Rapid Re-optimization of Prostate Intensity-Modulated Radiation Therapy Using Regularized Linear Programming

RAPID RE-OPTIMIZATION OF PROSTATE INTENSITY-MODULATED RADIATION THERAPY USING REGULARIZED LINEAR PROGRAMMING

BY

MARYAM KHALAJI,

A THESIS

SUBMITTED TO THE DEPARTMENT OF COMPUTING & SOFTWARE

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

© Copyright by Maryam Khalaji, March 2013

All Rights Reserved

Master of Science (2013) McMaster University

(Computing and Software) Hamilton, Ontario, Canada

TITLE:	Rapid Re-optimization of Prostate Intensity-Modulated Radiation Therapy Using Regularized Linear Programming
AUTHOR:	Maryam Khalaji B.A.Sc., (Software Engineering) University of Tehran, Tehran, Iran
SUPERVISOR:	Dr. Christopher Anand Dr. Marcin Wierzbicki
NUMBER OF PAGES:	vi, 64

Abstract

This thesis presents a new linear programming approach for re-optimizing a intensity modulated radiation therapy (IMRT) treatment plan, in order to compensate for inter-fraction tissue deformations. Different formulations of the problem involve different constraints, but a common constraint that is difficult to handle mathematically is the constraint that the dose be deliverable using a small number of multi-leaf collimator positions. MLC leaves are tungsten alloy attenuators which can be moved in and out to shape of the radiation aperture. Since leaves are solid, photon fluence profiles will follow a staircase function and this constraint is not convex, and difficult to formulate. In this thesis, we propose a relaxation of this constraint to the ℓ_1-norm of the differences between adjacent radiation fluxes. With the appropriate bound, this constraint encourages the dose to be deliverable with a series of shrinking or growing openings between the leaves. Such a solution can be made realizable by rounding, which is beyond the scope of this thesis.

This approach has been tested on an anonymized prostate cancer treatment plan with simulated deformations. Without rounding, solutions were obtained in five of nine cases, in less than 4 to 5 seconds of computation on a NEOS server. Solved cases demonstrated excellent target coverage (minimum dose in the target was 95% of the prescribed dose) and organ sparing (mean dose in normal tissues was below 25% of the prescribed dose).

Notation and abbreviations

IMRT Intensity Modulated Radiation Therapy

MLC Multi-leaf Collimator

CTV Clinical Target Volume

PTV Planning Target Volume

OAR Organs at Risk

DVC Dose-volume Constraint

CT Computed tomography

AMPL A Mathematical Programming Language

ODM Opening Density Matrix

CBCT Cone-beam Computed Tomography

ISL Inverse Square Law

Gy Gray

LP Linear Programming

KNITRO Nonlinear Interior point Trust Region Optimization

TERMA Total Energy Released per Unit Mass

Contents

List of Figures

List of Tables

Chapter 1

Introduction to IMRT

Prior to introducing the proposed optimization method, first the motivation for prostate cancer radiotherapy is clarified. This includes statistics from various references which shows the importance of this work. Having motivation explained, the basic principles of Intensity Modulated Radiation Therapy(IMRT) process are introduced. This involves a brief introduction to IMRT for prostate cancer treatment, as well as explaining the concept of inverse planning, IMRT delivery and tumour localization. The goal of this chapter is to ensure the reader will gave a good understanding of why and how IMRT is being used for treatment of prostate cancer.

1.1 Prostate Cancer

Other than skin cancer, prostate cancer is the most common cancer in American men. The American Cancer Society estimates for prostate cancer in the United States for 2013 are:

- About 238,590 new diagnoses

- About 29,720 deaths

Prostate cancer can be a serious disease, but most men diagnosed with prostate cancer do not die from it. In fact, more than 2.5 million men in the United States who have been diagnosed with prostate cancer at some point are still alive today [21]. Fortunately, development of new methods in radiotherapy improved treatment success rate while reducing side effects [24].

1.2 Prostate Cancer Radiation Therapy

The entire radiotherapy chain maybe summarized as follows:

- Patient is diagnosed
- CT scan is performed
- Tumour volume and radiosensitive organs are delineated
- A radiotherapy treatment plan is created
- The same treatment is delivered daily over six to eight weeks
- Patient attends follow up appintments

An ideal radiotherapy treatment plan delivers a high dose of radiation to the tumour while minimizing the dose delivered to the surrounding normal tissues such as the rectum, bladder and femoral heads. Typically, there is a direct relationship between the amount of radiation delivered to tumour and probability of tumour control. Practically, the deliverable tumour dose is often limited by the radiation tolerance of sensitive organs around it.

Intensity-Modulated Radiation Therapy(IMRT) enables the delivery of a highly-conformal dose of radiation to the target volume. This is achieved by modulating the intensity of radiation throughout the patient, improving the normal tissue sparing compared to conventional radiation therapy. Because of this, an increased dose of radiation can be delivered to the tumour using IMRT[25].

Conventional treatment-planning is performed using a trial and error approach. For example, a radiation treatment field is placed on the patient and the computed dose is evaluated by the radiation oncologist. If the dose is not acceptable, a new field or other modifications are made. This trial-and-error process is repeated until an acceptable dose distribution and coverage is achieved. This is time-consuming and does not result in an optimal plan.

IMRT combines two advanced concepts to deliver dose distribution that conforms tightly to the target [23]:

- Non-uniform intensity of the radiation beams
- Inverse treatment planning via computerized optimization

1.2.1 Tumour Localisation

To perform inverse planning, a CT image is first acquired. Critical organs and the tumour volume is then contoured on this image. The clinical target volume(CTV) is the gross tumour volume plus all areas that might be at an increased risk of subclinical involvement. The CTV for early stage prostate cancer includes the prostate gland and may extend into the seminal vesicles.

The planning target volume (PTV) is obtained by geometrically expanding the CTV

[17]. During planning, the PTV will be covered by the prescribed dose. This will ensure that the CTV receives the full dose even if there is motion between the planning CT geometry and the geometry during IMRT delivery. Such motion can occur due to set up errors and anatomical movement during organ filling. At Juravinski Cancer Centre, $PTV = CTV + 10mm$ expansion in all directions except in the posterior where the expansion is $7mm$.

The next phase of treatment planning is to contour the organs at risk(OAR). These healthy organs or tissues are in close proximity to the treatment volume and sensitive to radiation. For prostate cancer, the OARs are the bladder, rectum and femoral heads. Figure 1.1 shows this anatomy for one patient. Once the anatomy is contoured, the planner places beams to strategically irradiate the volume of interest and defines the appropriate doses to be achieved in the targets and OARs, including their relative importance factors.

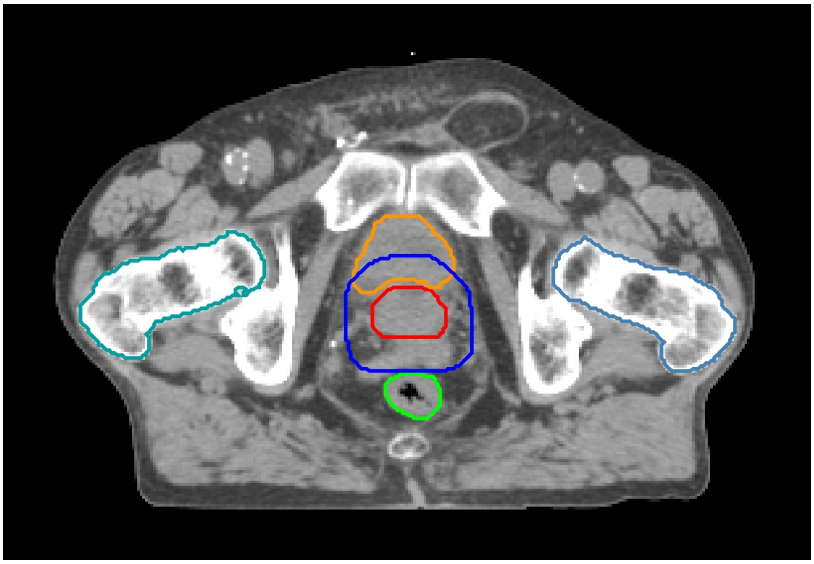


Figure 1.1: This is a transverse slice of a CT image. The coloured contours are as follows, red: CTV, blue: PTV, orange: bladder, green: rectum, remaining two regions: right and left femoral heads

1.2.2 Intensity Modulation

Variable radiation intensity is generated by subdividing each beam into many beamlets, each with an individual intensity level. This enables the delivery of a complex dose distribution. Figure 1.2 shows an example cross section of a beam, where the lowest intensity is shown in black and highest in white. This example demonstrates the complexity of the delivered intensities. Typically, low intensities correspond to sensitive tissue that should receive minimum amount of radiation and the high intensities correspond to the location of the tumour.

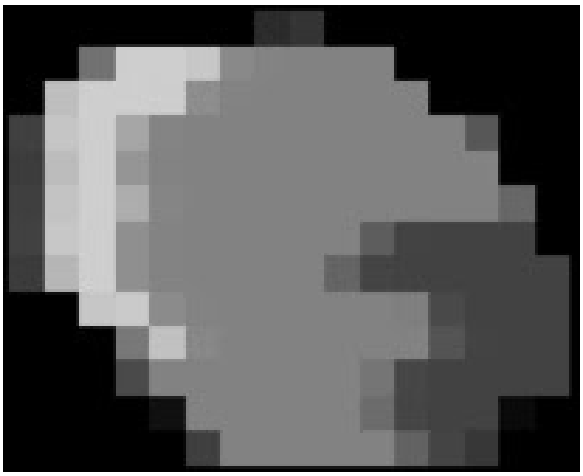


Figure 1.2: IMRT fluence map, divided into multiple beamlets with variable intensities. White and black corresponds to high and low fluence, respectively.

The use of several beams, each directed at the patient from a different angle, can build up a precise conformal dose distribution. Figure 1.3 illustrates the dose distribution achievable with IMRT.

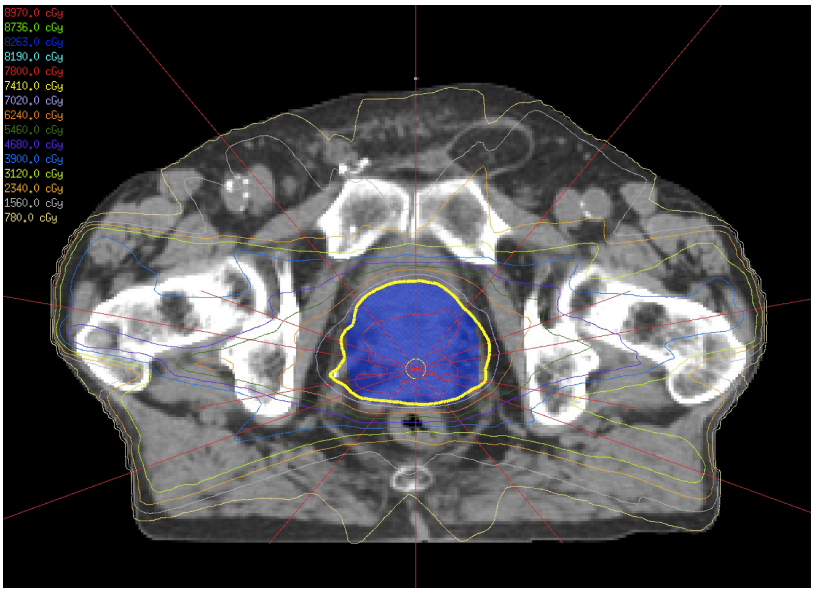


Figure 1.3: IMRT dose distribution. The blue colorwash is the PTV. Red lines indicate the beams, each directed at the isocentre(yellow point), Various isodose lines are shown corresponding to the doses indicated in the top left corner

1.2.3 Inverse Planning

As already mentioned, one important problem with conformal radiotherapy is that it is forward planned. The planner decides the number, shape, and orientation of the beams. Clearly, the outcome of this approach depends heavily on the planner's experience level. Furthermore, there is a limited number of degrees of freedom that a human can manually optimize. The result is a relatively simple plan with at most three-four beam directions and poor target volume conformality. Additional parameters can be included, but increased complexity increases planning time.

In inverse planning, a computerized optimization technique will determine the optimal treatment fields. For example, the Pinnacle Treatment Planning System(TPS)(Philips,

OAR	Dose to 30% of Volume(Gy)	Dose to 50% of Volume(Gy)	Dose to 5% of Volume(Gy)
Rectal Wall	< 71	< 53	_
Bladder Wall	< 71	< 53	_
Right Femoral Head	_	_	< 53
Left Femoral Head	_	_	< 53

Table 1.1: OARs dose constraints used for prostate IMRT at Juravinski Cancer Centre

Amsterdam, Netherlands) employs a simulated annealing heuristic to optimize beam intensities [10]. The planner provides the beam directions, prescribed tumour dose and limiting doses for normal tissue. At the Juravinski Cancer Centre, seven beams with angles $\theta = \{0, 40, 80, 110, 250, 280, 320\}$ are used for prostate cancer. The prescribed dose for early stage cancer is typically 76 to 78 Gy in 38 or 39 fractions. An acceptable plan delivers $\geq 95\%$ of the prescribed dose to $\geq 99\%$ of the PTV with the max dose $\leq 105\%$. The OAR dose limits employed at the Juraviski Cancer Centre are shown in table1.1. Having these inputs, the computer system adjusts the beam intensities to find a configuration which best matches the input parameters. The dose distribution is assessed and input parameters modified if the plan is not acceptable. The process is also time consuming but enables simultaneous optimization of many parameters. With experience, total planning time is reduced as centres develop standard procedures.

The optimization process maybe summarize as follows. Each beamlet that is traced through the patient produces an initial dose distribution. A small change is then made in the weight of a single beamlet and if it improves the distribution, it will be accepted. This process is repeated for all beamlets during a single iteration and should result in an improved plan. The iterative process is repeated for many cycles until improvements in the distribution are bellow tolerance. At this point optimal intensity across each beam that

produces the specified dose distribution was obtained.

As discussed, in IMRT inverse planning, fluence profiles are optimized to produce a desired dose distribution. These profiles are then converted to machine deliverable settings, with a possible degradation of plan quality. With the direct machine parameter optimization approach used in Pinnacle, machine settings are produced directly within the optimization process. Therefore, there is no need for conversion, filtering, weight optimizations or other kinds of post processing, and the plan quality will not degrade [10].

1.2.4 IMRT Delivery

A linear accelerator equipped with a multi-leaf collimator (MLC) is used for IMRT delivery. The MLC consists of movable tungsten leaves that can reduce the dose to 2% of the unshielded field. Figure 1.4 illustrates an example MLC configuration used during a prostate cancer case. At Juravinski Cancer Centre the Varian (Varian Inc, Palo Alto, USA) Millennium 120 MLC is used. The system consists of two banks of 60 leaves. The central 40 leaf pairs are 5*mm* wide for increased precision, while the remaining 20 pairs are 10*mm* wide. Leaves are positioned under computer control to create an irregular field that conforms to the tumour shape and shields normal tissues.

During delivery, the gantry rotates to each of the seven angles around the patient . At each angle, the MLC leaves are moved to create the first arrangement(control point) and the linear accelerator beam is turned on. The leaves are then repositioned and the second control point, or segment, is delivered. This process is repeated for all planned apertures and the treatment continues with the next gantry angle. The positions of the MLC and the amounts of radiation delivered by each control point are the same as obtained during treatment planning.

1.2.5 Tumour and Patient Movement

The PTV ensures acceptable coverage of the CTV throughout the treatment and as a result, takes into account all internal organ motions. The PTV margin also allows for daily variation in patient setup and machine precision. Large margins will ensure coverage but over-irradiate OARs while small margins will spare OARs but increase the risk of under-dosing the CTV. The appropriate choice of margin is especially important in IMRT due to the high dose gradient created away from the PTV. Any internal motion leads to significant CTV under-dosage and increases the chance of treatment failure.

Studies on prostate cancer have shown that the prostate can move up to 2 cm between treatments and the position is dependent on rectal filling [16]. Therefore, prostate IMRT is delivered under image guidance. At Juravinski Cancer Centre, two orthogonal projection images of the patient are acquired following patient setup for treatment. Radiopaque markers in the prostate gland are easily identifiable on the radiographs and their locations reconstructed in 3D. Marker locations on treatment are compared to the locations observed during treatment planning. Any difference indicated CTV motion. This is corrected by shifting the patient on the treatment couch in 3D and delivering the original treatment as planned. The limitation of this approach is the fact that deformable organ motion in the patient is corrected using a simple rigid couch shift.

Another approach involves the use of a linear accelerator with an on-board CBCT system. In this scenario, many projections of the patient are acquired and reconstructed into a 3D image. The CBCT image is then registered with the reference image, either manually or automatically. Again the problem is that a rigid couch shift is used to correct CTV motion caused by the internal anatomical deformations.

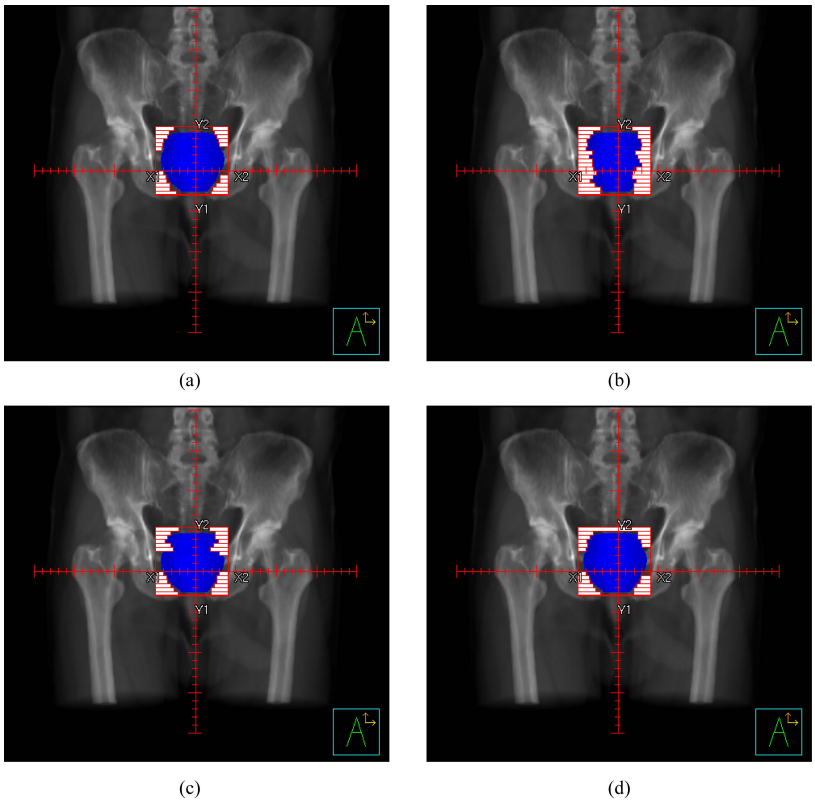


Figure 1.4: MLC configuration for the anterior-posterior beam(angle 0°)used during a prostate cancer case at Juravinski Cancer Centre. The blue surface is the PTV, white bars represent MLC leaves.

1.3 Thesis Goal

As discussed, the current approach to IMRT involves setting the patient for treatment, imaging(e.g. CBCT), rigidly shifting the couch and delivering the original treatment plan. The problem is that internal patient motion is non-rigid. Currently, the incompatibility between the anatomical deformations and rigid couch corrections is addressed by increasing the size of the PTV margin. This leads to an increased dose to sensitive organs. A method for

correcting IMRT treatments in response to anatomical deformation would enable the use of smaller PTV margins. This would reduce the volume of normal tissue exposed to high doses.

Daily motion can be accounted for by deformably registering the planning CT with the daily CBCT. The resulting transformation would then be used to obtain the daily representations of the PTV and OARs[26]. Once the contours corresponding to the current treatment fraction are known, a full IMRT re-optimization procedure can be undertaken. Therefore the goal of this thesis was to to develop a fast technique for re-optimizing a prostate IMRT plan based on the motion observed on daily CBCT imaging.

To achieve this goal, a linear programming approach is used in this thesis to minimize the number of MLC leaf movements, hence, increasing the speed of the optimization process. A fast re-optimization method is necessary since this will occur after the patient is setup for treatment. Thus computational efficiency is the primary concern as following imaging, the patient anatomy will continue to change.

To perform this development, data from the treatment of a prostate cancer patient were analyzed. The data comprised a planning CT and CTV, PTV and OAR contours. As discussed, there is no precise control over the internal anatomical motion. The proposed linear programming-based re-optimization method for prostate IMRT treatment planning, is meant to provide fast and precise control over treatment targeting on a daily base. Treatment modification in response to daily motion enables the use of a smaller PTV, reducing dose to nearby OARs and therefore, decreasing the probability of side effects.

Chapter 2

IMRT Optimization

2.1 Introduction

As mentioned in the first chapter, the goal of IMRT is to deliver the prescribed dose to the tumour. At the same time, the amount of radiation that is absorbed by healthy tissues should be limited. This is complicated by the fact that patient anatomy changes between treatment fractions. This thesis addresses the problem of re-optimization of prostate cancer IMRT in response to anatomical motion. As discussed, IMRT at the Juravinski Cancer Centre is delivered using a set of seven beam angles, $\theta = \{0, 40, 80, 110, 250, 280, 320\}$. To re-optimize the treatment, fluence maps for each of these seven directions have to be obtained for each treatment fraction.

A fluence map consists of many beamlets intensities and results in a complex dose distribution in the patient. Determining this map is difficult, since the PTV must receive the prescribed dose, while dose to critical organs should be minimized. Each critical organ has a specific tolerance level for the amount of radiation it can receive, and based on these levels, dose-volume constraints(DVCs), are defined. Usually DVCs are stated as a percentage

describing the portion of the organ that is allowed to receive a higher dose.

The DVCs represent the goals of the radiotherapy. They are used in an optimization procedure to obtain beam fluences that best achieve these goals. The fluence map may be represented mathematically using $w_{\theta,t}$, where θ is the beam angle, t is the spatial offset from the beam edge, and w represents the amount of photon fluence to be delivered. This notation represents a simplification of the $2D$ fluence map into $1D$ and therefore the reduction of the $3D$ patient to a single $2D$ slice. This simplification was considered acceptable given the initial development work described in this thesis.

Finally, having fluence maps $w_{\theta,t}$ calculated, a conversion into MLC leaf positions is required. For example, high fluence corresponds to MLC leaves staying open for a longer duration of the beam on time. The process of converting the fluence maps into the opening and closing movements of leaves is called leaf sequencing.

One last component required to re-optimize an IMRT treatment is the method of computing dose to the patient given the current beam fluence maps. This ties the parameters of the problem ($w_{\theta,t}$) with the DVCs.

2.2 Dose-Volume Constraints

Multiple beam angles are used in IMRT to concentrate the dose in the PTV, while sparing critical organs as best as possible. Despite this multi-beam strategy, non-negligible doses are delivered to healthy tissues. As a result, one important objective in this problem is to control the volume of healthy tissues that will be exposed to high doses. DVCs employed at the Juravinski Cancer Centre for prostate cancer are shown in table 1.1. The minimum dose constraint for PTV ensures the CTV will be irradiated to the prescribed dose and therefore ensures sufficient number of tumour cells will be erradicated. The maximum constraint on

the PTV avoids tissue necrosis in this region. This is important since, despite containing cancerous cells, the PTV also contains functional tissue. DVCs for rectum, bladder and femoral heads keep the risk of serious side effects below 5%. Possible side effects include rectal fistulae, bladder consticiton and bone necrosis. [1]

2.3 Fluence Map Optimization

The first step in IMRT fluence map generation is to discretize the beam aperture for each angle into many beamlets. The fluence is represented as $w_{\theta,t}$ for each angle θ and beam edge offset t. Ideally, one would set the size of the beamlet to the width of the MLC leaf (in the direction perpendicular to t) and to 1mm along the t axis. One mm would be appropriate since the MLC specification allows the leaf to be positioned within 1mm of the desired location. This would result in a beamlet size of 5mm (leaf width) $\times$ 1mm for the Varian Millennium 120 MLC. However such a dense representation of the fluence map results in significantly increased dose computation times. Typically, the beamlet is $5 \times 5mm^2$ during optimization[10] to reasonably balance computation speed with complexity.

The number of elements also depends on the number of gantry angles and the geometry of the patient. If it is detected that a ray from the source to a point in the patient does not have a significant effect on the treatment region (i.e. no intersection with the treatment region), it can be ignored during discretization. The total number of beamlets in a fluence map optimization problem is n, the sum of all beamlets from all gantry angles. The beamlet intensities that are to be obtained can be denoted using vector $w \in R_n$ instead of using $w_{\theta,t}$ for each beamlet. Also, since intensities cannot be negatives, $w \in R^n_+$.

In addition to discretizing the beam aperture, the patient must also be discretized. This three-dimensional volume of interest contains the PTV and nearby healthy tissues. As

already discussed, for the initial development discussed in this thesis, only a single $2D$ slice of the patient is considered(fluence maps are 1D). This $2D$ slice is a transverse section of the patient(slice axis is perpendicular to the patient's long axis). An example of a transverse slice obtained using CT imaging is shown in Figure1.1. This $2D$ slice is broken down into elements called voxels. Typically, the size of such an element is $2.5 \times 2.5mm^2$ for dose computation, and in general, differs from the voxel size of the CT image. Each voxel has coordinates $(i, j) \in R^2$ and the total number of voxels is m. The doses computed for each voxel are denoted by vector $d \in R^m_+$.

2.4 Dose Calculation

A common approach in IMRT for computing dose at each voxel is the pencil beam algorithm denoted as[11],

$$d_{i,j} = \sum_{\theta,t} D_{\theta,t,i,j} w_{\theta,t}, \tag{2.1}$$

where $D_{\theta,t,i,j}$ represents the amount of dose ($d_{i,j}$) deposited in voxel i, j per unit weight of beamlet θ, t. The values, $D_{\theta,t,i,j}$ are normally obtained by convolving the photon fluence per beamlet weight with a dose deposit kernel(units of Gy/ fluence). There are additional corrections that are required such as the decrease in fluence with increasing distance from the source, changes in fluence due to photon attenuation and variability in fluence across the beam due to linear accelerator factors(due to design, the fluence corss section of an open beam is not uniform). Figure 2.1 shows how each element $D_{\theta,t,i,j}$ relates a fluence map beamlet with a voxel in the discretized region of interest.

It is been assumed that D has no zero rows or columns. This means that all voxels receive some nonzero amount of radiation and every beamlet affects at least one voxel

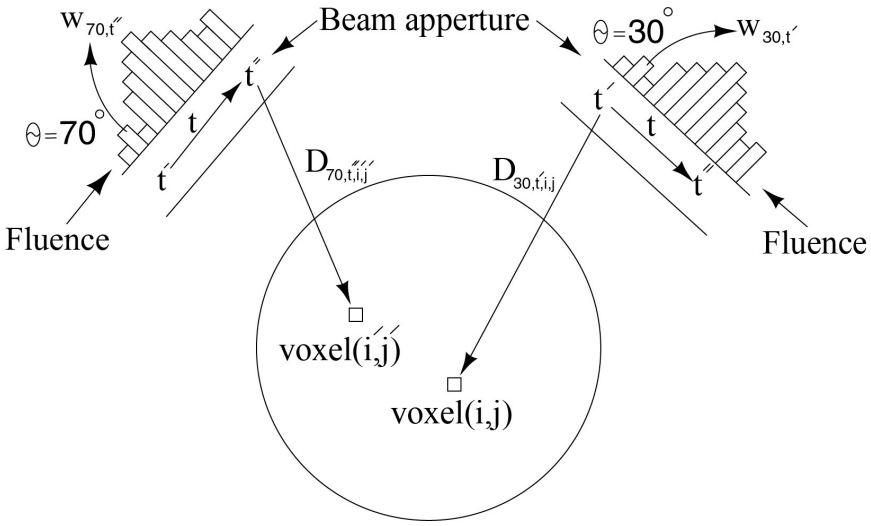


Figure 2.1: Fluence map beamlet and voxels in the discretized region of interest. The arrows labeled $D_{\theta,t,i,j}$ do not indicate a direct path of radiation but rather that the off-axis contribution is due to photon-electron transport and radiation scatter.

dose. Typically, $m \gg n$ with m on the order of 10^5 or larger and n around 300. Note that $D_{\theta,t,i,j}$ are always positive. Typically, for the prostate cancer IMRT, D would be very dense due to significant scattering of high energy photon beams.

The pencil beam dose calculation model can be considered as a first- order approximation. Radiation absorption as a function of the beamlet intensities can be modelled with linear Boltzmann transport equations [13]. Solving these equations is expensive, complicating implementation in the inner-most loop of an optimization procedure.

Monte Carlo photon transport is another accurate dose computation technique. However, this method is very slow prohibiting implementation in IMRT optimization. Some commercial treatment planning systems employ dose calculation engines with different levels of accuracy. In this way, dose calculations in early iterations, which are usually not required to be highly accurate, can be made less expensive [28], while more accurate (and more expensive) schemes can be used in later iterations. For example the Pinnacle

treatment planning system employs the convolution superpostion algorithm. In this approach, dose is computed by convolving the total energy released per unit mass (TERMA), which depends on photon fluence and patient attenuation, with a local dose deposition kernel, which is precomputed using Monto Carlo simulation. The collapsed cone convolution implementation in Pinnacle speeds up computation by collapsing the energy within some specified solid angle into a line[2].

2.5 Current Research

The fluence map optimization problem can be viewed as an inverse problem as a desirable dose distribution is required and one tries to determine the beamlet intensity vectors that give this distribution. There are different ways to formulate this into an optimization problem using various objective functions. At present, a common formulation is the "weighted least squares" model. If it is needed to have an objective function for each anatomical structure, then this problem would be a multi-objective optimization problem (for example [3], [12]). For multi-objective problems it is more likely that a single function does not exist but rather a finite set or even an infinite set of optima exist. The problem is thus transformed from an optimization problem to a decision making problem. Even if there is such a utility function, it is difficult to find this solution, since it is unknown what parameters and weights or even objective functions should be used to find the optimal beam fluences. The problem is that treatment planning systems try to be as simple as possible and having a large number of solutions is something that makes the planning more difficult.

Due to the difficulties in solving multi-objective optimization problems, the common

approach in IMRT fluence map optimization is to use a weighted least squares fitting strategy. There are many variations, but the typical form of the weighted least squares formulation is as follows [31].

For each voxel (i, j), try to fit the calculated dose value $d_{i,j}$ to some desirable value $b_{i,j}$. For a target voxel, this desirable value is the prescribed dose. If a calculated dose for a voxel representing healthy tissue, is less than its desirable value, then the corresponding error term is set to zero. This way, only doses higher than their desirable values are penalized. Each target and critical structure, is assigned a weight parameter that represents the priority of fitting its calculated doses to the desired dose. Different tradeoffs between structures can be made by adjusting these weights. For example, suppose there are four structures S_l , $l \in \{0,1,2,3\}$, where S_0 is the target structure and the other three are healthy tissues. The objective function in the weighted least squares formulation takes the form:

$$f(d) = \sum_{l=0}^{3} \omega_l f_l(d) \tag{2.2}$$

Where ω_l is the weight assigned to structure l and $f_l(d)$ are the objective functions for each structure defined as:

$$f_0(d) = \sum_{(i,j)\in S_0} (d_{i,j} - b_{i,j})^2, \tag{2.3}$$

$$f_l(d) = \sum_{(i,j)\in s_l} \max(0, d_{i,j} - b_{i,j})^2, l = 1,2,3 \tag{2.4}$$

In this case, the resulting error function $f(d)$ is a convex piece-wise quadratic function of d. The goal is to minimize $f(d)$ subject to the non-negativity of the doses $d_{i,j}$ [31]. Clearly, solutions to this weighted least squares model vary with different priority of

weights.

The challenging aspect of weighted least squares modelling is that it does not guarantee the solution will satisfy the prescribed DVCs. One approach to enforce DVCs is to by trial and error identify a set of weights that results in an acceptable solution. This trial and error process is time-consuming and operator dependent.

In addition to manipulating the weights, the other approach is to add some penalty to the weighted least squares objective to encourage but not impose the dose-volume constraint feasibility. For example[27], proposed a method to quickly modify a treatment plan in adaptive radiotherapy. The method is based on a Cimmino-type algorithm in linear programming. The fast convergence speed is achieved by over-relaxing the algorithm relaxation parameters, so that the over-relaxed Cimmino (ORC) algorithm can effectively approximate an unconstrained re-optimization process in adaptive radiotherapy.

However these penalty terms are non-convex and one has to deal with local minima. To address this problem, some formulations(e.g.[29]) add the option of using stochastic global optimization techniques to escape from the local minima such as genetic algorithms and simulated annealing. Their clinical objectives have been defined in terms of dose limits or in terms of limits on volumes receiving certain specified dose. The advantage of variant of Newtons methods they have adopted for optimizing the objective function is its speed of convergence.

However, the Pinnacle treatment planning system, uses a sequential quadratic programming algorithm instead of weighted least squares model. The approach is a gradient based algorithm. The process starts with pencil beam algorithm and simple intensity fluence map optimization. Results are then converted to leaf positions, hence then the parameters of the problem are segment weights and leaf positions. In each iteration, these are tweaked and

the change in the objective function is observed[10].

Quadratic objective functions have been accepted as a standard due to some important benefits. For example, as it is stated in [22], it can be proven that simple least square objective function will not result in a local minima. Despite benefits like what mentioned, there are also disadvantages for this types of functions which makes clinical planning based on them, more difficult and time-consuming. The most important weakness is mainly due to the fact that the weight factor in these functions, w, does not have any clinical meaning which makes it completely arbitrary. As a result it is required to try multiple clinical planning to select the final plan.

On the other hand, the gradient-based algorithms are able to deliver the optimum plan quite fast and satisfactory. Although these functions might result in local optima, they demonstrate better overall performance and seem to be a better choice for beam intensity optimization due to their fast response.

The fluence map optimization problem also attracts researchers attention in mathematical programming to formulate the problem into linear or mixed-linear program. For example [20] presents a novel LP model for solving the fluence-map optimization problem in IMRT treatment planning. Their approach to this problem overcomes the apparent limitations of LP by using linearizations of nonlinear convex penalty functions. In addition, they employ a new type of DVC that bounds the mean value of the tail of the differential dosevolume histogram of a structure. Also [9] presents computational approaches for optimizing beam angles and fluence maps in IMRT assuming that the number of angles to be used for the treatment is given by the treatment planner. In their approach mixed integer programming model and a linear programming model are used to find an optimal set of beam angles and their corresponding fluence maps. The mixed integer programming

model is solved using a branch and bound method while the linear programming model is solved using the interior point method.

A stated in the first chapter, in IMRT inverse planning, the planner provides the prescribed tumour dose and limiting dose for OARs and then a computerized technique will be used to optimize the fluence distributions in the beams. The problem arises when this resulting dose can not directly be delivered to the region of interest and a this distributions need to be converted to the control points. This conversion is done without taking the physicians preferences into account, and the resulting dose distribution may not be desirable. One method proposed to solve this problem is to calculate fluence directly from a set of control points described by leaf positions and segment weights. This method that is called direct machine parameter optimization (DMPO), basically skips one intermediate step and directly optimize machine parameters based on physician suggested dose distribution. However, DMPO has a greater degree of nonlinearity to be a convex problem. This will make this method subject to numerous linear constraints. As a result, a non-DMPO method was selected for the work described in this thesis.

Chapter 3

Proposed Optimization Approach

3.1 Introduction

In IMRT not only is the shape of the beam controlled, but combinations of open and closed MLC leaves modulate the intensity as well. In this section, a linear programming(LP) approach is proposed which allows optimization over beamlet fluence weights. The proposed method is a multi-step process. The difficulty is that the full problem is mixed-integer since the solution (fluence weights) must be realizable using a certain number of leaf movements.

If it is been chosen to vary the number of angles the problem becomes large and also not convex. In this problem, seven coplanar beams intersected at the iso-centre have been employed. Similar to the previous approaches, beamlets variables have been introduced that are not in the original problem, because this relaxed problem will be convex. Having a convex problem makes it much easier to solve. However, one drawback of using this approach is that the resulting solution will not be a solution to the original problem. What will happen is that it will give a solution that could be used only if there is no limits on the number of leaf positions. A tradeoff between solvability and deliverability exists. To

understand why the relaxed problem is so much easier to solve, consider that if the leaves are in a certain position, then changing the leaves only affects the tissues that are near the edges of the beam with little effect elsewhere. So if you look at it from a point of view of a gradient-based method, what derivatives indicate is that there is no way to change the dose for a voxel which is not near the leaf edge. Of course, with eight control points per beam, there exists a good chance that all areas in the PTV are sufficiently close to a leaf edge for the gradient methods. However, the nearby control point may not be the best to modify.

With beamlets, the dose at every voxel is sensitive to changes in the appropriate beamlet. It should be noted that even after this conversion, it still may be impossible to solve the problem, such that sufficient radiation is delivered to the tumour and OAR limits are not violated. However, even if there is no solution to the problem, the optimization results may tell how to replace constraints with penalties to find a solution.

Having relaxed the problem using beamlets and possibly penalties in place of constraints, it would be a rapidly solvable problem, but one which is not necessarily deliverable. Now that the radiation has been converted into beamlets, the next step is to translate these into leaf positions. Basically, try to find a beamlet solution which is easier to turn into leaf positions. The way doing that is that it's been observed if a set of beamlets comes from leaf positions, then the resulting fluence map would be piece-wise constant and the number of steps going up or down would be, at most, two times the number of leaf positions. This is because if by moving the leaves between every shot, will end up with two fluence steps, one for the leaf to the left and one for the leaf to the right, assuming that both left and right leaves never return to the same position. Although requiring at most a certain number of fluence steps would result in the mixed-integer problem tried to avoid, introducing regularization would favour fewer steps. For some choices of regularizer, the convex property of

the relaxed beamlet problem will be preserved, while trying to find the solution which is closer to something that is realizable with leaves.

Another way of saying a small number of steps is wanted is to say that a piece-wise constant fluence map is desired. Such a solution can be favoured by penalizing changes in neighbouring beamlet intensity using a difference norm, thereby indirectly penalizing required changes in the leaf positions. To make the problem as easy as possible to solve, an ℓ_2 difference norm can be used which is sum of the square of differences because it would be quadratic penalty. One problem with this approach is that it will penalize the large fluence changes more than small changes. There is no reason to to do this, because changing leaf positions is what is important and not the amount of the intensity change. Alternatively, the ℓ_1 difference norm can be used. This way, the total of all intensity changes will be penalized independent of how they spread out.

The last point to note is that although there is no constraint on the total dose, there is no reason to ever have more dose from any angle than the target dose in the tumour (after adjusting for attenuation and loses due to beam spread). A stronger constraint is chosen based on the ℓ_1 difference norms. If the radiation at a given angle is formed by nested(leaves moving monotonically toward centre) leaf positions, which corresponds to there being two intervals of beamlet intensities: a monotonically increasing interval followed by a monotonically decreasing interval, then the ℓ_1 difference must be less than two times the intensity required to produce the maximum dose in the tumour from the beam. If, however, the radiation has multiple maxima, corresponding to non-nested leaf positions, then this constraint may be stricter than necessary. So this one constraint which has been imposed, enforces both maximum radiation per angle and prevents excessive variation. Most importantly, it is convex. It is actually linear, so by formulating other constraints in

terms of simple inequalities, it still can be solved using a linear solver.

Although the l_1 and l_2 norms have computational advantages, being linear and quadratic, respectively, they are not the only norms available for penalty construction. For example, an $\ell_{1/2}$ difference norm can be added. The ℓ_p norms are convex between ℓ_1 and ℓ_∞, while those with the order less than one, are not convex. The non-convex norms have the property that they are smaller for a smaller number of large fluence steps compared to small fluence steps. If the number of steps was sufficiently reduced, the large steps would correspond to the desired number of leaf positions without actually imposing it. The rounding step could be avoided to produce piecewise-constant intensity profiles. To summarize, the $\ell_{1/2}$ norm should favour solutions with small numbers of discontinuities, but it is non-convex and may require significantly more computation.

Non-convex objectives and constraints may cause solvers to get stuck near an unacceptable local minimum. The $\ell_{1/2}$ norm can be added if it is in the right neighbourhood of solution, which can reasonably be assumed to be the case after solving a convex problem first. An additional problem is that adding a new constraint will reduce the feasible region, possibly making the solution to the convex problem infeasible. Adding a penalty will not cause this to happen. Since the relaxed problem is still being solved, it may still be infeasible in terms of the mixed-integer problem. However, by solving the convex problem and then adding another norm, such as $\ell_{1/2}$ difference norm, hopefully it will cause the solution to get closer to the desired solution, realizable by a finding number of leaf positions.

The new norm(s) would be in addition to the ℓ_1 difference norms. Because the ℓ_1 difference norms remain, new unwanted solutions are not allowed, because, for example, more radiation than the target dose is not wanted for the tumour anywhere.

Eventually, in the last step, which is left as future work, the resulting fluence maps

must be rounded, since a limited number of leaf movements is allowed to deliver these fluence maps. One can think of this step as a more restricted version of the original problem including integer constraints, but with the addition of some inequalities or objective functions which are defined to us as continuous functions. There are two approaches to rounding relaxed problems to obtain integer solutions:

Random Rounding is looking at all of the variables in this problem that are supposed to be integers and start by rounding them all down or up randomly. So if there are N variables, there would be 2^N different ways of rounding up and down. That might give us a better solution, but it is too time consuming to test all possible roundings, or even most of them.

Branch and Bound is a way of solving general mixed-integer problems without testing every rounding possibility of the relaxed continuous problem. In this method, first one variable is picked and one say different integer values are going to be defined for that one variable and going to solve the continuous problem in the other variables keeping that one fixed. It might be possible to show that for some values, the solution can never be as good as some feasible solution that already exits, so that possibility is eliminated, and the others are kept. For each value kept, another variable is chosen and the integer values it could take is considered. Continuing in this way, all possibilities are tested or eliminated.

Applying the latter to a leaf identification method, one would pick one fluence variable at a time, fix it and then (re)solve the smaller continuous problem containing one less variable, thereby constraining the position of one leaf at a time. If solutions are needed quickly, it does not need to follow exactly this prescription for branch and bound. One could pick leaves directly effect the tumour or critical tissue doses, and use branch and bound, and

later resort to a few tests of random rounding. It would not be practical to pick all the possible leaves one at a time, try to fix in the leaves and then re-optimize and fix another leaf and so on. In this way, one could use the structure of the problem to round individual fluence values until the solution is realizable via leaves.

Although the proposed algorithm will not necessarily result in the optimum solution for most cases, hopefully, that the solution obtained by rounding is within a small fraction of the optimal solution. In practice, what the optimal solution to the mixed-integer problem is never known. However, for the re-optimization problem, when one have a solution from a "gold-standard" solver, and it is known that the patient has changed, due to inflammation, or natural body processes, one can re-solve both the original problem, and the deformed problem, using the gold-standard and the rapid methods. A comparison of the two solutions for the original case and then the deformed case would provide evidence that the proposed method obtains near-optimal solutions.

3.2 Proposed Approach

3.2.1 Implementation

Because an iterative solution method is being proposed(including some branch and bound steps), it is important to know the problem can be sovled very quickly. Convex continuous optimization problems are fast, but no equally fast. In order to try different formulations (notably, with different norms $\ell_{1/2}, \ell_1, \ell_2$), The basic model has been implemented in the modelling language AMPL(a mathematical programming language)[Bell Labs, Madison, Wisconsin], which can interface with multiple solvers, which support different types of constraints and objectives. The disadvantage to using AMPL is that it does not support

recursive data structures (e.g. trees) which would make it easier to implement recursive algorithms suited to final rounding, although AMPL does support iteration, and enumeration over sets which simplify the specification of operations research problems.

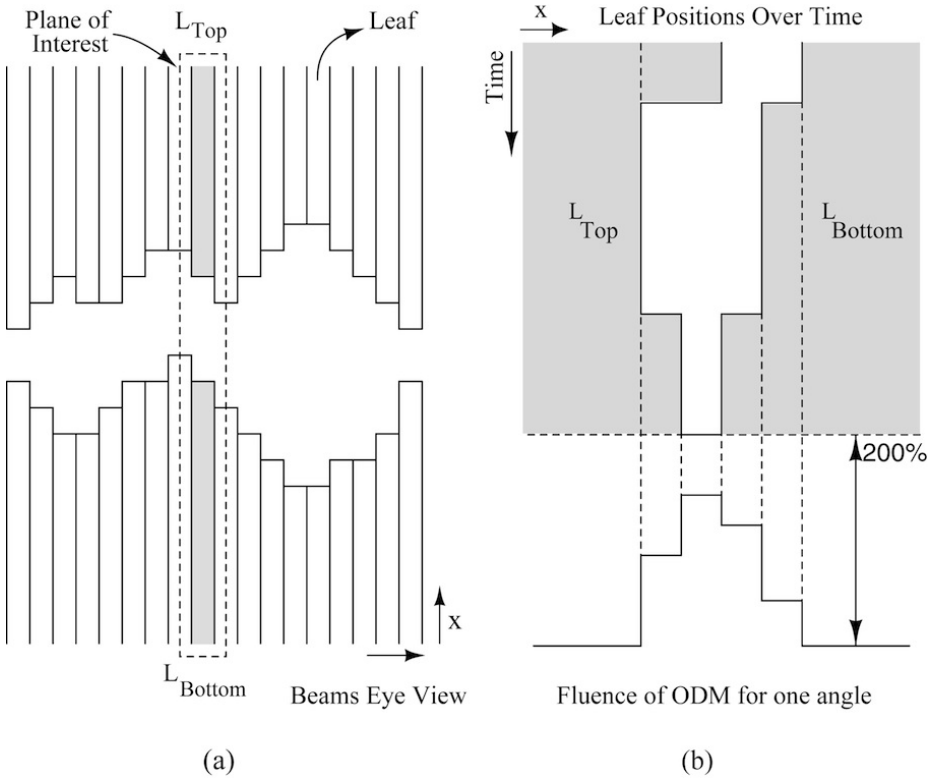


Figure 3.1: Panel (a) shows the beam's eye view with the irregular radiation aperture in the centre formed by several leaves. Panel (b) shows the positions of a single leaf pair over time and the resulting fluence profile.

As discussed, the ℓ_2 norm penalizes the large differences in the fluence. The ℓ_1 norm penalizes all the differences and the $\ell_{1/2}$ norm penalizes small differences. The equations for ℓ_2, ℓ_1, $\ell_{1/2}$, respectively,

$$\ell_2 = \| w \|_2 = \sqrt{\sum {w_i}^2} \tag{3.1}$$

$$\ell_1 = \| w \|_1 = \sum | w_i | \tag{3.2}$$

$$\ell_{1/2} = \| w \|_{1/2} = (\sum \sqrt{| w_i |})^2 \tag{3.3}$$

Note that the maximum of the ℓ_2 norm is also the maximum of the ℓ_2 norm squared, and, in fact, the square of the ℓ_2 norm is being used in optimization because it is a quadratic problem. Quadratic objectives are computationally easier to handle, as evidenced by the fact that many solvers only support quadratic objectives. Similarly, one would optimize with the square root of the $\ell_{1/2}$ norm, but this does not change the fact that $\ell_{1/2}$ difference norms are non-convex, so using only this type of norm is likely to result in a not good solution or no solution. The ℓ_1 norm does not need to be raised to a power, but it does contain a non-differentiable point at 0. Instead, the ℓ_1 norm can be replaced with an auxiliary variable and two inequalities, bounding the argument of the norm above and below. Thus problems with ℓ_1 norms are problem with linear constraints. Problems with only linear constraints are Linear Programs, which can be most easily solved among constrained problems.

3.2.2 LP Formulation and Variables

The possible use of multiple norms has been described. In this section, the formula for the ℓ_1-norm will be discussed in detail, because it was the most successful. Then, modifications to incorporate other norms will be discussed.

The general form of a linear problem is given by:

$$\begin{aligned} &\min c^T x \\ &s.t.\ Ax = b \\ &x \geq 0. \end{aligned} \tag{3.4}$$

Where $c^T x$ is the cost function to be minimized, $Ax = b$ and $x \geq 0$ are the constraints for the objective function and x holds the variables to be determined. Before declaring the constraints for the linear problem, one needed to define the set of variables and parameters applied for this problem. The variables are listed in the following description.

$w_{\theta,t}$ beamlet weight, proportional to photon fluence measured at a line of isocentre for a beam angle θ. t is the offset along the line.

$d_{i,j}$ dose, the radiation dose delivered to voxel (i, j)

$D_{\theta,t,i,j}$ finite size pencil beam dose distribution, represents the dose deposits in voxel i, j per unit weight of a beamlet θ, t

τ auxiliary variable, used to construct the ℓ_1-norm using only inequalities

where the relationship between the beamlets and dose in voxel (i, j) was described by equation (2.1) in chapter 2, which is restated here for convenience:

$$d_{i,j} = \sum_{\theta} \sum_{t} D_{\theta,t,i,j} w_{\theta,t} \tag{3.5}$$

In (2.1), $\theta \in \{10, 50, 90, 130, 170, 200, 340\}$ the set of gantry angles. Coordinates $(i, j) \in$ {(-53, -75), ..., (53, 75)}, are also voxel indices into the 2D slab representing the patient,

which has a 106-mm height and 150-mm width. For this development angles were modified from $\{0,40,80,110,250,280,320\}$ to $\{10,50,90,130,170,200,340\}$. Because seven angles are employed around the patient and, their directions are not critical from the optimization point of view, as long as they are well spread over 360°. In the future, the clinical angles should be used.

Now, the optimization problem can be defined, using the variables defined in this section. As already discussed the use of penalties, but not the objective. In fact, all of the clinical requirements are constraints, and the penalty is the only objective. The ℓ_1 norm is replaced by two inequalities, and the auxiliary variables $\tau_{\theta,t}$.

$$
\begin{aligned}
&\min \sum_{\theta,t} \tau_{\theta,t} \\
&\text{s.t. } w_{\theta,t} - w_{\theta,t+1} \leq \tau_{\theta,t}, \quad \forall \theta, t \\
&\quad -w_{\theta,t} + w_{\theta,t+1} \leq \tau_{\theta,t}, \quad \forall \theta, t \\
&\quad d_{i,j} = \sum_{t} D_{i,j,t,\theta} w_{\theta,t}, \quad \forall i, j \\
&\quad 0 \leq w_{\theta,t} \leq w_{max}, \quad \forall \theta, t \\
&\quad w_{\theta,t_{\max}} = 0, \quad \forall \theta \\
&\quad w_{\theta,t_{\min}} = 0, \quad \forall \theta \\
&\quad 95 \leq d_{ij} \leq 105, \forall (i,j) \in T \\
&\quad \sum_{\{(i,j)\in R\}} d_{ij} \leq M_r |R| \\
&\quad \sum_{\{(i,j)\in H\}} d_{ij} \leq M_h |H| \\
&\quad \sum_{\{(i,j)\in B\}} d_{ij} \leq M_b |B|
\end{aligned}
\tag{3.6}
$$

Additional indices on the constraints are:

T is the set of voxels belonging to the PTV;

R is the set of rectum voxels, with M_r maximum average dose;

H is the set of femoral heads voxels, with M_h maximum average dose;

B is the set of bladder voxels, with M_b maximum average dose.

Figure 3.2 illustrates a simplified tumour and critical structures around it. The tumour is coloured in red and sensitive tissues in green and blue, resulting in constraints

$$95\% \leq d_{i,j} \leq 105\%, \quad \forall (i,j) \in \{(4,4),(4,5),(4,6),(5,5),(5,6)\},$$

$$\sum_{(i,j)\in\{(3,4),(3,5)\}} d_{i,j} \leq 2M_{\text{blue}},$$

$$\sum_{(i,j)\in\{(5,4),(6,4)\}} d_{i,j} \leq 2M_{\text{green}}.$$

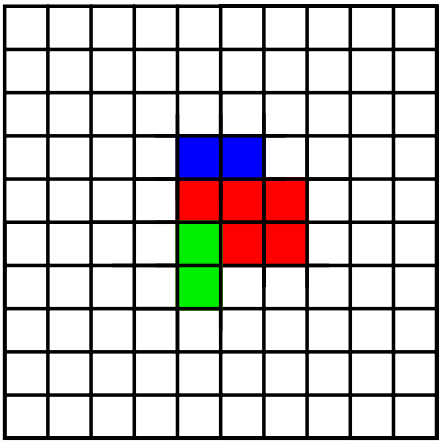


Figure 3.2: Simpler version of tumour and critical tissue sets

3.3 Dose Matrix

Since the propsed approach relies heavily on computed dosages, providing an accurate method for calculating dose will undoubtedly be an important research direction. Any questions of efficiency will almost certainly involve considering how the dose matrix D is computed and how accurately it is computed. For example, the current method ignores the effects of scattering or tissue attenuation.

A simple dose matrix is being used for these experiments which can be stored and rotated to obtain dose matrices for various angles around the phantom. Real people are, however, much more irregular and inhomogeneous which forces dose calculations to be even more expensive and makes the matrix D so large it typically cannot be stored ahead of time.

To circumvent this, many commercial IMRT planning systems calculate dose on the fly, requiring various levels of accuracy at different stages in the optimization to reduce computational cost. Although, it is not been considered how and when to compute approximations to D in this thesis, having a solid understanding of how this is done and how to improve its interface with the optimization algorithm will undoubtedly be a key factor in increasing algorithm performance.

3.3.1 Calculation Method

Radiation originates at the source. The position of the source is given by $(100 \times \cos(\theta), 100 \times \sin(\theta))$, because the source is 100cm away from the iso-center (where reference measurements are made). It has been treated as a point source and from the point source, rays diverge in all directions isotropically. As the radiation spreads out, the concentration of radiation will decrease. Since it is in three dimensions, the radiation decreases with one

over distance squared. Since the problem is modelled with beamlets, those beamlets are representing the radiation that would go through a plane at iso-center, that is perpendicular to the beam. The plane at isocentre is referred to as the opening density matrix (ODM). Figure 3.3 depicts this geometry.

Now, the dose is at each voxel has to be figured out. The photon fluence as a function of offset, depends on the ODM value but voxels are at different distances from the radiation source. So one needed to adjust the dose for voxels, which depends on the position of each voxel and its distance from iso-center. This can be figured out by solving a set of linear equations. It should be noted that the ray which leaves the source, goes through a position on the ODM before reaching the target voxel. The intersection of the source-target voxel line and the ODM can be obtained geometrically. This value can be corrected for inverse square law and then stored in the target voxel as the dose. This is a first order approximation of the dose that ignores scatter, beam attenuation, and the conversion of photons to electrons before dose is deposited in the tissue.

3.3.2 Geometric Factors

Figure 3.3 shows, the volume of interest in three dimensions (x, y, z). The isocentre is assumed to be at $(0, 0, 0)$ of the volume of interest and the source 100cm away from it. The ODM considered a (u, v) plane going through the iso-center. This plane which rotates in the same direction as the source, helps us finding dose values in other voxels in the volume of interest. Based on this, one needs the coordinates of (x, y, z) on the box of interest to be in terms of coordinates on the plane (u, v) which is the intersection of the beamlet with the ODM.

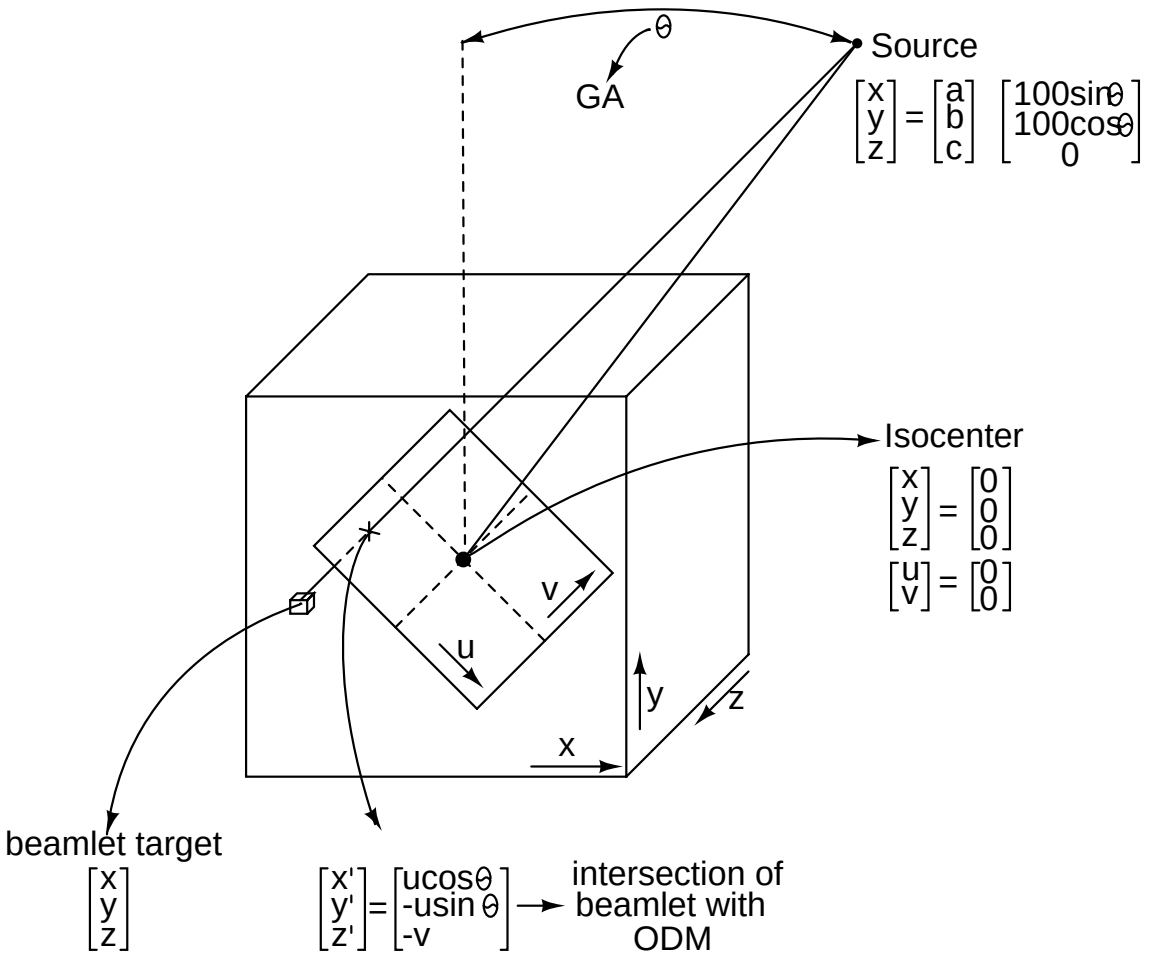


Figure 3.3: A simple view of ODM, source and region of interest which shows the conversion of $3D$ coordinates into $2D$ coordinates

The three points, (x, y, z), the ODM intersection and the source, are all on a line. Therefore the vector from the source to (x, y, z) is a scaling of the vector from the source to the

ODM intersection by the factor α as seen in the following equation,

$$\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} a \\ b \\ c \end{bmatrix} \right) \times \alpha = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} - \begin{bmatrix} a \\ b \\ c \end{bmatrix} \tag{3.7}$$

$$x' - a = (x - a) \times \alpha \tag{3.8}$$

$$y' - b = (y - b) \times \alpha \tag{3.9}$$

$$z' - c = (z - c) \times \alpha \tag{3.10}$$

α can be eliminated from (3.8),

$$\alpha = \left(\frac{u \cos\theta}{x - a} \right) \tag{3.11}$$

Then (3.11) is plugged into (3.9) and (3.10),

$$u = 100 \times \left(\frac{-\cos\theta(x - 100\sin\theta) + \sin\theta}{\sin\theta(x - 100\sin\theta) - \cos\theta} \right) \tag{3.12}$$

$$v = -z \times \left(\frac{u\cos\theta - 100\sin\theta)}{x - 100\sin\theta} \right) \tag{3.13}$$

Once (u, v) are known in terms of (x, y, z), the raw photon fluence at the target voxel can

be obtained using:

$$\text{intensity}_{(x,y,z)} = \text{ODM}_{(u,v)} \tag{3.14}$$

Since the coordinates (u,v) generated by the above equations, are not necessarily integers, bilinear interpolation is used to obtain $ODM(u,v)$.

Also the intensity has been corrected by considering the fact that radiation intensity will decrease on its path from source to the voxels using Inverse Square Law,

$$\text{ISL} = \left(\frac{100^2}{(x-a)^2 + (y-b)^2 + (z-c)^2} \right) \tag{3.15}$$

So equation (3.14) in more precise form is:

$$\text{intensity}_{(x,y,z)} = \text{ISL}_{(x,y,z)} \times \text{ODM}_{(u,v)} \tag{3.16}$$

The proposed method in this thesis is $2D(z=0)$. The intensity obtained for a pixel (x,y) will be used in (2.1) as below:

$$D_{\theta,t,i,j} = \text{intensity}_{(i,j,0)} \tag{3.17}$$

where i represents x, j represents y, t represents u, and θ dependence is explicitly included. It should be noted again that, this is a first order approximation that ignores beam attenuation, photon scatter, and conversion of photons to electrons.

Chapter 4

Validation

4.1 Input Data

To test the optimization program, data from the treatment of an anonymized prostate cancer patient were analyzed. The data comprised a planning CT, PTV and OAR contours. Figure 4.1 shows a screenshot of the anatomy with a 1cm grid overlaid. In this figure, the green circle at the bottom indicates the rectum, the blue oval is the PTV and the orange oval is the bladder. Also the two femoral heads are shown in blue/green colour. To simplify the problem, the volume of interest has been assumed as a rectangle. The goal is to deliver the prescribed dose to the PTV while all other organs receive the minimum dose. This is not a trivial task since all the organs surround the target area, and the bladder overlaps the PTV as shown in Figure 4.1.

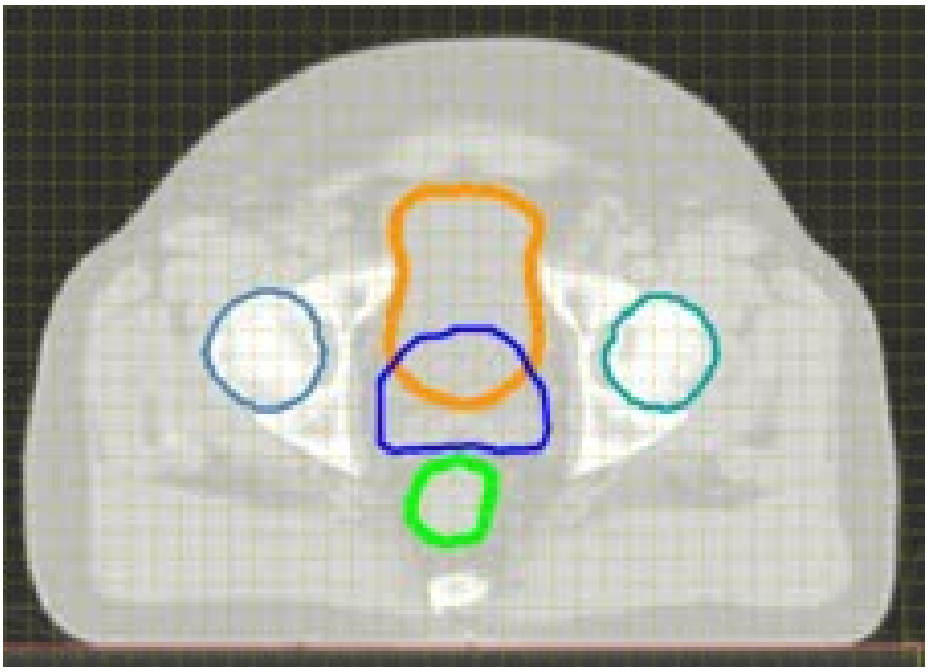


Figure 4.1: Transverse CT slice and contours corresponding to PTV(blue), bladder(orange), rectum(green) and femoral heads(green/blue). The volume of interest has been assumed as described in chapter 3.

4.2 Methods

As explained in the previous chapter, the requirements of treatment planning are captured as constraints in the linear programming approach. The constraints are specified as percentages of the target dose in the PTV. The PTV dose is constrained to the prescribed range of 95% to 105%, so that it would be effective in killing tumour cells. On the other hand, the rectum, bladder and femoral heads need to be spared. Since OARs and PTV have an overlap area, there is a trade-off between minimizing radiation dose to the OARs and maximizing the dose in the PTV at the same time. For this development, the mean dose in the

OARs was kept below 25%.

It should be noted that in this method of treatment, seven fixed angles are employed. As previously explained, in this optimization method, radiation angle will not be included as one of the parameters and, so only fluence maps representing the intensity of radiation coming from each angle will be reported as results. Figure 1.1 in chapter 1 shows radiation beams from seven clinically employed angles. Angles are $\theta = \{0, 40, 80, 110, 250, 280, 320\}$ which ensure the dose is concentrated in the PTV. A higher number of angles results in greater flexibility which makes the optimization problem's constraints easier to satisfy, but extend the treatment time.

4.2.1 Modelling Software

For modelling the optimization problem, NEOS online server is used [4, 8, 5]. It provides a web interface to different solvers. In this optimization two different solvers were used for the linear and the non-linear problem. The ℓ_1 problem is solved using MOSEK (software package for optimization problems, version 6.0) and the non-linear problem obtained by adding $\ell_{1/2}$ was solved using KINTRO (Nonlinear Interior point Trust Region Optimization, version 8.0) which are efficient and powerful environments for solving an optimization problem in AMPL format. To localize the contours, the geometry in Figure 4.1 was converted to a black and white image and then converted to a text file with several integer numbers demonstrating different tissue margins. This text file was loaded to the online server together with the AMPL modelling file. The computation of $D_{i,j,\theta,t}$ was also included in the AMPL code.

Case 1	Rectum is shifted down by 5mm, bladder size is reduced by 5mm
Case 2	Rectum is shifted down by 5mm, bladder is enlarged by 5mm
Case 3	Rectum is shifted up by5mm, bladder size is reduced by 5mm
Case 4	Rectum is shifted up by5mm, bladder is enlarged by 5mm
Case 5	Rectum is shifted left by 5mm, bladder size is reduced by 5mm
Case 6	Rectum is shifted left by 5mm, bladder is enlarged by 5mm
Case 7	Rectum is shifted right by5mm, bladder size is reduced by 5mm
Case 8	Rectum is shifted right by 5mm, bladder is enlarged by 5mm

Table 4.1: Eight different type of movements are indicated by eight case numbers.

4.2.2 Tissue Motion

Since the motivation for plan re-optimization comes from concerns about anatomical motion between treatment fractions, eight cases were optimized representing typical patient motion/deformation. These included both increases and decreases in bladder size (contour changed by 5 mm in the anterior/posterior direction), and displacements of the rectum by 5mm in anterior, posterior, left and right directions. As discussed in chapter 1, the maximum displacement of the prostate target is 20mm, so 5 mm was considered appropriate for this analysis. Deformations during treatment result from normal physiological function (including varying amounts of intestinal gas) and changes due to treatment (including swelling as a result of radiation damage). Table 4.1 enumerates the eight deformation experiments,

For all cases, the parameters for the optimization problem (3.6) were set up. The PTV and critical tissues were localized and also the information needed to define the constraints was extracted, namely the sets, T, B, R, and H for voxels in the PTV, bladder, rectum, and femoral heads, respectively. The modified data were sent to the NEOS online server .

4.2.3 Norm Analysis

In this section the effect of adding the $\ell_{1/2}$-difference norm as a penalty to the linear problem was investigated. As discussed, finding a small number of fluence steps is desired. So by adding $\ell_{1/2}$-difference norm, it is possible to get closer to the desired solution. Since by adding this norm as a constraint, the problem becomes non-convex, the ℓ_1 norm has been kept and the $\ell_{1/2}$ is added as a penalty to it. In theory, if current state is in a small neighbourhood of solution, addition of non-linear term should help to improve results.

4.3 Results

In addition to the one-dimensional radiation profiles for each angle, two-dimensional dose distributions are presented. Dose distributions, show the relative dose at all voxels in the region of interest. Both plots are expressed as percent of target dose in Gy, as were the constraints. The radiation oncologist sets the target dose in, however this does not change the problem from the point of view of evaluating this method, since relative dose is assumed to be proportional to Gy.

The plots are very useful for understanding the method and qualitatively evaluating it. Three numerical measures of performance are also presented. First, the calculated dosage at the the PTV and whether the radiation conforms to the target area. This parameter shows the effectiveness of this approach. Secondly, it is interesting to see whether the dosage at other body organs is below what is considered a safe limit. Finally the third parameter is the time that it takes for the solver to find an optimum solution. Basically this parameter shows whether this approach can practically be used for plan re-optimization while the patient waits in treatment position.

4.3.1 Original Data

Figure 4.2 shows the results obtained under the original patient geometry. As Figure 4.2 demonstrates, the LP approach conforms the dose to the PTV with little exposure of the critical tissues.

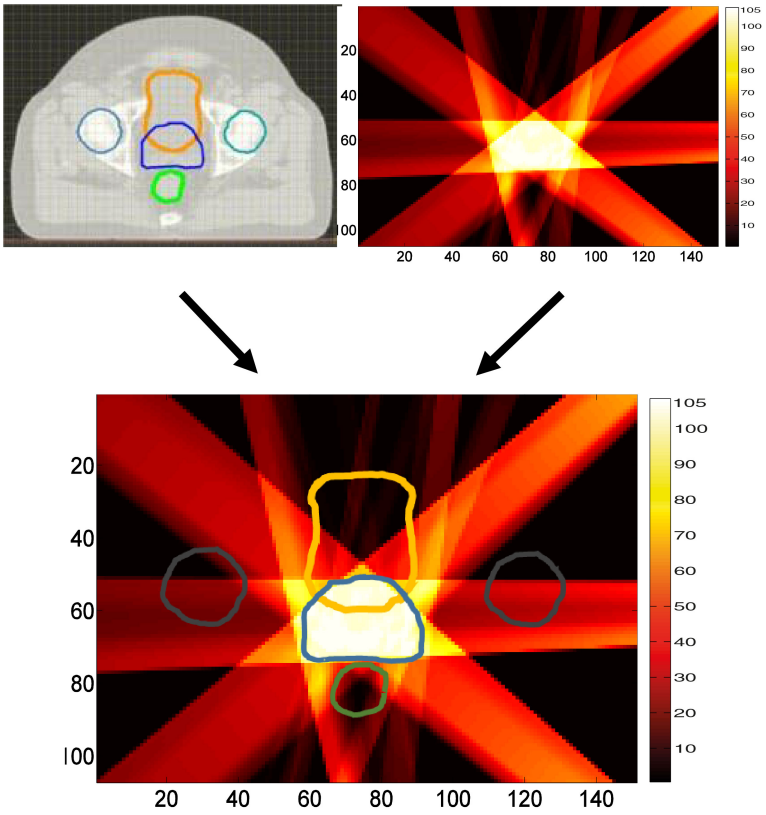


Figure 4.2: Top left: Input data. Top right: Optimizer output with the relative dose scale on the right. Low dose corresponds to black, high dose to white. Bottom: Overlay of dose with contours to visually verify the solution.

Although these plots show qualitative results where one can say that radiation is directed toward the PTV and fluence maps are determined correctly, to verify the proposed

Contour	Min	Mean	Max
PTV	95.0	99.6	105
Rectum	3.6	25.0	93.2
Bladder	0	25.0	97.1
Right Femoral Head	9.3	25.0	44.9
Left Femoral head	9.3	25.0	44.9

Table 4.2: Relative doses achieved by the optimizer for the original data. This information is shown in terms of minimum, mean and maximum doses received by the contour indicated.

approach, and eventually compare it to other approaches, numerical evaluations are required. Table 4.2 demonstrates the numbers that were obtained as output of the code and compares them to the minimum or maximum allowed for each contour.

Finally, the amount of time it takes for the solver to generate outputs is measured to show the practicality of this approach. The average time measured for a number of iterations made was between 4 to 5s using the NEOS online server.

Now that dose distribution is verified to be in the acceptable range for both the PTV and critical tissues, the next step in this approach is to generate radiation profiles for the multi-leaf collimator to setup the sequence of leaves. As there are seven different angles used in this project, the code should generate seven different sets of radiation profile instructions for these angles. Figure 4.3 graphs the radiation profiles of seven co-planar beams obtained using the proposed LP approach. In this figure the horizontal axis shows the offset which means the distance from the beam edge, and the vertical axis shows the amount of radiation that should be sent through each offset for the respective angle. It should be mentioned that these radiation profiles do no have to be generated in one iteration and they may be the result of multi-step collimator configurations that are summed.

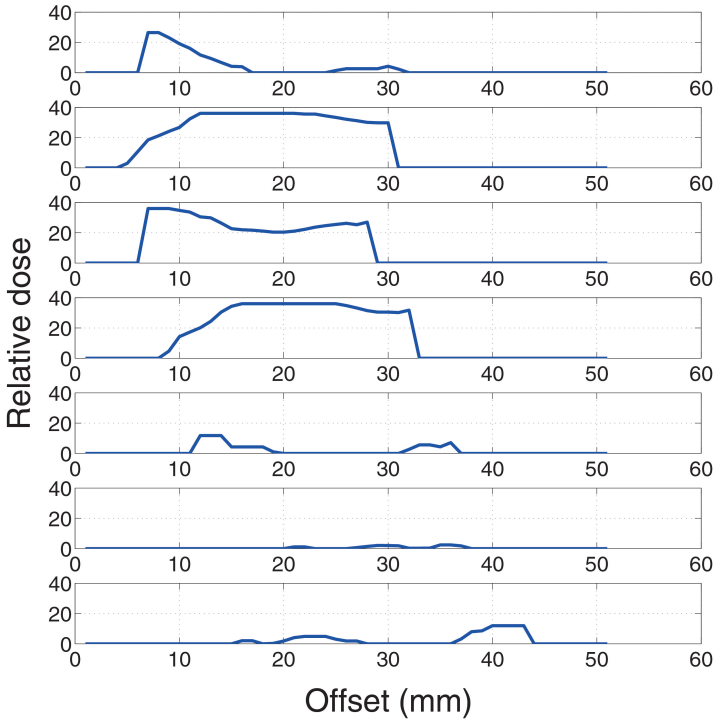


Figure 4.3: ODM profiles for the seven beam angles, 10 (top), 50, 90, 130, 170, 200, 340 (bottom)

4.3.2 Tissue Motion

Table 4.3 presents the numerical summaries of the eight treatment plans. Figures 4.4 and 4.5, use the same scheme as Figure 4.2 to demonstrate delivered dosage to each contour. In each figure, the top left part shows the dose distribution for the original case, the top right shows the distribution when an organ moves and the bottom part shows the difference in dose distribution between the top two. The difference is added mainly to illustrate the effectiveness of this method visually. In the difference plots, warm colours (red, orange, yellow) shows where the original dosage is larger than the one after the movement and the

#	PTV min/mean/max	Rectum min/mean/max	Bladder min/mean/max	R Femoral min/mean/max	L Femoral min/mean/max
1	95/98.46/105	4.9/25/51.26	0/25.6/98	11.2/24.6/47.2	11.1/24.6/48.3
2	95/98.43/105	4.94/25/51.35	0/24.4/97.9	11.2/24.6/47.16	11.1/24.6/48.3
3*	0/4.46/10	0/25.43/142.2	0/58/144.7	0/27.9/65.19	0/27.9/65.3
4*	0/4.34/10	0/25.51/142.2	0/53.24/143	0/27.9/65.2	0/27.93/65.3
5*	0/3.15/10	0/30.7/120.1	0/36.4/137.4	0/21.34/64.85	0/21.3/64.9
6*	0/2.16/10	0/31/120.42	0/32.7/137.4	0/21.3/64.85	0/21.3/64.9
7	95/98.3/105	1.8/20.42/51.3	0/24.4/89.2	3.18/24.8/53.7	3.1/24.73/53.5
8	95/98.24/105	1.8/20.52/51.6	0/24.6/98	3.2/24.8/53.5	3.1/24.73/53.4

Table 4.3: Relative doses for all tissues in terms of minimum, mean and maximum received. This table demonstrates good coverage for the PTV in case of rectum moving down and right, while sparing the normal tissues. The cases marked by a star have failed in the optimization process. As numbers for dose amounts indicates, when the PTV overlaps with the rectum, the DVCs can not be satisfied and the optimization problem failed.

cold colours (blue, green) show the opposite. As figure 4.4 demonstrates, in the case rectum shifted down, there is a decrease in radiation from down because rectum moves down, and an increase in radiation from bottom right that just hits the PTV above the rectum)

4.3.3 Norm Analysis

For the non-linear analysis, solver states might not be in the final solution neighbourhood and the program returned almost the same dose distribution and even increased the number of leaf positions required. It also took the solver more time (about 14s)and iterations to solve the new non-linear problem. Figure 4.6 is the KNITRO output demonstrating that the solver could not solve it within 50 iterations.

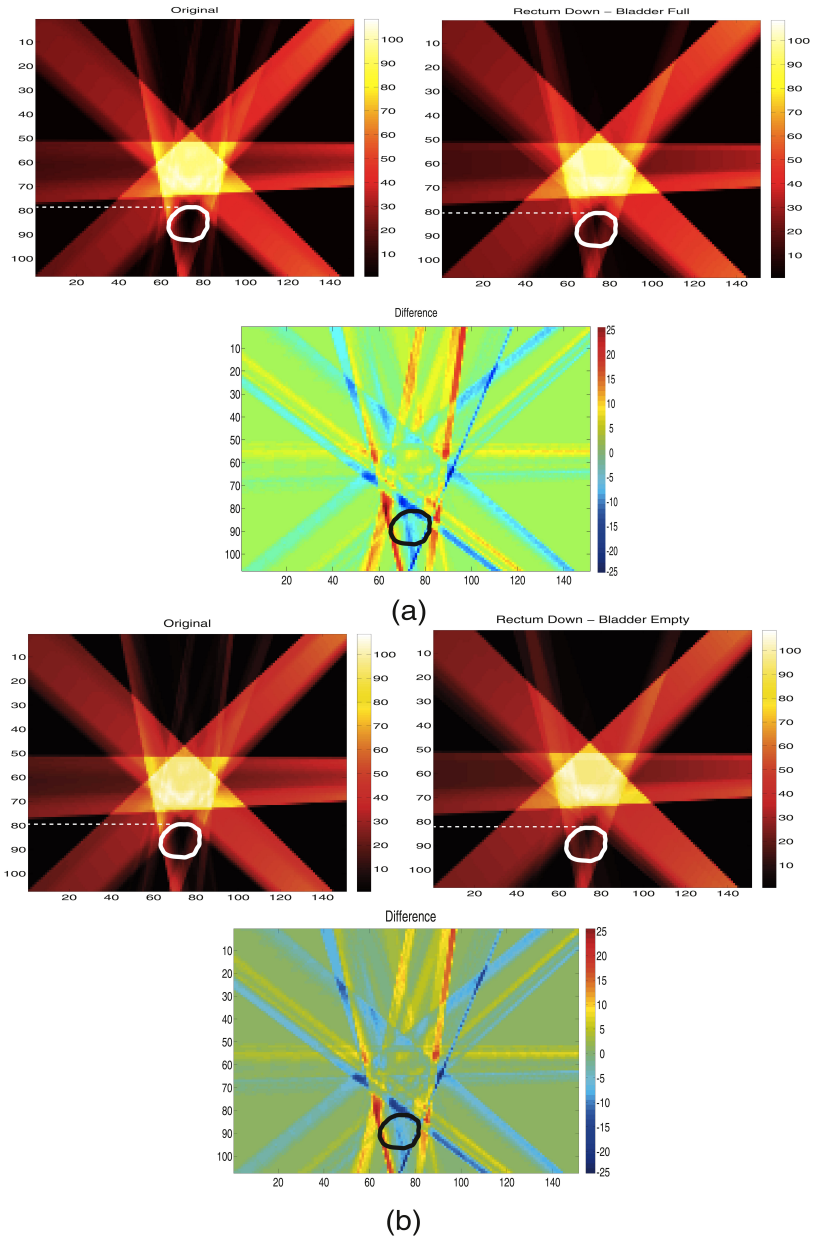


Figure 4.4: (a) Results obtained when the bladder size decreases. The top left picture shows the original dose while the top right shows the dose obtained when the rectum moves posteriorly. The bottom picture is the difference between the top two. Warm colours indicate the original dose was higher while cold colours indicate the opposite case. (b) Same results as in (a) but obtained when the bladder size increases.

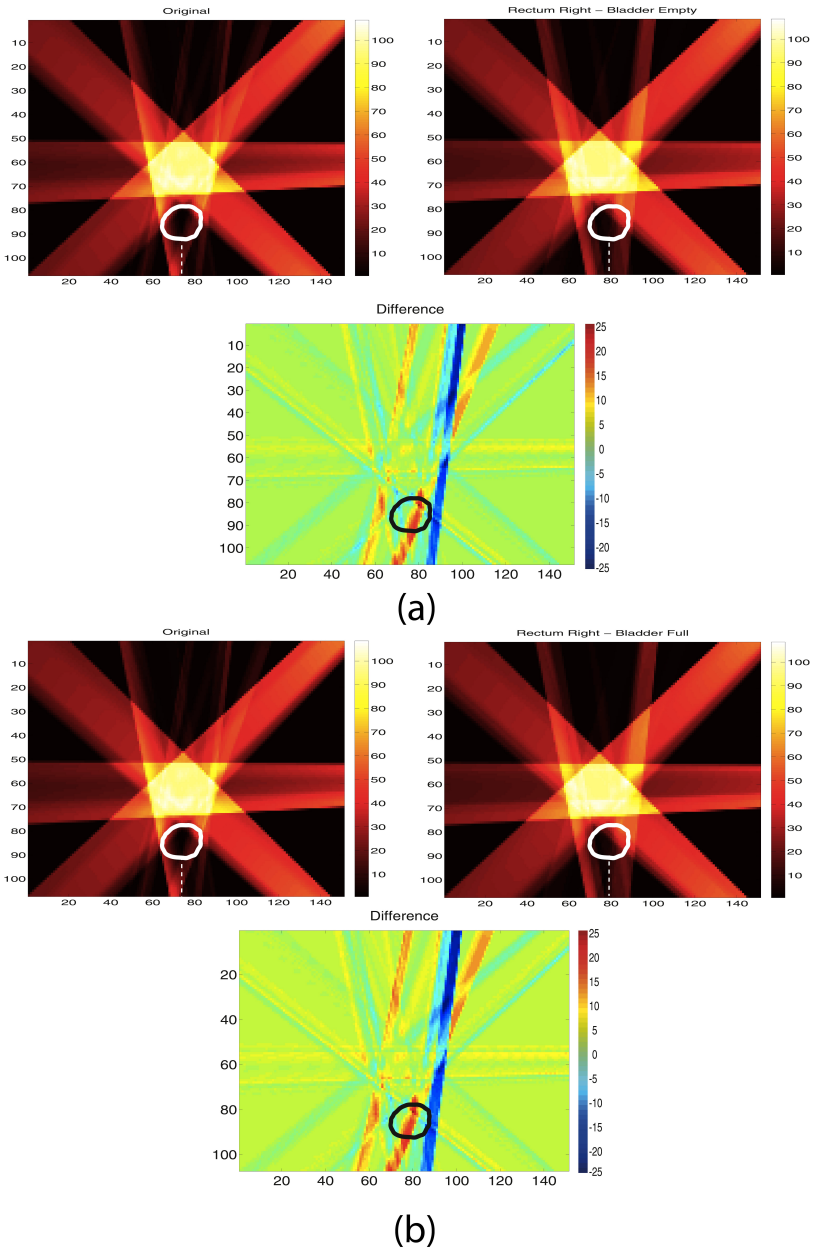


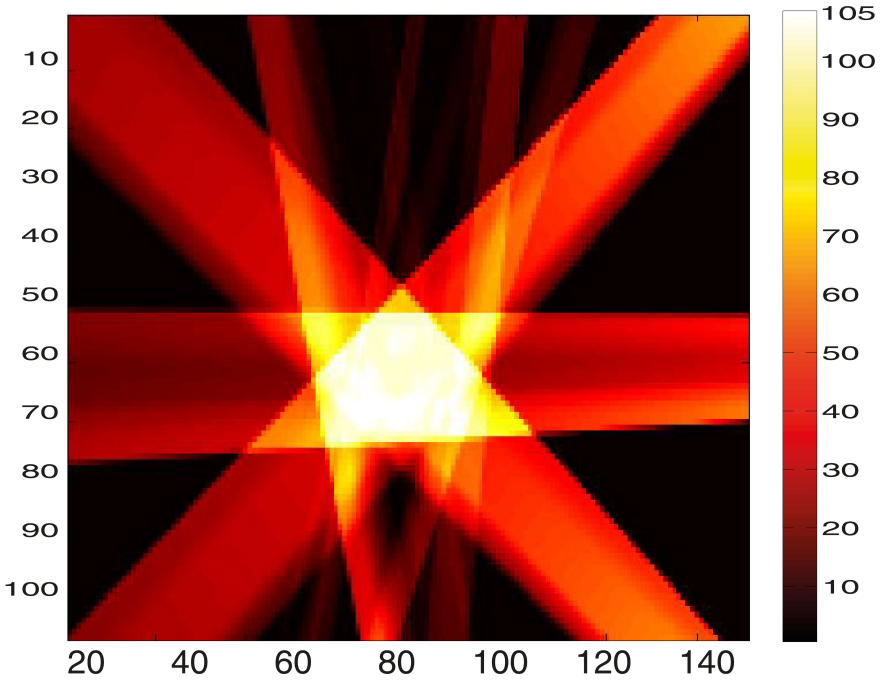
Figure 4.5: (a) Results obtained when the bladder size decreases. The top left picture shows the original dose while the top right shows the dose obtained when the rectum moves to the right. The bottom picture is the difference between the top two. Warm colours indicate the original dose was higher while cold colours indicate the opposite case. (b) Same results as in (a) but obtained when the bladder size increases.

```
Problem Characteristics
-----------------------
Objective goal:  Minimize
Number of variables:                 24004
    bounded below:                   19566
    bounded above:                       0
    bounded below and above:          4438
    fixed:                               0
    free:                                0
Number of constraints:               24001
    linear equalities:               16157
    nonlinear equalities:                0
    linear inequalities:              7844
    nonlinear inequalities:              0
    range:                               0
Number of nonzeros in Jacobian:     267018
Number of nonzeros in Hessian:        7847

  Iter      Objective      FeasError   OptError    ||Step||    CGits
--------  --------------  ----------  ----------  ----------  -------
       0    6.088804e+03   6.261e+00
      10    3.413655e+03   2.312e-01   1.579e-01   1.141e+01        1
      20    2.920304e+03   7.188e-02   5.808e-02   1.422e+01        1
      30    2.830406e+03   3.713e-02   3.650e-02   4.338e+00        1
      40    2.821485e+03   2.722e-02   5.513e-02   4.732e+00        2
      50    2.820973e+03   2.501e-02   3.363e-01   1.700e+00        2

EXIT: Iteration limit reached.

Final Statistics
----------------
Final objective value               =   2.82097323785708e+03
Final feasibility error (abs / rel) =   2.50e-02 / 4.00e-03
Final optimality error  (abs / rel) =   3.36e-01 / 1.70e-01
# of iterations                     =         50
# of CG iterations                  =         63
# of function evaluations           =         51
# of gradient evaluations           =         51
# of Hessian evaluations            =         50
Total program time (secs)           =      14.07218 (    14.070 CPU time)
Time spent in evaluations (secs)    =       0.28574

===============================================================================

KNITRO 8.0.0: Iteration limit reached.
objective 2820.973238; feasibility error 0.025
50 iterations; 51 function evaluations
```

Figure 4.6: KNITRO solution for the non-linear problem. It demonstrates it is not able to solve the problem within the given iterations.

Adding ℓ_2 difference norm to the problem gave smoother fluence maps as expected

and did not seem to affect the final dose distribution. Figure 4.7 demonstrates the dose distribution resulting from using ℓ_2 difference norm.



edge.jpg

Figure 4.7: Relative dose obtained using ℓ_2 difference norm. The dose distribution is almost the same except that the doses at the edges are smoother which requires more MLC control points to deliver.

4.4 Discussion

The proposed method using the ℓ_1 norm requires up to 5 s for the original data to produce the acceptable results shown in figure 4.2. As shown in table 4.2, the average dosage that is received by the tumour area is above the minimum required which shows the effectiveness

of this approach. Also the average dosage that is received by the other organs are all in the prescribed safe range. The only thing that should be noted in practical experiment is that the patient movement should be minimized during the 5 s processing time. In other words, after each movement, it takes the same amount of time to calculate radiation doses at each point. Finally, It should be mentioned that as an online solver is used, this time basically includes all the delays that are involved in that process as well as printing the results which would not be the case when a solver is embedded into a treatment planning system.

Although it was mentioned that the resulting dose can be generated in several steps, it is preferred to have a minimum number of leaf movements, or change in leaf configurations. This is basically what has been targeted in the optimization problem. In other words, it is tried to generate radiation profiles that are not only realizable using multi-leaf collimator leaves, but also require a minimum number of leaf movements to realize them. To do this, it is necessary to make sure that the fluence maps are piece-wise constant for each angle. Local maximum and minimums in the profile will result in a very high number of iterations of leaf configurations being required to generate such profile. As shown in Figure 4.3, all the profiles for each angle are appear piece-wise constant. The optimization method has resulted in profiles that may be generated using a few iterations in leaf configurations.

Since the motivation for plan re-optimization comes from concerns about patient motion between treatment fractions, eight cases were optimized representing typical patient deformations. The difference plots in figure 4.4 and figure 4.5 clearly shows that there is always a strip of intense warm or cold colour right at the edge of the organ that is moving which shows the change that is made by the program to re-optimized the treatment plan. It is interesting to note that when the rectum moves down (and away from the tumour), although the new plan reduces the maximum dose to the rectum, the average dose remains

constant. As the relative doses in table 4.3 demonstrates, the LP proposed approach performs well for shifting down and right cases, however failed in the other two motions. In the case the rectum is shifted up, it overlaps with the PTV area and satisfying both tissue constraints will be impossible and the optimizer will not be able to solve it. The reason that the shifting left case has been failed is likely due to the non-clinical angles employed in this analysis

Also rather than ℓ_1 difference norm, $\ell_{1/2}$ and ℓ_2 norms also were used to improve the solution. As shown in figure 4.6, using $\ell_{1/2}$ as a penalty to the ℓ_1 difference did not result in better solutions. As figure 4.7 shows, ℓ_2 norm also gave the same dose map as the ℓ_1 norm with smoother fluence profiles which requires smaller gradients and subsequently, less OAR sparing. Also more segments are required to realize this solution using the linear accelerator.

Chapter 5

Conclusion

As discussed, the current approach to IMRT involves setting the patient for treatment, imaging, rigidly shifting the couch and delivering the original treatment plan. The problem is that internal patient motion is non-rigid. Currently, the problem of internal deformations is addressed by increasing the size of the PTV margin which leads to delivering more radiation to sensitive organs. A fast method of re-optimizing IMRT treatments in response to anatomical deformation detected just prior to each treatment fraction would allow the PTV margin to be reduced. Therefore the goal of this thesis was to prototype a fast technique for re-optimizing a prostate IMRT plan based on the motion observed on daily CBCT imaging.

To achieve this goal, a rapid linear programming approach using was proposed in this thesis. The novelty in this approach was the use of regularization to minimize the necessary number of MLC leaf movements. Of the methods tried, ℓ_1-difference regularization was the most effective and was readily solvable. This case corresponds to a linear program. Other norms, including the ℓ_2 and $\ell_{1/2}$ in combination with the ℓ_1 norm, produced inferior solutions in longer solution times. As the numerical and graphical results indicate, the LP solution conforms radiation to the tumour shape, as expected, in the test problem.

Since the motivation for plan re-optimization comes from concerns about patient motion between treatment fractions, eight cases were optimized representing typical patient deformations. In the case when rectum is shifted up and has an overlapped area with the PTV, the re-optimziation plan failed and could not satisfy the DVCs. Although this optimization method could fail, one can always fall back to the original optimization plan provided by the currently-used software.

For full implementation on a linear accelerator, the resulting fluence profiles need to be rounded to get a piece-wise constant function that is realizable with a small number of leaf movements. Qualitatively, the LP solution satisfies this condition without requiring excessive processing time. However the $l_{1/2}$ difference norm which was expected to reduce the number of fluence steps to make it easier to round to leaves, did not work very well. The resulting non-linear problem was harder for KNITRO to solve and the quality of the solution was not as good.

Appendix A

AMPL Code

```
 param size := 11;
param max_rad := 1000;
param pi := acos(-1);
param max_angl := 17;
 param max_odm := 35;
 adding an array of arbitrary angles


  param odm_per_cm := 4;
  param dose_per_cm := 4;



param odm_radius_cm := 35;
param odm_per_cm := 4;
param dose_per_cm := 4;
param max_odm := odm_per_cm * odm_radius_cm*2;  to avoid error
```

```
param source_distance_cm := 100;
param n := 7;
set ANGLES := {10, 50, 90, 130, 170, 200,340};


var rad{angle in ANGLES, offset in {-max_odm..max_odm}} >= 0;

 param max_at_risk := 6;
 param min_at_tumor := 10;

param t{i in {-53..53},j in {-75..75}, angle in ANGLES} =
(i/dose_per_cm*cos(angle*pi/180)+j/dose_per_cm*sin(angle*pi/180))/
(i/dose_per_cm*cos(angle*pi/180)+j/dose_per_cm*sin(angle*pi/180)-
source_distance_cm);
param s {i in{-53..53}, j in {-75..75}, angle in ANGLES} :=
odm_per_cm * (-sin(angle*pi/180)*(i/dose_per_cm+t[i,j,angle]*
(source_distance_cm*cos(angle*pi/180)-i/dose_per_cm)) +
 cos(angle*pi/180)*(j/dose_per_cm+t[i,j,angle]*
 (source_distance_cm*sin(angle*pi/180)-j/dose_per_cm)));

param s_int {i in{-53..53}, j in {-75..75}, angle in ANGLES} :=
floor(s[i,j,angle]);
param s_frac {i in{-53..53}, j in {-75..75}, angle in ANGLES} :=
s[i,j,angle]- s_int[i,j,angle];
param s_frac1 {i in{-53..53}, j in {-75..75}, angle in ANGLES} :=
1 - s_frac[i,j,angle];
```

```
var dose {i in {-53..53}, j in {-75..75}};

param geom {i in {-53..53}, j in {-75..75}, angle in ANGLES} :=
source_distance_cm*source_distance_cm/
(((source_distance_cm*cos(angle*pi/180)- i/dose_per_cm)^2)+
(source_distance_cm*sin(angle*pi/180)-j/dose_per_cm)^2);

  var t1{angle in ANGLES, offset in{-max_odm..max_odm}} >= 0;

  subject to constT1 {{angle in ANGLES, offset in {-max_odm..max_odm}}:
   t1[angle,offset] >= rad[k,offset] - rad[k-1,offset];
  subject to constT2 {angle in {ANGLES+1}, offset in {-max_odm..max_odm}}:
   -t1[angle,offset] <= rad[angle,offset] - rad[angle-1,offset];

var t2{angle in ANGLES, offset in {-max_odm+1..max_odm}} >= 0;

subject to constT12 {angle in ANGLES, offset in{-max_odm+1..max_odm}}:
        t2[angle,offset] >= rad[angle,offset] - rad[angle,offset-1];
subject to constT21 {angle in ANGLES, offset in {-max_odm+1..max_odm}}:
        -t2[angle,offset] <= rad[angle,offset] - rad[angle,offset-1];


subject to const3 {i in {-53..53}, j in {-75..75}}:
dose[i,j] =sum {angle in ANGLES}geom[i,j,angle]*
((s_frac1[i,j,angle])*rad[angle,s_int[i,j,angle]]+(s_frac[i,j,angle])*
rad[angle,1+s_int[i, j, angle]]);
```

```
 //====================================
  param max_tumor;
  let max_tumor := {(i,j) in T} dose[i,j];

//=======================================
subject to const_68 {(i,j) in {
(-2,-4),(-2,-3),(-2,-2),(-2,-1),(-2,0), ...
}}: dose[i,j] >= 95;
//=========================================
subject to const1_68 {(i,j) in {
(-2,-4),(-2,-3),(-2,-2),(-2,-1),(-2,0), ...}}: dose[i,j] <= 105;



 subject to const1_68 {(i,j) in {
 (-24,18),(-24,19),(-24,20),(-24,21),(-24,22), ...
 }}: dose[i,j] >= 105;
param max_s := 25;
 subject to const_2:(sum {(i,j) in {
 (-25,-2),(-25,-1),(-24,-9),(-24,-8),(-24,-7), ...
 }}dose[i,j])*(0.00179856115) <= max_s;

subject to const_3:(sum {(i,j) in {
(-8,-35),(-8,-34),(-8,-33),(-8,-32),(-8,-31), ...
}} dose[i,j])*(0.00321543408) <= max_s;
```

```
subject to const_4:(sum {(i,j) in {
(-7,28),(-7,29),(-7,30),(-7,31),(-7,32), ...
}} dose[i,j])*(0.00355871886) <= max_s;

subject to const_5:(sum {(i,j) in {
(19,-4),(19,-3),(19,-2),(19,-1), ...
}} dose[i,j])*(0.00534759358) <= max_s;

 subject to l1odm{angle in ANGLES}:rad[angle,-max_odm] +
  rad[angle,max_odm] +
 sum{off in {-max_odm+1..max_odm}}t2[angle,off] <= 2*max_rad;

minimize amount:
( sum {angle in ANGLES, off in {-max_odm+1..max_odm}} t2[angle,off]) +
( sum {angle in ANGLES} (rad[angle,-max_odm] + rad[angle,max_odm])) ;

solve;
```

Bibliography

[1] Emami B, Lyman J, Brown A, Coia L, Goitein M, Munzenrider JE, Shank B, Solin LJ, and Wesson M. Tolerance of normal tissue to therapeutic irradiation. *Int J Radiat Oncol Biol Phys.*, 21:109–22, 1999.

[2] Saeed Ahmad Buzdar, M. Afzal, and Andrew Todd-Pokropek. Comparison of pencil beam and collapsed cone algorithms, in radiotherapy treatment planning for 6 and 10 mv photon. *J Ayub Med Coll Abbottabad*, 22(3):3521, 2010.

[3] C. Cotrutz, M. Lahanas, C. Kappas, and D. Baltas. A multiobjective gradient-based dose optimization algorithm for external beam conformal radiotherapy. *Phys. Med. Biol.*, 46:2161–2175, 2001.

[4] J. Czyzyk, M. Mesnier, and J. Mor. The neos server. *IEEE Journal on Computational Science and Engineering*, 3:68–75, 1998.

[5] E. Dolan. The neos server 4.0 administrative guide. ANL/MCS-TM 250, Mathematics and Computer Science Division, Argonne National Laboratory, May 2001.

[6] Matthias Ehrgott, i?gdem Gler, Horst W. Hamacher, and Lizhen Shao. Mathematical optimization in intensity modulated radiation therapy. *Springer Verlag*, 6:299–262, 2008.

[7] Petruzzelli GJ Emami B, Sethi A. Influence of mri on target volume delineation and imrt planning in nasopharyngeal carcinoma. *Int J Radiat Oncol Biol Physt*, 57(2):481–8, 2003.

[8] W. Gropp and J. Mor. Optimization environments and the neos server. *Approximation Theory and Optimization*, 6:67–182, 1997.

[9] A. Holder. Designing radiotherapy plans with elastic constraints and interior point methods. *Health Care and Management Science*, 6(1):5–16, 2003.

[10] Bjrn Hrdemark, Anders Liander, Henrik Rehbinder, and Johan Lf. Direct machine parameter optimization with raymachine in pinnacle. *RaySearch White Paper*, 2003.

[11] Bourland JD and Chaney EL. A finite-size pencil beam model for photon dose calculations in three dimensions. *Med Phys.*, 19:1401–12, 1992.

[12] M. Lahanas, M. Schreibmann, and D. Baltas. Multiobjective inverse planning for intensity modulated radiotherapy with constraint-free gradient-based optimization algorithms. *Phys. Med. Biol.*, 48(17):2843–2871, 2003.

[13] E. W. Larsen. Tutorial: The nature of transport calculations used in radiation oncology. *Transport Theory Statist. Phys.*, 26:739, 1997.

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

[15] J. Lf. *Development of a general framework for optimization of radiation therapy*. PhD thesis, Stockholm University, 2000.

[16] Van Herk M, Bruce A, Kroes AP, Shouman T, Touw A, and Lebesque JV. Quantification of organ motion during conformal radiotherapy of the prostate by three dimensional image registration. *Int J Radiat Oncol Biol Physt*, 33(5):1311–20, 1995.

[17] S L Morgan-Fletcher. Prescribing, recording and reporting photon beam therapy. ICUR 62, International Commission on Radiation Units and Measurements, 7910 Woodmont avenue Bethesda, Maryland 20814, USA, September 1999.

[18] R. Rardin, M. Langer, F. Preciado-Walters, and V. Thai. A coupled column generation, mixed integer approaches to optimal planning of intensity-modulated radiation therapy for cancer. *To appear in Mathematical Programming*, 2004.

[19] H. Romeijn, R. Ahuja, J. Dempsey, A. Kumar, and J. Li. A novel linear program- ming approach to fluence map optimization for intensity modulated radiation therapy treatment planning. *Phys. Med. Biol.*, 48:3521–3542, 2003.

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

[21] American Cancer Society. Statistics about prostate cancer, January 2013.

[22] Bortfeld T. Optimized planning using physical objectives and constraints. *Radiat Oncol*, 9:20–34, 1999.

[23] Woo SY Teh BS and Butler EB. Intensity modulated radiation therapy (imrt): a new promising technology in radiation oncology. *Oncologist*, 4(6):433–42, 1999.

[24] Dr Padraig Warde, Malcolm Mason, Keyue Ding, Peter Kirkbride, and Michael Brundage. Combined androgen deprivation therapy and radiation therapy for locally advanced prostate cancer: a randomised, phase 3 trial. *The Lancet*, 378:2104 – 2111, 2011.

[25] Peter White, Kit Chi Chan, Ka Wai Cheng, Ka Yiu Chan, and Ming Chun Chau2. Volumetric intensity-modulated arc therapy vs conventional intensity-modulated radiation therapy in nasopharyngeal carcinoma: a dosimetric study. *J Radiat Res*, 111:2161–2175, 2012.

[26] Marcin Wierzbicki, Bryan Schaly, Terry Peters, and Rob Barnett. Automatic image guidance for prostate imrt using low dose cbct. *Med Phys.*, 37(3):3677–86, 2010.

[27] C. Wu, R. Jeraj, W. Lu, and T. Mackie. Fast treatment plan modification with an over-relaxed cimmino algorithm. *Med. Phys.*, 31:191–200, 2004.

[28] Q. Wu, D. Djajaputra, M. Lauterbach, Y. Wu, and R. Mohan. A fast dose calculation method based on table lookup for imrt optimization. *Phys. Med. Biol.*, 48(12):159–166, 2003.

[29] Q. Wu and R. Mohan. Algorithms and functionality of an intensity modulated radio-therapy optimization system. *Med. Phys.*, 27(4):701–711, 2000.

[30] Q Jackie Wu, Danthai Thongphiew, Zhiheng Wang, Boonyanit Mathayomchan, Vira Chankong, Sua Yoo, W Robert Lee, and Fang-Fang Yin. On-line re-optimization of prostate imrt plans for adaptive radiation therapy. *Phys. Med. Biol.*, 53:673–691, 2008.

[31] Michael Merritt Yin Zhang. Fluence map optimization in imrt cancer treatment planning and a geometric approach. *Nonconvex Optimization and Its Applications*, 82:205–227, 2006.