Optimization of Beam Orientations in Intensity Modulated Radiation Therapy Planning

Dissertation

von

Ahmad Saher Azizi Sultan

Dedicated to the memory of my first teachers;

Prof. Mohammed Ibn Abd-Allah

and

my parents

Acknowledgments

Kaiserslautern, June 2006

Ahmad Saher Azizi Sultan

Abstract

For the last decade, optimization of beam orientations in intensity-modulated radiation therapy (IMRT) has been shown to be successful in improving the treatment plan. Unfortunately, the quality of a set of beam orientations depends heavily on its corresponding beam intensity profiles. Usually, a stochastic selector is used for optimizing beam orientation, and then a single objective inverse treatment planning algorithm is used for the optimization of beam intensity profiles. The overall time needed to solve the inverse planning for every random selection of beam orientations becomes excessive. Recently, considerable improvement has been made in optimizing beam intensity profiles by using multiple objective inverse treatment planning. Such an approach results in a variety of beam intensity profiles for every selection of beam orientations, making the dependence between beam orientations and its intensity profiles less important. This thesis takes advantage of this property to accelerate the optimization process through an approximation of the intensity profiles that are used for multiple selections of beam orientations, saving a considerable amount of calculation time. A dynamic algorithm (DA) and evolutionary algorithm (EA), for beam orientations in IMRT planning will be presented. The DA mimics, automatically, the methods of beam's eye view and observer's view which are recognized in conventional conformal radiation therapy. The EA is based on a dose-volume histogram evaluation function introduced as an attempt to minimize the deviation between the mathematical and clinical optima. To illustrate the efficiency of the algorithms they have been applied to different clinical examples. In comparison to the standard equally spaced beams plans, improvements are reported for both algorithms in all the clinical examples even when, for some cases, fewer beams are used. A smaller number of beams is always desirable without compromising the quality of the treatment plan. It results in a shorter treatment delivery time, which reduces potential errors in terms of patient movements and decreases discomfort.

Contents

Chapter 1

Introduction

Almost immediately after the discovery of X-rays, the biological effects of emitting radiation to human cells were recognized, therefore, being addressed and studied. The result was fortunate. It was found that, although high doses of radiation can kill both cancerous and normal cells, the cancerous cells are more sensitive to radiation than the normal cells. This is due to the fact that the repair mechanism of cancerous cells is less efficient than that of normal ones. Thus normal cells are more likely to fully recover from the effects of radiation [NCI97]. This natural fact, together with a good radiation treatment plan, has proved radiation therapy to be a successful tool in cancer treatments. Consequently, radiation therapy becomes one of the most-used treatment tools during the fight against cancer. Approximately, fifty percent of the people who are diagnosed with life threatening forms of cancer are treated with radiation, either exclusively or in combination with surgery and chemotherapy.

The majority of radiation therapy treatments are performed using high energetic photon beams generated by a linear accelerator and delivered to the patient from different directions (see figure 1.1). The aperture of each individual beam is created so that it adheres to the shape of the tumor. This can be done by using an automatic device called a multi-leave collimator (MLC) in which a number of parallel leaves are projected into the primary beam to create the required tumorodical shape [VRA]. For depiction see figure (1.2).

In conventional conformal radiation therapy, the adjustment of the MLC aperture can not be altered while the linear accelerator is turned on. Thus the resulting radiation will have uniform intensity throughout the beam aperture. In many cases, especially for irregularly shaped non-convex targets, this
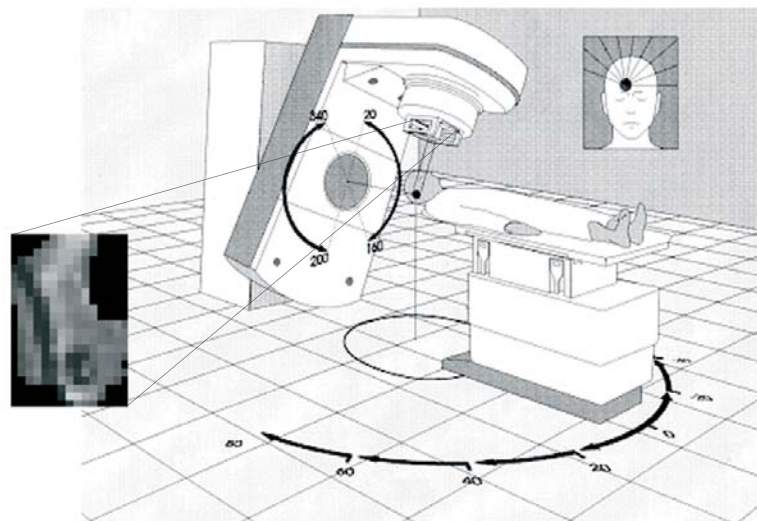
Figure 1.1: The gantry can move around the couch on which the patient lies and the couch itself may be moved to alter the directions

limits the possibilities of achieving dose distribution with a good degree of accuracy to the prescribed dose.

Recently, the above limitation is somewhat overcome by the invention of intensity-modulated radiation therapy (IMRT), where the MLC is equipped with what is called dynamic mode. Here the leaves can be moved, while the beam is turned on, to block (unblock) portions of the treatment field where less (more) intensity would be more accurate to the prescribed dose. Such a dynamic mode is capable of producing inhomogeneous beams, allowing the planner to realize complex shaped dose distributions. Thus IMRT is able to better conforme to high dose distributions and hence it results in more precise therapies [Web01].

Despite the use of IMRT technology, one third of the patients that are treated by radiotherapy die from receiving either too little radiation to cure the tumor or too much radiation to surrounding healthy tissue, or organs at risk (OAR), which leads to complications in the region and sometimes to death (see [HK02]). As a matter of fact, the goals of delivering a high dose of radiation to the tumor and, at the same time, a small dose to the
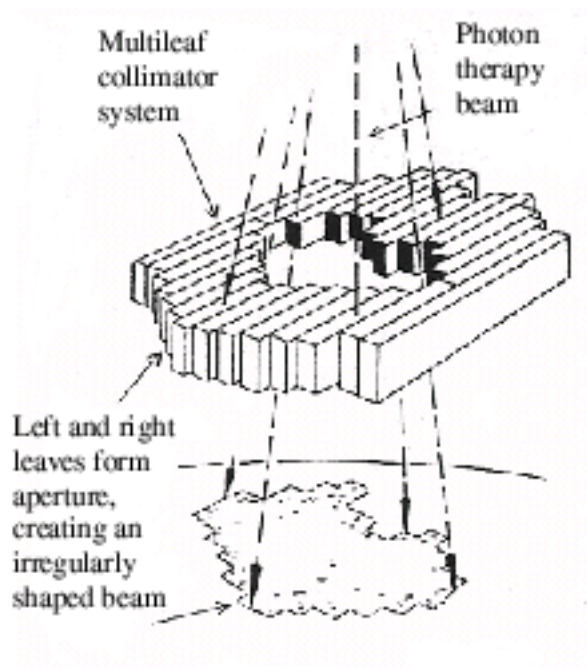
Figure 1.2: A Multileaf Colimator (MLC) system. (picture from [RADK03])

surrounding OARs are of a contradictory nature and hence not achievable simultaneously, especially in the case where some OARs are adjacent to the tumor. In such a case part of the adjacent OAR will receive a dose at least equal to the minimum target dose. Therefore the alternative realistic treatment objective can only be to find and deliver a curative dose to the tumor which spares as much as possible of the surrounding OARs. Realizing such a curative dose with the ideal clinical sparing of OARs for each individual patient, challenges the radiation therapy planners in their daily practice.

The IMRT treatment planning problem is solved when the optimal set of its parameters is found. Although many of these parameters, like the dose bounds, radiation modality, energies, fractionation, etc. do not lend themselves to mathematical approaches -they can simply be considered as given parameters in mathematical contexts- the problem of IMRT remains a large scale optimization problem of a highly complex nature. Solving it in reasonable time is still beyond the capacities of today's computer equipment. Therefore, in practice, the variables system of the IMRT problem is split into two classes: the *beam orientations* given by the number of beams with their positions (gantry and table angles), and the intensity distribution of

the beamlets, called also *intensity maps*.

During the last few decades the problem was often approached using a two stage algorithm in which a stochastic optimization technique is employed for the beam orientations, while an iterative inverse planning optimization scheme is used for the intensity maps calculations. Here the optimization process evaluates the plan by using a one-dimensional objective function and returns the plan with the best value. Such an approach has two major drawbacks: first any proper stochastic technique requires the optimization process which calculates the intensity maps to be performed many thousand times, each time for a different selection of beam orientations. The total time needed for this becomes excessive especially in the 3D situation. Secondly, using a one-dimensional objective function to measure the quality of the treatment plan, which is essentially of conflicting goals, is not proper. The resulting plan is often unsatisfactory with respect to clinical considerations.

Recently researchers in the field have admitted that the IMRT problem is a multi-criteria optimization problem and started formulating it using a multi-objective function. In contrast to the formal approach, where the treatment plan is represented by one single value, here it is represented by a vector containing several objectives (criteria), each corresponding to a certain organ. For each selection of beam orientations, optimality is no longer characterized by a unique solution of the intensity maps but rather by a set of efficient solutions. A solution is efficient, also called Pareto optimal, if any improvement in one criterion will worsen at least one of the other criteria. Such an approach allows the planner to search through a variety of Pareto optimal solutions and eventually decide upon the most desirable treatment plan for the corresponding beam orientations. Although this has led to a considerable success with respect to the configuration of the intensity maps, the problem of beam orientations in IMRT remains unsolved, and the choice of beam directions in IMRT is still a trial-and-error procedure based on intuition an empirical knowledge.

This thesis is focused on the problem of beam orientations in IMRT planning. It is organized in the following manner. Chapter 2, which is a complementary chapter, discusses in general the current state of the art in inverse planning radiation therapy. Chapter 3 introduces the problem of beam orientations and shows that that problem is non-convex when it is formulated as a continuous optimization problem and NP-hard if it is modeled as a combinatorial optimization problem. Chapter 4 utilizes the methods of beam's eye view and observer's view, which are recognized for beam orientation in con-

ventional conformal radiation therapy, in form of a fully automatic algorithm for beam orientations in IMRT. To illustrate the efficiency of the algorithm it has been applied to an artificial example where optimality is trivial, and to different clinical cases. In comparison to the standard equally spaced beam plans, improvements are reported in all examples, even with fewer number of beams. Chapter 5 introduces a dose volume histogram evaluation scheme for measuring the quality of a given IMRT plan and presents an accelerated evolutionary algorithm for optimizing beam configurations in IMRT. As the same as in chapter 4, similar results are reported when comparing the plan obtained by the evolutionary algorithm with the standard equally spaced beam plans. In Chapter 6 the results of comparing the evolutionary algorithm with the dynamic one are reported for different clinical cases. A summary and conclusion of what has been done in this thesis is found in chapter 7, and an outlook of what could be done is discussed in chapter 8.

Chapter 2

The multicriteria environment in intensity modulated radiation therapy planning

As we have seen in the introductory chapter, usually the optimization process in IMRT planning aims at finding mainly the intensity maps and, sometimes, beam orientations. In contrast to the main goal of this thesis, beam orientations are assumed to be given throughout this complementary chapter which is meant to present some modeling issues and give an over view on intensity maps optimization.

Moreover, as the title implies, this chapter is also meant to discover the invironmental framework of the problem of IMRT when it is modeled as a multicriteria optimization problem. The goal of such an approach is to generate a database containing a sufficient number of representatives of the Pareto optimal solutions that are of clinical interest. These solutions are then browsed using a specially designed navigation tool, allowing for an easy access over the database. The planner can then easily search through the database interactively until eventually he/she decides upon the most desirable treatment plan.

2.1   From IMRT to multicriteria optimization

The goal of the planners in IMRT is to find optimal compromises between a sufficiently high dose in tumor tissue that guarantee a high tumor control and dangerous overdose of critical structures in order to avoid high normal tissue complication problems. To achieve this goal it is required, in many

cases, to balance overdoses between the different critical structures where each of them has its own biological properties when it is being exposed to radiation. This demonstrates that the problem of IMRT planning has the same nature as the multicriteria optimization problem, which is, in general, formulated as follows (see e.g. [GSH99]):

$$\min F = (f_1, f_2, \ldots, f_m),$$
$$s.t. \quad x \in \mathcal{S}, \tag{2.1}$$

where $\mathcal{S}$ is a nonempty subset contained in $\mathbb{R}^n$, $n \in \mathbb{N}$, and $F : \mathcal{S} \to \mathbb{R}^m$, where also $m \in \mathbb{N}$. Moreover it is assumed that the set $\mathbb{R}^m$ is partially ordered by the binary relation $\leqslant$, and the minimality is given by the definition stated below.

Definition 1. *Let L be an arbitrary nonempty subset of $\mathbb{R}^m$.*
$x \in L$ is called a minimal element of L, if there is no $y \in L$ with $y \neq x$ and $y \leqslant x$.

Since F is a vector valued function, one speaks of a so called multi-objective optimization problem. The set $\mathcal{S}$ is called the constraints set (planning domain), $f_1, f_2, \ldots, f_m$ are called the indicator functions and F the objective function, which is, in our case, a way of evaluating the treatment plan.

2.1.1   Treatment plan evaluations

Usually the treatment plan is determind by its intensity vector $\mathbf{x}$ - or sometimes by its corresponding dose distribution $\mathbf{d(x)} = \mathbf{Ax}$ for a given dose matrix $\mathbf{A}$ - and represented by so called *dose volume histogram* (DVH) which depicts the percentage of the volume of a considered clinical structure that receives at least a certain dose over the relevant dose interval. For illustration see figure (2.1) in which a depiction of a DVH is represented by the solid curve. Usually the dose distribution of each clinical structure is represented by an individual DVH where the quality of a plan is finally judged by the shape of these histograms.

Consequently, treatment plan evaluation is, indeed, nothing more than finding a proper evaluation function that is able to quantify the quality of the corresponding DVHs efficiently.

Typical evaluation functions

In mathematical sense the problem of IMRT treatment planning is generally considered as a constraint optimization problem with either a single or multi
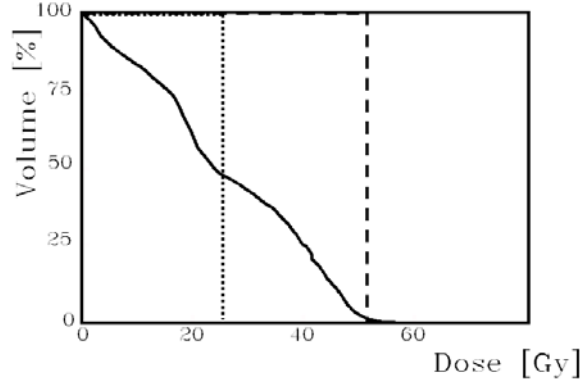
Figure 2.1: Exemplary DVH curve and the coresponding EUD with q about 1 and and q close to ∞

criteria objective function. However in both cases, the role of the objective function is to measure the quality or the goodness of a given treatment plan subject to some physical and clinical considerations. The objective function is an evaluation function f, that maps the dose distribution $\mathbf{d}(\mathbf{x})$ with respect to a certain clinical structure to a real value $f(\mathbf{d}(\mathbf{x}))$. For simplification, the critical structures in the body are discretized into small volume parts v_j called *voxels*. Standard evaluation functions include those that measure deviations from a desired dose value D^p in the target T:

$$f(\mathbf{d}(\mathbf{x})) = \max_{v_j \in T} |D^p - d(v_j)(\mathbf{x})|, \tag{2.2}$$

or

$$f(\mathbf{d}(\mathbf{x})) = w_T \sum_{v_j \in T} (D^p - d(v_j)(\mathbf{x}))^2, \tag{2.3}$$

where w_T is the target importance factor (see e.g. [SFOM99] and [PBX00]).

Similarly there are functions which penalize voxel doses exceeding an upper bound U in an organ at risk R:

$$f(\mathbf{d}(\mathbf{x})) = \left(|R|^{-1} \sum_{v_j \in R} \max \left(d(v_j)(\mathbf{x}) - U, 0 \right)^q \right)^{\frac{1}{q}}, \quad q \in [1, \infty). \tag{2.4}$$

where $|R|$ denotes the number of voxels in R and q is an organ dependent parameter. Another choice are functions which take account of the whole shape of a dose distribution by means of an equivalent uniform dose (EUD), cf. e.g.

[Bra84]. Using for example A. Niemierko's EUD [Nie97], the deviation for the organ at risk R from U is measured by

$$f(\mathbf{d}(\mathbf{x})) = \frac{1}{U}\Big(|R|^{-1}\sum_{v_j \in R} d(v_j)(\mathbf{x})^q\Big)^{\frac{1}{q}}, \quad q \in [1, \infty), \qquad (2.5)$$

The EUD measure for an organ or a clinical structure is derived by finding the homogeneous dose corresponding to the tumor control probability (TCP) or normal tissue complication probability (NTCP) for OAR in an inhomogeneous dose distribution. Figure 2.1 illustrates two possible EUD evaluations of a DVH. The dotted and the dashed lines are EUD measures with q about 1 and q close to ∞, respectively. Thus, the dotted line is equivalent to the original histogram under the EUD-measure using $q = 1$. Organs that work parallel (e.g. lungs, kidneys) are often evaluated using values for q that are close to 1, whereas serial organs such as the spinal chord are evaluated with relatively high values for q.

TCP and NTCP are statistical functions that measure the probability of destroying clonogenic cells of the tumor and the probability of damaging risk organs, respectively, dependent on dosages. Lower dose bound (prescribed dose D^p) for the target and upper dose bounds (U) for organs that guaranty a high probability of tumor control and a low probability of damaging risks can be derived using these function based on statistics gained from experiences with thousands of treated patients (see, e.g.[ELB⁺91]).

The basic property of the mentioned evaluation functions is convexity and a resulting optimization model can be guaranteed to be solved to optimality. The physician may nevertheless, for the sake of more efficiency, evaluate the plan on the basis of non-convex criteria like dose volume histogram constraints. Such evaluation functions will be discussed separately in chapter 6, where a DVH based evaluation function is introduced and used to solve the problem of beam orientaitons.

2.1.2 Pareto solutions and the planning domain

Recall the the multiciteria optimization problem in its general form

$$\begin{aligned}&\min F = (f_1, f_2, \ldots, f_K),\\ &s.t. \ \ x \in \mathcal{S}.\end{aligned} \qquad (2.1)$$

Pareto optimal solution is defined as follows ([GSH99]):

Definition 2. *$x \in \mathcal{S}$ is called an efficient solution or a Pareto optimal solution of the problem (2.1), if $F(x)$ is a minimal element of the image set $F(\mathcal{S})$.*

In another words, an element $x \in \mathcal{S}$ is a Pareto optimal solution of the problem (2.1) if there is no $y \in \mathcal{S}$ such that $f_k(x) \geqslant f_k(y)$ for all $k = 1, \ldots, K$, where strict inequality holds for at least one index k. Thus, the efficient solution property indicates that any improvement that could happen to some of the objectives f_k will worsen some of the others. The method used here is based on gathering preference information after the calculation of some Pareto solutions. In fact it is neither possible nor meaningful to calculate all efficient solutions. It is not possible because the Pareto-front is a coherent and therefore an infinite subset of the set of feasible solutions. It is also not meaningful because there are many Pareto solutions that are clinically irrelevant.

For instance, in the example given in Figure (2.2) one would not go from point A with dose levels of 11 Gy in the spinal cord and 13 Gy in the parotid gland to the upper left efficient solution with dose levels of 9 Gy (spinal cord) and 33 Gy (parotid gland). In other words, the 2 Gy dose reduction in the spinal cord at this low dose level is not worth the price of a 20 Gy dose increase in the parotid gland, which may cause xerostomia. We therefore try
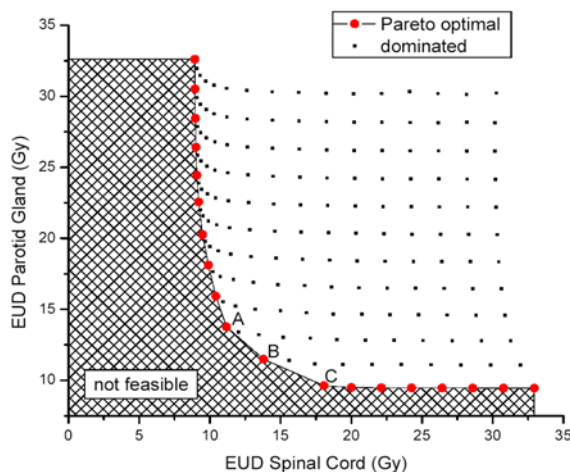
Figure 2.2: Exploration of the Pareto front for a head and neck case by brute force methods. Every square dot represents one treatment plan. A total of 16x16=256 plans was generated. The round dots represent the Pareto front for this case, i.e., the set of efficient treatment plans.

to focus on parts of the Pareto boundary that contain clinically meaningful plans. Since it is easier to classify clinical irrelevance than relevance, we try to exclude as many irrelevant plans as possible and call the remaining set of possible plans the *planning domain* $\mathcal{X}_u$.

To exclude plans that exceed the clinically acceptable values in the indicator functions, hard constraints are added. Let $\mathbf{F}$ be the vector on indicator functions and $(\mathbf{x})$ be the vector of beamlet intensities, the so called *intensity map*. Then these *box constraints*

$$\mathbf{F}(\mathbf{x}) \leq \mathbf{u}$$

for upper bound $\mathbf{u}$ should be set rather generously in order to allow a flexible range of solution outcomes from which experts may select an appropriate candidate. Of course, the more flexible this range is selected to be, the more calculations will be necessary. If chosen too strict, they may, however, lead to infeasibility. This serves as a first indication to the decision-maker that the initial appraisement of the situation was utopian. If after a relaxation of the box constraints there are still no feasible solutions, the oncologist may realize that more irradiation directions, i.e. more degree of freedom are needed to find a clinically acceptable solution and alter the geometry accordingly. In such cases optimizing beam orientations may play a considerable role in finding feasible solutions, even with less generous upper bound or fewer number of beams as it is demonstrated in the comming chapters.

2.1.3   Solution strategies

Once a planning domain is fixed, the problem to solve is given by

$$\begin{aligned} \mathbf{F}(\mathbf{x}) &\rightarrow \quad \min \quad \text{subject to} \\ \mathbf{x} &\in \quad \mathcal{X}_u, \end{aligned} \tag{2.6}$$

where

$$\mathcal{X}_u := \{\mathbf{x} \geq 0 \mid \mathbf{F}(\mathbf{x}) \leq \mathbf{u}\}$$

is the set of feasible intensity maps.

Usually multicriteria problems are solved by formulating possibly multiple scalarized versions of the problem. There are several standard methods along with their variations which can be used to solve the multicriteria problem and which exhibit different characteristics.

One possibility is to solve the *lexicographic problem*. It treats the objective functions one by one in a fixed order according to the preference of the decision-maker.

$$F_{(1)}(\mathbf{x}), \ldots, F_{(|\mathcal{K}|)}(\mathbf{x}) \quad \rightarrow \quad \text{lex min} \quad \text{subject to} \qquad (2.7)$$
$$\mathbf{x} \quad \in \quad \mathcal{X}_u$$

The theoretical justification to use the solution of problem (2.7) requires the decision maker to be able to arrange the objective functions according to their absolute importance. This implies that an indicator function $F_{(m)}$ that was ranked more important than some other criterion $F_{(n)}$, i.e. $m < n$, is infinitely more important to the decision maker [Mie99]. As this assumption cannot be made in clinical practice, this approach is not relevant to the IMRT optimization problem.

An applicable approach which has been widely used in commercial planning systems for the last few decades is the *weighted sum approach*. It lets the decision-maker choose positive weights w_k for a structure $k \in \mathcal{K}$ in the weighted scalarization problem

$$F_w(\mathbf{x}) = \sum_{k \in \mathcal{K}} w_k F_k(\mathbf{x}) \quad \rightarrow \quad \text{min} \quad \text{subject to} \qquad (2.8)$$
$$\mathbf{x} \quad \in \quad \mathcal{X}_u.$$

For convex multicriteria problems every set of positive weights yields a Pareto optimal plan and every Pareto optimal plan is an optimum of (2.8) for an appropriate set of non-negative weights (for more details see [Mie99]).

In this approach, if a large weight factor w_k for a structure $k \in \mathcal{K}$ is used more emphasis is placed on the dose delivered to this structure. Although this ensure a kind of comparability between the different objective functions, it however does not give the decision-maker any hint that can help him in finding suitable set of weight factors. In deed, these weights are usually determined by a trail and error procedure which is a mothod we label the "Human Iteration Loop" depicted in figure (2.3). This strategy has several pitfalls. First, it is impossible to ask the decision-maker for optimization parameters that leads to an ideal solution in the first iterate. An iterative adjustment of the parameters converges to a solution that hardly incorporates any wishes that are not explicitly modelled.

Weights in a scalarization approach, for example, are nothing but an attempt to translate the decision-maker's ideal into arbitrary weights. This
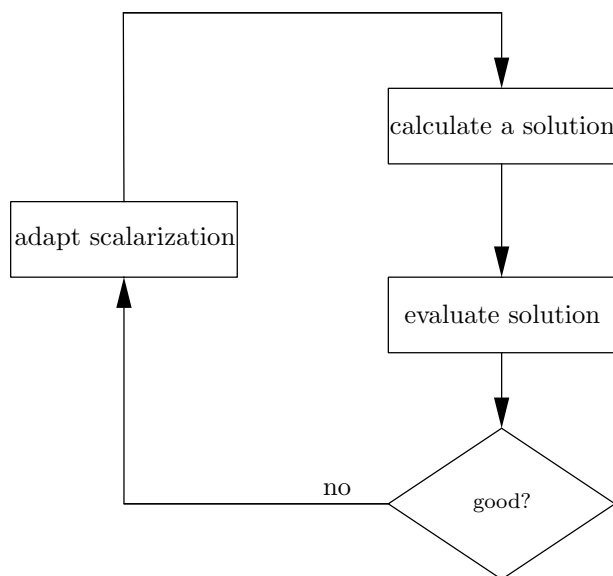
Figure 2.3: Human iteration loop described as method of *successive improvements*

artificial nature results from having to quantify global trade-off rates between different objectives that are often of very different nature. The fact that no information on efficiency can be conveyed allows only maximum time or a maximum number of repetitions as stopping rules for this process. These arguments render the Human Iteration Loop a futile effort. A truly multicriteria decision-making framework which is able to depict more information, allowing for flexible modeling becomes urgent.

Recently, to overcome the drawbacks mentioned above, optimizing the intensity maps for multicriteria IMRT is performed using different approaches in which the Pareto boundary is approximated and stored in a form of database that guarantees an easy access to the planner. Since the intensity maps calculations is beyond the scope of this dissertation, only a summary of one of these approaches will be mentioned in order to give the reader sort of complete understanding of the problem of beam orientations in multicriteria IMRT. For more details on this topic see e.g. [KMS+06].

The approach to be mentioned is the *compromise programming* or the *weighted metric* approach ([Mie99], [Yu73], [Zel73]). Here, a reference point, outside the feasible region, is chosen and the distance to it is minimized in a suitable metric. The ideal point - the point given by the minima of the individual indicator functions - or some utopia point - a point that is smaller than the ideal point in each component - can be used as a reference

point. The different components of $\mathbf{F}(\mathbf{x})$ are scaled to get different solutions. Alternatively, the metric can be varied, or both. The solutions obtained are guaranteed to be Pareto optimal if the metric is chosen appropriately and the scaling parameters are positive. A popular choice is the Tchebycheff problem.

$$\max_{k \in \mathcal{K}}\{\sigma_k F_k(\mathbf{x})\} \quad \rightarrow \quad \min \quad \text{subject to} \tag{2.9}$$
$$\mathbf{F}(\mathbf{x}) \quad \leq \quad \mathbf{u}$$
$$\mathbf{x} \quad \in \quad \mathcal{X}_u$$

Note that scaling here is not the same as choosing weights for a problem like the weighted scalarization (2.8) above. The scaling coefficients σ_k contain information about the willingness to deteriorate relative to the specified reference point. Hence, deviations from the treatment goals can be much better controlled by the reference point methods than by (2.8), since the solutions obtained by varying weights provide no information about the trade-offs of the objectives, see [DD97].

An initial set of representatives of the Pareto set $\mathcal{X}_{\mathrm{Par}}$, where

$$\mathcal{X}_{\mathrm{Par}} := \{\mathbf{x} \in \mathcal{X}_u \mid \mathbf{x} \text{ is Pareto optimal}\},$$

is obtained by solving problem (2.9): *extreme compromises*. The extreme compromises balance so-called *active indicator functions* only and Pareto optimality is enforced.

Let $\emptyset \neq \mathcal{M} \subseteq \mathcal{K}$ be the indices of the active indicators. Let $\mathcal{N}_0 := \mathcal{M}, \mathbf{u}^{(0)}$ be the bounds of the planning domain and $i = 0$. Then the solution of the following procedure is called extreme compromise for the active indicators $\mathcal{M}$.

1. Solve:

$$\max_{j \in \mathcal{N}_i}\{F_j(\mathbf{x})\} \quad \rightarrow \quad \min \quad \text{subject to} \tag{2.10}$$
$$\mathbf{F}(\mathbf{x}) \quad \leq \quad \mathbf{u}^{(i)}$$
$$\mathbf{x} \quad \in \quad \mathcal{X}_u$$

2. Let $\mathbf{y}^{(i)}$ be the optimal objective value of the above optimization problem and $\mathcal{O}_i \subseteq \mathcal{N}_i$ be the subset of indicators that cannot be further

improved. Then let

$$
\begin{aligned}
\mathbf{u}_j^{(i+1)} &:= \mathbf{y}^{(i)} \text{ for } j \in \mathcal{N}_i, \\
\mathcal{N}_{i+1} &:= \mathcal{N}_i \setminus \mathcal{O}_i \text{ and} \\
i &:= i + 1
\end{aligned}
$$

3. If $\mathcal{N}_i \neq \emptyset$, goto 1.

4. Repeat the above procedure with $\mathcal{N}_{i+1} := \mathcal{K} \setminus \mathcal{M}$, $\mathbf{u}^{(i+1)} := \mathbf{u}^{(i)}$ and $i := i + 1$.

The approach is a kind of mixture between weighted metric and lexicographic optimization. The solution for $\mathcal{M} = \mathcal{K}$ is also a solution of the Tchebycheff problem (2.9). It can easily be shown, that the extreme compromises are Pareto optimal. With this setting, we will find therapy plans that give an equibalanced solution for active indicator functions $F_k \in \mathcal{M}$, while the remaining non-active functions with indices $k \in \mathcal{N}$ satisfy the relaxed condition of staying in $\mathcal{X}_{\text{Par}}$. The extreme compromises are typically not of high clinical relevance, but they can serve, in our case as an approximation of the Pareto boundary, since they mark a relatively large region within $\mathcal{X}_{\text{Par}}$ in which the interesting solutions are contained. What is left, in order to convey the set $\mathbf{F}(\mathcal{X}_{\text{Par}})$ to the planner, is filling the gaps in between the approximated Pareto boundary by a good discrete approximation of Pareto solutions. The number of points needed to cover the Pareto boundary, which is in general a $|\mathcal{K}| - 1$ dimensional manifold, with a grid of maximum distance ρ is at least

$$
O\left((1/\rho)^{|\mathcal{K}|-1} \right),
$$

Such a number of points is neither tractable nor actually needed for the considered problem. In order to overcome the need for a fine grid optimized interpolation could be used to yield a continuous approximation of the Pareto boundary. More about the interpolated points and their optimization methods used to approximate that Pareto boundary in IMRT planning are to be found in [KMS+06].

The computation of the extreme compromises and the intermediate points are technical and done without human interaction, so that the calculation and generating the database could for example be performed over night allowing for interactive database navigation the next working day.

2.2 Database navigation

When the plan database computation is finished, a vital part of the planning is not yet accomplished. The planner still has to search through the database to decide on the most desirable plan. Since inspecting a plan usually involves sifting through all its slices a manual inspection of all plans contained in the database is infeasible. In particular, when considering additionally convex combination of plans and thus an infinite number of plans, manual inspection is not an option. The mechanisms used by the navigation tool allows for human interaction that is a distinct improvement compared to something like the human iteration loop described in section (2.1.3).

Figure (2.4) shows the user interface screen, for the navigation tool, which is divided into two parts. The left hand side visualizes the database as a whole and embeds the current solution into the database. The right hand side displays the current plan's dose volume histogram and the dose distribution on transversal, frontal and sagittal slices.

The star on the left hand side is composed of axes for the different indicator functions. The indicator functions associated to the organs at risk are combined into the radar plot, where as the indicators associated with the tumor volumes are shown as separate axes. The interval on the axes corresponds to the range of values contained in the database for the respective indicator functions. The white polygon marks the indicator function values for the currently selected plan.

The shaded area represents the *planning horizon*. It is subdivided into the *active* and the *inactive* planning horizon. The former is bounded on each axis by the maximum and minimum values implied by the currently set restrictions and the latter is the currently excluded rang contained in the database. Note that the line connecting the minimum values of the active planning horizon is the current ideal point estimate and the line connecting the maximum values is the current estimate for the Nadir point.

The line representing the currently selected plan has a handle bar called *selector* at each intersection with an axis and triangles for the tumor related axes. Both can be grabbed with the mouse and moved to ease the selection process performed by the planner. The right hand side of the screen displays the plans that are found while the selector is moved. The axes also contain *restrictors* represented by brackets that can also be grabbed with the mouse and moved to change the upper bounds for the corresponding indica-
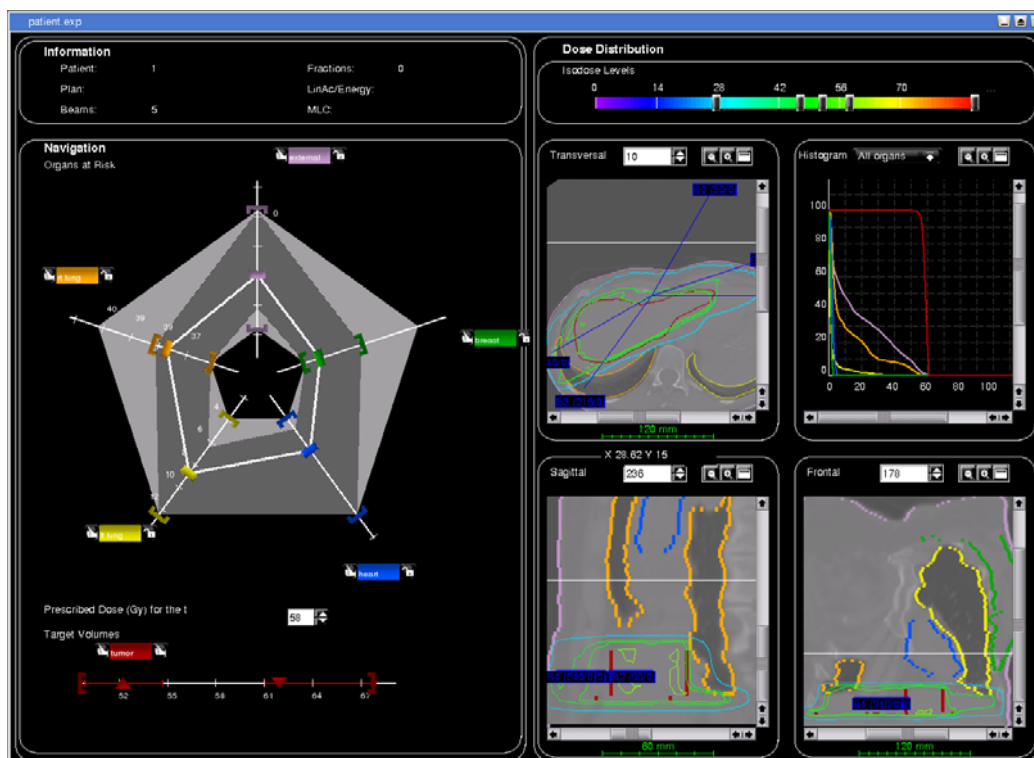
Figure 2.4: The navigation screen. The star on the left hand side with the active and inactive planning horizon and the current solution. On the right the dose visualization and the dose volume histogram.

tor. When the planner moves a restrictor, the active and inactive planning horizon are updated simultaneously.

The visualization is updated several - usually around seven - times a second when a selector or restrictor is moved, so the user gets immediate feedback about the consequences of his or her decisions. Instead of waiting for the consequences of a parameter adjustment the planner immediately sees the outcome of each differnet movement. This process offers a level of interactivity which provides the planner with an easy selection mechanism during the search for a good treatment plan.

Chapter 3

The problem of beam orientations in focus

In chapter 2 we have seen that the problem of IMRT is usually treated using two approaches. The traditional one was based on optimizing a single objective function which evaluates the treatment plan and assigns a number to it. Mainly due to the contradictory goals of treatment planning, this single number was not able to determine the quality of the plan properly. Consequently, the result of such an approach is often suboptimal with respect to clinical consideration. The most recent approach, which is more accurate to the structure of the problem, models it as a multi-criteria optimization problem with conflicting objectives. However, both approaches calculate beam intensity profiles under the assumption of a fixed preselected set of beam orientations. In fact, optimization of beam orientations may play a significant role during the search for the best possible compromise in IMRT [PLB+01]. However, prefixing the set of beam orientations seems to be prevalent in proposed methods. This is due to the fact that the problem of beam orientations is of high mathematical complexity, when modeled as a continuouse optimization problem, or high difficulty level when approached by the use of combinatorial optimization, as demonstrated below.

3.1 Complexity of beam orientations

3.1.1 Beam orientation is not convex

Unfortunately, the influence of a set of beam orientations on the final dose distribution can not be determined unless an optimal solution of the intensity maps is found. Moreover, any proper coupling between beam orientations

and intensity maps to formulate the complete problem will indeed turn the objective function in both of the above mentioned approaches into a nonconvex function, preventing the optimization process from the usage of gradient methods which might very likely, in such a case, be trapped in a local minimum ([BS93]). To examine this closely, let us consider the following example illustrated in figure (3.1).
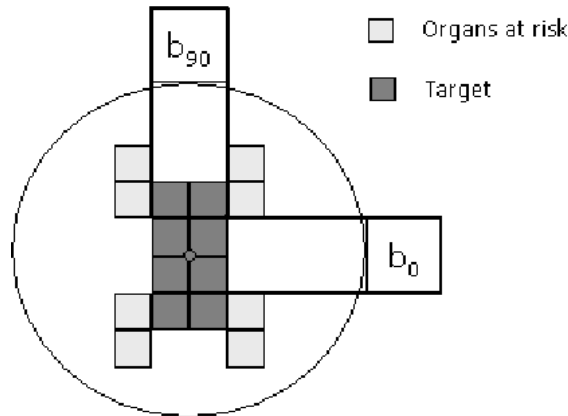
Figure 3.1: Optimization of beam orientation is highly nonconvex.

The depicted example consists of a rectangular shaped target placed at the center of a circular phantom in addition to four other smaller rectangular structures, constituting organs at risk, placed at the corners of the target. For simplicity reasons, we assume that the optimization process is carried out for this special example under the following conditions:

- The set of beam orientations consists of only one single beam, so we are seeking the best beam instead of a combination of beams.

- The beam head consists of only one single pixel.

- The beam width is fixed as it is shown in figure (3.1).

- The beam head may be arranged at any angle on the circular phantom. In other words we assume that the beam space set is given as follows:

$$\mathbb{B} = \{b_\theta : 0° \leq \theta < 360°, \theta \in \mathbb{R}\}.$$

Obviously the beam direction labeled by b_{90} is an optimal solution whereas any optimization approach that is based on some gradient descent method might result in the beam direction labeled by b_0 which is indeed a local minimum in this case. Notice that any search algorithm would have to make a big jump to leave this local optimum.

Contemplating the complexity of the problem of beam orientation of such a simple example gives us an idea about the degree of complexity of the original problem, which usually deals with combinations of beam directions that consist of multiple intensity modulated pixels rather than one single beam with one pixel. For a numerical example see section (An artificial coplanar case study).

Consequently, the problem of beam orientations is rendered to empirical and heuristic search methods. Almost all of these approaches consider a discretized beam space. This is justified from practial and technical points of view. Empirically, it was shown in practice that the radiation beams which are generated by close beam heads have similar results with only marginal slight differences ([EJ03]). Technically, the treatment units are often designed in such a way they can only perform one degree changes of the beam head position ([EJ03]). Of course, such a discretization would make a considerable advantage for any search mechanism since it replaces the continuouse feasible region by a finite set of feasible solutions. The problem of beam orientations can be, in general, formulated as follows:

Definition 3. *For a given finite beam space* $\mathbb{B} = \{b_1, b_2, \ldots, b_n\}$ *where* $n \in \mathbb{N}$*, and a given collection of subsets* $\mathfrak{B} \subset 2^{\mathbb{B}}$ *with a vector function,* $C : \mathfrak{B} \to \mathbb{R}^m$*, where* $m \in \mathbb{N}$*, The **generic beam orientations** (GBO) problem is:*

$$\min\{C(B) : B \in \mathfrak{B}\}. \tag{GBO}$$

The objective function C is actually meant to define a propper quantative quality measurement to evaluate the treatment plan which is produced by the set of beam orientations $B \in \mathfrak{B}$. In general, it is a multi-criteria objective similar to the one that appears in the multicriteria optimization problem (2.1).

At the first sight when looking at the GBO problem, combinatorial optimization would be the first thing that comes to mind to solve it (Appendix A contains a formal definition of the multicriteria combinatorial optimization problem). This gives a rise to the question, how difficult is solving the GBO

problem?

The rest of this chapter is devoted to answer this question in terms of its computational complexity. First, a simplified introduction to complexity theory is introduced, and then it is shown that the problem of beam orientations belongs to the class of $\mathcal{NP}-hard$ problems.

3.2 Hardness of beam orientations

3.2.1 An introduction to complexity theory

Since complexity theory is beyond the scope of this thesis, we will make this introductory section as concise as possible by introducing a general terminology that avoids the rigor found in computer science texts, which would require many definitions and concepts such as Turing machine, language, and so forth.

General terminology

Generally speaking, computational complexity theory, when applied to a given class $\mathcal{C}$ of "reasonable problems"[1] , attempts to classify each given element $P \in \mathcal{C}$ according to its degree of difficulty. Mainly the classification is based on distinguishing between two subclasses: $\mathcal{E} \subseteq \mathcal{C}$ contains "easy" problems and $\mathcal{D} \subseteq \mathcal{C}$ contains "difficult" problems.

To develop such a terminology of classification in a mathematical sense we introduce the following definition:

Definition 4. *Let R be a transitive binary relation defined on a class (set) $\mathcal{C}$. Moreover let $\mathcal{E}$ and $\mathcal{D}$ be defined as follows:*

$$\mathcal{E} = \{P \in \mathcal{C} : (P, Q) \in R, \forall Q \in \mathcal{C}\}.$$

$$\mathcal{D} = \{P \in \mathcal{C} : (Q, P) \in R, \forall Q \in \mathcal{C}\}.$$

*The relation R is said to be a **reduction relationship** on the class $\mathcal{C}$ if both of the subsets $\mathcal{E}$ and $\mathcal{D}$ are simultaneously not empty.*

For the definition of the binary relation and its related properties see Appendix A.

---

[1]In general the class $\mathcal{C}$ could be any given set. We have just found it appropriate to call it this way.

Appropriately enough, a problem $P \in \mathcal{C}$ is said to be an easy problem if $P \in \mathcal{E}$, and it is called difficult if $P \in \mathcal{D}$. Moreover, if $(P, Q) \in R$ then we say that P reduces to, or is not more difficult than, Q.

Let R be a reduction relationship defined on a given class (set) $\mathcal{C}$. This immediately leads to the following propositions:

Proposition 1. *Suppose that $P, Q \in \mathcal{C}$ it follows:*

(i) *If $Q \in \mathcal{E}$ and $(P, Q) \in R$, then $P \in \mathcal{E}$.*

(ii) *If $P \in \mathcal{D}$ and $(P, Q) \in R$, then $Q \in \mathcal{D}$.*

Proof. (i): For all $C \in \mathcal{C}$ we have $(Q, C) \in R$ since $Q \in \mathcal{E}$ by hypothesis. Moreover since $(P, Q) \in R$ and R is transitive, it follows that $(P, C) \in R$ for all $C \in \mathcal{C}$ and hence $P \in \mathcal{E}$.

(ii): Analogous to (i).

$\square$

Proposition 2. *The following statements are equivalent:*

(i) $\mathcal{E} \cap \mathcal{D} \neq \emptyset$.

(ii) $\mathcal{E} = \mathcal{C}$.

(iii) $\mathcal{C} = \mathcal{D}$.

(iv) $R = \{(P, Q) : (P, Q) \in \mathcal{C} \times \mathcal{C}\}$.

(v) R *defines an equivalence relation on $\mathcal{C}$.*

Proof. (i) $\Rightarrow$ (ii): Since $\mathcal{E} \cap \mathcal{D} \neq \emptyset$, then $\exists P \in \mathcal{C}$ such that $P \in \mathcal{E}$ and $P \in \mathcal{D}$. This leads to $(C, P) \in R$ for all $C \in \mathcal{C}$. As $P \in \mathcal{E}$, from proposition (1-(i)) it follows that $C \in \mathcal{E}$ for all $C \in \mathcal{C}$. Hence $\mathcal{C} \subseteq \mathcal{E}$. Furthermore $\mathcal{E} \subseteq \mathcal{C}$ by definition. Thus $\mathcal{C} = \mathcal{E}$.

(ii) $\Rightarrow$ (iii): Consider any $C \in \mathcal{C}$. As (ii) holds, $\forall Q \in \mathcal{C}$ we have $Q \in \mathcal{E}$. Hence $\forall Q \in \mathcal{C}$ we have $(Q, C) \in R$. Thus $C \in D$ and hence $\mathcal{C} \subseteq \mathcal{D} \Rightarrow$ (iii) holds.

(iii) $\Rightarrow$ (iv): For any $P, Q \in \mathcal{C}$, as $\mathcal{D} = \mathcal{C}$, it follows that $P, Q \in \mathcal{D}$. This implies that $(P, Q) \in R$ for any pair $(P, Q) \in \mathcal{C} \times \mathcal{C}$.

(iv) $\Rightarrow$ (v): Trivial.

(v) $\Rightarrow$ (i): From the definition of R we have $\mathcal{E} \neq \emptyset$ and $\mathcal{D} \neq \emptyset$ as well. Thus $\exists P, Q \in \mathcal{C}$ such that $P \in \mathcal{E}$ and $Q \in \mathcal{D}$. This leads to $(P, Q) \in R$. But R is symmetric by hypothesis. This yields $(Q, P) \in R$. From proposition (1-(ii)) it follows that $P \in \mathcal{D}$ as well. Hence $P \in \mathcal{E} \cap \mathcal{D} \neq \emptyset$.

□

From a practical point of view, the principal result of this mathematical setting is the notion of an $\mathcal{NP}$-*complete* problem. Intuitively, an $\mathcal{NP}$-*complete* problem is a computational problem that is as difficult as any reasonable problem. From the terminology presented above one would conceive that the class of $\mathcal{NP}$-*complete* problems will most probably, have similar characteristics as the class $\mathcal{D}$. Then what are the classes $\mathcal{C}$ and $\mathcal{E}$ from a practical point of view? In order to introduce these concepts subject to a precise formulation, we need some additional definitions.

Decision problems and classes $\mathcal{NP}$, $\mathcal{P}$ and $\mathcal{NPC}$

For a given combinatorial optimization problem

$$\max\{f(x) = cx : x \in S\},$$

the associated *decision problem* is:

Given a constant $k \in \mathbb{Z}$, does there exist $x \in S$ such that $f(x) \geq k$?

The range of such a problem is given by the set $\{Yes, No\}$. If the answer of an instance X of a decision problem P is Yes, then there must be an $\acute{x} \in S$ for which $f(\acute{x}) \geq k$. Such an $\acute{x}$ is called *certificate* whereas X is called *Yes-input* for the decision problem P. If all instances, or none of them, of a given problem P are Yes-input, then P is called *trivial*.

Definition 5. *Given an instance $X = \{c, k,$ and a representation of $S\}$, of a decision problem, the length of the binary representation of X is called the **length of the input** X and denoted $|X|$.*

Definition 6. *Given a decision problem P and an algorithm A for solving it, let $T_A(X)$ be the number of elementary calculation steps required to execute the algorithm A for the given input (distance) X. Then $T_A(n) = \sup\{T_A(X) : |X| = n\}$ defines the **computational time** of the algorithm A. An algorithm A is **polynomial**, for the problem P If $T_A(n)$ is bounded from above by a polynomial ie. $T_A(n) = O(n^p)$ for some positive integer p.*

Now it is time to define the class $\mathcal{NP}$ (of reasonable problems $\mathcal{C}$).

Definition 7. *$\mathcal{NP}$ is the class of decision problems with the property that: for any instance for which the answer is Yes, there is a certificate that can be proved in polynomial time.*

Definition 8. *$\mathcal{P}$ is the class of decision problems in $\mathcal{NP}$ for which there exists a polynomial time algorithm.*

Definition 9. *Given $P, Q \in \mathcal{NP}$, P is called **polynomially reducible** to Q ($P \propto Q$) if there is a polynomial algorithm A that transforms any given instance X of P to an instance $A(X)$ of Q such that $A(X)$ is a Yes-input for Q, if and only if X is a Yes-impute for P.*

Definition 10. *The class of $\mathcal{NP}-$**complete**($\mathcal{NPC}$) is the subset of problems $P \in \mathcal{NP}$ such that any problem $Q \in \mathcal{NP}$ is polynomially reducible to P.*

Definition 11. *An optimization problem is called $\mathcal{NP}-$**hard** if the corresponding decision problem is $\mathcal{NP}-$complete.*

Cook [Coo71] showed that the class of $\mathcal{NPC}$ problems is not empty by showing the existence of an $\mathcal{NPC}$ problem. This results in the following corollaries:

Corollary 1. *Let $\mathcal{T}$ be the subset of problems $T \in \mathcal{NP}$ that are trivial. The polynomial reduction ($\propto$) defines a reduction relationship on the class $\mathcal{NP} - \mathcal{T}$.*

Proof. : Trivial. $\qquad\square$

Corollary 2. *Suppose that $P, Q \in \mathcal{NP}$ it follows:*

(i) If $Q \in \mathcal{P}$ and $P \propto Q$, then $P \in \mathcal{P}$.

(ii) If $P \in \mathcal{NPC}$ and $P \propto Q$, then $Q \in \mathcal{NPC}$.

Proof. It follows directly from corollary (1) and proposition (1). $\qquad\square$

Corollary 3. *If $\mathcal{P} \cap \mathcal{NPC} \neq \emptyset$, then $\mathcal{P} = \mathcal{NP}$.*

Proof. It follows from corollary (1) and proposition (2). $\qquad\square$

The list of problems that are belonging to class $\mathcal{NPC}$ is now enormous. Corollary (2) often plays the role of enlarging it. This is usually done by reducing a well-known $\mathcal{NPC}$ problem to the problem in focus. In a moment,

we will experience such a reduction scheme that results in adding the decision problem of beam orientations into the above mentioned list. In order to do this the 0-1 KNAPSACK[2]problem, which is known as an $\mathcal{NPC}$ problem, is reduced to a special case of the decision problem of beam orientations.

Definitions (from 5 to 11) were taken from [Wol98], but they can also be found in most books that address computational complexity. For a general book that covers this topic see [Pap94].

3.2.2   Example construction and reduction

Consider a circular target volume located at the center of a larger phantom circle on which the beams are spread evenly (standard equally spaced beam separation). Assume the irradiated volume contains only one organ at risk R, with arbitrary shape, adjacent to the target volume see figure (3.2).
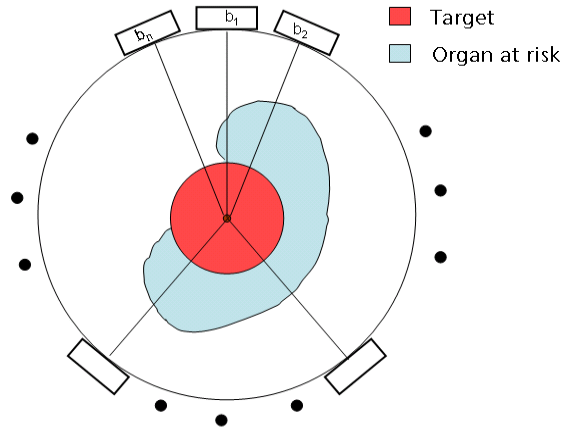
Figure 3.2: Beam orientations as a binary optimization problem.

The beam space is given by the set $\mathbb{B} = \{b_1, b_2, \ldots, b_n\}$ where every beam b_i is assigned a given intensity, resulting in a certain target dose t_i and organ at risk dose r_i. These doses are represented by the following weight (intensity) functions:

$$f_t : \mathbb{B} \to \mathbb{Z}, \text{ where } f_t(b_i) = t_i,$$

$$f_r : \mathbb{B} \to \mathbb{Z}, \text{ where } f_r(b_i) = r_i.$$

---

[2]See lemma 1 for the definition of 0-1 KNAPSACK problem.

Defining $x \in \{0,1\}^n$ by

$$x_i = \begin{cases} 1, & \text{if } b_i \in B \\ 0, & \text{otherwise} \end{cases}$$

for all $i = 1, \ldots, n$, every subset $B \in \mathfrak{B} \subset 2^{\mathbb{B}}$ can be then identified by a binary vector x.

Finally, defining $C : \mathfrak{B} \to \mathbb{Z}^2$ by

$$C(B) = (C_t(B), C_r(B)),$$

where $C_t(B) = -\sum_{i=1}^n t_i x_i = -tx$ and $C_r(B) = \sum_{i=1}^n r_i x_i = rx$, the simplified beam orientations (SBO) problem of this example then reads

$$\min_{x \in X}(-tx, rx), \tag{SBO}$$

where $X = \{x \in \{0,1\}^n : x \text{ corresponds to some } B \in \mathfrak{B}\}$.

The related decision problem (DSBO) is then defined as:

$$\text{Is there } x \in X \text{ such that } -tx \leq k_1 \text{ and } rx \leq k_2, \tag{DSBO}$$

for the given constraints $k_1, k_2 \in \mathbb{Z}$.

Lemma 1. *The (DSBO) problem is $\mathcal{NP}-$complete.*

Proof. We show that 0-1 KNAPSACK $\propto$ DSBO. First we state the definition of the 0-1 KNAPSACK problem.

Definition. Given $V, P \in \mathbb{Z}$ and $v = (v_1, \ldots, v_n)$, $c = (c_1, \ldots, c_n) \in \mathbb{Z}^n$, the 0-1 KNAPSACK problem reads

$$\text{Is there } x \in \{0,1\}^n \text{ such that } cx \geq P \text{ and } vx \leq V?$$

Reduction. Let $k_1 := -P$, $t := c$, $k_2 := V$ and $r := v$. It follows

$$cx \geq P \Longleftrightarrow tx \geq -k_1 \Longleftrightarrow -tx \leq k_1,$$

$$vx \leq V \Longleftrightarrow rx \leq k_2,$$

$\square$

The consequence of lemma 1 is made formal in the next theorem:

Theorem 1. *The problem of beam orientation is $\mathcal{NP}-$hard.*

3.3 Conclusion

It was shown above that the problem of beam orientations is neither convex, when formulated as a continuouse optimization problem, nor polynomially solvable, when modeled as a combinatorial optimization problem. Thus, it has been rendered to the usage of heuristic methods such as artificial neural networks ([HABY99], [RWO99] and [RWO98]), simulated annealing (see e.g. [PLB$^+$01], [ROW99], [SMW$^+$97], [BS93] or [PBX00]), and genetic algorithms ([Ezz96] and [HBM98]) in addition to exhaustive search method ([AHSB99]). Most of these studies, equip the traditional approach with some stochastic optimization algorithm where the optimization of the intensity maps is usually performed for every individual random selection of beam orientations. This requires performing the optimization process which calculates the intensity maps many thousand times, each time for a different selection of beam orientations (see figure 3.3).
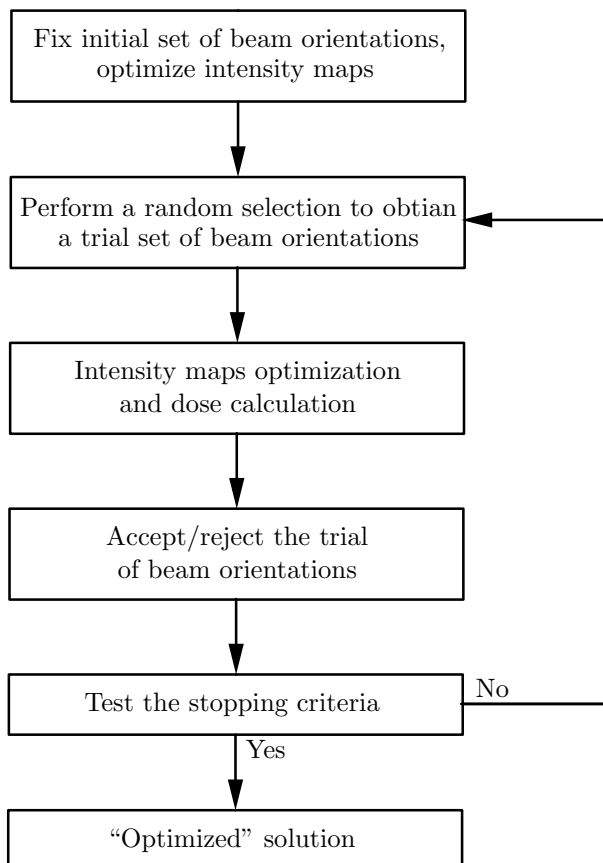
Figure 3.3: Flow chart of a usual optimization approach for beam orientations in IMRT.

The time needed for this purpose, especially in the 3D situation, is not acceptable in practice. Note that here is no place for the human iteration loop which is usually required for intensity maps optimization. Discarding the human iteration loop might result in solutions that are clinically irrelevant (see chapter 2).

Moreover, it is not proper to apply the stochastic methods in the multi-criteria case where intensity map optimization results in a set of Pareto solutions for each set of beam orientations. The planner then needs to search through the set of Pareto optimal manually to decide upon the most desirable solution. This manual search which is required for every set of beam orientations is inconvenient for computerized beam orientations. Hence, in real life, beam orientations in IMRT is still a time consuming trial-and-error search scheme based on intuition and empirical knowledge.

Nevertheless, to facilitate an automated treatment delivery, usually (see e.g. [BSP97], [Bor95], [Bra95] or [GLB94]), an *isocentric* model is used for the choice of beam's setup geometry, i.e., the central rays of the irradiation beams meet in one single point, the *isocenter of irradiation* (see figure 3.4).

Figure 3.4: Intensity maps for different beam directions intersecting at the tumor

Furthermore in most cases coplanar beams are used (see also figure 3.4). The treatment can then be delivered completely by just rotating the gantry

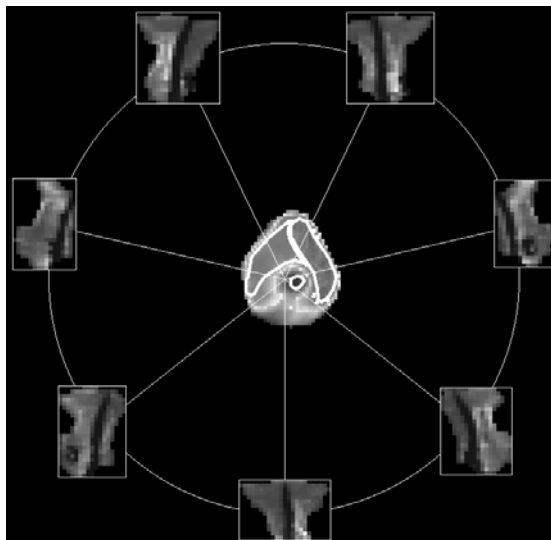around the patient without the need to rotate or translate the treatment couch between different beams. Practically, this leads to shorter overall treatment times which are desirable both for stressing the patient as little as possible and for minimizing the treatment costs.

Typically, in the standard IMRT plans, the set of beam orientations is composed of an odd number of equispaced coplanar beams [Bor99]. Using such beam configurations gave a rise to the question of how many equiangular coplanar intensity modulated beams (IMBS) are required to obtain an optimum IMRT treatment plan [SMW+97]. Webb [Web01] has adopted a realistic conclusion concerning this issue. He said:
"No-one has ever made the ultimate inverse-planning investigation to determine the truly optimum solution because this is, I maintain, impossible job. The reason is that all the variables in an optimization problem are coupled–beam energy, orientation, bixelation, beamwieghts, beashapes etc. A *truly optimal* plan would be that which resulted from allowing a completely free variation in all these parameters, some thing never done in practice. Hence, what usually happens is that workers fix some parameters and search for the best values of others left free. A consequence is that there is an incomplete message delivered about how many beams are needed. It depends on what else you choose to fix and what problem is under consideration"[Web01].

Consequently, in order to optimize the intensity maps within a practical time, the optimization process of the intensity maps is usually performed after fixing all the other parameters including beam orientations (seen chapter 2). Contrary to this approach, in this thesis, the optimization of beam orientations will be carried out under the assumption of pre-fixed intensity maps and a fixed number of beams as well. In either approach, however, the choice of the values for the fixed parameters is critical and has to be clever enough in order to produce clinically relevant solutions (plans) that are somewhat comparable or at least not too far from the truly optimal plans. The thesis will present two different algorithms that approximates the intensity maps for the purpose of optimizing the set of beam orientations in IMRT. The first one approximates beam intensity profiles iteratively instead of doing it for every selection of beam orientations. In the second one the beam intensity profiles for all the beams are firstly approximated and kept fixed throughout the hole running time of the algorithm. In both techniques the intensity maps are calculated, at most, few times, saving a considerable amount of calculation time. A general flow chart of the optimization process of such approaches is shown in figure (3.5).

```
┌─────────────────────────────────┐
│     Initial intensity maps,     │
│   special optimization approach │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│   beam orientations optimization,│
│        by some algorithm        │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│  Optimized set of beam orientations │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│     Intensity maps optimization,│
│        multicriteria approach   │
└─────────────────────────────────┘
                 │
                 ▼
┌─────────────────────────────────┐
│   Approximation of the Pareto front │
└─────────────────────────────────┘
```
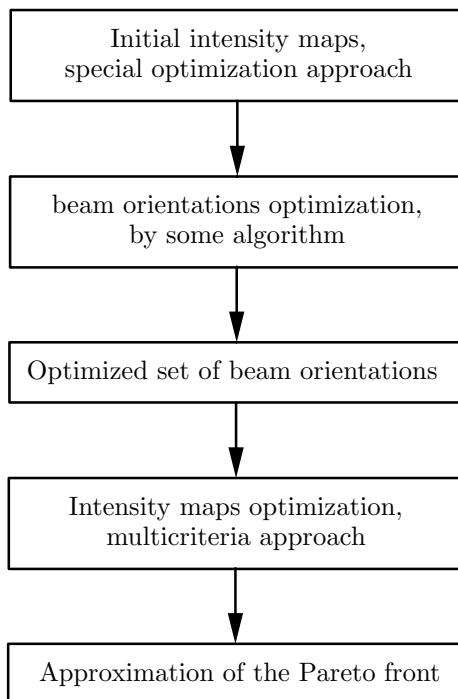
Figure 3.5: Flow chart of the proposed optimization approach for beam orientations in IMRT.

Unlike the flow charts of the usual beam orientations approaches (see e.g. figure 3.3), here the optimization of the intensity maps is independent from the optimization of beam orientations, saving considerable amount of calculations, allowing for reasonable time performance. A property makes the presented approaches distinguishable from the other usual ones.

Finally, an important point that deserves being mentioned here is that there is no unique solution of the IMRT planning problem. This is due to the fact that any solution of the problem depends on the applied mathematical model as well as the applied constraints. Each research group has its own model which emphasizes some features of the optimization while considering only simplified assumptions about others. For example, some groups focus on dose optimization whereas others go for optimizing the biological outcome. After all, it is difficult or even improper to compare different approaches between various groups. However, in beam orientations optimization, it is common to compare the plan resulting from the optimized beam orientations with the standard IMRT plans which are usually composed of a set of an odd number of coplanar equispaced beams (See e.g. [PLB+01], [PX1a], [PX1b],

[PX02], [GWR04] and [WZD$^+$04]). Similarly, to illustrate the efficiency of the algorithms that will be presented in this thesis, the resulting plans are are compared with the standard IMRT plans. Furthermore the results of both algorithms are compared with each another.

Chapter 4

A dynamic algorithm for systematic selection of beam orientations in IMRT

This chapter is devoted for presenting a dynamic algorithm (DA) for systematic selection of beam orientations in IMRT treatment. The algorithm approximates beam intensity profiles iteratively instead of doing it for every selection of beam orientation, saving a considerable amount of calculation time. Every iteration goes from an N-beam plan to a plan with $N+1$ beams. Beam selection is based on a score function that minimizes the deviation from the prescribed dose, in addition to a reject-accept criterion. Although the algorithm does not solve the problem completely, its deviation score function provides the planner with an efficient beam placement guideline which is an alternative to the trial-and-error procedure.

4.1   A mathematical model

The idea behind our model comes from the most recognized technique for beam orientation in conventional conformal radiation therapy. The main features of this technique are summarized by the following two stages:

- **Beam's eye view (BEV):** This stage is meant to identify the appropriate beam directions that are potentially promising for producing a good treatment plan, if they are enclosed in the beam orientation set. Beam directions are labeled good or bad depending only on their geometrical aspects. The planner views the surface model from the position of the radiation source to find out which organs are hit by

the beam being examined. The main criterion is to enclose the target volume completely by the beam and avoiding enclosing organs at risk simultaneously. If it is not possible to find such a beam then the planner goes for minimizing the partial volume of the organ at risk which is covered by the beam. Often modifying the beam portal would be helpful in order to minimize the dose directed towards an organ at risk. Adjusting the beam portal to the form of the target volume could be done using the BEV method. The best beam in the view of the planner is chosen to be the first in the beam orientation set. For the last decade, many modifications where applied to the BEV method, most of them share the binary fashion quality measure to evaluate a given beam direction depending on whether the beam hits organ at risk ([CSP$^+$92], [CRRM99], [GAR$^+$83] and [MCV$^+$92]).

- **Observer's view (OV):** Further appropriate beam directions are found interactively with the Observer's view, where the information of the patient model is provided to the planner using the same scheme as in the BEV method but from an arbitrary point of view enabling him to figure out the volumes of the patient model where the irradiation beams overlap ([SM01]).

Our model, which is fully automated will adapt the same strategy of the BEV and OV methods. First it looks for the best possible irradiation beam for the IMRT planning and then it goes for further appropriate beam directions.

4.1.1   Automated beam's eye view

Unlike the BEV method where the evaluation of a beam direction is carried out by a human, the evaluation in the automated beam's eye view (ABEV) technique is totally computerized. It is based on a score function that minimizes the deviation from the prescribed dose in the target region taking into account the intensity modulation which resulted from solving a specially designed optimizing problem for approximating beamlet intensity profiles. Mathematically speaking, the aim of the ABEV tool is to solve the following problem:

Problem 1. *For a given beam space, $\mathbb{B} = \{b_1, \ldots, b_n\}$ where $n \in \mathbb{N}$, sort the elements of the set $\mathbb{B}$ according to their decreasing performace in producing a good IMRT treatment plan, assuming that the IMRT treatment will be carried out using only one single beam direction.*

In order to have a sorted set of the beam space $\mathbb{B}$ each beam direction is evaluated and ranked using the following two steps procedure:

- **Optimization of beam intensity profiles:** In this step a weighted least square objective function is used for optimizing the delivery of radiation therapy, although a variety of objectives are available and could also be used for the same purpose ([SFOM99]). In the case of the least square fitting the optimizer seeks to minimize the weighted squared differences between the prescribed and the actual doses. Thus for every beam direction $b \in \mathbb{B}$ the intensity of each individual beamlet from the considered beam b is approximated using the following optimization problem:

$$\min \qquad \frac{w_T}{N_T} \sum_{i \in T}(D^p - xd_i)^2 + \sum_{k \in \mathcal{K}} \frac{w_{R_k}}{N_{R_k}} \sum_{i \in R_k}(xd_i)^2 \qquad (4.1)$$
$$\text{subject to} \qquad\qquad x \geq 0,$$

  where w_T and w_{R_k} are the importance factors assigned to the target and organs at risk respectively, D^p is the target prescribed dose, V is the volume occupied by the beamlet, and N_V is the number of pixels belonging to the volume V.

- **Beam evaluation:** After having approximated the intensity profiles of all the beams, the dose distribution of each individual beam direction $b \in \mathbb{B}$ is calculated and its deviation score S_b is evaluated using the following quadratic function

$$S_b = \sum_{i \in T}(D^p - D_i)^2, \qquad (4.2)$$

  where $D_i = \sum_j x_j d_{ij}$ is the total calculated dose value absorbed by voxel i and delivered by the beam direction b.

Now the beam space $\mathbb{B}$ is easily sorted in a decreasing order according to the beams performance by ordering the beams in an increasing order with respect to their deviation score values.

4.1.2   Automated observer's view

As we have seen above, the goal of the OV method is to find a good beam to be added to the already chosen set of beam orientations. Thus the problem of the automated observer's view (AOV) can mathematically be formulated as follows:

Problem 2. *For a given beam space, $\mathbb{B} = \{b_1, \ldots, b_n\}$ where $n \in \mathbb{N}$ and a given set $B \subset \mathbb{B}$, find the beam $b^* \in B^{\complement}$, where $B^{\complement} = \mathbb{B} - B$, such that the orientation set $B \cup \{b^*\}$ will be able to produce the best IMRT treatment plan among those plans that are produced by the orientation sets $B \cup \{b\}$ for all $b \in B^{\complement}$.*

In order to solve this problem we utilize the methods used in the ABEV tool to let it adhere to the goals of the problem. The utilized methods are summarized:

- **Optimization of beam intensity profiles:** Obviously, from the construction of the considered problem, selecting a new beam to be merged into the set B depends heavily on the beams that belong to the set B. More precisely, it is based on how the dose distribution generated by the beam being examined fits into the dose distribution that is produced by the set B. The best fitted beam is the one that reduces the gap between the dose produced by the set B and the prescribed dose to its minimum. Of course this is subject to the beam intensity profile. The utilized version of the optimization problem (4.1) was especially designed to adhere to the treatment goals taking into account the dose distribution that is contributed by the set B. Thus every beamlet from each individual beam direction $b \in B^{\complement}$ is approximated by solving the problem

$$\min \frac{w_T}{N_T} \sum_{i \in T} [D^p - (D_i^d + x d_i)]^2 + \sum_{k \in \mathcal{K}} \frac{w_{R_k}}{N_{R_k}} \sum_{i \in R_k} (D_i^d + x d_i)^2 \qquad (4.3)$$

  where D_i^d corresponds to the dose delivered to voxel i by the beams that belong to the set B. If the resulting intensity x of the considered beamlet is negative, it is set to zero.

- **Merged beam evaluation:** After having approximated the intensity profiles of all the beams $b \in B^{\complement}$, the dose distribution of each orientation set $B \cup \{b\}$ for all $b \in B^{\complement}$ is calculated and its deviation score $S_{B \cup \{b\}}$ is evaluated using a similar evaluation function to the one that has been used to evaluate the ABEV,

$$S_{B \cup \{b\}} = \sum_{i \in T} (D^p - D_i)^2, \qquad (4.4)$$

  where D_i is the total calculated dose value absorbed by voxel i and delivered by the beams belonging to the set $B \cup \{b\}$.

Now the beams belonging to the set $B^{\complement}$ are easily sorted in a decreasing order according to their performance when they are merged with the orientation set B.

4.1.3 Insight

As one can see, the deviation score function in equation (4.2)

$$S_b = \sum_{i \in T} (D^p - D_i)^2,$$

does not take into account the dose delivered to organs at risk. Actually, it is nothing more than the sum of the squared deviation between the prescribed and delivered dose taken only over the target volume. Thus, the higher the delivered dose that does not exceed the prescribed dose the better is the deviation value, even in the case where an organ at risk is highly overdosed. This seems contradictory to the goals of the treatmet plan. Indeed, such a case will never occur since the optimization problem of the intensity profiles in equation (4.1) takes care of reducing the dose delivered to the target, if the corresponding beamlet hits an organ at risk, especially when a relatively large weight is assigned to this organ. To obtain a better understanding of how the optimization of beam intensity is cooperating with the deviation score function to model the beam orientation process, let us consider the special case illustrated in figure (4.1), where the target is given by the rect-
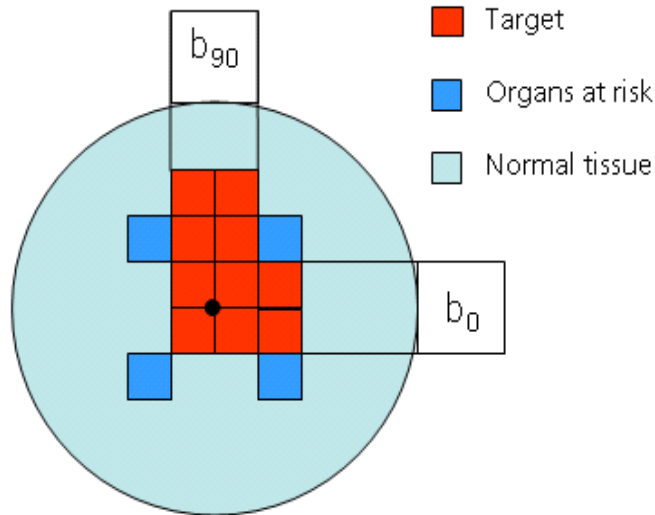

Figure 4.1: A simple artificial example demonstating the processing progress of the ABEV and AOV methods.

angular shape located nearly in the middle of the circle, all other rectangular

shape structures replaced around the target represent an organ at risk and the remaining area inside the circle is normal tissue. We assume that the beams may be arranged at any integer angle on the depicted circle. Thus the beam space is given by the set $\mathbb{B} = \{b_0, b_1, \ldots, b_{359}\}$. Moreover, we consider in this special case that the beam width is equal to the width of two voxels and the irradiation will be carried out using only two beams.

In the above example, one can easily conceive that the optimal two-beam orientation geometry will be given by the set $B = \{b_0, b_{90}\}$. Note that beam b_{270} is not as good as b_{90} since the later one is closer to the target. This empirical judgment is well described by the ABEV and AOV methods presented above. To see this recall the beamlet intensity optimization problem (4.1) which can be rewritten as follows:

$$
\min_{x \geq 0} \left(\frac{w_T}{N_T} \sum_{i \in T} d_i^2 + \sum_{k \in \mathcal{K}} \frac{w_{R_k}}{N_{R_k}} \sum_{i \in R_k} d_i^2 \right) x^2 -
$$
$$
\left(\frac{2w_T}{N_T} \sum_{i \in T} D^p d_i \right) x + \frac{w_T}{N_T} \sum_{i \in T} (D^p)^2.
\tag{4.5}
$$

The optimal solution x^* then can be found by taking the derivative and equalize it with 0. Thus,

$$
x^* = \frac{\frac{w_T}{N_T} \sum_{i \in T} D^p d_i}{\frac{w_T}{N_T} \sum_{i \in T} d_i^2 + \sum_{k \in \mathcal{K}} \frac{w_{R_k}}{N_{R_k}} \sum_{i \in R_k} d_i^2}
\tag{4.6}
$$

As it can be seen from the denominator of equation (4.6), the second term, which corresponds to the organs at risk, imposes a limit on the beamlet intensity. Thus, considering our example depicted in figure (4.1), it becomes obvious that the beamlets belonging to the beam b_{90} and beam b_{270} will be assigned the highest intensity values among all the beamlets. The deviation score of a beam depends only on the dose delivered to the tumor. The higher the beamlet intensity values, the higher the delivered dose to the target and hence the deviation score becomes less. Thus beams b_{90} and b_{270} will be assigned the minimum deviation score values and ranked at the top of the sorted list. Moreover, the beamlet that has to travel a longer distance to reach the tumor will then have a smaller numerator in the fraction of equation (4.6). Consequently it will be assigned a lower beamlet intensity which yields a lower dose distribution on the target and hence a greater deviation value. Thus the ABEV method will definitely prefer beam b_{90} over beam b_{270}.

After determining the first beam $B = \{b_{90}\}$ it is now the time for starting the AOV process where the optimal solution of the beamlet optimization problem (4.3) is given by the following equation:

$$x^* = \frac{\frac{w_T}{N_T} \sum_{i \in T}(D^p - D_i^d)d_i - \sum_{k \in \mathcal{K}} \frac{w_{R_k}}{N_{R_k}} \sum_{i \in R_k} D_i^d d_i}{\frac{w_T}{N_T} \sum_{i \in T} d_i^2 + \sum_{k \in \mathcal{K}} \frac{w_{R_k}}{N_{R_k}} \sum_{i \in R_k} d_i^2}, \qquad (4.7)$$

if the numerator is positive, and $x^* = 0$ otherwise. In this equation, the already delivered dose D_i^d plays a considerable role in reducing the intensity of the beamlet being optimized, especially if the voxels that lie across the beamlet entrance zone were already relatively highly irradiated by the previously chosen beams belonging to the set B. Consequently the considered beamlet will not contribute with a considerable "target dose" and hence will not considerably reduce the value of the deviation score of the beam to which it belongs. Actually this is how the final beam orientation set B avoids having closed beams. For the same reason beams that are potential candidates for producing high dosages on some organ at risk are very likely avoidable as well. For instance if we come back to our example in figure (4.1), all the beamlets that belong to beam b_{270} will be assigned low intensities because of the already delivered doses emitted by the beamlets belonging to beam b_{90}. Consequently the deviation score $S_{B \cup b_{270}}$ will remain close to S_B which in turn causes b_{270} to sink and prevent it from being ranked at the top of the ordered list. Insted $S_{B \cup b_0}$ will win the minimum deviation value since the terms concerning organs at risk in equation (4.7) will be at their minimum for beam b_0 allowing for maximum beamlet intensities and hence a maximum dose contribution over the target.

Thus the ABEV and AOV methods will suggest $B = \{b_{90}, b_0\}$ as a two-beam irradiation geometry for the example depicted in figure (4.1). This result is as the same result as our empirical knowledge has suggested. The AOV method plays the role of pointing out the beams that best complete the dose already delivered to the target with the least compromises concerning organs at risk. This is done with the guidance of the optimization problem (4.3) or equation (4.7).

We should emphasize here that although the construction of the optimization problems (4.1) and (4.3) was gained empirically through the intuitive consideration to capture the main features of the treatment goals, the obtained intensity profiles are intended to be used only during the search for an optimal set of beam orientation. Indeed they might be very different from the final optimal intensity profiles for IMRT treatment which are calculated

using a totally different technique as we have seen in chapter 2 .

Now it is time to present the systematic algorithm which results a promising beam orientation set for producing a good treatment plan.

4.2 A dynamic algorithm

For a given beam space, $\mathbb{B} = \{b_1, \ldots, b_n\}$ where $n \in \mathbb{N}$, and an integer number q, we are seeking a set of beam orientation $B \subseteq \mathbb{B}$, with $|B| = q$, which is able to produce intensity profiles with dose distribution as close as possible to the desired prescribed dose. For technical reasons we assume that the orientation set B is an ordered set. The algorithm is based on two major stages repeated recursively until the desired number of beams is achieved. The first stage is performed at the beginning of each iteration where the intensity for each beamlet is approximated. Then comes the role of the second stage in which the complement of the orientation set $B^{\complement}$ is exhaustively searched in order to find the best beam to be moved into the set B. To do this, the dose distribution for the set B coupled with each individual beam $b \in B^{\complement}$ is calculated and its deviation score $S_{B \cup \{b\}}$ is evaluated using equation (4.4). We will see how these two stages are cooperating with each another through the algorithm presented below.

4.2.1 Algorithm

Stage 0: **Input**

Beam space $\mathbb{B}$ and the required number of beams q.

Stage 1: **Intensity profiles**

a) *Set* $\acute{B} \leftarrow B$ *and* $\bar{B} \leftarrow \mathbb{B}$.
   Set $B \leftarrow \emptyset$.

b) WHILE ($\bar{B} \neq \emptyset$)
       IF ($\acute{B} \neq \emptyset$) THEN
          Let $b :=$ *the first beam in the ordered set* $\acute{B}$,
       ELSE
          Let $b :=$ *an element of the set* $\bar{B}$,
       END IF
       Let $C := \dfrac{\sum_{j=|\acute{B}|+1}^{|B|+|\acute{B}|+1} j}{\sum_{j=1}^{|B|+|\acute{B}|+1} j}$ (I).
       FOR (*Each beamlet belonging to b)*
         DO

$$y := \frac{\frac{w_T}{N_T}\sum_{i \in T}(C.D^p - D_i^d)d_i - \sum_{k \in \mathcal{K}}\frac{w_{R_k}}{N_{R_k}}\sum_{i \in R_k}D_i^d d_i}{\frac{w_T}{N_T}\sum_{i \in T}d_i^2 + \sum_{k \in \mathcal{K}}\frac{w_{R_k}}{N_{R_k}}\sum_{i \in R_k}d_i^2}\quad\text{(II)}$$

         IF ($y > 0$) THEN
           Intensity $x = y$
         ELSE
           Intensity $x = 0$
         END IF
        END DO
       END FOR
       IF ($b \in \acute{B}$) THEN
        Delete beam b from $\acute{B}$
        Add beam b to the orientation set B
       END IF
       Delete beam b from $\bar{B}$
    END WHILE

Stage 2: **Sorting and stopping criterion**

a) *Set* $B^{\complement} \leftarrow \mathbb{B} - B$.

b) FOR (*Each beam* $b \in B^{\complement}$)
    DO
        Set $B_b \leftarrow B \cup \{b\}$.
        Calculate the dose distribution of the set B_b.
        Compute the deviation score $S(B_b)$ of the set B_b where

$$S(B_b) = \sum_{i \in T}(D^p - D_i)^2.$$

    END DO
  END FOR

c) *Let* $\mathcal{B} := \{B_b : b \in B^{\complement}\}$.

d) *Sort the element of the set* $\mathcal{B}$ *in a non-decreasing order according to their deviation values* $S(B_b)$, *where* $b \in B^{\complement}$.

e) *Let Accepted := 0.*

f) WHILE (*Accepted = 0*)
    DO
        Let $B_1 :=$ *the first element in the ordered set* $\mathcal{B}$.
        Let $B^* :=$ *the ordered set of* B_1.
        Recompute the intensity profiles of the beams belonging to the set B^*.
        Calculate the Dose distribution of B^* *and then* $S(B^*)$.
        IF ($S(B^*) > S(B_1)$) THEN
          Delete B_1 *form* $\mathcal{B}$.
        ELSE
          Accepted = Accepted +1.
        END IF
    END DO
  END WHILE

g) *Set* $B \leftarrow B^*$.

Stopping criterion

a) IF ($|B| = q$) THEN
    Output the orientation set B.
  ELSE
    GO TO ***stage 1.***
  END IF

b) END ALGORITHM.

4.2.2   Iterative analysis

During the first iteration, which is a direct application of the ABEV method as we will see soon, the orientation set B is kept empty the way it is defined in stage 0, and hence $D_i^d = 0$. Moreover, since $B = \emptyset$, and hence $\acute{B} = \emptyset$ as well, we have from (I) in stage 1

$$C = \frac{\sum_{j=|\acute{B}|+1}^{|B|+|\acute{B}|+1} j}{\sum_{j=1}^{|B|+|\acute{B}|+1} j} = \frac{\sum_{j=0+1}^{0+0+1} j}{\sum_{j=1}^{0+0+1} j} = 1,$$

where C is defined as the portion of the prescribed dose D^P to be considered while optimizing the intensity profiles of the currently processed beam. The role of this value C will be explained with more details a bit later. Yet since $D_i^d = 0$ and $C = 1$, equation (II) in stage 1 during the first iteration solves merely the optimization problem (4.1)

$$\min \qquad \frac{w_T}{N_T} \sum_{i \in T} (D^p - xd_i)^2 + \sum_{k \in \mathcal{K}} \frac{w_{R_k}}{N_{R_k}} \sum_{i \in R_k} (xd_i)^2$$

$$\text{subject to} \qquad\qquad x \geq 0,$$

for every individual beamlet in each beam $b \in \bar{B} = \mathbb{B}$. After that comes stage two where the dose distribution for every beam $b \in B^{\complement} = \mathbb{B} - \emptyset$ is calculated and its deviation score $S(b)$ is evaluated. Beams are then sorted in non-decreasing order according to their deviation scores. Actually the order found in this first iteration is important for reordering purposes concerning further iterations. Indeed it is the criterion according to which the sets B, $\acute{B}$ and B^* are reordered whenever they appear non empty during the algorithm. Choosing the beam with the minimum deviation score to be added to our orientation set B will terminate this iteration. In a sense, this iteration is somewhat similar to the BEVD and pBEV techniques ([PX1a] and [PX1b]), but with different optimization goals, and hence different objective and score functions.

Every further iteration is meant to move the beam which contributes with the maximum target "missing dose" from the set $B^{\complement}$ into the set B. To see how this is done let us have a closer look at the iterative stages. Each new iteration starts with stage 1 in which the intensity profile for every beam $b \in \mathbb{B}$ is modified. The modification procedure is built in a constructive manner described below. From step a in stage 1, the sets $\mathbb{B}$ and B are copied to the sets $\bar{B}$ and $\acute{B}$ respectively. The set B is then freed to become empty, but only until the intensity profile of the first ordered beam in the set $\acute{B}$ is

recomputed. When step b from stage 1 starts, it processes first the beams belonging to the set $\acute{B}$ taking into account these two points:

- The optimization problem (4.3),

$$\min_{x \geq 0} \ \frac{w_T}{N_T} \sum_{i \in T} [D^p - (D_i^d + xd_i)]^2 + \sum_{k \in \mathcal{K}} \frac{w_{R_k}}{N_{R_k}} \sum_{i \in R_k} (D_i^d + xd_i)^2,$$

considers the beamlet being processed as the last to be used to irradiate the tissue. This often results in a relatively high intensity, especially if there are still more beams to be added, which is the case when the orientation set B is still not yet completed. Consequently, a low intensity will be assigned to the beamlets which intersect the current one and belong to the beams that will be processed later. For an efficient intensity profiles the algorithm considers a "beam share system" of the prescribed target dose. This is established by multiplying the prescribed dose D^p in the above optimization problem by the value C which is given by equation (I) as follows:

$$C = \frac{\sum_{j=|\acute{B}|+1}^{|B|+|\acute{B}|+1} j}{\sum_{j=1}^{|B|+|\acute{B}|+1} j}.$$

For instance, assume that $B = \{b\}$ and the algorithm is about to start its second iteration. In this case B is first copied to $\acute{B}$ and then it is freed. Thus $\acute{B} = \{b\}$ and $B = \emptyset$. And since the condition $\acute{B} \neq \emptyset$ is satisfied in step (1b), beam b will be the first to enter the optimization process with delivered dose $D_i^d = 0$ and a portion of the prescribed dose given by

$$C = \frac{\sum_{j=1+1}^{0+1+1} j}{\sum_{j=1}^{0+1+1} j} = \frac{2}{3}$$

Thus, equation (II) solves the problem

$$\min \qquad \frac{w_T}{N_T} \sum_{i \in T} (\frac{2}{3} D^p - xd_i)^2 + \sum_{k \in \mathcal{K}} \frac{w_{R_k}}{N_{R_k}} \sum_{i \in R_k} (xd_i)^2$$

$$\text{subject to} \qquad\qquad\qquad x \geq 0,$$

leaving the remaining portion of the prescribed dose for the contribution of the upcoming beam. After modifying the intensity of beam b, it is deleted from $\acute{B}$ and added back to its origin the set B. Now, because $\acute{B} = \emptyset$ every beam from the set $\bar{B}$ will enter the optimization process with $C = 1$ and D_i^d equal to the dose account for the first beam $b \in B$ where B is no longer empty during the current iteration.

- At the beginning of this study, the set B was assumed to be an ordered set. Actually, besides the share system presented above, it can easily be seen that the solutions in equation (II) depend heavily on the dose delivered by the beams belonging to the set B. Moreover, computing the intensity profile of a beam $b \in B$ requires knowing the intensity profiles of all its successor beams belonging to the set B. Thus different orders of these successors will result in different solutions. To overcome this situation, the set B was assumed to be an ordered set sorted according to the order found during the first iteration.

Second, after having reprocessed the last beam in $\acute{B}$, stage 1 starts processing the beams belonging to $B^{\complement}$ one by one but this time without giving any importance to their order. Here all the beams belonging to the set $B^{\complement}$ are treated equally with respect to C, which is equal to one, and D_i^d which accounts for the beams belonging to the orientation set B.

Now it is time to start the second phase of the current iteration where the dose distribution for every set $B_b = B \cup \{b\}$, $b \in B^{\complement}$, is computed and the value of its deviation score $S(B_b)$ is calculated. Finally, and unlike in iteration one, where the beam with the minimum deviation value is directly moved into the set B, the selection criteria is slightly more complicated. Here the best beam b^*, which corresponds to the beam with the minimum deviation score $S(B_{b^*})$, is not moved into B immediately. Instead, it is tested by a reject-accept procedure that accomplishes the following three steps:

1. Assuming $B = \{b_1, b_2, \ldots, b_k\}$ where $k < q$, the set B_{b^*} then is precisely written as follow:

$$B_{b^*} = B \cup \{b^*\} = \{b_1, b_2, \ldots, b_k, b^*\}.$$

   Since B is an ordered set we know that $S(b_1) \leq S(b_2) \leq \ldots \leq S(b_k)$. But the inequality $S(b_k) \leq S(b^*)$ does not necessarily hold. This gives rise to the definition of the following set:

$$B_{b^*}^o = \{b_1, b_2, \ldots, b_{l-1}, b^*, b_l, \ldots, b_{k-1}, b_k\},$$

   where $S(b_1) \leq S(b_2) \leq \ldots \leq S(b_{l-1}) \leq S(b^*) \leq S(b_l) \leq \ldots \leq S(b_{k-1}) \leq S(b_k)$. Thus $B_{b^*}^o$ is nothing more than the set B_{b^*} after having it reordered.

2. Since the calculation of the intensity profiles cosideres an ordered set, if the order of the sets $B_{b^*}^o$ and B_{b^*} are different, the intensity profiles for those beams belonging to the ordered set $B_{b^*}^o$ are recalculated with

respect to the order of B^o_{b*}. Now the dose distribution of the set B^o_{b*} is computed and its deviation score $S(B^o_{b*})$ is evaluated.

3. The decision whether b^* is accepted or rejected depends on how the deviation value $S(B^o_{b*})$ differs from $S(B_{b*})$. If $S(B^o_{b*}) \leq S(B_{b*})$, then b^* is accepted. Otherwise it is rejected to give the chance for the next best to enter the reject-accept procedure.

The first beam that passes the reject-accept procedure is added to the orientation set B, and the algorithm starts a new iteration. It terminates when the recommended number of beams is achieved. Figure (4.2) contains a flow chart of the presented DA.
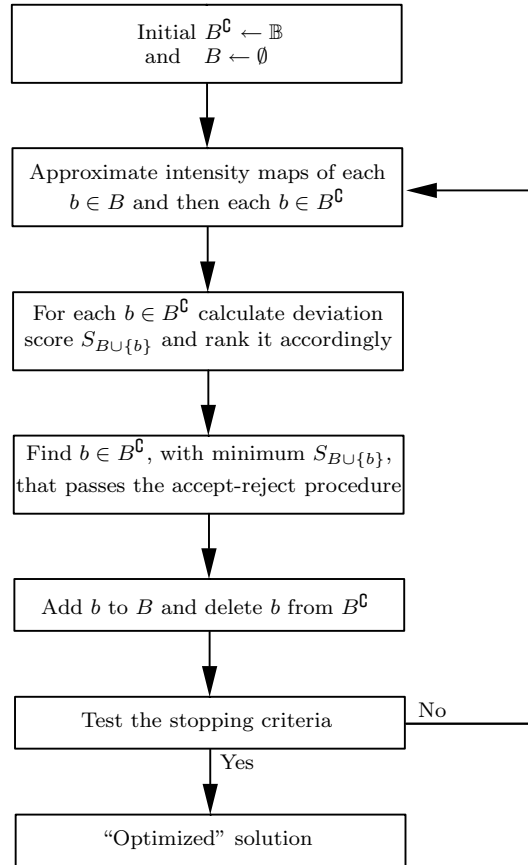

Figure 4.2: Flow chart of the dynamic algorithm.

One point that deserves mentioning here is that this algorithm could easily be constructed in such a way that it performs the reject-accept procedure several times until a couple of beams are accepted. The beam with the best

score, among the accepted ones, is then moved into the set B. These accepted beams could form a guideline for the planner just in case some alternative solutions are required.

Now after obtaining a set of beam configuration, the final optimal intensity profiles for IMRT treatment are computed using a multi-criteria optimization solver developed at ITWM (See chapter 2).

To illustrate the efficiency of the algorithm, it has been applied to one artificial example, where optimality is trivial, and three clinical cases: a prostate carcinoma, a tumor in the head and neck region and a paraspinal case. In all of the real clinical cases, the judgment of the algorithm's efficiency was based on the comparison between 2 types of optimization: (1) beam intensity profiles were optimized for coplanar equidistant beams; (2) beam intensity profiles were optimized for the coplanar orientation set which was found by the presented algorithm. To evaluate the quality of the treatment plan, a depiction of the dose distribution and dose-volume histogram (DVH) was used for the target and each of organs at risk.

4.3   Examples

4.3.1   An artificial coplanar case study

The construction of this simple example was motivated to highlight the adherance of the algorithm to the human intuitive and logical sense. It consists of a cross shaped target surrounded by four squared critical structures located at the angles formed by the cross in a circular phantom. An illustration of the example is depicted in figure (4.3).

In a coplanar case, every beam direction could be represented by its gantry angle which is assumed to be an integer in this example. Thus, the beam space set $\mathbb{B}$ is written as follows:

$$\mathbb{B} = \{b_i : 0 \le i < 360, i \in \mathbb{N}\}.$$

Importance factors and all other parameters are given in table (4.1).

The result of applying the dynamic algorithm on the prescribed sample example is shown in figure (4.4) which contains four curves representing the deviation score for beam directions through four consecutive iterations. As one can see, the curve representing iteration one shows global minima at
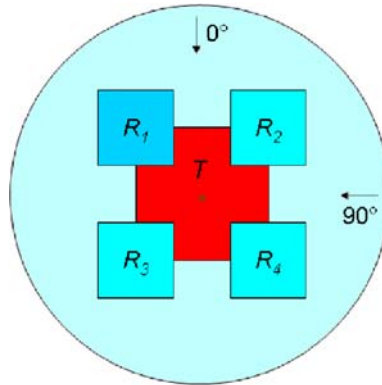
Figure 4.3: A simple artificial example.

Organ / Structur	w	Upper bound	Lower bound
Target	1.0		50
Organ at risk 1	0.6	20	
Organ at risk 2	0.5	25	
Organ at risk 3	0.5	25	
Organ at risk 4	0.5	25	
Unclassified tissue	0.2	50	

Table 4.1: Importance factors and ideal doses chosen for the target and organs at risk.

gantry angles 90° and 180°. This makes intuitive sense since theses beams are the most exposed beams to the target with minimum dose delivered to organ at risk R_1 which has the lowest upper bound and highest importance factor among all organs at risk. Thus, the first iteration results in the beam orientation set $B = \{90°\}$. The second curve resulted from iteration two shows a gap at 90° and a global minimum at gantry angle 180°. Discarding 90° is very logical since it has been already chosen and chosing gantry angle 180° is intuitivily obvious. The remaining curves coresponding to iterations three and four show global minimum at 270° and 0° respectivily. Thus the algorithm results in the following set of beam orientations:

$$B = \{90°, 180°, 270°, 0°\}.$$

One point that deserves mentioning here is that increasing the importance factor for the unclassified tissue avoids having opposite directions in the final resulted beam configration.
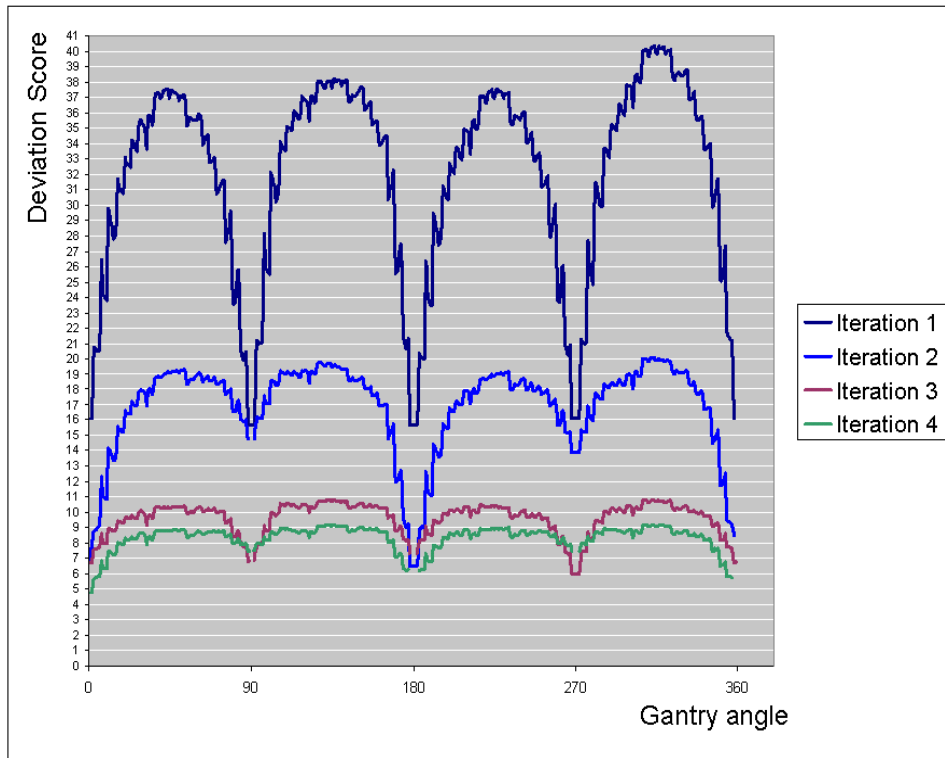
Figure 4.4: Deviation scores for gantry angles resulted from applying the DA on the sample example.

4.3.2   Prostate carcinoma

In this study, the irradiated volume contained a two tumor parts (boost and target) grown in the prostate area and surrounded by four organs at risk; bladder, rectum, right femur and left femur. The couch angle was set to $0°$, whereas the gantry angle was allowed to vary from $0°$ to $360°$ in $5°$ of increments. Five equiangular spaced beams plan were used in comparison with the five coplanar beams plan optimized using the algorithm presented above. The dose distributions and the dose volume histograms of both plans are shown in Fig (4.5) and Fig. (4.6) respectively.

In comparison with the plan obtained using equiangular beams, significant improvements in the dose distribution regarding all clinical structures were reported. On the one hand, as one can see from Fig (4.6), the use of the optimized beams yielded a slightly better dose uniformity in the target. On the second hand the values of the multicriteria objective function were all
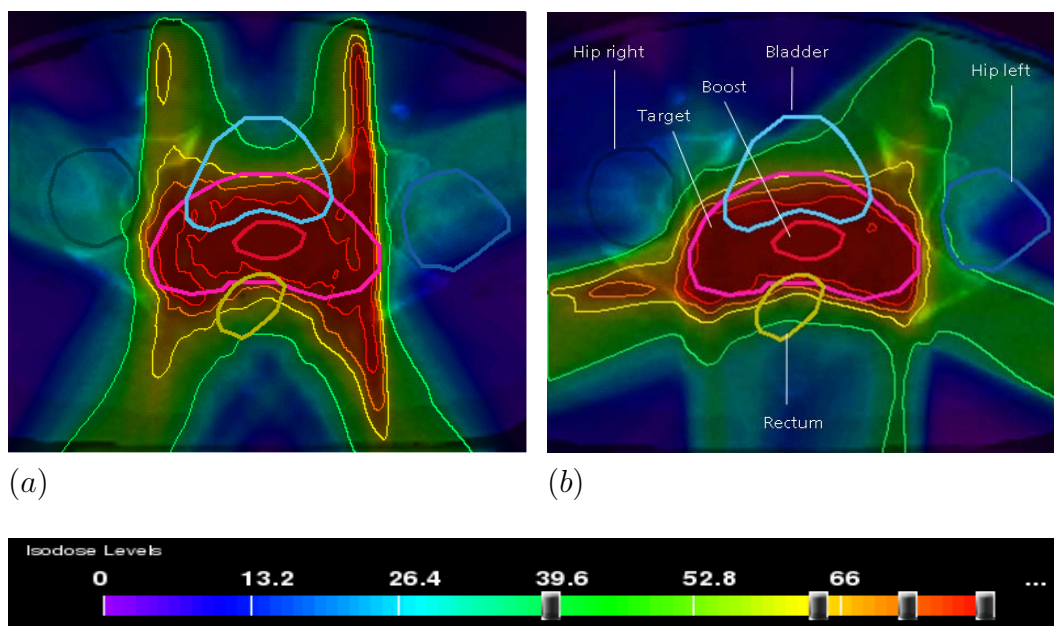
Figure 4.5: The dose distribution of two multicriteria IMRT prostate tumor treatments corresponding to (a) five equiangular spaced beams and (b) five optimized beams obtained by the DA.
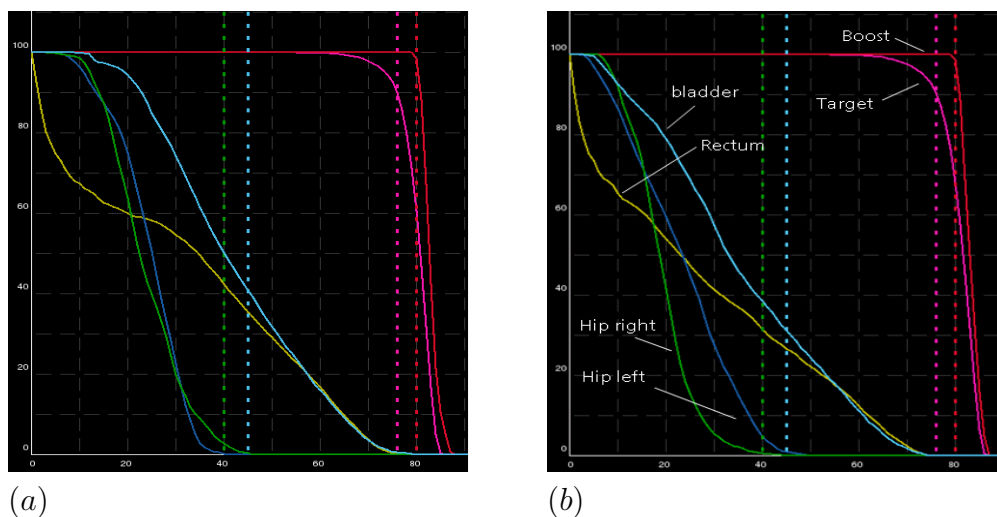

Figure 4.6: Dose volume histograms of two multicriteria IMRT prostate cancer treatments corresponding to (a) five equiangular spaced beams and (b) five beams optimized by the DA.

reduced by 5%, 4%, 6% and 7% for the bladder, rectum, left femur and right femur, respectively. Moreover, the shape of the dose distribution resulting from the optimized beams is better fitted to the shape of the tumor than the one obtained by the equiangular beams as demonstrated in Fig (4.5).

4.3.3 Head and neck tumor

The irradiated volume in this case contained again two tumor parts, spreading this time in the head and neck region among three organs at risk; spinal cord, brain stem, and parotid gland. Unlike the prostate case where the comparison was performed on two sets of beam orientations with the same number of beams; in this case a seven coplanar equiangular beams plan was compared with the plan obtained by using only six coplanar beams optimized using the DA. The dose distributions of these two plans and their dose volume histograms are shown in Fig. (4.7) and Fig. (4.8) respectively.
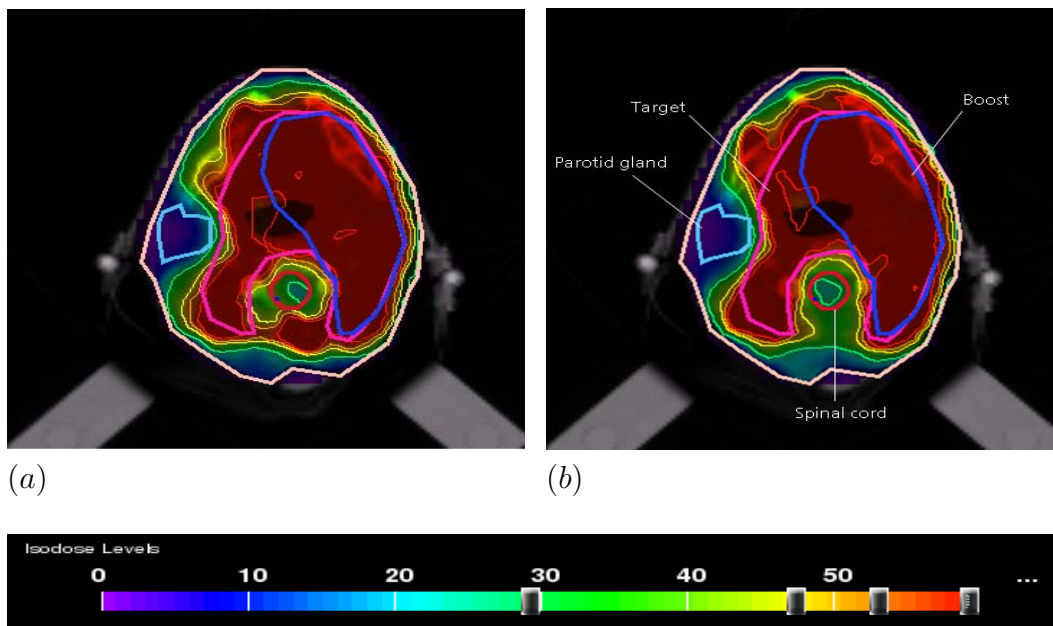

Figure 4.7: The dose distribution of two multicriteria IMRT treatments, for head and neck tumor, corresponding to (a) seven equiangular spaced beams and (b) six beams obtained by the DA.

As can be seen from Fig (4.8), the use of the optimized beams resulted in a reduction of the maximum dose to the spinal cord and brain stem against margin maximum dose increment delivered to the parotid gland. One can also see from Fig. (4.7) that some unnecessary hot spots in the irradiated
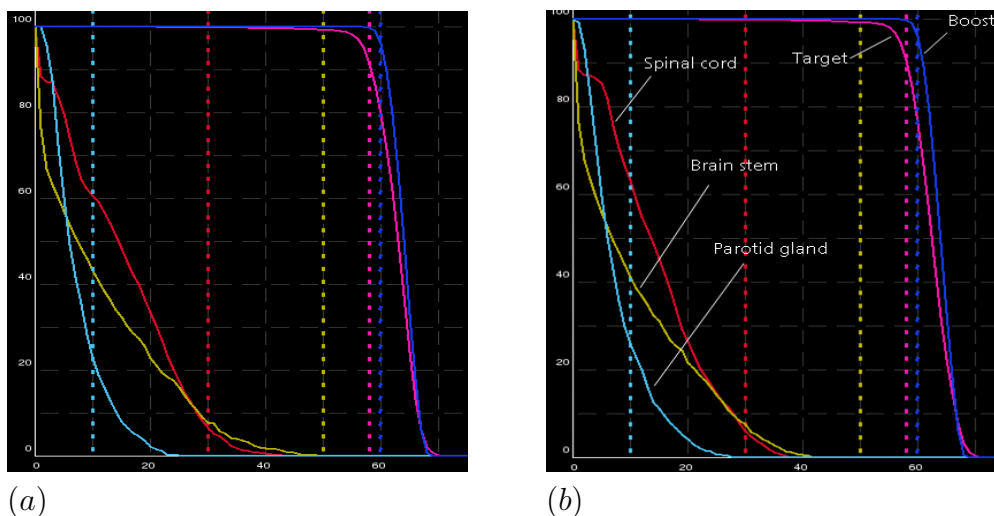
Figure 4.8: Dose volume histograms of two multicriteria IMRT treatments, for head and neck tumor, corresponding to (a) seven coplanar equiangular spaced beams and (b) six beams obtained by the DA.

volume were eliminated, allowing for better tumor dose shaping than the one produced using the equiangular spaced beams. Consistent results were reported by the numerical realization where the values of the multicriteria objective function were improved by 4% and 5% for the spinal cord and the brain stem, respectively, against worsening the parotid gland by 9%.

Although optimizing the set of beam orientations in this case has not considerably improved the treatment plan, having a similar treatment plan by using a fewer number of beams is a matter of considerable importance with respect to clinical and practical aspects. It is always desirable to reduce the number of irradiating beams to its minimum without compromising the quality of the treatment. In fact, using a large number of beams may have the undesirable consequence of spreading low doses to a large volume of the normal tissue which in turn tends to increase the risk of second cancers (see e.g [Hal06] or [HW03]). It may also increase treatment delivery time which may increase patient discomfort in addition to the possibility of increasing potential error in terms of patient movement.

4.3.4 Paraspinal case

In this example, the irradiated volume contained one target spreading around the spinal cord among three organs at risk; esophagus, left lung and right

lung. Similar to the head and neck example, in this case a seven coplanar equiangular beams plan was compared with the plan obtained by using six coplanar beams optimized using the DA. The dose distributions of these two plans and their dose volume histograms are plotted in Fig. (4.9) and Fig. (4.10) respectively.
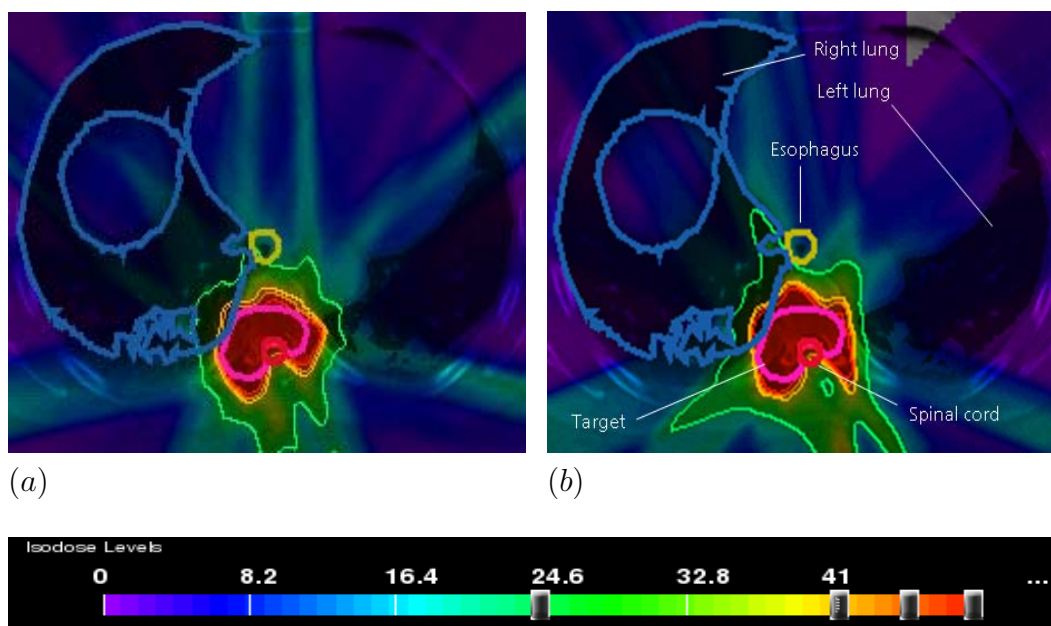

Figure 4.9: The dose distribution of two multicriteria IMRT treatments for paraspinal tumor, corresponding to (*a*) seven equiangular spaced beams and (*b*) six optimized beams obtained by the DA.

As can be seen from these figures, significant improvments were reported in two rogans at risk; esophagus and left lung where it is hard to note any differences regarding the DVHs coresponding to right lung, the spinal cord or the target. Improvemnts regarding the numerical reilization were reported as follows: 12% for the esophagus and 33% for the left lung where as all the others values were almost equal.
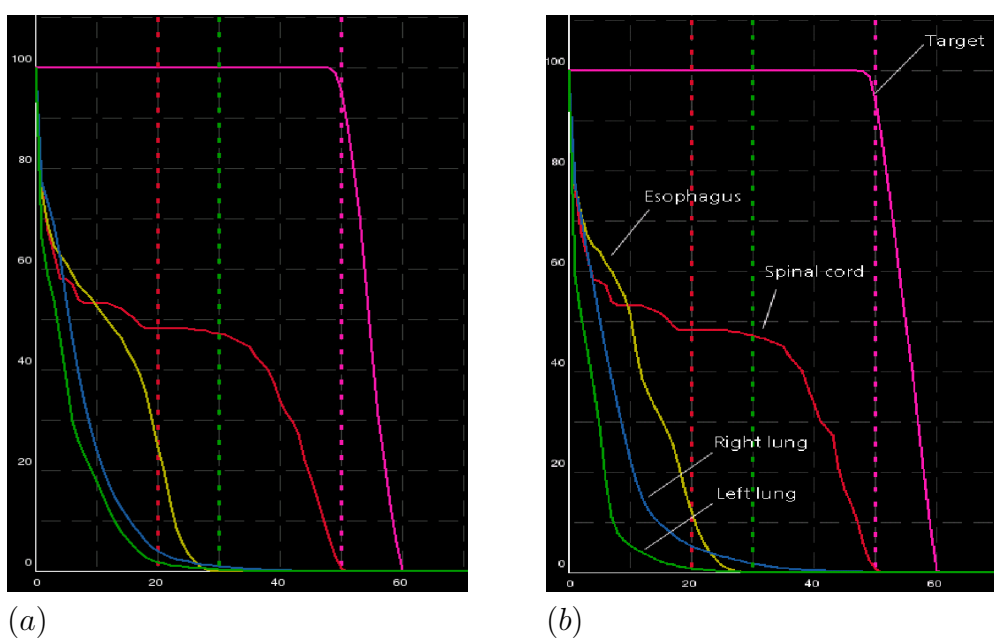
Figure 4.10: Dose volume histograms of two multicriteria IMRT treatments for paraspinal tumor, corresponding to (a) seven coplanar equiangular spaced beams and (b) six beams obtained by the DA.

Chapter 5

DVH-Evaluation scheme and an evolutionary algorithm for beam orientations in IMRT

5.1 Introduction and motivation

As we have seen from the preceeding chapters, the process of designing a satisfactory treatment plan for IMRT is difficult. A considerable part of this difficulty is due to the unintuitive prerequisite of defining appropriate criteria for plan evaluation. For better understanding it is helpful to distinguish between *plan(s) judgment*, and *plan optimization*. Plan(s) judgment implies some dose distribution evaluation scheme in which the planer quantifies the clinical quality of some given plan(s). This is usually utilized, in addition to the isodose displays, by the use of the DVH concept which summarizes different important aspects of the dose distribution that are difficult to be conceived by only considering the isodose displays. Actually the DVH is nothing more than condensing the information present in the dose distribution by neglecting the geometrical information. However, the DVH concept has been recognized as a quantitative measurement tool for dose distribution evaluation, see e.g. [Cha88], and consequently plan judgment. If we restrict ourselves to mathematical terms, where the problem of IMRT is, in general, considered as a constraint optimization problem, plan optimization means the identification of the plan for which the objective function attains its optimal value.

Yet for a successful IMRT treatment plan it is required to formulate a proper objective function such that its optimal value is achievable in practi-

cal calculation time and additionally corresponds to a plan that is assessable to be of a high clinical quality. Unfortunately, the reality is different. Many times it happens that the optimal solution of the plan optimization does not even correspond to an acceptable treatment plan. This often occurs because plan optimization is not always able to capture the criteria that are used by the clinicians when they evaluate and judge plans.

To parallelize plan optimization with plan judgment criteria there is a need to incorporate the DVH concept properly into the optimization process such that plan optimization results in optimal solutions of high clinical quality.

This chapter is devoted to present a mathematical model for optimization of beam orientations in IMRT that is totally based on a DVH evaluation scheme.

5.2 A Mathematical model

5.2.1 The DVH concept as a function

The DVH depicts for each volume of interest (VOI) the volume percentage v that receives a certain dose d as a function of the dose, cf. figure (5.1).
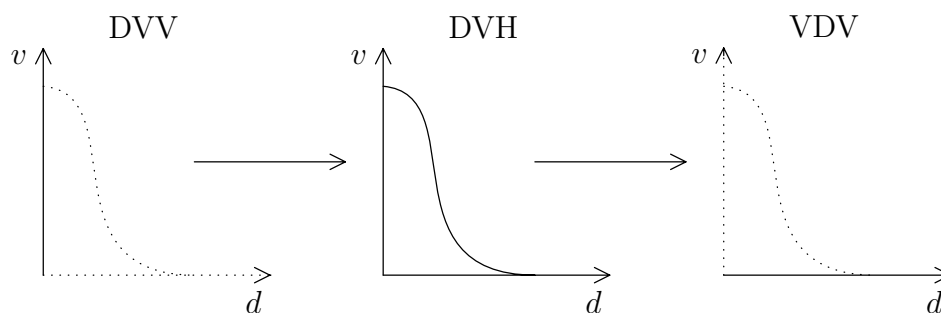

Figure 5.1: Represents the DVV (discrete), DVH (continouse) and the VDV (discrete) depicted from left to right respectively.

In order to plot the DVH of some VOI first a nonnegative real vector, we call DVV, is calculated. The *dose volume vector* (DVV) represents the discrete form of the DVH. For some VOI let $M = \{0, 1, 2, \ldots, m\}$ be a finite set with m representing the smallest integral that is greater than or equal to the maximum dose the VOI has received. Thus the DVV is an n-dimensional

vector, where $n = m + 1$, with the ith component is equal to the volume percentage that receives the dose $i - 1$. Obviously the first component is equal to 100 which then decreases as the component's index increases until it becomes 0 for the nth component. Although the dose variable $d \in M$ is an integral here, the corresponding volume v_d could obtain any real value from the interval $[0, 100]$ with the property $v_d \geq v_{d+1}$.

Now the DVH is easily obtained by just connecting the points of the d-axis and the points representing the DVV, see figure (5.1).

Another entity, we will also use to evaluate the treatment plan is the *volume dose vector* (VDV) which represents the discrete form of the (inverse of the) DVH. In other words it is a 101-dimensional vector with i^{th} component – where $i = 0, 1, \ldots, 100$ stand for an integral volume persentage of the VOI – gives the maximum dose d as a function of the volume persentage i receiving at most the dose d. One can easily obtain the VDV from the DVH, see also figure (5.1).

5.2.2 DVH evaluation scheme

Generally, the reaction of the tumor or organ at risk (OAR) is a function of the volume or percentage of the volume subject to each level of dose it receives. In fact, physicians quantify clinical structures or organs at risk in two groups according to the way they react when they are exposed to radiation. In some organs, destroying one individual functional unit would be enough to disable the whole organ. Such organs, known as chain organs, can handle relatively high dose of radiation. But strictly speaking, even one single hot spot in such an organ has to be avoided. Conversely, organs belonging to the other group, known as parallel organs, show distinct volume effect. The tolerance dose for such organs depends on the fraction of the irradiated volume. The tolerance dose is considerabley high if a small fraction is irradiated compared with the tolerance dose for larger irradiated fractions. For more about chain and parallel organs see [GN88], [WTM87], or [Wol84].

A successful DVH evaluation has somehow to take into account the aspects mentioned above. We propose the following DVH evaluation scheme:

$$f(DVH) = \alpha \sum_{d=1}^{m} dv_d + (1 - \alpha) \sum_{v=1}^{100} vd_v, \qquad (5.1)$$

where $0 \leq \alpha \leq 1$ is a tissue-specific parameter, depending on the biological properties of the individual specified tissue. Indeed, biological studies beyond

the scope of this thesis have to be done in order to determine an accurate α for the different clinical structures. However we believe that an accurate α for the target is 0 or close to it and for organs at risk, especially for those that are classified as chain organs, would be 1 or close to it. Note that for α close to 0 more emphasis will be placed on the volume and minimizing (maximizing) the function $f(DVH)$ will be reflected by minimizing (maximizing) the volume irradiated. On the other hand, if α close to 1 more emphasis will be placed on the maximum dose and hence minimizing the function $f(DVH)$ will naturally minimize the maximum dose received by the organ at risk. An accurate α balances the optimization process between the dose and volume according to the biological properties of the corresponding clinical structure.

5.2.3   Treatment plan evaluation

To parallelize plan optimization with plan judgment a given treatment plan, represented by its dose volume histograms (DVH_T for the target and DVH_{R_k}, where $k = 1, \ldots, K$, corresponding to the different OARs), can be evaluated by the following objective function:

$$F = W_T \ f(DVH_T) - \sum_{k=1}^{K} W_{R_k} \ f(DVH_{R_k}), \qquad (5.2)$$

where f is the DVH evaluation function defined in equation (5.1) and W_T, W_{R_k} for $k = 1, \ldots, K$ are nonnegative weights (importance factors) corresponding to the target and different OARs respectively.

5.2.4   Intensity maps approximation

In this section the DVH evaluation scheme is utilized for optimizing (approximating) the intensity maps. For one unit of intensity each individual beamlet results in several mini-DVHs (mDVHs), corresponding to the different VOIs. Note that in such a case the maximum received dose D usually is not more than 1 unite of intensity since it is only a portion of the intensity, see figure (5.2).

Another value which is of our interest is the maximum portion V of the VOI that receives a dose greater than 0, see also figure (5.2).
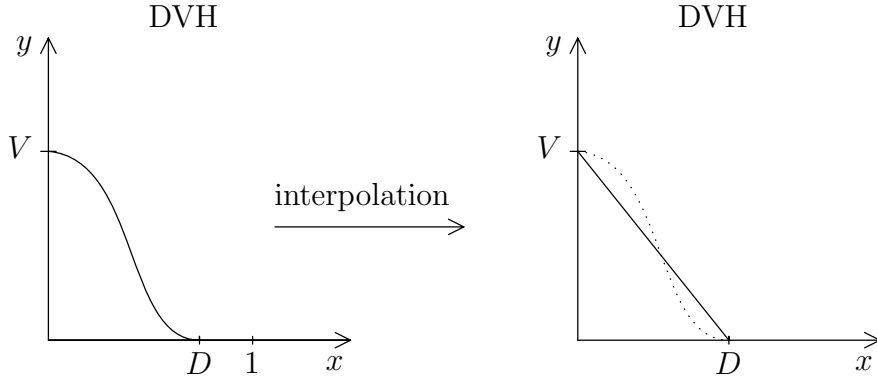
Figure 5.2: Represents, on the left hand side, a mini-DVH for one single beamlet with one unit of intensity. The right hand side depicts an approximation of the DVH obtained by linear interpolation.

Mini-DVH approximation

To utilize the mDVH in our mathematical context, we approximate it by a linearly interpolated function $h(x)$ between the two points $(D, 0)$ and $(0, V)$ as shown in figure (5.2). Thus the mDVH is represented by the approximation function

$$y = h(x) = \frac{-V}{D}x + V, \text{ where } x \in [0, D]. \tag{5.3}$$

Obviously the function h is bijective[1] and its inverse is given by

$$x = h^{-1}(y) = \frac{-D}{V}y + D, \text{ where } y \in [0, V]. \tag{5.4}$$

Applying the DVH evaluation scheme (5.1), using its continuous formulation, to the mDVH presented by $h(x)$ results in

$$f(mDVH) = \alpha \int_0^D xh(x)dx + (1 - \alpha) \int_0^V yh^{-1}(y)dy. \tag{5.5}$$

By substituting $h(x)$ and $h^{-1}(y)$ from equations (5.3) and (5.4) respectively and performing some elementary calculations we obtain

$$f(mDVH) = \frac{1}{6}\alpha V D^2 + \frac{1}{6}(1 - \alpha)V^2 D, \tag{5.6}$$

or simply

$$f(mDVH) = \alpha V D^2 + (1 - \alpha)V^2 D. \tag{5.7}$$

So far it was assumed that the mDVH is produced by a beamlet having one unit of intensity. One can easily verify that for a beamlet with intensity

---

[1] A bijective mapping between two sets is one-to-one and onto.

x –from now and on x is a variable representing the intensity of a beamlet and no longer the x axis in figure (5.2)– the evaluation function, denoted $f_x(mDVH)$ in this case, obtains finally the form

$$f_x(mDVH) = \alpha V(Dx)^2 + (1-\alpha)V^2(Dx). \tag{5.8}$$

Beamlet intensity optimization

As it was shown in chapter 1, the aperture of a radiation beam is created in such a way that it resembles the shape of the tumor. Thus each individual beamlet will certainly hit the tumor and hence produce some tumorodical mini dose volume histogram (mDVH$_T$). Additionally, it might also hit an OAR, producing other mDVH$_R$ concerning the corresponding organ. Setting in equation (5.8) α to 0 for the target and to 1 for organs at risk, as it was suggested in section (5.2.2), results in the evaluation function

$$f_x(mDVH_T) = V_T^2 D_T x \tag{5.9}$$

for the target and

$$f_x(mDVH_R) = V_R D_R^2 x^2 \tag{5.10}$$

for an organ at risk.

Furthermore, delivering a high dose, equal to the prescribed dose, to the target volume is usually done by using a set of beams in which each beam contributes to the irradiated volume by some portion of the prescribed dose. This indeed puts an upper bound U on the intensity variable x.

On the one hand the objective of delivering a high dose to the tumor proposes maximizing the following target objective

$$f_x(mDVH_T) = V_T^2 D_T x, \tag{5.11}$$

where $x \in [0, U]$.

On the other hand, for sparing an OAR from unnecessary overdosage we minimize the objective

$$f_x(mDVH_R) = V_R D_R^2 x^2, \tag{5.12}$$

where $x \geq 0$.

Consequently, the intensity optimization problem can be defined as follows:

$$\min \; F(x) = \sum_{k=1}^{K} w_{R_k} V_{R_k} D_{R_k}^2 x^2 - w_T V_T^2 D_T x,$$
$$x \in [0, U], \tag{5.13}$$

where $w_{R_k}, k = 1, \ldots, K$, and w_T are nonnegative weights (importance factors) corresponding to the different OARs and the target respectively.

If we define $\bar{x}$ by the equation

$$\bar{x} = \frac{w_T V_T^2 D_T}{2 \sum_{k=1}^{K} w_{R_k} V_{R_k} D_{R_k}^2},$$

then the optimization problem (5.13) has the solution:

$$x = \begin{cases} \bar{x}, & \text{if } \bar{x} \leq U; \\ U, & \text{otherwise.} \end{cases} \tag{5.14}$$

5.2.5 Optimization of beam orientations

As is was mentioned in chapter 3, the optimization of beam orientions is carried out under the assumption of fixed intensity maps. These intensity maps are approximated by solving the optimization problem (5.13) for every beamlet in each individual beam. Thus every combination of beams will produce a certain plan with its corresponding DVHs for the different VOIs. Hence for a given beam space $\mathbb{B} = \{b_1, \ldots, b_n\}$, where $n \in \mathbb{N}$, analogously to equation (5.2), our objective function $F : 2^{\mathbb{B}} \to \mathbb{R}$ reads:

$$F(B) = W_T \; f(DVH_T) - \sum_{k=1}^{K} W_{R_k} \; f(DVH_{R_k}), \tag{5.15}$$

Additionally, for the methods presented here, it is assumed that the number of beams belonging to the set of beam orientations, denoted q, is given in prior. Thus the problem of beam orientation becomes as follows:

Problem 3. *For a given beam space $\mathbb{B} = \{b_1, \ldots, b_n\}$, where $n \in \mathbb{N}$, find the orientation set $B^* \in \mathfrak{B} = \{B \in 2^{\mathbb{B}} : |B| = q \in \mathbb{N} \text{ and } q < n\}$ for which*

$$F(B^*) = max\{F(B) : B \in \mathfrak{B}\}.$$

While problem (3) constrains the feasible solutions to a fixed number of beams, the solutions' hyperspace remains huge and the choice of an exhaustive search maintains impractical. For instance, an exhaustive search of the best 6-beam configuration in a pool of 72 different beams would require intercomparison of $\binom{72}{6} = 156,238,908$ IMRT plans. The rest of this chapter is devoted to present an evolutionary algorithm suitable for approximating the optimal solution of the problem in consideration.

5.3   An evolutionary Algorithm

5.3.1   Introduction

Evolutionary algorithms (EA) are a family of computational models inspired by evolution. They mimic aspects of natural selection and differential reproduction. While a thorough review of the EAs and their strategies are beyond the scope of this thesis, the fundamental principles are summarized. In our context, when the DVHs of q different beams are merged together to compose a potential plan, such a plan is considered as an *individual*. A collection of such individuals produces a *population*. Every individual in the population is ranked according to its *fitness* which is a kind of scoring technique based mainly on the objective function.

EAs, in addition to the solution space which is the space of the actual solutions, use another space called the search space and contains coded solutions, or *genotypes*, of the actual solutions called *phenotypes*. Thus each individual consists of a genotype belonging to the search space and a corresponding phenotype belonging to the solution space. Genotypes must be mapped onto their phenotypes before the fitness of each solution is evaluated.

The basic idea of EAs is that individuals in the population are recombined to produce new individuals using reproduction operators that mimic those that occur in nature, such as mutation, recombination and selection. While the reproduction operators conduct the genotypes randomly and blindly, the selection of the phenotypes guided by their fitnesses determines which genotypes will be promoted for reproduction. Hence, evolutionary search is unquestionably pushed towards areas in the search space that contain better solutions, making EAs a powerful model for solving large combinatorial optimization problems.

The rest of this chapter will be devoted to implement an evolutionary al-

gorithm suitable for the problem of beam orientation presented above (problem 3).

The terminology presented here can be found in most books that address the topic of EAs or genetic algorithms (See e.g. [Ben99] or [Mic98]).

5.3.2 Implementation of evolutionary algorithm

The purpose of the EA is to select a set of beams, each with the corresponding intensity maps, which together constitute a plan. Each plan produces a particular DVHs which can then be evaluated and hence a fitness value is assigned accordingly. The important features that recognize an EA include the representation scheme which encode the plan as a string, the selection procedure, and the reproduction operators.

Representation

A typical representation for combinatorial problems, where one has to determine which of a number of items he wishes to choose, is the fixed-length vector of bits $\{0, 1\}$ [MF00]. The number of items or parameters at hand determines the length of the vector and 1 indicates that the indexed item is chosen whereas 0 indicates that it is discarded.

However, in our EA an individual plan is represented by a set containing the beam indices that identify which beams are in the plan. Although this representation is nonstandard, one can easily see that it is bijective to the fixed-length vectors representation. For instance, from a given set of 20 beams where 5 different beams are to be chosen, the fixed-length representation given by the binary string

$$(0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0)$$

is equivalent to the following set representation:

$$\{2, 6, 10, 14, 18\}$$

and vice versa.

Thus the choice of either representation offers no unique advantage over the other one (see [FG97]) except that the set representation appears to be more intuitive to the structure of our problem and hence more recommended [MF00].

Initial population

When the initial population is created, the number of beams for each individual is initialized according to the required number of beams q, see problem 3. The beam indices are then selected uniformly at random from those beams available in the space $\mathbb{B} = \{b_1, \ldots, b_n\}$.

Finally, one more solution, called *potential solution* (PS), is added to the initial population. This PS is found by the following procedure:

Stage 0: Initialization

a) [Intensity maps]

    FOR (*Each beam $b \in \mathbb{B}$*)
        FOR (*Each beamlet belonging to b*)
            Solve the optimization problem
$$\min \ F(x) = \sum_{k=1}^{K} w_{R_k} V_{R_k} D_{R_k}^2 x^2 - w_T V_T^2 D_T x,$$

$$x \in [0, U].$$

        END FOR
    END FOR

b) $B \leftarrow \emptyset.$
  $B^{\complement} \leftarrow \mathbb{B}.$

Stage 1: Find the PS

WHILE ($|B| \neq q$)
        Solve the optimization problem
$$\max_{b \in \mathbb{B}} \ F(B \cup \{b\}) = W_T \ f(DVH_T) - \sum_{k=1}^{K} W_{R_k} \ f(DVH_{R_k}).$$

    $B \leftarrow B \cup \{b\}.$
    $B^{\complement} \leftarrow B^{\complement} \setminus \{b\}.$
END WHILE
RETURN B
END

As one can see, the idea of this procedure is taken from the dynamic algorithm presented in section (4.2). The benefit of including the PS to the initial population becomes clear in section (5.4) which contains some analytical results of the EA implementation.

Parents selection

After a generation is created, each genotype (set of beams B) is mapped to the corresponding phenotype (corresponding plan P_B) which is then evaluated using the function

$$F(P_B) = W_T \; f(DVH_T) - \sum_{k=1}^{K} W_{R_k} \; f(DVH_{R_k})$$

as introduced in section (5.2.3). The generation is then ranked from 1 to N, where 1 refers to the fittest solution and N to the one with the worst fitness.

The selection procedure for this EA is rather simple. To induce a selective pressure towards the evolution of fitter solutions some of the parents are chosen by utilizing the *elitist selection* in which only the fittest solutions are being selected for parenthood. Furthermore, in order to have a kind of diversity between the individuals several solutions are chosen uniformly at random from the remaining solutions.

From a parent to a family

Yet each selected parent is ready for generating its own family while going through the following reproduction operators:

- *Cloning:* In order to avoid losing good genetic material (e.g. good beam indices), the first child in the family is a duplication or cloning of its parent. This guarantees that the next generation will contain solutions that are at least as good as the parents.

- *Recombination:* In EAs, recombination refers usually to a sexual reproduction operator, that shuffles together genetic material from selected parents (two or more) to produce new genotypes of offspring. Hence first of all a partner for the operated parent is found. This is simply done by selecting a partner from the elite parents uniformly at random. Furthermore, when passing the genes, which are the beam indices in our case, to the children, these genes that appear in both of the genotypes of the parents are copied to the genotypes of the generated children. These genotypes are then completed randomly from the shuffled remaining genes. For illustation, see the example shown below:

  parent A = { 7, 12, 20, 25, 35, 41, 60 }
  parent B = { 1, 19, 15, 20, 27, 35, 50 }

similar genes = { 20, 35 }
remaining genes = { 27, 12, 1, 60, 50, 41, 15, 7, 19, 25 }
child X = { 20, 35, 27, 12, 1, 60, 50 }
child Y = { 20, 35, 41, 15, 7, 19, 25 }

Note that it is difficult to identify the goodness of a single beam index. In fact, the quality of a beam index is relative to the remaining beam indices that appear in the considered genotype. Hence, it would not be clever to divide a component consisting of two (or more) beam indices that are good with respect to each another into parts where each is passed to a different child. Such good components are easily recognized since they usually appear in different elites. This is how they are recognized in the recombination operator and hence passed immediately to the children.

- *Mutation:* Mutation is a random change in one or more genes of a selected genotype.
  In the representation used here, mutation has special importance since it is the only way that brings new genetic material into the population. Mutating the genotype of the selected parnet is accomplished by the following procedure:

  1- Select a couple of genes for mutation uniformly at random.

  2- For each selected gene b, mutate subject to the function

  $$b \leftarrow [b + N(0, \sigma)] \bmod n,$$

  where $N(0, \sigma)$ is a Gaussian random variable with mean zero and standard deviation σ and n is the number of available beams.

  3- Add the resulted genotype to the parent's family.

Moreover, since the beams are located on a sphere, it very likely to happen the case that replacing a certain beam direction by its symmetric position, with respect to the center of the sphere, results in a better plan. Hence the following mutation is also used in the implementation of the EA:

  1- Select a gene b for mutation uniformly at random.

  2- Mutate b symmetrically.

  3- Add the resulting genotype to the parent's family.

Hence the above reproduction operators cloning, recombination and mutation make out of each selected parent a family with several children, each of which is a potential candidate for being promoted to the next generation.

Parent replacement

After a family is generated, each of its genotypes is mapped to the corresponding phenotype for evaluation purposes and then the parent is replaced by the elitist child in the family. All other family members are discarded after the replacement.

Twin replacement

The mechanism of the above reproduction operators allows twins to appear in the same generation, a phenomenon that happens more likely in the advanced generations. To maintain genetic diversity along the progress of the algorithm, in every generation all the twins except one are replaced by totally new individuals generated spontaneously at random. This allows for inserting new solutions into the population and in the same time keeps the number of its individuals unchanged.

5.4 Results

To illustrate the efficiency of the presented approach and the performance of the implementation of the EA they are applied to the data of three clinical cases: a prostate carcinoma, a paraspinal case and a tumor in the head and neck region. These three examples are the same clinical examples that have been introduced and used in chapter 4. For complete descriptions of these examples, see section 4.3.

5.4.1 Parameters values

The performance of an EA depends on the choice of its parameters, such as population size, the number of the selected parents, stopping criteria, etc. The parameter values used for the above examples were chosen empirically, according to the size of the data structure of the problem such as the total number of available beams and the required number of beams, as follows:

1- Population size = 20.

2- The number of elitist solutions selected for parenthood = 6.

3- The number of random solutions selected for parenthood $= 3$.

4- In the Gaussian random variable $N(0, \sigma)$: $\sigma = 5$.

5- Number of generations $= 100$.

5.4.2   Numerical results of the EA

The robustness of the optimization process of the EA was judged by the consistency of the results. In order to measure the consistency, ten runs of the EA were performed and the mean and the standard deviation of the results were reported. To examine the effect of inserting the PS into the initial population, the EA was tested again, but this time without including the PS into the initial population.

Figures (5.3), (5.4) and (5.5) plot the maximum score of each generation of the EA for the three clinical cases mentioned above. The maximum scores represent the average in generation over ten runs.
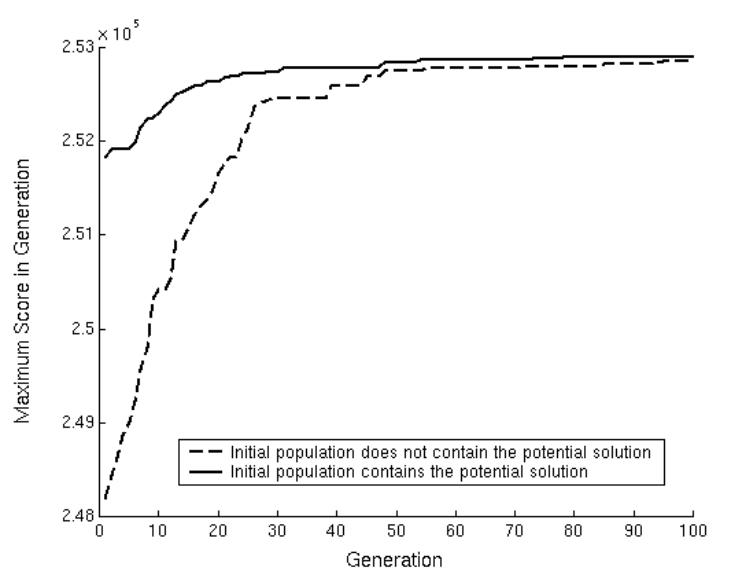

Figure 5.3: Results of the average scores over ten runs for the prostate example.

As it can be seen from these figurers, the best scores are consistently achieved earlier when the PS was inserted into the initial population.
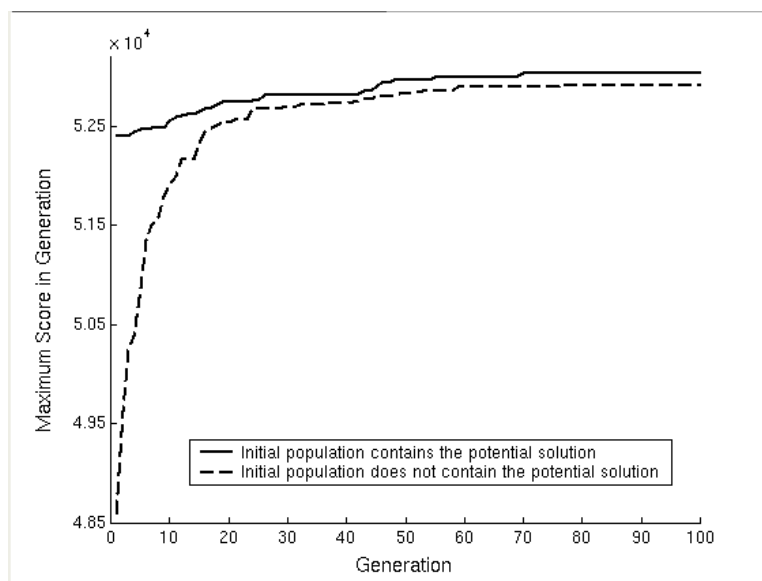
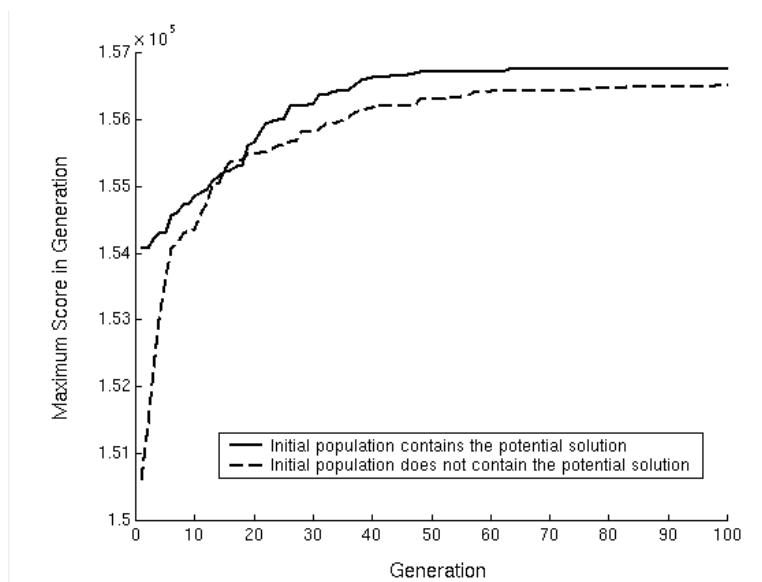Figure 5.4: Results of the average scores over ten runs for the paraspinal case.

Figure 5.5: Results of the average scores over ten runs for the head and neck example.
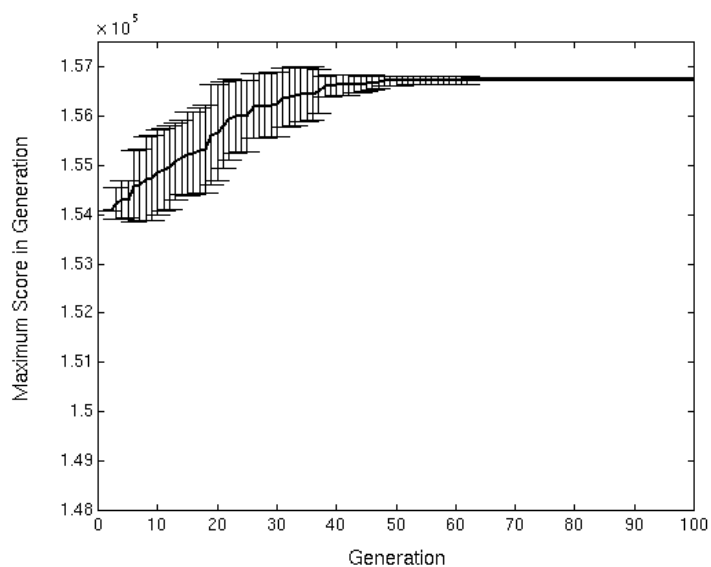
Figure 5.6: Results of the mean best score for each generation over ten runs with the variation represented by the standard deviation error bars. This data is for the head and neck example with initial population contains the potential solution.
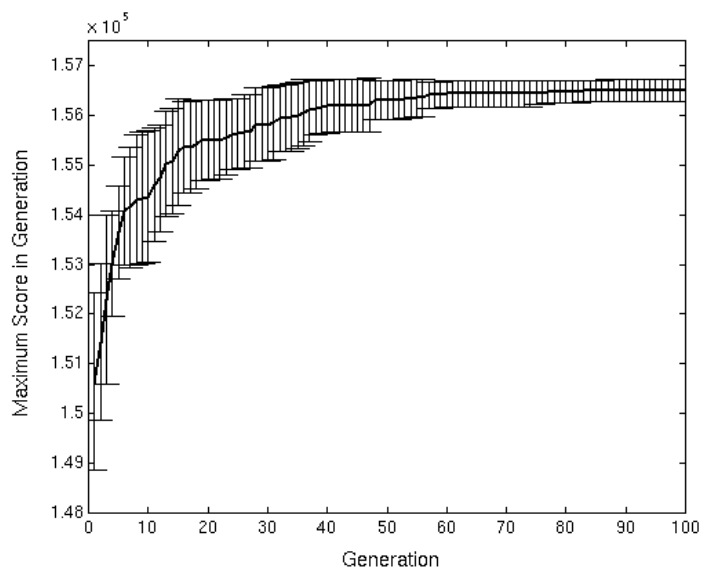

Figure 5.7: Results of the mean best score for each generation over ten runs with the variation represented by the standard deviation error bars. This data is for the head and neck example with initial population contains no potential solution.

Figure 5.5 is of particular interest. Although the curves report somewhat similar values between generation ten and generation twenty, the solid line representing the results of the runs with the potential solution included shows a faster convergence towards optimality. This indicates that the potential solution contains combination of beams that is at least close to combinations of a high quality. Hence the pressure of the optimization process towards areas of good solutions increases.

Furthermore, figures (5.6) and (5.7) show the results of computing the variance in the scores for the two cases with and without potential solution inserted into the initial population respectively. For the purpose of brevity, only the results of the head and neck example were depicted. These two figures demonstrate that inserting the potential solution into the initial population helps reducing the variance in earlier generations, allowing for a faster convergence.

5.4.3   Beam orientations results

Similarly to chapter 4, the judgment of the efficiency of the presented model was based on the comparison between two types of optimization: (1) beam intensity profiles were optimized for coplanar equiangular beams; (2) beam intensity profiles were optimized for the coplanar orientation set obtained by the EA. To evaluate the quality of the treatment plan, a depiction of the dose distribution and DVHs were used for the target and each of the organs at risk.

Prostate carcinoma

A five coplanar equiangular beams plan were used in comparison with the five coplanar beams plan optimized by using the EA presented above. The dose distributions and the dose volume histograms of these two plans are shown in figures (5.8) and (5.9).

It can be seen from figure (5.9) that the plan obtained by the optimized beam orientations has notably improved the bladder, rectum and right femur. These improvements were also reproted by the numerical realization of the values of the objective function which were reduced by 2%, 4%, 7% and 3% for the bladder, rectum, left femur and right femur respectively. Moreover, the shape of the dose distribution obtained using the optimized beams is fitted better to the shape of the tumor than the one obtained by the equiangular spaced beams. This can be easily seen in figure (5.8).
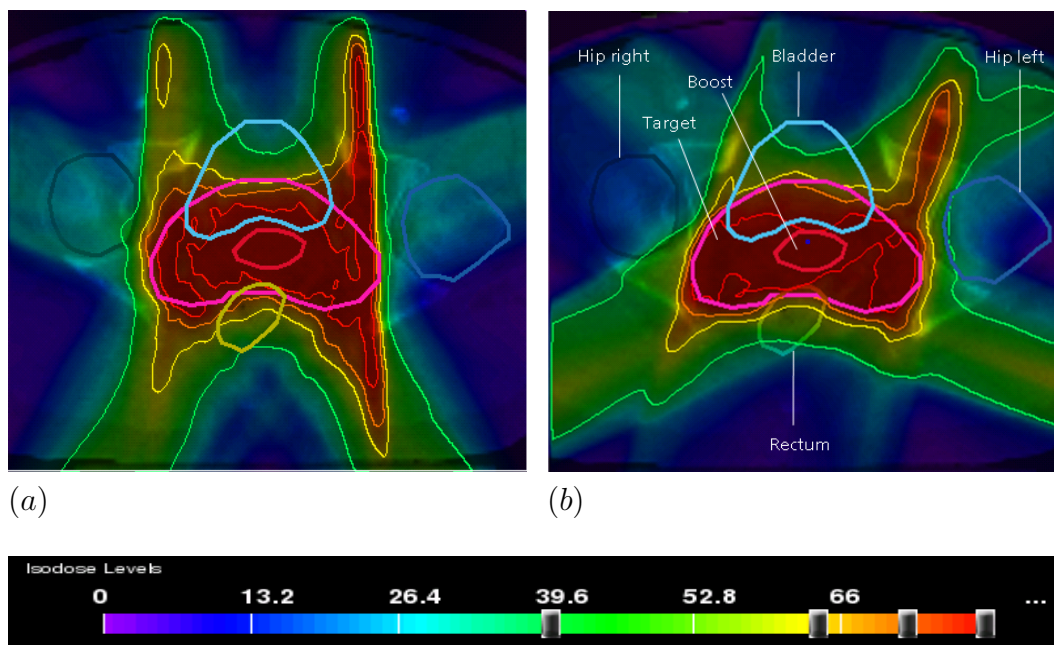
Figure 5.8: The dose distribution of two multicriteria IMRT prostate tumor treatments corresponding to (*a*) five equiangular spaced beams and (*b*) five optimized beams obtained by the methods and evolutionary algorithm presented above.
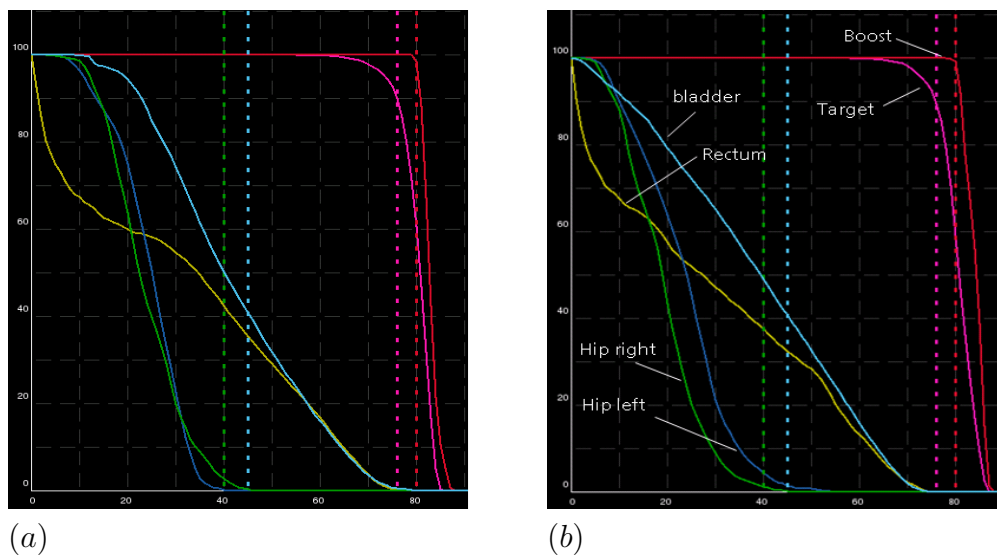

Figure 5.9: Dose volume histograms of two multicriteria IMRT prostate cancer treatments corresponding to (*a*) five equiangular spaced beams and (*b*) five optimized beams.

Paraspinal case

In this case, a seven coplanar equally spaced beams plan was compared with the plan obtained by using six coplanar beams optimized using the EA. The dose distributions of these two plans and their dose volume histograms are plotted in Figures (5.10) and (5.11) respectively.
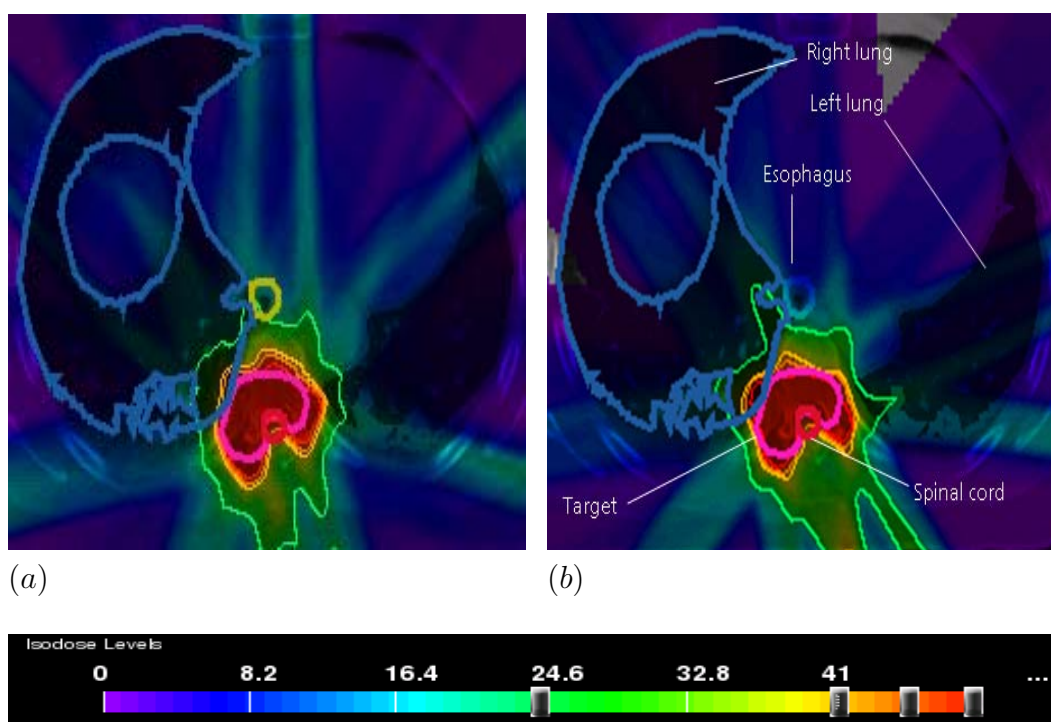

Figure 5.10: The dose distribution of two multicriteria IMRT treatments for paraspinal tumor, corresponding to (a) seven equiangular spaced beams and (b) six optimized beams obtained by the evolutionary algorithm.

From these figures, a significant improvement with respect to the esophagus can be easily seen. Improvements regarding the values of the objective function were also reported. The values of the objective function were reduced by 16% and 4% with respect to the esophagus and right lung respectively, where as all the others values were almost equal.
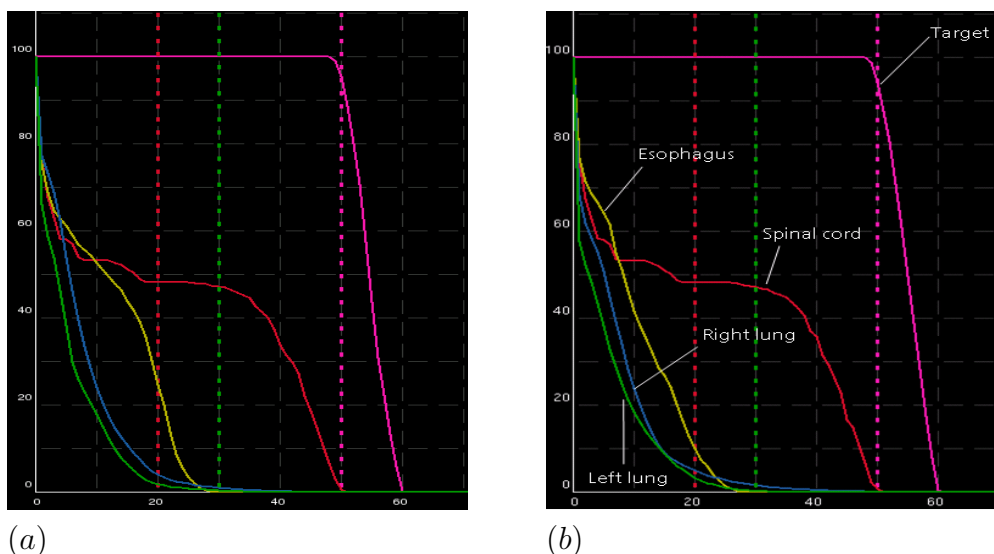
(a)   (b)

Figure 5.11: Dose volume histograms of two multicriteria IMRT treatments for paraspinal tumor, corresponding to (a) seven coplanar equiangular spaced beams and (b) six optimized beams.

Head and neck tumor

Similarly to the paraspinal case, the comparison in this example was also between a seven coplanar equally spaced beams plan and the plan obtained by using only six, rather than seven, coplanar beams optimized by the EA. The dose distributions of these two plans and their dose volume histograms are shown in Figures (5.12) and (5.13) respectively.

As can be seen from figure (5.13), the use of the optimized beams resulted in significant improvement with respect to the brain stem against only margin deuteration concerning the parotid gland and target dose uniformity. Furthermore, figure (5.12) shows that some unnecessary hot spots in the irradiated volume were eliminated, allowing for better dose shaping over the target. The numerical realization of the objective function reported in this case was improved by 33% for the brain stem against worsening the parotid gland by 20%.

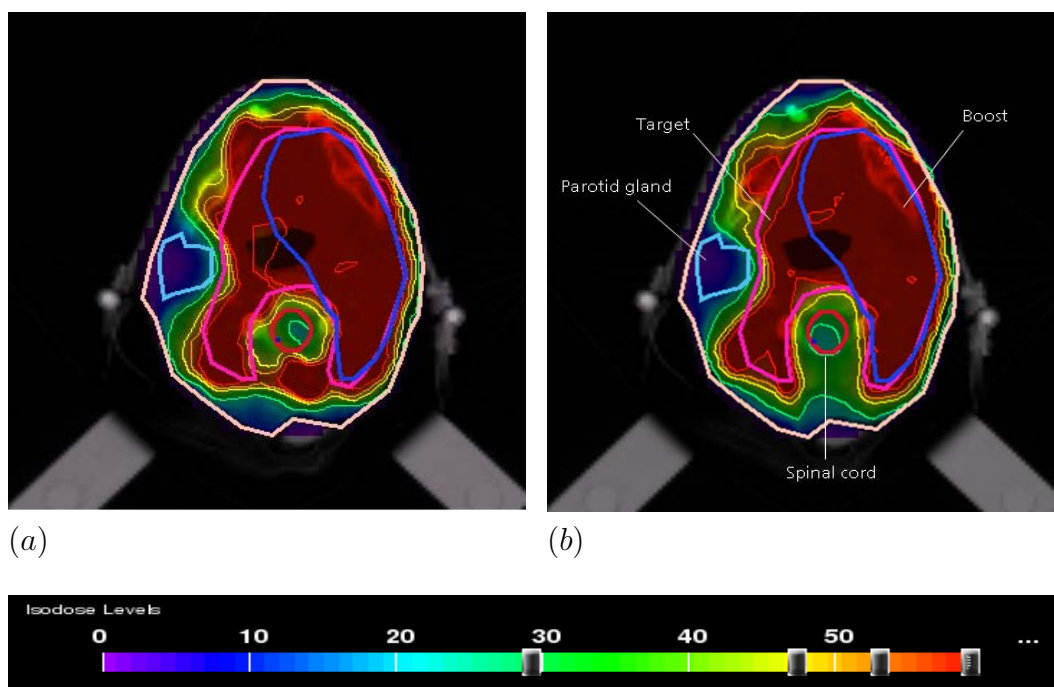Figure 5.12: The dose distribution of two multicriteria IMRT treatments, for head and neck tumor, corresponding to (a) seven equiangular spaced beams and (b) six optimized beams obtained by the presented EA.
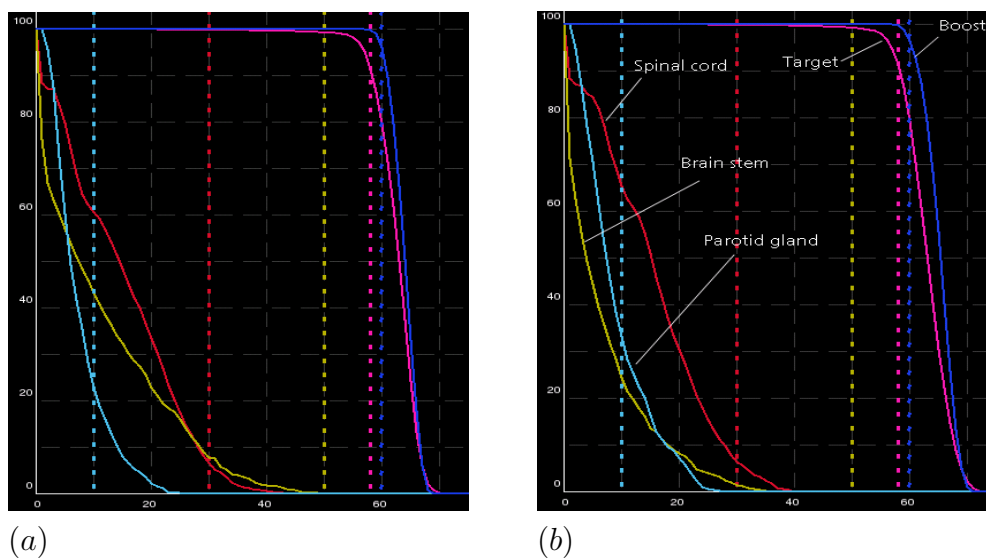

Figure 5.13: Dose volume histograms of two multicriteria IMRT treatments, for head and neck tumor, corresponding to (a) seven coplanar equiangular spaced beams and (b) six optimized beams obtained by the presented EA.

Chapter 6

Comparison between the dynamic and evolutionary algorithms

In this thesis, two algorithms with totally different approaches have been developed to solve the problem of beam orientations in IMRT. As usual, to examine the efficiency of these algorithms, the resulting plan of each of them was then compared with the plan resulting from using equidistant beam orientations. The goal of this chapter is to test both algorithms against each another and investigate the possibility of any correlation. This is done with the three clinical cases that have been used along this thesis. For each clinical example, the results of the both algorithms are depicted and discussed qualitatively.

6.1   Results

6.1.1   Prostate carcinoma

In this case the DA and EA have resulted the following sets of beam orientations (angles):

$$B_{DA} = \{\mathbf{50°}, \mathbf{115°}, 180°, \mathbf{250°}, 270°\}$$

$$B_{EA} = \{15°, \mathbf{60°}, \mathbf{110°}, \mathbf{245°}, 310°\}.$$

The dose distributions and the dose volume histograms of these two plans are shown in figures (6.1) and (6.2) respectively.
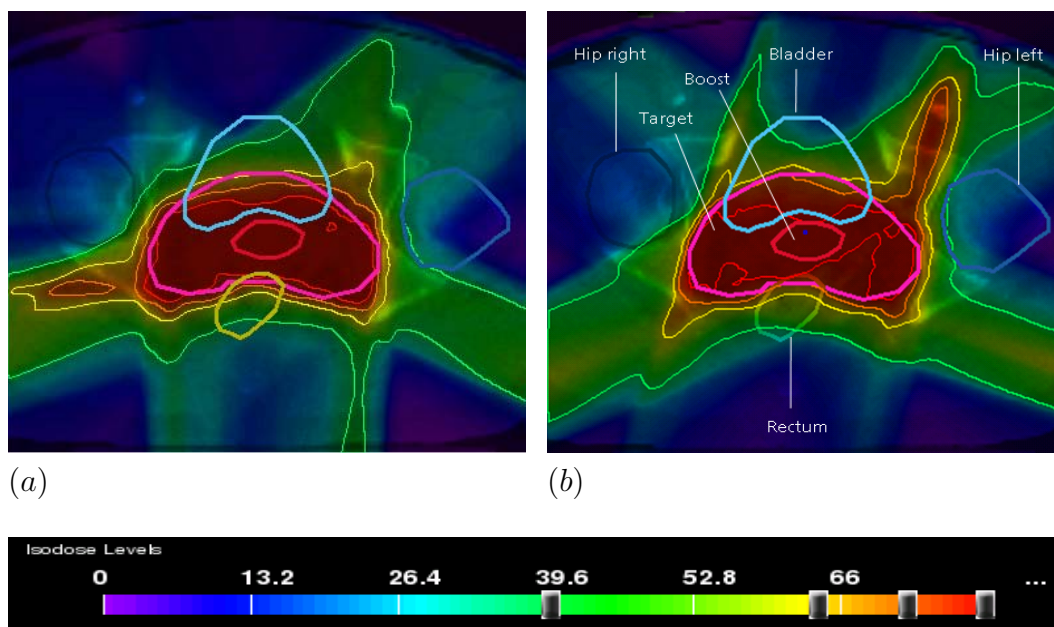
Figure 6.1: The dose distribution of two multicriteria IMRT prostate tumor treatments corresponding to (a) five optimized beams obtained by the DA and (b) five optimized beams obtained by the EA.
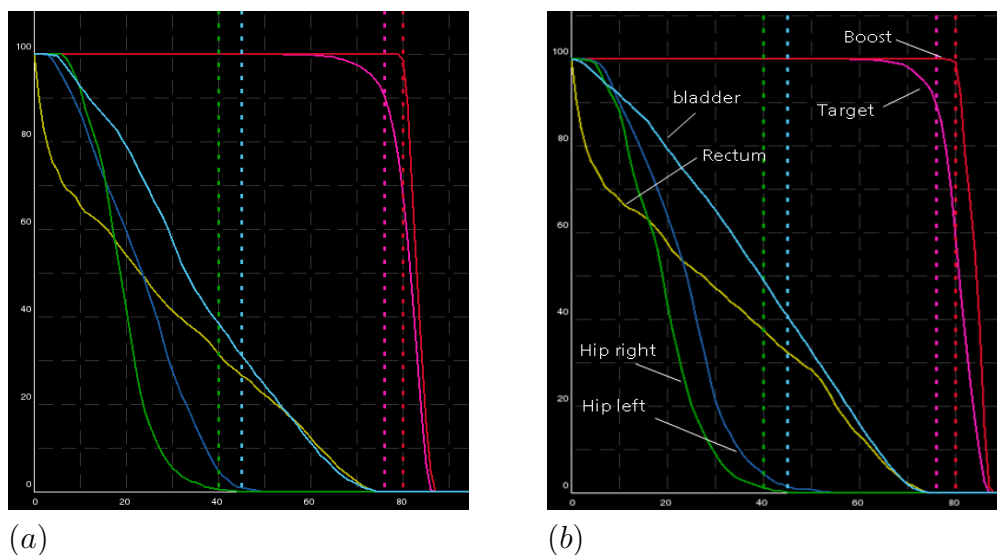

Figure 6.2: Dose volume histograms of two multicriteria IMRT prostate cancer treatments corresponding to (a) five optimized beams obtained by the DA and (b) five optimized beams obtained by the EA.

While figure (6.1) demonstrates that the target dose coverage is well established in the two plans, figure (6.2) shows that the plan obtained by the DA results in minor improvements, regarding the bladder, rectum, and right femur over the corresponding ones that are obtained by the plan resulting from the EA. These little improvements were also repotred by the numerical realization of the values of the objective function which were reduced by 4%, 1%, and 3% for the bladder, rectum, and right femur respectively.

6.1.2   Paraspinal case

In this case the beam orientations set resulted by the DA is given as follows:

$$B_{DA} = \{110°, \mathbf{145°}, \mathbf{180°}, \mathbf{225°}, 250°, 350°\},$$

where as the optimal set of beam orientation obtained by the EA reads

$$B_{EA} = \{70°, 95°, \mathbf{140°}, \mathbf{175°}, \mathbf{225°}, 265°\}.$$

The dose distributions and dose volume histograms of the plans obtained by these two sets of beam orientations are plotted in Figures (6.3) and (6.4) respectively.

Figure (6.3) shows that the dose distribution of the plan resulting from the orientation set B_{EA} is only a bit better concentrated in the target than the dose distribution of the plan obtained by the beam orientations B_{EA}. Moreover, plan (b) in figure (6.4) demonstrates that the totally spared volume of the right lung has increased. In addition to this, an improvement in DVH of the esophagus was reported against a noticeable worsening regarding the DVH of the left lung. Numerically, the values of the objective function of plan (b) compared with plan (a) were reduced by 4% with respect to the esophagus and 3% regarding the right lung. On the other hand, the value regarding the left lung was worsened by 50%. Note that these percentage changes in the values of the objective function are not explicit for plan evaluation. In this case for instance, it is difficult to decide which plan, (a) or (b), is more preferable. After all, the decision here is completely planner-dependent.

(a)   (b)

Figure 6.3: The dose distribution of two multicriteria IMRT treatments for paraspinal tumor, corresponding to (a) six optimized beams obtained by the DA and (b) six optimized beams obtained by the EA.

(a)   (b)

Figure 6.4: Dose volume histograms of two multicriteria IMRT treatments for paraspinal tumor, corresponding to (a) six optimized beams obtained by the DA and (b) six optimized beams obtained by the EA.

6.1.3   Head and neck tumor

In this case the beam orientations set resulted by the DA reads

$$B_{DA} = \{\mathbf{100°}, \mathbf{175°}, \mathbf{200°}, \mathbf{235°}, 310°, 345°\},$$

whereas the optimal set of beam orientations obtained by the EA is given as
follows:

$$B_{EA} = \{ \overbrace{\mathbf{20°}}^{+180°=\mathbf{200°}}, \mathbf{100°}, 135°, \mathbf{175°}, \mathbf{230°}, 295°\}.$$

Figures (6.5) and (6.6) depict the dose distributions and dose volume histograms of the plans resulted from the above two sets of beam orientations.
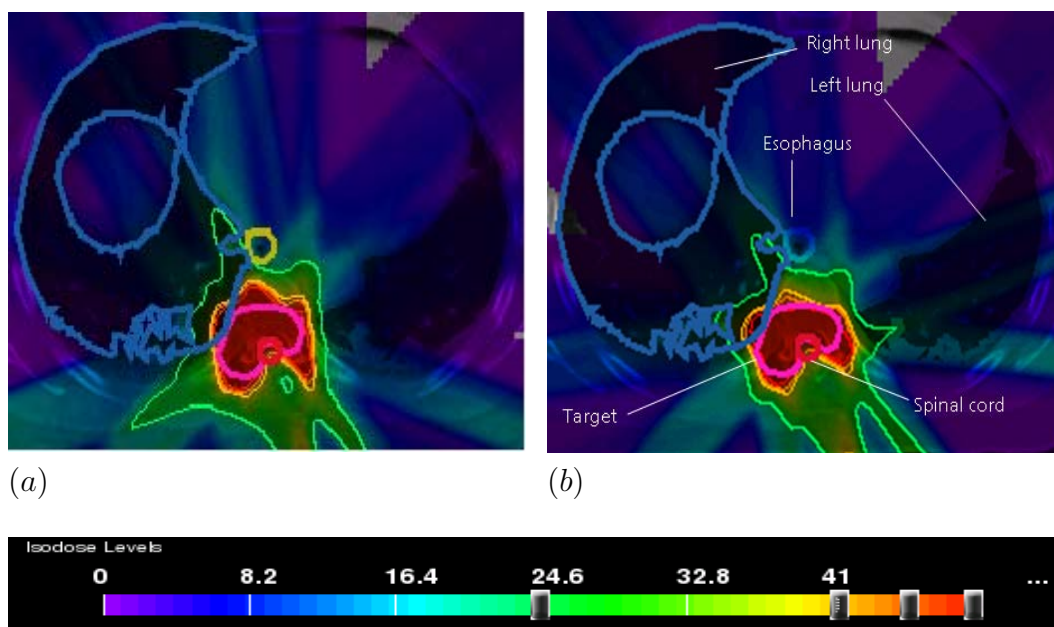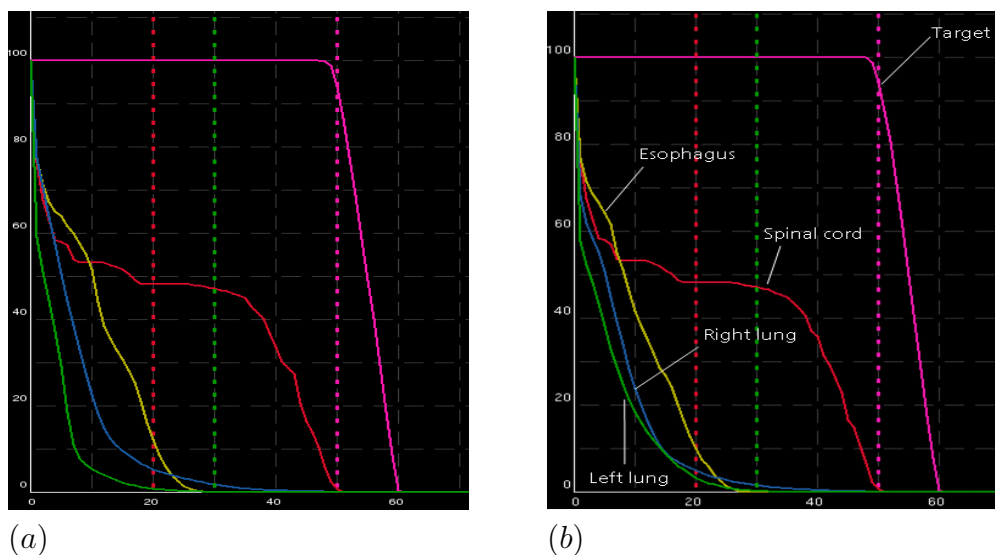

(a) (b)

Figure 6.5: The dose distribution of two multicriteria IMRT treatments, for head and neck tumor, corresponding to (a) six optimized beams obtained by the DA and (b) six optimized beams obtained by the EA.

While figure (6.6) shows improvement, reported for plan (b) over plan (a), regarding the brain stem, it on the other hand shows some worsening regarding the parotid gland and target dose uniformity. From figure (6.5), it is difficult to note any major differences between the two plans. The numerical realizations reported were improvement by 29% with respect to the value corresponding to the brain stem against worsening the values corresponding to the spinal cord and parotid gland by 4% and 10%, respectively.
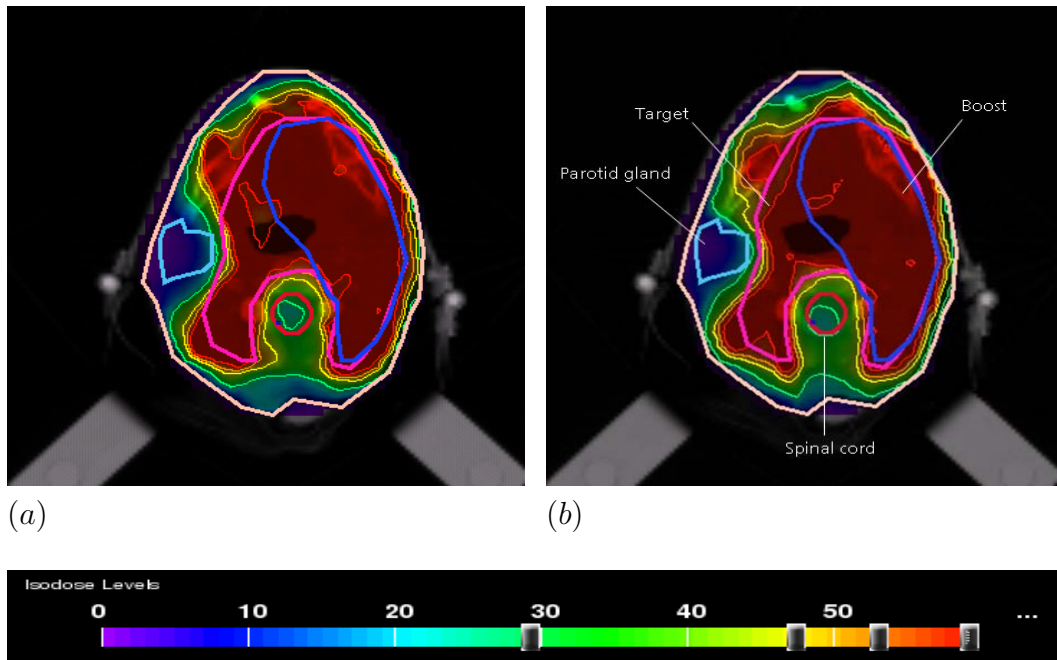
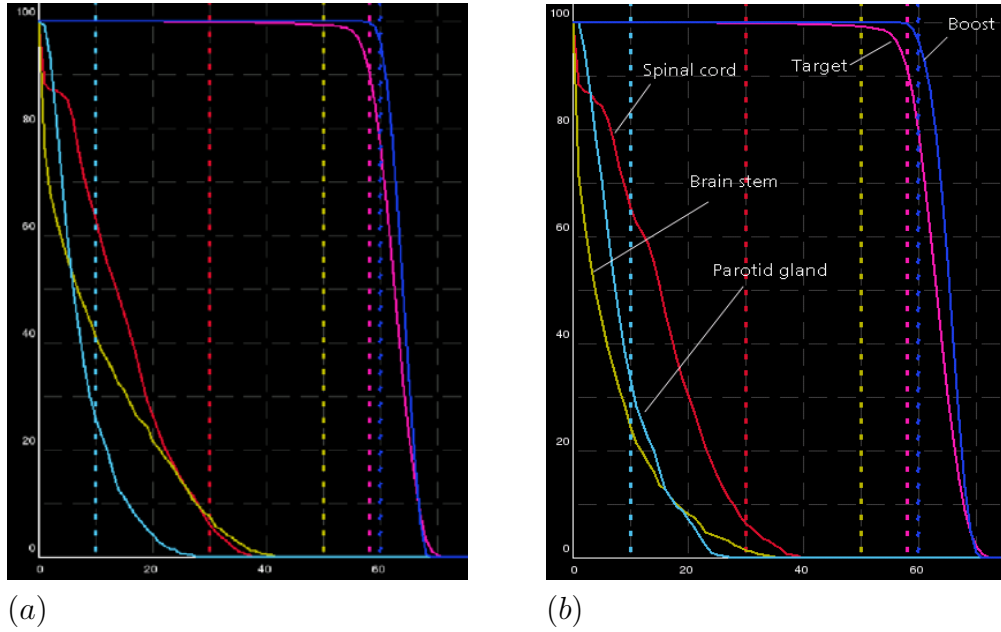(a)   (b)

Figure 6.6: Dose volume histograms of two multicriteria IMRT treatments, for head and neck tumor, corresponding to (a) six optimized beams obtained by the DA and (b) six optimized beams obtained by the EA.

6.2   Discusion

This chapter was devoted to compare the performances of the dynamic and evolutionary algorithms presented in chapters 4 and 5 of this thesis. Although, the two algorithms are essentially different in that each of them uses its own completely different methods, the results of applying the algorithms to the different clinical examples were widely comparable. This is mainly due to the degree of similarity which was observed in the outputs of these two algorithms. Surprisingly enough, for each of the three clinical cases, the EA has resulted in a set of beam orientations in which at least three beams were similar or close to some corresponding beams that can be found among the beams resulted by the DA. For instance in the head and neck case, the beam angles $\{100°, 175°\}$ have appeared in both of the orientation sets B_{DA} and B_{EA}. Also the beam angle $235° \in B_{DA}$ lies just next to the beam angle $230°$ which belongs to the set B_{EA}. In addition to this, the angle $200°$ belonging to B_{DA} is exactly the opposite direction of the beam angle $20°$ which belongs to the orientation set B_{EA} obtained by the EA.

Moreover, the result of achieving highly correlated plans with different

beam configurations indicate that in general there exists a variety of good beam configurations. This supports what has been mentioned by Meedt [MAN03] and Hou [HWCG03]. They both have indicated that beam orientations optimization in IMRT leads to a large number of equivalent beam orientations.

Chapter 7

Summary

This thesis was focused on the optimization problem of beam orientations in IMRT planning. It was shown that this problem is neither convex, when formulated as a continuous optimization problem, nor polynomially solvable, when modeled as a combinatorial optimization problem. Hence the problem was rendered to the usage of heuristic methods. Usually, a stochastic selector is used for optimizing beam orientations, and then a single objective inverse treatment planning algorithm is used for the optimization of beam intensity profiles. Unfortunately, the quality of a set of beam orientations depends heavily on its corresponding beam intensity profiles. The overall time needed to solve the inverse planning for every random selection of beam orientations becomes excessive. Hence, the choice of beam orientations in IMRT treatment planning is still a trial-and-error procedure based on intuition and empirical knowledge, although the optimum beam configuration may sometimes be counterintuitive [SMW+97].

Recently, considerable improvement has been made in optimizing beam intensity profiles by using multiple objective inverse treatment planning. Such an approach results in a variety of beam intensity profiles for every selection of beam orientations, making the dependence between beam orientations and its intensity profiles less important. The thesis took advantage of this property to accelerate the optimization process through an approximation of the intensity profiles that can be used for multiple selections of beam orientations, saving a considerable amount of calculation time. Two algorithms, DA and EA, for beam orientations in IMRT planning were put forward. The DA is an automatic imitation of the beam's eye view and observer's view methods which are well-recognized techniques in conventional conformal radiation therapy. Every iteration goes from an N-beam plan to a plan with $N + 1$ beams. Beam selection criteria are based on a score func-

tion that minimizes the deviation from the prescribed dose, in addition to a reject-accept criterion. The EA was based on a DVH evaluation function which was introduced as a mechanism to minimize the deviation between the mathematical and clinical optima which might be in general very different, depending on the chosen mathematical model.

After applying the algorithms to different clinical examples, the resulting plans were compared with the corresponding plans obtained by the standard equally spaced beam orientations. It was shown that both algorithms are able to produce superior beam orientations to the equally spaced beams in clinically efficient time, even when, for some cases, a smaller number of beams were used.

It may be argued that there is no evidence that the algorithms converge to a global optimum, especially in the case of the DA when a large number of beams is required. In practice this is not necessarily so, because it was found that the results of applying the algorithms to the different clinical examples were widely comparable, indicating that there exists a variety of good beam configurations. And since IMRT can compensate for modest imperfections in beam directions, it is not clear whether achieving the global optimum in beam orientations optimization is clinically important. Moreover, it has been argued that the smaller the number of beams, the more necessary it is to optimize beam orientations (see e.g [WZD+04]). In general, when a relatively large number of beams is used, the final IMRT plan may be less sensitive to the selection of beam orientations [SMW+97]. However, using a large number of beams may have the undesirable consequence of spreading low doses to a large volume of the normal tissue which in turn tends to increase the risk of second cancers (see e.g [Hal06] or [HW03]). It may also increase treatment delivery time which may increase patient discomfort in addition to the possibility of increasing potential error in terms of patient movement. Indeed, it is always desirable to reduce the number of beams to its minimum without compromising the quality of the treatment plan.

Chapter 8

Further research

8.1 Improving the DA

As an attempt to improve the DA, a current study based on merging the presented DA with a multicriteria solver is in progress. It is believed that this study, which aims at optimizing beam orientations in IMRT, may produce plans of high quality. In fact, a good approximation of the intensity profiles is crucial for optimizing the orientation set accurately. For this reason, the multicriteria solver is intended to be used once at the end of every iteration where the decision on the most desirable intensity profiles is left for the planner. It is expect that this human guidance will help the algorithm to minimize the deviation from the optimal clinical solution efficiently.

8.2 A two phases algorithm

As was concluded in chapter 6, the resulting beam orientations sets of the two algorithms presented in this thesis were almost similar in at least three beam directions per solution. It was observed that these three beams were often easily detected by the first three iterations performed by the DA, even without the need for the reject-accept criterion. Obtaining them again by the EA motivates trying the following two phases algorithm:

Phase 1: Apply the DA to find a three beams orientations set.

Phase 2: Fix the beams found in phase 1 and apply the EA to find the remaining beams.

It is expected from such an algorithm to accelerate the EA considerably without compromising the quality of the final beam orientations.

8.3   Biological parameters

In chapter 5, the milestone of the EA was the DVH-evaluation scheme which
was introduced to evaluate the clinical impact of the irradiated dose on a
certain organ or clinical structure. As it was pointed out, a propper math-
ematical model has to take into account the biological properties of the in-
dividual specified tissue when it is exposed to radiation. Hence, the tissue-
specified parameter α was incorporated into the model. To more fully exploit
the potential of the presented model, biological investigations are needed to
determine the accurate values of the parameter α for the different clinical
structures.

Appendix A

Complementary definitions

A.1 Multicriteria combinatorial optimization problems

Let $\mathbb{B}$ be a finite set $\mathbb{B} := \{b_1, b_2, \ldots, b_n\}$ and let several weight functions $w_j : \mathbb{B} \to \mathbb{Z}$ be given, yielding several objective functions $f_i, i = 1, \ldots, m$. A combinatorial optimization problem is given by a feasible set

$$\mathfrak{B} \subset 2^{\mathbb{B}}$$

defined as a subset of the power set of $\mathbb{B}$ and an objective function $f(B) = (f_1(B), \ldots, f_m(B))$ defined for $B \in \mathfrak{B}$ to be minimized (see e.g. [Ehr00]). It can therefor be written in the common notaion

$$\min_{B \in \mathfrak{B}} f(B).$$

A.2 Binary relation and some properties

Given a set $\mathcal{C}$, a *binary relation* on $\mathcal{C}$ is defined as a subset of $\mathcal{C} \times \mathcal{C}$. IF $(P, Q) \in R$ where $P, Q \in \mathcal{C}$, we say "P is related to Q under R".

A relation R defined on a set $\mathcal{C}$ is said to be *reflexive* if for all $P \in \mathcal{C}$, $(P, P) \in R$ holds. It is *symmetric* if for all $P, Q \in \mathcal{C}$, and $(P, Q) \in R$ it follows $(Q, P) \in R$ as well. It is *transitive* if for all $O, P, Q \in \mathcal{C}$, $(O, P) \in R$ and $(P, Q) \in R$ implies $(O, Q) \in R$.

Finally a relation is said to be an *equivalence relation* if it is reflexive, symmetric and transitive.

Appendix B

Abbreviations

ABEV   Automated beam's eye view.
AOV   Automated observer's view.
BEV   Beam's eye view.
DA   Dynamic algorithm.
DSBO   Related decision problem of SBO.
DVH   Dose volume histogram.
DVV   Dose volume vector.
EA   Evolutionary algorithm.
EUD   Equivalent uniform dose.
GBO   Generic beam orientations.
IMRT   Intensity-modulated radiation therapy.
MLC   Multi-leave collimator.
$\mathcal{NP}$ See definition 7 in page 28.
$\mathcal{NPC}$ $\mathcal{NP}$-complete.
$\mathcal{NP}$-complete See definition 10 in page 28.
$\mathcal{NP}$-hard   See definition 11 in page 28.
NTCP   Normal tissue complication probability.
$\mathcal{P}$ Class of decision problems that can be solved in polynomial time.
OAR   Organ at risk.
OV   Observer's view.
TCP   Tumor control probability.
SBO   Simplified beam orientations problem.
VDV   Volume dose vector.
VOI   Volume of interest.

Bibliography

[AHSB99]   M. Asell, S. Hyodynmaa, S. Soderstrom, and A. Brahme, *Optimal electron and combined electron and photon therapy in the phase space of complication-free cure*, Phys Medic Biol **44** (1999), 235–252.

[Ben99]   Peter J. Bentley, *Evolutionary design by computers*, Morgan Kaufmann, 1999.

[Bor95]   T Bortfeld, *Dosiskonformation in der Tumortherapie mit externer ionisierender Strahlung: Physikalische Möglichkeiten und Grenzen*, Habilitationsschrift, Deutsches Krebsforschungszentrum, Heidelberg (1995).

[Bor99]   _____, *Optimized planning using physical objectives and constraints*, Semin Radiat. Oncol. **9** (1999), 20–34.

[Bra84]   A Brahme, *Dosimetric precision requirements in radiation therapy*, Acta Radiol Oncol **23** (1984), 379–391.

[Bra95]   _____, *Treatment optimization using physical and radiobiological objective functions*, Radiation Therapy Physics (A Smith, ed.), Springer, Berlin, 1995.

[BS93]   T Bortfeld and W Schlegel, *Optimization of beam orientation in radiation therapy: Some theoretical consideration*, Phys Med Biol **38** (1993), 291–304.

[BSP97]   T Bortfeld, J Stein, and K Preiser, *Clinically relevant intensity modulation optimization using physical criteria*, Proceedings of the XIIth ICCR (D.D. Leavitt and G. Starkschall, eds.), Salt Lake City, 1997.

[Cha88]   G Y T Chan, *Dose volume histogram in treatment planning*, Int. J. Radiat. Oncol. Biol. Phys. **14** (1988), 1319–1320.

[Coo71]   S. A. Cook, *The complexity of theorem-proving procedures*, Proceedings of the 3rd Annual ACM Symposium on the Theory of Computing (1971), 151–158.

[CRRM99]   BC Cho, WH Roa, D Robinson, and B Murray, *The development of target-eye-view maps for selection of coplanar or noncoplanar beams in conformal radiotherapy treatment planning*, Med Phys **26** (1999), 2367–2372.

[CSP⁺92]   GT Chen, DR Spelbring, CA Pelizzari, JM Balter, LC Myrianthopoulos, S Vijayakumar, and H Halpern, *The use of beam's eye view volumetrics in the selection of non-coplanar radiation portals*, Int J Radiat Oncol Biol Phys **23** (1992), 153–163.

[DD97]   I Das and J Dennis, *A closer look at the drawbacks of minimizing weighted sums of objectives for pareto set generation in multicriteria optimization problems*, Structural Optimization **14** (1997), no. 1, 63–69.

[Ehr00]   M Ehrgott, *Multicriteria optimization*, Springer, 2000.

[EJ03]   M Ehrgott and R Johnson, *Optimisation of beam directions in intensity modulated radiation therapy planning*, OR Spectrum **25** (2003), 251–264.

[ELB⁺91]   B Emami, J Lyman, A Brown, L Coia, M Goitein, JE Munzenrieder, B Shank, LJ Solin, and M Wesson, *Tolerance of normal tissue to therapeutic irradiation*, Int J Radiat Oncol Biol Phys **21** (1991), 109–122.

[Ezz96]   Gary A. Ezzell, *Genetic and geometric optimization of three-dimensional radiation therapy treatment planning*, Medical Physics **23** (1996), 291–304.

[FG97]   D. B. Fogel and A. Ghozeil, *A note on representations and variation operators.*, IEEE Transaction on Evolutionary Computation **1** (1997), no. 2, 159–161.

[GAR⁺83]   M Goitein, M Abrams, D Rowell, H Pollari, and J Wiles, *Multi-dimensional treatment planning: Ii. beam's eye-view, back projection, and projection through ct sections*, Int J Radiat Oncol Biol Phys **9** (1983), 789–797.

[GLB94]   A Gustafsson, BK Lind, and A Brahme, *A generalized pencil beam algorithm for optimization of radiation therapy*, Med Phys **21** (1994), no. 3, 343–356.

[GN88]   M Goitein and A. Niemierko, *Biological based models for scoring treatment plans*, Scandinavian Symposium on Future Directions of Computer-Aided Radiotherapy (1988).

[GSH99]   T Gal, T J Stewart, and T Hanne, *Multicriteria decision making: Advanced in mcdm models, algorithms, theory, and applications*, Kluwer Academic Publishers, 1999.

[GWR04]   S Gaede, E Wong, and H Rasmussen, *An algorithm for systematic selection of beam directions for IMRT*, Medical Physics **31** (2004), no. 2, 376–388.

[HABY99]   M. E. Hosseini-Ashrafi, H. Bagherebadian, and E. Yahaqi, *Pre-optimization of radiotherapy treatment planning: an artificial neural network classification aided technique*, Phys Med Biol **44** (1999), 1513–1528.

[Hal06]   EJ. Hall, *Intensity-modulated radiation therapy, protons, and the risk of second cancers*, Int J Radiat Oncol Biol Phys **65** (2006), no. 1, 1–7.

[HBM98]   O. C. L. Haas, K. J. Burnham, and J. A. Mills, *Optimization of beam orientation in radiotherapy using planar geometry*, Phys Med Biol **43** (98), 2179–2193.

[HK02]   H.W. Hamacher and K.-H. Küfer, *Inverse radiation therapy planning - a multiple objective optimization approach*, Discrete Applied Mathematics **118** (2002), 145–161.

[HW03]   EJ. Hall and C. Wuu, *Radiation-induced second cancers: The impact of 3d-crt and imrt*, Int J Radiat Oncol Biol Phys **56** (2003), 83–88.

[HWCG03]   Q. Hou, J. Wang, Y. Chen, and JM Galvin, *Beam orientation optimization for imrt by a hybrid method of the genetic algorithm and the simulated dynamics*, Med Phys **30** (2003), 2360–2367.

[KMS+06]   Karl-Heinz Küfer, Michael Monz, Alexander Scherrer, Philipp Süss, Fernando Alonso, Ahmad Saher Azizi Sultan, Thomas

Bortfeld, and Christian Thieke, *Multicriteria optimization in intensity modulated radiotherapy planning*, Handbook of Optimization in Medicine (Panos M. Pardalos and H. Edwin Romeijn, eds.), Kluwer Academic Publishers, to appear in 2006.

[MAN03]   G. Meedt, M. Alber, and F. Nsslin, *Non-coplanar beam direction optimization for intensity-modulated radiotherapy*, Phys Medic Biol **48** (2003), 2999–3019.

[MCV⁺92]   LC Myrianthopoulos, GT Chen, S Vijayakumar, HJ Halpern, DR Spelbring, and CA Pelizzari, *Beam's eye view volumetrics: an aid in rapid treatment plan development and evaluation*, Int J Radiat Oncol Biol Phys **23** (1992), 367–375.

[MF00]   Z. Michalewicz and D. b. Fogel, *How to solve it: Modern heuristics*, Springer-Verlage Berlin Heidelberg., 2000.

[Mic98]   Z. Michalewicz, *Genetic algorithms + data structures = evolution programs*, Springer, Berlin., 1998.

[Mie99]   K Miettinen, *Nonlinear multiobjective optimization*, Kluwer, Boston, 1999.

[NCI97]   National Cancer Institute NCI, *Radiation therapy and you-a guide to self-help during treatment*, http://www.cancer.gov, 1997.

[Nie97]   A Niemierko, *Reporting and analyzing dose distributions: a concept of equivalent uniform dose*, Med Phys **24** (1997), 103–110.

[Pap94]   C.H. Papadimitriou, *Computational complexity*, Addison-Wesley, 1994.

[PBX00]   A. Pugachev, A.L. Boyer, and L. Xing, *Beam orientation optimization in intensity-modulated radiation therapy planning*, Med Phys **27** (2000), 1238–1245.

[PLB⁺01]   A. Pugachev, J.G. Li, A.L. Boyer, S. Hancock, Q.T. Le, S.S. Donaldson, and L. Xing, *Role of beam orientation optimization in intensity-modulated radiation therapy*, Int J Radiat Oncol Biol Phys **50** (2001), 551–560.

[PX1a]   A. Pugachev and L. Xing, *Pseudo beam's-eye-view as applied to beam orientation selection in intensity-modulated radiation therapy*, Int J Radiat Oncol Biol Phys **51** (2001a), 1361–1370.

[PX1b]   A. Pugachev and L Xing, *Computer-assisted selection of coplanar beam orientation in intensity-modulated radiation therapy*, Phys Med Biol **46** (2001b), 2467–2476.

[PX02]   ————, *Incorporating prior knowledge into beam orientation optimization in IMRT*, Int J Radiat Oncol Biol Phys **54** (2002), 1565–1574.

[RADK03] EH Romeijn, RK Ahuja, JF Dempsey, and A Kumar, *A column generation approach to radiation therapy planning using aperture modulation*, technical report, Department of Industrial and Systems Engineering, University of Florida, Gainesville, Florida (2003).

[ROW99]   C.G. Rowbottom, M. Oldham, and S. Webb, *Constrained customization of non-coplanar beam orientations in radiotherapy of brain tumours*, Phys Med Biol **44** (1999), 383–399.

[RWO98]   Carl Graham Rowbottom, Steve Webb, and Mark Oldham, *Improvements in prostate radiotherapy from the customization of beam directions*, Med Phys **25** (1998), 1171–1179.

[RWO99]   ————, *Beam-orientation customization using an artificial neural network*, Phys Med Biol **44** (1999), 2251–2262.

[SFOM99] David M. Shepard, Michael C. Ferris, Gustavo H. Olivera, and T. Rockwell Mackie, *Optimizing the delivery of radiation therapy to cancer patients*, SIAM **41** (1999), no. 4, 721–744.

[SM01]   W Schlegel and A Mahr, *3D Conformal Radiation Therapy - Multimedia Introduction to Methods and Techniques*, Multimedia CD-ROM, Springer, 2001.

[SMW+97] J. Stien, R. Mohan, X.H. Wang, T. Bortfeld, Q. Wu, K. Preiser, C.C Ling, and W. Schlegel, *Number and orientations of beams in intensity-modulated radiation treatments*, Med Phys **24** (1997), 149–160.

[VRA]   Valley Radiotherapy Associates VRA, http://www.valley-radiotherapy.com.

[Web01]   S Webb, *Intensity-modulated radiation therapy*, IOP Publishing Ltd, 2001.

[Wol84]   A. Wolbarst, *Optimization of radiation therapy II: The critical-voxel model*, Internatonal Journal of Radiation Oncology, Biology, Physics **10** (1984), 741–745.

[Wol98]   Laurence A. Wolsey, *Integer programming*, John Wiley and Sons, 1998.

[WTM87]   H. Withers, J. Taylor, and B. Maciejewski, *Treatment volume and tissue tolerance*, Internatonal Journal of Radiation Oncology, Biology, Physics **14** (1987), 751–759.

[WZD+04]   Xiaochun Wang, Xiaodong Zhang, Lei Dong, Helen Liu, Qiuwen Wu, and Radhe Mohan, *Development of methods for beam angle optimization for IMRT using an accelerated exhaustive search strategy*, Internatonal Journal of Radiation Oncology, Biology, Physics **60** (2004), no. 4, 1325–1337.

[Yu73]   PL Yu, *A class of solutions for group decision problems*, Management Science **19** (1973), no. 8, 936–946.

[Zel73]   M Zeleny, *Compromise programming*, Multiple Criteria Decision Making (J.L. Cochrane and M. Zeleny, eds.), University of South Carolina Press, Columbia, South Carolina, 1973, pp. 262–301.

Curriculum Vitae

Personal data

Name:	Ahmad Saher Azizi Sultan
Date and of birth:	7th February 1973
Place of birth:	Aleppo, Syria
Nationality:	Syrian
Email address:	ahmadsaher_azizisultan@yahoo.com

Education

October 2006	Ph.D. in Mathematics with specialization in *Optimization* at the university of Kaiserslautern (Germany)
WT 99/00-ST 02	Diplom (diploma) in applied mathematics at the university of Kaiserslautern (Germany)
WT 90/91-ST 98	Part time bachelor in pure mathematics at Aleppo university (Syria)
August 1990	High School degree (Syria)

Professional career

01.04.2000 - 31.12.00	Graduate assistant at Fraunhofer institute IESE (Institut Experimentells Software Engineering)
01.01.2001 - 31.10.02	Graduate assistant at Fraunhofer institute ITWM (Institut fuer Techno- und Wirtschaftsmathematik)
	Teaching assistant at the department of mathematics, university of Kaiserslautern:
WT 02/03	Algorithms and Probability
WT 03/04 and ST 05	Scheduling Algorithms
Academic year 98-99	Postgraduate lecturer at the department of mathematics, Aleppo University (Syria)