Winter 12-15-2015

Automating Intensity Modulated Radiation Therapy Treatment Planning by using Hierarchical Optimization

Paras Babu Tiwari
Washington University in St. Louis

Recommended Citation

Tiwari, Paras Babu, "Automating Intensity Modulated Radiation Therapy Treatment Planning by using Hierarchical Optimization" (2015). *Engineering and Applied Science Theses & Dissertations*. 140.
https://openscholarship.wustl.edu/eng_etds/140

WASHINGTON UNIVERSITY IN ST. LOUIS

School of Engineering and Applied Science
Department of Computer Science and Engineering

Dissertation Examination Committee:
Yixin Chen, Chair
Roger Chamberlain
Sanmay Das
Joseph O. Deasy
Eric E. Klein

Automating Intensity Modulated Radiation Therapy Treatment Planning by using Hierarchical Optimization

by
Paras Babu Tiwari

A dissertation presented to the
Graduate School of Arts & Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

December 2015
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank my advisors Prof. Yixin Chen and Prof. Joseph O. Deasy for supporting me to complete the dissertation project. I could not have finished the dissertation work unless they had passionately supported and encouraged me during the more challenging moments.

The dissertation is supported by Enid A. Haupt Endowed Chair in Medical Physics, CCF-1215302, IIS- 1343896, and DBI-1356669 grants from the National Science Foundation and Microsoft Research new faculty fellowship. I am thankful to all financial sources for the dissertation work.

I am grateful to the Memorial Sloan Kettering Cancer Center team for providing feedback on my work during weekly meetings. I would like to thank Margie Hunt, Dr. Linda Hong, Dr. Gikas Mageras, Dr. James Mechalakos, and Dr. Aditya Apeta for their valuable time.

I am thankful to Prof. Roger Chamberlain, Prof. Sanmay Das, and Prof. Eric E. Klein for serving on my dissertation committee. I am thankful to Prof. Raj Jain for serving as the academic advisor in the beginning of my doctoral program.

I would like to thank Dr. Jung Hun Oh and the administrative staff of Memorial Sloan Kettering Cancer Center (MSKCC) for their time and support while I was an intern in MSKCC.

I would like to thank the administrative staff of the Department of Computer Science, Washington University in St. Louis for their administrative support during my graduate study.

I am thankful to Mark Bober and his team of Engineering IT, Washington University in St. Louis for helping me to run the experiments in the Engineering Cloud Cluster.

I would like to thank my colleague Dr. Yao Xie with whom I have had stimulating discussions about the dissertation work.

Finally, I would like to thank my family and friends for supporting me in my dissertation work.

Paras Babu Tiwari

Washington University in St. Louis

December 2015

Dedicated to my family.

.

ABSTRACT OF THE DISSERTATION

Automating Intensity Modulated Radiation Therapy Treatment Planning by using Hierarchical Optimization

by

Paras Babu Tiwari

Doctor of Philosophy in Computer Science

Washington University in St. Louis, 2015

Professor Yixin Chen, Chair

The intensity modulated radiation therapy (IMRT) optimizes the beam's intensity to deliver the prescribed dose to the target while minimizing the radiation exposure to normal structures. The IMRT optimization is a complex optimization problem because of the multiple conflicting objectives in it. Due to the complexity of the optimization, the IMRT treatment planning is still a trial and error process. Hierarchical optimization was proposed to automate the treatment planning process, but its potential has not been demonstrated in a clinical setting. Moreover, hierarchical optimization is slower than the traditional optimization. The dissertation studied a sampling algorithm to reduce the hierarchical optimization time, customized an open source optimization solver to solve the nonlinear optimization formulation and demonstrated the potential of hierarchical optimization to automate the treatment planning process in a clinical setting. We generated the treatment plans of 31 prostate patients by hierarchical optimization using the same criteria as used by planners to prepare the treatment plans at Memorial Sloan Kettering Cancer Center. We found that hierarchical optimization produced the same or better treatment plans than that produced by a planner using the Eclipse treatment planning system. Therefore, the dissertation demonstrated that hierarchical optimization could automate the

treatment planning process and shift the paradigm of the treatment planning from manual trial and error to an ideal automated process.

Chapter 1: Introduction

Intensity-modulated radiation therapy (IMRT) is widely used to deliver radiation therapy [1]–[3]. The IMRT optimization has multiple conflicting goals, and it requires substantial time and trial and error iterations to find an acceptable treatment plan. The most important goal is to deliver the prescribed dose to target structures and then to minimize dose to normal organs. To reflect the importance of a goal in the optimization, a weight is assigned to each goal. The planner chooses these weights based on experience and runs the treatment planning application to find a preferred dose distribution. Often, it is difficult to find a set of weights that result in a good dose distribution. Therefore, the optimization is run several times and the weight parameters are tweaked to find an acceptable treatment plan.

There has been substantial research in automating the treatment planning process using hierarchical optimization [4]–[6]. Hierarchical optimization assigns priorities to goals of the optimization and optimizes those goals based on priority. The purpose of prioritization is to ensure that the higher priority goals are not sacrificed to improve on the lower priority goals. Wilkens and Clark, et al.[4], [7] demonstrated the potential of hierarchical optimization by automatically creating treatment plans for head and neck and prostate plans. However, the treatment plan automation based on hierarchical optimization has not been extensively studied in a clinical setting. The clinical setting has complex beam arrangements, finer beam resolutions and rigorous dose distribution requirements. Therefore, it is essential to demonstrate that hierarchical optimization could generate the clinically acceptable plans before drawing any conclusions. Moreover, hierarchical optimization is slower than traditional optimization because hierarchical optimization increases the problem size.

The dissertation addresses three needs in the treatment planning optimization: 1) we explored the sampling method to reduce the hierarchical optimization time, 2) we customized Interior Point Optimizer, an open source solver that can solve general nonlinear function, to solve the hierarchical optimization problem[1], 3) we formulated the mathematical model to automate the treatment planning process in the clinical setting and automatically prepared the treatment plans of 31 prostate patients using the formulated mathematical model. The present study, therefore, demonstrated the potential of hierarchical optimization to shift the treatment planning paradigm from manual trial and error to an ideal automated process.

1.1 Background

The history of radiation therapy can be traced back to 1895 when Wilhelm Conrad Röntgen discovered the X-ray [2]. The discovery of the X-ray led to widespread use of radiation in diagnosis and treatment in the medical field. Emil Grubbe, a medical student from Chicago, experimented with X-rays as treatment for breast cancer in 1896 [3]. Grubbe’s treatment greatly benefited patients and radiation therapy was widely used to treat cancer patients in the US and Europe.

Although X-rays gained widespread popularity, there were several challenges in using them to treat cancer. The first challenge is that the X-ray kills both tumor and normal cells. There was limited knowledge about the impact of X-rays on normal cells. The limited knowledge about X-rays led to a high patient mortality rate. Later, it was realized that radiation was better tolerated if administered in small amounts over a long period. This idea is called fractionation and was primarily due to Prof. Claude Regaud from the Paris Institute. The second challenge was that the X-rays of that era had low energy[4]. The energy of X-rays was in the range of 180-200KV and such low energy X-rays were not powerful enough to treat tumors situated deep inside tissue.

The subsequent research led to the development of the machine that could produce super voltage X-rays. William David Coolidge built the first supervoltage X-rays machine that could produce an X-ray of approximately 500kV-2MV energy in 1927[5].

The scientific advancement led to the development of the megavoltage linear accelerator[6], [7]. The Varian Medical System began to build the first commercial linear accelerator in 1958. Figure 1-1 shows the modern linear accelerator built by Varian Medical Systems. The modern linear accelerator uses microwave technology that can produce high energy X-rays of 6-18 MV. The electrons produced by the electron gun are accelerated through the waveguide. The accelerated electrons collide with the heavy metallic target to produce a high energy X-ray. Finally, the X-ray particle is reshaped using the multi-leaf collimator to produce the dose distribution that matches the patient geometry[8].

In recent years, there has been substantial research on producing dose distribution that conforms to patient geometry[9]–[14]. The linear accelerator is equipped with multi-leaf collimators that can shape the radiation to match patient geometry[15]. The multi-leaf collimator consists of the individual leaves that can be opened and closed to produce the desired dose distribution (Figure 1-2). The multi-leaf collimator is controlled by the treatment planning application, which uses mathematical optimization to determine the leaves’ position. The research on developing the treatment planning application that can automatically produce conformal dose distribution is still an active research area.

Figure 1-1 The modern linear accelerator developed by Varian Medical Systems. The modern linear accelerator can produce an X-ray of 6-18MV [16].

Figure 1-2 The multi-leaf collimator of the modern linear accelerator. The multi-leaf collimator shapes the radiation to produce the dose distribution that matches with patient geometry [16].

1.1.1 Conformal Radiation Therapy

The conformal radiation therapy is a type of external radiation therapy that uses advanced technology to produce the dose distribution that conforms to patient geometry[17]. The research in medical physics and computer science has advanced imaging [18], planning[19] and delivery technologies[20]. Modern imaging technology can generate the three-dimensional image of the patient. The three-dimensional image provides accurate information about the location, size, and shape of the target and normal structures[18], [21]–[25]. Planning technology has advanced as a result of imaging technology advancement. The treatment planning application can calculate the accurate dose to structures, visualize the dose distribution, and perform optimization due to the geometrical information obtained from the three-dimensional image of the patient[26]–[29]. The planner can produce conformal dose distribution using the modern treatment planning application. Finally, the delivery techniques have also advanced as a result of imaging and planning technology advancement[30]. Modern linear accelerators are equipped with multi-leaf collimators that can shape the radiation to match patient geometry.

Three-dimensional conformal radiation therapy (3DCRT) was the first attempt to produce conformal dose distribution. In 3DCRT, the three-dimensional image of the patient is generated using the CT/PET scanner. The images produced from the scanner are imported into the treatment planning application. The treatment planning application provides the functionalities such as contouring the structures, discretizing the structure into voxels, calculating dose to each voxel of the structures, and navigating to the different slices of images. The patient geometry can be visualized to the physician using the treatment planning application. The physician chooses the beam direction that delivers the prescribed dose to the target and minimizes the radiation exposure to normal structures. The planner inputs the chosen beam direction in the treatment planning application and calculates the dose to each structure. The physician provides the feedback to the plan prepared by the planner, and the planner modifies the plan based on physician feedback. The process is repeated until the desired dose distribution is found.

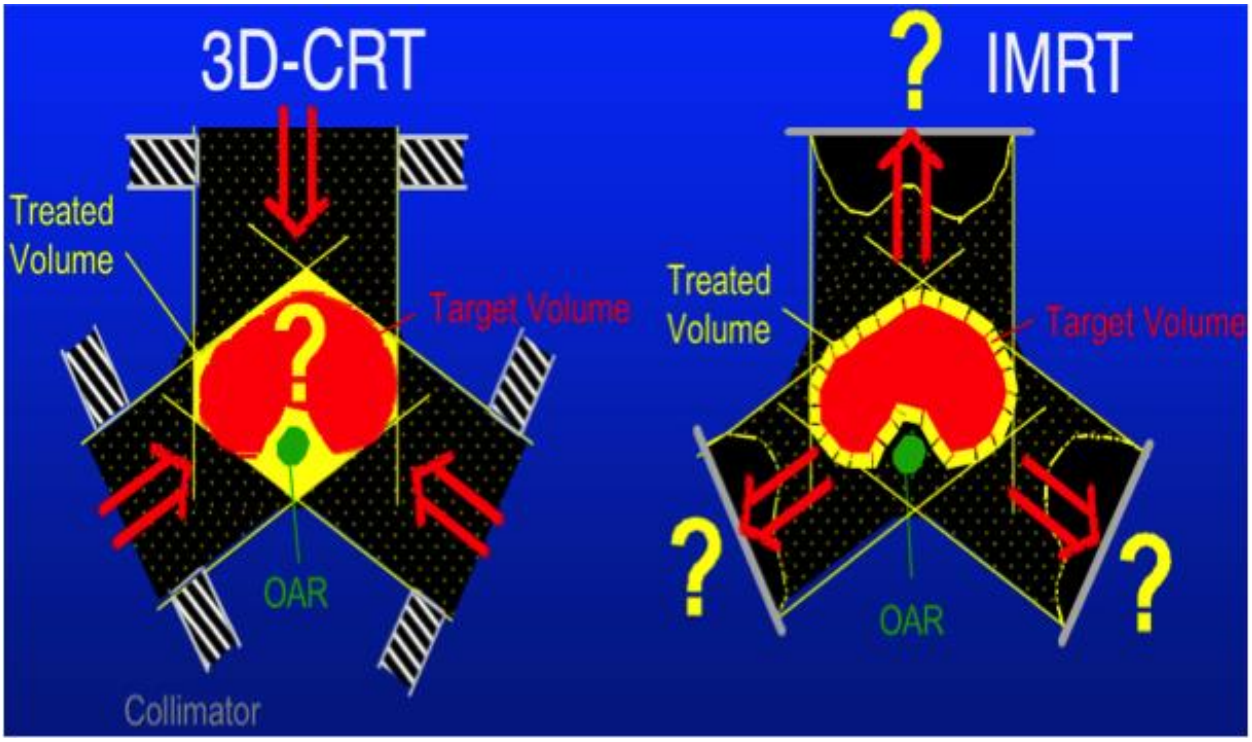


Figure 1-3 The 3DCRT vs. IMRT dose distribution. The 3DCRT uses beams of uniform intensity, while IMRT uses beams of variable intensity for treatment. Therefore, the dose distribution produced in IMRT better conforms to the target structure than that produced in 3DCRT [31].

Although 3DCRT improved the quality of radiation therapy, it still has limitations. The 3DCRT uses the uniform intensity beam for treatment. The uniform intensity beam produces the uniform dose around the target. Therefore, the 3DCRT cannot spare the normal structures located near the target structure (Figure 1-3). Intensity modulated radiation therapy is the advanced form of the three-dimensional conformal radiation therapy[32]. Unlike 3DCRT, the intensity of each beam could be different in IMRT. Intensity modulated radiation therapy modulates the beams to produce the dose distribution that matches the complex target shape (Figure 1-3). To modulate the beams, they are divided into beamlets and a mathematical model is formed that relates the intensity of the beamlets to the quality of the dose distribution to structures[33], [34]. Finally, the mathematical model is optimized to obtain the beamlets' intensity that produces the desired dose distribution.

Let x be the beamlet weight, $f_i(x)$ be the function defined for the i^{th} structure, and then intensity modulated radiation therapy solves the following function to find the optimal beamlet intensity:

$$Minimize \sum_{i \in S} w_i f_i(x) .$$

The w_i is the weight parameter that reflects the importance of the i^{th} structure in the overall optimization.

1.1.2 Challenges in Intensity Modulated Radiation Therapy

Intensity Modulated Radiation Therapy (IMRT) has multiple conflicting goals, and the IMRT optimization seeks to find the dose distribution that produces the best trade-off in the conflicting goals. Traditional IMRT optimization assigns a weight to each goal and the planner can change the weight parameters to reflect the importance of the goal in the optimization. But the planner doesn't know the exact weight parameters that produce the best trade-off in the dose distribution. Figure 1-4 shows the typical process used by the planner to find the value of the weight parameters. The planner guesses the weight parameters and runs the optimization to optimize dose distribution. The optimized dose distribution is examined to determine the quality of dose distribution. Typically, the dose distribution of the first round of the optimization does not satisfy the clinical criteria. The planner tweaks the weight parameters and reruns the optimization until the desired dose distribution is found. Thus, treatment planning became a trial and error process rather than an ideal automated process because of the trial and error iterations in determining the weight parameters of the optimization.

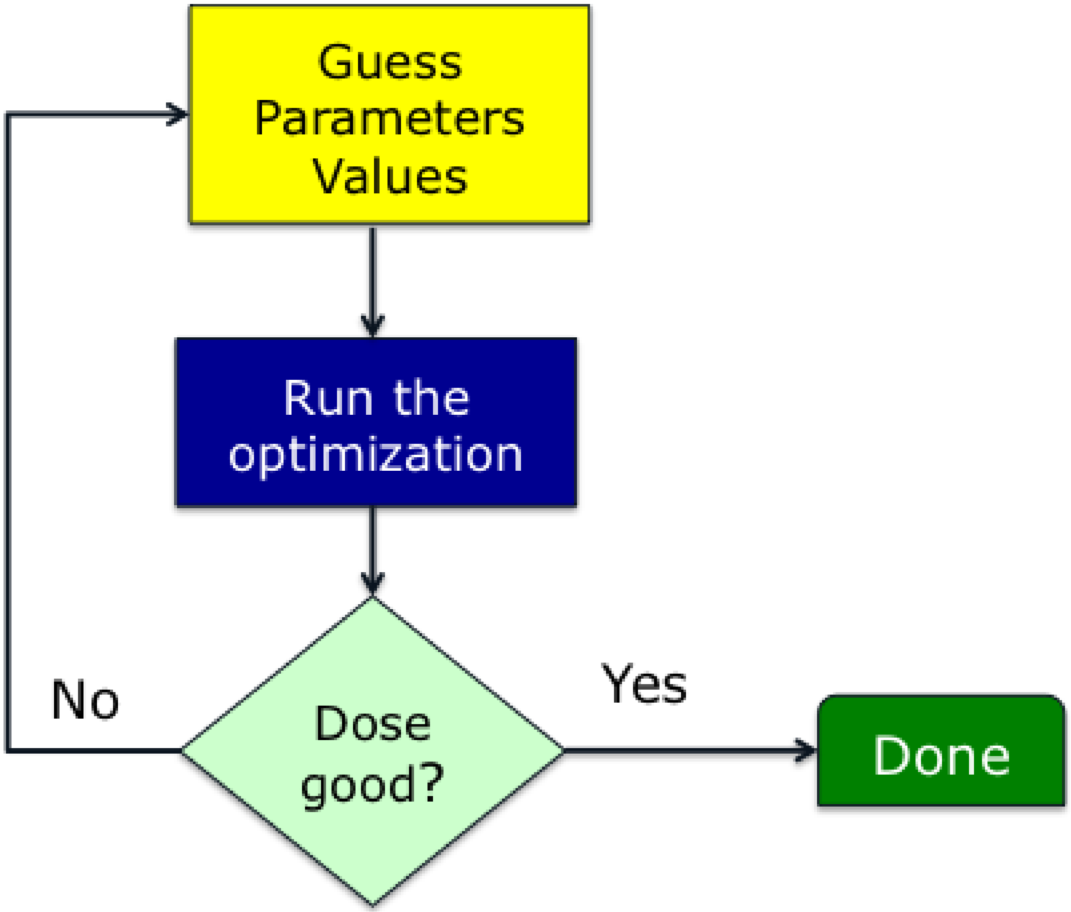


Figure 1-4 The treatment planning optimization workflow to determine the value of the weight parameters. The planner guesses the weight parameters and runs the optimization to optimize dose distribution. Typically, treatment planning optimization requires multiple iterations to produce the desired dose distribution. The planner tweaks the weight parameters and reruns the optimization until the desired dose distribution is found.

1.2 Multi-Objective Optimization Problem

Intensity Modulated Radiation Therapy is a multi-objective optimization problem. IMRT optimization has an objective function for each structure included in the planning and optimizes the objectives to produce the desired dose distribution. The multi-objective optimization does not have a unique optimal solution[35]–[39]. Therefore, the optimal solution of the multi-objective optimization problem is defined in terms of the Pareto optimality[40].

Mathematically, the multi-objective optimization problem can be defined as:

Let x =[x_1 ,x_2…, x_n] be the n-dimensional decision vector and $f_1(x)$,$f_2(x)$…$f_m(x)$ be the m objective functions, then the multi-objective optimization problem can be formulated as:

$$Minimize[f_1(x), f_2(x), \dots, f_m(x)],$$

$$Subject\ to:$$

$$x \in S\ ,$$

Where S is the set of constraints and defines the feasibility region of the optimization problem.

The optimal solution of the multi-objective optimization problem is defined in terms of the Pareto optimality.

Let x^1 and x^2 be two solutions of the multi-objective optimization problem, then the solution x^2 is said to be Pareto dominant over the x^1 if:

1. $f(x^2) \leq f(x^1)$ for all functions in the optimization
2. $f(x^2) < f(x^1)$ for at least one function in the optimization

A solution x* is said to be Pareto optimal if it is Pareto dominant over all solutions of the multi-objective optimization problem. We cannot improve on an objective without making another objective worse off when we have a Pareto optimal solution.

The set of all Pareto optimal solutions forms a Pareto frontier. Figure 1-5 shows the Pareto frontier for two objective functions. We assumed that the optimization problem minimizes the function $f_1(x)$ and $f_2(x)$. The set of points {a,b,c,d,e} defines the Pareto frontier of the optimization problem.

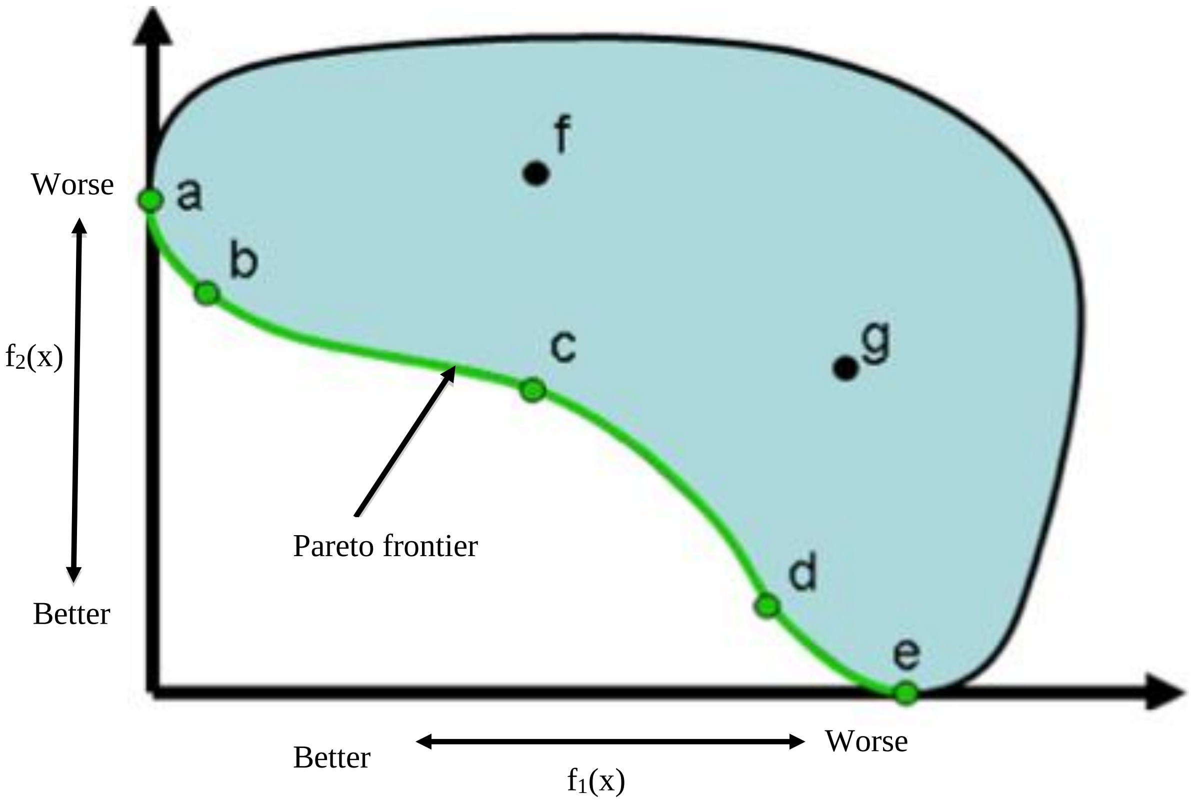


Figure 1-5 The optimal solution of the multi-objective optimization problem. The x-axis shows the first objective and the y-axis shows the second objective. The Pareto frontier displays all of the Pareto optimal solutions of the $f_1(x)$ and $f_2(x)$[41].

1.3 Goal Programming

There has not been a direct method to solve the multi-objective optimization problem. Therefore, the multi-objective optimization problem must be transformed into a form that can be solved using the existing optimization algorithms. Goal programming transforms the multi-objective optimization problem into a form that can be solved using the existing optimization algorithms [42]–[55]. The goal programming approach assigns the numeric value for each objective in the optimization. The optimization finds the solution that minimizes the deviation from the stated goal. Let f(x) is an objective function, "d" is a numerical value for the objective function, and then there are three types of goals:

1) $f(x)>=d$ sets the lower limit in the objective function f(x) and is called lower limit, one-sided goal.
2) $f(x)<=d$ sets the upper limit for the objective function and is called upper limit one-sided goal function.
3) $d1<=f(x)<=d2$ sets the upper and lower limit in the objective function and is called two-sided goal.

The IMRT optimization was solved using the goal programming approach. The objective functions of IMRT optimization are turned into goals by assigning a numerical value to the objective function. The goal for the target structure is to produce the prescribed dose to each voxel of the structure, while the goal for the normal structure is to push the dose as low as possible. Therefore, the target voxel that receives more than the prescribed dose is penalized in the optimization. Similarly, the dose of normal structure is minimized to push the dose as low as possible.

There are two ways to solve goal programming problems based on how the priorities of the goals are handled:

- Non-preemptive goal programming
- Pre-emptive goal programming

1.3.1 Non-Preemptive Goal Programming

Non-preemptive goal programming assigns almost the same priority to all of the goals in the optimization[56], [57]. The goals are weighted and combined to form a single objective function, which is minimized to find an optimal solution. The problem designer should have a good understanding of the problem domain to assign the weight to each goal of the optimization. For instance in IMRT optimization, the target goals are assigned higher weight than the normal structure. Mathematically, non-preemptive goal programming is formulated as follows:

$$Minimize \sum_{i=1}^{n} w_i(f_i(x) - d_i)^2 ,$$

d_i is the desired value of the function $f_i(x)$ and w_i is the weight assigned to the i^{th} goals.

1.3.2 Preemptive Goal Programming

Preemptive goal programming prioritizes the goals of the optimization and solves the goal based on its priority level [57]. If there is more than one goal of the same priority, they can be combined as in the non-preemptive method. The idea behind the preemptive goal programming is that the higher priority goal can’t be sacrificed to improve on the lower priority goals. The result of the higher priority goals is converted into the constraint while solving the lower priority goals. For instance in IMRT optimization, the target goal has the highest priority, the rectum has the second priority and the bladder and the femur have the third priority. Preemptive goal programming optimizes the target structure goal. The result of the target structure goal is converted into the constraint while optimizing for the rectum structure goal. Finally, the result of the target and the rectum structure goal is converted into the constraint while optimizing for the bladder and the femur goals.

In this dissertation, we used hierarchical optimization to solve the multi-objective IMRT optimization problem. Hierarchical optimization is the preemptive goal programming approach. The preemptive approach is an attractive choice to solve the IMRT optimization as it has the potential to automate the treatment planning process. However, the problem size of the preemptive approach is larger than that of the non-preemptive approach, and thus increases the computational time. We will discuss the sampling approach in reducing the optimization time in the next chapter.

1.4 Outline

The dissertation is organized into five chapters. Chapter 2 discusses boundary and interior voxel sampling to reduce the hierarchical optimization time. We demonstrated that the optimization time could be reduced by including all of the boundary voxels and sampling the interior voxels of a structure in a quasi-random fashion. Chapter 3 discusses customizing the high-performance nonlinear optimization interior point solver (IPOPT) to solve the hierarchical optimization problem. In Chapter 4, we formulated the mathematical model needed to automate the treatment planning process. We automatically prepared the treatment plans of 31 prostate patients using the mathematical formulation and compared the treatment plans prepared using the hierarchical and Eclipse treatment planning systems. Finally, Chapter 5 summarizes the dissertation work.

Chapter 2: Sampling

As we discussed in the previous chapter, hierarchical optimization increases the problem size by introducing new constraints in the problem formulation. The increase in the problem size increases the computational time of the optimization. In this chapter, we proposed a sampling algorithm to reduce the computational time of hierarchical optimization.

2.1 Introduction

Intensity-modulated radiation therapy (IMRT) offers many additional degrees of freedom compared to conformal radiotherapy techniques (namely, the "beamlet" intensity and variables), thereby enabling a high degree of control over the shape and intensity of the high dose region within a patient [58]–[62]. This is reflected in the typical computational formulation of the planning problem in a large number of "beamlet" intensity variables that together comprise deliverable beams of radiation. Unfortunately, the added degrees of freedom greatly increase the complexity of treatment planning[63]–[66]. Conventional treatment planning algorithms attempt to cope with this complexity through an array of structure-specific weights and dosimetric goals that the user modifies during planning [67]. Currently, the IMRT treatment planning has not yet approached the ideal of automated, ‘inverse planning,’ and is typically a time-consuming, "trial-and-error" process with user-modified input parameters. Some approaches to improve this process have been proposed[68]–[75]. In particular, newer methods [69] based on goal programming methodologies that avoid the need for tuning parameters to a given patient. Another approach is to parallelize a simplified (quadratic) formulation of the planning problem, which might be particularly well-suited to fast, online adaptations of treatment plans [76]. But these improved methods still require the solution of large-scale optimization algorithms. Therefore, time efficiency remains an unachieved, but essential computational goal if it can be obtained at a negligible cost to accuracy.

During inverse planning for IMRT, radiation beams at a given gantry angle are further converted into discrete "beamlets," resulting in a beam intensity profile that can be decomposed into sub-fields delivered using a multi-leaf collimator. In state of the art algorithms, the dose that each beamlet contributes to each voxel is computed once for fluence of arbitrary intensity and stored in a large "influence matrix," which typically has millions of rows (number of voxels in the dose calculation volume) and thousands of columns (the number of beamlets), resulting in billions of elements. During optimization, the computer must multiply the influence matrix by the beamlet weight vector hundreds of times. Memory requirements are a challenge as well because accessing large memory blocks outside of on-chip cache is a slow process [77]–[80]. Additionally, since the solution time usually increases as a polynomial of the problem size, large influence matrices typically correspond to much longer solution times. Therefore, methods to reduce the size of the problem statement are needed.

The general idea of voxel sampling has been explored previously by Martin, et al. [81], who used sampled voxels to compute estimates of the gradient and objective functions to solve the optimization problem. Because doses for nearby voxels are related, some groups have utilized voxel clustering. For example, Scherrer, et al. [82] introduced an adaptive voxel clustering method to increase the speed of optimization; however, the method is complex and ensuring final solution quality can be difficult. The Computational Environment for Radiotherapy Research (CERR), in the ORART Toolbox graphical user interface (GUI) [83], provides for gridded down-sampling of low-dose beamlet scatter values. Rocha, et al. [84] studied the impact of different sampling rates in treatment planning using sampling methods available in CERR. Increasing the sampling rate in CERR results in larger voxel size and thus deteriorates the dose distribution in organs. Lu, et al. [85] studied the performance of random and grid sampling methods. Using various metrics they determined that grid sampling is relatively efficient and accurate. Another commonly used approach eliminates low-dose contributions from a given beamlet in the influence matrix [71], [86]. Voxel sampling has also been used in beam angle optimization. Ferris, et al. [87] developed an adaptive algorithm that uses three phases for beam angle optimization. The first phase uses coarse sampling to determine the most promising beam angles and is refined in subsequent phases.

Most of the sampling methods discussed in the literature are for unconstrained, optimization problem formulation. Some methods require multiple passes to refine the sampling process. These methods may not be suitable for the large-scale, constrained, optimization problem that often arises while automating the treatment plan.

In this paper, we propose a voxel sampling method; referred to as "Boundary and Interior Sampling," or "BIS," to reduce the computational time of the constrained IMRT optimization problem. We used a prioritized prescription algorithm[88] to demonstrate the BIS method. The prioritized prescription formulation is attractive because it offers greater control over the problem statement and the potential for more automation than the weighted sum objective function approach. The purpose of prioritization is also to ensure that higher priority dose-volume planning goals are not sacrificed to improve lower priority goals. The algorithm has four steps. The first optimizes dose to the target structures while sparing key sensitive organs from radiation. In the second, the algorithm finds the best beamlet weight to reduce toxicity risks to normal tissue while holding the objective function achieved in the first step as a constraint, with a small amount of allowed slip. Likewise, the third and fourth steps introduce lower priority normal tissue goals and beam smoothing, as further discussed below.

2.2 Methods

We tested the Boundary and Interior sampling (BIS) algorithm on seven different anonymized/non-identifiable head and neck treatment planning datasets. The treatment plans, after importing into CERR, had slices of 256 x 256 with voxels of size 0.2 × 0.2 cm; the slice thickness was 0.5 cm. We used seven equispaced beams of energy 6MV to create the treatment plan.

2.2.1 Optimization Formulation

The optimization formulation was based on hierarchical optimization[88]. Hierarchical optimization prioritizes objectives from most to least important and solves them one at a time. The optimization has four steps as discussed below.

Step I

The goal of the first step is to achieve good target coverage while satisfying the hard constraints in the normal structures. Let T be the set of target structures, R^I be the set of normal structures, D_i^{pre} be the prescription dose to the target structure i, and $D_j(w)$ be the actual dose to the voxel j of the target structure.

The dose to target structures was optimized in the first step in the presence of the hard constraints of the normal structures. Let S be the set of structures in the first step. Then,

$$S = \{i \mid i \in T \cup R^I\} \quad (2\text{-}1)$$

The dose $D_j(w)$ to the voxel j of a structure is given by

$$D_j(w) = \sum_{i=1}^{n} A_{ij} * w_j \qquad \forall i \in S \qquad (2\text{-}2)$$

Matrix A contains the dose to each voxel of structures due to the unit beamlet fluence and is called the influence matrix. The influence matrix is of the size m by n, where m indicates the total number of voxels in the structures and n indicates the total number of beamlets in the treatment.

Let V_i be the set of voxels before sampling, and V_i' be the set of voxels after sampling for the structure i. The optimization formulation is as follows:

Minimize
$$F^I(w) = \sum_{i \in T} \left(\frac{1}{|V_i'|} \sum_{j \in V_i'} [D_j(w) - D_i^{pre}]^2 \right) + \sum_{i \in T} t_i^2, \tag{2-3}$$

Subject to
$$D_i^{pre} - D_j(w) \le t_i \qquad \forall i \in T, \tag{2-4}$$

$$0.05 * \mathrm{D}_i^{pre} \le t_i \qquad \forall i \in T, \tag{2-5}$$

$$D_i^{min} \le D_j(w) \le D_i^{max} \qquad \forall i \in T, j \in V_i', \tag{2-6}$$

$$D_j(w) \le D_i^{max} \qquad \forall i \in R^I, j \in V_i', \tag{2-7}$$

$$w_i \le w^{max}. \tag{2-8}$$

The under-dose region is undesirable in the target structure. The under-dose region of the target structure was penalized by putting the upper bound (t_i) on the deviation of the target dose in (2-4) and minimizing the upper bound, t_i, in the objective function. Finally, the maximum dose to the target and normal structures was constrained in (2-6) and (2-7).

Step II

The goal of the second step is to minimize dose to the brainstem while preserving the solution of the first step. We used the mean dose to α% of the brainstem voxels to reduce dose in the high dose region of the brainstem. Similarly, we used the constraints to preserve the dose to the 95% region of the target structure. Although the “no-sampling” (NS) approach uses only the constraints to preserve the target dose, the BIS approach uses the additional penalty term in the objective function to maintain the homogeneity of the target dose. In NS, the constraints were applied to each voxel of the target structures to maintain homogenous dose distribution to target structures, whereas, in the BIS method, the constraints were applied only to the sampled voxels of target structures. The BIS method had no constraints to maintain homogeneous dose to non-sampled regions of target structures; thus BIS produced non-homogenous dose distribution to target structures. To prevent dose deviations in the non-sampled region of target structures, the targets voxels dose deviates from the maximum to the minimum and is penalized in the objective function.

Let R^{II} denote the set of structures in the second step, F^{I} denote the objective function of the first step, and w^{I} denote the optimal beamlet weight of the first step. The optimization formulation is given by:

Minimize

$$F^{II}(w) = \sum_{i \in R^{II}} \left(y_i^{\alpha} + \frac{1}{(1-\alpha)\left|V_i'\right|} \sum_{j=1}^{|V_i'|} p_{ji}^{\alpha} \right) + \sum_{i \in T} \sum_{j \in (V_i - V_i')} \Big(\gamma * \left(\min(0, D_i^{max} - D_j)^2 + \varphi * \max(0, D_i^{min} - D_j \right)^2 \Big), \quad (2\text{-}9)$$

Subject to

$$D_j - z_{ij} = 0 \qquad \forall \mathrm{i} \in R^{II}, \mathrm{j} \in V_i', \quad (2\text{-}10)$$

$$p_{ji}^{\alpha} \geq 0, \quad (2\text{-}11)$$

$$p_{ji} + y_i^{\alpha} - z_{ji} > 0 \qquad \forall i \in R^{II}, \mathrm{j} \in V_i', \quad (2\text{-}12)$$

$$F^{I}(w) \leq (1+s) * F^{I}(w^{I}) \quad (2\text{-}13)$$

$$D_i^{min} \leq D_j(w) \leq D_i^{max} \qquad \forall i \in T, j \in V_i', \tag{2-14}$$

$$D_j(w) \leq D_i^{max} \qquad \forall i \in R^I, j \in V_i', \tag{2-15}$$

$$w_i \leq w^{max}. \tag{2-16}$$

The formulation of step II is based on the mean-tail-dose formulation[89], [90]. The mean-tail-dose formulation introduces the artificial variable p, y and z to minimize the mean dose to the hottest α% region ((2-9) - (2-12)). The parameters γ and φ in (2-9) control the deviation of the target dose from the maximum and minimum dose.

The slip parameter "s" was used in (2-13) to relax the target dose distribution. The target dose was relaxed by almost 60% in the second step. The maximum dose to target and normal structures was constrained by the equations (2-14) and (2-15).

Step III

The third step minimizes the average dose to normal structures while preserving the solution of the first and second step. Following is the optimization formulation for the third step:

Minimize

$$F^{III}(w) = \sum_{i \in R^{III}} \frac{1}{|V_i'|} \sum_{j \in V_i'} D_j(w) + \sum_{i \epsilon T} \sum_{j \in (V_i - V_i')} \Big(\gamma * \big(\min(0, D_i^{max} - D_j)^2 + \varphi * \max(0, D_i^{min} - D_j)^2\Big), \tag{2-17}$$

Subject to

$$D_j - z_{ij} = 0 \qquad \forall \mathrm{i} \in R^{II}, \mathrm{j} \in V_i', \tag{2-18}$$

$$p_{ji}^{\alpha} \geq 0, \tag{2-19}$$

$$p_{ji} + y_i^{\alpha} - z_{ji} > 0 \qquad \forall i \in R^{II}, \mathrm{j} \in V_i', \tag{2-20}$$

$$\sum_{i \in R^{II}} \left(y_i^{\alpha} + \frac{1}{(1-\alpha)|V_i'|} \sum_{j=1}^{|V_i'|} p_{ji}^{\alpha} \right) \leq M_{\alpha}^{*i} \qquad \forall i \in R^{II}, \tag{2-21}$$

$$F^I(w) \leq (1+s)^2 * F^I(w^I), \tag{2-22}$$

$$D_i^{min} \leq D_j(w) \ \leq D_i^{max} \qquad \forall i \in T, j \in V_i', \tag{2-23}$$

$$D_j(w) \ \leq D_i^{max} \qquad \forall i \in R^I, j \in V_i', \tag{2-24}$$

$$w_i \leq w^{max}. \tag{2-25}$$

The constraints (2-18) to (2-21) preserve the mean dose to the hottest α% region of the normal structure obtained in the second step (M_{α}^{*i}). Mathematically M_{α}^{*i} is given by,

$$M_{\alpha}^{*i} \ = \sum_{i \in R^{II}} \left(y_i^{*\alpha} + \frac{1}{(1-\alpha)|V_i'|} \sum_{j=1}^{|V_i'|} p_{ji}^{*\alpha} \right) \qquad \forall i \in R^{II} \tag{2-26}$$

Similarly, the constraints (2-21) and (2-22) preserved the dose to the target structure obtained in the first step. The slip parameter was doubled to relax the target dose constraints so that the dose to the normal structure could be improved.

Step IV

The fourth step smooths the beam profiles, removes the hot spots outside the target structure and produces the deliverable plan. The optimization formulation is as follows:

Minimize
$$F^{IV}(w) = \sum_{i=1}^{n} w_i^2, \tag{2-27}$$

Subject to

$$\sum_{i \in R^{II}} \left(y_i^\alpha + \frac{1}{(1-\alpha)|V_i'|} \sum_{j=1}^{|V_i'|} p_{ji}^\alpha \right) \leq M_\alpha^{*i} \quad \forall i \in R^{II}, \tag{2-28}$$

$$D_j - z_{ij} = 0 \qquad \forall \mathrm{i} \in R^{II}, \mathrm{j} \in V_i, \tag{2-29}$$

$$p_{ji}^\alpha \geq 0 \tag{2-30}$$

$$p_{ji} + y_i^\alpha - z_{ji} > 0 \qquad \forall i \in R^{II}, \mathrm{j} \in V_i', \tag{2-31}$$

$$F^I(w) \leq (1+s)^3 * F^I(w^I), \tag{2-32}$$

$$D_i^{min} \leq D_j(w) \ \leq D_i^{max} \qquad \forall i \in T, j \in V_i', \tag{2-33}$$

$$D_j(w) \ \leq D_i^{max} \qquad \forall i \in R^I, j \in V_i', \tag{2-34}$$

$$\sum_{i \in R^{III}} \frac{1}{|V_i|} \sum_{j \in V_i} D_j(w) \leq M_{III}^*, \tag{2-35}$$

$$w_i \leq w^{max}. \tag{2-36}$$

Note that the constraints (2-28) - (2-34) in the fourth step are same as in the third to preserve the solution of the first and second steps. The only difference is in the slip parameter, which was increased to $(1+s)^3$ to relax the target dose constraint. The fourth step has added a constraint (2-35) to preserve the average dose to the normal structures (M_{III}^*) obtained in the third. Mathematically, the average dose to the normal structure (M_{III}^*) is given by,

$$M_{III}^{*} = \sum_{i \in R^{III}} \frac{1}{|V_i|} \sum_{j \in V_i} D_j(w*) \tag{2-37}$$

2.3 Smaller Slip value for Sampling

We found that the slip value, a parameter used in hierarchical optimization, should be kept smaller in the BIS formulation than in the NS formulation. Hierarchical optimization uses constraints to preserve the solution of the higher priority goals while solving for the lower priority goals. The result of the higher priority goal is multiplied by a slip factor 1+s (s>0) to relax constraints so that the solution of the lower priority goals may be improved. Since BIS degrades the solution of higher priority goals compared to NS, the value of "s" was decreased from 1 to 0.6 to preserve the solution of higher priority goals.

2.4 Sampling scheme

Because photon beams must pass through boundaries and boundaries are of a lower dimension than volumes, they are attractive for inclusion in non-uniform sampling schemes. We, therefore, include all boundary voxels in our sampling methods. A voxel is considered a boundary voxel if there are no voxels surrounding it in at least one of the four directions in 2D (left, right, up, down) from the same region. Otherwise, the voxel is considered an interior voxel.

A k-means clustering algorithm was used to sample inner-voxels of targets and critical normal structures (i.e., structures like the brainstem, which have a hard maximum dose constraint) in quasi-random fashion. The k-means clustering algorithm forms clusters of voxels that are similar to each other and, therefore, requires a similarity measure; the influence matrix was used to estimate voxel similarity.

If p is the percentage of inner-voxels that will be selected in quasi-random fashion, then the number of voxels to be sampled from each cluster is the rounded integer of the number of voxel points divided by p. Thus, the actual sampling rate is discontinuous as a function of p. This discontinuity is ameliorated for larger sampling rates of p by using a larger cluster size, but is unavoidable for smaller values of p, if we remain consistent with a quasi-random coverage of the structure. Given the goal of achieving quasi-uniform sampling, the number of clusters is reduced when the sampling rate is high. This leads to relatively uniform coverage, irrespective of whether there are fewer or a larger number of clusters with a lower sampling rate.

In practice, the number of clusters is calculated by dividing the total number of voxels in a region of interest by the minimum number of voxels in each cluster. The minimum number of voxels in each cluster equals the percentage of inner-voxels (p) that are to be sampled. Therefore, the number of clusters for a structure is given by:

$$\text{Number of clusters} = {|V_i|}/{p} \qquad \text{(2-38)}$$

2.5 Experimental Results

We ran the BIS and NS algorithms in seven head and neck treatments plans to compare the dose quality, computational time and memory required for the optimization process. We found that BIS could reduce the computational time and memory while maintaining the dose quality.

2.5.1 Determining the best Sampling Rate

We plotted the average root mean square error and the optimization run time to determine the best inner voxel sampling rate (Figure 2-1). Including a small fraction of interior voxels (around 5%) decreased the root mean square error from 2.5% to less than 1%. As seen in Figure 2-1, although the sampling rate was increased, the average root mean square error was almost the same when the sampling rate was greater than 5%. Therefore, we chose the sampling rate based on the optimization run time.

We observed that the sampling rate of 10% took the least amount of time to solve the optimization problem. The optimization time was the highest when no inner voxels were included in the optimization because our method controls the dose to the non-sampled regions of target structures to produce a homogenous dose distribution after the first step. When target dose was optimized with the boundary voxels and no inner voxels (i.e., a sampling rate of 0%), a large fraction of the compute time in later steps was spent adjusting the dose to the non-sampled regions of target structures. However, the time to fix the target dose was small when a fraction of the inner voxels was included in the first step (Figure 2-1).

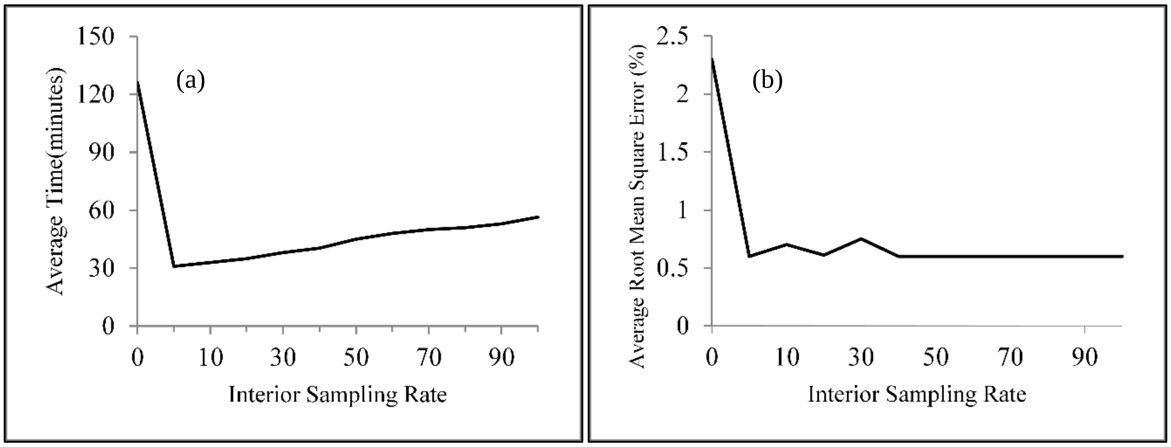


Figure 2-1 IMRT optimization time with the different sampling rates, (a) the average root mean square error of the targets and brainstem for the sampling rate 0 to 100, (b) The boundary voxels were always included, and inner voxels were sampled at the different sampling rates. The 10% sampling rate gave the best trade-off in the optimization time and the solution quality. Therefore, we chose a sampling rate of 10% for the optimization.

2.5.2 Run Time of BIS and No-Sampling Methods

We used the time reported from the optimization solver to compare the run time of BIS and the no-sampling method. We found that BIS produced different speed-ups depending on the treatment data. The minimum speedup was a factor of 1.80 for treatment plan 3; the maximum was 2.64 for treatment plan 5. Thus, BIS produced an average speed-up of 2.04 compared with standard sampling (Table 2-1).

Table 2-1 The speed-up in the optimization time using the boundary and interior sampling (BIS) method (for sampling rate 10%) compared with the no-sampling method is tabulated; the sampling method was faster than the no-sampling method.

	Time (Minutes)		
Plan	Sampling	No-sampling	Speedup
1	5.11	11.56	2.26
2	16.83	25.41	1.50
3	4.62	8.33	1.80
4	8.74	17.21	1.96
5	3.8	10.04	2.64
6	6.53	15.05	2.30
7	13.67	25.01	1.82

2.5.3 Dose Distribution of BIS and No-Sampling Methods

We plotted dose-volume histograms and dose-color-wash displays (Figure 2-3 to Figure 2-8) for head and neck sites to compare the dose obtained from BIS and no-sampling (NS) methods. The first step in the process optimized the dose to target structures. To compare the doses to target structures, we computed D95 of each of the target structures using the treatment plan obtained from BIS and NS.

The D95 obtained from BIS was almost the same as that obtained from NS. Out of seven head and neck treatment plans, one treatment plan had lower, one had equal, and five treatment plans had higher D95 from BIS than from NS. The D95 obtained from BIS was 0.29% lower than that obtained from the NS for one treatment plan. Similarly, on average, the D95 obtained from BIS was 0.36% higher than that obtained from the NS for the five treatment plans. The results show that the target coverage from the sampling is comparable with that of NS.

The second step of hierarchical optimization minimized the mean dose of the hottest 1% (MOHx) of the brainstem voxels. The dose to the brainstem obtained from BIS was comparable to that obtained from the NS. On average, the MoHx obtained from BIS was 0.6% higher than that obtained from the NS for the two treatment plans. The results show that the dose to the brainstem obtained from BIS was equivalent to that of NS.

The third step of hierarchical optimization minimized the mean dose to normal structures. There were a number of normal structures in the optimization and we chose the spinal cord as a representative structure to show the dose to the normal structure. The dose to the spinal cord obtained from BIS was comparable to that obtained from NS. On average, the mean dose to spinal cord obtained from BIS was 5.19% less than that obtained from NS. The result indicates that BIS minimized radiation exposure to the normal structures.

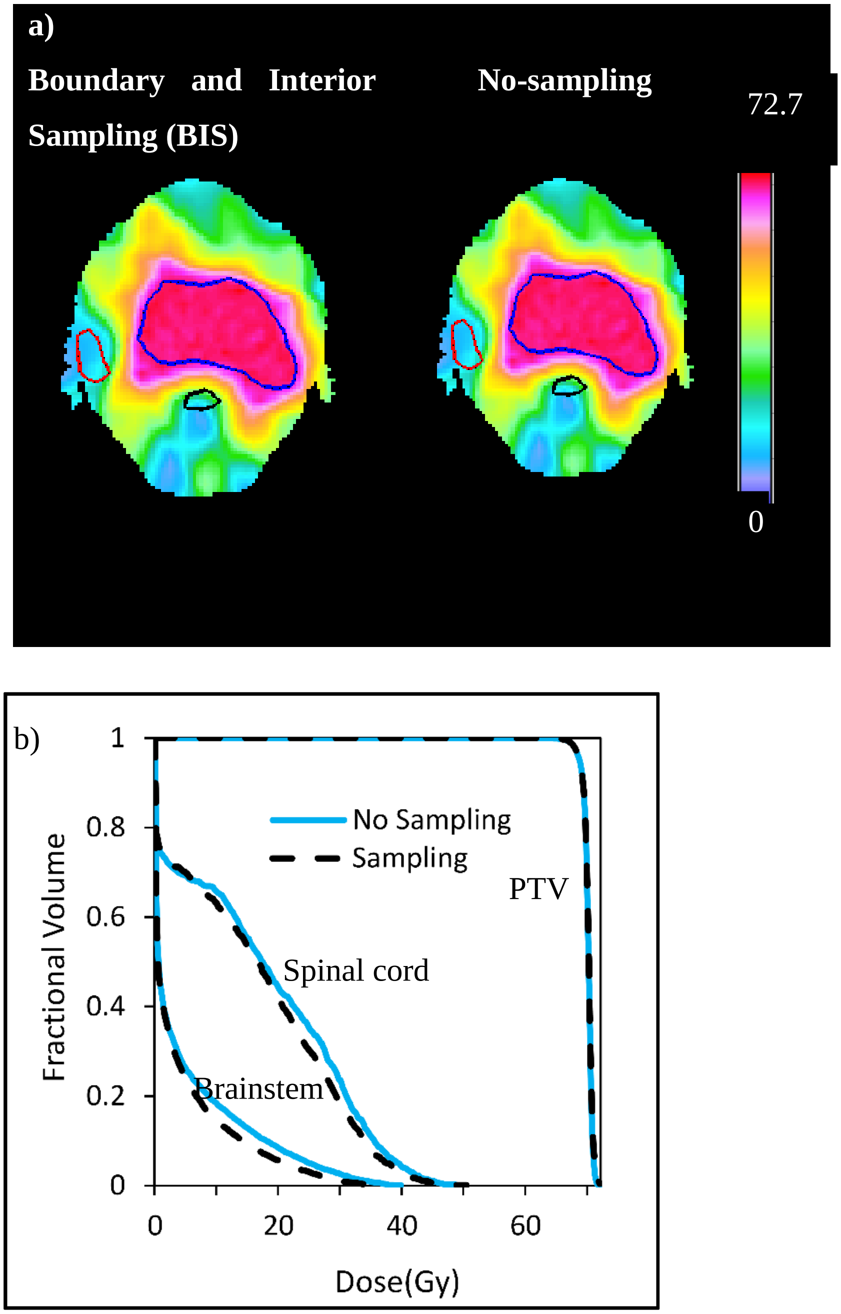


Figure 2-2 a) Dose color wash and b) dose-volume histogram produced in the boundary and interior sampling (BIS) and no-sampling methods for head and neck treatment plan 1. The BIS and no-sampling methods produced almost the same dose distribution. Therefore, the dose-volume histogram of BIS and no-sampling almost overlap and the dose color wash produced in BIS and no-sampling have similar colors.

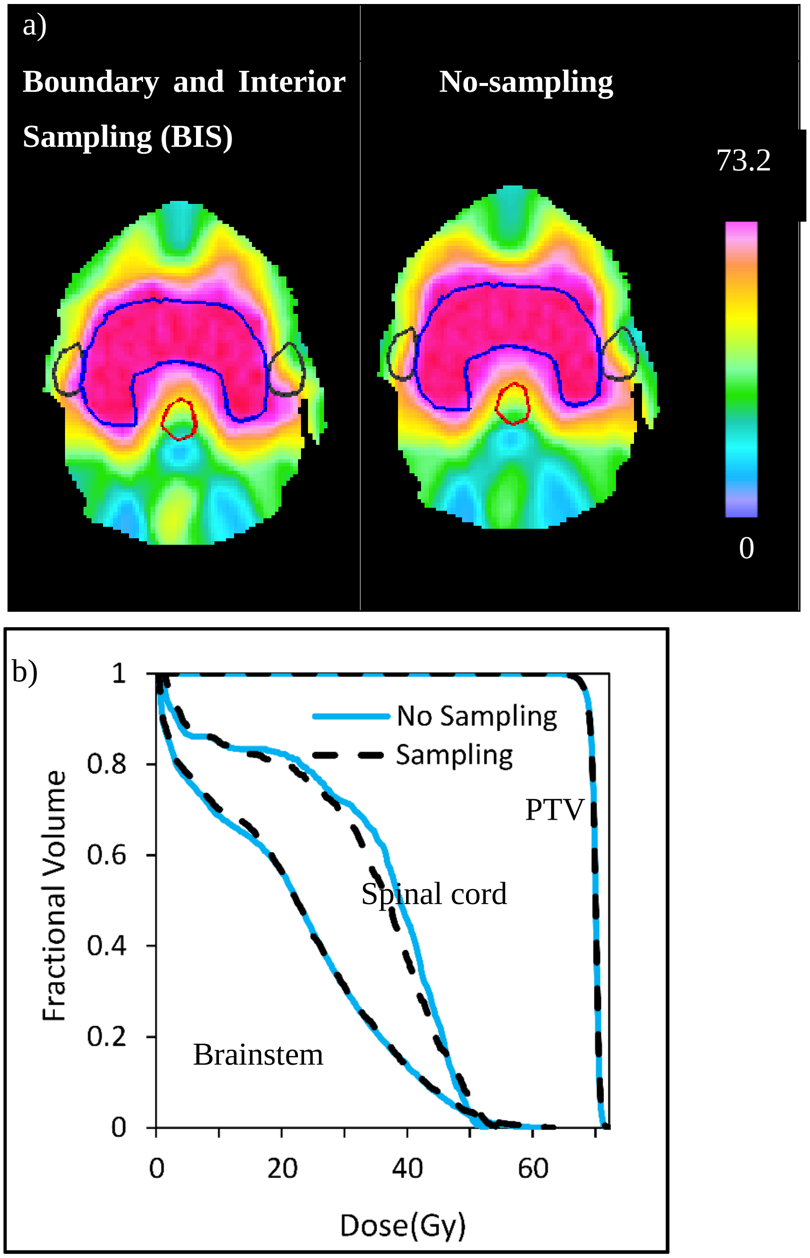


Figure 2-3 a) Dose color wash and b) dose-volume histogram produced in the boundary and interior sampling (BIS) and no-sampling methods for head and neck treatment plan 2. The BIS and no-sampling methods produced almost the same dose distribution. Therefore, the dose-volume histogram of BIS and

no-sampling almost overlap and the dose color wash produced in BIS and no-sampling have similar colors.

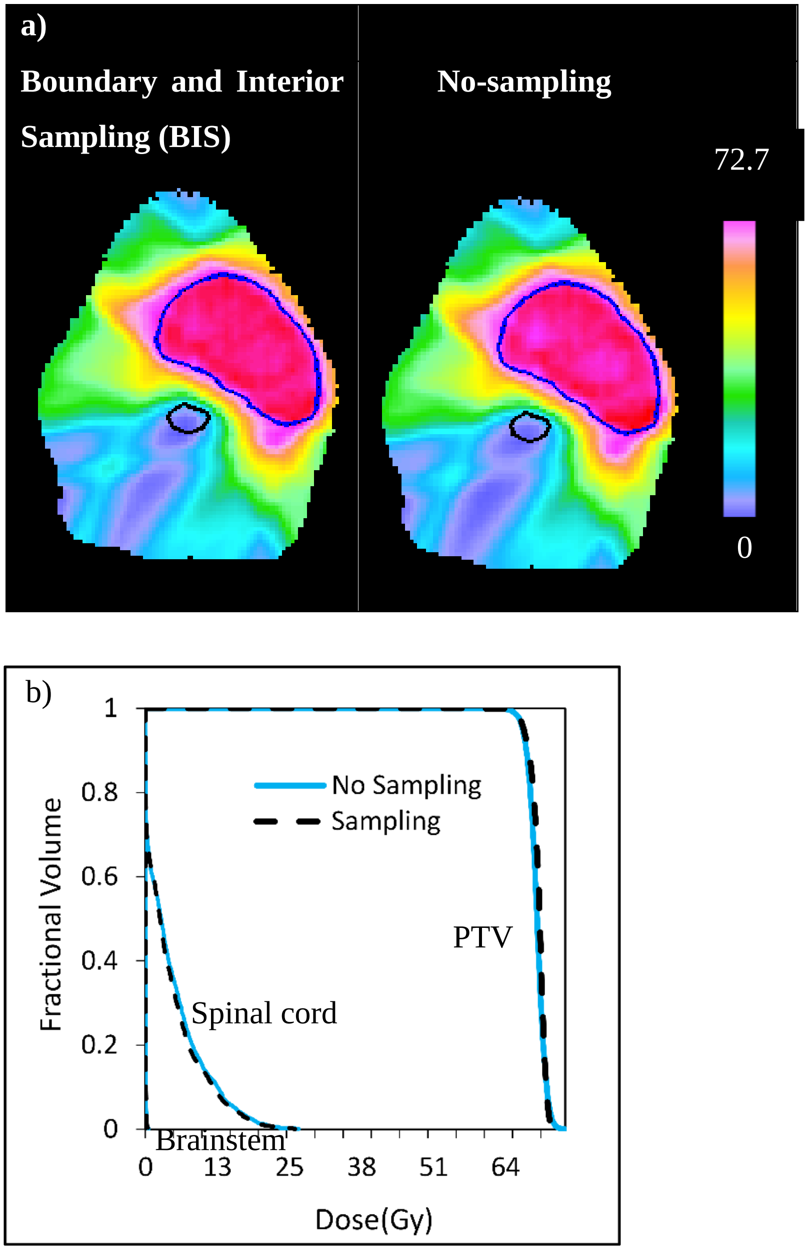


Figure 2-4 a) Dose color wash and b) dose-volume histogram produced in the boundary and interior sampling (BIS) and no-sampling methods for head and neck treatment plan 3. The BIS and no-sampling methods produced almost the same dose distribution. Therefore, the dose-volume histogram of BIS and

no-sampling almost overlap and the dose color wash produced in BIS and no-sampling have similar colors.

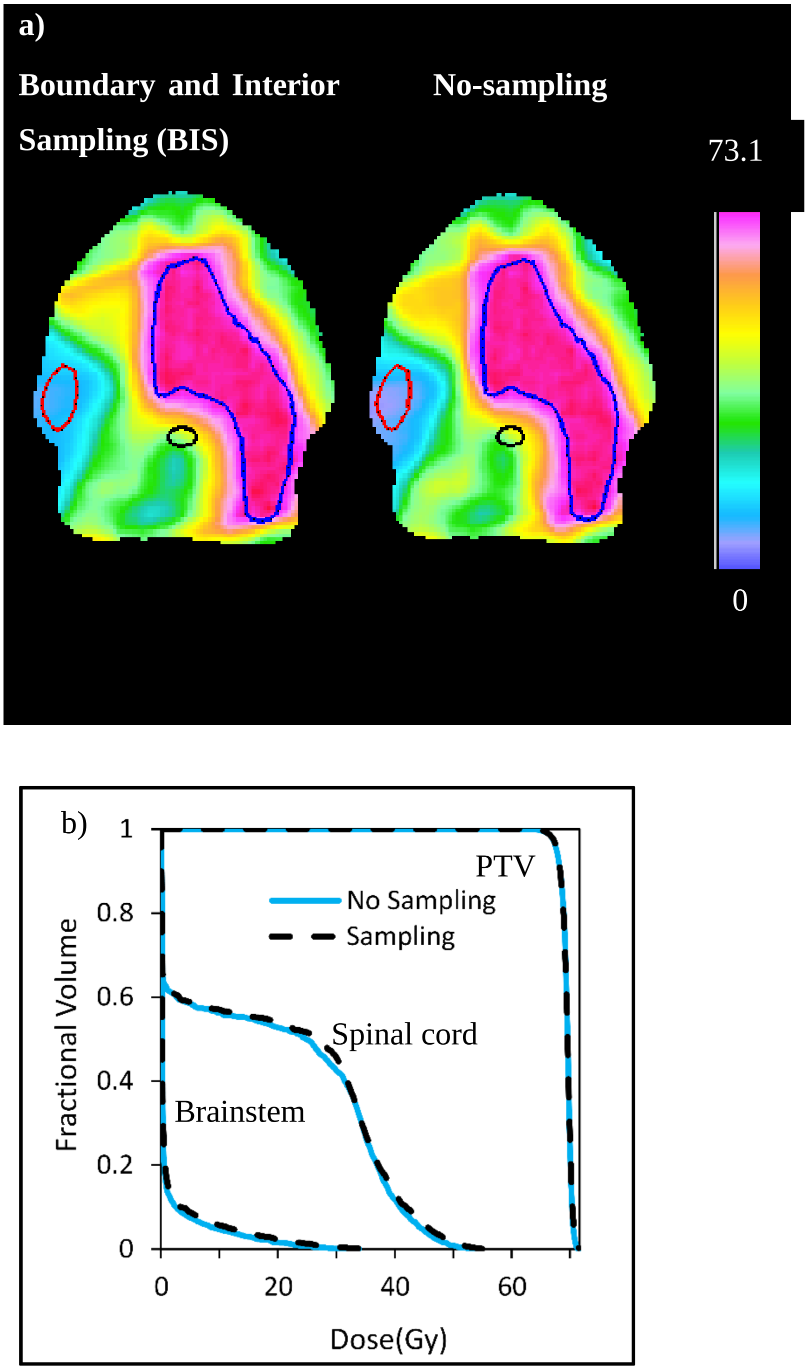


Figure 2-5 a) Dose color wash and b) dose-volume histogram produced in the boundary and interior sampling (BIS) and no-sampling methods for head and neck treatment plan 4. The BIS and no-sampling methods produced almost the same dose distribution. Therefore, the dose-volume histogram of BIS and

no-sampling almost overlap and the dose color wash produced in BIS and no-sampling have similar colors.

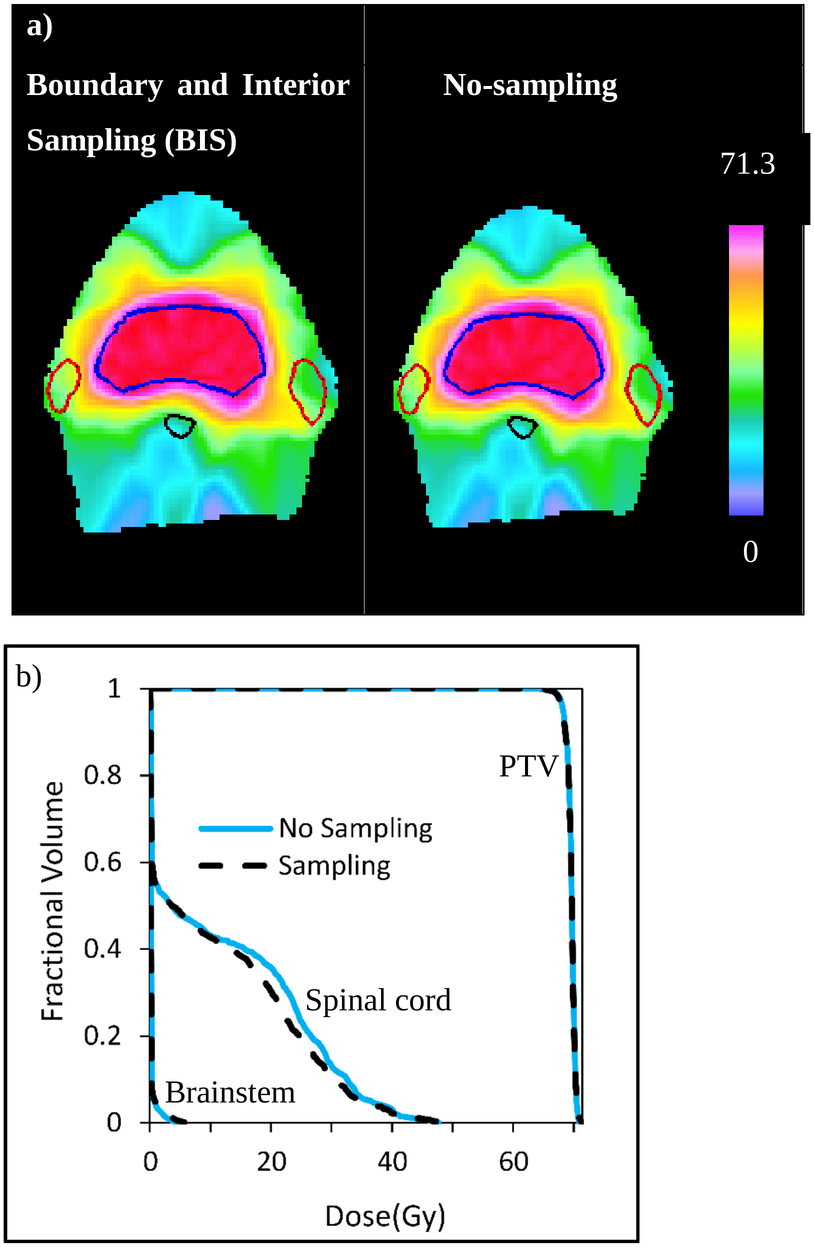


Figure 2-6 a) Dose color wash and b) dose-volume histogram produced in the boundary and interior sampling (BIS) and no-sampling methods for head and neck treatment plan 5. The BIS and no-sampling methods produced almost the same dose distribution. Therefore, the dose-volume histogram of BIS and no-sampling almost overlap and the dose color wash produced in BIS and no-sampling have similar colors.

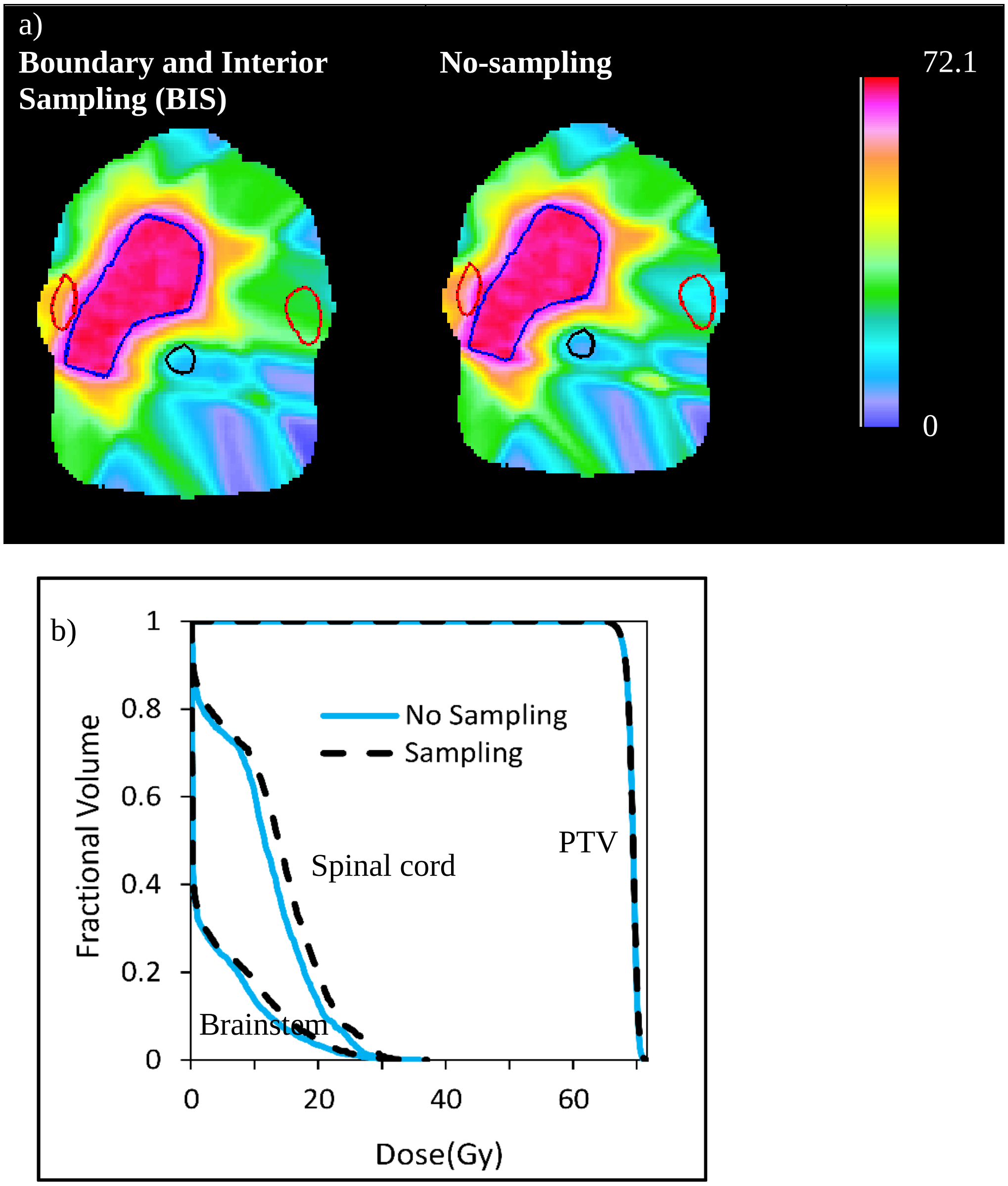


Figure 2-7 a) Dose color wash and b) dose-volume histogram produced in the boundary and interior sampling (BIS) and no-sampling methods for head and neck treatment plan 6. The BIS and no-sampling methods produced almost the same dose distribution. Therefore, the dose-volume histogram of BIS and

no-sampling almost overlap and the dose color wash produced in BIS and no-sampling have similar colors.

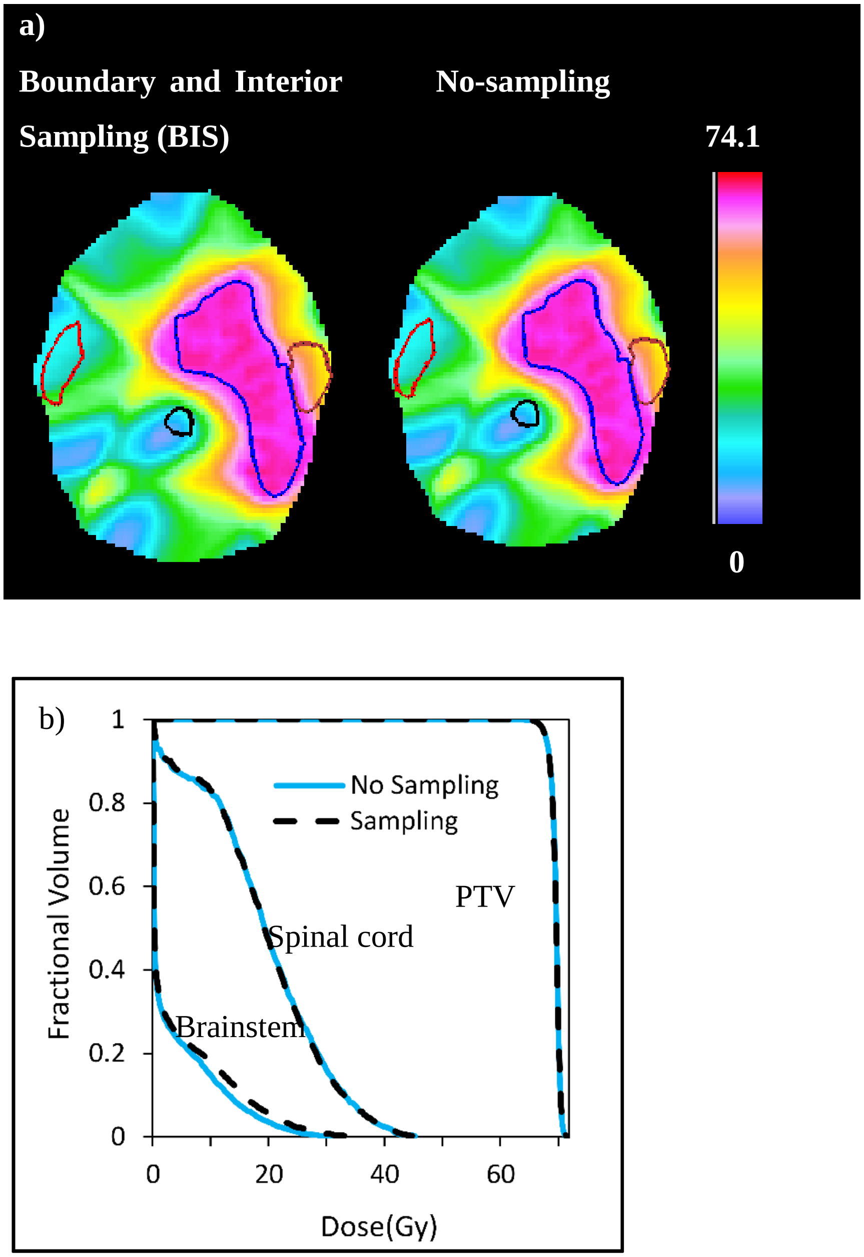


Figure 2-8 a) Dose color wash and b) dose-volume histogram produced in the boundary and interior sampling (BIS) and no-sampling methods for head and neck treatment plan 7. The BIS and no-sampling methods produced almost the same dose distribution. Therefore, the dose-volume histogram of BIS and no-sampling almost overlap and the dose color wash produced in BIS and no-sampling have similar colors.

2.6 Memory Used by BIS and No-Sampling Methods

We collected the computational memory used by BIS and NS to compare the memory usage of these methods. The BIS method reduced the problem size, which ultimately decreased the memory used in the optimization by 40% (

Figure 2-9).

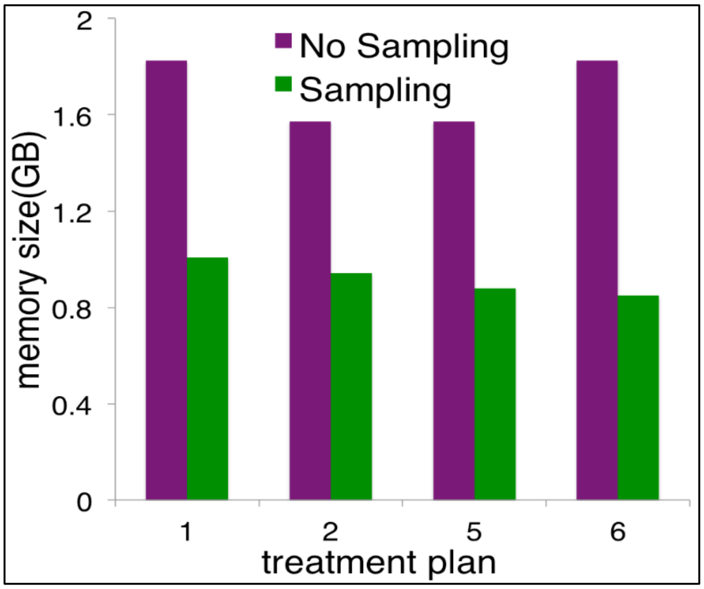


Figure 2-9 Memory used by the boundary and interior sampling (BIS) and no-sampling methods in IMRT optimization. The BIS method has less voxels compared to the NS methods. Therefore, BIS used less memory than the NS method.

2.7 Discussion

This study demonstrates that Boundary and Interior Sampling (BIS) could significantly reduce the time and memory used by large-scale, constrained, IMRT optimization without substantially degrading dose distributions to organs. The average speedup of the sampling was 2.04 compared to that of no sampling (NS). The difference in D95 to the target structure by the sampling was less than 1%, and sampling produced a lower dose to normal organs than that produced by NS. We hypothesized that the most important voxels for controlling interior dose distributions were actually the boundary voxels. Optimal results were obtained by including approximately 10% of inner voxels sampled using quasi-random cluster sampling to reduce the occurrence of large "holes" in the regions of interest.

The sampling method proposed by Martin, et al.[81] used the steepest descent algorithm[91] to solve the optimization problem. However, the steepest descent algorithm cannot solve the constrained optimization problem. One way to use the steepest descent algorithm to solve the constrained optimization is to project the gradient into the feasible space to satisfy the constraints in the problem. The projected gradient may not give a good decrease in the objective function and the algorithm might require large iterations to converge to an optimal solution. In contrast, BIS solved the large-scale constrained optimization problem using the open source IPOPT [1] optimization solver.

Scherrer, et al. [82] proposed a voxel clustering method that requires knowledge about the solution space to cluster the voxels. They solved an approximation of the original problem to probe the solution space, which is a challenging task. Boundary and Interior sampling does not require complex initialization steps and can directly solve the original problem.

The sampling method proposed by Ferris, et al. [87] solves the problem in three phases. They used coarse sampling in the first phase and refined the sampling in subsequent phases. The Ferris, et al. method requires multiple passes to solve the optimization problem, but our sampling method does not; thus making it an attractive choice to solve a multi-step optimization problem.

In summary, the Boundary and Interior Sampling (BIS) method was 1.8 to 2.64 times faster than the standard approach with NS. As a consequence, BIS used around 40-50% of the memory needed without sampling. Dose distribution metrics (such as D95) computed using BIS were comparable (typically within 1%) with that of the NS method. It is recommended to use a full dose calculation after optimization to create reportable metrics.

2.8 Conclusions

The computational complexity of the IMRT treatment planning problem, especially promising but complex approaches such as prioritized prescription optimization, motivated our proposed sampling method. We have demonstrated, for a range of treatment plans that a combination of boundary and partial, quasi-random, interior sampling results in a high-quality IMRT solution.

Chapter 3: Integrating High-Performance Nonlinear Solver

3.1 Introduction

The previous chapter studied the sampling method to reduce the computational time of hierarchical optimization. Hierarchical optimization is a large-scale nonlinear optimization problem. The efficient nonlinear solver is needed to solve such a large-scale optimization problem. This chapter focuses on adopting the general purpose open source nonlinear solver to solve the hierarchical optimization problem.

There are several commercial and open-source treatment planning applications to perform IMRT optimization [92]–[94]. Most of these applications use the weighted sum of objective functions for optimization. The weighted sum approach of solving the treatment planning optimization makes the treatment planning process a manual trial and error process. Therefore, there has been considerable research in automating the treatment planning process using hierarchical optimization [69], [72], [73]. However, automating the treatment planning application requires the ability to solve the large-scale constrained optimization problem, which is why we adopted the open source Interior Point Optimization (IPOPT) solver. We studied the memory access pattern of the large-scale matrices and reduced the computational time required to solve the optimization problem. Our application is integrated with the Computational Environment for Radiotherapy Research (CERR), a widely used software platform to develop and share radiation therapy treatment planning research[95]; thus providing a complete flow of information for the IMRT treatment planning process. We found that improved memory access reduced memory latency, which ultimately decreased the IMRT optimization time.

3.2 Methods

We developed a software application in C++ to solve the large-scale nonlinear hierarchical optimization and investigated the impact of the memory access pattern in run time of the optimization. We prepared treatment plans of seven head and neck plans using the software application.

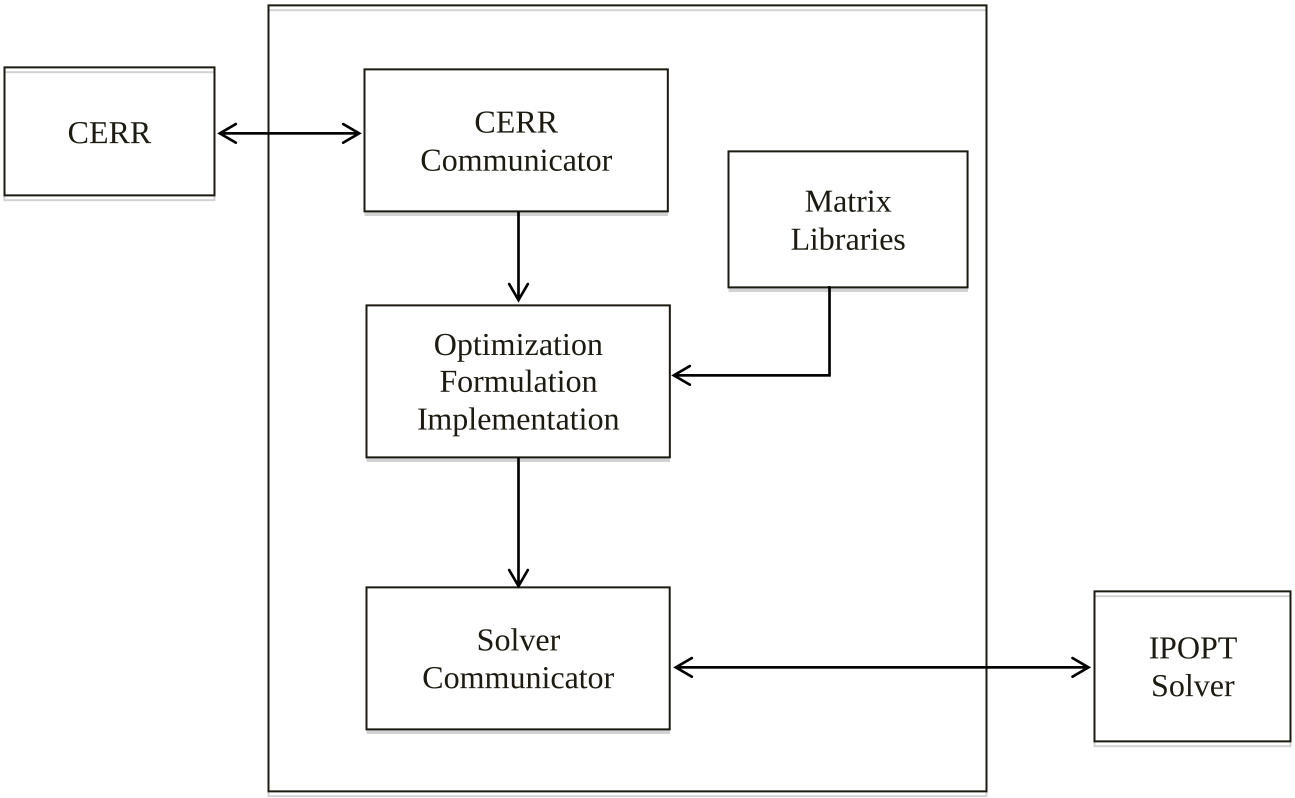


Figure 3-1. Illustration of the treatment planning application. The application communicates with Computational Environment for Radiotherapy Research (CERR) to read input data for optimization and uses an Interior Point Optimizer (IPOPT) solver to perform the optimization.

The application has four different modules: the CERR communicator, Optimization Formulation Implementation, Matrix Libraries and the Solver Communicator (Figure 3-1). The CERR communicator reads the influence matrix, which relates beamlet fluence to voxel dose as generated in the CERR using the ORART toolbox. The optimization formulation module implements the hierarchical optimization formulation, which is a preemptive goal programming formulation[88]. The optimization problem was solved using the IPOPT[96] solver.

3.3 Optimization Formulation

The optimization formulation is based on hierarchical optimization as discussed in section 2.1. The hierarchical optimization formulation prioritizes the treatment planning goals and solves one goal at a time. There are four steps in the optimization. The first step optimizes the dose to target structure. The second optimizes dose to critical structures while using constraints to preserve the dose to the target structures. The later steps of optimization introduce lower priority goals such as optimizing dose to normal structure and smoothing fluence profile.

Step I

The first step produces good target coverage while satisfying maximum dose constraints of normal structures. Let T be the set of target structures, R^I be the set of normal structures, and V_i be the set of voxels in the structure i, then the set of all structures "S" in the first step of optimization is given by:

$$S = \{i \mid i \in T \cup R^I\} \qquad (3\text{-}1)$$

To calculate dose to each structure, structures are discretized into voxels and beams are discretized into beamlets. If there are n number of beamlets in the treatment planning, then the dose to voxel j of a structure i is given by:

$$D_j(w) = \sum_{i=1}^{n} A_{ij} * w_j \qquad \forall i \in S \qquad (3\text{-}2\ )$$

The matrix "A" is m by the n influence matrix. The number of rows in the matrix equal to the number of voxels in the structures and the number of columns equal to the number of beamlets in the treatment.

The first step minimizes the difference between the actual dose and the prescribed dose to produce good target coverage. The formulation is given below:

$$Minimize\ F^{I}(w) = \sum_{i\in T}\left(\frac{1}{|V_i|}\sum_{j\in V_i}[D_j(w) - D_i^{pre}]^2\right) + \sum_{i\in T} t_i^2, \tag{3-3}$$

$Subject\ to$

$$D_i^{pre} - D_j(w) \le t_i \qquad \forall i \in T, \tag{3-4}$$

$$0.05 * D_i^{pre} \le t_i \qquad \forall i \in T, \tag{3-5}$$

$$D_i^{min} \le D_j(w) \le D_i^{max} \qquad \forall i \in T, j \in V_i', \tag{3-6}$$

$$D_j(w) \le D_i^{max} \qquad \forall i \in R^{I}, j \in V_i', \tag{3-7}$$

$$w_i \le w^{max}\ . \tag{3-8}$$

Step II

Step II minimizes the mean dose to the hottest α%(MoHα) region of critical structures. The MoHα formulation computes the mean dose to the hottest α% by using the artificial variables p, z, and y in the formulation. The formulation of Step II is given by:

$$Minimize \qquad F^{II}(w) = \sum_{i\in R^{II}}\left(y_i^{\alpha} + \frac{1}{(1-\alpha)|V_i'|}\sum_{j=1}^{|V_i'|} p_{ji}^{\alpha}\right), \tag{3-9}$$

$Subject\ to$

$$D_j - z_{ij} = 0 \qquad \forall i \in R^{II}, j \in V_i, \tag{3-10}$$

$$p_{ji}^{\alpha} \ge 0\ , \tag{3-11}$$

$$p_{ji} + y_i^{\alpha} - z_{ji} > 0 \qquad \forall i \in R^{II}, j \in V_i, \tag{3-12}$$

$$F^I(w) \leq (1+s) * F^I(w^I), \tag{3-13}$$

$$D_i^{min} \leq D_j(w) \leq D_i^{max} \qquad \forall i \in T, j \in V_i', \tag{3-14}$$

$$D_j(w) \leq D_i^{max} \qquad \forall i \in R^I, j \in V_i', \tag{3-15}$$

$$w_i \leq w^{max}. \tag{3-16}$$

The objective function of the first step was allowed to degrade, by a small factor (1+s), as shown in equation (3-13). The small factor was called slip, and the value of slip (1+s) was two. Thus, slip effectively doubled the optimal value of the first step objective function.

Step III

The third step minimizes the average dose to normal structures while using constraints to preserve the result of the first and second steps. The optimization formulation is given by:

$$Minimize \qquad F^{III}(w) = \sum_{i \in R^{III}} \frac{1}{|V_i|} \sum_{j \in V_i} D_j(w) , \tag{3-17}$$

$Subject\ to$

$$\sum_{i \in R^{II}} \left(y_i^{\alpha} + \frac{1}{(1-\alpha)|V_i'|} \sum_{j=1}^{|V_i'|} p_{ji}^{\alpha} \right) \leq M_{\alpha}^{*i} \qquad \forall i \in R^{II}, \tag{3-18}$$

$$D_j - z_{ij} = 0 \qquad \forall i \in R^{II}, j \in V_i, \tag{3-19}$$

$$p_{ji}^{\alpha} \geq 0, \tag{3-20}$$

$$p_{ji} + y_i^{\alpha} - z_{ji} > 0 \qquad \forall i \in R^{II}, j \in V_i, \tag{3-21}$$

$$F^I(w) \leq (1+s)^2 * F^I(w^I), \tag{3-22}$$

$$D_i^{min} \leq D_j(\mathrm{w}) \leq D_i^{max} \qquad \forall i \in T, j \in V_i, \tag{3-23}$$

$$D_j(w) \leq D_i^{max} \qquad \forall i \in R^I, j \in V_i, \tag{3-24}$$

$$w_i \leq w^{max}. \tag{3-25}$$

Note that the result of the first step was preserved using constraints (3-22) – (3-24). The slip value was increased from (1+s) in the second step to (1+s)^2 in the third. The equations (2-18) to (2-21) constrained the mean dose to the hottest α% region of the critical structure obtained in the second step. The optimal value of the mean dose to the hottest α% region of the critical structure is given by:

$$M_\alpha^{*i} = \sum_{i \in R^{II}} \left(y_i^{*\alpha} + \frac{1}{(1-\alpha)|V_i'|} \sum_{j=1}^{|V_i'|} p_{ji}^{*\alpha} \right) \quad \forall i \in R^{II} \tag{3-26}$$

Step IV

The fourth step smooths the beam profile to remove hot regions outside the target structures while preserving the result of the first, second and third steps. The optimization formulation is given by:

$$Minimize \qquad F^{IV}(w) = \sum_{i=1}^{n} w_i^2, \tag{3-27}$$

$Subject\ to$

$$\sum_{i \in R^{II}} \left(y_i^{\alpha} + \frac{1}{(1-\alpha)|V_i'|} \sum_{j=1}^{|V_i'|} p_{ji}^{\alpha} \right) \leq M_\alpha^{*i} \quad \forall i \in R^{II}, \tag{3-28}$$

$$D_j - z_{ij} = 0 \qquad \forall i \in R^{II}, j \in V_i, \tag{3-29}$$

$$p_{ji}^{\alpha} \geq 0, \tag{3-30}$$

$$p_{ji} + y_i^{\alpha} - z_{ji} > 0 \qquad \forall i \in R^{II}, j \in V_i, \tag{3-31}$$

$$F^I(w) \leq (1+s)^3 * F^I(w^I\ ), \tag{3-32}$$

$$D_i^{min} \leq D_j(w)\ \leq D_i^{max} \qquad \forall i \in T, j \in V_i, \tag{3-33}$$

$$D_j(w)\ \leq D_i^{max} \qquad \forall i \in R^I, j \in V_i, \tag{3-34}$$

$$\sum_{i \in R^{III}} \frac{1}{|V_i|} \sum_{j \in V_i} D_j(w) \leq M_{III}^*, \tag{3-35}$$

$$w_i \leq w^{max}\ . \tag{3-36}$$

Constraints (3-28) to (3-34) were the same as that used in the third step to preserve the result of the first and second steps. The only difference is in the slip value, and the slip was increased from $(1+s)^2$ in the third step to $(1+s)^3$ in the fourth. The equation (3-35) constrained the optimal value of the mean dose to the normal structure obtained in the third step. The optimal value of the third step is denoted by M_{III}^*. Mathematically,

$$M_{III}^*\ = \sum_{i \in R^{III}} \frac{1}{|V_i|} \sum_{j \in V_i} D_j(w\,*) \tag{3-37}$$

3.4 Improving the Performance of the Application

The performance of the treatment planning optimization module was improved using the following two methods:

Storage and access of the influence matrix: We analyzed the memory access pattern of the influence matrix and optimized the memory access pattern to reduce the application run time.

Computing the exact Hessian matrix: We computed the exact Hessian matrix of the optimization model that reduced the application run time.

3.4.1 Storage and Access of the Influence Matrix

Figure 3-2 shows the memory hierarchy in the computer system. Both memory size and access time increase as we move further away from the processor. Memory and hard disk access time are around 50 and 10^{12} times larger (respectively) than that of cache access time[77]. Cache sizes typically range from kilobytes to megabytes; memory size is a few gigabytes, and the hard disk size is in terabytes. In our system, the L1 cache was 64KB, the L2 cache was 256KB and the L3 cache was 8192KB. When the processor requires an item, it searches the cache. If the item is in the cache, then it is referred to as a cache “hit” and the requested item is served from the cache. On the other hand, if the item is not in the cache then it would be a cache “miss,” and the requested item is served from the lower level memory in the hierarchy. If there are many cache misses, the performance of the system is severely impaired, as it requires significantly more time to access an item from a lower level memory in the hierarchy.

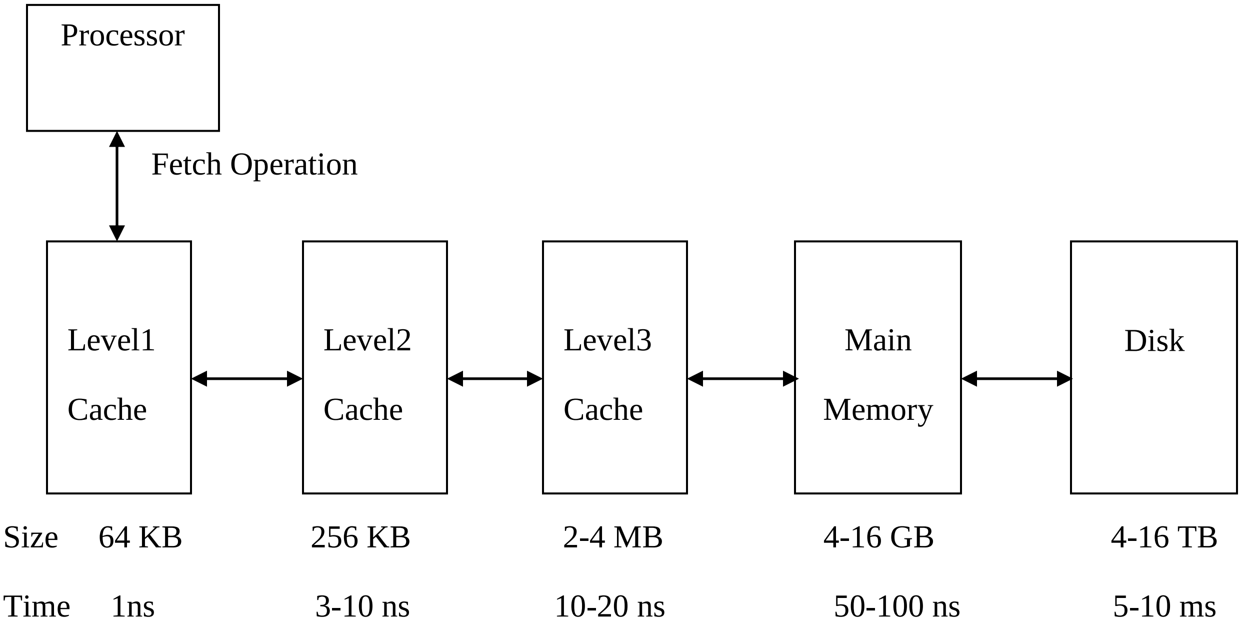


Figure 3-2 Organization of a typical computer memory pipeline. The size and access time of memory (ns, nanoseconds; ms, milliseconds) increases as we move further away from the processor.

We used the uBlas[97] library for matrix operations and stored the matrix in row-major order. Row-major ordering stores each row of the matrix end-to-end in memory. We accessed the influence matrix in two different ways: row- and column-wise to study the impact of the memory access pattern on the performance of the treatment planning application. In row-wise access, we accessed one row at a time, while in column-wise access, we accessed one column at a time.

3.5 Computation of Hessian Matrix

IPOPT[96] is a general purpose interior point optimization solver. The algorithm starts with some initial guess and computes the search direction based on gradient and second derivative (Hessian) information to find the next point in the iteration.

In more detail, given the optimization problem

$$\begin{aligned} &Minimize\ \ f(x) \qquad\qquad \forall x \epsilon R^n, \\ &Subject\ to \qquad\qquad\qquad\qquad\qquad\qquad \text{(3-38)} \\ &c(x) = 0, \\ &x \geq 0. \end{aligned}$$

The Lagrangian function is defined as

L(x,y,z) = f(x) + λ c(x) + z x,

$\lambda \in R^m, z \in R^n$ are Lagrange multipliers for the equality constraints c(x) and the bound constraints x ≥ 0, respectively.

Let (x_k,λ_k,z_k) be the current solution at some iteration "k. " The IPOPT solves the following equation to find the search direction:

$$\begin{bmatrix} W_k & A_k & -I \\ A_k^T & 0 & 0 \\ Z_k & 0 & X_k \end{bmatrix} = -\begin{bmatrix} \nabla f(x_k) + A_k\lambda_k - z_k \\ c(x_k) \\ X_k Z_k e - \mu_j e \end{bmatrix} \qquad \text{(3-39)}$$

$A_k = \nabla c(x_k)$, $W_k = \nabla^2_{xx} L(x_k, \lambda_k, z_k)$ is Hessian of the Lagrangian function L(x_k,λ_k,z_k), and X_k = diag(x_k). We solved the hierarchical optimization problem using both the quasi-Newton and Newton methods and recorded the run time of the application. In the quasi-Newton method, the IPOPT automatically approximates the Hessian matrix from the Jacobian matrix, whereas in Newton's method, the user is required to provide the exact Hessian matrix.

3.6 Experimental Results

We compared the run times when the influence matrix was accessed in two different ways: row- and column-wise. Table 3-1 shows run time for row- vs. column-wise access of the influence matrix. Compared with column-wise access, the row-wise access of the influence matrix reduced run time by an average factor of 400. Column-wise access means a full beamlet set is mixed with many regions of interest (ROIs), whereas row-wise access results in the ability to compute the full dose to a given ROI after loading only a fraction of the influence matrix. As discussed [77], modern processors use spatial locality to increase the probability of a cache hit. Spatial locality is based on the idea that a program is more likely to access memory locations close to the currently accessed location in the near future. In our implementation, the column-wise access incurred more cache misses than those incurred by the row-wise access.

Table 3-1 Impact of row-wise and column-wise access of the influence matrix on the run time of the treatment planning application. We accessed the influence matrix one row at a time in row-wise access, and one column at a time in column-wise access.

	Time (minutes)		
Plan	**Column-wise Access**	**Row-wise Access**	**Speedup**
1	583.99	1.26	462.9
2	423.10	1.08	389.98
3	121.59	0.19	639.59
4	509.06	1.66	306.40
5	538.66	1.36	393.44

We observed that the IPOPT solver was faster when the exact Hessian matrix was used instead of an approximation. The IPOPT solver needs a second order derivative of the Lagrange function, which can either be approximated or evaluated by a user. Figure 3-3 shows the run times when the exact or the approximate Hessian matrices were used in the IPOPT solver. Our observation indicates that the IPOPT solver took a large number of iterations when the Hessian matrix was approximated (Table 3-2). Consequently, the IPOPT was four times faster when the Hessian matrix was exact compared with an approximation. The Quasi-Newton method made a small improvement in the objective function after a few optimization iterations (Figure 3-4). We studied the impact of prematurely terminating the optimization formulation on the dose distribution of treatment plans. We set the maximum number of iterations to 22 for the Quasi-Newton and solved the optimization problem. e) Dose volume histogram of plan number 4

f) Dose volume histogram of plan number 5

Figure 3-5 shows that the prematurely terminating optimization sacrificed the dose distribution to structures. We therefore computed the exact Hessian matrix of the optimization problem and provided it to the IPOPT solver. The exact Hessian matrix reduced the number of iterations in the optimization thus reducing the run time of the application.

Table 3-2: Number of iterations required for the optimization using the Quasi-Newton and Newton methods. The Quasi-Newton method made a smaller improvement in the objective function after a few iterations and didn't terminate, even after running the optimization for the maximum number of iterations.

	Number of Iterations	
Plan	**Quasi-Newton**	**Newton**
1	3000	28
2	3000	38
3	3000	28
4	3000	29
5	3000	60

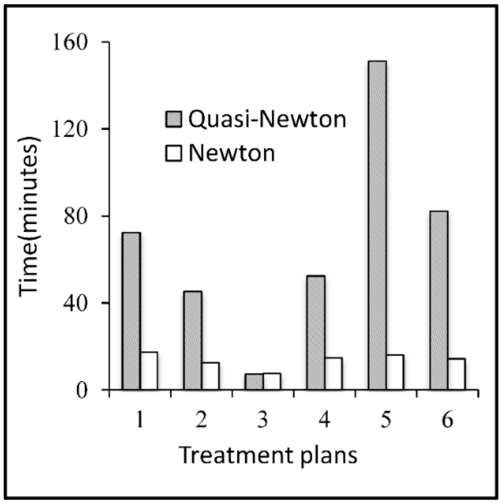


Figure 3-3 Optimization time using Newton vs. Quasi-Newton method.

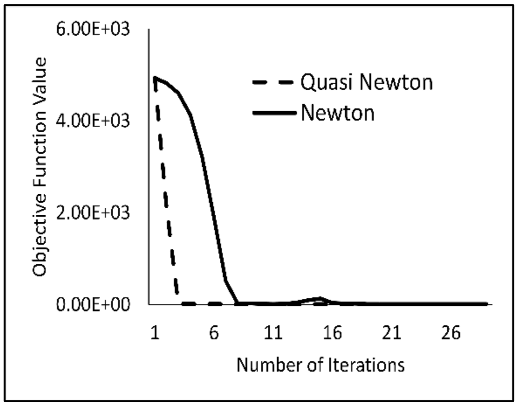
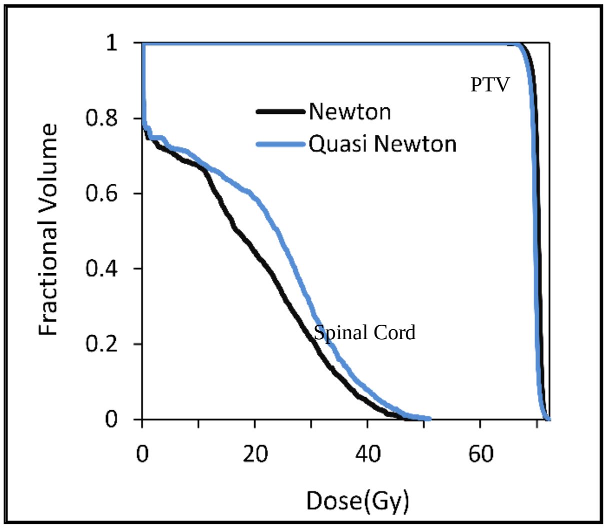
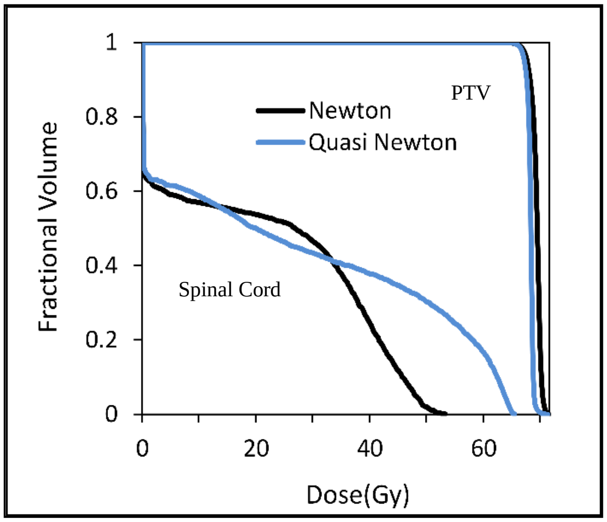


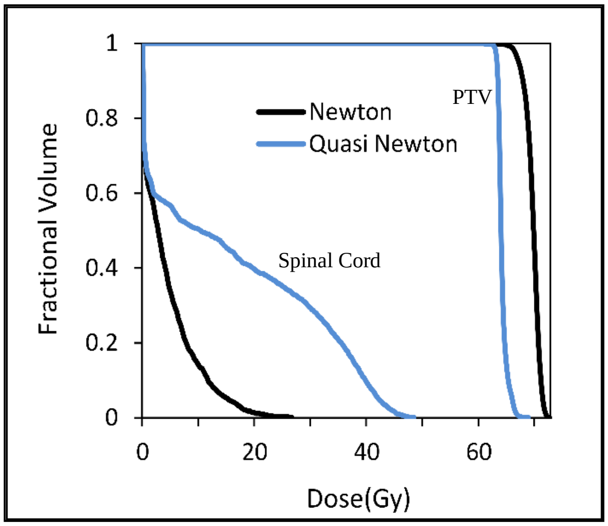
Figure 3-4 Convergence of Newton vs. Quasi-Newton method. The Quasi-Newton method made a small improvement in the objective function after a few iterations.



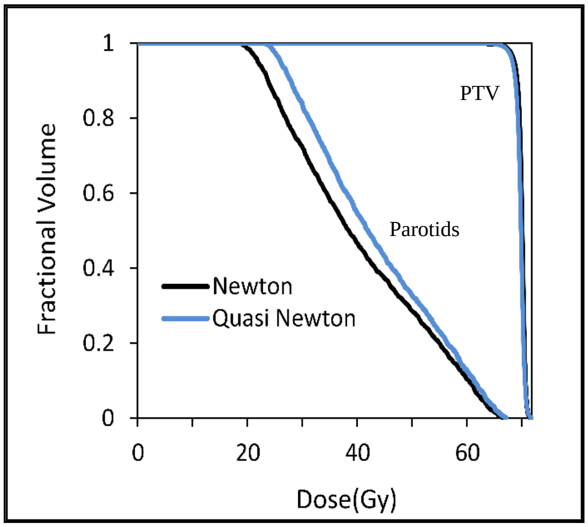
a) Dose volume histogram of plan number 1



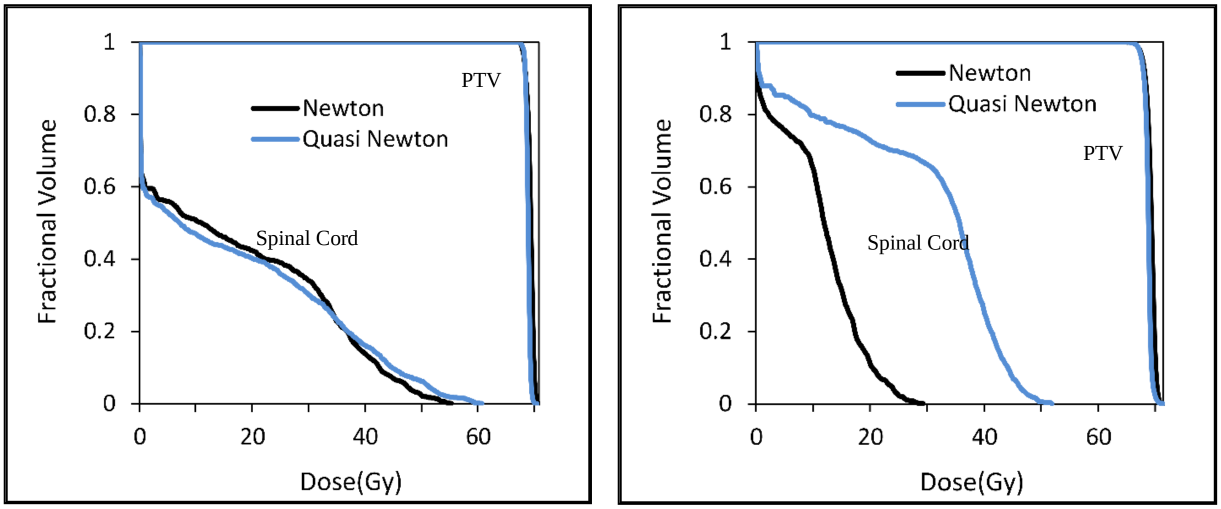
b) Dose volume histogram of plan number 2



c) Dose volume histogram of plan number 3



d) Dose volume histogram of plan number 3



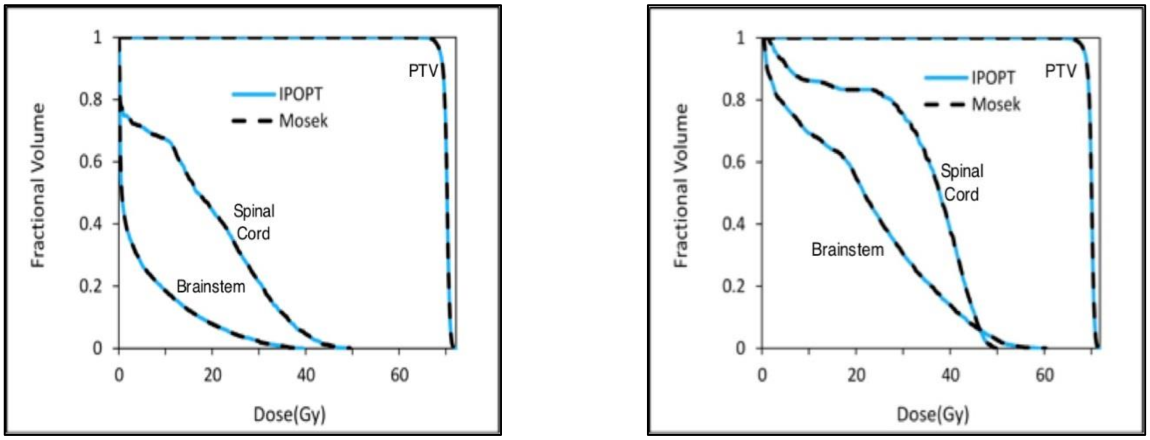
e) Dose volume histogram of plan number 4

f) Dose volume histogram of plan number 5

Figure 3-5 Dose-volume histogram demonstrating the effect of prematurely terminating the Quasi-Newton optimization. We set the maximum number of optimization iterations to 22 for the Quasi-Newton method based on Figure 3-4. It is impossible to predetermine the number of optimization iterations that produces good dose distribution for all treatment plans.

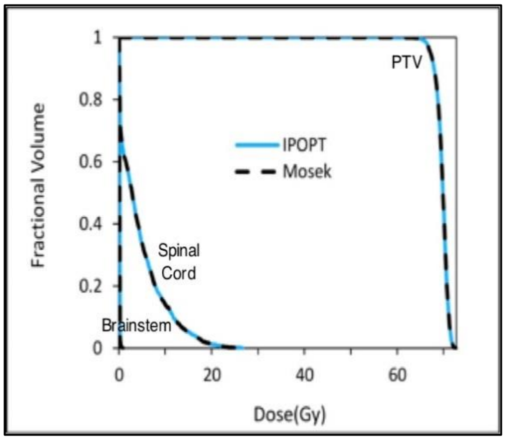
We compared the dose quality and run time of the IPOPT-based treatment planning application with the Mosek-based treatment planning application developed by Clark et. al.[88]. Figure 3-6 shows the dose-volume histograms (DVHs) and Figure 3-7 shows the dose color wash of head and neck treatment plans computed using the Mosek[98] and IPOPT solvers. Although there were a number of structures included in the optimization, we chose to show only a few of them in the histogram for the sake of clarity. The x-axis shows the dose distribution to the organs and the y-axis shows the fractional volume that received the specified dose. The dose distribution produced with the IPOPT solver (solid line) overlapped with that produced using the Mosek solver (dashed line) for each structure, indicating that both solvers produced nearly identical dose distributions.

Figure 3-8 demonstrates the step-by-step change in the dose distribution to structures using hierarchical optimization. In our formulation, we optimized dose to the target structure in the first step, the dose to the brainstem in the second and the dose to the spinal cord in the third . The target dose was allowed to deteriorate by a "slip" factor in each step. Therefore, the target DVH deteriorates in each step, compared to the previous step, as shown in Figure 3-8.The DVH of the brainstem and spinal cord were better after the second and third step of the optimization, respectively.



a) Dose volume histogram of plan number 1

b) Dose volume histogram of plan number 2



c) Dose volume histogram of plan number 3

Figure 3-6 Dose-volume histogram (DVH) produced in intensity-modulated radiation therapy treatment plans for a patient with head and neck cancer, optimized using the IPOPT-based (solid line) and Mosek-based (dashed line) treatment planning applications. We chose to show three out of seven treatment plans as the other plans have the same DVH pattern.

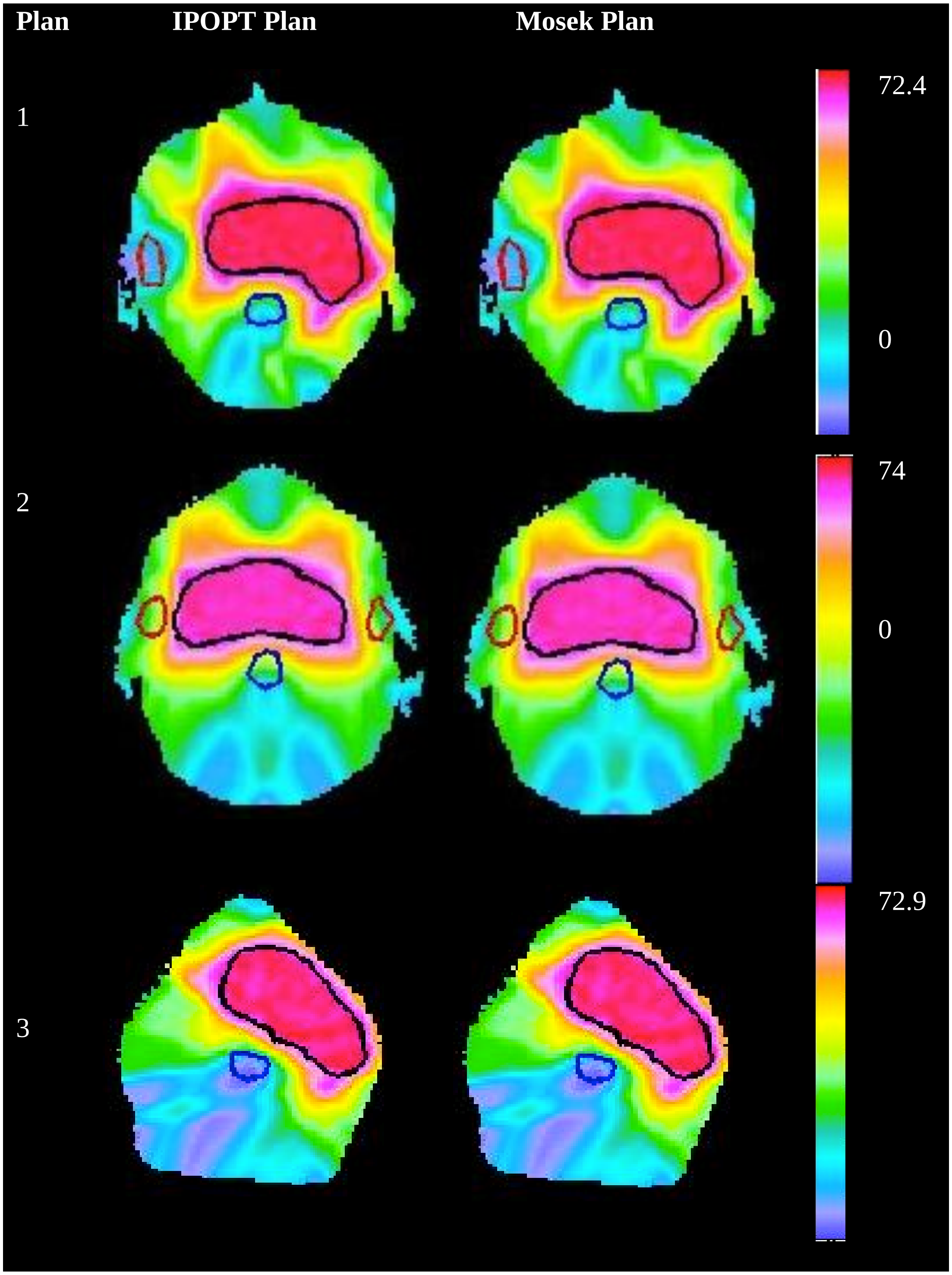


Figure 3-7 Dose color wash showing the similarity of treatment plans produced by the IPOPT and Mosek solvers. We chose to show three out of seven head and neck treatment plans as the other plans have the same dose color wash pattern.

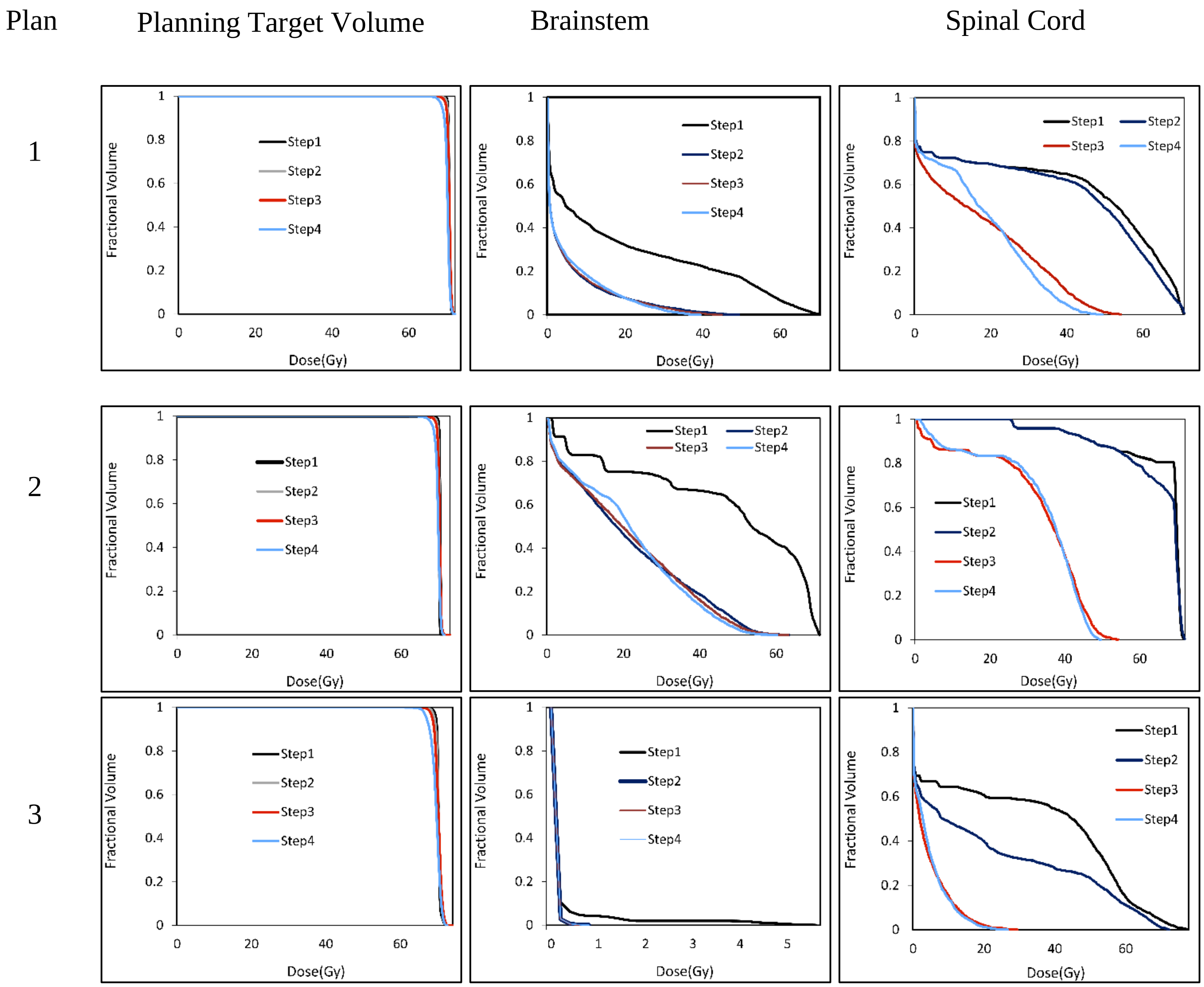


Figure 3-8 The dose-volume histogram (DVH) of the Planning Target Volume (PTV), Brainstem, and Spinal Cord produced in hierarchical optimization of three head and neck treatment plans. The optimization was solved using the IPOPT solver. We have optimized the dose to PTV in the first step, to the brainstem in the second and to the spinal cord in the third. The DVH shows that PTV coverage in the fourth step was less than that in the first. Low coverage was due to the "slip" parameter used in later optimization steps to relax the PTV dose constraints.

Table 3-3 shows the run times for IMRT optimization using either the IPOPT- or Mosek-based solvers for the seven head and neck plans. With a median increase in speed by a factor of 21, the IPOPT-based application was significantly faster than the Mosek-based application.

Table 3-3. Run time of the IPOPT- and Mosek-based prioritized optimization.

	Time (minutes)		
Plan	IPOPT-based Application	Mosek-based Application	Speedup
1	17.39	356.65	20.5
2	12.63	220.74	17.47
3	7.51	164.84	21.92
4	14.74	223.69	15.16
5	16.08	379.15	23.57
6	9.93	319.59	32.16
7	14.31	934.18	65.25

3.7 Discussion and Conclusion

We studied the memory access pattern of the influence matrix and demonstrated that optimized access could significantly reduce the IMRT optimization time.

We developed an efficient, open-source, and flexible treatment planning application, using hierarchical optimization, which automates IMRT treatment planning. Therefore, the application can find the best treatment plan in a single run without manual intervention. Moreover, the application is integrated with the CERR, a commonly used open-source tool for importing, contouring, and viewing treatment plans, which performs dose-volume and dose-surface analysis. Therefore, the application provides a complete suite for the treatment planning process. Finally, the open source model will allow new optimization models with minimal modifications. A user can provide model-specific information to include the new model and can reuse the built-in hooks to manage communication between the CERR and the IPOPT solver.

Chapter 4: A Comparative Study of Hierarchical and Eclipse Treatment Planning Systems

In chapter 3, we demonstrated that the Interior Point Optimizer (IPOPT) can be used to solve the nonlinear constrained hierarchical optimization formulation. In this chapter, we formulate the mathematical model to automate the treatment planning process and use the IPOPT solver to solve the optimization formulation.

4.1 Introduction

The dose-volume histogram (DVH) relates radiation dose to the volume of a structure[99]. The dose-volume metrics are based on the dose-volume histogram and are standard metrics used in clinics to evaluate the treatment plan. The commonly used dose-volume metrics are the dose received by x% volume of a structure (Dx), and the dose to at least x% volume of a structure (Vx). Although physicians and treatment planners widely use dose-volume metrics, these metrics can't be directly used in the optimization formulation because they are non-convex and the optimization formulation based on the dose-volume metrics has multiple local minima[100]–[103]. Therefore, the linear and quadratic surrogate functions are used to preserve the convexity of the IMRT optimization formulation.

The commercial treatment planning system uses the weighted sum of the linear and quadratic functions to optimize the dose distribution. The weighted sum optimization formulation requires trial-and-error iterations to find the best trade-off in dose distribution. Hierarchical optimization was proposed to eliminate trial-and-error iterations in the treatment planning process. The previous study demonstrated that hierarchical optimization can be used to automate the treatment planning process[72], [88]. However, these studies didn't prepare the treatment plans in a clinical setting. It is essential to prepare the treatment plans using clinical criteria to demonstrate the potential of hierarchical optimization. Therefore, we automatically prepared treatment plans using hierarchical optimization in a clinical setting. We found that the linear and quadratic models were not enough to automatically produce treatment plans. Therefore, we proposed a nonlinear optimization formulation to automate the treatment planning process. We trained the model in six prostate plans and tested it in 25 prostate plans randomly selected from prostate patients treated at Memorial Sloan Kettering Cancer Center (MSK). Finally, we compared the automatically produced plans with the plans produced by the planners at MSK using the Eclipse treatment planning system.

4.2 Methods

We randomly selected 31 prostate patients treated at MSK to prepare treatment plans using the hierarchical optimization model. The hierarchical optimization model was trained in six prostate plans and tested in 25 plans.

4.2.1 Application Work Flow

Figure 4-1 shows the flow diagram of the automated treatment planning process based on hierarchical optimization. Treatment plans were exported from the Eclipse treatment planning system in the Digital Imaging and Communications in Medicine (DICOM) format and imported into the Computational Environment for Radiotherapy Research (CERR). The beam geometry and the contours of structures are obtained from the imported plan. The new beams were created in CERR interface using the beam geometry obtained from the imported plan, and the dose was calculated using the the voxel-based Monte Carlo (VMC++) dose calculation engine. The hierarchical optimization was run based on the clinical criteria specified by the physicist of Memorial Sloan Kettering Cancer Center. The Monte Carlo dose calculation engine doesn’t incorporate the machine parameters and, therefore, the dose distribution of the hierarchical optimization was an ideal dose distribution. The ideal fluence map was exported from the CERR, imported into the Eclipse application, and the dose was calculated using the Eclipse dose calculation engine to incorporate the machine parameters. Dose distribution was deteriorated by incorporating the machine parameters. Therefore, the plan was re-optimized to prepare the final deliverable plan.

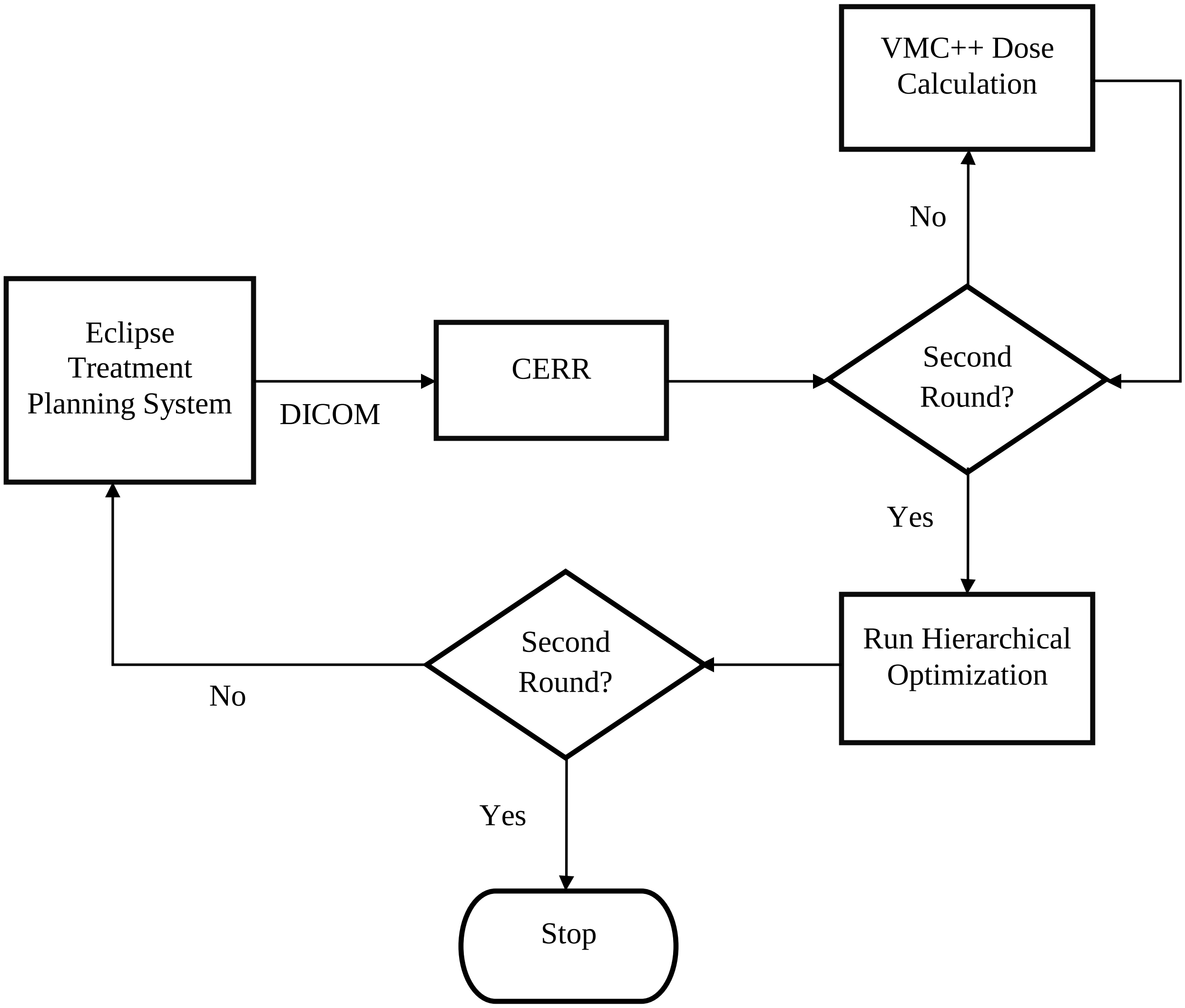


Figure 4-1 The flowchart of the automated treatment planning system. We exported treatment plans from the Eclipse treatment planning system and imported them into CERR. The first round of optimization used the voxel-based Monte Carlo (VMC++) dose calculation engine. Therefore, dose distribution was ideal. The second round of optimization incorporated the machine parameters and produced the deliverable dose distribution.

4.2.2 Dose calculation

We used the voxel-based Monte Carlo (VMC++) application to calculate the dose to the structures. VMC++[104] is an open source Monte Carlo[105] dose calculation engine and the CERR provides the interface to set the parameters for the VMC++ application. The VMC++

takes substantial time to calculate the dose and, therefore, the dose computation time was reduced by calculating the dose in parallel. The procedure Parallel-Dose-Computation takes the parameter IM, which contains beam information and the parameters for the VMC++ algorithm. The new structure "vmcparam" and "beams" were created to store beam information of each beam separately. The computation of each beam was carried out in parallel, but the computations of the beams’ beamlets were carried out sequentially.

Algorithm 4-1 Pseudocode of dose computation in parallel

```
Parallel-Dose-Computation(IM)
    vmcparam = Create a structure of size equals to the number of beams(beaminfo.length)
    beams = IM.beam
    for i=1 to beaminfo.length
        vmcparam(i).number_of_particles = IM.vmcparam(i).number_of_particles
        Vmcparam(i).number_of_batches = IM.vmcparam(i).number_of_batches
    parfor i=1 to beaminfo.length
        for j=1 to beams(i).numbeamlet
            dosefile = dosefile_beam_i_beamlet_j
            Vmc++(beams(i),vmcparam(i),dosefile)
            dose = Readdosefile(dosefile)
            beams(i).beamlets(j) = dose
```

Dose calculation requires some parameters, namely voxel size, the parameters for the VMC++ algorithm, and beam geometry information. The voxel size was approximately 0.2 by 0.2 by 0.2 cm in x, y and z directions and there were 256 slices in x and ydirections. Table 4-1 shows the parameters for the VMC++ algorithm. The parameter value can be set from graphical user interface (GUI) of the CERR application. The default value was used for the parameters that are not shown in Table 4-1. Finally, Table 4-2 shows the beam parameters. We used photon beam size 0.2 by 0.5 cm to calculate the dose. The beam parameters were the same as that used by the planner at MSK to prepare treatment plans.

Table 4-1 The parameter for the VMC++ algorithm. The parameters can be set using the CERR graphical user interface (GUI) interface.

S.No.	Parameters Name	Values
1.	Number of particles	50,000
2.	Number of batches	10
3.	Score dose to water	Yes
4.	Scatter method	Exponential
5.	Skip	1
6.	Dimension	60
7.	Skip	1
8.	Dose term	Gauss primary and scatter
9.	Cut-off distance	4
10.	Random step	30

Table 4-2 The parameters for the beams used in dose calculation. The beam parameters were the same as those used by the planners at Memorial Sloan Kettering Cancer Center to prepare treatment plans.

S.No.	Parameters Name	Values
1.	Beam Modality	Photons
2.	Beam energy	15 MV
3.	Beam delta x	0.25
4.	Beam delta y	0.5

4.2.3 Clinical criteria for treatment plans

The clinical criteria used in hierarchical optimization were the same as those used by the planner at MSK to prepare treatment plans. The criteria specify the dose requirement to target and normal structures (Table 4-3). The target prescription dose was 72Gy and the dose to 95% of the target structure must equal the prescription dose. 53% volume of the rectum and bladder must receive less than 47Gy of the dose. Similarly, the maximum dose to the rectum and bladder must be less than 103% and 105% of the prescription dose.

Table 4-3 Clinical criteria used in hierarchical optimization. The criteria specify the dose requirement of the prostate plan. The target dose is given at the top rows of the table. The prescription dose was 72Gy and the rectum and bladder must satisfy dose-volume constraints. The criteria have been used to treat patients at Memorial Sloan Kettering Cancer Center.

Prostate 7200cGy	Target Criteria			
	PTV $D_{95\%}$ ~ 100%	Cases with large amount of bladder-PTV overlap may result in somewhat inferior coverage		
	PTV $V_{95\%} \geq 99\%$			
	PTV $D_{max} \leq 110\%$			
	PTV D_{min} near 100%			
	Normal Tissue Criteria			
	Structures	Total Dose or Volume ≤	To:	Comments
	Rectal Wall	53% of total volume	V_{47Gy}	
		103.5%	Max Point Dose	Avoid hot spots in the rectal wall
	Rectoe**	103.5%	Max Point Dose	**one needs to look at this specifically for nodal cases
	Bladder Wall	53%(63%) of total volume	V_{47Gy}	*Planner should first try to meet the V47Gy ≤ 53%. If this is not possible a value of 63% may be used
		105%(108%)	Max point dose	Avoid hot spots in the bladder wall

	Femoral Heads	68Gy	Max Point Dose	

4.2.4 Optimization Formulation

Hierarchical optimization assigns priority to optimization goals and solves the goals based on the assigned priority. Table 4-4 shows the goals of hierarchical optimization and the priority of each goal. The target structure is in priority level 1; the rectum is in priority level 2, and the bladder and femur are in priority level 3. Moreover, the rectum and bladder hard constraints are in priority level 1, achieving good dose fall-off is in priority level 4, producing a smooth beam profile is in priority level 5, and incorporating the deliverability issue is in priority level 6. Hierarchical optimization has six steps corresponding to the six priority levels. Higher priority goals are optimized before lower priority goals. After optimizing higher priority goals, they are converted into constraints and lower priority goals are optimized in the presence of higher priority goal constraints.

Table 4-4 The optimization goals and their priorities. Goals are optimized based on the assigned priority. Priority level one goals are optimized in the first step, priority level two goals are optimized in the second step and so on.

S.No.	Goals	Priority Level
1.	PTV $D_{95\%}$ ~ 100%	1
2.	Max point dose in rectum $\leq$ 103.5%	1
3.	Max point dose in bladder $\leq$ 108%	1
4.	V47Gy of rectum $\leq$ 53% of total volume	2
5.	V47Gy of bladder $\leq$ 53% of total volume	3
6.	Max point dose in Femoral Heads	3
7.	Dose should fall-off outside the	4

	target	
8.	Beamlet should be smooth	5
9.	Incorporate the deliverability issue	6

We have formulated the mathematical model needed to automate the treatment planning process using the hierarchical optimization framework. The optimization has six steps as discussed below.

Step 1

The first step optimizes the three highest priority goals as shown in Table 4-4. The first goal was to achieve the dose to 95% of the target structure as close as the 100% of the prescription dose. The objective function (4-1) minimizes the difference between the actual and prescribed dose. The variable "t," in the objective function, penalized the under-dose region in the target structure. The two other goals specify the maximum dose to the bladder and rectum and maximum dose constraints (4-3) were used to achieve the goal.

Let T be the set of target structures, R^I is the set of the normal structures, D_i^{pre} be the prescription dose to target structure i, V_i be the set of voxels in structure i, D_j is the dose to voxel j of a structure, then the first step formulation is given by:

$$\textbf{Minimize} \qquad \sum_{i \in T} \frac{1}{|V_i|} \sum_{j \in V_i} (D_j - D_i^{pre})^2 + t_i^2, \tag{4-1}$$

Subject to:

$$D_j - D^p \leq t_i \qquad \forall j \in V_i, i \in T, \tag{4-2}$$

$$D_j \leq D_i^{max} \qquad i \in R^I, \forall j \in V_i, \tag{4-3}$$

$$D_j(w) \ \leq D_i^{max} \qquad \forall i \in T, j \in V_i. \tag{4-4}$$

Let A be the influence matrix, w be the beamlets weight vector, and n be the number of beamlets in the optimization, then the dose to voxel j of the target structure is given by:

$$D_j = \sum_{j \in V_i} \sum_{k=1}^{n} A_{jk} * w_k \qquad \forall i \in T \tag{4-5}$$

Step 2

The second step optimizes the rectal dose. The dose-volume constraint of the rectum requires that 53% rectal volume must receive less than the 47Gy of the dose. The dose-volume constraints cannot be incorporated directly into the optimization without losing the convexity of the problem. Therefore, we used the linear combination of the generalized equivalent uniform dose (gEUD) [106]–[110] and the mean rectal dose to enforce the dose-volume constraint to the rectum. The gEUD has a parameter called "a" and the parameter can be tuned to emphasize the high dose region of the rectum. Therefore, the rectum high dose region was controlled by the gEUD with a large "a" value and the low dose region was controlled by the mean dose. The objective function of the second step is given by:

$$\gamma * gEUD(a)\ of\ the\ rectum + \rho * mean\ dose\ to\ rectum \tag{4-6}$$

$\gamma, a, and\ \rho$ are the model’s parameters, which we empirically determined the value of by training the model in the six training prostate plans (Table 4-5).

Table 4-5 The value of the model parameters for step 2. The parameter value was determined by training the model in the six prostate cases.

Parameter	Value
a	20
γ	0.5
ρ	0.5

Let R^{II} be the set of structure in the second step, V_i be the set of voxels of the structure i, D_j be the dose to the voxel j, then the second step formulation is given by:

$$\boldsymbol{Minimize}\quad 0.5 * \sum_{i\in R^{II}} \left({}^{1}\!/_{|V_i|} * \sum_{j\in V_i} D_j^a \right)^{1/a} + 0.5 * \sum_{i\in R^{II}} {}^{1}\!/_{|V_i|} * \sum_{j\in V_i} D_j, \tag{4-7}$$

$\boldsymbol{Subject\ to}$:

$$\sum_{j\in V_i} \min(0, D_j - D_i^{pre})^2 \le \sum_{j\in V_i} \min(0, D_j^I - D_i^{pre})^2 \qquad \forall i \in T\,, j \in V_i, \tag{4-8}$$

$$\frac{1}{|V_i|}\sum_{j\in V_i} (D_j - D_i^{pre})^2 \le s * \frac{1}{|V_i|}\sum_{j\in V_i} (D_j^I - D_i^{pre})^2 \qquad \forall i \in T, j \in V_i, \tag{4-9}$$

$$D_j \le D_i^{max} \qquad i \in T, \forall j \in V_i, \tag{4-10}$$

$$D_j \le D_i^{max} \qquad i \in R^I, \forall j \in V_i. \tag{4-11}$$

The dose to jth voxel of the rectum D_j is given by:

$$D_j = \sum_{j\in V_i}\sum_{k=1}^{n} A_{jk} * w_k \qquad \forall i \in R^{II} \tag{4-12}$$

Similarly, Let w^I be the beamlet weight found in the first step, then the first step dose to the voxel j of the target structure (D_j^I) is given by:

$$D_j^I = \sum_{j \in V_i} \sum_{k=1}^{n} A_{jk} * w_k^I \qquad \forall i \in R^{II} \tag{4-13}$$

The result of the first step should be preserved while optimizing for the second step goals. The target dose of the first step was preserved by using constraints (4-9) and (4-10). However, the rectal dose cannot be reduced if the hard constraint is used to preserve the target dose. Therefore, we applied a slip in constraint (4-9) to relax the target dose. Compared to the plan produced by the planner, the hierarchical system plan had lower D95 to the target structure when a large slip was used and a higher dose in the rectum when the smaller slip was used. We noticed that the planner tends to sacrifice the homogeneity of the target dose to improve on the rectal dose. Therefore, we introduced constraint (4-8) to preserve the sum of the dose to the target voxels that was less than the prescription dose. Constraint (4-8) ensured that no substantial cold spots were introduced to the target structure in the second step. This enabled us to use a large slip to reduce the rectal dose while not degrading the D95 to the target structure.

Step 3

The third step optimizes the bladder dose under the constraints to preserve the previous step's result and to limit the maximum dose to the femur. The clinical criteria state that the maximum dose to the femur must be less than 68Gy and the dose to 53% volume of the bladder must be less than 47Gy. The objective function of the third step minimizes the linear combination of the gEUD and the mean dose to the bladder. Mathematically,

$$\gamma * gEUD(a)\ of\ the\ bladder + \rho * mean\ dose\ to\ the\ bladder$$

γ, a and ρ are the model's parameters, which were empirically determined by training the model in the six prostate plans. The table below shows the value of the model parameters:

Table 4-6 The value of the model parameters of step 3. The parameter value was determined by training the model in the six prostate cases.

Parameter	Value
a	20
γ	0.9
ρ	0.1

Let R_b and R_f be the bladder and femur structures that are being optimized in the third step, then the optimization formulation is given by:

$$\boldsymbol{Minimize}\ \ 0.9 * \textstyle\sum_{i\in R_b} \left({}^{1}\!/_{|V_i|} * \sum_{j\in V_i} D_j^a \right)^{1/a} + 0.1 * {}^{1}\!/_{|V_i|} \sum_{i\in R_b} \sum_{j\in V_i} D_j, \qquad \text{(4-14)}$$

$$\boldsymbol{Subject\ to}: \qquad D_j \leq D_i^{max} \qquad \forall i \in R_f, j \in V_i, \qquad \text{(4-15)}$$

$$0.5 * \sum_{i\in R^{II}} \left({}^{1}\!/_{|V_i|} * \sum_{j\in V_i} D_j^a \right)^{1/a} + 0.5 * \sum_{i\in R^{II}} {}^{1}\!/_{|V_i|} * \sum_{j\in V_i} D_j \qquad \text{(4-16)}$$

$$\leq\ 0.5 * \sum_{i\in R^{II}} \left({}^{1}\!/_{|V_i|} * \sum_{j\in V_i} {D_j^{II}}^a \right)^{1/a} + 0.5 * \sum_{i\in R^{II}} {}^{1}\!/_{|V_i|} * \sum_{j\in V_i} D_j^{II},$$

$$\sum_{j\in V_i} \min(0, D_j - D_i^{pre})^2 \leq \sum_{j\in V_i} \min(0, D_j^I - D_i^{pre})^2 \qquad \forall i \in T\,, j \in V_i, \qquad \text{(4-17)}$$

$$\frac{1}{|V_i|} \sum_{j\in V_i} (D_j - D_i^{pre})^2 \leq\ s^2 * \frac{1}{|V_i|} \sum_{j\in V_i} (D_j^I - D_i^{pre})^2 \qquad \forall i \in T, j \in V_i, \qquad \text{(4-18)}$$

$$D_j \leq D_i^{max} \qquad i \in T, \forall j \in V_i, \qquad \text{(4-19)}$$

$$D_j \leq D_i^{max} \qquad i \in R^I, \forall j \in V_i. \qquad \text{(4-20)}$$

Note that constraint (4-16) preserves the rectal dose, constraints (4-17), (4-18) and (4-19) preserve the target dose and constraint (4-20) limits the maximum dose in the bladder and rectum.

Step 4

The goal of step 4 was to squash the hot spot outside the target structure and produce good dose falloff. We created two new structures: *rind*, which encompass the region 2cm away from the target structure and *rind'*, which encompasses the region 2cm within the target structure. The fourth step’ s objective function minimizes the general equivalent uniform dose (gEUD) to *rind* and gEUD and the mean-square dose to the *rind’* structure. The *rind* and *rind’* include all the voxels in the patient scan. Therefore, we down-sampled the *rind* and *rind’* structures by a sampling rate of 2 to make the optimization computationally feasible.

The hot spot outside the target structure could not be squashed by only minimizing the objective function of the rind structure. The constraints to preserve the bladder dose forbid the squashing of all the hot spots outside the target structure. The bladder mean dose must be allowed to increase in order to squash the hot spot outside the target structure. However, we didn’t know the minimum increase in the bladder dose needed to squash all the hot spots. Therefore, we introduced a variable epsilon "ϵ" in the objective and constraint for the bladder term. The optimizer determines the minimum increase in the bladder mean dose needed to squash all the hot spots outside the target structure.

The gEUD and the mean-square dose term should be scaled to the same magnitude. We calculated the gEUD and mean-square dose using the third step result and divided the gEUD by the mean-square dose to get the scaling factor $\emptyset$.

Let D_j^{III} be the dose to the voxel j obtained in the third step, then the parameters’ value are calculated as follows:

$$\emptyset = \frac{\sum_{i\in rind} {}^{1}/_{|V_i|} * \sum_{j\in V_i} {D_j^{III}}^{2}}{\sum_{i\in rind\cup rind'} \left({}^{1}/_{|V_i|} * \sum_{j\in V_i} {D_j^{III}}^{a}\right)^{1/a}} \qquad (4\text{-}21)$$

$$\boldsymbol{Minimize}\text{: } \epsilon + \emptyset * \sum_{i \in rind \cup rind'} \left({}^{1}\!/\!_{|V_i|} * \sum_{j \in V_i} D_j^a \right)^{{}^{1}\!/\!_{a}} + \sum_{i \in rind'} {}^{1}\!/\!_{|V_i|} * \sum_{j \in V_i} D_j^2, \tag{4-22}$$

$\boldsymbol{Subject\ to}$:

$$\sum_{i \in R_b} {}^{1}\!/\!_{|V_i|} \sum_{j \in V_i} D_j \leq \sum_{i \in R_b} {}^{1}\!/\!_{|V_i|} \sum_{j \in V_i} D_j^{III} + \epsilon, \tag{4-23}$$

$$D_j \leq D_i^{III^{max}} \qquad \forall i \in R_f, \forall j \in V_i, \tag{4-24}$$

$$D_j^{max} \leq D_i^{III^{max}} \qquad \forall i \in T, j \in V_k, k \in rind', \tag{4-25}$$

$$\begin{aligned} 0.5 * \sum_{i \in R^{II}} \left({}^{1}\!/\!_{|V_i|} * \sum_{j \in V_i} D_j^a \right)^{{}^{1}\!/\!_{a}} + 0.5 * \sum_{i \in R^{II}} {}^{1}\!/\!_{|V_i|} * \sum_{j \in V_i} D_j \\ \leq 0.5 * \sum_{i \in R^{II}} \left({}^{1}\!/\!_{|V_i|} * \sum_{j \in V_i} D_j^{III^a} \right)^{{}^{1}\!/\!_{a}} + 0.5 * \sum_{i \in R^{II}} {}^{1}\!/\!_{|V_i|} * \sum_{j \in V_i} D_j^{III}, \end{aligned} \tag{4-26}$$

$$\sum_{j \in V_i} \min(0, D_j - D_i^{pre})^2 \leq \sum_{j \in V_i} \min(0, D_j^I - D_i^{pre})^2 \qquad \forall i \in T, j \in V_i, \tag{4-27}$$

$$\frac{1}{|V_i|} \sum_{j \in V_i} (D_j - D_i^{pre})^2 \leq s^3 * \frac{1}{|V_i|} \sum_{j \in V_i} (D_j - D_i^{pre})^2 \qquad \forall i \in T, j \in V_i, \tag{4-28}$$

$$D_j \leq D_i^{max} \qquad i \in T, \forall j \in V_i, \tag{4-29}$$

$$D_j \leq D_i^{max} \qquad i \in R^I, \forall j \in V_i. \tag{4-30}$$

Constraints (4-31) to (4-32) are the same as those in step 3 formulation, and they are used to preserve the result of steps 1 and 2. Note that the slip value was increased from s^2 to s^3 to relax the target dose constraint.

Step 5

Step 5 smooths the beam profile to produce the deliverable dose distribution. We applied a Laplacian smoothing function for this purpose.

Let w_i be the weight of the i^{th} beamlet, $w_i(x_k,y_k)$ be the beamlet at the (x_k,y_k) co-ordinate then the Laplacian smoothing is given by:

$$L_i(w) = -4w_i(x_k, y_k) + w_{i+1}(x_{k+1}, y_k) + w_{i+2}(x_k, y_{k+1}) + w_{i+3}(x_{k-1}, y_k) + w_{i+4}(x_k, y_{k-1}) \tag{4-33}$$

If there are n beamlets, then the objective function is given by:

$$\boldsymbol{Minimize} \qquad \textstyle\sum_{i=1}^{n} L_i(w)^2, \tag{4-34}$$

$\boldsymbol{Subject\ to}$:

$$\sum_{i\in rind'} {}^{1}\!/_{|V_i|} * \sum_{j\in V_i} D_j^2 \le \sum_{i\in rind'} {}^{1}\!/_{|V_i|} * \sum_{j\in V_i} {D_j^{IV}}^2, \tag{4-35}$$

$$\sum_{i\in rind\cup rind'} \left({}^{1}\!/_{|V_i|} * \sum_{j\in V_i} D_j^a \right)^{1/a} \le \sum_{i\in rind\cup rind'} \left({}^{1}\!/_{|V_i|} * \sum_{j\in V_i} {D_j^{IV}}^a \right)^{1/a}, \tag{4-36}$$

$$0.5 * \sum_{i\in R^{II}} \left({}^{1}\!/_{|V_i|} * \sum_{j\in V_i} D_j^a \right)^{1/a} + 0.5 * \sum_{i\in R^{II}} {}^{1}\!/_{|V_i|} * \sum_{j\in V_i} D_j \le 0.5 * \sum_{i\in R^{II}} \left({}^{1}\!/_{|V_i|} * \sum_{j\in V_i} D_j^{IV} \right)^{1/a} + 0.5 * \sum_{i\in R^{II}} {}^{1}\!/_{|V_i|} * \sum_{j\in V_i} D_j^{IV}, \tag{4-37}$$

$$\sum_{j\in V_i} \min(0, D_j - D_i^{pre})^2 \le \sum_{j\in V_i} \min(0, D_j^I - D_i^{pre})^2 \qquad \forall i \in T, j \in V_i, \tag{4-38}$$

$$\frac{1}{|V_i|}\sum_{j\in V_i}(D_j - D_i^{pre})^2 \leq \; s^3 * \frac{1}{|V_i|}\sum_{j\in V_i}(D_j^I - D_i^{pre})^2 \qquad \forall i \in T, j \in V_i, \tag{4-39}$$

$$D_j \leq D_i^{max} \qquad i \in T, \forall j \in V_i, \tag{4-40}$$

$$D_j \leq D_i^{max} \qquad i \in R^I, \forall j \in V_i, \tag{4-41}$$

$$\sum_{i\in R_b} {}^{1}\!/_{|V_i|} \sum_{j\in V_i} D_j \leq \sum_{i\in R_b} {}^{1}\!/_{|V_i|} \sum_{j\in V_i} D_j^{IV}, \tag{4-42}$$

$$D_j \leq D_i^{IV^{max}} \qquad \forall i \in R_f, \forall j \in V_i, \tag{4-43}$$

$$D_j \leq D_i^{IV^{max}} \qquad \forall i \in T, j \in V_k, k \in rind'. \tag{4-44}$$

Let V_i be the set of voxels of a structure i, then the dose to the voxel j (D_j^{IV}) is given by:

$$D_j^{IV} = \sum_{j\in V_i}\sum_{k=1}^{n} A_{jk} * w_k^{IV} \;, \tag{4-45}$$

w_k^{IV} is the optimal beamlet weight found in the fourth step.

Similarly, $D_i^{IV^{max}}$ is the maximum dose of a structure and is given by:

$$D_i^{IV^{max}} = \max\{D_j^{IV} : j \in V_i\}$$

Step 6

The CERR dose calculation algorithm does not take the deliverable parameters into consideration. The beam profile produced from step 5 was ideal, but the linear accelerator can't deliver it. Therefore, the beam profile was exported into text format from the CERR application and imported into the Eclipse treatment planning system. Finally, the dose was computed using the Eclipse dose calculation algorithm to produce the deliverable plans.

The dose distribution of treatment plans deteriorated when the dose was computed using the Eclipse dose calculation algorithm. The new hot and cold spots were introduced in the target structure and the dose to 95% of the target structure was substantially degraded. We introduced a new step in the optimization to correct the deteriorated dose distribution. The dose was corrected by assigning a new prescription dose to each voxel of the target structure. The prescription dose was decreased if the Eclipse system produced a higher dose than that of the hierarchical and increased otherwise.

Let $D_j^{Eclipse}$ be the dose to the voxel j of the target structure after running through the Eclipse, and $D_j^{hierarchical}$ be the dose to the voxel j of the target structure before running through the Eclipse, then the optimization formulation is given by:

$$\boldsymbol{Minimize} \sum_{i \in T} \frac{1}{|V_i|} \sum_{j \in V_i} (D_j - D_i^{pre} + D_j^{Eclipse} - D_j^{hierarchical})^2, \tag{4-46}$$

$\boldsymbol{Subject\ to}$:

$$\sum_{i\in rind'} {}^{1}/_{|V_i|} * \sum_{j\in V_i} D_j^2 \le \sum_{i\in rind'} {}^{1}/_{|V_i|} * \sum_{j\in V_i} D_j^{V^2}, \tag{4-47}$$

$$\sum_{i\in rind\cup rind'} \left({}^{1}/_{|V_i|} * \sum_{j\in V_i} D_j^a \right)^{1/a} \le \sum_{i\in rind\cup rind'} \left({}^{1}/_{|V_i|} * \sum_{j\in V_i} D_j^{V^a} \right)^{1/a}, \tag{4-48}$$

$$\sum_{i=1}^{n} L_i(w)^2 \le \sum_{i=1}^{n} L_i(w^V)^2, \tag{4-49}$$

$$0.8 * \sum_{i\in R^{II}} \left({}^{1}/_{|V_i|} * \sum_{j\in V_i} D_j^a \right)^{1/a} + 0.2 * \sum_{i\in R^{II}} {}^{1}/_{|V_i|} * \sum_{j\in V_i} D_j \tag{4-50}$$

$$\le 0.8 * \sum_{i\in R^{II}} \left({}^{1}/_{|V_i|} * \sum_{j\in V_i} D_j^{V^a} \right)^{1/a} + 0.2 * \sum_{i\in R^{II}} {}^{1}/_{|V_i|} * \sum_{j\in V_i} D_j^V,$$

$$\sum_{i\in R_b} {}^{1}/_{|V_i|} \sum_{j\in V_i} D_j \le \sum_{i\in R_b} {}^{1}/_{|V_i|} \sum_{j\in V_i} D_j^V. \tag{4-51}$$

Let w^V be the optimal beamlet weight found in the fifth step, then the dose to voxel j obtained from the fifth step is given by:

$$D_j^V = \sum_{j\in V_i} \sum_{k=1}^{n} A_{jk} * w_k^V \tag{4-52}$$

4.3 Results

The hierarchical system produced the same or better quality treatment plans than those produced by the planner using the Eclipse treatment planning system. We randomly selected 31 prostate plans treated at MSK, trained the hierarchical model in six prostate plans, and tested the model in 25 plans. The dose to 95% region (D95) of the target structure from the hierarchical system was scaled to match the D95 of the target structure obtained from the Eclipse treatment planning system (Figure 4-2 to Figure 4-7). The mean dose to the rectum and bladder were substantially lower for the treatment plans prepared using the hierarchical system than that prepared using the Eclipse treatment planning system. The hierarchical system produced a 10% lower average mean dose to the rectum and bladder than that of the Eclipse treatment planning system in 31 prostate plans (Table 4-7,

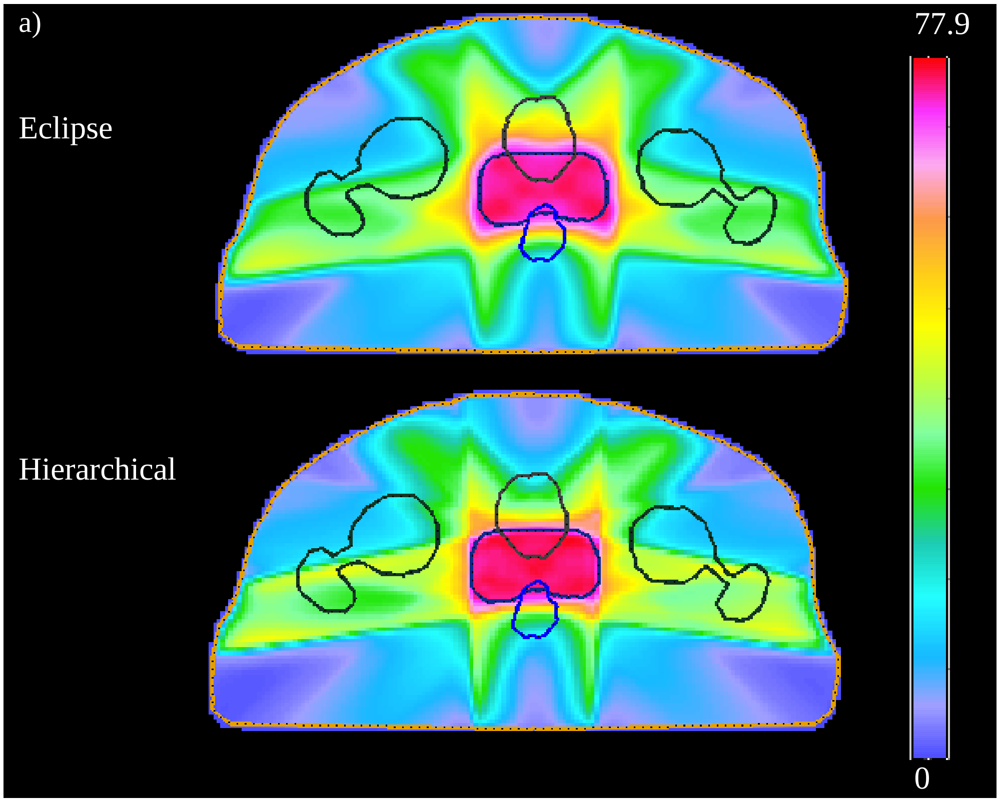
Table 4-11, Table 4-13). Similarly, the maximum dose to the rectum and bladder was almost same from both systems across all treatment plans (

Table 4-12 and

Table 4-14). The result suggests that the hierarchical system could automatically produce dose distribution comparable with the treatment plans produced by the planner using the Eclipse treatment planning system.

4.4 Training the Model

We trained the model in six plans to determine the objective functions, constraints, and model parameters needed to produce treatment plans. The dose distribution from the hierarchical system was comparable or better than that of the Eclipse treatment planning system. The hierarchical system substantially reduced the dose in the rectum and bladder (Figure 4-2 to Figure 4-7). The rectal volume that receives at least 47Gy (V47) from the hierarchical system was almost 30% lower than that of the Eclipse system in three out of six training plans (Table 4-7). The hierarchical system increased V47 by 0.05% in three plans, but the increase was negligibly small. Similarly, the bladder V47 from the hierarchical system was almost 11% lower on average than that from the Eclipse system in five out of six plans (Table 4-7). The target D95 was the same from both systems because of scaling. The dose color-wash shows that no hot spot was produced outside the target structure and that the dose decreased as the distance from the target structure increased. Therefore, the hierarchical model seems to produce treatment plans automatically without manual intervention.



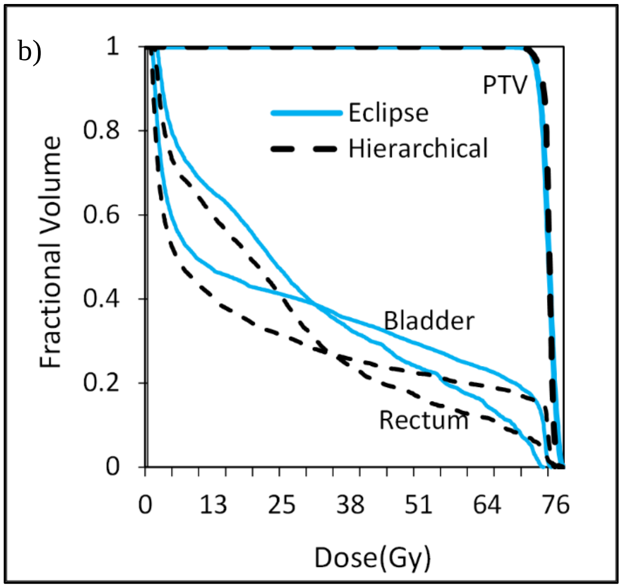
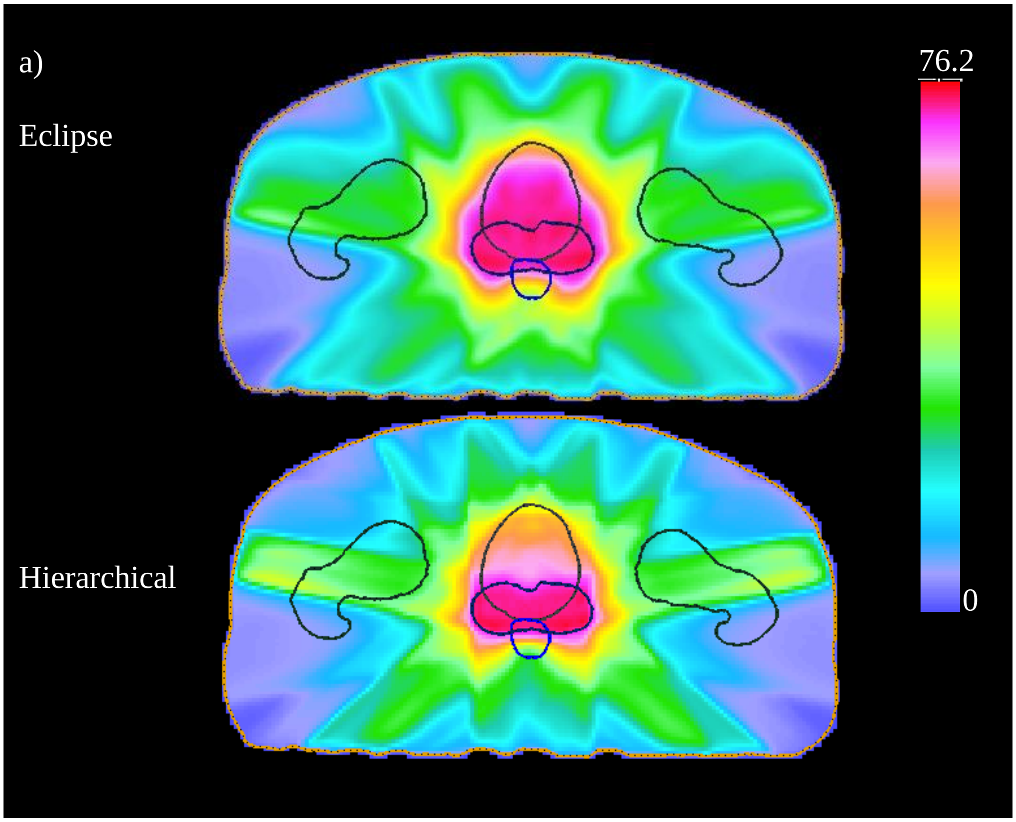


Figure 4-2 Dose distribution of prostate treatment plan 2. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



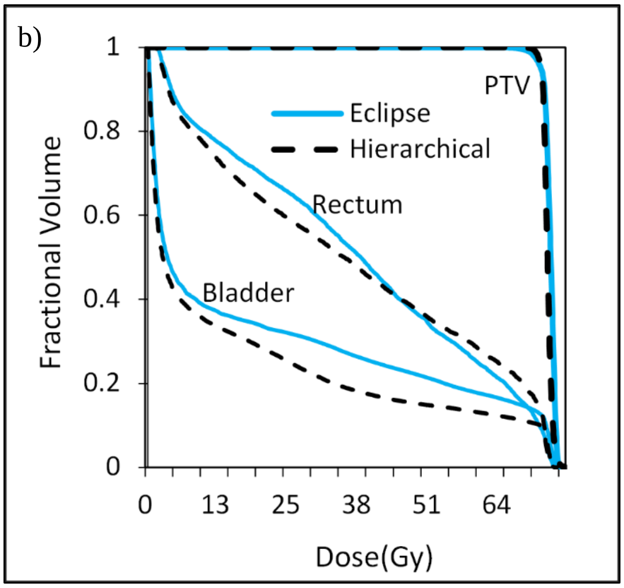
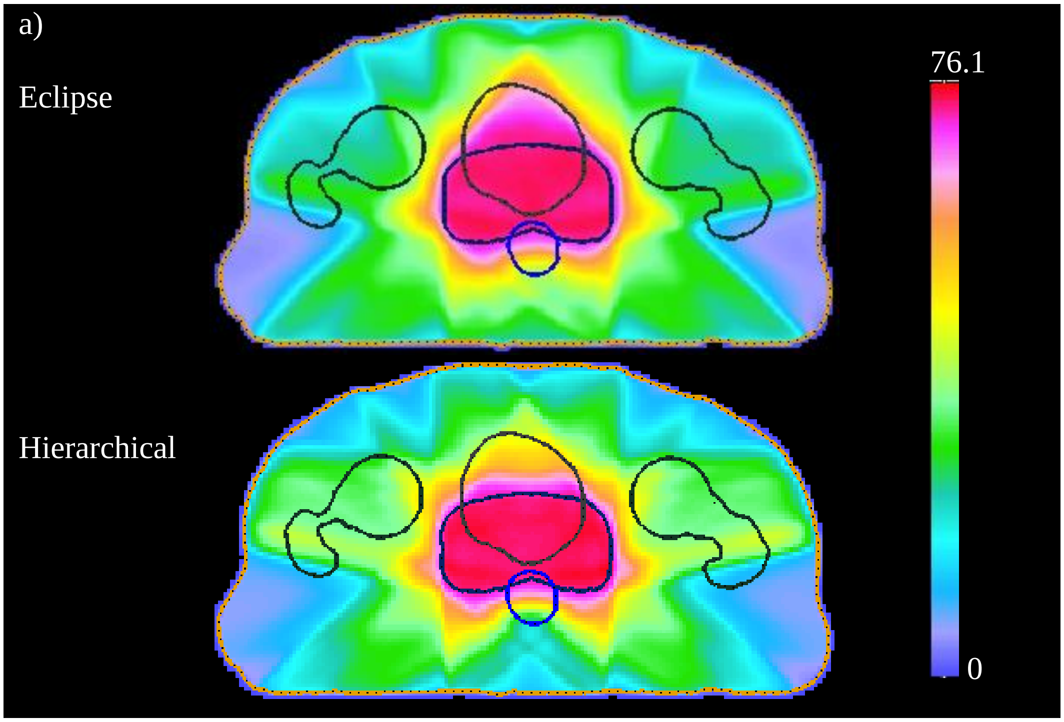


Figure 4-3 Dose distribution of prostate treatment plan 5. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



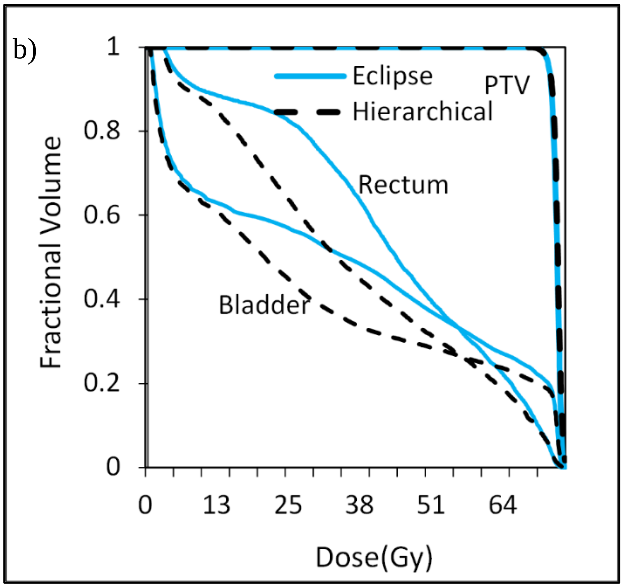
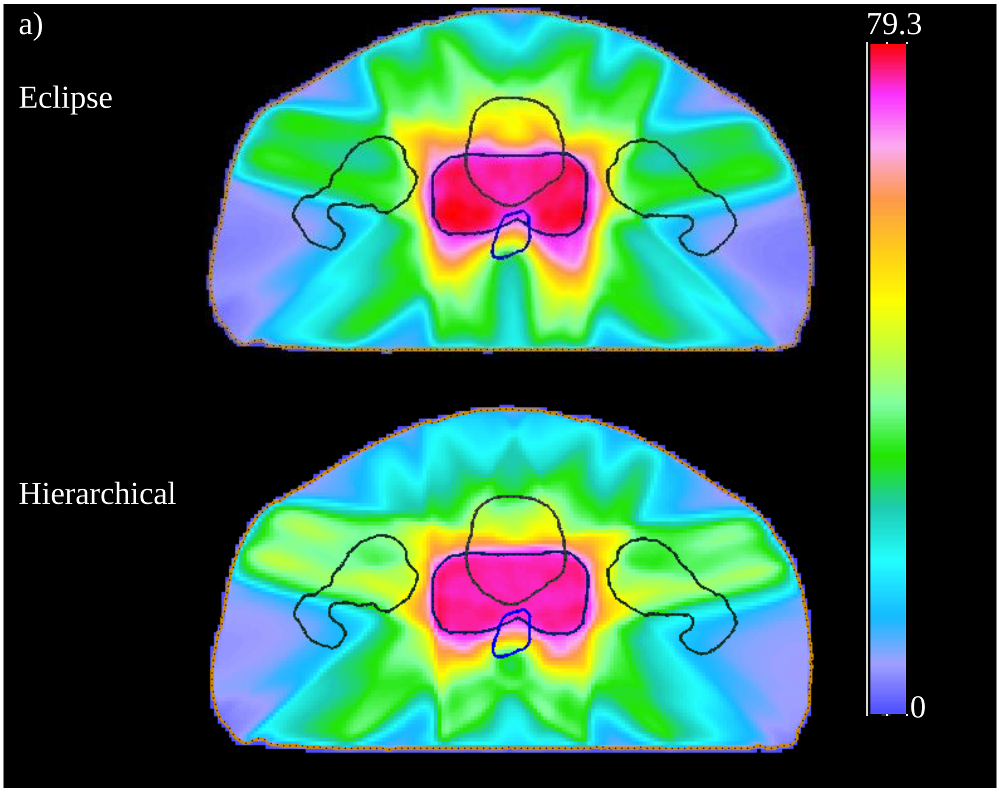


Figure 4-4 Dose distribution of prostate treatment plan 6. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



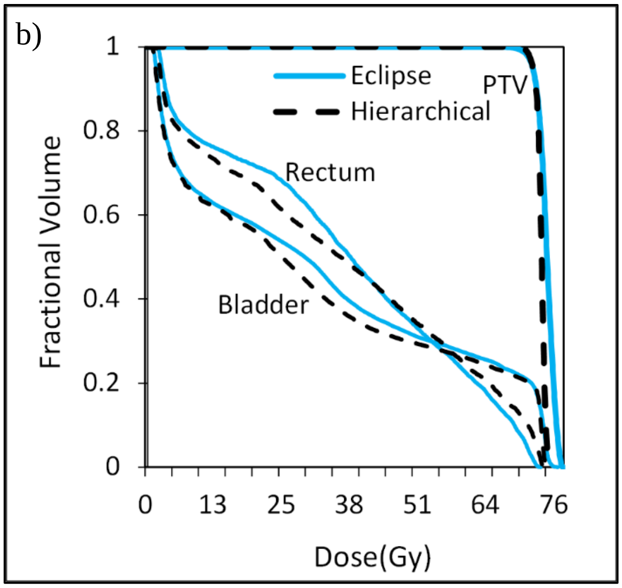
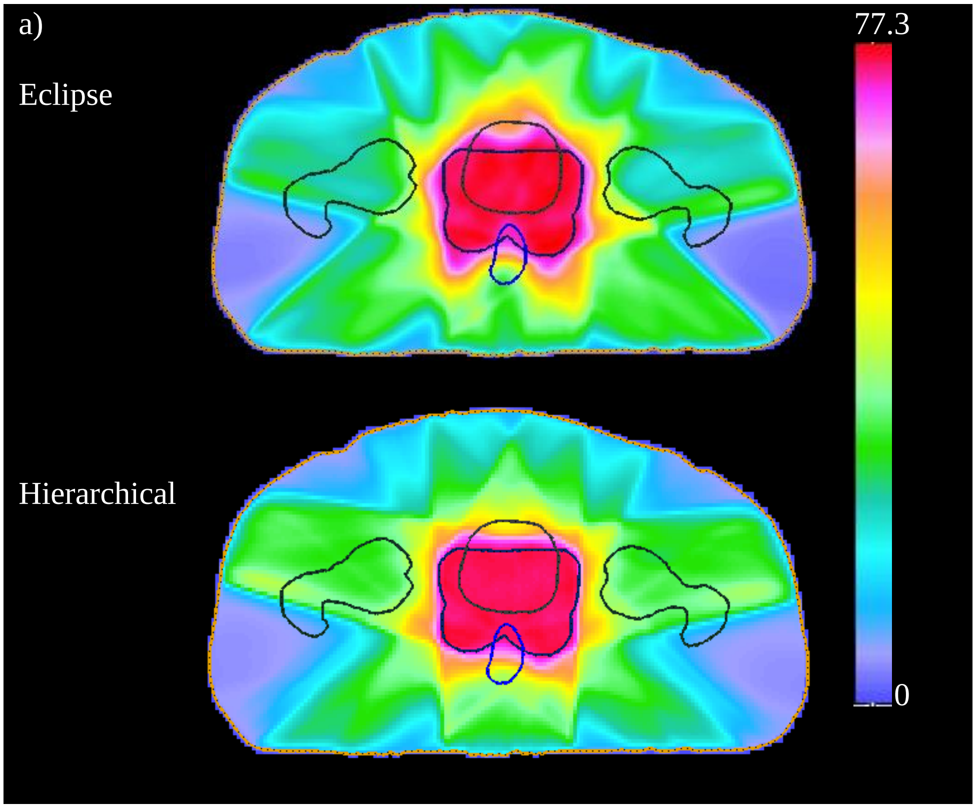


Figure 4-5 Dose distribution of prostate treatment plan 7. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. Dose-volume



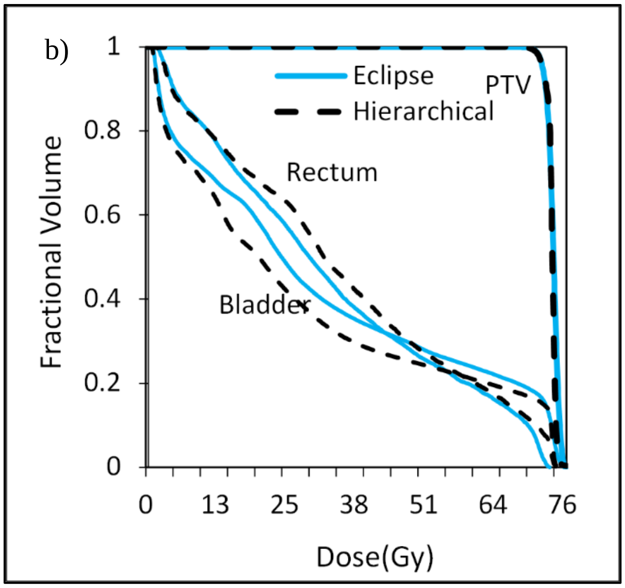
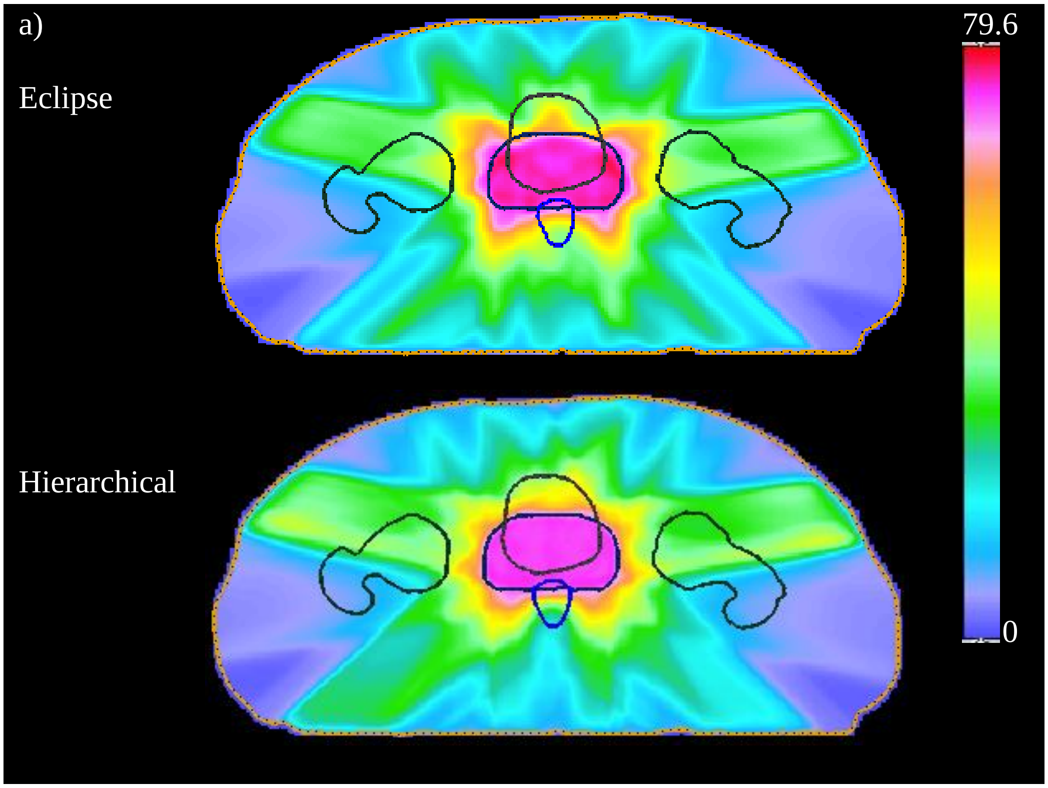


Figure 4-6 Dose distribution of prostate treatment plan 8. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in



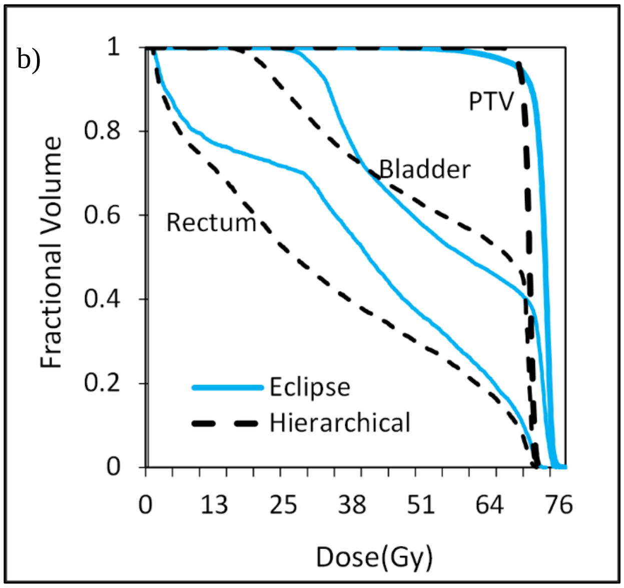


Figure 4-7 Dose distribution of prostate treatment plan 10. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.

Table 4-7 The volume that receives at least 47Gy (V47) of the rectum and bladder of the hierarchical and Eclipse treatment plans. The hierarchical system produced the lower V47 in more treatment plans than that produced by the Eclipse treatment planning system.

Plan Number	Rectum V47 of hierarchical plans in % (a)	Rectum V47 of Eclipse plans in % (b)	Ratio a/b	Bladder V47 of hierarchical plans in % (c)	Bladder V47 of Eclipse plans in % (d)	Ratio c/d
2	28.4	40	0.71	27.6	33.7	0.81
5	52.85	52.56	1.005	21.19	24.67	0.85
6	40.13	52.13	0.76	31.94	35.97	0.88
7	51.98	50.89	1.02	36.55	38.04	0.96
8	39.5	37.44	1.05	34	37.07	0.91
10	42.64	56.48	0.75	70.37	63.88	1.10

4.5 Testing the Model

We prepared 25 prostate treatment plans using the hierarchical system to demonstrate the potential of hierarchical optimization to automate the treatment planning process. The treatment plans were evaluated using two different methods: 1) the hierarchical treatment plans were presented to the physicist of MSK, the physicist categorized the dose to target and normal structures into one of four categories: inferior, same, better and superior, based on their impression of the hierarchical and Eclipse plans, 2) the common dosimetry criteria were compared for treatment plans produced from the hierarchical and Eclipse treatment planning systems.

Based on the physicist’s impression and the common dosimetry criteria, the target dose from the hierarchical system seems better than that of the Eclipse system in the majority of treatment plans. The target dose was inferior in three treatment plans, the same in 12, better in seven and superior in three (Table 4-8, Figure 4-2 to Figure 4-32). Table 4-9 and Table 4-10 show dose to 99% of the target structure (D99), dose to 5% of the target structure (D05), and maximum and minimum dose to the target structure obtained from the hierarchical and Eclipse treatment planning systems. The average ratio of the D99 and D05 obtained from the hierarchical system to the Eclipse system was close to 1, suggesting that both systems produced the same target coverage. Similarly, the average ratio of the maximum dose to the target structure obtained from the hierarchical system to the Eclipse system was close to 1, suggesting that the maximum dose obtained from the hierarchical system was the same as that obtained from the Eclipse system. Finally, the minimum dose to the target structure obtained from the hierarchical system was almost 5% higher than that obtained from the Eclipse system. The higher minimum dose suggests that the hierarchical system didn't produce any cold spots in the target structure.

The rectal dose of the hierarchical system seems better than that of the Eclipse system in the majority of treatment plans. The rectal dose was inferior in four treatment plans, the same in four, better in six and superior in 11 (Table 4-8, Figure 4-2 to Figure 4-32).

Table 4-11 shows the ratio between the rectum V47 of the hierarchical and Eclipse systems. The average ratio of V47 of the rectum was 0.87, suggesting that rectum V47 of the hierarchical system was lower than that of the Eclipse system. Similarly,

Table 4-12 shows the ratio between the rectum maximum dose of the hierarchical and the Eclipse systems. The average ratio was 1, which implied that both systems produced the same maximum dose. Finally,

Table 4-12 shows the ratio between the mean rectal dose of the hierarchical and Eclipse systems. The average ratio of the mean rectal dose was 0.91, suggesting that the mean rectal dose of the hierarchical system was lower than that of the Eclipse system. The physicist's impression and the dosimetry metrics supports that the rectal dose of the hierarchical system was better than that of the Eclipse treatment planning system.

The bladder dose of the hierarchical system was substantially lower than that of the Eclipse treatment planning system. The bladder dose was inferior in one treatment plan, the same in three, better in 13 and superior in eight (Table 4-8) (Figure 4-2 to Figure 4-32). Table 4-13 shows the ratio between the bladder V47 of the hierarchical and Eclipse systems. The average ratio of V47 of the bladder was 0.91, suggesting that bladder V47 of the hierarchical system was lower than that of the Eclipse system. Similarly, Table 4-14 shows the ratio between the bladder maximum dose of the hierarchical and Eclipse systems. The average ratio was 1, which implied that both systems produced the same maximum dose. Finally, Table 4-14 shows the ratio between the mean bladder dose of the hierarchical and Eclipse systems. The average ratio of the mean dose to the bladder was 0.92, suggesting that the bladder mean dose of the hierarchical system was lower than that of the Eclipse system. The result suggests that the bladder dose of the hierarchical system was better than that of the Eclipse treatment planning system.

The clinical criteria require that the femur maximum dose must be less than 68Gy. The dose to femur obtained from the hierarchical system satisfied the maximum dose criteria (Table 4-15). However, the mean dose to the femur of the hierarchical system was almost 10% higher than that of the Eclipse system (Table 4-15). The femur was considered a lower priority structure in the optimization. Thus, the mean dose to the femur increased due to a decrease in the mean dose to the bladder and rectum. Nevertheless, the femur satisfied the maximum dose constraint.

Table 4-8 Categorization of the dose to target and normal structure into one of four categories: inferior, same, better, and superior. Physicists from Memorial Sloan Kettering Cancer Center categorized dose distribution into one of the four categories based on the overall impression of dose distribution of the treatment plans obtained from the hierarchical and Eclipse treatment planning systems.

Plan Number	Target Dose	Rectum Dose	Bladder Dose
12	Superior	Better	Same
13	Same	Superior	Better
15	Same	Better	Superior
16	Same	Inferior	Better
18	Superior	Better	Superior
19	Superior	Inferior	Superior
20	Same	Same	Better
21	Same	Superior	Superior
22	Same	Superior	Superior
24	Better	Better	Better
26	Same	Inferior	Better
27	Better	Superior	Inferior
28	Better	Better	Better
29	Better	Inferior	Superior
30	Better	Superior	Superior
31	Same	Same	Better
32	Inferior	Superior	Same
33	Same	Superior	Superior
34	Better	Better	Same
35	Inferior	Superior	Better
36	Better	Same	Better
37	Same	Superior	Better
38	Same	Superior	Better
39	Same	Superior	Better
40	Inferior	Same	Better

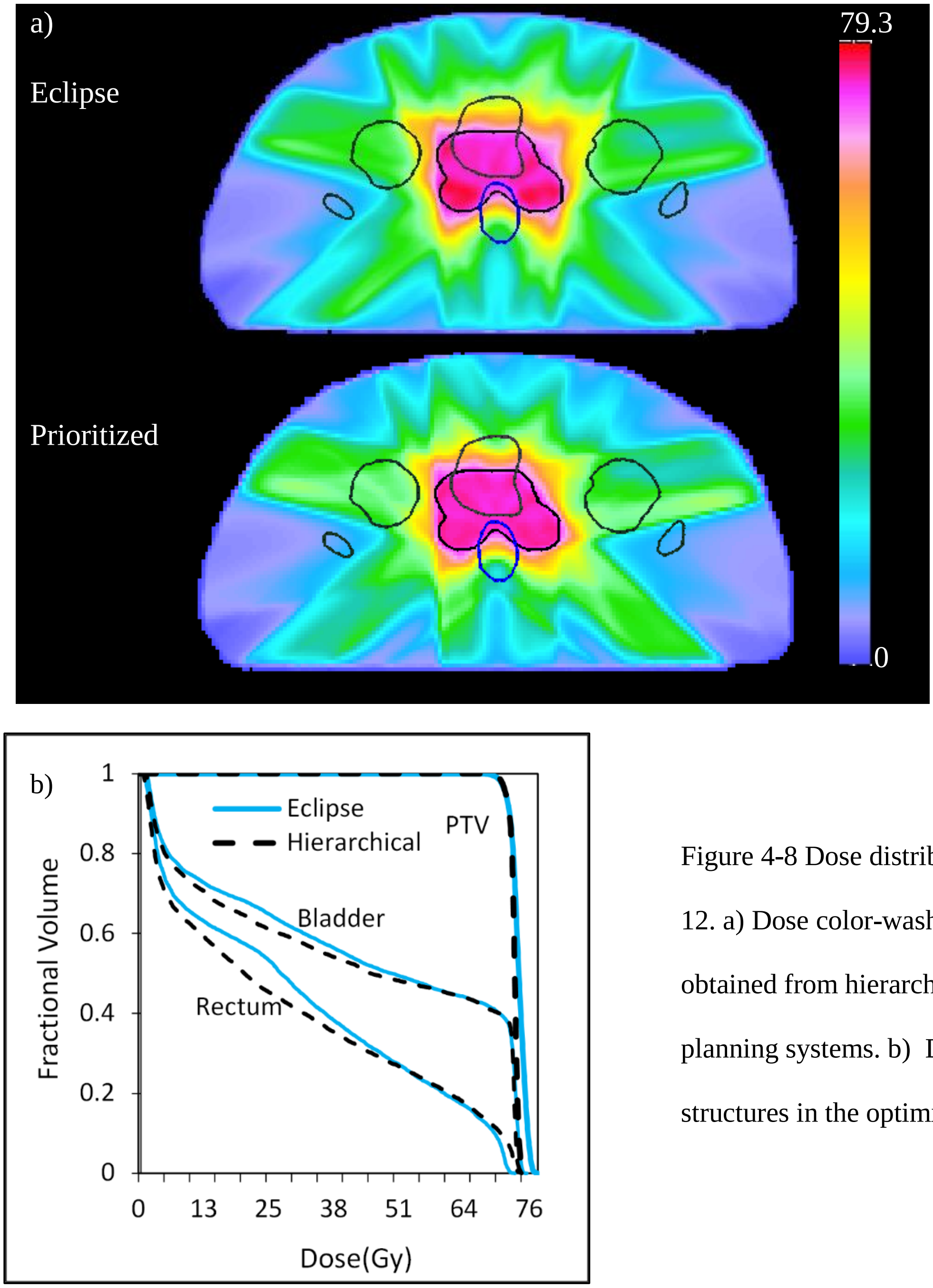


Figure 4-8 Dose distribution of prostate treatment plan 12. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.

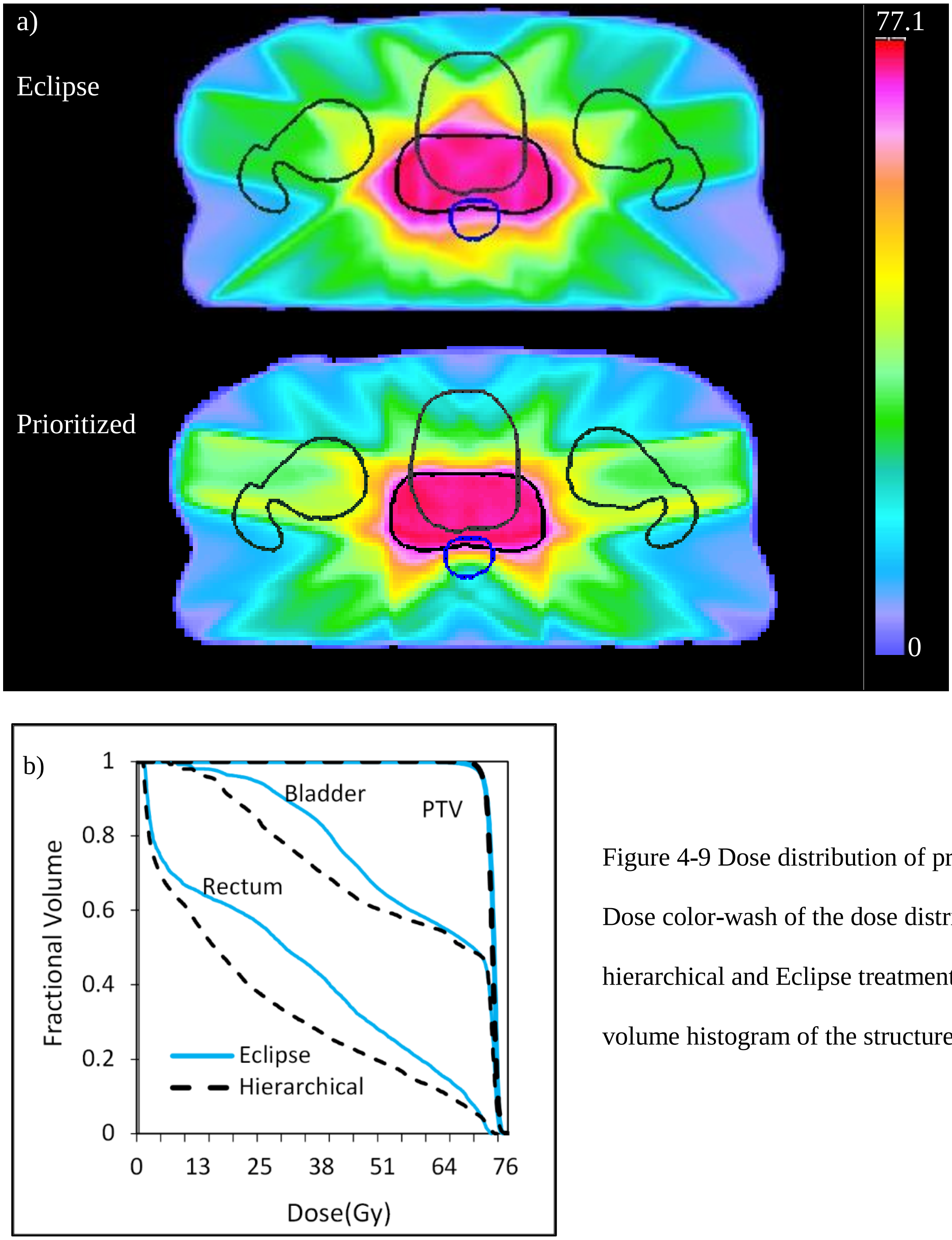
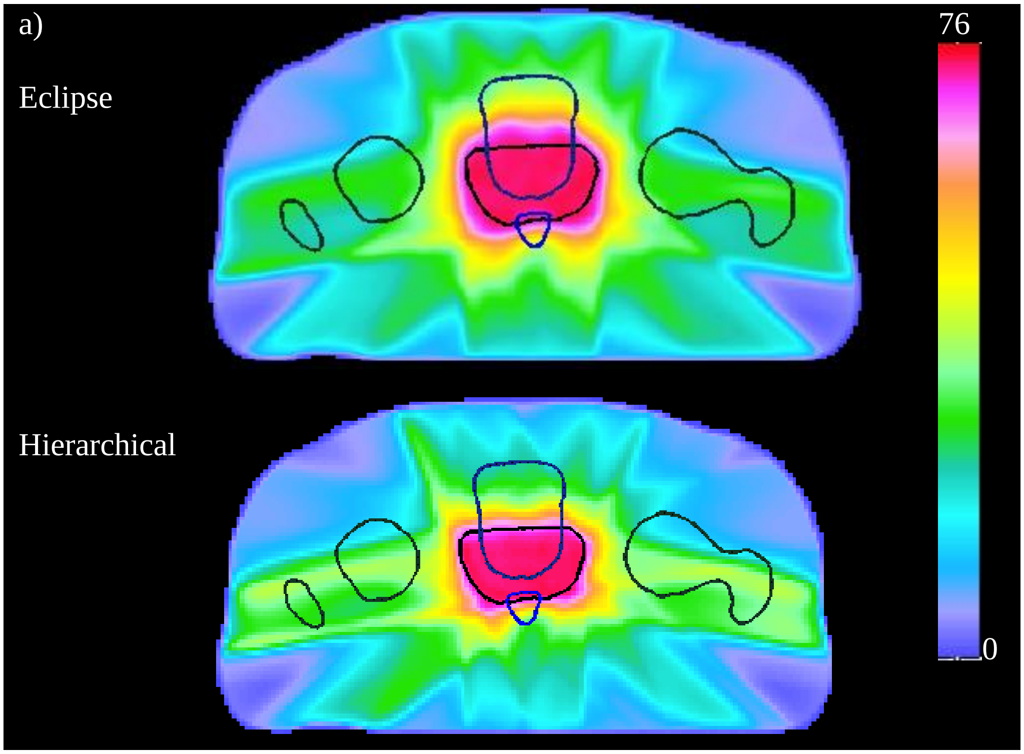


Figure 4-9 Dose distribution of prostate treatment plan 13. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



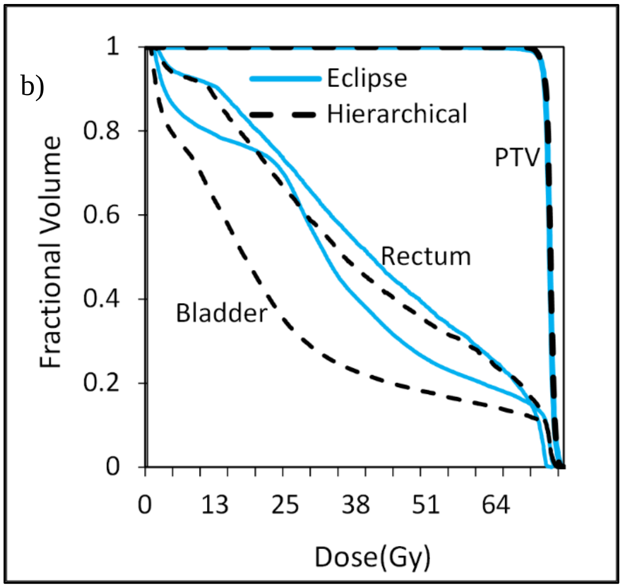
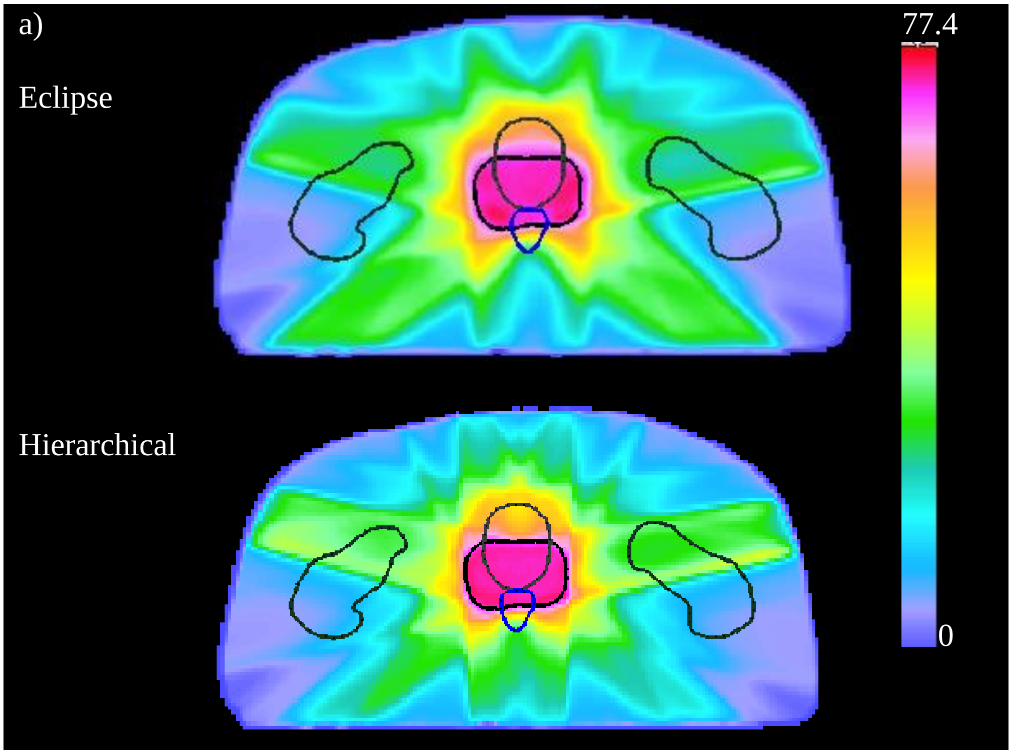


Figure 4-10 Dose distribution of prostate treatment plan 15. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



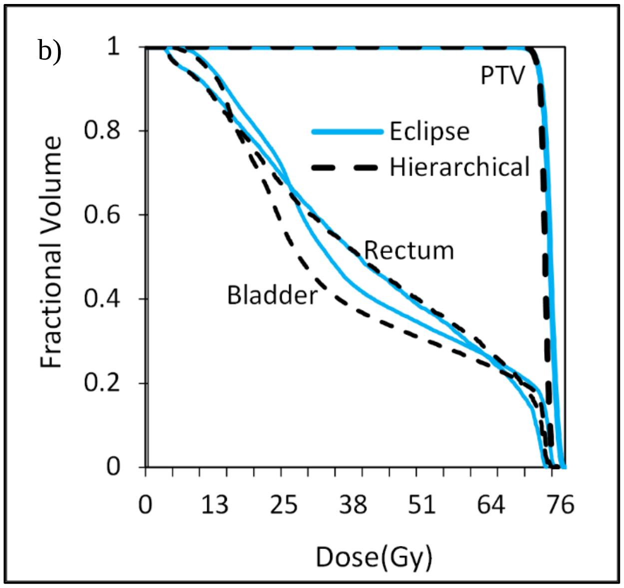
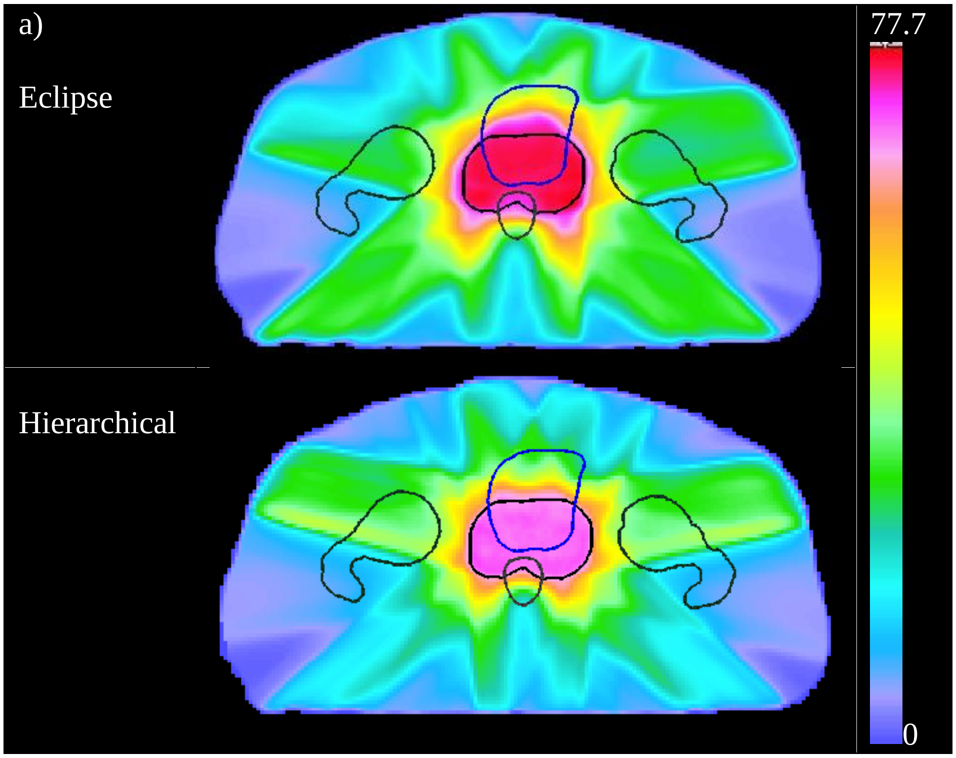


Figure 4-11 Dose distribution of prostate treatment plan 16. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



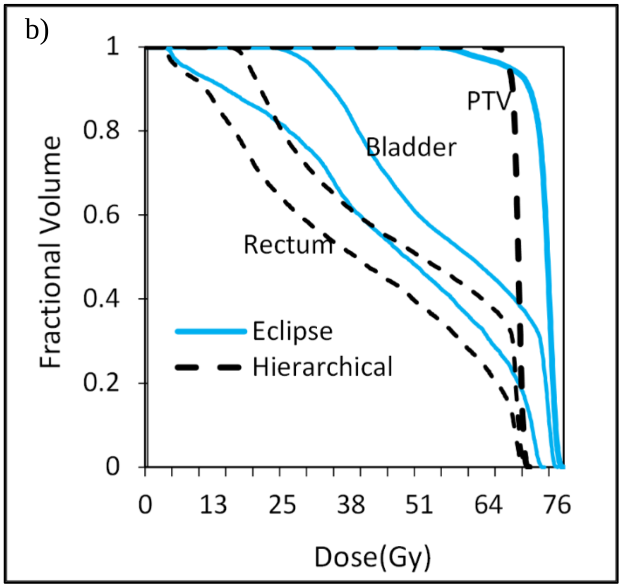
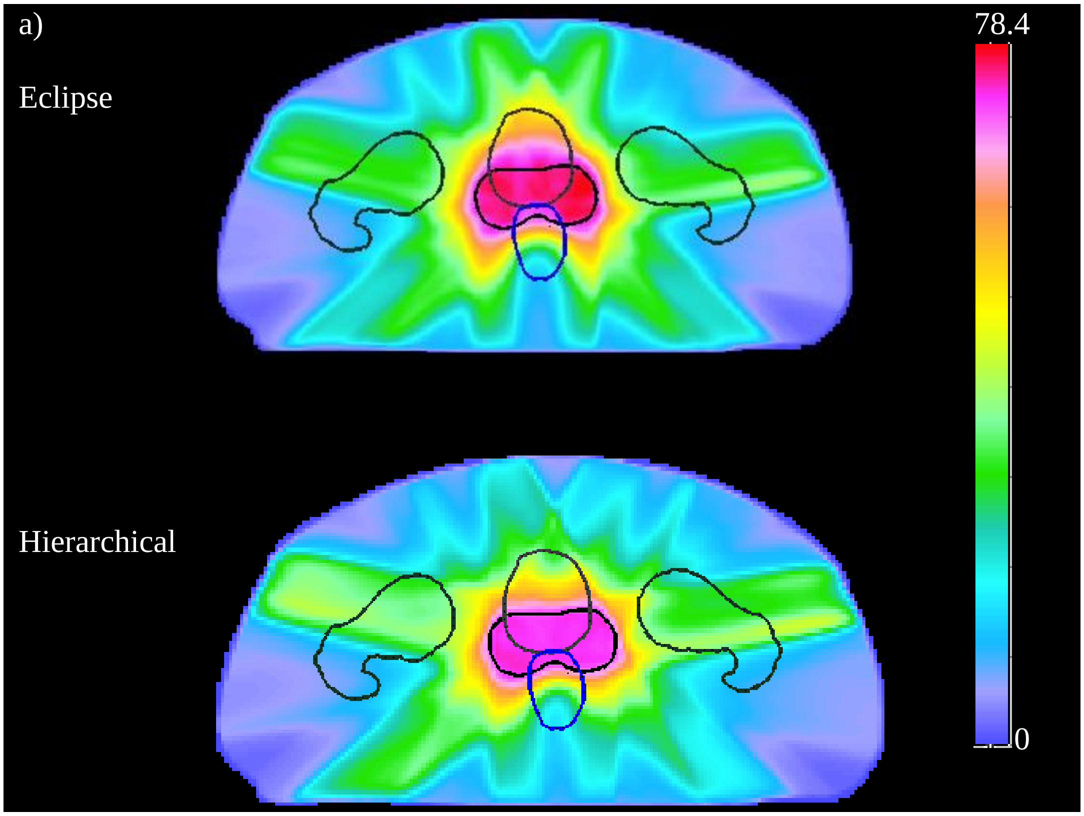


Figure 4-12 Dose distribution of prostate treatment plan 18. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



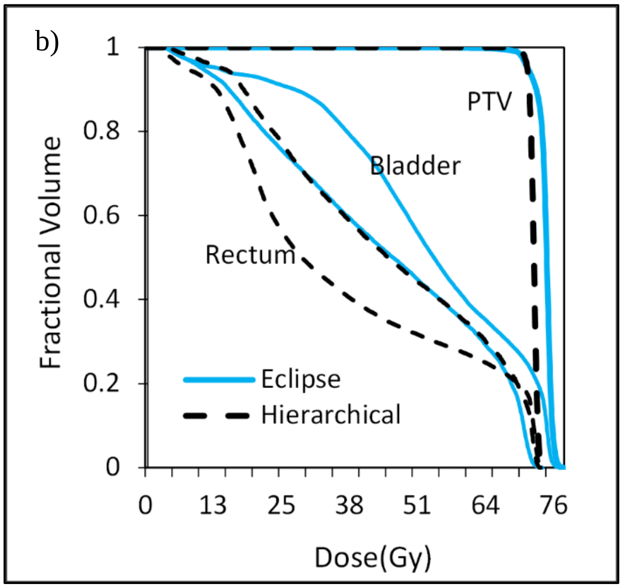
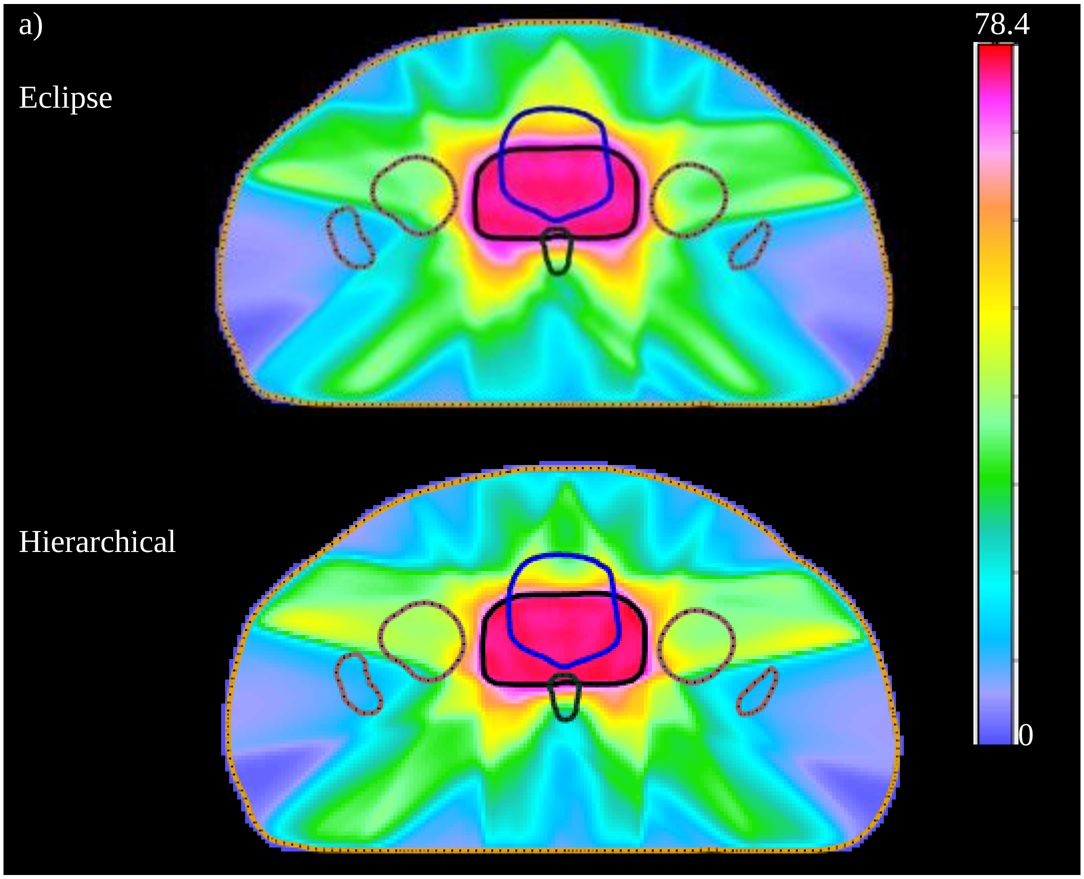


Figure 4-13 Dose distribution of prostate treatment plan 19. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



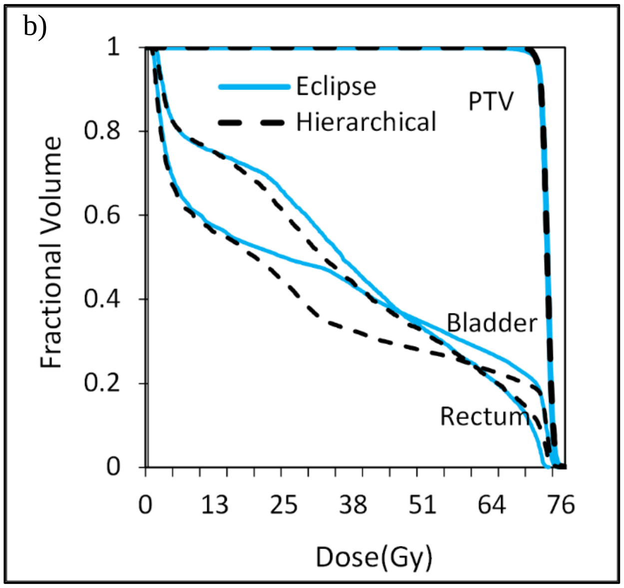
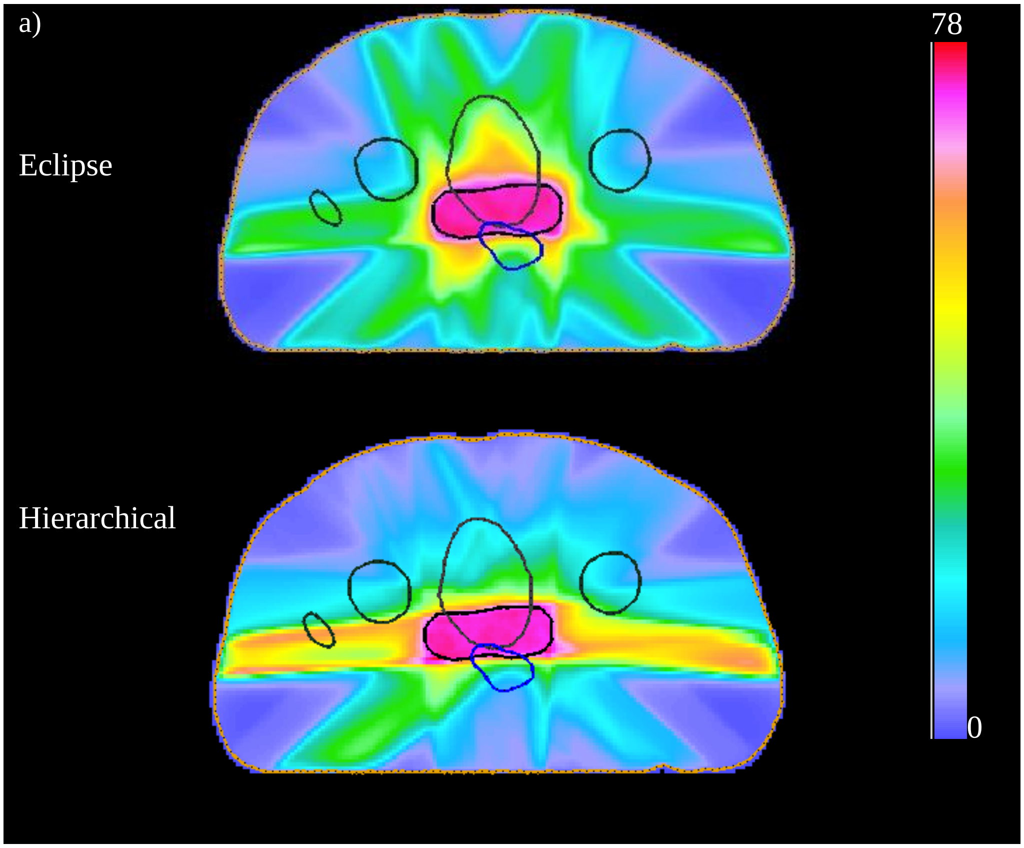


Figure 4-14 Dose distribution of prostate treatment plan 20. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems.) Dose-volume histogram of the structures in the optimization.



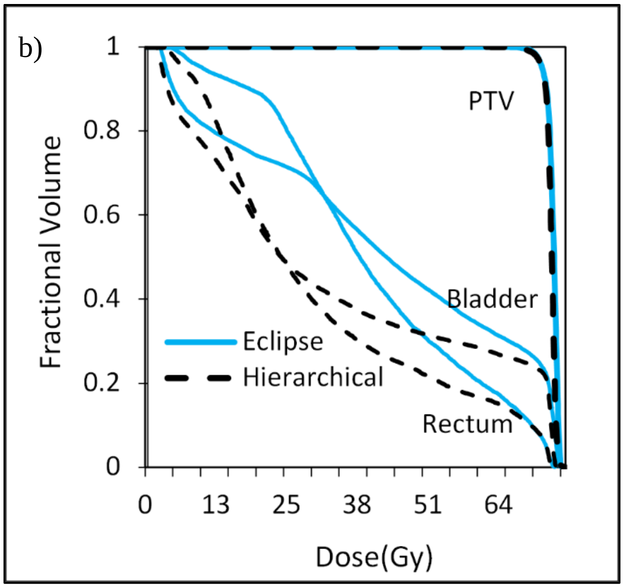


Figure 4-15 Dose distribution of prostate treatment plan 21.) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.

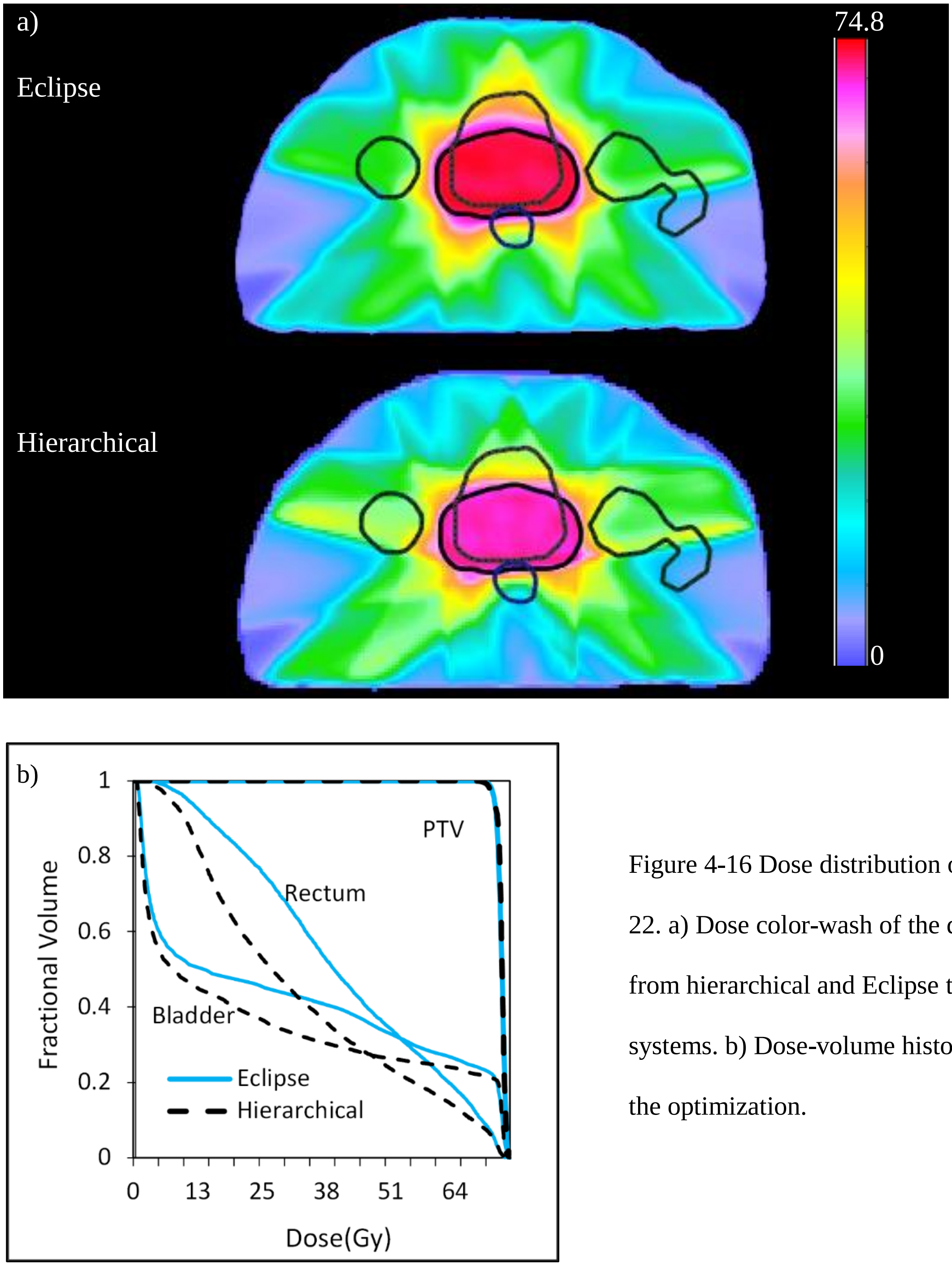
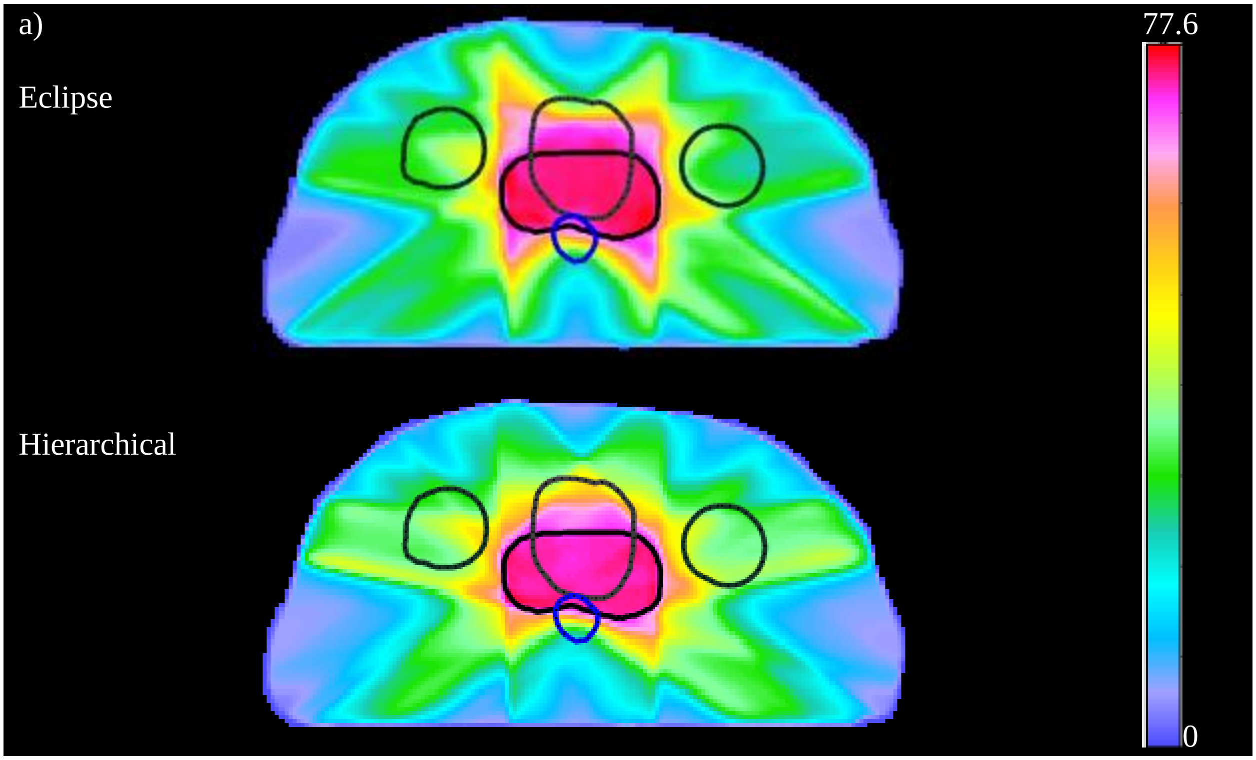


Figure 4-16 Dose distribution of prostate treatment plan 22. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



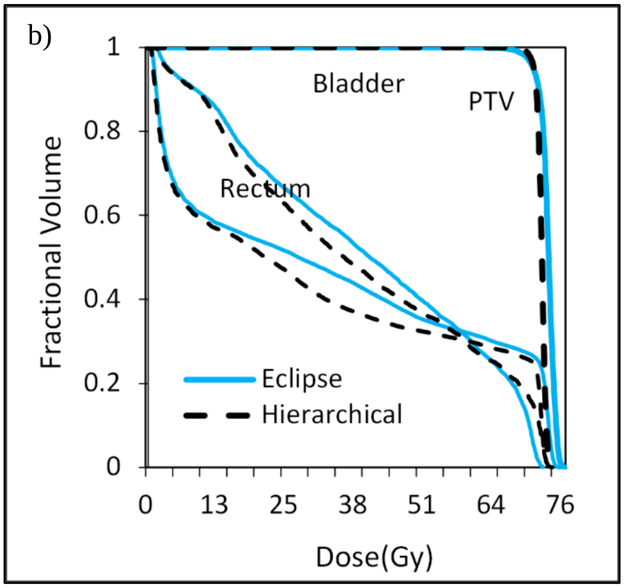
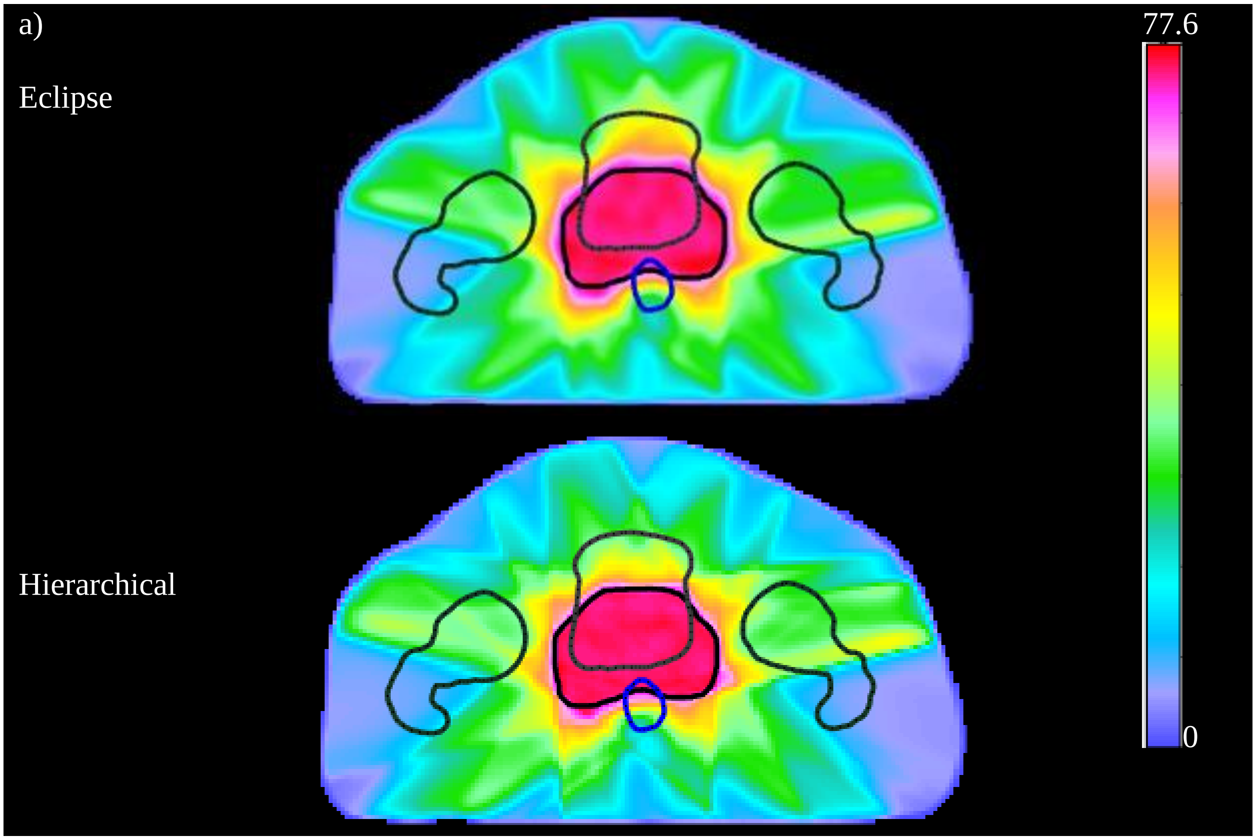


Figure 4-17 Dose distribution of prostate treatment plan 24. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems.) Dose-volume histogram of the structures in the optimization.



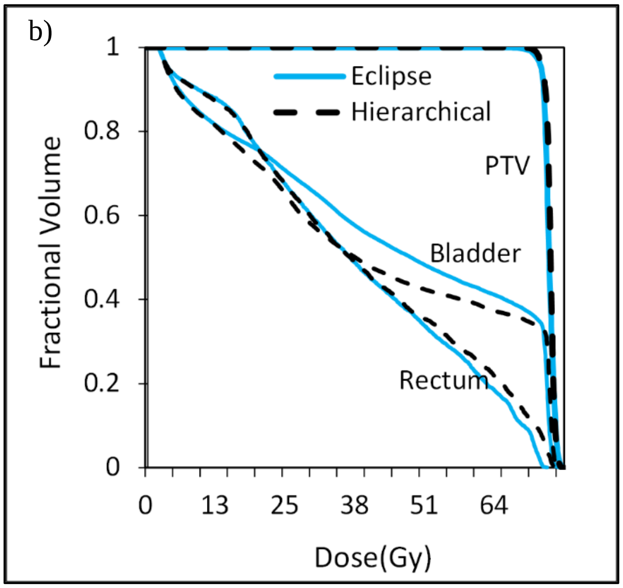
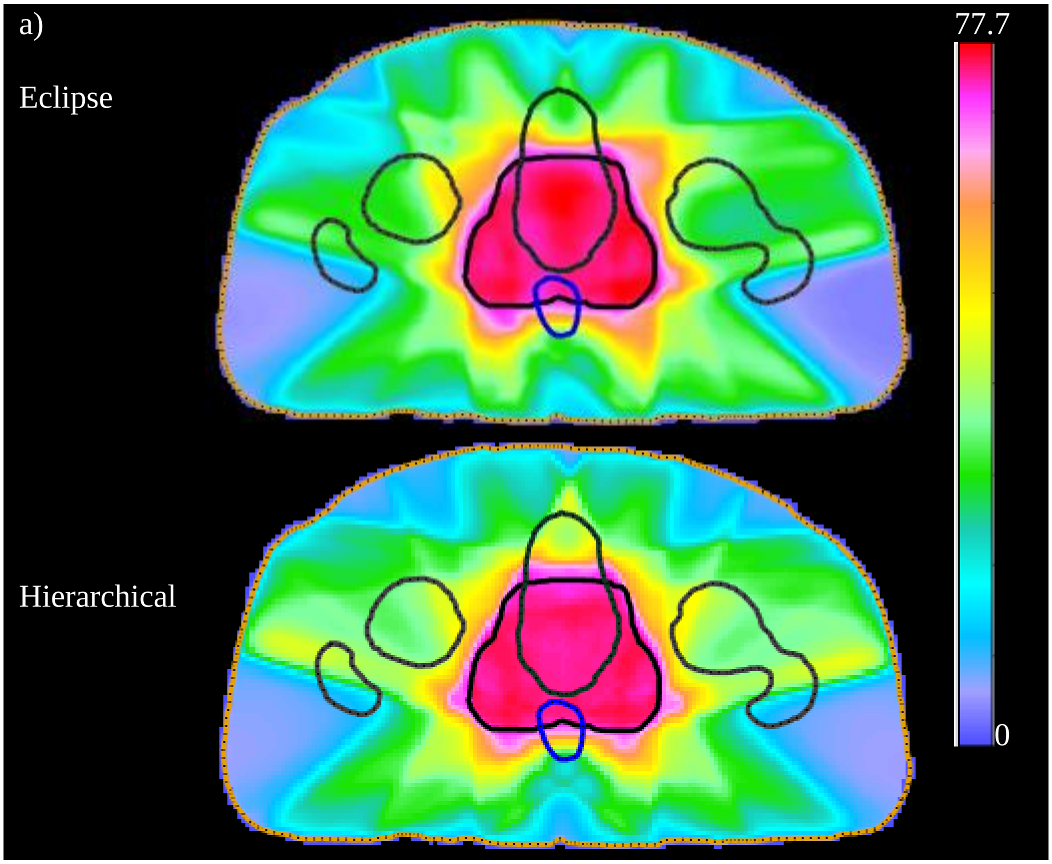


Figure 4-18 Dose distribution of prostate treatment plan 26. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



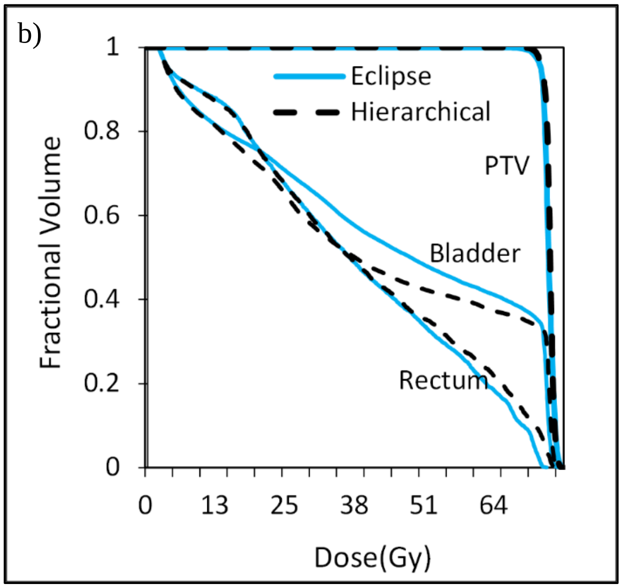
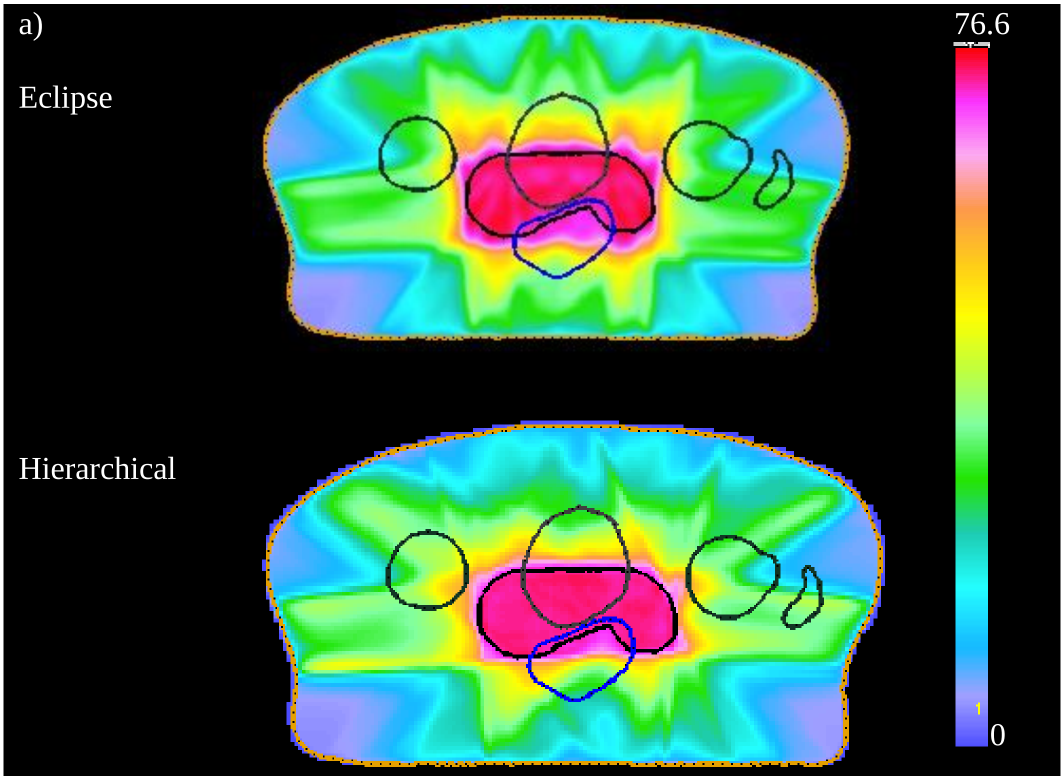


Figure 4-19 Dose distribution of prostate treatment plan 27. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



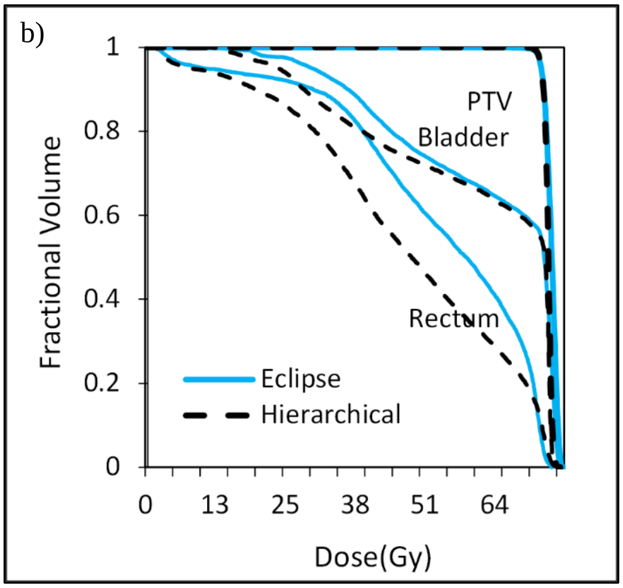
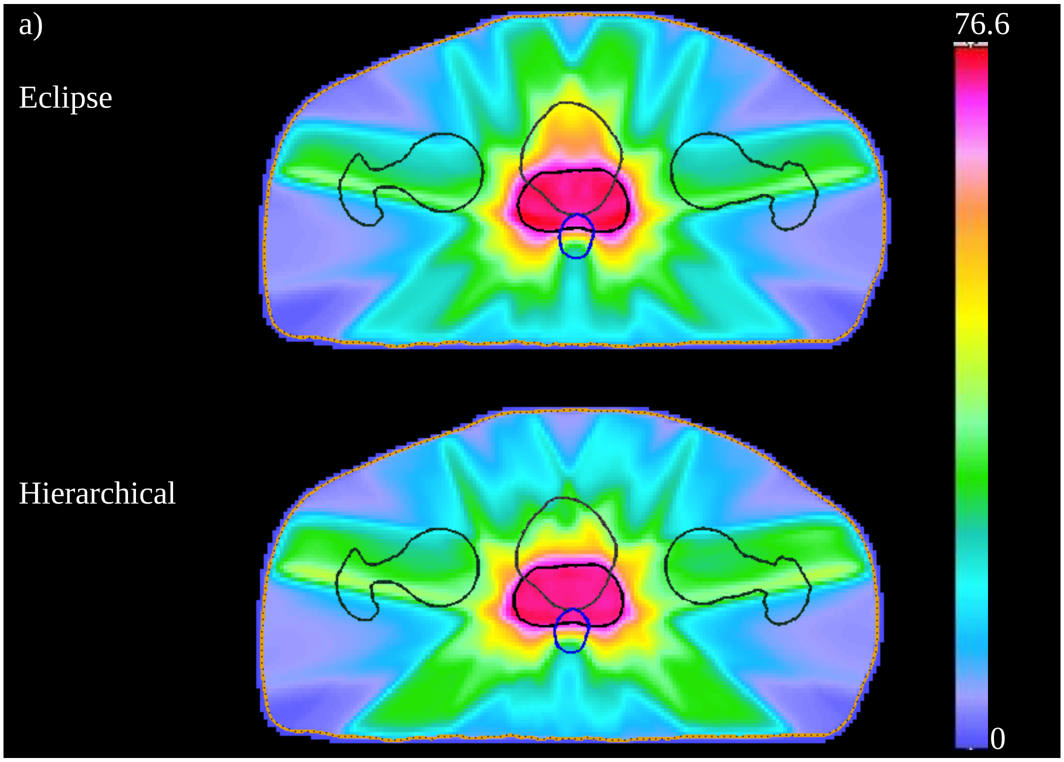


Figure 4-20 Dose distribution of prostate treatment plan 28. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems.) Dose-volume histogram of the structures in the optimization.



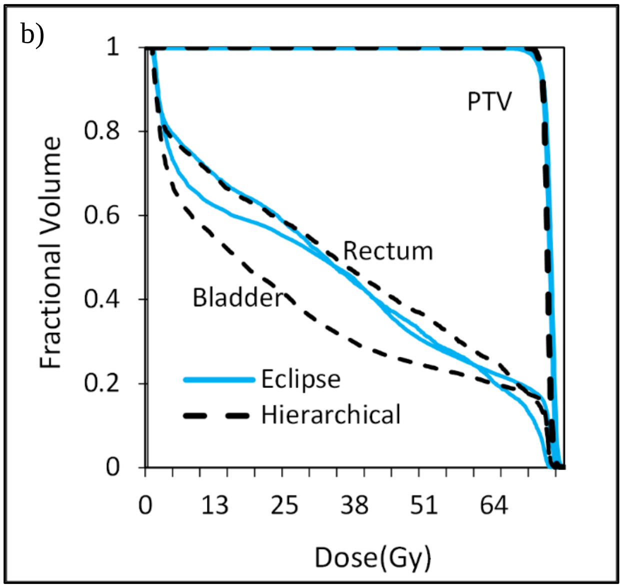


Figure 4-21 Dose distribution of prostate treatment plan 29. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.

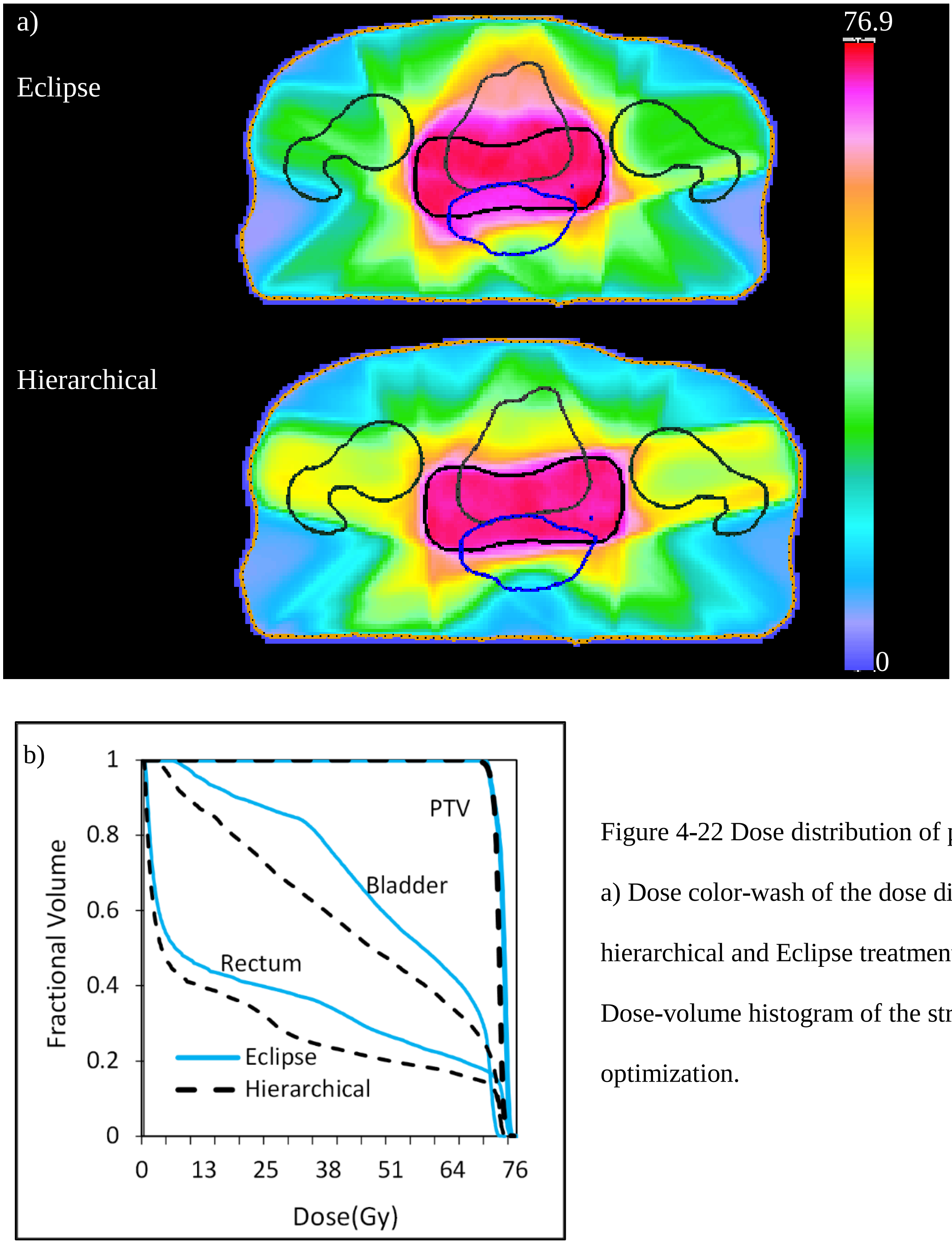
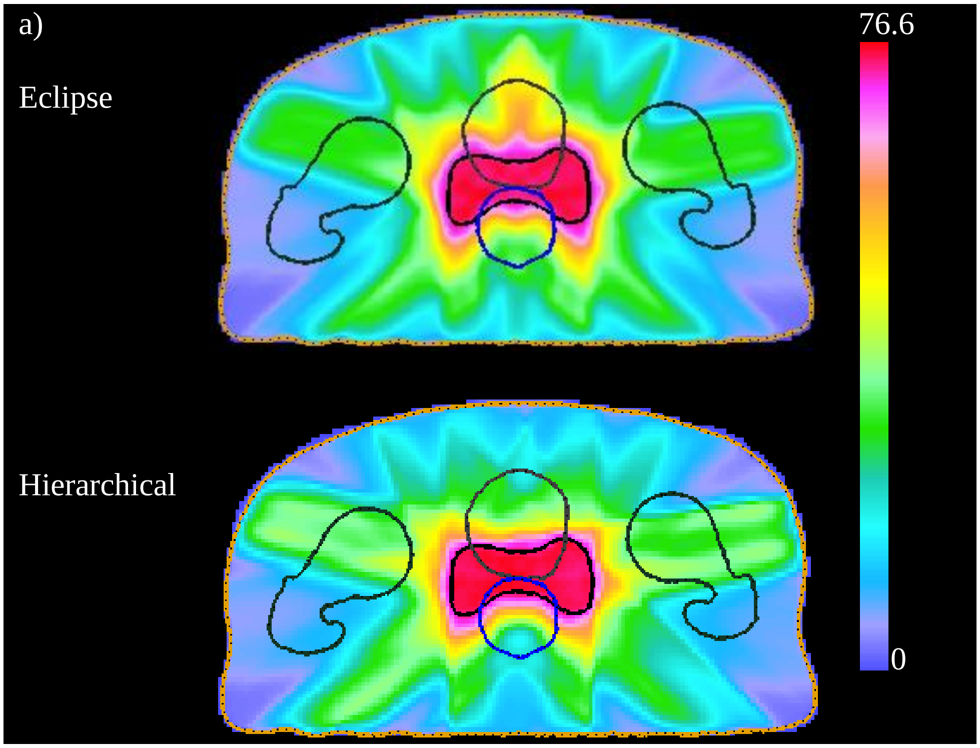


Figure 4-22 Dose distribution of prostate treatment plan 30. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems.) Dose-volume histogram of the structures in the optimization.



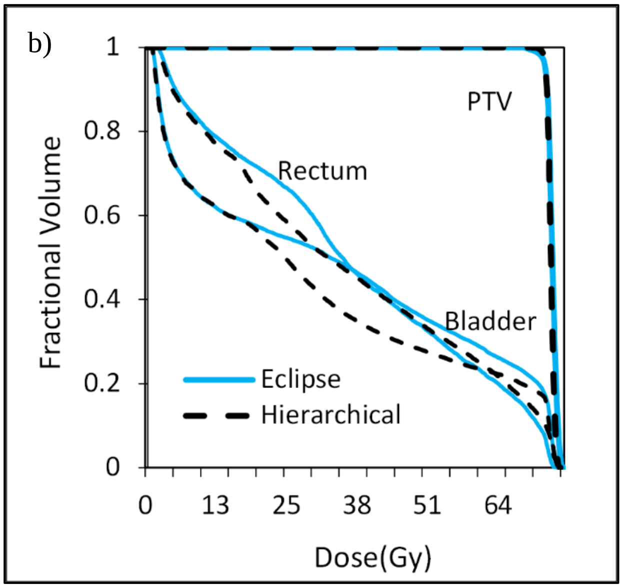
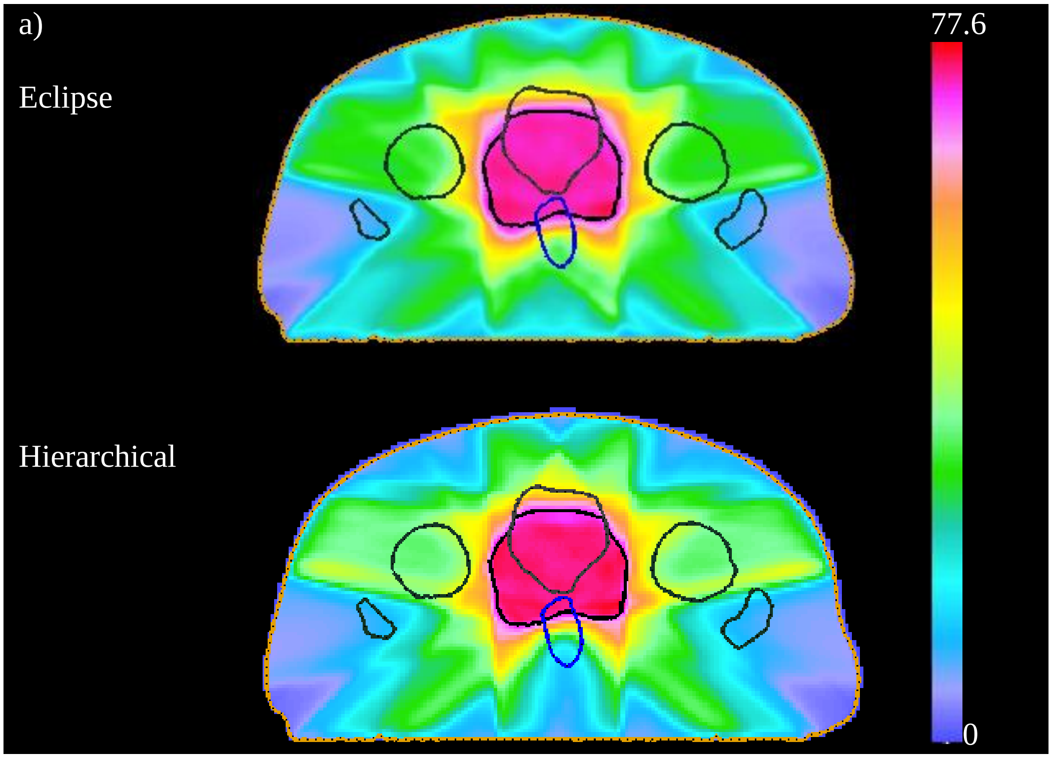


Figure 4-23 Dose distribution of prostate treatment plan 31. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems.) Dose-volume histogram of the structures in the optimization.



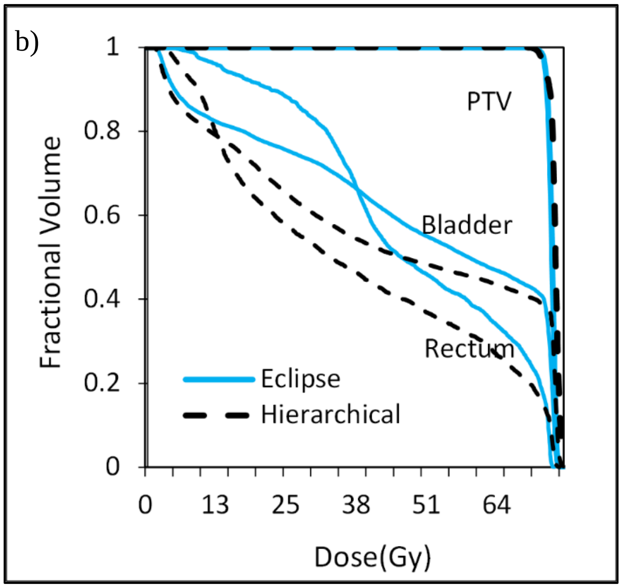
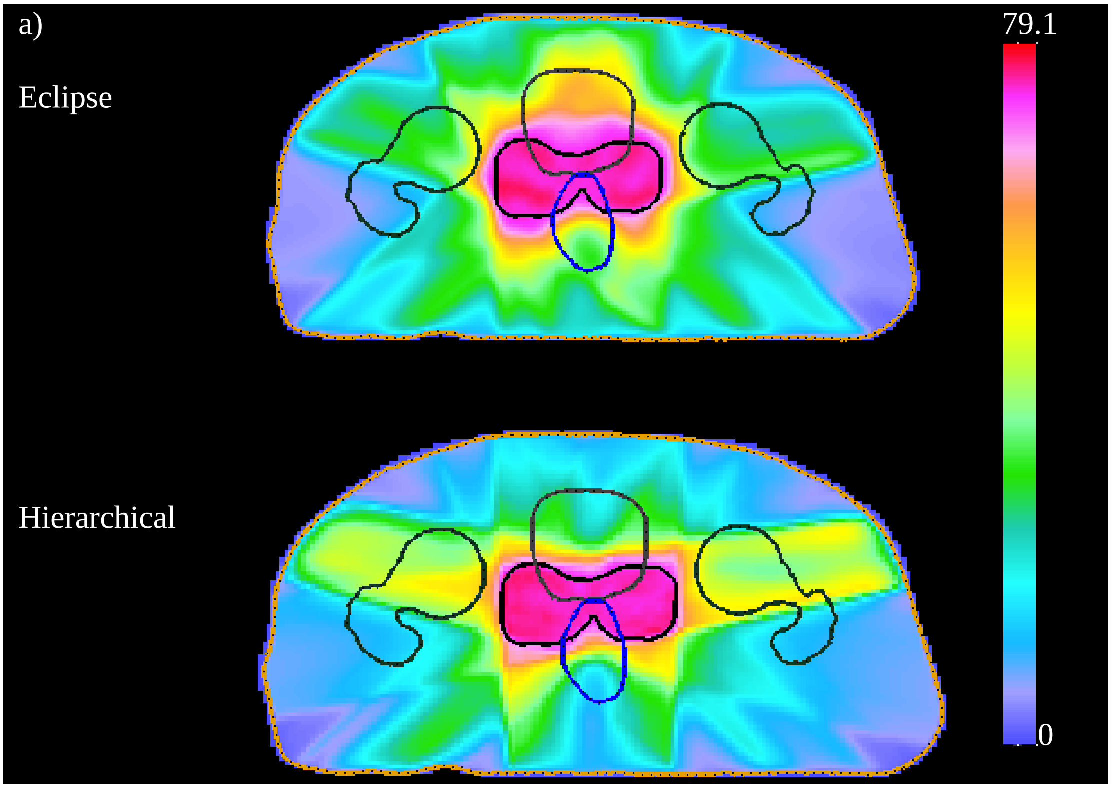


Figure 4-24 Dose distribution of prostate treatment plan 32. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



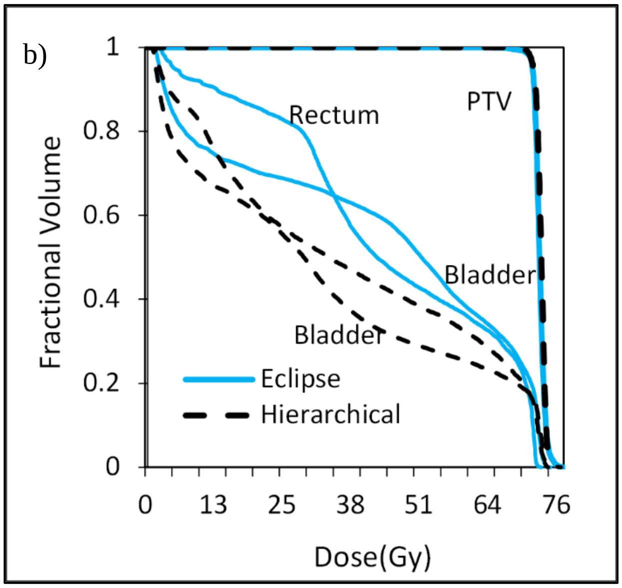
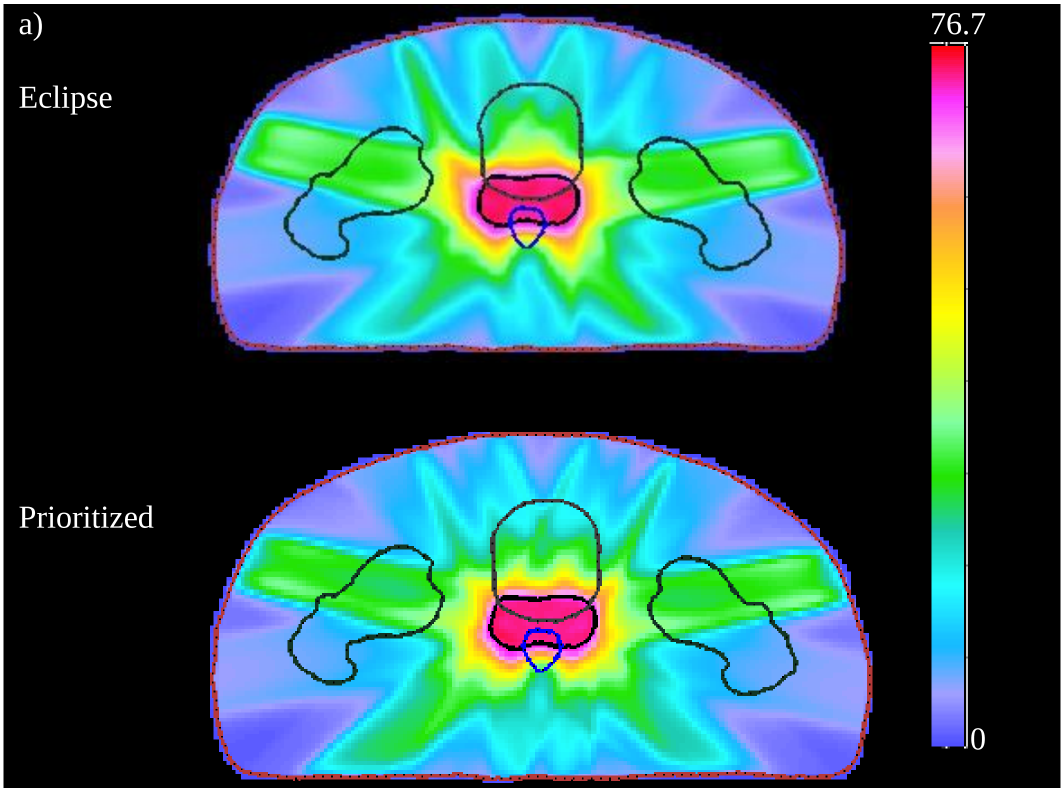


Figure 4-25 Dose distribution of prostate treatment plan 33.

a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems.) Dose-volume histogram of the structures in the optimization.



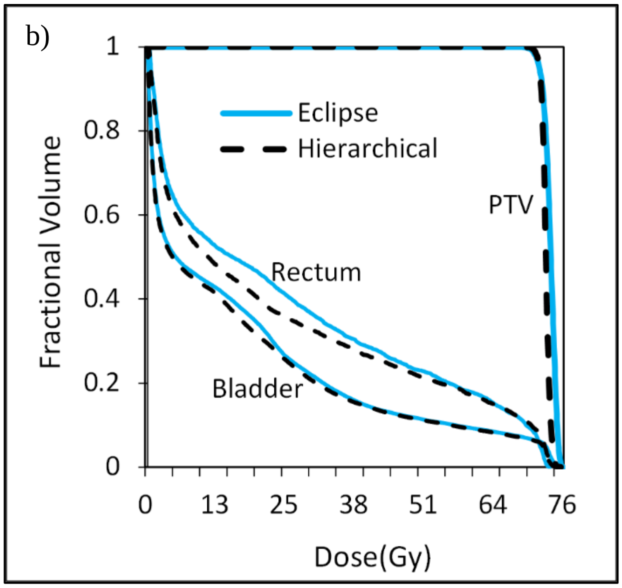
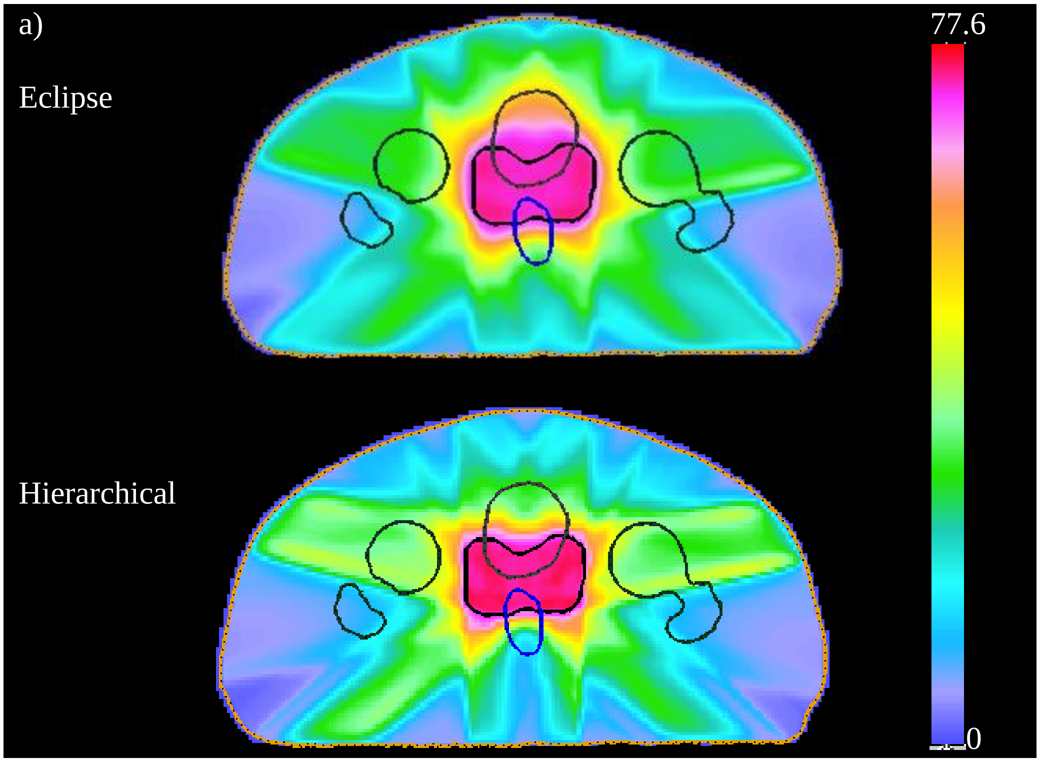


Figure 4-26 Dose distribution of prostate treatment plan 34. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization



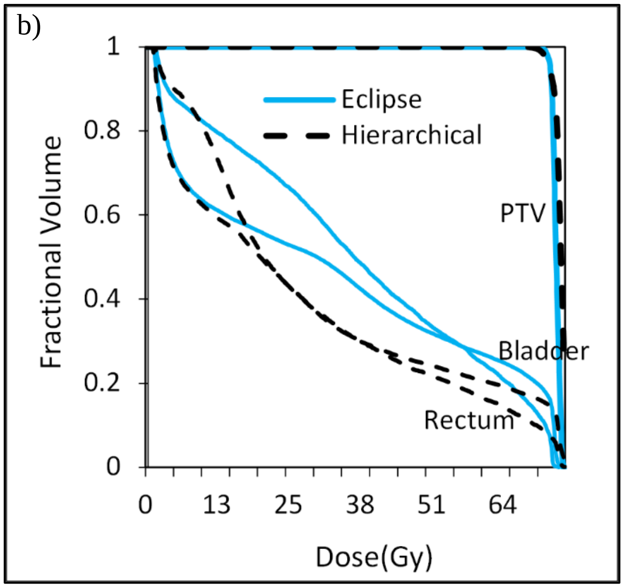
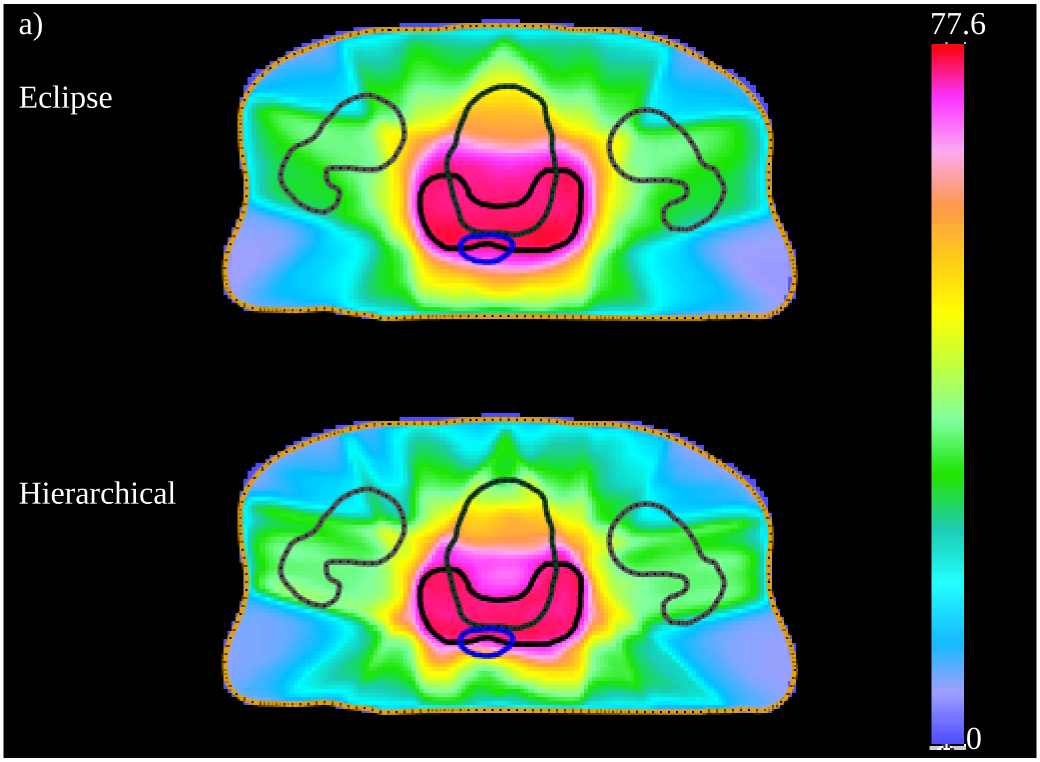


Figure 4-27 Dose distribution of prostate treatment plan 35. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



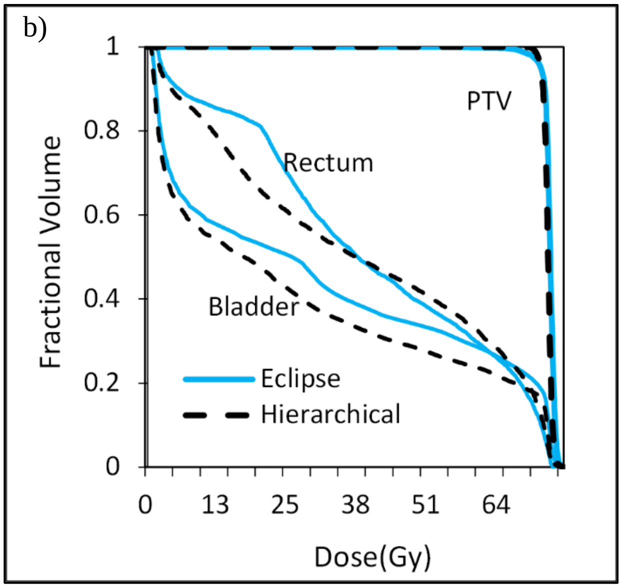
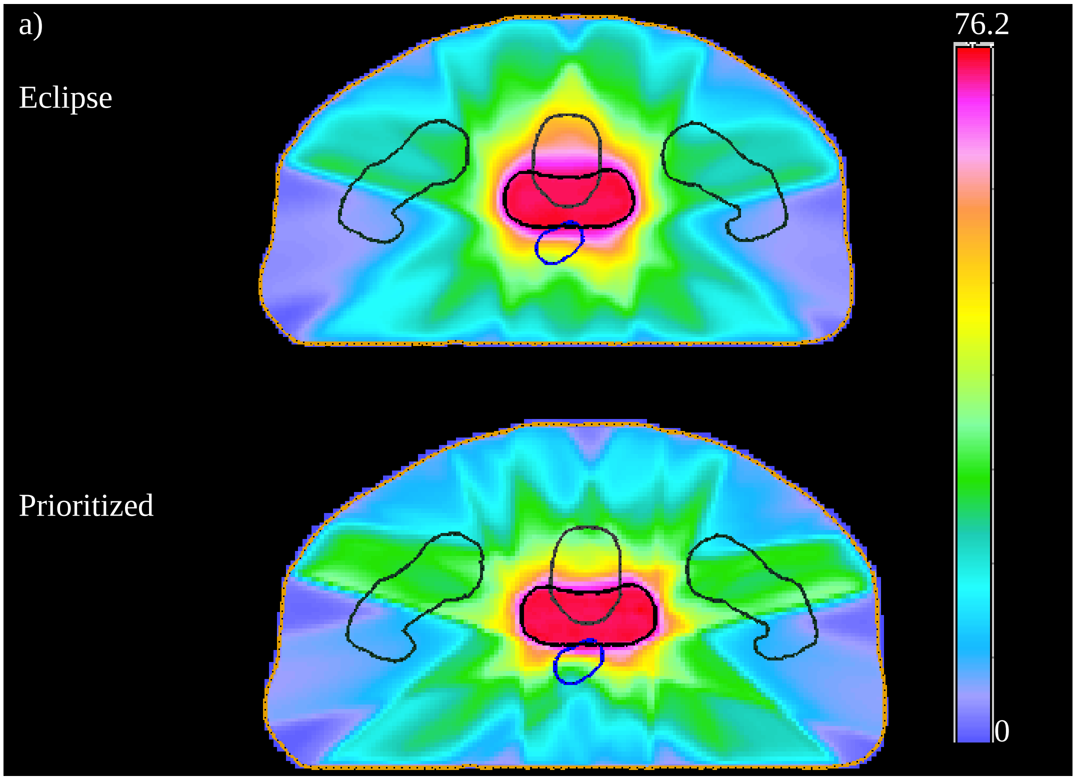


Figure 4-28 Dose distribution of treatment plan 36. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the



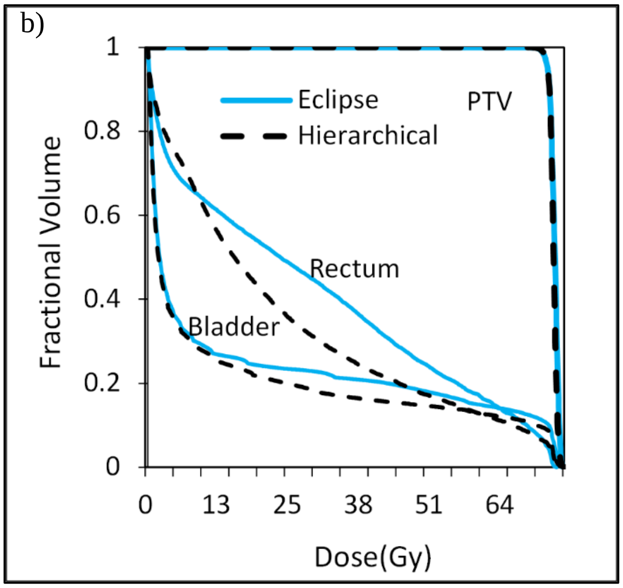
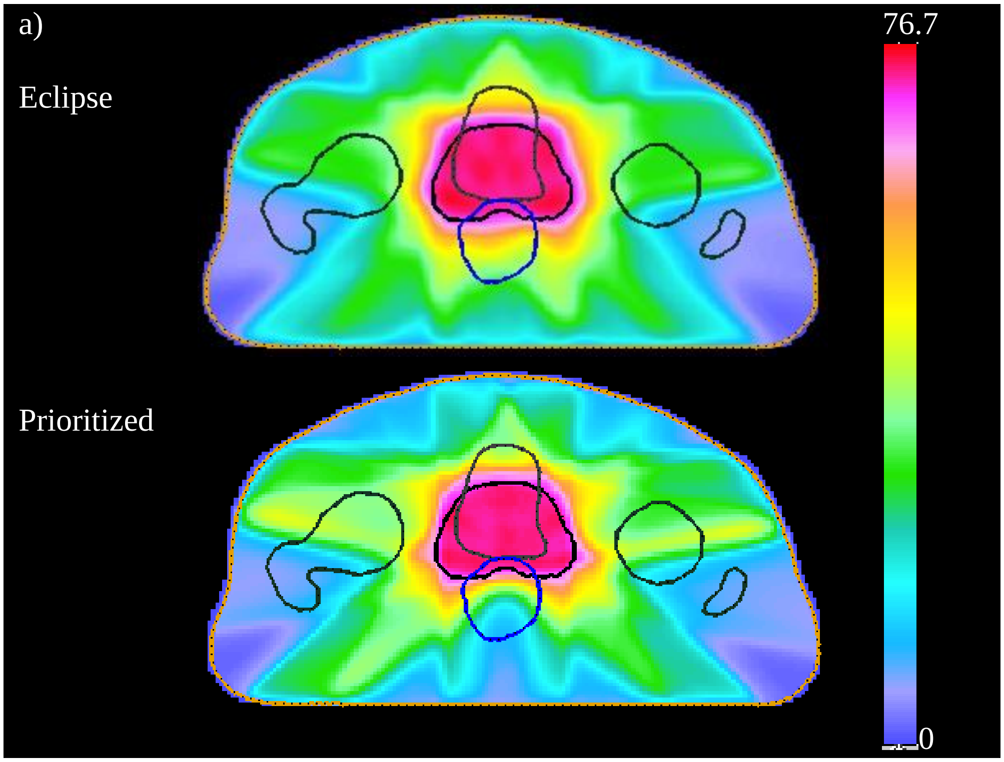


Figure 4-29 Dose distribution of prostate treatment plan 37. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



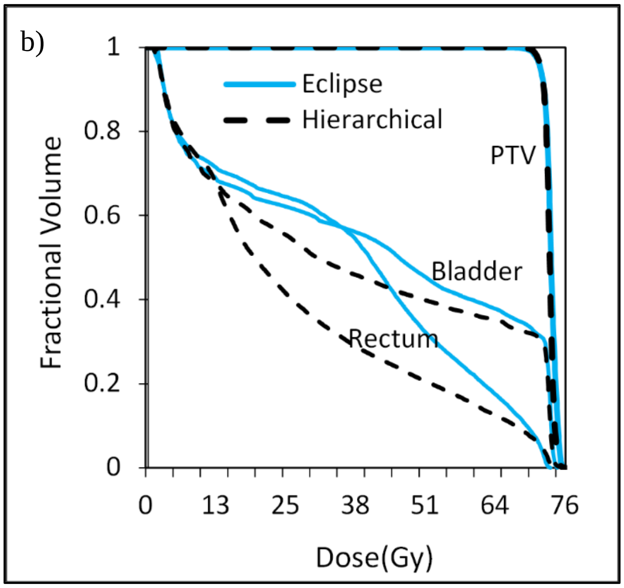
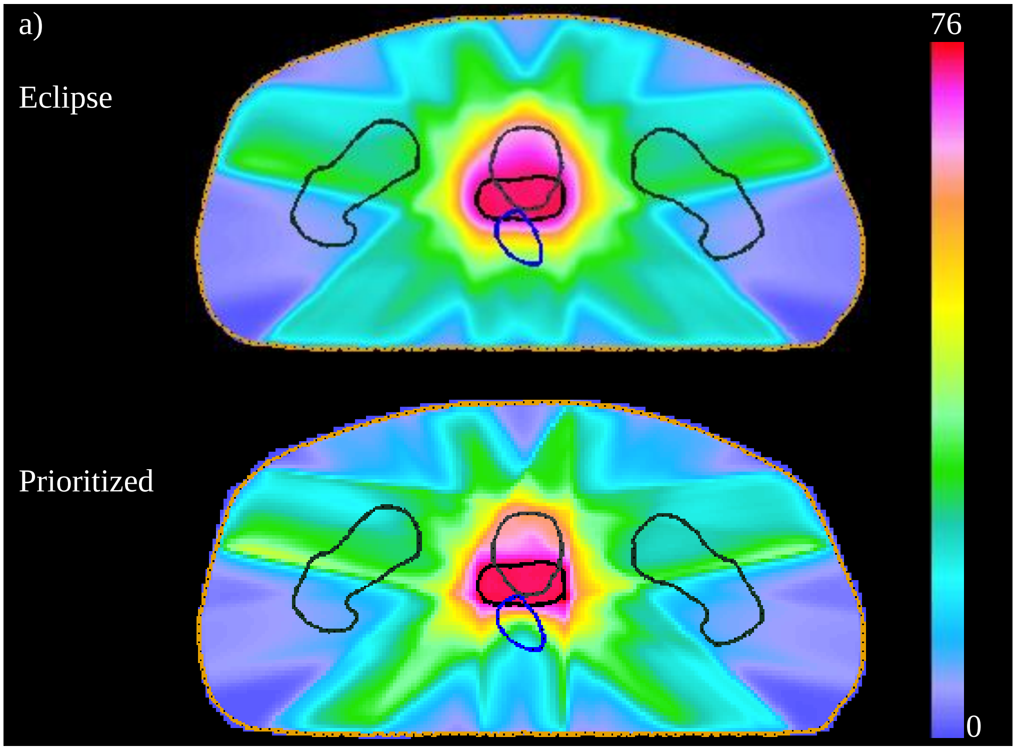


Figure 4-30 Dose distribution of prostate treatment plan 38. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



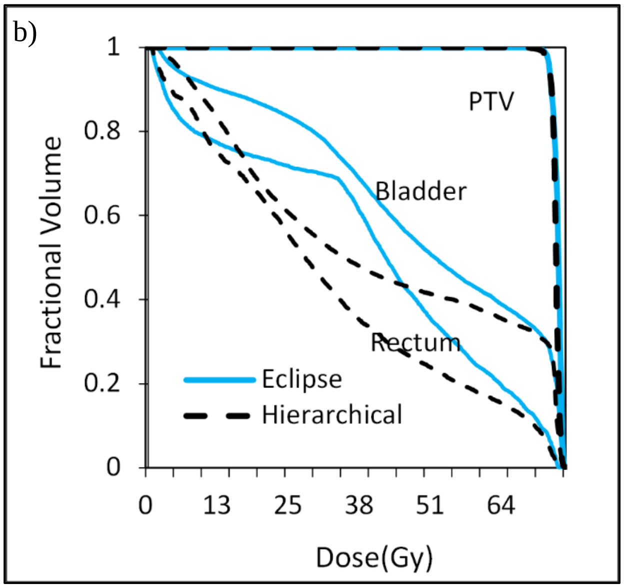
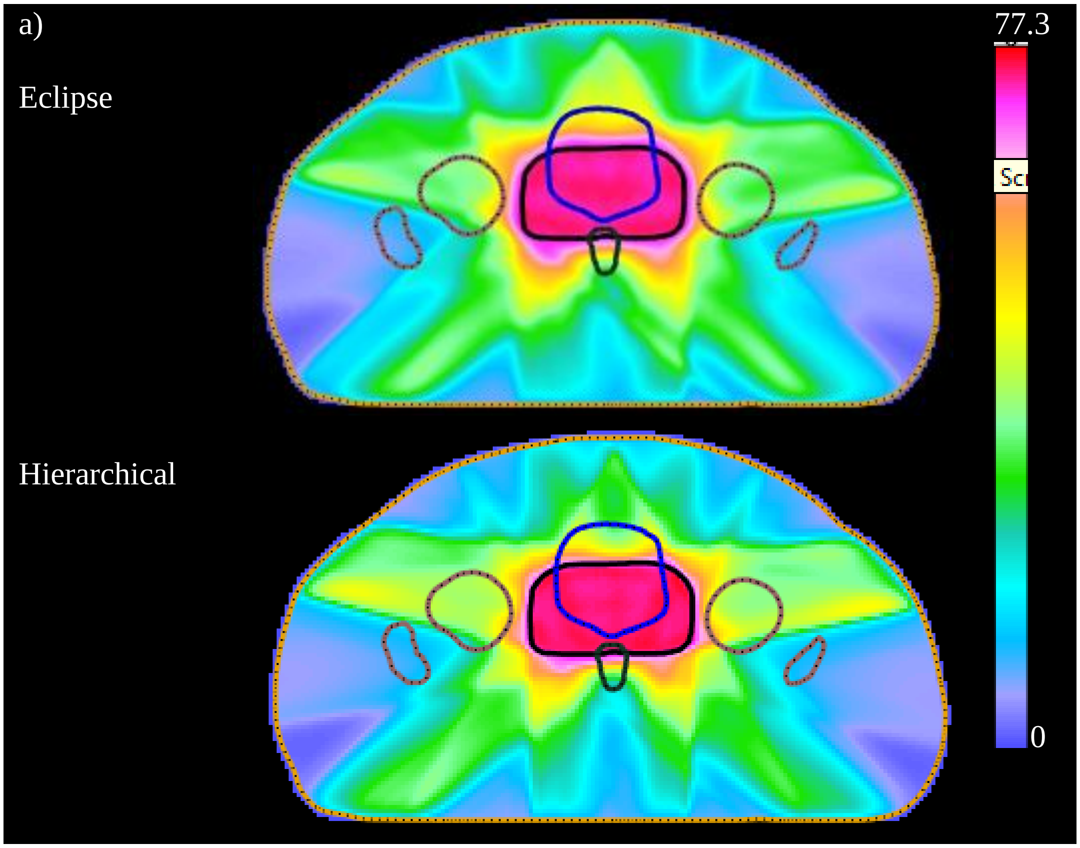


Figure 4-31 Dose distribution of prostate treatment plan 39. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.



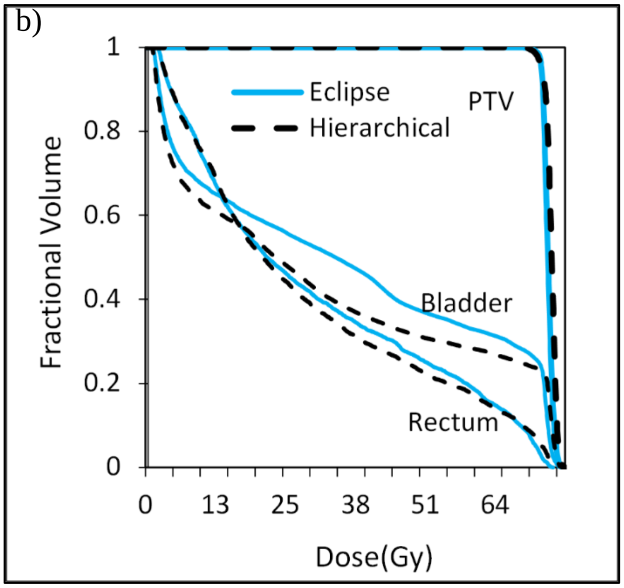


Figure 4-32 Dose distribution of prostate treatment plan 40. a) Dose color-wash of the dose distribution obtained from hierarchical and Eclipse treatment planning systems. b) Dose-volume histogram of the structures in the optimization.

Table 4-9 D99 and D05 of the target structure from the hierarchical and Eclipse treatment planning systems in percentage of the prescription dose. The prescription dose was 72Gy. The average ratio of D99 and D05 was approximately 1, suggesting that both systems produced almost the same target coverage.

Plan Number	Target D99 of Hierarchical Plans in % (a)	Target D99 of Eclipse Plans in % (b)	Ratio a/b	Target D05 of Hierarchical Plans in % (c)	Target D05 of Eclipse Plans in % (d)	Ratio c/d
12	97.59	98.13	0.99	103.81	106.5	0.97
13	97.51	96.77	1	104.7	104.4	1
15	97.83	97.83	1	103.6	103.45	1
16	98.16	98.16	1	103.35	105.28	0.98
18	97.9	86.78	1.12	104.22	111.81	0.93
19	98.46	97.16	1.01	102.96	106.78	0.96
20	97.68	97.68	1	104.13	103.84	1
21	97.3	97.3	1	103.71	103.84	0.99
22	97.9	98.75	0.99	103.46	103.36	1
24	97.27	97.27	1	103.76	106.31	0.97
26	97.66	97.66	1	103.82	104	0.99
27	97.42	88.34	1.10	104.68	107.24	0.97
28	98.34	98.34	1	103.08	104.44	0.98
29	98.27	97.09	1.01	103.31	104.29	0.99
30	97.55	99.18	0.98	104.05	105.51	0.98
31	98.42	97.98	1	102.79	103.52	0.99
32	97.51	99	0.98	104.15	102.81	1.01
33	97.54	98.27	0.99	103.9	103.9	1
34	98.38	98.38	1	103.33	104.83	0.98
35	97.41	98.99	0.98	104.51	102.72	1.01
36	98.2	95.03	1.03	103.33	103.66	0.99
37	98.61	98.83	0.99	102.97	102.97	1
38	97.96	97.96	1	104.08	104.67	0.99
39	98.12	99	0.99	102.99	103.23	0.99
40	97.65	98.86	0.98	104.38	103.58	1
Average			1.01			0.99

Table 4-10 Maximum and minimum dose to the target structure of hierarchical and Eclipse treatment plans in percentage of the prescription dose. The prescription dose was 72Gy. The average ratio of the maximum dose was 1, suggesting that both systems produced almost the same maximum dose. The hierarchical plans have a higher minimum dose to target structure than that of the Eclipse plan. The higher minimum dose implies that the hierarchical plan didn't produce the under dose region in the target structures.

Plan Number	Target Maximum Dose of Hierarchical Plans in % (a)	Target Maximum Dose of Eclipse Plans in % (b)	Ratio a/b	Target Minimum Dose of Hierarchical Plans in % (c)	Target Minimum Dose of Eclipse Plans in % (d)	Ratio c/d
12	105.7	108.7	0.97	85	91.9	0.92
13	107.4	107.1	1	86.6	81.8	1.05
15	106.4	105.3	1.01	89.9	78.9	1.13
16	104.9	106.7	0.98	90.8	90	1
18	105.8	113.6	0.93	86.2	61.9	1.39
19	105.2	109.3	0.96	93	70.5	1.31
20	106	105.3	1	89.1	79	1.12
21	107.1	105.3	1.01	80.8	85.2	0.94
22	108.7	104.3	1.04	86.9	94.4	0.92
24	106.5	108.4	0.98	89.5	76.6	1.16
26	105.9	105.9	1	89.6	86.8	1.03
27	106.7	109	0.97	84.4	80.7	1.04
28	105.2	106	0.99	86.6	85.2	1.01
29	107.3	106.3	1	92	84.8	1.08
30	107	107.8	0.99	87.6	96.8	0.90
31	104	104.8	0.99	92.3	93.3	0.98
32	106.9	104.6	1.02	90.5	93.1	0.97
33	106.1	108.2	0.98	90	81.8	1.10
34	105.7	106.2	0.99	89.2	93	0.95
35	107.3	103.4	1.03	89.4	94	0.95
36	106.5	105.5	1	91.5	79.4	1.15
37	104.8	104.1	1	93.2	93.6	0.99
38	108	106.2	1.01	91.2	90.4	1
39	105.8	103.9	1.01	91.5	95	0.96

40	107.1	106	1.01	89.3	92.8	0.96
Average			0.99			1.05

Table 4-11 Rectal volume that receives at least 47Gy (V47) of the hierarchical and Eclipse treatment plans. The majority of hierarchical treatment plans have lower V47 than that of the Eclipse plans.

Plan Number	Rectum V47 of hierarchical plans in % (a)	Rectum V47 of Eclipse plans in %(b)	Ratio a/b
12	47.46	48.95	0.96
13	37.26	53.05	0.70
15	41.74	45.88	0.90
16	51.56	50.67	1.01
18	47.29	55.16	0.85
19	51.56	50.62	1.01
20	47.64	49.8	0.95
21	25.5	36.2	0.70
22	31.1	44.24	0.70
24	44.32	48.52	0.91
26	41.76	41.2	1.01
27	42.56	50.45	0.84
28	39.16	48	0.81
29	54.46	50.63	1.07
30	55.74	70.06	0.79
31	49.48	46.35	1.06
32	41.91	50.76	0.82
33	45.54	50.6	0.9
34	44.25	47.08	0.93
35	30.75	48.2	0.63
36	54.69	51.87	1.05
37	31.05	42.8	0.72
38	35.26	42.77	0.82
39	25.91	47.91	0.54
40	34.2	38.37	0.89
Average			0.87

Table 4-12 Maximum and minimum dose to the rectum structure of the hierarchical and Eclipse treatment plans in percentage of the prescription dose. The prescription dose was 72Gy. The average ratio of the maximum dose was 1, suggesting that both systems produced almost the same maximum dose. The hierarchical plans have a lower minimum dose than that of the Eclipse plans.

Plan Number	Rectum Maximum Dose of Hierarchical Plans in % (a)	Rectum Maximum Dose of Eclipse Plans in % (b)	Ratio a/b	Rectum Mean Dose of Hierarchical Plans in % (c)	Rectum Mean Dose of Eclipse Plans in % (d)	Ratio c/d
12	104.5	102.9	1.01	62.1	65	0.95
13	104.6	102.6	1.01	54.8	69.2	0.79
15	104.5	102.5	1.01	59.8	62.8	0.95
16	104	102.2	1.01	66.4	65.8	1
18	104.3	108.2	0.96	62.4	72.1	0.86
19	103.9	103.8	1	67.5	65.1	1.03
20	107.8	102.7	1.04	64.5	66.4	0.97
21	104.1	103.3	1	45.9	59.8	0.76
22	105.2	103.1	1.02	50.8	62.6	0.81
24	104.9	103.3	1.01	59.7	61.9	0.96
26	104.1	102	1.02	59.9	59.1	1.01
27	104.5	104.4	1	59.8	68.1	0.87
28	104.1	102.9	1	61.8	70.1	0.88
29	107.3	103.4	1.03	68.6	65.9	1.04
30	104.8	104.4	1	69.4	79.9	0.86
31	103.9	103	1	65.2	64.3	1.01
32	105.3	102.2	1.03	56.5	71.4	0.79
33	107.8	102.9	1.04	58.8	69.9	0.84
34	105.7	103	1.02	61	64	0.95
35	105.3	102	1.03	49.9	64.3	0.77
36	105.8	103.4	1.02	66.5	68	0.97
37	104.2	103	1.01	50.2	60.5	0.82
38	105.7	103	1.02	53.1	64.7	0.82
39	104	102.6	1.01	52	67.6	0.76
40	104	103	1	52.8	53.6	0.98
Average			1.01			0.90

Table 4-13 The volume that receives at least 47Gy (V47) of the bladder of the hierarchical and Eclipse treatment plans. The majority of hierarchical treatment plans have lower V47 than that of the Eclipse plans.

Plan Number	Bladder V47 of hierarchical plans in % (a)	Bladder V47 of Eclipse plans in % (a)	Ratio a/b
12	50	51.03	0.97
13	61.31	60.96	1
15	27.39	32.07	0.85
16	38.54	40.62	0.94
18	59.34	65.31	0.90
19	41.43	49.25	0.84
20	37.51	41.35	0.90
21	36.67	41.18	0.89
22	32.15	36.47	0.88
24	38.36	41.1	0.93
26	47.07	52.38	0.89
27	57.74	55.72	1.03
28	62.21	61.58	1.01
29	32.36	36.75	0.88
30	27.33	34.18	0.79
31	32.34	38.88	0.83
32	52.09	52.76	0.98
33	36.34	51.81	0.70
34	18.86	18.86	1
35	31.73	37.45	0.84
36	35.91	38.72	0.92
37	20.31	23.03	0.88
38	43.33	46.47	0.93
39	45.29	48.49	0.93
40	36.15	39.88	0.90
Average			0.91

Table 4-14 Maximum and minimum dose to the rectum structure of the hierarchical and Eclipse treatment plans in percentage of the prescription dose. The prescription dose was 72Gy. The average ratio of the maximum dose was 1, suggesting that both systems produced almost the same maximum dose. Hierarchical plans have a lower minimum dose than that of Eclipse plans.

Plan Number	Bladder Maximum Dose of Hierarchical Plans in % (a)	Bladder Maximum Dose of Eclipse Plans in % (b)	Ratio a/b	Bladder Mean Dose of Hierarchical Plans in % (c)	Bladder Eclipse Mean Dose in % (d)	Ratio c/d
12	105.6	106.1	0.99	58.5	60.1	0.97
13	106.6	105.7	1	74.5	77.2	0.96
15	104.5	105.3	0.99	39.9	48.3	0.82
16	104.3	104.8	0.99	54.9	57.8	0.94
18	105.5	112	0.94	74.3	82.5	0.90
19	104.6	108.3	0.96	56.3	67.3	0.83
20	106	105.3	1	45.4	48.3	0.93
21	104.2	104.3	0.99	50.1	57.2	0.87
22	104.5	104	1	39.9	44.6	0.89
24	105.1	107.1	0.98	47.2	50	0.94
26	105.1	104	1.01	58.9	61.8	0.95
27	106.3	107.6	0.98	71.8	70.7	1.01
28	103.7	105.1	0.98	76.3	79.3	0.96
29	104.4	104.8	0.99	41.6	47	0.88
30	105	106.6	0.98	34.8	40.5	0.85
31	104	104.2	0.99	43.3	47.6	0.90
32	105.1	104	1.01	61.9	64	0.96
33	104.4	105.1	0.99	48.1	58.2	0.82
34	104.7	105.8	0.98	27.9	28.3	0.98
35	107.1	103.2	1.03	42.6	47.1	0.90
36	106.5	105.1	1.01	43.8	46.8	0.93
37	102.8	103.8	0.99	25.3	27.6	0.91
38	106.4	104.4	1.01	53	55.9	0.94
39	103.5	103.6	0.99	57.7	64.5	0.89
40	105.6	104.7	1.0	46.2	49.3	0.93

Table 4-15 Maximum and average femur dose of the hierarchical and Eclipse treatment plans in percentage of the prescription dose. The prescription dose was 72Gy. The average ratio of the maximum dose was 1, suggesting that both systems produced almost the same maximum dose. Hierarchical plans have a higher minimum dose to femur than that of Eclipse plans.

Plan Number	Femur Maximum Dose of Hierarchical Plans in % (a)	Femur Maximum Dose of Eclipse Plans in % (b)	Ratio a/b	Femur Mean Dose of Hierarchical Plans in % (c)	Femur Mean Dose of Eclipse Plans in % (d)	Ratio c/d
12	65.6	75.7	0.86	15	14.8	1.01
13	82.4	80.2	1.02	27.7	26.4	1.04
15	62.7	55.5	1.12	19.3	16.5	1.16
16	64.6	66.4	0.97	15	13.1	1.14
18	72.7	83.5	0.87	19.1	17	1.12
19	81.4	67.8	1.20	25.1	22.4	1.12
20	81.7	70.6	1.15	27	24.6	1.09
21	89.4	62.7	1.42	34.1	20.4	1.67
22	76.4	71.4	1.07	20.3	18.2	1.11
24	75.8	77	0.98	24.3	22.7	1.07
26	68.5	72.2	0.94	16.9	16.6	1.01
27	84.4	87.6	0.96	28.3	23.2	1.21
28	85.6	85.1	1.0	19.7	18.2	1.08
29	56.7	58.9	0.96	16.9	15	1.12
30	94.2	90.7	1.03	24.2	21.1	1.14
31	63.8	62.8	1.01	13.2	12.7	1.03
32	75.4	75.2	1.0	16.7	15.1	1.10
33	75.6	72.5	1.04	22.4	18.4	1.21
34	56.6	60.6	0.93	14	14.2	0.98
35	66.1	67.9	0.97	20	19.4	1.03
36	63.9	71.8	0.88	23.2	22.5	1.03
37	51.9	50.1	1.03	12.9	11.6	1.11
38	70.2	61	1.15	24.5	19.7	1.24
39	58.1	41.8	1.38	20.8	17.7	1.17

40	83.2	78.7	1.05	28.1	26.4	1.06
Average			1.04			1.12

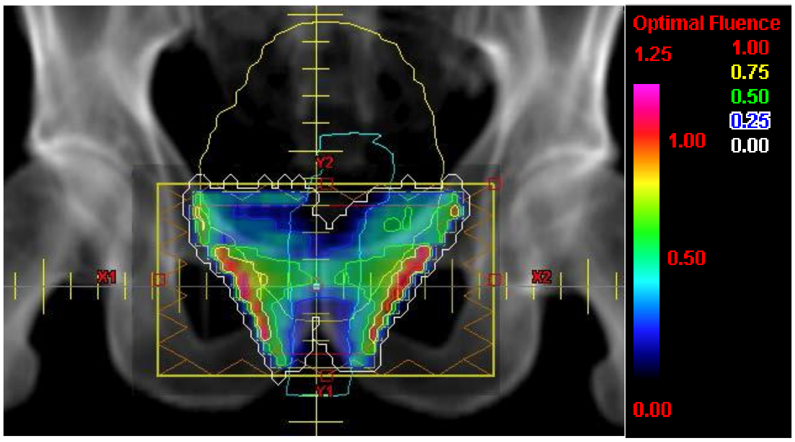
4.5.1 Deliverability of the Treatment plans

The linear accelerator can’t deliver the beam profile with a large monitor unit (MU). We found that the smaller value of the Laplacian function corresponded to the smooth beam profile and lower MU. Therefore, we minimized the Laplacian function in the fifth step of optimization to smooth the beam profile (Figure 4-33 and Figure 4-35). The hierarchical method produced a smaller MU in some treatment plans and a larger one in others. Table 4-16 shows the ratio between the MU of the hierarchical and Eclipse systems. The average MU ratio was 1.27, suggesting that the MU of the hierarchical system was larger than that of the Eclipse system. Although MU of the hierarchical system was larger than that of the Eclipse system, it was still within a reasonable range. Therefore, the hierarchical system did not substantially compromise the deliverability of the treatment plans.

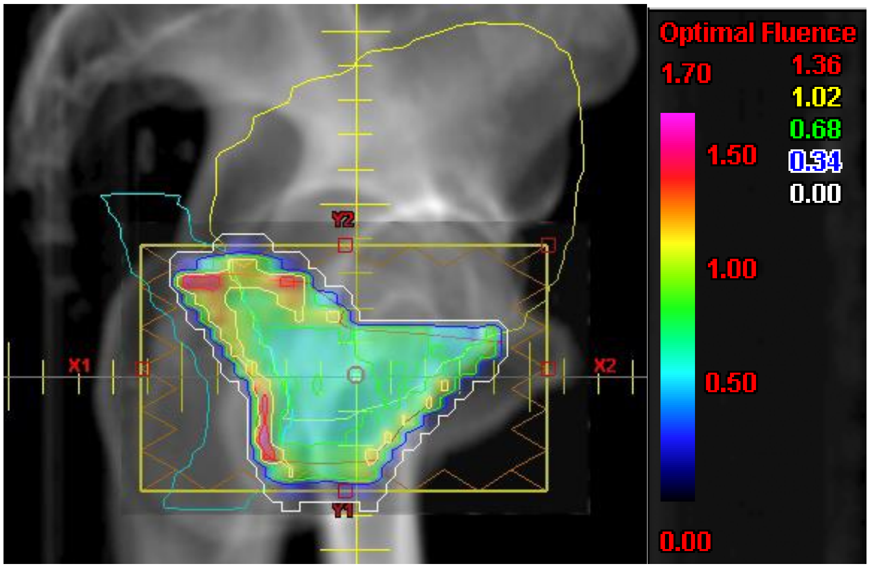
Table 4-16 Monitor unit of the beam profile of treatment plans prepared using the hierarchical and Eclipse treatment planning systems.

Plan Number	Monitor Unit of Hierarchical Plans (a)	Monitor Unit of Eclipse Plans(b)	Ratio a/b
12	1189	1121	1.06
13	1307	1063	1.23
15	1176	958	1.23
16	1211	1002	1.21
18	1247	1290	0.97
19	1075	1101	0.98
20	1716	1235	1.39
21	1516	869	1.74
22	1123	823	1.36
24	1257	1394	0.90
26	1145	875	1.31

27	1278	1199	1.07
28	1587	1575	1.01
29	1181	1022	1.16
30	1457	941	1.55
31	1102	974	1.13
32	1142	919	1.24
33	1244	916	1.36
34	1132	938	1.21
35	1206	793	1.52
36	1131	861	1.31
37	1160	755	1.54
38	1210	805	1.50
39	1333	693	1.92
40	1265	979	1.29
Average			1.28

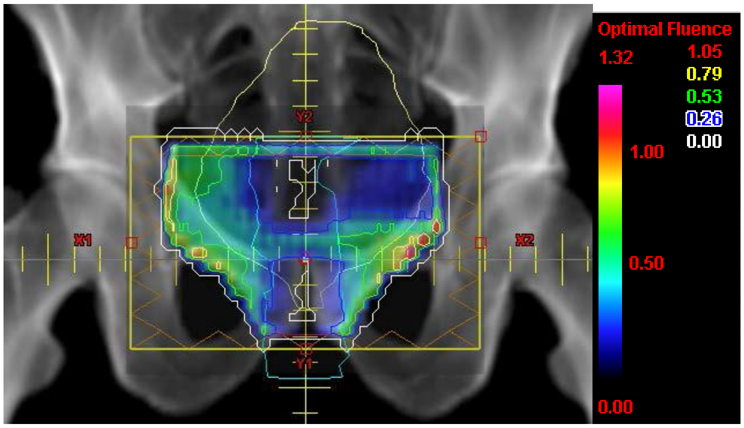


a) Beam profile of posterior-anterior beam

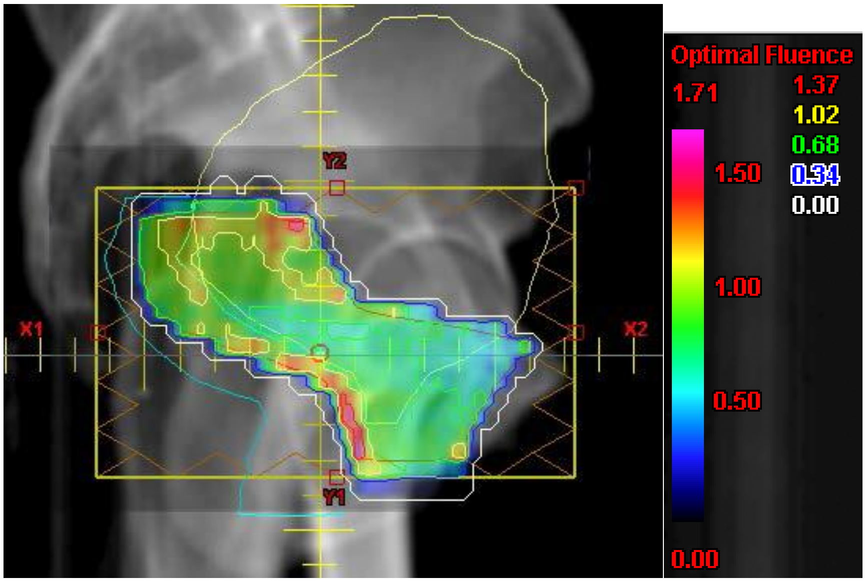


b) Beam profile of right anterior oblique beam

Figure 4-33 Beam profile of treatment plan 5 prepared using the hierarchical system.

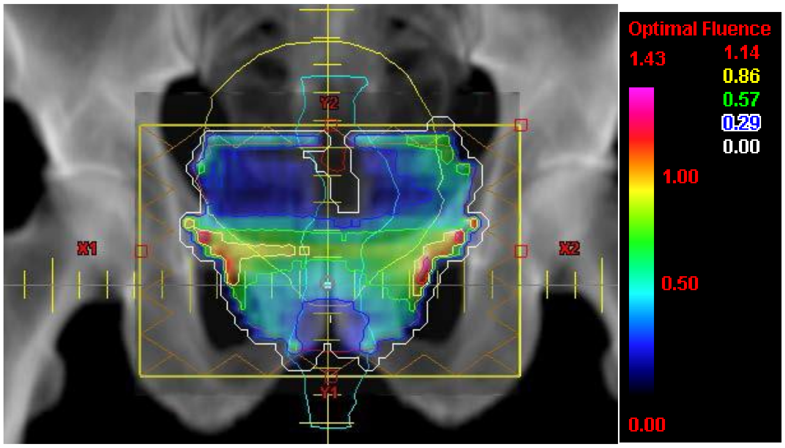


a) Beam profile of posterior-anterior beam

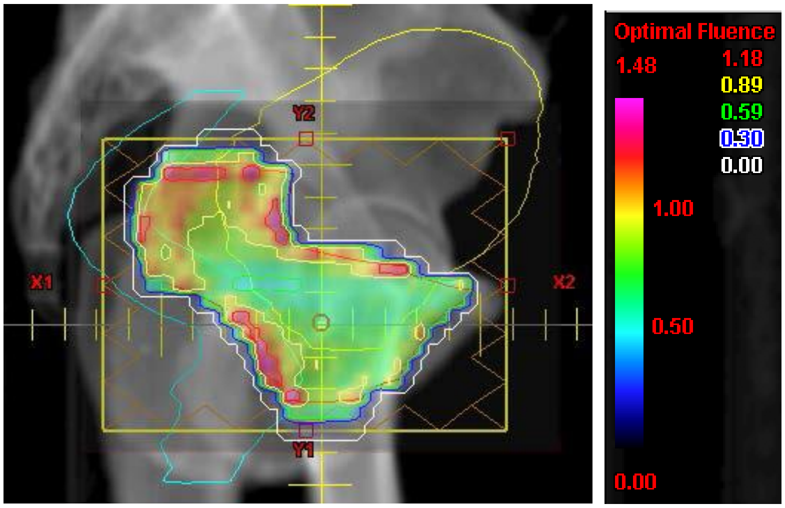


b) Beam profile of right anterior oblique beam

Figure 4-34 Beam profile of treatment plan 6 prepared using the hierarchical system.



a) Beam profile of posterior-anterior beam



b) Beam profile of right anterior oblique beam

Figure 4-35 Beam profile of treatment plan 8 prepared using the hierarchical system.

4.6 Discussion and Conclusion

Our findings demonstrate that treatment plans could be automatically produced without trial-and-error iterations. We compared the treatment plans of 31 prostate patients prepared using the hierarchal system with that prepared by the planner using the Eclipse treatment planning system. The result showed that the hierarchical system produced better treatment plans in the majority of cases. The hierarchical system reduced the mean dose to the bladder and rectum by 10%, on average, while the dose to 95% of the target structure was the same from both systems.

Similarly, the dose to target was homogenous from the hierarchical system than from that of the Eclipse system in the majority of the treatment plans.

Different methods have been proposed to automate treatment planning. One such method is maintaining the database of the previously treated patients and selecting a patient from the database to prepare the plan of a new patient [111]. This method requires maintenance of a large database and the ability to find a previous patient that is closely related to the new patient within the database. Moreover, the quality of the new treatment plan depends on the quality of the old plan. By contrast, the hierarchical method does not depend on the old treatment plan's quality. It formalizes the intuitive notion of priority that the physician prescribed while creating treatment plans. Therefore, the hierarchical method could produce the best, mathematically guaranteed treatment plans with a fixed structure priority.

Another approach to automate the treatment planning process is exploring the Pareto surface to find the best Pareto optimal solution [74], [112]–[114]. The method is attractive because the planner can navigate different optimal solutions and evaluate the trade-off in dose distribution to structures. However, the computation of the whole Pareto surface is not feasible and the algorithm may not display all possibilities. On the other hand, the hierarchical method improves on each goal as much as possible. Therefore, the hierarchical method could give better visibility to the trade-off in dose distribution of different structures than the Pareto navigation method.

Our observations of dose distribution in the hierarchical system are in accordance with the observations of other research groups. We observed that the hierarchical system produced a lower dose in the rectum and bladder and a higher dose in the femur. The rectum and bladder had higher priority in optimization than the femur. Therefore, the lower priority goal was sacrificed to improve on the higher priority goals. The finding was in accordance with Breedveld, et al. [69] who also needed to relax lower priority goals to improve on higher priority goals.

One concern of the hierarchical method is that higher priority goals might dominate lower priority goals and there would be little room to improve on lower priority goals[43], [51]. The hierarchical method uses the "slip" parameter from the second through fifth steps of optimization. The "slip" parameter increases the size of the feasibility space and thus provides room for improvement in lower priority goals. The "slip" relaxes on higher priority goals while optimizing lower priority goals. In our optimization, we used the same value of "slip" for all prostate plans. This is in contrast to the lexicographic method proposed by Jee, et al. [73] where they relax higher priority goals, in a different extent, while optimizing lower priority goals. Future work in hierarchical optimization could focus on studying the behavior of the optimal solution with larger "slip" on the higher priority goal.

In summary, the hierarchical method could be used to automate the treatment planning process in a clinical setting. The treatment plans produced by the hierarchical method were better than those produced by the planner in the majority of cases. The automatically produced treatment plans have uniform dose in the target structure and lower dose in the rectum and bladder while having the same target coverage. The automation of treatment planning has the potential to change the quality of radiation therapy treatment planning. The method could explore the solution space that a planner cannot explore using the trial-and-error method and thus could produce better treatment plans. Moreover, the method could open the door for adaptive radiation therapy, which requires automatic re-planning to incorporate changes inpatient geometry while treating the patient.

Chapter 5: Conclusion

This dissertation investigates hierarchical optimization to automate the Intensity Modulated Radiation Therapy (IMRT) treatment planning process. The state-of-the-art treatment planning application requires trial-and-error iterations to produce clinically acceptable treatment plans. The planner cannot produce the best dose distribution by using the trial-and-error iterations because there is no way to determine the exact sequence of iterations that will lead to the best dose distribution. Hierarchical optimization has the potential to shift the treatment planning paradigm from a trial-and-error manual process to an ideal automated process. The hierarchical-based treatment planning application could probe a larger search space than that probed by planners using trial-and-error iterations. Therefore, the hierarchical system produces better treatment plans than those produced by planners using the trial and error method. Better treatment plans will ultimately improve the quality of radiation therapy treatment.

Hierarchical optimization uses large-scale, nonlinear, constrained optimization formulation to automate the treatment planning process. This type of large-scale, nonlinear optimization is computationally intensive and slower than the traditional IMRT optimization problem. Moreover, the mathematical model proposed in previous studies has not been used to prepare treatment plans in a clinical setting. Therefore, this dissertation addresses three issues of hierarchical optimization: 1) voxel sampling to reduce optimization time was explored, 2) the open source optimization solver was customized to solve the large-scale, nonlinear, constrained hierarchical optimization problem, 3) the mathematical model was formulated to prepare treatment plans automatically.

Hierarchical optimization has thousands of variables and millions of constraints in it. The optimization solver takes substantial time to solve a large-scale, constrained optimization problem. Therefore, we proposed a voxel sampling algorithm to reduce optimization time. The algorithm categorizes the voxel of a structure into the boundary and interior voxels. A voxel is called a boundary voxel if there is no voxel in at least one of four directions. Similarly, a voxel is called an interior voxel if voxels surround it in all directions. Dose to boundary voxels are difficult to optimize because the dose starts to change from the boundary voxel of a structure. Therefore, the sampling algorithm includes all boundary voxels of a structure and samples the interior voxels using the K-mean clustering algorithm. Hierarchical optimization was run on the sampled voxels to prepare the treatment plans. We found that the sampling sped up the optimization by 2-3 times without sacrificing the quality of the treatment plans. The average dosimetric error occurred when the sampling was less than 1%. Therefore, the boundary and interior voxel sampling is an attractive method to reduce optimization time without compromising the quality of the dose distribution.

The previous study solved the hierarchical optimization formulation using the Mosek solver. Mosek solver has two problems: 1) it cannot solve the nonlinear, constrained optimization problem, 2) it takes substantial time to solve the hierarchical optimization problem. The Mosek solver cannot be customized because it is a commercial solver and doesn’t provide access to the source code. Therefore, we customized an open source Interior Point Optimization solver to solve the large-scale, constrained optimization problem.

The Interior Point Optimization Solver (IPOPT) is an open source solver that can solve the linear and nonlinear, convex, and non-convex, large-scale optimization problem. The solver is written in C++ and uses the Interior Point Algorithm to solve the constrained optimization problem. The hierarchical optimization module was developed in C++ and uses the IPOPT functions to solve the optimization problem. The optimization module is integrated with CERR (Computational Environment for Radiotherapy Research), a widely used software platform for developing and sharing radiation therapy treatment planning research, and thus provides the complete suite of the treatment planning application.

The hierarchical optimization module was sped up by optimizing the memory access of the large sparse matrix. The CERR application precomputes the dose to each structure and stores it in the sparse matrix called the "influence matrix." The hierarchical optimization repeatedly accesses the influence matrix to compute the dose for each structure. We studied the memory access pattern of the influence matrix and modified the code to reduce memory latency in accessing the influence matrix. We solved the hierarchical optimization formulation proposed by Clark, et al.[88] by using the IPOPT solver and compared the run time and dose quality of the IPOPT and Mosek-based applications. The dose distribution of the treatment plans were exactly the same, but the IPOPT-based application was almost 15-65 times faster than the Mosek-based application.

It’s essential to prepare treatment plans in a clinical setting using hierarchical optimization to demonstrate its viability. Therefore, we trained the hierarchical optimization model in six prostate plans to formulate the optimization model and automatically prepared 25 prostate plans. The treatment plans were prepared using the same clinical criteria used by the planner to prepare treatment plans at Memorial Sloan Kettering Cancer Center. We found that the nonlinear, constrained optimization model is needed to prepare treatment plans automatically. The dose distribution of the majority of automatically prepared plans was better than that prepared by a planner using the Eclipse treatment planning system.

Future studies of hierarchical optimization could focus on preparing the treatment plans of head and neck, lung and other types of cancer. It’s essential to demonstrate that hierarchical optimization is broad enough to produce automatic treatment plans for different types of cancer. The formulation of prostate plans could be extended to demonstrate that treatment plans can be automatically generated for different types of cancer.

The hierarchical optimization problem is computationally intensive because of the thousands of variables and millions of constraints in it. The computational time of hierarchical optimization could be reduced by using parallel computation. The optimization performs the substantial matrix operations in each iteration to find the optimal solution. The matrix operations could be performed in graphics processing units (GPUs) to reduce hierarchical optimization time.

Uncertainty arises in treatment planning due to the motion of structures and differences in patient setup during multiple treatments of the patient. Effectiveness of treatment plans is diminished due to these uncertainties. Treatment planning should incorporate uncertainties into the optimization to produce robust treatment plans. The robust plan incorporates uncertainties during optimization and therefore is less sensitive to the changes that come during treatment plan delivery. Hierarchical optimization can easily be extended to incorporate robustness in treatment planning. A new step can be added to handle the robustness in hierarchical optimization. The ability to prepare a robust, automatic plan will improve the quality of radiation therapy treatment plans.

In conclusion, this dissertation demonstrated that treatment plans could be prepared automatically without manual intervention. Automatic treatment planning shifts the paradigm of radiation therapy from manual trial and error to an ideal automated process. The elimination of trial and error iterations will improve patient care quality of because treatment planning will no longer depend on planner expertise. The automated system will open the door for adaptive radiation therapy as treatment plans can be automatically re-planned in real time to adapt to changes in patient geometry.

References

[1] A. Wächter and L. T. Biegler, “On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming,” *Math. Program.*, vol. 106, no. 1, pp. 25–57, Mar. 2006.

[2] “The First X-ray, 1895 | The Scientist Magazine®,” *The Scientist*. [Online]. Available: http://www.the-scientist.com/?articles.view/articleNo/30693/title/The-First-X-ray--1895/. [Accessed: 01-Sep-2015].

[3] B. H. Orndoff, “Emile Herman Grubbe,” *Radiology*, vol. 75, no. 3, pp. 473–474, Sep. 1960.

[4] S. Forshier, *Essentials of Radiation, Biology and Protection*. Cengage Learning, 2008.

[5] O. of the H. Secretary and N. A. of Sciences, *Biographical Memoirs*. National Academies Press, 1982.

[6] U. Linz, Ed., *Ion Beam Therapy*, vol. 320. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012.

[7] D. I. Thwaites and J. B. Tuohy, “Back to the future: the history and development of the clinical linear accelerator,” *Phys. Med. Biol.*, vol. 51, no. 13, p. R343, Jul. 2006.

[8] D. Greene and P. C. Williams, *Linear Accelerators for Radiation Therapy, Second Edition*. CRC Press, 1997.

[9] L. K. Mell, J. C. Roeske, and A. J. Mundt, “A survey of intensity-modulated radiation therapy use in the United States,” *Cancer*, vol. 98, no. 1, pp. 204–211, Jul. 2003.

[10] T. Bortfeld, “IMRT: a review and preview,” *Phys. Med. Biol.*, vol. 51, no. 13, p. R363, Jul. 2006.

[11] G. Bauman, R. B. Rumble, J. Chen, A. Loblaw, and P. Warde, “Intensity-modulated Radiotherapy in the Treatment of Prostate Cancer,” *Clin. Oncol.*, vol. 24, no. 7, pp. 461–473, Sep. 2012.

[12] J. G. Mechalakos, J. St. Germain, and C. M. Burman, “Results of a one year survey of output for linear accelerators using IMRT and non-IMRT techniques,” *J. Appl. Clin. Med. Phys.*, vol. 5, no. 1, May 2004.

[13] W. P. M. Mayles, “Survey of the Availability and Use of Advanced Radiotherapy Technology in the UK,” *Clin. Oncol.*, vol. 22, no. 8, pp. 636–642, Oct. 2010.

[14] T. Frenzel and A. Krüll, “The use of IMRT in Germany,” *Strahlenther. Onkol.*, Mar. 2015.

[15] A. Brahme, “Multi leaf collimator,” US4672212 A, 09-Jun-1987.

[16] *Clinac® iX*.

[17] W. Schlegel and A. Mahr, *3D Conformal Radiation Therapy: Multimedia Introduction to Methods and Techniques*, 2nd ed. Springer Publishing Company, Incorporated, 2007.

[18] K. Doi, “Diagnostic imaging over the last 50 years: research and development in medical imaging science and technology,” *Phys. Med. Biol.*, vol. 51, no. 13, p. R5, Jul. 2006.

[19] L. Rothenberg, R. Mohan, J. V. Dyk, B. Fraass, and T. Bortfeld, “WE-G-16A-01: Evolution of Radiation Treatment Planning,” *Med. Phys.*, vol. 41, no. 6, pp. 515–515, Jun. 2014.

[20] C. J. Karzmark, “Advances in linear accelerator design for radiotherapy,” *Med. Phys.*, vol. 11, no. 2, pp. 105–128, Apr. 1984.

[21] F. A. M. Jr, P. W. Wiest, J. A. Locken, and C. AKelsey, “CT scanning: patterns of use and dose,” *J. Radiol. Prot.*, vol. 20, no. 4, p. 353, Dec. 2000.

[22] W. F. McGuirt, K. Greven, D. Williams, J. W. Keyes, N. Watson, J. O. Cappellari, and K. R. Geisinger, “Pet scanning in head and neck oncology: A review,” *Head Neck*, vol. 20, no. 3, pp. 208–215, May 1998.

[23] P. Bratslavets, A. Kalniņš, I. Pļaviņa, A. Popov, B. Rapoport, A. Tāle, and B. Žeigurs, “The Advancement of Imaging Science and Technology,” *Int Acad. Publ*, p. 474, 1990.

[24] S. Voros, “What are the potential advantages and disadvantages of volumetric CT scanning?,” *J. Cardiovasc. Comput. Tomogr.*, vol. 3, no. 2, pp. 67–70, Mar. 2009.

[25] B. Huang, M. W.-M. Law, and P.-L. Khong, “Whole-Body PET/CT Scanning: Estimation of Radiation Dose and Cancer Risk,” *Radiology*, vol. 251, no. 1, pp. 166–174, Apr. 2009.

[26] J. Sievinen, W. Ulmer, and W. Kaissl, “AAA photon dose calculation model in Eclipse.”

[27] “Monaco® - Radiobiological Planning for IMRT, VMAT and SBRT | Elekta,” *Elekta AB*. [Online]. Available: http://www.elekta.com/healthcare-professionals/products/elekta-software/treatment-planning-software/planning-software/monaco.html?utm_source=monaco5&utm_medium=redirect&utm_campaign=redirects. [Accessed: 01-Sep-2015].

[28] “RayStation Version 4.0 Treatment Planning System Released,” *Imaging Technology News*. [Online]. Available: http://www.itnonline.com/content/raystation-version-40-treatment-planning-system-released. [Accessed: 01-Sep-2015].

[29] J. E. M Kowski, “SU-E-T-505: BrainLab Plan Comparisons: Brain Scan Pencil Beam versus IPlan Monte Carlo.,” *Med. Phys.*, vol. 39, no. 6, p. 3821, 2012.

[30] D. Shepard, M. Ferris, G. Olivera, and T. Mackie, “Optimizing the Delivery of Radiation Therapy to Cancer Patients,” *SIAM Rev.*, vol. 41, no. 4, pp. 721–744, Jan. 1999.

[31] T. Bortfeld, “Current IMRT Optimization Algorithms:Principles,Potential and Limitations,” Massachusetts General Hospital, Northeast Proton Therapy Center,Boston.

[32] A. J. Mundt and J. C. Roeske, *Intensity Modulated Radiation Therapy: A Clinical Perspective*. PMPH-USA, 2005.

[33] A. B. Wolbarst, E. S. Sternick, B. H. Curran, R. J. Kosinski, and A. Dritschilo, “A FORTRAN program for the optimization of radiotherapy treatment planning using the complication probability factor (CPF),” *Comput. Programs Biomed.*, vol. 11, no. 2, pp. 99–104, Apr. 1980.

[34] B. E. Bjärngard, “Optimization in radiation therapy,” *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 2, no. 3–4, pp. 381–382, Apr. 1977.

[35] K. Deb and K. Deb, “Multi-objective Optimization,” in *Search Methodologies*, E. K. Burke and G. Kendall, Eds. Boston, MA: Springer US, 2014, pp. 403–449.

[36] C. C. Coello, G. B. Lamont, and D. A. van Veldhuizen, *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer Science & Business Media, 2007.

[37] D. A. V. Veldhuizen and G. B. Lamont, “Multiobjective Evolutionary Algorithms: Analyzing the State-of-the-Art,” *Evol. Comput.*, vol. 8, no. 2, pp. 125–147, Jun. 2000.

[38] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. John Wiley & Sons, 2001.

[39] D. W. Corne, N. R. Jerram, J. D. Knowles, M. J. Oates, and M. J, “PESA-II: Region-based Selection in Evolutionary Multiobjective Optimization,” in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO’2001*, 2001, pp. 283–290.

[40] J. G. Lin, “Multiple-objective problems: Pareto-optimal solutions by method of proper equality constraints,” *IEEE Trans. Autom. Control*, vol. 21, no. 5, pp. 641–650, Oct. 1976.

[41] M. Vukov, R. Quirynen, and K. Leuven, *A Toolkit for Automatic Control and Dynamic Optimization.*

[42] Y.-M. Wang and C. Parkan, “A preemptive goal programming method for aggregating OWA operator weights in group decision making,” *Inf. Sci.*, vol. 177, no. 8, pp. 1867–1877, Apr. 2007.

[43] J. L. Ringuest and T. R. Gulledge, “A Preemptive Value-Function Method Approach for Multiobjective Linear Programming Problems,” *Decis. Sci.*, vol. 14, no. 1, pp. 76–86, Jan. 1983.

[44] J. P. Ignizio, “A Review of Goal Programming: A Tool for Multiobjective Analysis,” *J. Oper. Res. Soc.*, vol. 29, no. 11, pp. 1109–1119, Nov. 1978.

[45] C. Romero, “A survey of generalized goal programming (1970–1982),” *Eur. J. Oper. Res.*, vol. 25, no. 2, pp. 183–191, May 1986.

[46] J. Kornbluth, “A survey of goal programming,” *Omega*, vol. 1, no. 2, pp. 193–205, Apr. 1973.

[47] J. P. Ignizio, “Generalized goal programming An overview,” *Comput. Oper. Res.*, vol. 10, no. 4, pp. 277–289, 1983.

[48] A. Charnes and W. W. Cooper, “Goal programming and multiple objective optimizations: Part 1,” *Eur. J. Oper. Res.*, vol. 1, no. 1, pp. 39–54, Jan. 1977.

[49] M. Tamiz, D. Jones, and C. Romero, “Goal programming for decision making: An overview of the current state-of-the-art,” *Eur. J. Oper. Res.*, vol. 111, no. 3, pp. 569–581, Dec. 1998.

[50] C. Romero, *Handbook of Critical Issues in Goal Programming*. Elsevier, 2014.

[51] J. S. Hokey Min, “On the Origin and Persistence of Misconceptions in Goal Programming,” *J. Oper. Res. Soc. - J OPER RES SOC*, vol. 42, no. 4, pp. 301–312, 1991.

[52] A. Baykasoğlu, “Preemptive goal programming using simulated annealing,” *Eng. Optim.*, vol. 37, no. 1, pp. 49–63, Jan. 2005.

[53] N. Freed and F. Glover, “Simple but powerful goal programming models for discriminant problems,” *Eur. J. Oper. Res.*, vol. 7, no. 1, pp. 44–60, May 1981.

[54] M. Zeleny, “The pros and cons of goal programming,” *Comput. Oper. Res.*, vol. 8, no. 4, pp. 357–359, 1981.

[55] S. I. Gass, “The setting of weights in linear goal-programming problems,” *Comput. Oper. Res.*, vol. 14, no. 3, pp. 227–229, 1987.

[56] C. Romero, “Naive Weighting in Non-Preemptive Goal Programming,” *J. Oper. Res. Soc.*, vol. 36, no. 7, pp. 647–648, Jul. 1985.

[57] H. D. Sherali and A. L. Soyster, “Preemptive and nonpreemptive multi-objective programming: Relationship and counterexamples,” *J. Optim. Theory Appl.*, vol. 39, no. 2, pp. 173–186, Feb. 1983.

[58] L. Arbea, L. I. Ramos, R. Martínez-Monge, M. Moreno, and J. Aristu, “Intensity-modulated radiation therapy (IMRT) vs. 3D conformal radiotherapy (3DCRT) in locally advanced rectal cancer (LARC): dosimetric comparison and clinical implications,” *Radiat. Oncol.*, vol. 5, no. 1, p. 17, Feb. 2010.

[59] G. Luxton, S. L. Hancock, and A. L. Boyer, “Dosimetry and radiobiologic model comparison of IMRT and 3D conformal radiotherapy in treatment of carcinoma of the prostate,” *Int. J. Radiat. Oncol.*, vol. 59, no. 1, pp. 267–284, May 2004.

[60] L. K. Schubert, V. Gondi, E. Sengbusch, D. C. Westerly, E. T. Soisson, B. R. Paliwal, T. R. Mackie, M. P. Mehta, R. R. Patel, W. A. Tomé, and G. M. Cannon, “Dosimetric

comparison of left-sided whole breast irradiation with 3DCRT, forward-planned IMRT, inverse-planned IMRT, helical tomotherapy, and topotherapy," *Radiother. Oncol.*, vol. 100, no. 2, pp. 241–246, Aug. 2011.

[61] M. T. Vlachaki, T. N. Teslow, C. Amosson, N. W. Uy, and S. Ahmad, "IMRT versus conventional 3DCRT on prostate and normal tissue dosimetry using an endorectal balloon for prostate immobilization," *Med. Dosim.*, vol. 30, no. 2, pp. 69–75, 2005.

[62] B. Longobardi, E. De Martin, C. Fiorino, I. Dell'oca, S. Broggi, G. M. Cattaneo, and R. Calandrino, "Comparing 3DCRT and inversely optimized IMRT planning for head and neck cancer: Equivalence between step-and-shoot and sliding window techniques," *Radiother. Oncol.*, vol. 77, no. 2, pp. 148–156, Nov. 2005.

[63] Z. Jiang, M. A. Earl, G. W. Zhang, C. X. Yu, and D. M. Shepard, "An examination of the number of required apertures for step-and-shoot IMRT," *Phys. Med. Biol.*, vol. 50, no. 23, p. 5653, Dec. 2005.

[64] E. Ludlum and P. Xia, "Comparison of IMRT planning with two-step and one-step optimization: a way to simplify IMRT," *Phys. Med. Biol.*, vol. 53, no. 3, p. 807, Feb. 2008.

[65] M. Broderick, M. Leech, and M. Coffey, "Direct aperture optimization as a means of reducing the complexity of Intensity Modulated Radiation Therapy plans," *Radiat Oncol*, vol. 4, no. 8, pp. 1–7, 2009.

[66] M. M. Matuszak, E. W. Larsen, and B. A. Fraass, "Reduction of IMRT beam complexity through the use of beam modulation penalties in the objective function," *Med. Phys.*, vol. 34, no. 2, pp. 507–520, Feb. 2007.

[67] T. Bortfeld, J. Bürkelbach, R. Boesecke, and W. Schlegel, "Methods of image reconstruction from projections applied to conformation radiotherapy," *Phys. Med. Biol.*, vol. 35, no. 10, pp. 1423–1434, Oct. 1990.

[68] X. Zhang, X. Li, E. M. Quan, X. Pan, and Y. Li, "A methodology for automatic intensity-modulated radiation treatment planning for lung cancer," *Phys. Med. Biol.*, vol. 56, no. 13, pp. 3873–3893, Jul. 2011.

[69] S. Breedveld, P. R. M. Storchi, P. W. J. Voet, and B. J. M. Heijmen, "iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans," *Med. Phys.*, vol. 39, no. 2, pp. 951–963, Feb. 2012.

[70] H. Yan, F.-F. Yin, H. Guan, and J. H. Kim, "AI-guided parameter optimization in inverse treatment planning," *Phys. Med. Biol.*, vol. 48, no. 21, pp. 3565–3580, Nov. 2003.

[71] J. L. Bedford and S. Webb, "Elimination of importance factors for clinically accurate selection of beam orientations, beam weights and wedge angles in conformal radiation therapy," *Med. Phys.*, vol. 30, no. 7, pp. 1788–1804, Jul. 2003.

[72] J. J. Wilkens, J. R. Alaly, K. Zakarian, W. L. Thorstad, and J. O. Deasy, "IMRT treatment planning based on prioritizing prescription goals," *Phys. Med. Biol.*, vol. 52, no. 6, p. 1675, Mar. 2007.

[73] K.-W. Jee, D. L. McShan, and B. A. Fraass, "Lexicographic ordering: intuitive multicriteria optimization for IMRT," *Phys. Med. Biol.*, vol. 52, no. 7, p. 1845, Apr. 2007.

[74] E. Schreibmann, M. Lahanas, L. Xing, and D. Baltas, "Multiobjective evolutionary optimization of the number of beams, their orientations and weights for intensity-modulated radiation therapy," *Phys. Med. Biol.*, vol. 49, no. 5, pp. 747–770, Mar. 2004.

[75] C. Wu, G. H. Olivera, R. Jeraj, H. Keller, and T. R. Mackie, "Treatment plan modification using voxel-based weighting factors/dose prescription," *Phys. Med. Biol.*, vol. 48, no. 15, pp. 2479–2491, Aug. 2003.

[76] Chunhua Men 1 , Xuejun Gu 1 , Dongju Choi 2 , Amitava Majumdar 2 , Ziyi Zheng 3 , Klaus Mueller 3 and S. B. Jiang, “GPU-based ultra fast IMRT plan optimization,” vol. 54, PHYSICS IN MEDICINE AND BIOLOGY, 2009.

[77] J. L. Hennessy and D. A. Patterson, *Computer Architecture: A Quantitative Approach*. Elsevier, 2012.

[78] S. A. Przybylski, *Cache and Memory Hierarchy Design: A Performance-directed Approach*. Morgan Kaufmann, 1990.

[79] K. R. Kaplan and R. O. Winder, “Cache-based Computer Systems,” *Computer*, vol. 6, no. 3, pp. 30–36, Mar. 1973.

[80] N. P. Jouppi, “Improving direct-mapped cache performance by the addition of a small fully-associative cache and prefetch buffers,” in *, 17th Annual International Symposium on Computer Architecture, 1990. Proceedings*, 1990, pp. 364–373.

[81] B. C. Martin, T. R. Bortfeld, and D. A. Castanon, “Accelerating IMRT optimization by voxel sampling,” *Phys. Med. Biol.*, vol. 52, pp. 7211–7228, Dec. 2007.

[82] A. Scherrer, K.-H. Küfer, T. Bortfeld, M. Monz, and F. Alonso, “IMRT planning on adaptive volume structures—a decisive reduction in computational complexity,” *Phys. Med. Biol.*, vol. 50, pp. 2033–2053, May 2005.

[83] J. O. Deasy, A. I. Blanco, and V. H. Clark, “CERR: a computational environment for radiotherapy research,” *Med. Phys.*, vol. 30, no. 5, pp. 979–985, May 2003.

[84] H. Rocha, J. M. Dias, B. C. Ferreira, and M. Carmo Lopes, “Influence of Sampling in Radiation Therapy Treatment Design,” *Comput. Sci. Its Appl. - ICCSA 2011*, vol. 6784, pp. 215–230, 2011.

[85] X.-Q. Lu and L. M. Chin, “Sampling techniques for the evaluation of treatment plans,” *Med. Phys.*, vol. 20, no. 1, pp. 151–161, 1993.

[86] J. Deasy, E. K. Lee, T. Bortfeld, M. Langer, K. Zakarian, J. Alaly, Y. Zhang, H. Liu, R. Mohan, R. Ahuja, A. Pollack, J. Purdy, and R. Rardin, “A collaboratory for radiation therapy treatment planning optimization research,” *Ann. Oper. Res.*, vol. 148, no. 1, pp. 55–63, Nov. 2006.

[87] M. C. Ferris, R. Einarsson, Z. Jiang, and D. Shepard, “Sampling issues for optimization in radiotherapy,” *Ann. Oper. Res.*, vol. 148, pp. 95–115, Oct. 2006.

[88] V. H. Clark, Y. Chen, J. Wilkens, J. R. Alaly, K. Zakaryan, and J. O. Deasy, “IMRT treatment planning for prostate cancer using prioritized prescription optimization and mean-tail-dose functions,” *Linear Algebra Its Appl.*, vol. 428, no. 5–6, pp. 1345–1364, Mar. 2008.

[89] R. K. A. H Edwin Romeijn, “A novel linear programming approach to fluence map optimization for intensity modulated radiation therapy treatment planning,” *Phys. Med. Biol.*, vol. 48, no. 21, pp. 3521–42, 2003.

[90] H. E. Romeijn, R. K. Ahuja, J. F. Dempsey, and A. Kumar, “A New Linear Programming Approach to Radiation Therapy Treatment Planning Problems,” *Oper. Res.*, vol. 54, no. 2, pp. 201–216, Mar. 2006.

[91] D. P. Bertsekas, *Nonlinear programming*. Athena Scientific, 1999.

[92] M. A. Hunt and C. M. Burman, “Treatment planning considerations using IMRT,” *A Practical Guide to Intensity Modulated Radiation Therapy*, pp. 103–21, 2003.

[93] I. J. Kalet, R. S. Giansiracusa, C. Wilcox, and M. Lease, “Radiation Therapy Planning: an Uncommon Application of Lisp,” *Lisp in the Mainstream: Proceedings of the Conference on the 40th Anniversary of Lips*, 1998.

[94] S. Davis, A. Alexander, M. Renaud, M. Hobson, E. Soisson, R. Ruo, and J. Seuntjens, "SU-E-T-698: Clinical Evaluation of the McGill Monte Carlo Treatment Planning System (MMCTP) for Fixed-Field IMRT and RapidArc Plans," 2011, vol. 38, pp. 3650–3651.

[95] J. O. Deasy, A. I. Blanco, and V. H. Clark, "CERR: A computational environment for radiotherapy research," *Med. Phys.*, vol. 30, no. 5, pp. 979–985, 2003.

[96] A. Wachter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Math Program*, vol. 106, no. 1, pp. 25–57, May 2006.

[97] J. Walter and K. Mathias, *uBlas. Boost C++ library available from www.boost.org/doc/libs*. 2006.

[98] Mosek, "The MOSEK optimization toolbox for MATLAB." Mosek, -.

[99] R. E. Drzymala, R. Mohan, L. Brewster, J. Chu, M. Goitein, W. Harms, and M. Urie, "Dose-volume histograms," *Int. J. Radiat. Oncol.*, vol. 21, no. 1, pp. 71–78, May 1991.

[100] T. Halabi, D. Craft, and T. Bortfeld, "Dose–volume objectives in multi-criteria optimization," *Phys. Med. Biol.*, vol. 51, no. 15, pp. 3809–3818, Aug. 2006.

[101] J. O. Deasy, "Multiple local minima in radiotherapy optimization problems with dose–volume constraints," *Med. Phys.*, vol. 24, no. 7, pp. 1157–1161, Jul. 1997.

[102] H. E. Romeijn and J. F. Dempsey, "Intensity modulated radiation therapy treatment plan optimization," *TOP*, vol. 16, no. 2, pp. 215–243, Nov. 2008.

[103] J. Llacer, J. O. Deasy, T. R. Bortfeld, T. D. Solberg, and C. Promberger, "Absence of multiple local minima effects in intensity modulated optimization with dose volume constraints," *Phys. Med. Biol.*, vol. 48, no. 2, pp. 183–210, Jan. 2003.

[104] I. Kawrakow, "VMC++, Electron and Photon Monte Carlo Calculations Optimized for Radiation Treatment Planning," in *Advanced Monte Carlo for Radiation Physics, Particle Transport Simulation and Applications*, D. A. Kling, P. F. J. C. Baräo, P. M. Nakagawa, P. L. Távora, and P. P. Vaz, Eds. Springer Berlin Heidelberg, 2001, pp. 229–236.

[105] N. Metropolis and S. Ulam, "The Monte Carlo Method," *J. Am. Stat. Assoc.*, vol. 44, no. 247, pp. 335–341, Sep. 1949.

[106] Q. Wu, R. Mohan, and A. Niemierko, "IMRT optimization based on the generalized equivalent uniform dose (EUD)," in *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2000*, 2000, vol. 1, pp. 710–713 vol.1.

[107] B. Choi and J. O. Deasy, "The generalized equivalent uniform dose function as a basis for intensity-modulated treatment planning," *Phys. Med. Biol.*, vol. 47, no. 20, pp. 3579–3589, Oct. 2002.

[108] Q. Wu, D. Djajaputra, H. H. Liu, L. Dong, R. Mohan, and Y. Wu, "Dose sculpting with generalized equivalent uniform dose," *Med. Phys.*, vol. 32, no. 5, pp. 1387–1396, May 2005.

[109] Q. Wu, R. Mohan, A. Niemierko, and R. Schmidt-Ullrich, "Optimization of intensity-modulated radiotherapy plans based on the equivalent uniform dose," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 52, no. 1, pp. 224–235, Jan. 2002.

[110] C. Thieke, T. Bortfeld, A. Niemierko, and S. Nill, "From physical dose constraints to equivalent uniform dose constraints in inverse radiotherapy planning," *Med. Phys.*, vol. 30, no. 9, pp. 2332–2339, Sep. 2003.

[111] M. Zarepisheh, T. Long, N. Li, Z. Tian, H. E. Romeijn, X. Jia, and S. B. Jiang, "A DVH-guided IMRT optimization algorithm for automatic treatment planning and adaptive radiotherapy replanning," *Med. Phys.*, vol. 41, no. 6, p. 061711, Jun. 2014.

[112] D. Craft, T. Halabi, H. A. Shih, and T. Bortfeld, “An Approach for Practical Multiobjective IMRT Treatment Planning,” *Int. J. Radiat. Oncol.*, vol. 69, no. 5, pp. 1600–1607, Dec. 2007.

[113] D. Craft, T. Halabi, and T. Bortfeld, “Exploration of tradeoffs in intensity-modulated radiotherapy,” *Phys. Med. Biol.*, vol. 50, no. 24, pp. 5857–5868, Dec. 2005.

[114] D. L. Craft, T. F. Halabi, H. A. Shih, and T. R. Bortfeld, “Approximating convex Pareto surfaces in multiobjective radiotherapy planning,” *Med. Phys.*, vol. 33, no. 9, pp. 3399–3407, Sep. 2006.