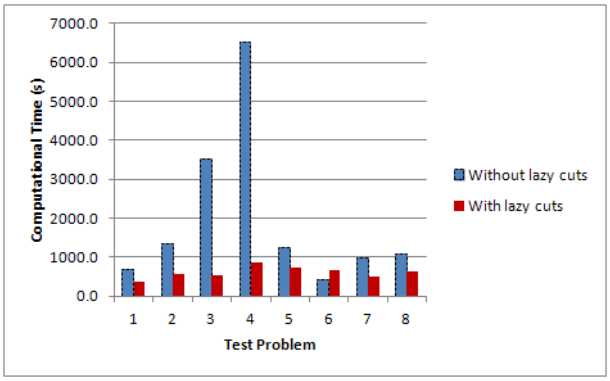
Figure 3.4(a) illustrates the absorbed dose (Gy) in the body using the original field of radiation and optimal field of radiation. It shows that by using our proposed method we can reduce the absorbed dose in the body significantly. Figure 3.4(b) compares the computational time (sec) before adding lazy cuts to CPLEX and after adding lazy cuts. It shows significant reduction of the computational time in all the test problems except one.

3.6 Improving Current Solution

Results after adding strong valid inequalities from lemmas (2)-(7) show the improvement in the computational time and optimality gap. To speed up the computation, we performed a preprocessing search algorithm to determine the feasible geometric increments for a



(a) Absorbed Dose Reduction



(b) Computational Time

Figure 3.4: The percentage of the improvement in the objective function

known set of primary and secondary angles. This method reduces the feasible increments using the geometry of the ROI. Adding this search algorithm in the preprocessing phase reduces the number of binary variables and corresponding constraints and speeds up the computational time. We adopt the set of commonly used C-arm angles as shown in Table 3.2. We solved test problems for different set of angles and present the computational time to get the optimal solutions after adding this algorithm. We summarize the number of variables, constraints, and computational times to reach the optimality by comparing with and without the lazy cuts added in Table 3.3. The last column indicates the percentage of the computational time improvement. With the cuts added to the algorithm, most of the

Primary Angle	Secondary Angle	Number of cine runs	Percentage of usage
-20	0	71	8%
-10	0	36	4%
0	0	537	65%
15	0	71	8%
20	0	40	5%
30	0	61	7%
15	10	20	2%
-20	20	12	1%

Table 3.2: Frequently used C-arm angles in vascular interventional procedures (total of 864 cine runs)

Primary angle	Secondary angle	No. of variables	No. of constraints	Time (s)		Improvement percentage
				No cut	Added cuts	
-20	20	81294	165426	672.4	376.4	44.02%
-20	0	89211	186921	1354.7	571.6	57.81%
-10	0	87764	152552	3514.9	550.1	84.35%
0	0	86019	154609	6514.8	856.7	86.85%
15	0	82179	167220	1241.3	724.1	41.67%
15	10	80080	159597	420.1	681.3	-62.18%
20	0	88725	172384	981.6	498.7	49.20%
30	0	81053	151238	1082.3	640.2	40.85%

Table 3.3: Comparison of computational time with and without lazy cuts for a set of angulations

computational times are reduced signifficantly to be within 15 minutes, which is the reasonable time frame to be prepared for the preparation of an procedure since multiple anglulations are to be computed prior to the exam. Note that in the angle (15, 10), the computational time is increased by adding additional cuts since CPLEX can solve the problem in about 7 minutes. So the searching and adding additional cuts in this case increases the computational time. When the voxel size of the phantom is refined to a smaller size, the size of the resulting formulation increases dramatically. The algorithm is even hard to load the model due to the memory limits. With additional memory or parallel computing capability, the solution can be further improve.

Primary angle	Secondary angle	No Collimator adjustment		Collimators adjusted		Dose reduction percentage
		Dose (Gy)	Field size (cm^2)	Dose (Gy)	Field size (cm^2)	
-20	20	0.0885	37.45	0.0036	9.95	95.93%
-20	0	0.2321	37.45	0.0380	15.65	83.63%
-10	0	0.4168	37.45	0.0242	5.27	94.19%
0	0	0.3548	37.45	0.0343	19.75	90.33%
15	0	0.2751	37.45	0.0084	5.85	96.95%
15	10	0.1524	37.45	0.0358	10.97	76.51%
20	0	0.1784	37.45	0.0198	18.29	88.90%
30	0	0.1689	37.45	0.0172	20.92	89.82%

Table 3.4: Comparison of comparison of the absorbed dose using the original radiation field and after adding the algorithm

To demonstrate the radiation reduction, we demonstrate the reduced absorbed in the critical organs by 1) not using collimators and 2) not placing the x-ray tube at the optimal location. In Table 3.4, we compare the absorbed dose in the breasts with the optimal solution from the model and the dose without using collimators, where the x-ray tube is located at the same location with the optimal solution. The dose reduction ranges from 75% to 95%.

It suggests that the use of collimators in such an FGI procedures in most of the angulation is an efficient tool to reduce the radiation dose in the critical organs. By examining the solutions, the collimators are positioned to ensure the field of view as small as possible to include the region of interest, which is not surprising to reduce the absorbed in the critical organs since it minimizes the radiation emission. In Table 3.5, we demonstrate the radiation reduction of the optimal X-ray tube location as comparing to alternative locations. Both sets of experiments test the radiation with and without collimator. These experiments demonstrate that radiation reduction is largely achieved by using collimators properly. The x-ray tube location, however, also reduces the absorbed dose even with collimators. Since our study is a prior exam preparation, with the

Primary angle	Secondary angle	X-ray tube			Absorbed dose (Gy)	
		x	y	z	No collimators adjustment	Adjusted collimators
15	0	2	7	9	0.2618	0.0106
		2	7	11	0.2751	0.0084
		2	7	13	0.2842	0.0093
30	0	3	13	4	0.1689	0.0172
		3	15	5	0.2022	0.0215
		3	17	9	0.2648	0.0236

Table 3.5: Comparison of the absorbed dose using the different geometry settings. Asterisk symbol indicates the optimal solution for the particular angulation

optimization model, the physician can be suggested that for each angulation, the best x-ray tube location and the placement of collimators, which minimize the field of size to include the region of interests. The optimal energy level will be translated back to mA level for the equipment setting.

Chapter 4

APPROXIMATE GRADIENT ALGORITHM

4.1 Introduction

Computational time and the limit of the memory are major challenges when solving large-scale Mixed Integer Linear Programming (MILP) problems [70]. In the branch and cut algorithm, the size of the branch and cut tree grows fast and it may even be hard to find a feasible solution. In section 3.3, we considered a MIP model which determines the geometry of the table, energy generated at the X-ray tube and size of field of radiation for every set of known angles (primary and secondary angles). The average number of binary variables for the problem is about 89,000 and the average number of constraints is 187,000. The size of the problem grows with refining the size of the voxels and increasing the number of beamlets going to the image. Besides, it is critical to consider the angulation (orientation) of the c-arm. From the experiments in Table 3.4, we can see the different angles may result in significant dose reduction. Including the angles as decision variables will lead to a Mixed Integer Nonlinear Programming (MINLP) model and add up to the complexity of the problem. Besides, the commercial implementations of branch-and-cut algorithm can only be applied to MINLP where the nonlinear continuous part is convex. The scale that can be solved within reasonable time is much smaller in compare to MILP. Using a fixed set of angles is a limitation for the problem and exploring continues geometry parameters may provide better solution quality. As mentioned in the literature, most of the studies consider a set of angles and solve the problem for each set and find the best angle [51, 52, 53]. Some others start with a subset of the commonly used angles and apply a greedy method or optimization problem to select the best angle [54].

In this chapter, we would like to explore a systematical way to determine the table location and angulations together to obtain a global optimal solution. We propose a

method to lower the complexity of the problem 3.3 by fixing certain decision variables and explore more angles. To improve the computation time and to be able to find the optimal set of angles, we use an approximate gradient approach. We partially relax the MIP formulation in section 3.3 to explore the continuous geometry and relax the limitation of the discrete geometry movements. In the sub routine, we treat the geometry of the table and angles as a set of known parameters. The size of the problem which was about 130,000 binary variables and 220,000 constraints is refined. Thus, number of binary variables are 77,000 and number of constraints are 4,800. Then the model determines the required energy and field size for each given set of geometry (angles and table increments). We specify a time limit to solve this MIP problem for each iteration of the approximate gradient algorithm.

4.2 Classical Gradient Rules and Literature Review

In gradient methods, if the multivariable function $F(x)$ is differentiable in a neighborhood of a point a, then $F(x)$ decreases if we move from a in the direction of the negative gradient of F at a, $-\nabla F(a)$. If and only if

$$b = a - \alpha \nabla F(a) \tag{4.1}$$

for $\alpha \to 0$ a small enough number, then $F(a) \geq F(b)$. Thus, we can start with a guess x_0 for a local minimum of F, and considers the sequence of $x_0, x_1, x_2, \ldots$ such that $x_{t+1} = x_t - \alpha_t \nabla F(x_t)$, $t \geq 0$. We have $F(x_0) \geq f(x_1) \geq F(x_2) \geq \ldots$ so hopefully the sequence (x_t) converges to the desired local minimum. Note that the value of the step size α is allowed to change at every iteration. The convergence to a local minimum is guaranteed if the function of F is Lipschitz continuous [71]. If the function is not differentiable, we use subgradient methods. A step size α_t should satisfy the following conditions to ensure the convergence of the algorithm [72],

$$\lim_{t\to\infty} \alpha_t = 0 \tag{4.2}$$

$$\sum_{t=1}^{\infty} \alpha_t = \infty \tag{4.3}$$

$$\alpha_t \geq 0 \quad \forall\, t \tag{4.4}$$

Using a diminishing step size and normalized gradients, ensures the convergence of the algorithm for the convex functions [73].

In this study, the objective function of the MIP problem introduced in section 3.3 is used to compute the gradient direction. It is comprised of binary variables, and it is neither a convex function nor a differentiable function. Thus, we cannot use any of gradient or subgradient methods directly. We use an approximate gradient method to approximately calculate the gradient vector. The vector used for the iteration $t+1$ of the approximate gradient method is defined as

$$\vec{\omega}_{t+1} = \vec{\omega}_t - \alpha_t\, \vec{g}_t \tag{4.5}$$

Since we do not have the convexity/differentiable function for the MIP problem, we approximate the gradient vector using the objective function of the MIP problem. The gradient vector is

$$\vec{g'}_t = (g_t'^1, \ldots, g_t'^5) \tag{4.6}$$

where

$$g_t'^i = (f(\vec{\omega}_t + \Delta_i \vec{e}_i) - f(\vec{\omega}_t))/\Delta_i \tag{4.7}$$

and

$$\vec{g}_t = \vec{g'}_t / ||\vec{g'}_t||_2 \tag{4.8}$$

The value of Δ_i is determined based on the corresponding increment for the vector of $\vec{\omega}_t$ at iteration t and $\vec{e}_i$ is the eigenvector.

Since the objective function of the MIP problem is not a convex function, we stop the algorithm when we have,

$\lim_{t\to\infty} |f(\vec{\omega}_t) - f(\vec{\omega}_{t-1})| < \varepsilon$.

4.2.1 Gradient and Subgradient Methods in IMRT

Some studies use a modified subgradient algorithm for the nonconvex functions [74, 75, 76, 77].

There are studies which apply gradient and subgradient methods for the Intensity Modulated Radiation Therapy (IMRT) [78, 79, 80, 81, 82]. Gradient algorithms are the most commonly employed search methods in the routine optimization of IMRT plans [78]. Michalski et al. [79] present a novel formulation of a dose volume constraint satisfaction search for the discretized radiation therapy model. Problem is solved using the simultaneous version of the cyclic subgradient projection algorithm which is a deterministic iterative method designed for solving the convex feasibility problems.

Censor et al. [80] propose an optimization algorithm which minimizes a weighted proximity function that measures the sum of the squares of the distances to the dose constraint sets. In most IMRT inverse problems, it is not possible to satisfy all of the constraints. Thus, a method is find plans that satisfy the constraints as much as possible. To do this, they introduce proximity functions that measure the distance of a point to the set. The goal is to minimize a weighted sum of the proximity functions. A projected gradient method is used to optimize the problem. Projection algorithms are frequently used to find a point that belongs to the intersection of closed convex sets. A projection onto a set is the point within that set that is closest to the current point. In projection algorithms, a point is repeatedly projected onto each set, according to some algorithmic rules, until a point in all of the sets is found.

Zhang et al. [78] compare the relative speed of different gradient algorithms, to investigate the strategies for accelerating the optimization process, to assess the validity of these strategies, and to study the convergence properties of these algorithms for dose-volume and biological objective functions. They implement Newton's, conjugate gradient (CG), and the steepest decent (SD) algorithms for dose-volume- and equivalent uniform dose-based objective functions. They approximate the second derivative matrix (Hessian) by its diagonal. They also implement the standard SD algorithm and the CG algorithm with "line minimization". A variation of the CG algorithm which is called "scaled conjugate gradient" (SCG) is investigated. A hybrid optimization strategy is evaluated in which approximations to calculated dose distributions are used during most of the iterations. Newton's method outperforms other algorithms in terms of computational time. The SCG algorithm can speed up the standard CG algorithm by at least a factor of 2. Having the same initial conditions, all algorithms converge essentially to the same plan.

A technical note introduces a local search technique for the beam angle optimization [81]. For a set of know angles, $f(\theta)$ indicates the objective cost of the optimized IMRT plan for the set of θ. Gradient of the function with respect to θ is used for the local search. A combination of the gradient method and duality theory is used to proceed.

Another study uses a hybrid approach for the beam angle optimization problem which automatically selects the beam angles and computes the beam intensities [82]. This method applies a Simulated Annealing procedure (SA) with a Gradient Descent method (GD). It consists of an iterative approach which alternates few steps of GD for quickly finding a local minimum, with few steps of SA for jumping out of the local minima then it starts to search in a different part of the solution space.

We discussed different studies in IMRT which use gradient or subgradient methods to solve the different problems in IMRT. Most of these studies take the angles as the input vector for the subgradient method and solve a sub problem which is either MIP or LP. We use an approximate gradient method for the MIP problem discussed in chapter 3 and solve a subroutine which treats the geometry parameters and angles as continues parameters.

4.3 Mathematical Model

The model developed in chapter 3, included the geometry of the table as discrete decision variables. Since the image originally is discretized to the pixels, we used a discretized geometry which is also practical in clinical use. Also we used a set of commonly used angles to run the problem. In this chapter, we would like to investigate continues geometry of the table and angles to further improve the computational time and solution quality. We propose a partially relaxed formulation of the MIP problem given in section 3.3.3 where the constraints corresponding to the discrete geometry are relaxed. The approximate gradient algorithm allows taking continuous positions and using fixed geometry in the formulation. The model is solved at each iteration of the approximate gradient algorithm as a subroutine within a specified time limit. We refer to the constraints which are the same as constraints in model 3.3.3 by the reference number to those constraints.

We define parameter $\beta_{(i_p,j_p,l_p),i'j'}$ as follows,

$$\beta_{(i_p,j_p,l_p),i'j'} = \begin{cases} 1 & \text{if beam } i'j' \text{ hits the body voxel of } (i_p, j_p, l_p) \\ 0 & \text{otherwise} \end{cases}$$

Other parameters and variables are not mentioned in this section, are defined in chapter 3.

$$f(\omega_t) = \textbf{Min} \sum_{V_{(i_p,j_p,l_p)} \in C_o} dose_{V_{(i_p,j_p,l_p)}} =$$

$$\sum_{(i_p,j_p,l_p):V_{(i_p,j_p,l_p)} \in C_o} \sum_{i'=1}^{n} \sum_{j'=1}^{n} \frac{\Gamma_{(i_p,j_p,l_p),i'j'}(1 - e^{-\mu_{(i_p,j_p,l_p)}d})}{m_{V_{(i_p,j_p,l_p)}}} \tag{4.9}$$

Subject to

- Constraints (3.4), (3.5), (3.6), (3.7), (3.8), (5.13), (5.14), (5.15), (5.16), (5.17), (5.18), (5.19), (5.20), (3.26), (5.21), (5.22), (5.23), (5.24), (5.25), (5.26), (5.27)

- Constraint to linearize $\psi_{(i_p,j_p,l_p),i'j'}$ equation : $\lfloor \frac{\beta_{(i_p,j_p,l_p),i'j'} + \gamma_{i'j'}}{2} \rfloor$

$$\psi_{(i_p,j_p,l_p),i'j'} \geq \beta_{(i_p,j_p,l_p),i'j'} + \gamma_{i'j'} - 1 \quad \forall\, (i_p,j_p,l_p) \in VP,\ i',\ j' = 1,..,n \tag{4.10}$$

$$\psi_{(i_p,j_p,l_p),i'j'} \leq \gamma_{i'j'} \quad \forall\, (i_p,j_p,l_p) \in VP,\ i',\ j' = 1,..,n \tag{4.11}$$

$$\psi_{(i_p,j_p,l_p),i'j'} \leq \beta_{(i_p,j_p,l_p),i'j'} \quad \forall (i_p,j_p,l_p),\in VP,\ i',\ j' = 1,..,n \tag{4.12}$$

- Variable range

$$\psi_{(i_p,j_p,l_p),i'j'} \in \{0,1\} \tag{4.13}$$

$$\Gamma_{l_p,i'j'} \geq 0 \tag{4.14}$$

$$\gamma_{i'j'} \in \{0,1\} \tag{4.15}$$

This model presents a formulation for a set of known geometry parameters and angles. It minimizes the dose deposited in the critical organs and determines the required energy and field size with respect to the image quality. Constraint 4.10 ensures that if a beam exists and hits the body voxel, the binary variable of $\psi_{(i_p,j_p,l_p),i'j'}$ is taking a value of one. Constraints 4.11 and 4.12 enforce the value of $\psi_{(i_p,j_p,l_p),i'j'}$ be zero when beam $i'j'$ does not intercept with voxel (i_p,j_p,l_p).

4.4 Algorithm

Here we first explain some common rules of the gradient method and modifications made in our approximate gradient algorithm.

We start with a feasible solution which is selected randomly from a set of known feasible solutions. We run the approximate gradient algorithm for each of the feasible solutions in the set for a few runs (3-5) and whichever achieved the lowest value in terms of the objective function is the one we pick as start point. Then, we use the best given feasible solution in each iteration to proceed to the next iteration of the gradient algorithm. A vector $\vec{\omega}_t$ is updated at each iteration of the gradient algorithm which consists of the primary and secondary angles and table increments in x, y and z directions.

We use norm-2 for normalizing the gradient direction. For any vector x, the norm-2 is defined as follows,

$||x||_2 := \sqrt{\sum_i (x_i^2)}$

4.4.1 Notations

t: iteration index

T: maximum number of iterations

$f(\vec{\omega}_t)$: objective function iteration t for the MIP problem 4.3 by using the input vector of $\vec{\omega}_t$

$\vec{\omega}_{best}$: the best local optimal point achieved in the iterations

$f(\vec{\omega}_{best})$: the best objective function achieved within iterations

$\vec{\omega}_0 = (\theta_{1,0}, \theta_{2,0}, x_0, y_0, z_0)$: starting point, where $\theta_{1,0}$ and $\theta_{2,0}$ indicate the primary and secondary angles and x_0, y_0 and z_0 indicate the table increments respectively

$\vec{\omega}_t$: vector of geometry in iteration t

M: large value, comparable to the maximum energy

$\xi_{[v][i][j]}$: indicates the intercepted voxel v of body with beam i and j

$\vec{g}_t$: gradient vector

α: multiplier

h: number of infeasible consequitive iterations

$\xi_{[v][i][j]}$: indicator for the voxel v of the body intercepted with beam i and j

h: number of infeasible consecutive iterations

h_t: number of infeasible consecutive iterations at iteration t

p: indicated number of iterations to average the objective function value

v: the lower bound set for the objective function which is acceptable in practice

4.4.2 Procedural Rules of the Algorithm

We use an approximate gradient direction $\vec{g}_t$ which is calculated as equation 4.8 using $g_t'^i$ values calculated by equation 4.7. We set $\alpha_t = \frac{1}{t}$.

We update the value of $\vec{\omega}_{t+1}$ based on the formula given in equation 4.5 and approximate gradient calculated in equation 4.8.

In some steps of the approximate gradient method, the step size may result in infeasible problem. In order to jump out of the infeasible solutions, we use a Fibonacci search to reduce the step size and go back to a feasible solution (step 5.1.). We compute norm and gradient for the two consequitive iterations. Then we update the vector of ω_t where the multiplier of α_t starts with $\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \frac{1}{5}, \ldots$ based on the Fibonacci series to find a feasible solution.

We also use another method to reject feasible solutions with a large gap from the best feasible solution. In this method, after a specified number of iterations (p iterations), we use the average objective function to reject the feasible solutions larger than this value. We continue using this method at every $(i+1)p$ iterations of the problem, where p is the

interval and i is the indicator of the corresponding interval. We continue using this method until we meet the stoping conditions as follows,

4.4.3 Stopping rules

- Number of iterations $\geq T$ (here we used $T = 1000$ iterations)
- Achieving two consecutive objective functions with a $|f(\omega_t) - f(\omega_{t-1})| < \varepsilon$
- Achieving a value for the objective function which is sufficiently small in practice (here we set Δ for as an acceptable lower bound for this objective function comparison)

4.4.4 Gradient Approximation Algorithm

A.1 : Steps of Gradient Approximation Algorithm

---

1. Initialize function $f(\vec{\omega}_0)$ with a large value of M. Set t = 0, $f(\omega_{best}) = \infty$, $f(\vec{\omega_{-1}}) = -M$. We also set $k = 0$ which indicates the frequency of lowering the bound for the objective function. Read initial vector of $\vec{\omega}_0 = (\theta_{1,0}, \theta_{2,0}, x_0, y_0, z_0)$ as input.
2. Prepare the input data to compute matrices of intersection of beams with body voxels, indices of the body voxels, indices of ROI and indices of the critical region with respect to the first voxel of the body using the vector of $\vec{\omega}_t$. Compute the values for the $\xi_{[v][i][j]}$, $\forall\, v \in Voxels\ of\ Body, i,\ j = 1 \ldots n$ which indicates the voxel v of the body intercepted with beam i and j.
3. If $t \leq T$, go to step 3, otherwise stop.

4. Solve MIP model of formulation 4.3 and calculate $f(\vec{\omega}_t)$ for a time limit of 500 seconds.

5. If MIP is feasible and has found a feasible solution within 500 seconds, go to step 5.2., otherwise go to step 5.1..

   5.1. Modify the step size α_t by Fibonacci search. Calculate the vector of $\vec{g'}_t$, $\vec{\omega}_{t+1}$ using equations 4.6, 4.7, and 4.5, and set $\vec{\omega}_t = \vec{\omega}_{t+1}$. Go to step 4.

   5.2. Solve MIP for $f(\vec{\omega}_t + \Delta_i \vec{e}_i)$, $\forall\, i = 1, \ldots, 5$.

   5.3. If the MIP problem $\forall\, i$ is feasible, go to step 5.4., otherwise go to step 5.5..

   5.4. If number of iterations t is equal to $kp + p + 1$, then go to step 5.6., otherwise go to step 5.8..

   5.5. Modify the value of the Δ_i, $\forall\, i$ where $f(\vec{\omega}_t + \Delta_i \vec{e}_i)$ is infeasible. Go to step 5.2..

   5.6. Set $\delta = \sum_{t=1}^{(k+1)p} f(\vec{\omega}_t)/((k+1)p)$. Set $k = k+1$.

   5.7. If $f(\vec{\omega}_t) \leq \delta$, go to step 5.8., otherwise go to step 5.11.3..

   5.8. Calculate $\alpha_t = 1/t$.

   5.9. Calculate $\vec{g'}_t$ and $\vec{\omega}_{t+1}$ using equations 4.6, 4.7, and 4.5.

   5.10. If $f(\vec{\omega}_t) - f(\vec{\omega}_{t-1}) \leq \varepsilon$, stop, otherwise go to step 5.11..

   5.11. If $f(\vec{\omega}_t) \leq f(\vec{\omega}_{best})$, then go to step 5.11.1., otherwise go to step 5.11.3..

      5.11.1. Set $f(\vec{\omega}_{best}) = f(\vec{\omega}_t)$, and $t = t+1$.

      5.11.2. If $f(\vec{\omega}_{best}) \leq \nu$, stop, otherwise go to step 4.

      5.11.3. Set $\vec{\omega}_t = \vec{\omega}_{best}$. Go to step 5.2..

Figure 4.1 demonstrates the flow chart for the approximate gradient algorithm. To ensure the image quality, all the voxels of the ROI should intercept with pencil beams. We check this condition in step 2 of the algorithm. If it is not feasible, we modify the step size. We make the step size smaller by using a Fibonacci search 5.1.. After each change, we calculate the updated ω_t and check if the problem is feasible or not. Anytime that we set $f(\omega_t)$ to M, we calculate the norm and gradient using the best (f_{best}) objective function achieved up to the current iteration 5.11.1..

4.5 Computation Results

We solve testing problems presented in section 3, Table 3.4 and compare the results from the computational aspects with results for the test problems solved using the branch-and-cut algorithm in section 3.5. The IBM ILOG CPLEX11.0 and Concert25 library packages are used to solve the MIP problem in 4.3. The runs are implemented by a desktop computer having Intel(R) Xeon(R) CPU and 4GB of RAM. We pick a feasible solution as a start point by randomly selecting from a set of known feasible solutions. We run the gradient algorithm for each of the available feasible solutions in the set for a few runs (3-5) and we pick the one achieving the lowest objective function. We set a time limit of 200 seconds and an optimality gap of the $\varepsilon < 0.0001$ to stop running the MIP model 4.3. Whichever of these condition is achieved first, the algorithm stops running the MIP model, updates the gradient, ω_t and continues. The approximate gradient algorithm stops using the rules given in section 4.4. The setting of the rules are: i)$f(\omega_t) - f(\omega_{t-1}) \leq \varepsilon$ where ε is set to 0.0001 for our experiments, ii) $f_{best} \leq \Delta$, where Δ is set to 0.005, and iii) maximum number of iterations T which is set to 1000 for this study.

We compared the values for the absorbed dose in the critical organs, using the approximate gradient algorithm and branch-and-cut in Table 4.1. Most of the test problems show a comparable objective value. The computational times (s) are compared

for using two algorithms in each test problem. The two last columns represent the computational results corresponding to each setting. The first one is the absorbed dose in the critical organ (Gy) and the next one is the computational time in seconds. We

Test Problem	Algorithm	Primary angle	Secondary angle	Increment Units			Computation	
				x	y	z	(Gy)	(s)
1	B&C	-20	20	-4	-8	1	0.0036	376.4
	gradient	-20	20.3	-5.8	-8.8	-2.7	0.0041	264.1
2	B&C	15	0	2	7	11	0.0084	724.1
	gradient	14.9	-0.3	2.1	8	12.6	0.0069	328.7
3	B&C	15	10	-2	10	15	0.0358	681.3
	gradient	14.9	10.4	-2.4	12	21.1	0.0037	174.6
4	B&C	30	0	3	13	4	0.0172	640.2
	gradient	29.7	-0.3	3.2	14.6	0.1	0.0164	279.5
5	B&C	-10	0	1	-6	15	0.0242	550.1
	gradient	-9.8	-0.3	2.1	-5.7	21.3	0.0086	308.7
6	B&C	20	0	-1	8	3	0.0198	498.7
	gradient	19.5	-0.6	3	8	-2	0.0023	643.8
7	B&C	0	0	0	1	28	0.0343	856.7
	gradient	0	-0.6	-0.1	-1	35	0.0277	218.3
8	B&C	-20	0	3	-11	7	0.0380	571.6
	gradient	-19.9	-0.5	0	-9.8	7.5	0.0213	221.4

Table 4.1: Comparison of the geometry parameters given by branch-and-cut (B&C) and updated by approximate gradient algorithm

compared the geometry parameters for the best objective function given by the approximate gradient algorithm with the best objective function given by the branch-and-cut algorithm in Table 4.1. The values for the gradient algorithm are continues but the results show that continues geometry parameters have close values to the discrete geometry. This indicates that approximate gradient algorithm provides comparable results to the branch-and-cut algorithm. Therefore, it can be useful if we refine the size of the problem for the smaller voxels that the commercial solvers have limitation for handling a large number of variables and constraints when the size of the branch-and-cut tree may become very large and the solver runs out of the memory.

Using approximate gradient algorithm with some modifications such as approximate gradient calculation in different iterations and applying rejection rules for the infeasible solutions and feasible solutions with large optimality gap provides comparable results in terms of the objective function value with previous modified branch-and-cut method. This method proceeds faster since there are fewer number of binary variables. Approximate gradient algorithm provides a better objective function in most of the settings and for one setting it shows a higher value which is still small enough for the objective function. It also provides the computational time for each algorithms which is smaller for most of the cases when an approximate gradient algorithm is used.

4.6 Summary

In chapter 3, we proposed a model which minimizes the absorbed dose in the critical organs with respect to the geometry in a discretized space using a set of known angles. The discretized geometry enables us to model problem as a MILP model and use branch-and-cut to find solutions. However; it also limits us to certain discrete locations which may bypass some opportunities to improve the objective function and explore the region. In this chapter, we investigated the use of an approximate gradient approach to have the capability of finding a set of angle and table geometry simultaneously which minimizes the absorbed dose in the critical organs with respect to the image quality. This approach uses a continuous setting for the geometry parameters and solves a sub routine MIP model which determines the energy and field of radiation. Using approximate gradient algorithm presents a combined method which is capable of searching in a continuous space and updates angles and locations simultaneously. Our results show that we can improve the computational time and solution quality using approximate gradient method although the global optimum is not guaranteed. By refining the grid size of the original MIP model presented in 3, the branch-and-cut method cannot handle the number

of variables and constraints. However, the approximate gradient method uses a refined MIP model which can be handled by branch-and-cut method for smaller grid size.

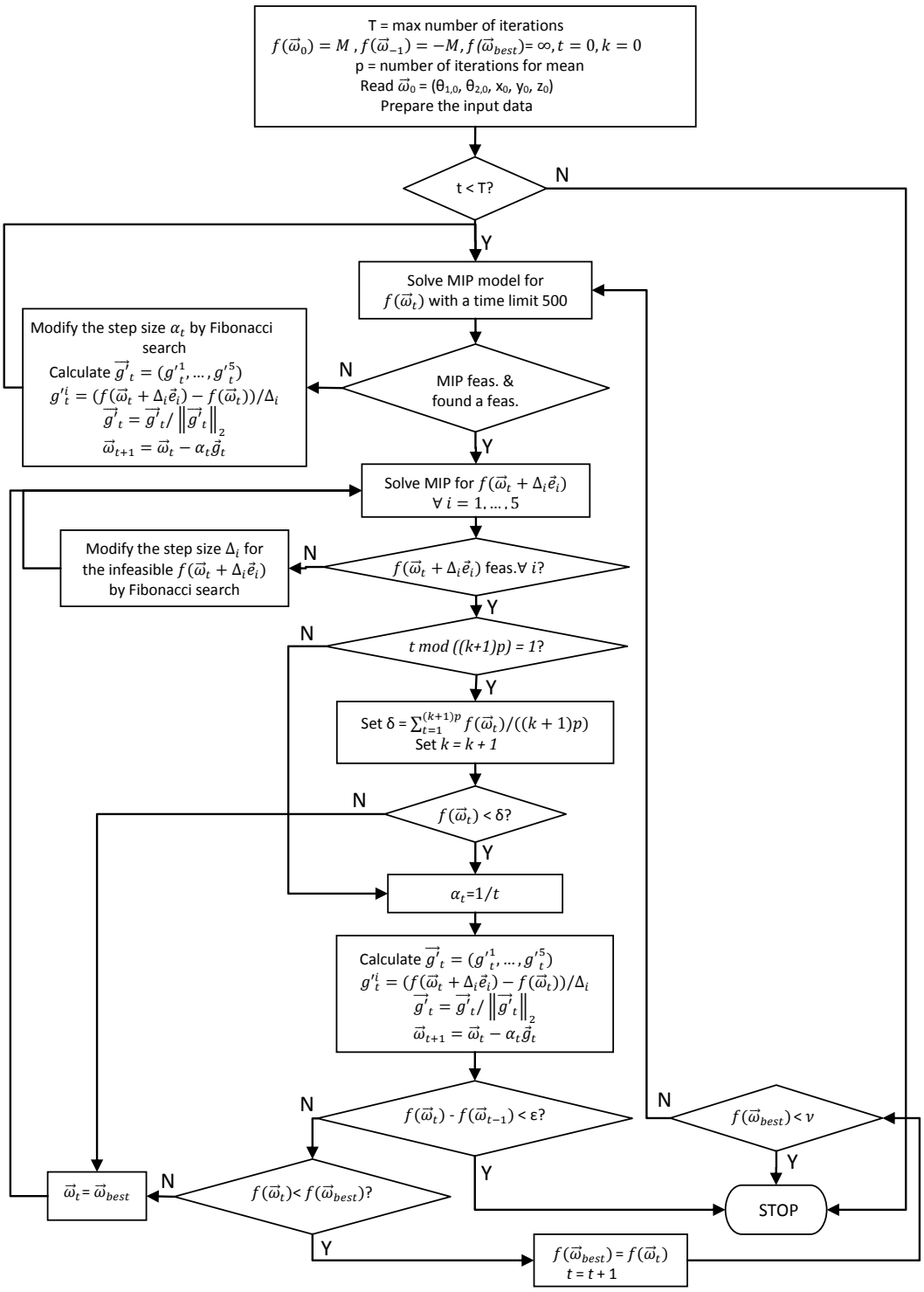


Figure 4.1: Flow chart for the approximate gradient algorithm

Chapter 5

ROBUST SOLUTION WITH INTRAFRACTIONAL MOTION

5.1 Introduction

In section 3, we discussed a deterministic model which minimizes the absorbed dose in the critical organs of the body with respect to the image quality. In this formulation, we assumed that patient position and relative position of the organs are deterministic. However, several sources of uncertainty exist in the radiation therapy and radiology. For example, when a patient is breathing, changes occur in the position of the internal organs relative to the beams; it results in uncertainty for the treatment planning. Other sources of uncertainty are imaging errors, physician error, intrafraction patient motion, registration errors, organ shrinkage and possible microscopic extensions of the tumor [83].

A major challenge in radiology and radiation therapy treatment planning is managing uncertainty, including uncertainty induced by intrafraction motion [84]. Breathing causes the change in the relative size and positions of some organs. There are some studies which address some possible uncertainties in IMRT [81, 83, 84, 85, 86].

Since the sources of uncertainties are not considered in our deterministic formulation, the absorbed dose calculated may be comprised of the errors caused by uncertainties. A deterministic model is solved for one setting of the input data scenarios and the solution is corresponding to one scenario. Such approaches that only evaluate decisions using one data scenario may fail. For other decision makers especially the medical decision makers would like to avoid the risk and are interested in having the information of any possible outcome. Thus a decision maker is not just looking for an optimal solution for a specific data scenario not even the most likely to occur scenario, but a decision which performs best across all scenarios. Robust decision is an approach to determine a solution that performs well across all scenarios. We aim to produce decisions

with reasonable objective values under any input data scenario to our optimization model over a pre-specified planning horizon. One possible criteria is *minimax* which minimizes the largest possible loss across the scenarios [87].

Here, we define two terms of "solution robust" and "model robust". "Solution robust" is used when a solution of the optimization model is robust with respect to the optimality and remains close to optimal for any input data scenario. If a solution remains almost feasible for the input data scenarios, it is called "model robust" [88].

A scenario based approach is usually used to represent the input data uncertainty to the decision models. In robust optimization, we do not need to have the probability distribution and since the distribution function may not be known we can not use stochastic programming. If we have the probability distribution, we can use stochastic programming to address uncertainty in the problems.

The mathematical model presented in chapter 3 is a deterministic model and does not address the uncertainties. In this chapter, we consider breathing as a source of uncertainty for the organ relative positions and size of the organs. We consider input data scenarios based on this source of uncertainties. Within the process of breathing the organs change continuously. Here, we only consider discrete scenarios which are like snap shots of this process. By adding up the number of scenarios, we are able to simulate this process closer to reality. We refine the mathematical model 3.3 so that it minimizes the maximum absorbed dose for the critical organs with respect to the image quality across all input data scenarios.

5.2 Literature Review

In this section, we introduce some definitions of a robust approach, and then we discuss studies in radiation therapy where some sources of uncertainty exist in the problem.

5.2.1 *Definition of the Robust Approach*

The robustness approach determines appropriate input data instances for the situation without assigning probabilities, and then attempts to find the decision which performs well in the worst case. A scenario based method can be used to represent the uncertainty of the input data in our decision model. Each scenario indicates a potential realization of the parameters in the decision model that we do not know the probability of occurrence. By using a scenario based approach, the decision maker can accommodate the effects of different factors on input data simultaneously [87].

We explain some notations and decision used in robust optimization and what we use in our problem. Let S be a set of input data scenarios over a specified horizon. Let X be the set of decision variables and D be the set of input data. The notation D^s denotes the instance of the input data corresponding to scenario s. F_s denotes the set of all feasible decisions for scenario s, and assume function $f(X,D^s)$ is used for the quality evaluation of decision $X \in F_s$. The optimal scenario X_s^* corresponding to input data D^s is the solution of the following deterministic optimization problem [87],

$$z^s = f(X_s^*, D^s) = min_{X \in F_s}\ f(X, D^s) \tag{5.1}$$

The decision X_A which minimizes the maximum total cost among all feasible solutions is called the absolute robust decision.

$$z_A = \max_{s \in S}\ f(X, D^s) = \min_{X \in \cap_{s \in S} F_s}\ \max_{s \in S}\ f(X, D^s) \tag{5.2}$$

These decisions are conservative and they are based on the worst case occurrence. Another robust criterion is Robust Deviation where the performance of the decision in each scenario is evaluated against the best possible decision for that scenario. The deviation of the performance of a scenario from the best performance is recorded for all

the scenarios. Relative Robustness is another criterion which is the percentage deviation from the optimal performance in a scenario and the robustness indicator is the worst observed percentage deviation from the optimal for the evaluated decision across all scenarios. The absolute robust criterion is more related to the decisions which are conservative and the major concern is to hedge against the worst possible happening.

The problem we are considering minimizes the absorbed dose for the critical organs with respect to the image quality. If we consider breathing as a source of uncertainty which results in the change of the relative positions of the organs, then we need to ensure that the absorbed dose in the critical organs is minimized in any of the known positions for the organs. Therefore, we use the absolute robust decision for our problem to analyze the worst case scenario [87].

5.2.2 Uncertainty in the Radiology and Radiation Therapy Problems

There are several uncertain parameters involved in radiology and radiation therapy. In this section, we introduce studies which discuss uncertainties involved in these procedures and reformulate the problem to address the uncertainties. Chan et al. [84] analyze the uncertainty in the probability mass function (pmf) which describes breathing motion in radiation therapy, and show how to incorporate it into the inverse planning optimization to produce a robust treatment plan. The focus of the analysis is on the case where there is no uncertainty in the objective function. A model of data uncertainty is introduced which describes the possible variations in the breathing pmf derived from breathing motion data. This study demonstrates the potential of using a robust optimization method in IMRT treatment planning to avoid overdosing healthy tissues while maintaining tumor coverage in the presence of uncertainty. It provides flexibility for the treatment planner to make suitable decisions regarding trade-offs of conflicting objectives.

Uncertainty resulted from the organ motions in IMRT is also discussed by study [85]. They propose a second-order cone programming model which is a probabilistic model for the IMRT inverse problem and show that under certain assumptions it is identical to a robust optimization model.

Robust optimization techniques are used to formulate an IMRT treatment planning problem with uncertainty in the dose matrices. This uncertainty is due to both dose calculation errors and interfraction positional uncertainty of tumor and organs. The original linear programming formulation becomes a second-order cone program after uncertainty is taken into the formulation. Dose matrices are composed of elements which describe the amount of radiation deposited by a unit weight (intensity) from a particular beamlet into a particular voxel in the treatment region. The dose matrix is calculated by techniques such as convolution/superposition and Monte Carlo simulation, given the properties of the treatment region and the beamlet source and orientation. These techniques provide approximate results, and hence are one source of uncertainty in the problem. Another source of uncertainty is change of the position in the internal tissues between treatment fractions. The probability distributions for both types of uncertainties are assumed to be known. The formulation is a second-order cone programming and is solved using sequential linear programming (interfractional motion) [83].

Bortfeld et al. [86] study the presence of the breathing motion for lung and liver tumors. They formulate a model of motion uncertainty using probability density functions that describe breathing motion which provides a robust formulation for IMRT. The model uses real patient data and the robustness of the resulting solutions is measured on a clinical lung example. The model minimizes the total dose delivered with respect to the adequate dose be received at the tumor.

We discussed studies which address uncertainty in dose computation resulted from intrafraction motion. They use probability functions to address the intrafraction motion.

5.3 Mathematical Model

Three critical elements of the robustness approach for decision making are as follows [87],

1. Using a Scenario Planning Approach to data uncertainty for the decision situation
2. Choice of appropriate Robustness Criteria
3. Development of a Decision Model (Robust Optimization)

A decision maker is required to identify predetermined elements of the environment, and formalize the perceived connection among events and factors that raises the uncertainty in the decision making environment. Scenario generation and robust decision making require the intuition of the environment.

5.3.1 Scenario Design

We construct our scenarios based on the intrafraction motion of the organs within breathing process. We use CT slices to indicate the positions of the organs and comparing a set of CT images when a patient is breathing, we determine the interval for the entire motion. We apply a discrete scenario approach that indicates relative position of the critical organs and ROI for specific points within the interval. This method facilitated incorporating scenarios into a mathematical model. Figure 5.1 depicts the relative position of the organs and the motion considered for constructing the scenarios. An interval is specified for the intrafraction motion in z direction. However, other possibilities to include the motion can be addressed in future studies.

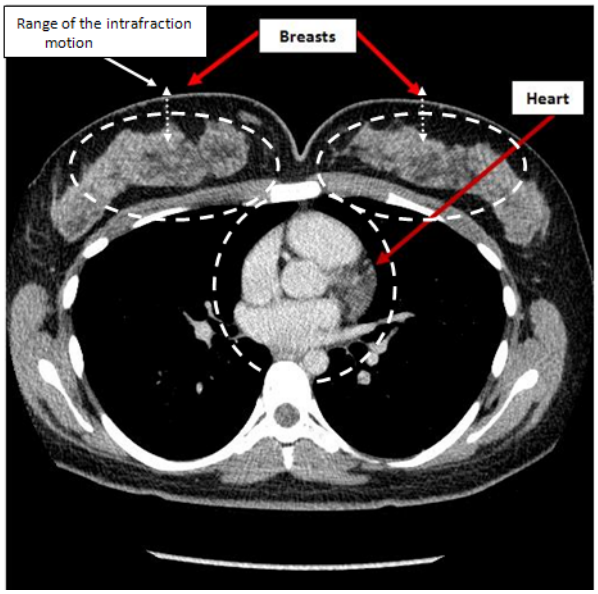


Figure 5.1: Demonstration of the relative positions of the heart and breasts. The scenarios are planned based on the intrafraction motions possible in the breasts and heart.

As an alternative to the deterministic model, we incorporate the scenario information into a robust model 5.3.4, so that the constraints reflect voxel location uncertainty.

In this section, we formulate the robust model which consists of indicated scenarios for the relative positions of the ROI and critical organs. The goal is minimizing the absorbed dose in the critical organs such as all the constraints corresponding to all scenarios are satisfied. Since there are parts of the formulation which are similar to the model in section 3.3, we refer to those constraints using the equation number. Set of constraints corresponding to the pencil beams remain the same as there is no effect by different scenarios on them. Variables and constraints corresponding to the body voxels and energy at each level are affected and for each scenario, there is one set of these variables and constraints.

5.3.2 Notations

We introduce the notations used in this section. However, we do not repeat those previously mentioned in section 3.

U: the set of all possible scenarios

k: index corresponding to the scenario

$\bar{C_o}$: the union of C_o^k for all k

$$\xi_{k,(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} = \begin{cases} 1 & \text{if there exist increments } i,\ j,\ l \text{ where} \\ & \hat{i} = i_p + i,\ \hat{j} = j_p + j,\ \hat{l} = l_p + l, \text{ for scenario } k \\ 0 & \text{otherwise} \end{cases}$$

5.3.3 Variables

We introduce notations used for declaring decision variables including continuous and binary variables which are refined from the model in section 3.3,

$E_{k,0}$ = Initial energy of each beam at source for scenario k

$$\Psi_{k,(i_p,j_p,l_p),i'j'} = \begin{cases} 1 & \text{if beam } i'j' \text{ intercepts pseudo voxel } (i_p, j_p, l_p) \text{ of phantom :} \\ & \gamma_{i'j'} = 1 \text{ and } x_i = 1,\ y_j = 1,\ z_l = 1, \text{ for scenario } k \\ 0 & \text{otherwise} \end{cases}$$

$$\beta_{k,(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} = \begin{cases} 1 & \text{if voxel } (i_p,j_p,l_p) \text{ is matching with pseudo voxel } \hat{i}\hat{j}\hat{l} \\ & \text{after increments } i,\ j,\ l \text{ in scenario } k \\ & \text{where } i=\hat{i}-i_p,\ j=\hat{j}-j_p,\ l=\hat{l}-l_p \\ & \text{and beam } i'j' \text{ intercepts with this pseudo voxel} \\ 0 & \text{otherwise} \end{cases}$$

$E_{k,l_p,i'j'} =$ Energy at layer l_p of the phantom when beam $i'j'$ is passing in scenario k

$E_{k,SD,i'j'} =$ Energy at the $i'j'$ pseudo pixel of the image receptor for scenario k

$$\tau_{k,i'j'} = \begin{cases} E_{k,SD,i'j'} & \text{if } \psi_{k,(i_p,j_p,l_p),i'j'} = 1\ , (i_p,j_p,l_p) \in ROI \\ 0 & \text{otherwise} \end{cases}$$

$$\Gamma_{k,(i_p,j_p,l_p-1),i'j'} = \begin{cases} E_{k,l_p,i'j'} & \text{if } \psi_{k,(i_p,j_p,l_p-1),i'j'} = 1 \\ 0 & \text{otherwise} \end{cases}$$

5.3.4 Optimization Model

The objective function for our robust approach is minimizing the absorbed dose in the critical organs across all scenarios. We can write it as follows,

$$\textbf{Min}\{\max_{k\in U} \sum_{V_{k,(ip,jp,lp)}\in C_o^k} dose_{V_{k,(ip,jp,lp)}}\} \tag{5.3}$$

Since the function of 5.3.4 is a nonlinear function, we turn it to a linear function by introducing ν and define it as follows,

$$\nu = \{\max_{k\in U} \sum_{V_{k,(ip,jp,lp)}\in C_o^k} dose_{V_{k,(ip,jp,lp)}}\} \tag{5.4}$$

Now the robust optimization model can be written as a MILP as follows,

$$\textbf{Min } v \tag{5.5}$$

Subject to

- Constraints (3.4), (3.5), (3.6), (3.7), (3.8),(3.16), (3.17), (3.18), (3.26)

$$v \geq \sum_{V_{k,(i_p,j_p,l_p)} \in C_o^k} dose_{V_{k,(i_p,j_p,l_p)}} \quad \forall\, k \in U \tag{5.6}$$

- Constraint to linearize $\psi_{k,(i_p,j_p,l_p),i'j'}$ equation :$\lfloor \frac{\beta_{k,(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} + \gamma_{i'j'}}{2} \rfloor$

$$\psi_{k,(i_p,j_p,l_p),i'j'} \geq \beta_{k,(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} + \gamma_{i'j'} - 1 \qquad \forall\, (i_p,j_p,l_p) \in VP,$$
$$(\hat{i},\hat{j},\hat{l}) \in S_{i'j'},\ s.t.\ \hat{i} = i + i_p,\ \hat{j} = j + j_p,\ \hat{l} = l + l_p,$$
$$i', j' = 1,..,n,\ k \in U \tag{5.7}$$

$$\psi_{k,(i_p,j_p,l_p),i'j'} \leq \gamma_{i'j'} \qquad \forall\, (i_p,j_p,l_p) \in VP,$$
$$i',\ j' = 1,..,n,\ k \in U \tag{5.8}$$

$$x_i + y_j + z_l \geq 3\, \beta_{k,(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \qquad \forall (i_p,j_p,l_p) \in VP,$$
$$\hat{i} - i_p = i,\ \hat{j} - j_p = j,\ \hat{l} - l_p = l,\ i', j' = 1,\ldots,n,\ k \in U \tag{5.9}$$

$$\beta_{k,(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \leq \xi_{k,(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \qquad \forall (i_p,j_p,l_p) \in VP,$$
$$\hat{i} - i_p = i,\ \hat{j} - j_p = j,\ \hat{l} - l_p = l, i', j' = 1,\ldots,n,\ k \in U \tag{5.10}$$

$$\psi_{k,(i_p,j_p,l_p),i'j'} \leq \sum_{(\hat{i},\hat{j},\hat{l}) \in S_{i'j'}} \beta_{k,(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} \qquad \forall (i_p,j_p,l_p),\ \in VP,$$
$$\forall i'j' = 1,\ldots,n,\ \forall\, k \in U \tag{5.11}$$

$$\beta_{k,(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} - (x_i + y_j + z_l + \gamma_{i'j'}) \geq \xi_{k,(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'} - 4$$
$$\forall (i_p,j_p,l_p) \in VP,\ \hat{i} - i_p = i,\ \hat{j} - j_p = j,\ \hat{l} - l_p = l,\ \forall\, i'j' = 1,\ldots,n \tag{5.12}$$

- Constraint to linearize objective function

$$\Gamma_{k,(i_p,j_p,l_p),i'j'} \geq E_{k,l_p,i'j'} - M(1-\psi_{k,(i_p,j_p,l_p),i'j'}) \qquad \forall\, i'j',\ (i_p,j_p,l_p) \in \bar{C}_o \tag{5.13}$$

- Constraints to connect different levels of energy of a pencil beam

$$E_{k,l_p-1,i'j'} \leq E_{k,l_p,i'j'} + M \sum_{i_p \in VP} \sum_{j_p \in VP} \psi_{k,(i_p,j_p,l_p-1),i'j'}$$
$$\forall\, i',j' = 1,\ldots,n,\ (i_p,j_p,l_p) \in VP,\ k \in U \tag{5.14}$$

$$E_{k,l_p,i'j'} \geq E_{k,l_p-1,i'j'} e^{-\mu_{k,(i_p,j_p,l_p-1)}d} - M(1-\psi_{k,(i_p,j_p,l_p-1),i'j'})$$
$$\forall\, i',j' = 1,\ldots,n,\ (i_p,j_p,l_p) \in VP \tag{5.15}$$

$$E_{k,l_p,i'j'} \leq E_{k,l_p-1,i'j'} e^{-\mu_{k,(i_p,j_p,l_p-1)}d} + M(1-\psi_{k,(i_p,j_p,l_p-1),i'j'})$$
$$\forall\, i',j' = 1,\ldots,n,\ (i_p,j_p,l_p) \in VP \tag{5.16}$$

$$\sum_{i_p \in VP} \sum_{j_p \in VP} \psi_{k,(i_p,j_p,l_p),i'j'} \leq 1$$
$$\forall\, i',j' = 1,\ldots,n,\ \ (i_p,j_p,l_p) \in VP \tag{5.17}$$

- Constraints to specify the range for energy at each level

$$E_{k,l_p,i'j'} \leq E_{k,l_p-1,i'j'} \qquad \forall\, i',j' = 1,\ldots,n,\ l_p \in VP \tag{5.18}$$

$$n^2\, E_{k,0,i'j'} - E_0 = 0 \qquad \forall\, i',j' = 1,\ldots,n \tag{5.19}$$

$$E_{k,SD,i'j'} = E_{k,l_p,i'j'} \qquad \forall\, i',j' = 1,\ldots,n,\ l_p \in VP \tag{5.20}$$

- Image quality constraint

$$K_1'' \sum_{i'=1}^{n} \sum_{j'=1}^{n} \alpha_{i'j'} \leq \sum_{i'=1}^{n} \sum_{j'=1}^{n} \tau_{k,i'j'} \tag{5.21}$$

$$\sum_{i'=1}^{n} \sum_{j'=1}^{n} \tau_{k,i'j'} \leq K_2'' \sum_{i'=1}^{n} \sum_{j'=1}^{n} \alpha_{i'j'} \tag{5.22}$$

- Image quality complementary constraint for linearizing and avoiding division by zero

$$\tau_{k,i'j'} \leq E_{k,SD,i'j'} \qquad \forall\, i', j' = 1,\ldots,n \tag{5.23}$$

$$\tau_{k,i'j'} \leq M\alpha_{i'j'} \qquad \forall i', j' = 1,\ldots,n \tag{5.24}$$

$$\tau_{k,i'j'} \geq E_{k,SD,i'j'} - M\,(1-\alpha_{i'j'}) \qquad \forall i', j' = 1,\ldots,n,\ M = E_{max} \tag{5.25}$$

$$\alpha_{i'j'} \geq \psi_{k,(i_p,j_p,l_p),i'j'} \qquad \forall i', j' = 1,\ldots,n,\ (i_p,j_p,l_p) \in ROI \tag{5.26}$$

$$\alpha_{i'j'} \leq \sum_{i_p,j_p,l_p \in ROI} \psi_{k,(i_p,j_p,l_p),i'j'} \qquad \forall i', j' = 1,\ldots,n \tag{5.27}$$

- Variable range

$$\psi_{k,(i_p,j_p,l_p),i'j'} \in \{0,1\} \tag{5.28}$$

$$\Gamma_{k,l_p,i'j'} \geq 0 \tag{5.29}$$

$$\tau_{k,i'j'} \geq 0 \tag{5.30}$$

Here we minimize the maximum dose absorbed in the critical organs across all the input data scenarios which indicate different positions of the ROI and critical organs and also different sizes of the organs. The index of k in each constraint indicates the variable corresponding to input scenario k.

The formulation of the absorbed dose for each input scenario k is computed as follows,

$$\sum_{V_{k,(i_p,j_p,l_p)} \in C_o^k} dose_{V_{k,(i_p,j_p,l_p)}} = \sum_{(i_p,j_p,l_p):V_{k,(i_p,j_p,l_p)} \in C_o^k} \sum_{i'=1}^{n} \sum_{j'=1}^{n} \frac{\Gamma_{k,(i_p,j_p,l_p),i'j'}(1-e^{-\mu_{k,(i_p,j_p,l_p)}d})}{m_{V_{k,(i_p,j_p,l_p)}}} \tag{5.31}$$

5.4 Methodology

The size of the mathematical model presented in section 5.3.4 increases by adding scenarios. For three scenarios, the problem has 242,000 variables and 468,000 constraints.

The commercial solvers (CPLEX and Concert) cannot handle the size of the problem. Thus, we outline the Bender's decomposition method which is used to solve the robust model. Then we present computational results for the robust model.

5.4.1 Decomposition Method

Bender's decomposition is a common method for solving large-scale MIP problems [89]. The general MIP is defined as follows,

$$z = \min\ cx + hy \tag{5.32}$$

$$s.t. \quad Ax + Gy \geq b \tag{5.33}$$

$$x \in X \subseteq Z_+^n,\ y \in R_+^p \tag{5.34}$$

The integer variables x can be viewed as complicating variables and by fixing x, we have a linear program. If $\bar{x} \in X$ denotes the fix values for the integer variables, we can write the original problem as follows,

$$z_{LP}(x) = \min\ hy \tag{5.35}$$

$$s.t. \quad Gy \geq b - Ax \tag{5.36}$$

$$y \in R_+^p \tag{5.37}$$

The dual of the LP problem is written as follows (sub problem),

$$w_{DP}(u) = \max\ u(b - Ax) \tag{5.38}$$

$$uG \leq h \tag{5.39}$$

$$u \in R_+^m \tag{5.40}$$

By the dual polyhedron, we can characterize whether LP(x)is infeasible, unbounded or bounded. Let $\{u^k \in R_+^m :\ k \in K\}$ be the set of extreme points of $Q = \{u \in R_+^m : uG \geq h\}$ and let $\{v^j \in R_+^m : j \in J\}$ be the set of extreme rays of $\{u \in R_+^m : uG \geq 0\}$. If $Q \neq \emptyset$, then

$\{v^j \in R^m_+ : j \in J\}$ is also the set of extreme rays of Q. The $z_{LP}(x)$ can be characterized as follows [70],

- If $Q = \emptyset$, then $z_{LP}(x) = -\infty$ if $v^j(b - Ax) < 0, \quad \forall\ j \in J$ and $z_{LP}(x) = \infty$ otherwise.
- If $Q \neq \emptyset$, then $z_{LP}(x) = \min_{k \in K} u^k(b - Ax) < \infty$ if $v^j(b - Ax) \geq 0,\ \forall\ j \in J$ and $z_{LP}(x) = \infty$ otherwise.

Thus, when $Q \neq \emptyset$, MIP can be written as follows,

$$z = \min_x (cx + \min_{k \in K} u^k(b - Ax)) \tag{5.41}$$

$$s.t. \quad v^j(b - Ax) \leq 0 \quad \forall\ j \in J \tag{5.42}$$

$$x \in X \tag{5.43}$$

and it can be reformulated as follows (Master problem),

$$z = \min\ \eta \tag{5.44}$$

$$s.t. \qquad \eta \geq cx + u^k(b - Ax) \quad \forall\ k \in K \tag{5.45}$$

$$v^j(b - Ax) \leq 0 \quad \forall\ j \in J \tag{5.46}$$

$$x \in X,\ \eta \in R^1 \tag{5.47}$$

Since we have a large number of constraints, using Bender's algorithm, we only add constraints correspoding to the extreme points of the dual problem which is solved as a sub problem.

5.4.2 Steps of Bender's Decomposition

A.2: Steps of the Bender's decomposition algorithm

Initialize the problem by setting iteration $t = 1$ and choosing $\bar{x} \in X$. Set $LB = -\infty$ and $UB = \infty$.

while $UB - LB > \varepsilon$ **do**

solve sub problem: $w_{DP}(u) = \max\{u(b - Ax) | uG \leq h,\ u \in R_+^m\}$

if unbounded

get unbounded ray v and add cut $v^j(b - Ax) < 0$ to master problem

else

get extreme point u and add cut $\eta \leq cx + u^k(b - Ax)$ to master problem if violated

set $UB = \min\{UB, cx + u^k(b - Ax)\}$

end if

solve master problem: $\min_x\{\eta | cuts, x \in X\}$

set $LB = \bar{\eta}$

end while

5.4.3 Bender's Reformulation

We present the dual subproblem and master problem of the mathematical model 5.3.4 which are used in Bender's algorithm. We consider the set of binary variables $x = (\gamma_{i'j'},\ \psi_{k,(i_p,j_p,l_p),i'j'},\ \beta_{k,(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'},\ x_i,\ y_j,\ z_l,\ \alpha_{i'j'})$ where $x \in B$ to be set for the initial feasible solution for problem 5.3.4. We denote these values by

$\bar{\gamma}_{i'j'}$, $\bar{\psi}_{k,(i_p,j_p,l_p),i'j'}$, $\bar{\beta}_{k,(i_p,j_p,l_p),(\hat{i},\hat{j},\hat{l}),i'j'}$, $\bar{x}_i$, $\bar{y}_j$, $\bar{z}_l$, $\bar{\alpha}_{i'j'}$ to represent a feasible solution to the problem. We denote continuous variables by $y = (\tau_{i'j'}, \Gamma_{k,l_p,i'j'}, E_{k,SD,i'j'}, E_{k,l_p,i'j'})$ where $y \in R_+^p$.

The formulation for the relaxation LP problem is as follows,

$$\textbf{Min } \boldsymbol{\nu} \tag{5.48}$$

Subject to

$$\nu - \sum_{V_{k,(i_p,j_p,l_p)} \in C_o^k} dose_{V_{k,(i_p,j_p,l_p)}} \geq 0 \quad \forall\, k \in U \tag{5.49}$$

- Constraint to linearize objective function

$$\Gamma_{k,(i_p,j_p,l_p),i'j'} \geq E_{k,l_p,i'j'} - M(1 - \bar{\psi}_{k,(i_p,j_p,l_p),i'j'}) \qquad \forall\, i'j',\ (i_p,j_p,l_p) \in \bar{C}_o \tag{5.50}$$

- Constraints to connect different levels of energy of a pencil beam

$$E_{k,l_p-1,i'j'} - E_{k,l_p,i'j'} \leq M \sum_{i_p \in VP} \sum_{j_p \in VP} \bar{\psi}_{k,(i_p,j_p,l_p-1),i'j'}$$
$$\forall\, i', j' = 1,\ldots,n,\ (i_p,j_p,l_p) \in VP,\ k \in U \tag{5.51}$$
$$E_{k,l_p,i'j'} - E_{k,l_p-1,i'j'} e^{-\mu_{k,(i_p,j_p,l_p-1)}d} \geq -M(1 - \bar{\psi}_{k,(i_p,j_p,l_p-1),i'j'})$$
$$\forall\, i', j' = 1,\ldots,n,\ (i_p,j_p,l_p) \in VP \tag{5.52}$$
$$E_{k,l_p,i'j'} - E_{k,l_p-1,i'j'} e^{-\mu_{k,(i_p,j_p,l_p-1)}d} \leq M(1 - \bar{\psi}_{k,(i_p,j_p,l_p-1),i'j'})$$
$$\forall\, i', j' = 1,\ldots,n,\ (i_p,j_p,l_p) \in VP \tag{5.53}$$

- Constraints to specify the range for energy at each level

$$E_{k,l_p,i'j'} - E_{k,l_p-1,i'j'} \leq 0 \qquad \forall\, i', j' = 1,\ldots,n,\ l_p \in VP \tag{5.54}$$

$$n^2\, E_{k,0,i'j'} - E_0 = 0 \qquad \forall\, i', j' = 1,\ldots,n \tag{5.55}$$

$$E_{k,SD,i'j'} - E_{k,l_p,i'j'} = 0 \qquad \forall\, i', j' = 1,\ldots,n,\ l_p \in VP \tag{5.56}$$

- Image quality constraint

$$\sum_{i'=1}^{n}\sum_{j'=1}^{n} \tau_{k,i'j'} \geq K_1'' \sum_{i'=1}^{n}\sum_{j'=1}^{n} \bar{\alpha}_{i'j'} \tag{5.57}$$

$$\sum_{i'=1}^{n}\sum_{j'=1}^{n} \tau_{k,i'j'} \leq K_2'' \sum_{i'=1}^{n}\sum_{j'=1}^{n} \bar{\alpha}_{i'j'} \tag{5.58}$$

- Image quality complementary constraint for linearizing and avoiding division by zero

$$\tau_{k,i'j'} - E_{k,SD,i'j'} \leq 0 \qquad \forall\, i', j' = 1,\ldots,n \tag{5.59}$$

$$\tau_{k,i'j'} \leq M\bar{\alpha}_{i'j'} \qquad \forall i', j' = 1,\ldots,n \tag{5.60}$$

$$\tau_{k,i'j'} - E_{k,SD,i'j'} \geq -M\,(1-\bar{\alpha}_{i'j'}) \qquad \forall i', j' = 1,\ldots,n,\ M = E_{max} \tag{5.61}$$

- Variable range

$$E_{k,SD,i'j'} \geq 0 \tag{5.62}$$

$$E_{k,l_p,i'j'} \geq 0 \tag{5.63}$$

$$\Gamma_{k,l_p,i'j'} \geq 0 \tag{5.64}$$

$$\tau_{k,i'j'} \geq 0 \tag{5.65}$$

The problem is bounded since the constraints for the image quality have lower and upper bounds which indicates the feasible region of the problem is restricted. We define dual

variables corresponding to each constraint of sub problem 5.4.3 and let u denote set of dual variables and $\bar{u_k}$ be solution to the dual of sub problem 5.4.3 in kth iteration. We denote the right hand side of the constraints in sub problem 5.4.3 by b. If the coefficient of these constraints are denoted by A, then, a restricted master problem is formulated as follows,

Min ζ

Subject to

$\zeta \geq (b - Ax)\bar{u_k}, \quad \forall\, k \in K$

$x \in \mathbb{B}$

Now using the algorithm given in section 5.4.2, we can solve the problem and add cuts to the master problem when corresponding constraint 5.4.3 is violated.

5.4.4 Computational Results

The mathematical model presented for the robust optimization in section 5.3.4 is larger in terms of number of variables and constraints in compare with model presented in chapter 3. Even for small number of scenarios, the desktop computer having Intel(R) Xeon(R) CPU and 4GB of RAM cannot handle the variables and constraints to build the model in CPLEX. We use decomposition methods to overcome the memory limitations. We use the Bender's decomposition approach which is in general a row generation method and adds constraints to the problem whenever they are violated as described in section 5.4.

We use a margin for the ROI which is an expanded region around the nominal position of the ROI in which we are confident ROI will be irradiated. A treatment plan will be said to be robust if all of the constraints in the subsequent formulation are satisfied [86]. A set of scenarios is constructed using CT images and the intrafraction motion considered for the organs. These scenarios alter the input data used to construct the model and calculate the absorbed dose in the voxels. An interval of 2cm is considered for the

motion in z direction. We study the problem for three scenarios which indicate the position of the critical organ within the specified interval. The robust approach indicates a larger absorbed dose in the critical organs in compare with deterministic method for each set of predefined angulations. However, the robust approach is a worst case study which ensures the minimum dose across all scenarios. Table 5.1 shows the results for comparison of the objective function obtained from the robust optimization approach and objective function of each individual scenario. The values of each scenario is the optimal solution for that scenario. We also compute the deviation of the robust objective function from the maximum value of the scenarios. These results indicate that robust optimization may increase the maximum objective function scenario by 26% to ensure the feasibility for all the scenarios. We also present the computational results where we feed the robust solution

	Scenario 1	Scenario 2	Scenario 3	Robust	Max Dev.
(-20,20)	0.0024	0.0067	0.0054	0.0078	69.23%
(-20,0)	0.0397	0.0284	0.0214	0.0418	48.80%
(-10,0)	0.0251	0.0165	0.0178	0.0282	41.49%
(0,0)	0.0363	0.0314	0.0296	0.0440	32.73%
(15,0)	0.0045	0.0104	0.0097	0.0131	65.65%
(15,10)	0.0069	0.0086	0.0109	0.0127	45.67%
(20,0)	0.0271	0.0188	0.0116	0.0321	63.86%
(30,0)	0.0220	0.0148	0.0193	0.0266	44.36%

Table 5.1: Comparison of the absorbed dose (Gy) for each individual scenario and the robust model (Max Dev. indicates the deviation percentage of the robust objective function from the max objective function across the scenarios)

in each scenario and compare the objective function of each scenario using robust solution with its optimal solution. Using this comparison, we can show how much robust solution is away from the best solution for each scenario. Table 5.2 depicts the computational results where the value presented for each scenario is the subtraction of the optimal objective function of the scenario from the objective function calculated using robust solution. We compute maximum deviation from optimal for each set of angles. Table 5.2

Angles	Scenario 1	Scenario 2	Scenario 3	Max Dev.
(-20,20)	0.0018	0.0008	0.0017	23.08%
(-20,0)	0.0006	0.0028	0.0142	33.97%
(-10,0)	0.0016	0.0050	0.0016	17.73%
(0,0)	0.0017	0.0090	0.0019	20.45%
(15,0)	0.0062	0.0011	0.0028	47.33%
(15,10)	0.0047	0.0017	0.0012	37.01%
(20,0)	0.0023	0.0041	0.0161	50.16%
(30,0)	0.0031	0.0043	0.0043	16.17%

Table 5.2: Deviation of the optimal objective function (Gy) from the robust solution using the setting of the robust solution for each individual scenario

indicates a maximum deviation of 0.0161 Gy for the difference of the robust solution from the optimal absorbed dose in each scenario. We also use the optimal solution of scenario 1 and plug it in corresponding problems of scenario 2 and 3. Then we report feasibility of the problem, absorbed dose (Gy) and percentage of the image area which is missed by using the field of radiation given in scenario 1. These results are reported in Table 5.3.

Angles	Test	Scenario 1	Scenario 2	Scenario 3
(0,0)	Feasibility test	Optimal	Infeasible	Infeasible
	Dose-Scen1 (Gy)	0.0363	0.0395	0.0357
	Missed area of the image	0%	12.5%	18.75%
	Dose-Robust (Gy)	0.0380	0.0404	0.0315
(20,0)	Feasibility test	Optimal	Infeasible	Infeasible
	Dose-Scen1 (Gy)	0.0271	0.0249	0.0307
	Missed area of the image	0%	18.75%	31.25%
	Dose-Robust (Gy)	0.0294	0.0229	0.0277
(15,10)	Feasibility test	Optimal	Feasible	Infeasible
	Dose-Scen1 (Gy)	0.0069	0.0112	0.0138
	Missed area of the image	0%	6.25%	12.5%
	Dose-Robust (Gy)	0.0116	0.0103	0.0121

Table 5.3: Deviation of using the optimal objective function from the scenario 1 in the other scenarios

We optimize the MIP problem for scenario 1 and using the expanded ROI which is feasibility condition in robust model. Then we plug in the solution of that in problems

Angle	Computed item	Scenario 1	Scenario 2	Scenario 3
(0,0)	Dose-Scen1 (Gy)	0.0369	0.0427	0.0368
	Dose-Robust (Gy)	0.0380	0.0404	0.0315
	Deviation	-2.98%	5.39%	14.40%
(20,0)	Dose-Scen1 (Gy)	0.0288	0.0273	0.0315
	Dose-Robust (Gy)	0.0294	0.0229	0.0277
	Deviation	-2.08%	16.12%	12.06%

Table 5.4: Solving problem for scenario 1 with expanded ROI and plugging in the solution in problems solve for the scenario 2 and 3

with scenario 2 and 3. The results are reported in Table 5.4. Presented results for two set of angles show that the optimal absorbed dose for scenario 1 is smaller than robust solution and the optimal absorbed dose given by scenario 2 and 3 are higher than robust solution and the deviation is up to 16% in these examples. The deviation shows that using a robust model can be beneficial and reduce the absorbed dose. We present computational results for using Bender's decomposition method to solve the robust model in Table 5.5. We set a time limit of 2000 seconds to run the algorithm. The optimality gap varied from 5% to 69% in different settings we run the algorithm. However, the solution to the robust model provides the minimum absorbed dose with respect to the image quality satisfaction for all the scenarios. To improve the optimality gap and solution quality, we can apply

Primary angle	Secondary angle	Aborbed dose (Gy)	Number of iterations	Optimality gap
0	0	0.0440	15	17.5%
-10	0	0.0282	21	10.99%
20	0	0.0321	19	15.58%
-20	0	0.0418	14	5.02%
15	0	0.0131	23	65.65%
30	0	0.0266	13	17.29%
15	10	0.0127	19	45.67%
-20	20	0.0078	14	69.23%

Table 5.5: Computational results for the Bender's decomposition method used for the robust model 5.3.4

other decomposition methods and implementations to use the structure of the problem more efficiently.

5.5 Summary

We presented a robust mathematical model to minimize the maximum absorbed dose in the critical organs with respect to the image quality in all the input data scenarios. We constructed our input data scenarios by considering the uncertainty of the relative position of the organs resulted from breathing. Our method addresses the uncertainty for the motion of the organs and change in the size of the organs. Results show an increase of the absorbed dose in the critical organs against each scenario up to 26%. However, the robust formulation hedges against motion uncertainty as determined by the scenarios.

There are other sources of the uncertainty which are involved in the radiology procedures and we did not consider in our robust optimization model because of the computation limitations. However, it is worth to find more efficient computational methods so that we can consider other sources of variability and include more possible scenarios. Hybrid decomposition methods which combine two or more algorithms may work more efficiently if sub problems are run simultaneously using parallel processors. We can also use subgradient methods which are based on partial relaxations of the mathematical model but have the advantage of exploring continuous variables. We can run each scenario in a subroutine independent of the other scenarios and then proceed with feasibility search using subgradient algorithm.

Chapter 6

CONCLUSION

we formulate the radiation reduction problem for an FGI procedure as a mixed integer linear program. Derived and implemented cutting planes show the efficiency improvements of the computational time. The radiation reduction can be achieved to over 70% in the computational experiments. The current results are based on theoretical computations. The dose reduction performance can be validated by using body phantom in the lab or Monte Carlo simulation. Even though the computational time is reduced significantly by adopting lazy cuts, the size of the problem may result in computational difiiculties when the size of the ROI or critical organ is large. It is also possible that some of the cuts which are more efficient are not captured from the pool by the algorithm.

The mathematical model presented in chapter 3 optimizes the problem for each given set of angles. Including angles in the model, results in a MINLP which enhances the complexity of the computation. To optimize the angulation and table locations simultaneously, we proposed a modified subgradient algorithm which treats angles and table locations as a vector and updates this vector at each iteration of the algorithm by using an estimate of subgradient. Each iteration of the subgradient method solves a partially relaxed MIP model which determines the field of radiation and energy and explores the continuous geometry. The results of this method demonstrates that this algorithm can reduce the absorbed dose in the critical organs by slightly change in the geometry. This method has the capability of exploring continuous geometry which might not be explored using the proposed discrete method. The computational results indicate further reduction (up to 80%) of the absorbed dose in compare with previous method.

We provided a robust optimization model to handle motion uncertainties within a breathing process. We applied scenario based method and minimax function for the optimization.

6.1 Future Research

The mathematical model presented in chapter 3 is a large-scale MIP and computation arises when size of the problem increases. We proposed some facet defining inequalities and preprocessing algorithm to speed up the computation and enhance the solution quality. However, more polyhedral analysis and investigating more heuristics is valuable and will be in our future steps.

Validation of the mathematical models provided in this dissertation can be achieved through a complete measurement of the absorbed dose and image processing. This will enable the capabilities for vendors adoption of real-time automatic dose reduction.

We proposed a modified subgradient method which shows some improvement on the computational complexities of the problem. This method can be improved more to provide a bound for the objective function. We can also analyze the convergence of the method using different step sizes and projection methods. This method can be used if we refine the size of the voxels and use smaller voxels which cannot be handled by the original model when using a branch-and-cut method.

A robust optimization problem is presented to indicate how we can handle uncertainties in the model. More sources of the uncertainty can be accounted into the model. More investigation is required to find more efficient algorithms and decomposition methods. A parallel processing approach can be used if we decompose problem to independent sub problems which can be solved simultaneously. Applying a modified

subgradient method to the robust model is another method which can improve the computational time by customizing the subgradient algorithm to the robust model.

REFERENCES

[1] D. J. Stoddard, "A Compelling Reason Why John Gofman is Correct in Identifying Medical X-rays as a Major Cause of Cancer." http://www.healpain.net/articles/reason.html, 2001.

[2] ICRP, "International commission on radiological protection," *Annals of ICRP*, vol. 26, 1977. Pergamon Press, Oxford.

[3] ICRP, "Recommendations of the international commission on radiological protection, ICRP publication 60," *Annals of ICRP*, vol. 21, 1991.

[4] National Research Council. Committee to Assess Health Risks from Exposure to Low Level of Ionizing Radiation, *Health Risks from Exposure to Low Levels of Ionizing Radiation: BEIR VII Phase 2*, vol. 7. 500 Fifth Street NW Washington DC 20001: THE NATIONAL ACADEMIES PRESS, 2006.

[5] J. A. Kaufman, J. A. Reekers, J. P. Burnes, A. Al-Kutoubi, C. A. Lewis, B. W. Hardy, S. Kuribayashi, and S. Sharma, "Global statement defining interventional radiology," *Journal of Vascular and Interventional Radiology*, vol. 21, pp. 1147–1322, August 2010.

[6] International Atomic Energy Agency (IAEA), "Radiation protection in diagnostic and interventional radiology. L 20: Optimization of protection in digital radiology," 2012. IAEA Training Material on Radiation Protection in Diagnostic and Interventional Radiology.

[7] American Colledge of Radiology (ACR) and Radiological Sociatey of North America (RSNA), "Radiologyinfo.org." http://www.radiologyinfo.org/en/info.cfm?pg=imrt, February 10, 2012.

[8] United States Nuclear Regulatory Commission (USNRC) Technical Training Center, *Biological Effects of Radiation.* USNRC Technical Training Center, 2011. Reactor Concepts Manual. http://www.nrc.gov/reading-rm/doc-collections/fact-sheets/bio-effects-radiation.html.

[9] J. T. Bushberg, J. Seibert, E. M. Leidholdt, and J. M. Boone, *The Essentials Physics of Medical Imaging*. 530 walnut street Philadelphia Pa 19106 USA: LIPPINCOTT Williams and Wilkins, second ed., 2002.

[10] B. Lindell, H. Dunster, and J. Valentin, "International commission on radiological protection: History, policies, procedures," *Swedish Radiation Protection Institute, SE*, vol. 171, pp. 1–12, 1998.

[11] D. J. Brenner, "Effective dose a flawed concept that could and should be replaced," in *ICRP*, 2011.

[12] Food and Drug Administration, "Public health advisory: Avoidance of serious x-ray-induced skin injuries to patients during fluoroscopically-guided procedures," tech. rep., Food and Drug Administration, September 1994.

[13] L. K. Wagner, M. D. McNeese, M. V. Marx, and E. L. Siegel, "Severe skin reactions from interventional fluoroscopy: Case report and review of the literature," *Radiology*, vol. 213, pp. 773–776, 1999.

[14] Joint Commission, "Radiation overdose as a reviewable sentinel event," tech. rep., Joint Commission, July 2006. http://www.jointcommission.org/NR/rdonlyres/10A599B4-832D-40C1-8A5B-5929E9E0B09D/0/Radiation_Overdose.pdf.

[15] W. Teeuwisse, J. Geleijns, J. Broerse, W. Obermann, and E. Van Persijn Van Meerten, "Patient and staff dose during CT guided biopsy, drainage and coagulation," *British journal of radiology*, vol. 74, no. 884, pp. 720–726, 2001.

[16] V. Tsapaki, C. Triantopoulou, P. Maniatis, S. Kottou, J. Tsalafoutas, and J. Papailiou, "Patient skin dose assessment during CT-guided interventional procedures," *Radiation protection dosimetry*, vol. 129, no. 1-3, pp. 29–31, 2008.

[17] A. Wambersie, J. Zoetelief, H. G. Menzel, and H. Paretzke, "The ICRU (international commission on radiation units and measurements): Its contribution to dosimetry in diagnostic and interventional radiology," *Radiation Protection Dosimetry*, vol. 117, no. 1-3, pp. 7–12, 2005.

[18] Office of Public Affairs, "Biological effects of radiation," tech. rep., United States Nuclear Regulatory Commission, October 2011. Fact Sheet.

[19] A. Miller, G. Howe, G. Sherman, J. Lindsay, M. Yaffe, P. Dinner, H. Risch, and D. Preston, "Mortality from breast cancer after irradiation during fluoroscopic examinations in patients being treated for tuberculosis," *New England Journal of Medicine*, vol. 321, no. 19, pp. 1285–1289, 1989.

[20] B. McParland, "A study of patient radiation doses in interventional radiological procedures," *The British Journal of Radiology*, vol. 71, no. 842, pp. 175–185, 1998.

[21] P. Dendy, "Radiation risks in interventional radiology," *The British Journal of Radiology*, vol. 81, pp. 1–7, 2008.

[22] J. Valentin, "Avoidance of radiation injuries from medical interventional procedures, ICRP publication 85," *Annals of the ICRP*, vol. 30, no. 2, p. 7, 2000.

[23] N. Andrieu, D. F. Easton, J. Chang-Claude, M. A. Rookus, R. Brohet, E. Cardis, A. C. Antoniou, T. Wagner, J. Simard, G. Evans, S. Peock, J.-P. Fricker, C. Nogues, L. V. Veer, F. E. V. Leeuwen, and D. E. Goldgar, "Effect of chest x-rays on the risk

of breast cancer among brca1/2 mutation carriers in the international brca1/2 carrier cohort study: A report from the embrace, genepso, geo-hebon, and ibccs collaborators group," *Journal of Clinical Oncology*, vol. 24, pp. 3361–3366, July 2006.

[24] E. Vano and K. Faulkner, "ICRP special radiation protection issues in interventional radiology, digital and cardiac imaging," *Radiation Protection Dosimetry*, vol. 117, no. 1-3, pp. 13–17, 2005.

[25] M. Kiss, D. Sayers, and Z. Zhong, "Measurement of image contrast using diffraction enhanced imaging," *Physics in Medicine and Biology*, vol. 48, no. 3, pp. 325–340, 2003.

[26] The International Organization for Standardization (ISO) 15739, "Photography-electronic still-picture imaging noise measurements," May 2003.

[27] L. A. Wolsey, *Integer Programming*. John Wiley and Sons, 1998.

[28] M. Sandborg, D. Dance, J. Persliden, and G. Carlsson, "A monte carlo program for the calculation of contrast, noise and absorbed dose in diagnostic radiology," *Computer methods and programs in biomedicine*, vol. 42, no. 3, pp. 167–180, 1994.

[29] T. O. Marshall, *Accuracy and Precision of Thermoluminescence Dosimetry*. London: HPA: Hufton, 1984.

[30] N. Meriç, "Calculation of radiation dose to the lens of the eye using monte carlo simulation," *Applied Radiation and Isotopes*, vol. 55, no. 4, pp. 557–560, 2001.

[31] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, *et al.*, "Geant4a simulation toolkit," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250–303, 2003.

[32] H. Jiang and H. Paganetti, "Adaptation of geant4 to monte carlo dose calculations based on CT data," *Medical physics*, vol. 31, no. 10, pp. 2811–2818, 2004.

[33] M. Khodaverdi, A. Chatziioannou, S. Weber, K. Ziemons, H. Halling, and U. Pietrzyk, "Investigation of different microct scanner configurations by geant4 simulations," *Nuclear Science, IEEE Transactions on*, vol. 52, no. 1, pp. 188–192, 2005.

[34] T. Aso, A. Kimura, T. Yamashita, and T. Sasaki, "Optimization of patient geometry based on CT data in geant4 for medical application," in *Nuclear Science Symposium Conference Record, 2007. NSS'07. IEEE*, vol. 4, pp. 2576–2580, IEEE, 2007.

[35] J. Briesmeister, *MCNP: A general Monte Carlo N-particle transport code, version 5*. Los Alamos NM, USA, January 2004. http://mcnp-green.lanl.gov/.

[36] F. Schultz, J. Geleijns, F. Spoelstra, and J. Zoetelief, "Monte carlo calculations for assessment of radiation dose to patients with congenital heart defects and to staff during cardiac catheterizations," *British journal of radiology*, vol. 76, no. 909, pp. 638–647, 2003.

[37] M. Tapiovaara, M. Lakkisto, and A. Servomaa, *PCXMC: a PC-based Monte Carlo program for calculating patient doses in medical x-ray examinations*. Finnish Centre for Radiation and Nuclear Safety, 1997.

[38] D. Pelowitz, "MCNPX users manual version 2.5.0," *Los Alamos National Laboratory LA-CP-05-0369*, 2005.

[39] S. Correa, E. Souza, A. Silva, R. Lopes, and H. Yoriyaz, "Dose-image quality study in digital chest radiography using monte carlo simulation," *Applied Radiation and Isotopes*, vol. 66, no. 9, pp. 1213–1217, 2008.

[40] R. Kramer, H. Khoury, J. Vieira, E. Loureiro, V. Lima, F. Lima, and G. Hoff, "All about FAX: a female adult voxel phantom for monte carlo calculation in radiation protection dosimetry," *Physics in medicine and biology*, vol. 49, pp. 5203–5216, 2004.

[41] K. Bacher, E. Bogaert, R. Lapere, D. De Wolf, and H. Thierens, "Patient-specific dose and radiation risk estimation in pediatric cardiac catheterization," *Circulation*, vol. 111, no. 1, pp. 83–89, 2005.

[42] A. Bozkurt and D. Bor, "Simultaneous determination of equivalent dose to organs and tissues of the patient and of the physician in interventional radiology using the monte carlo method," *Physics in medicine and biology*, vol. 52, no. 2, p. 317, 2007.

[43] R. Kramer, H. Khoury, and J. Vieira, "Caldose xa software tool for the assessment of organ and tissue absorbed doses, effective dose and cancer risks in diagnostic radiology," *Physics in Medicine and Biology*, vol. 53, no. 22, p. 64376459, 2008.

[44] International Electrotechnical Commission, "Medical electrical equipment-part 2-44: Particular requirements for the safety of x-ray equipment for computed tomography," Tech. Rep. 60601-2-44, International Electrotechnical Commission (IEC), November 2002. International Standard.

[45] G. Starck, L. Lonn, A. Cederblad, E. Forssell-Aronsson, L. Sjostrom, and M. Alpsten, "A method to obtain the same levels of CT image noise for patients of various sizes, to minimize radiation dose," *British journal of radiology*, vol. 75, no. 890, pp. 140–150, 2002.

[46] E. Nickoloff, A. Dutta, and Z. Lu, "Influence of phantom diameter, kVp and scan mode upon computed tomography dose index," *Medical physics*, vol. 30, pp. 395–402, 2003.

[47] M. Kalra, M. Maher, T. Toth, B. Schmidt, B. Westerman, H. Morgan, and S. Saini, "Techniques and applications of automatic tube current modulation for CT1," *Radiology*, vol. 233, no. 3, pp. 649–657, 2004.

[48] J. Purdy, J. Michalski, J. Bradley, S. Vijayakumar, C. Perez, and S. Levitt, "Three-dimensional treatment planning and conformal therapy," *Technical Basis of Radiation Therapy*, pp. 179–202, 2006.

[49] S. Yamamoto, T. Horiuchi, J. Sekiguchi, S. Wada, M. Komizu, and T. Yamaguchi, "Design of a dicom image-based program for estimating patient exposure dose in computed tomography," *Technology and Health Care*, vol. 15, no. 2, pp. 147–156, 2007.

[50] W. Huda, K. Ogden, and M. Khorasani, "Effect of dose metrics and radiation risk models when optimizing CT x-ray tube voltage," *Physics in Medicine and Biology*, vol. 53, pp. 4719–4732, 2008.

[51] G. J. Lim, J. Choi, and R. Mohan, "Iterative solution methods for beam angle and fluence map optimization in intensity modulated radiation therapy planning," *OR Spectrum*, vol. 30, pp. 289–309, 2008.

[52] M. Ehrgott, Ç. Güler, H. Hamacher, and L. Shao, "Mathematical optimization in intensity modulated radiation therapy," *4OR: a Quarterly Journal of Operations Research*, vol. 6, no. 3, pp. 199–262, 2008.

[53] D. Shepard, M. Ferris, G. Olivera, and T. Mackie, "Optimizing the delivery of radiation therapy to cancer patients," *Siam Review*, vol. 41, no. 4, pp. 721–744, 1999.

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

[55] H. Romeijn, R. Ahuja, J. Dempsey, A. Kumar, and J. Li, "A novel linear programming approach to fluence map optimization for intensity modulated radiation therapy treatment planning," *Physics in Medicine and Biology*, vol. 48, no. 21, pp. 3521–3542, 2003.

[56] D. M. Aleman, D. Glaser, H. E. Romeijn, and J. F. Dempsey, "Interior point algorithms: guaranteed optimality for fluence map optimization in imrt," *Physics in Medicine and Biology*, vol. 55, no. 18, pp. 5467–5482, 2010.

[57] F. Carlsson, *Utilizing problem structure in optimization of radiation therapy*. PhD thesis, Matematik, Kungliga Tekniska högskolan, Stockholm, Sweden, 2008.

[58] F. Carlsson and A. Forsgren, "Iterative regularization in intensity-modulated radiation therapy optimization," *Medical physics*, vol. 33, no. 1, pp. 225–234, 2006.

[59] G. Bednarz, D. Michalski, P. Anne, and R. Valicenti, "Inverse treatment planning using volume-based objective functions," *Physics in medicine and biology*, vol. 49, no. 12, pp. 2503–2514, 2004.

[60] G. Bednarz, D. Michalski, C. Houser, M. Huq, Y. Xiao, P. Anne, and J. Galvin, "The use of mixed-integer programming for inverse treatment planning with pre-defined field segments," *Physics in medicine and biology*, vol. 47, no. 13, pp. 2235–2245, 2002.

[61] 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*, vol. 18, no. 4, pp. 887–893, 1990.

[62] Z. C. Taşkın, J. C. Smith, and H. E. Romeijn, "Mixed-integer programming techniques for decomposition imrt fluence maps using rectangular apertures," *Annals of Operations Research*, pp. 1–20, June 2010.

[63] E. Lee, T. Fox, and I. Crocker, "Simultaneous beam geometry and intensity map optimization in intensity-modulated radiation therapy," *International Journal of Radiation Oncology* Biology* Physics*, vol. 64, no. 1, pp. 301–320, 2006.

[64] E. Lee, T. Fox, and I. Crocker, "Integer programming applied to intensity-modulated radiation therapy treatment planning," *Annals of Operations Research*, vol. 119, no. 1, pp. 165–181, 2003.

[65] J. Dai and Y. Zhu, "Conversion of dose–volume constraints to dose limits," *Physics in medicine and biology*, vol. 48, no. 23, pp. 3927–3941, 2003.

[66] 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*, vol. 23, pp. 957–964, 1996.

[67] X. Wang, X. Zhang, L. Dong, H. Liu, Q. Wu, and R. Mohan, "Development of methods for beam angle optimization for imrt using an accelerated exhaustive search strategy," *International Journal of Radiation Oncology* Biology* Physics*, vol. 60, no. 4, pp. 1325–1337, 2004.

[68] Y. Watanabe, "Derivation of linear attenuation coefficients from CT numbers for low-energy photons," *Physics in Medicine and Biology*, vol. 44, no. 9, pp. 2201–2211, 1999.

[69] X. G. Xu and K. F. Eckerman, *Handbook of Anatomical Models for Radiation Dosimetry*. Series in Medical Physics and Biomedical Engineering, CRC Press Taylor and Francis Group, 2010.

[70] G. Nemhauser and L. Wolsey, *Integer and combinatorial optimization*. John Wiley and Sons, 1999.

[71] D. Bertsekas and J. Tsitsiklis, "Gradient convergence in gradient methods with errors," *SIAM Journal on Optimization*, vol. 10, no. 3, pp. 627–642, 2000.

[72] D. Bertsekas, *Nonlinear Programming*. Cambridge, MA.: Athena Scientific, 2nd ed., 1999.

[73] N. Z. Shor, *Minimization Methods for Non-differentiable Functions*. Springer, 1985. Springer Series in Computational Mathematics.

[74] R. Burachik, C. Kaya, and M. Mammadov, "An inexact modified subgradient algorithm for nonconvex optimization," *Computational Optimization and Applications*, vol. 45, no. 1, pp. 1–24, 2010.

[75] A. Bagirov and A. Ganjehlou, "An approximate subgradient algorithm for unconstrained nonsmooth, nonconvex optimization," *Mathematical Methods of Operations Research*, vol. 67, no. 2, pp. 187–206, 2008.

[76] R. Freund, "Subgradient optimization, generalized programming, and nonconvex duality." Massachusetts Institute of Technology, 2004.

[77] N. Z. Shor, *Nondifferentiable Optimization and Polynomial Problems*. Netherlands: Kluwer, 1998.

[78] 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*, vol. 31, p. 1141, 2004.

[79] D. Michalski, Y. Xiao, Y. Censor, and J. Galvin, "The dose–volume constraint satisfaction problem for inverse treatment planning with field segments," *Physics in medicine and biology*, vol. 49, no. 4, p. 601, 2004.

[80] Y. Censor, T. Bortfeld, B. Martin, and A. Trofimov, "A unified approach for inversion problems in intensity-modulated radiation therapy," *Physics in Medicine and Biology*, vol. 51, no. 10, p. 2353, 2006.

[81] D. Craft, "Local beam angle optimization with linear programming and gradient search," *Physics in medicine and biology*, vol. 52, no. 7, p. N127, 2007.

[82] D. Bertsimas, V. Cacchiani, D. Craft, and O. Nohadani, "A hybrid approach to beam angle optimization in intensity-modulated radiation therapy," *Computers & Operations Research*, 2012.

[83] A. Olafsson and S. Wright, "Efficient schemes for robust imrt treatment planning," *Physics in medicine and biology*, vol. 51, no. 21, p. 5621, 2006.

[84] T. Chan, T. Bortfeld, and J. Tsitsiklis, "A robust approach to imrt optimization," *Physics in medicine and biology*, vol. 51, no. 10, p. 2567, 2006.

[85] 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*, vol. 50, no. 23, p. 5463, 2005.

[86] T. Bortfeld, T. Chan, A. Trofimov, and J. Tsitsiklis, "Robust management of motion uncertainty in intensity-modulated radiation therapy," *Operations Research*, vol. 56, no. 6, pp. 1461–1473, 2008.

[87] P. Kouvelis and G. Yu, *Robust discrete optimization and its applications*, vol. 14. Springer, 1996.

[88] J. Mulvey, R. Vanderbei, and S. Zenios, "Robust optimization of large-scale systems," *Operations research*, vol. 43, no. 2, pp. 264–281, 1995.

[89] J. Benders, "Partitioning procedures for solving mixed-variables programming problems," *Numerische mathematik*, vol. 4, no. 1, pp. 238–252, 1962.

APPENDIX A

GLOSSARY

GLOSSARY

Abbreviation	Description	Page
CT	Computed Tomography	page 1
FGI	Fluoroscopy Guided Interventional Radiology	page 1
IMRT	Intensity Modulated Radiation Therapy	page 1
Gy	Gray	page 2
FDA	Food and Drug Administration	page 4
ICRP	International Commission on Radiological Protection	page 11
SNR	Signal-to-Noise Ratio	page 13
MIP	Mixed-integer Programming	page 13
DICOM	Digital Imaging and Communications in Medicine	page 19
GEANT	GEometry ANd Tracking	page 19
MCNP	Monte Carlo N-Particle	page 20
mA	mili Amperage: Charge/Current	page 21
kV	kilo voltage	page 21
CTDI	Computed Tomography Dose Index	page 22
C-G	Chvatal-Gomory	page 46