UCL Université catholique de Louvain

École polytechnique de Louvain (EPL)

Real-time replanning for proton therapy

Dissertation presented by
Charlotte CIRRIEZ

for obtaining the Master's degree in
Biomedical Engineering

Supervisor(s)
John LEE , Edmond STERPIN

Reader(s)
Ana M. BARRAGAN, Florence GIESEN , Rudi LABARBE

Academic year 2015-2016

Abstract

Proton therapy is a very promising type of external radiotherapy that is not yet used as its full potential because of its high sensitivity to errors. Especially, as the treatment is usually delivered in a fractionated way, it is sensitive to the patient inter-fraction motion. Ideally, the treatment plan should therefore be adapted before every treatment session. However, regular treatment planning is cumbersome and is subject to tedious quality assurance protocols in order to be accepted by the physician and physicist in charge. In this thesis, we propose and test a strategy for replanning in real time a proton therapy treatment plan. The goal is for the adapted plan to preserve the quality and robustness of the initial plan.

Acknowledgements

First of all, I would like to thank my supervisors, John Lee and Edmond Sterpin, for their precious advice, guidance and support that helped me produce this thesis, as well as for building in me a deep interest in proton therapy.

I will never thank Ana Barragan enough for all the time she spent helping me through every bug I had when using MIROpt and for answering to all my e-mails with the same patience and kindness.

Many thanks also to Kevin Souris for sharing his knowledge in physics and for helping me finish the thesis in time despite the endless simulations I had to perform.

Thanks to Geoffrey Schepmans for his collaboration and especially for sharing the implementation of the smart initialization.

Finally, I am thankful to my friends and family for their support, with a special mention to my flatmates who still managed me despite the different stages of stress I was going through as the deadline approached.

Contents

1 Introduction

Radiotherapy is one of the main options for treating cancers, among other methods such as chemotherapy, surgery and hormonotherapy. External beam radiation therapy, such as conventional radiotherapy or proton therapy, is typically indicated in the case of non-systemic tumors since it allows to target a specific region. Radioactive particles transfer energy to the tissues, resulting in an absorbed dose, measured in Gray (Gy = Joule/kg). They damage cancer cells by breaking their DNA chains but are also toxic for healthy tissues. Therefore, radiotherapy aims at delivering a sufficient dose to the tumor volume while sparing surrounding healthy tissues [1].

Conventional radiotherapy nowadays uses photon beams to treat cancer patients. However, using proton beams instead can improve the trade-off between delivering the prescribed dose to the tumor while minimizing the dose delivered to healthy tissues. In fact, more and more proton therapy centers are being created, showing a growing interest in the field [1].

After a short build-up region, a photon beam entering a material shows a linear-exponential decreasing energy deposition as a function of depth, as illustrated in Figure 1.1. As a result, the entrance dose (the dose deposited before the target) is always higher than the dose delivered to the target, unless we use multiple beam directions. Moreover, there is an exit dose, deposited to the tissues located after the target [1].

Protons, on the other hand, are charged particles which behave differently (see Figure 1.1 also). Energy deposition is roughly constant with depth before a sharp peak at the end of the range, i.e. where protons stop. This concentration in a well-defined region is called the Bragg peak and has a widely recognized potential. It implies that no dose is transferred to organs located behind the target, and the maximum dose can be transferred to the target with a single beam direction [1].

While protons show obvious advantages over photons, they are seldom used in the clinical environement. Indeed, they are two drawbacks for protons. The first one is the cost, which is much higher for equipments and operations in proton therapy than for X-ray based facilities [1].

The second drawback is the range uncertainty. As the range of the proton depends on the beam energy and the density of the materials in the beam path, knowing the structure of the tissues at all times is essential in order to accurately predict the position of the Bragg peak. Thus, the accuracy of the range is affected by uncertainties (in imaging, in patient set up, in dose calulation,...) that may occur during many steps of the treatment. Because of these, it is difficult to design clinical studies to assess the performance of proton therapy, as it is not yet used at its full potential, i.e. with limited uncertainties [1].

Nevertheless, proton therapy centers are getting more affordable thanks to technical progress and research in proton therapy aims at reducing range uncertainties.

For instance, robust treatment planning can help manage range uncertainties. A typical treatment planning workflow consists basically in imaging the patient, then simulating and optimizing

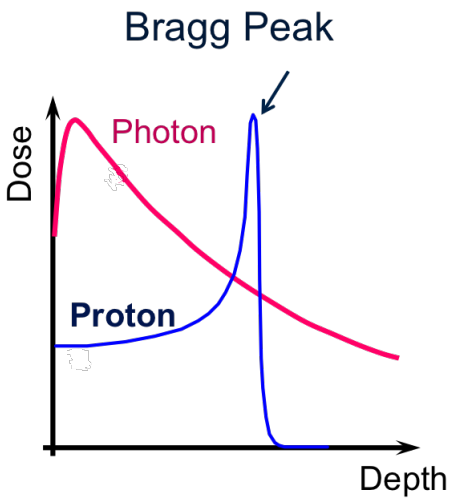


Figure 1.1: Energy deposition as a function of depth in tissue for a mono-energetic photon beam (in pink) or a mono-energetic proton beam (in blue). Credit : [2].

the dose delivery in order to find the best set of parameters that will yield a dose distribution consistent with the prescribed clinical specifications. Robustness is achieved through the integration of uncertainties into the treatment plan optimization process by defining scenarios of possible deviations from a nominal case. The plan is then optimized in order to respect the prescribed clinical specifications in all possible scenarios : it becomes robust against uncertainties. This method has shown good results [3], however the planning time increases with the number of scenarios.

A few days may occur between imaging for treatment planning and the beginning of the treatment. Moreover, the dose is most of the time fractionated and delivered one fraction per treatment session, therefore the treatment can last up to 7-8 weeks. In consequence, the images taken for treatment planning are much likely to be obsolete after some time, when taking into account morphological changes that can occur during the treatment, like tumor shrinkage, tumor motion, weight loss, variable filling of cavities (i.e. sinuses, bladder), etc. As shown on Figure 1.2, this may result in significant changes in the tissue densities and thus in proton range. The treatment plan should therefore be adapted to the new geometry. However, treatment planning, and especially robust planning, requires a non-negligible time and each new plan has to be approved by a physician and a physicist through a time-consuming quality assurance protocol. Therefore it is not imaginable to re-plan the treatment while the patient is on the table.

Image-guided radiotherapy includes on-board imaging, typically 3D computed tomography (CT), which allows to correct the position of the patient with respect to the treatment plan [4]. However a motion such as the one illustrated on Figure 1.2 can not be corrected by a simple shift. Therefore we need a real-time adaptive treatment as described in Figure 1.3 [4] for conventionnal radiotherapy: the treatment should be adapted in a few minutes based on the newest CT scan.

This thesis presents a strategy for adapting a proton therapy treatment in real time. The objective is that the adapted plan should keep the same robustness and quality as the initial plan, without the need of a new round of physician and physicist approval procedures.

We propose a method based on the concept of image registration and we define a new optimization compared to a fresh treatment planning. Image registration allows to define the geometrical mapping from the CT scan used for original treatment planning onto the new CT

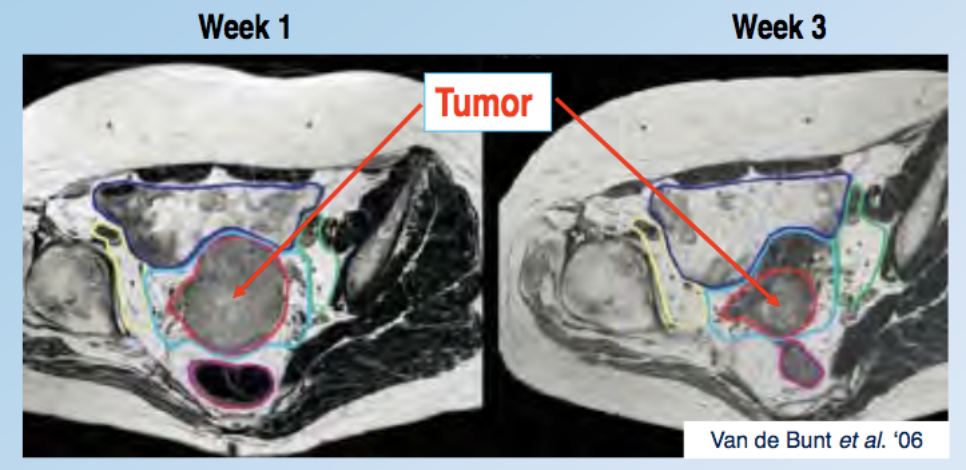


Figure 1.2: Image of a patient during first and third week of treatment. Regions of interests are delineated. From week 1 to week 3, some organs have taken more space as the tumor has shrunk, while others are in a different state of filling. Credit: [4].

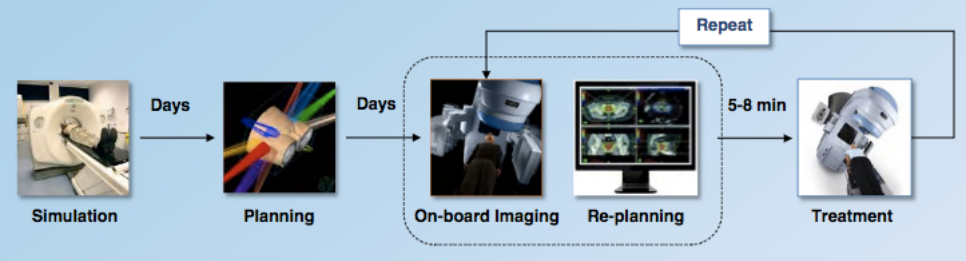


Figure 1.3: Workflow of an adaptive radiotherapy treatment. Credit: [4].

acquired just before a treatment session. This transformation is applied to the original dose distribution, resulting in a so-called reference dose. The optimization allows to find the planning parameters that will replicate the reference dose.

The thesis is organized as follows. The first chapter provides the context of the thesis which is essentially centered on proton therapy. The second chapter presents a review of the literature, and describes in detail our methods for registration, optimization and other concepts needed to understand the rest of the thesis. The third chapter explains the method in detail. The fourth chapter states the results, performed on two cases. Finally, the thesis ends with a discussion on those results according to the objectives stated above.

2 Context

This chapter aims to provide a context for the thesis. After giving more informations on radiation therapy in general, the technique of pencil beam scanning is described, then a specific treatment planning system is presented, and uncertainties are further defined in order to give the ground for a robust planning.

2.1 Generalities on radiation therapy

As mentionned before, radiation therapy relies heavily on medical imaging tools such as computed tomography (CT), which allows to see inside the patient by producing cross-sectionnal images, as illustrated in Figure 2.1. A tri-dimensionnal image can be reconstructed from a large series of 2D images. The intensity of each voxel (the smallest element of a 3D image) is proportional to the radiodensity of the tissues inside it. For useful image display, radiodensity is given in Hounsfield units (HU), that range from about -1000 HU to 3000 HU. Water in standard temperature and pression conditions corresponds to 0 HU, and air to -1000 HU [5].

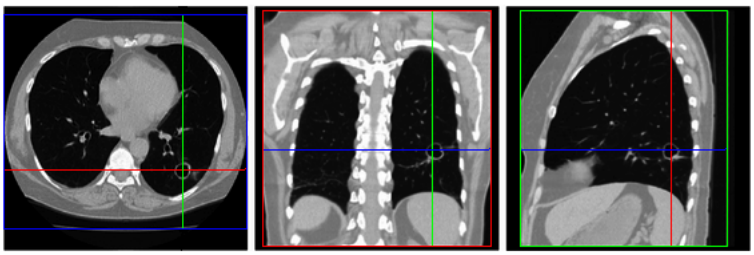


Figure 2.1: Axial (blue), coronal (red) and sagittal (green) sections of a CT scan showing lungs. Credit : [6]. Colored lines show the links between each views, i.e. the coronal view corresponds the slices shown in red in the other sections.

In order to ease the exchange and processing of those images, an international standard has been developped: DICOM (digital imaging and communications in medecine). It defines the formats for medical images. Thus, among other things, it provides a reference frame for the CT coordinates which is illustrated in Figure 2.2 [7].

Such images from the patient's CT scans allow to detect the tumor and plan the treatment. As it was illustrated in Figure 1.2, the tumor, as well as surrounding organs, are delineated. The structures defined this way are called regions of interest (ROI) as they play a role in the treatment planning. Among ROIs, organs that cannot absorb too much dose are called organs at risk (OAR).

For the tumor, three volumes are considered : the gross target volume (GTV), the clinical target volume (CTV) and the planning target volume (PTV). They are illustrated in Fig. 2.3. The GTV corresponds to the tumor as it can be seen on the CT. The CTV is an expansion of

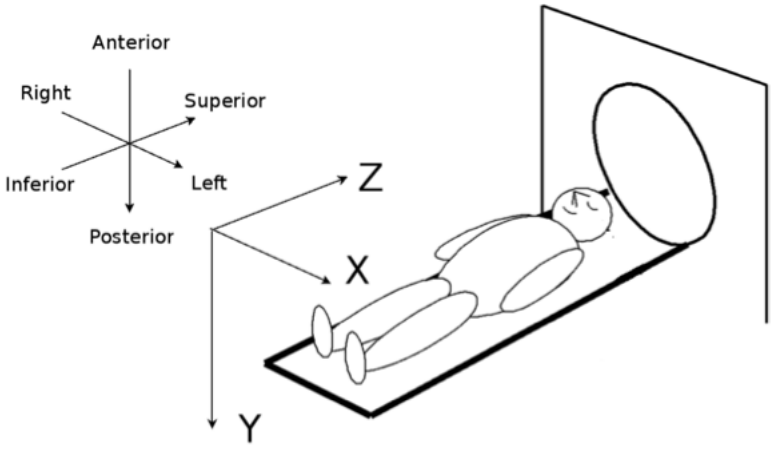


Figure 2.2: DICOM coordinate system. Credit : [8]

the GTV taking into account the microscopic tumor spread that is not imageable. The PTV further expands the CTV and takes into account uncertainties [9].

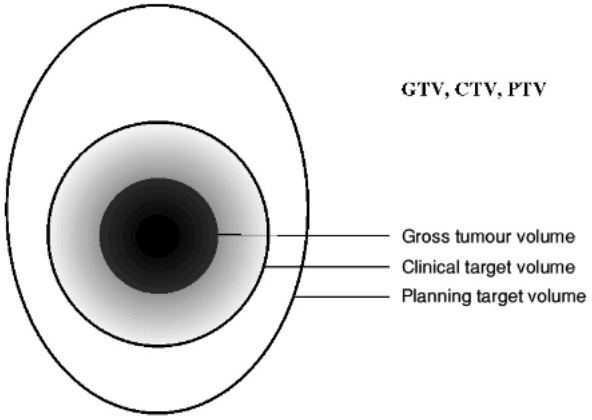


Figure 2.3: Illustration of the three volumes surrounding the tumor : GTV, CTV and PTV. Credit : [9]

Treatment planning is then performed in a few steps :

1. *Delineation of ROIs* : The physician must draw the contour of the tumor and organs manually on all slices, although research aims at developping automatic segmentation techniques in order to speed up this process. Then each structure is labeled, and the tumor is expanded into the appropriate volumes.

2. *Prescription* : The physician defines the prescription to the target volume, the constraints to the OARs and the fractionation scheme considered (i.e. the number of treatment fractions) [10].

3. *Optimization* of the parameters of dose delivery in order for the treatment to be consistent with the prescription and dose constraints.

4. *Simulation* of dose delivery in the patient's geometry. This helps ensure that the parameters have been well optimized.

During the treatment, many checks are done to ensure that the patient will be exactly in the same position relative to the beam delivery unit. Molds or masks may be constructed for a

patient to help them staying still. As mentionned before, on-board imaging devices are also used in the treatment room to help achieve precise patient positionning [10].

External beam radiation therapy can come under many forms. Mainly the type of particles used and the optimization method vary. The type of therapy prescribed depends on a lot of factors such as the type of cancer, its size and location, etc.

In conventional radiotherapy, electrons are accelerated by a linear accelerator (LINAC) and strike a target (typically tungsten) to produce a beam of X-rays. The most common method for delivering photons is 3D conformal radiation therapy (3D-CRT), using rotating gantries with collimators and computer software to shape the beams so they match precisely the target. Another method is called intensity modulated radiation therapy (IMRT). It uses tiny radiation beam-shaping devices (collimators) to deliver single doses of radiation, which intensities can be optimized in order to fit the tumor [10].

In proton therapy, the particle accelerator is either a cyclotron or synchrotron. They use strong electric fields to accelerate protons up to an energy of 230 or 250 MeV and strong magnetic fields to keep them moving in a circular path. The main difference between the two types is that cyclotrons use a constant magnetic field and frequency electric field, while synchotrons use varying magnetic and electric fields [1].

When exiting the accelerator, particles are guided through a beamline to the treatment room. There, two methods can be used for delivering protons : passive scattering or pencil beam scanning (PBS). Passive scattering uses scattering materials in order to spread out the beam in three dimensions and to match the target. PBS uses a very narrow beam which is scanned accross the target volume [1].

Electron beams can also be used to irradiate superficial tumors [10].

2.2 Pencil Beam Scanning

The state-of-the-art method for delivering proton therapy treatments is PBS. As mentionned above, a very narrow beam is scanned accross the tumor volume. To cover the tumor in the direction parallel to the beam, the energy is changed (typically between 70 to 250 MeV) to place the Bragg peak at the right depth. Scanning is thus used for conforming the tumor volume in directions perpendicular to the beam axis [1].

Energy modulation can be done in two ways. If the accelerator is a synchrotron, protons can be ejected with a specific energy. However, most facilities use a cyclotron, which accelerates all particles up to a specific energy. The energy is lowered afterwards by guiding the beam through a material of specific thickness, called a range shifter. [11].

For the transverse directions (directions perpendicular to the beam direction), two dipole scanning magnets are used to bend the beam. Basically, it works as follows : the beam in position A is targeting a spot in the tumor volume defined by energy an E_A (the energy needed by the beam to reach the spot in depth) and by "transverse" coordinates (x_A, y_A). Those are expressed in beam eye view coordinates (BEV), which have a z-axis parallel to the beam direction and are centered on the beam isocenter, which could be chosen as the center of the target for instance. This is illustrated in Figure 2.4.

The beam stays in position A for a certain amount of time t_A which is proportional to the number of protons delivered, and thus to the absorbed dose. Then, the beam is turned off and placed at an adjacent position B, and so on until the whole volume is covered [1]. Figure

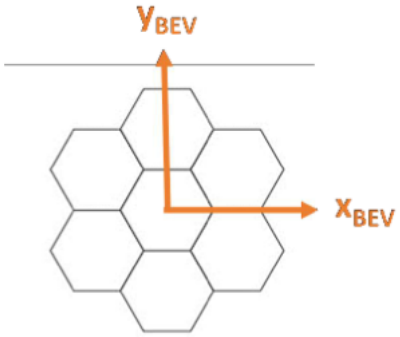


Figure 2.4: BEV coordinate system. Credit : [8]

2.5 illustrates the pencil beam going through a tumor. A pencil beam reaching for one spot is sometimes refered to as a beamlet.

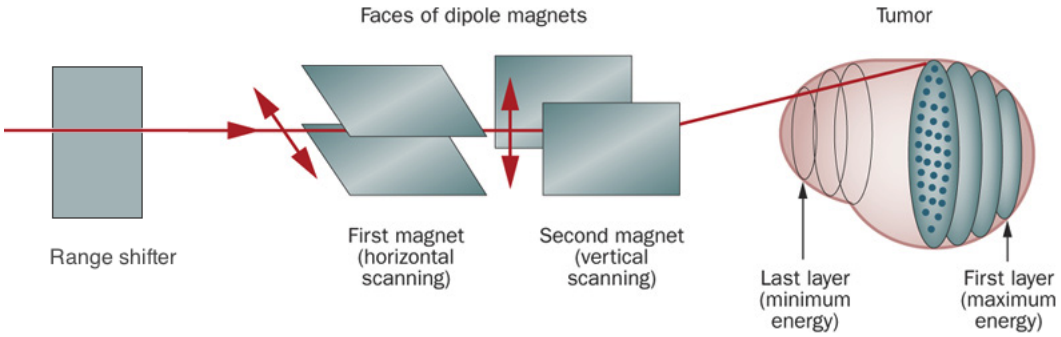


Figure 2.5: Illustration of pencil beam scanning. The range shifter reduces the beam energy according to its thickness. One magnet controls the horizontal position while the other controls the vertical position in order to reach each spot in the tumor. Credit : [12]

The depth-dose distribution will be determind by the superposition of the Bragg peaks, creating a so-called spread-out Bragg peak (SOBP), as illustrated in Figure 2.6.

An important variable is the time during which the beam is on in a certain position (i.e. t_A), since it determines the intensity of dose delivery at that spot (or the height of the corresponding Bragg peak as in Figure 2.6). This parameter gives a degree of freedom to the treatment planning, which can be optimized in order to yield the best possible dose trade-off. With this degree of freedom, the treatment method is also called intensity-modulated proton therapy (IMPT).

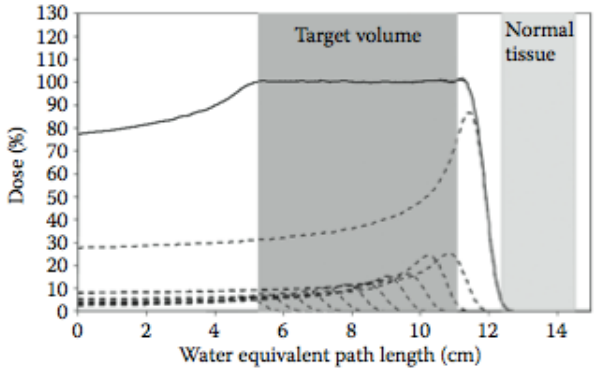


Figure 2.6: Schematic view of SOBP construction, showing the SOBP depth–dose distribution (solid line) and the component Bragg peaks (dashed lines). Credit : [1]

2.3 Treatment planning

This subsection presents a treatment planning system (TPS) for proton therapy using PBS called MIROpt (MIxed Robust Optimizer for PBS) and developped by Ana Maria Barragan Montero in the MATLAB environment [8]. The diagram presented in Figure 2.7 summarizes the main steps.

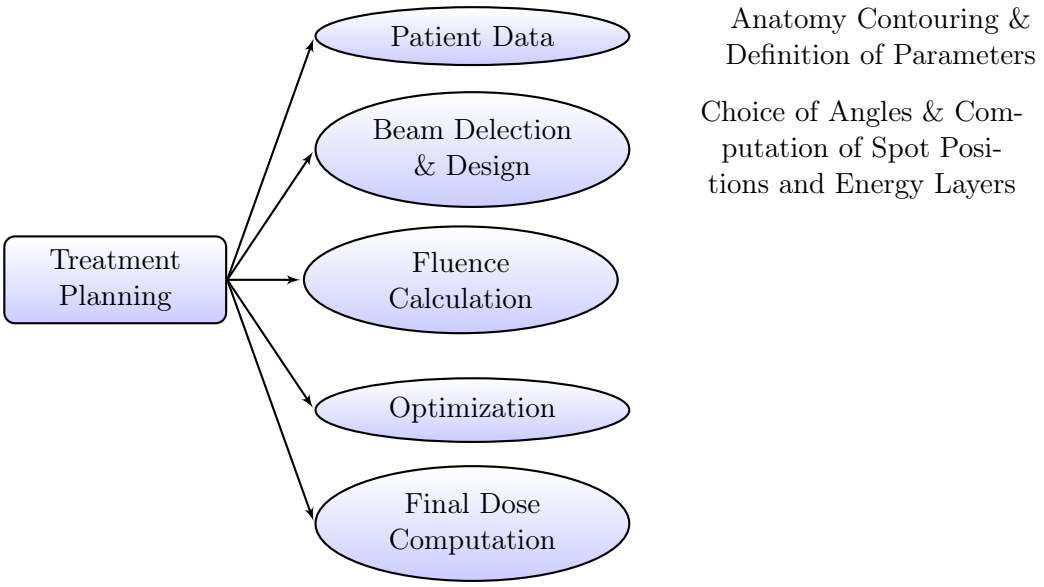


Figure 2.7: Summary of the main steps in the TPS MIROpt.

Patient data

The user imports a 3D CT in DICOM coordinates along with the contours of the ROIs. They can define the maximum or minimum dose allowed in each of them [8]. Since robust planning is used, the tumor should be represented by the CTV and not the PTV. This will be further discussed in section 2.4.

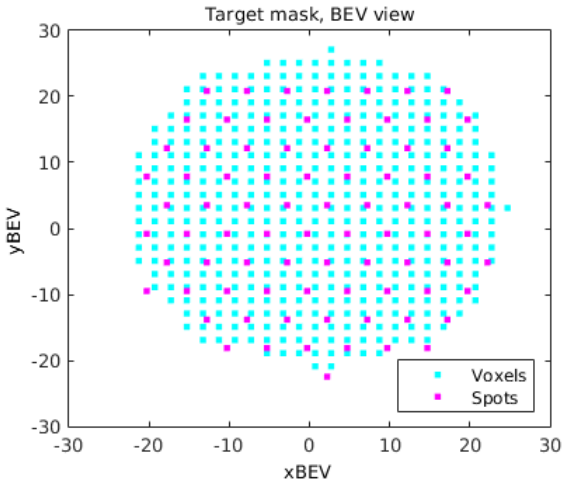


Figure 2.8: Illustration of a target mask for a spherical target in BEV coordinates, with voxel spacing 2×2 mm and spot spacing 5 mm on a hexagonal grid. The spots are represented by the pink dots and voxels by light blue dots.

Beam selection and design

The user must specify the incident angle of each beam. In MIROpt, this angle lies in the xy plane in DICOM coordinates.

For each beam direction, the positions of all Bragg peaks within the target must be determined. As mentionned before, a spot position is defined by an energy and two transverse coordinates. The spacing between those positions will depend on the characheristics of the beam delivery system : the resolution of the energy selection, the maximum deliverable energy, the width of the beam, etc [1].

From the CT, a target mask is created: it is a 3D array of the same size as the CT, in which all voxels inside the target are assigned an intensity equal to one (see Figure 2.8). This mask is a little expanded in order to account for penumbra and possible errors [8].

Then, the spots are placed:

- In the beam direction, the target is divided into discrete energy layers according to a user-defined energy spacing.
- In the transverse directions, the target is divided according to a hexagonal spotgrid as shown by the pink dots on Figure 2.8.

As mentionned in section 2.2, this spotgrid is expressed in BEV coordinates (Figure 2.4) whose origin is the center of the target. In axial view, each spot (that is, a Bragg peak position) has six neighbors [8].

The discretized energy layers are obtained by computing the lowest energy needed and the highest energy needed to cover the expanded target mask. Then, the layer spacing given by the user will determine the number of energy layers. But before that, the energy needed to reach each voxel in the expanded target mask must be computed. The primary information necessary to compute these energies is the intensity of each voxel in the CT (in HU). A simple linear interpolation can convert HU into stopping power ratio (SPR), provided the scanner calibration curve and the list of materials [8]. The SPR refers to the ability of the voxel to slow down the proton beam relative to water [13]. For each beam direction, SPR values are converted into water equivalent path length (WEPL), which would be the distance reached by the beam if it were going through water only. This is done through a ray-tracing algorithm as illustrated in

Figure 2.9. From each voxel, the inverse beam direction is followed until the whole CT has been crossed, and the traversed voxels are given a WEPL value. WEPL values W are then converted into energy E by means of the equation

$$E = \exp(3,464048 + 0,561372013\log(W) - 0,004900892(\log(W)^2 + 0,001684756748\log(W)^3)) \tag{2.1}$$

which is a fitting from NIST[1]/ICRU [2] database values [14].

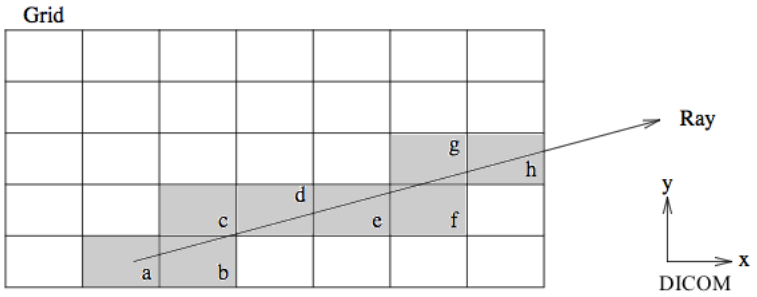


Figure 2.9: Illustration of the raytracing algorithm. From one voxel (a), the beam direction is followed and all voxels in the beam path are associated with a WEPL value. Credit: [15]

Fluence calculation

The relative fluence of each Bragg peak is how each of them contribute to the dose [16]. This is computed by simulating the behavior of protons within the patient's geometry. The section 3.1 will describe methods used to achieve such simulations. The information can then be stored in a matrix $P \in \mathbb{R}^{N\times n}$ called dose contribution matrix or beamlets matrix [8], with N the number of voxels in the CT and n the number of spots. Thus, one column b_i of this matrix refers to a beamlet reaching the ith spot (as described in subsection 2.2), and contains the dose delivered to each voxel by this beamlet, given that it is turned on for a unitary time.

Optimization and final dose calculation

Finally, the dose is computed by multiplying the dose contribution matrix P with a weight vector $x \in \mathbb{R}^n$. Each element of this vector is associated to a spot and is proportional to the time during which the corresponding beamlet should be in position. Hence, the dose to each voxel $D \in \mathbb{R}^N$ is given by

$$D = P \cdot x \tag{2.2}$$

Therefore, it is very important to optimize the vector x in order to obtain the prescribed dose in the target volume while respecting the constraints to OARs. This problem is known as inverse planning [17]. How this is done in MIROpt is explained in detail in section 3.2.

2.4 Uncertainties

As introduced before, proton beam delivery is subject to uncertainties than can arise at different steps of the process and that may induce huge errors in the treatment plan. Indeed, the main advantages of proton therapy can turn into disadvantages when we are dealing with uncertainties, as illustrated in Figure 2.10. When introducing a 2 cm air gap before the tumor in the beam path, the SOBP is shifted by 2 cm in the beam direction. The same perturbation does not have such a strong effect on a photon beam.

[1]National Insititute of Standards and Technology.
[2]International Commission on Radiation Units and Measurements.

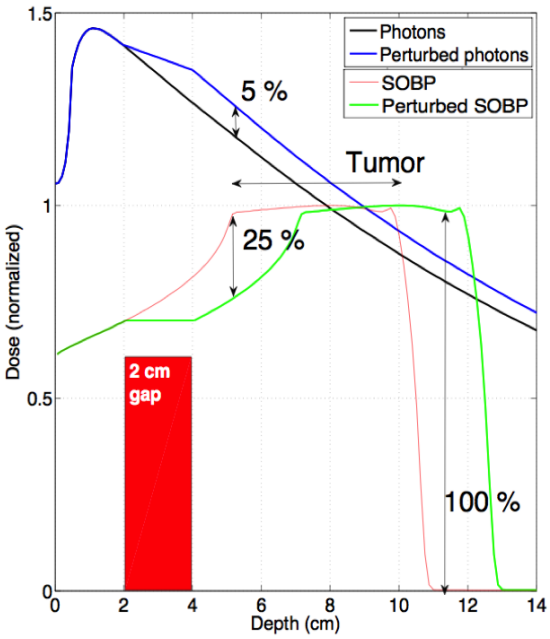


Figure 2.10: Dose-depth curve comparison between a photon beam and a spread-out Bragg peak (SOBP) before and after perturbation introduced by a 2 cm air gap in front of the tumor. Credit : [2].

An important cause of range uncertainty is the uncertainty in computing the SPR. However, the contribution of other several factors, such as biological uncertainties, or the fact that approximations of physical models are used in order to compute the dose, is non-negligible [1]. Another type of uncertainties to take into account are the set up errors, arising from patient positionning, immobilization of the patient, etc [1], but also from movements occurring during treatment delivery and imaging artefacts. In MIROpt, set up errors are divided into two subtypes to account for those two error possibilities : systematic set up errors, and random set up errors [8]. Some sources of uncertainties are summarized in Table 2.1.

In conventional radiotherapy, robustness against errors is achieved by expanding the CTV with a margin of a few millimeters (the PTV). The underlying assumption is the so-called "static dose cloud approximation", which states that uncertainties might shift the dose but the distribution itself will not change. However, this approach is no longer valid for proton therapy, especially

Source of uncertainty	Magnitude
Stopping powers	±1.3 mm
Range reproductibility	±1 mm
Accessories	±1 mm
CT	±3.5 % of range
Patient setup	±1.5 mm
Intrafraction patient motion	Variable
CT artifacts	Variable
Target coverage	±1-3 mm
Dosimetry	±1-3 mm and ±1-3 %

Table 2.1: Summary of estimated uncertainties in treatment planning and delivery of nonmoving targets with proton therapy. Credit : [1]

for PBS, which implies the need of using robust optimization to achieve robust plans [18].

As introduced before, the idea of robust planning is to include errors in the optimization process by specifying intervals of possible deviations for all three types of errors, i.e. range errors, systematic and random set up errors.

Systematic set up errors are simulated by moving the spot weights to adjacent positions [8]. This is not similar to using margins since shifting a spot laterally will induce a shift of the corresponding beam which will then go through a volume of different density. As a consequence, the spot will be also shifted longitudinally [19]. In MIROpt, the user defines the maximum possible shift in the three directions.

Random set up errors are simulated by sampling random shifts from a gaussian probability distribution with a standard deviation specified by the user [8].

Range errors are simulated by scaling the mass density of the treatment volume uniformly by a percentage, also specified by the user [8].

The optimization process then aims to minimize the consequences of the worst case scenario [19]. How this is done is presented in detail in section 3.2.

Scenarios are made of coherent combinations of error parameters. There are three density scalings corresponding to overshoot, undershoot and the nominal case. Therefore, three contribution matrices have to be computed. As the set up error is assumed to be isotropic, and the scanning grid is hexagonal, each beam has seven possible shifts positions (a spot can stay in position or move to one of its six neighbors). The same dose contribution matrix can be used for different shifts [19]. However, as different spot positions are used for each scenarios, the number of spots in a robust planning is way bigger than for regular planning.

In conclusion, robust optimization increases the computation time in three ways : three beamlets matrices must be computed, more spot positions are considered, and more constraints are integrated in the optimization.

In addition to robust optimization, progress in fast and accurate medical imaging can improve patient positionning, which will reduce set up errors. Moreover, extended accuracy in computing the dose contribution matrix may also reduce uncertainties in region of high heterogeneities.

3 Literature review

3.1 Dose calculation

As stated before (see section 2.3), dose calculation consists in computing the dose contribution matrix. In order to do so, an estimation of the behavior of protons within the CT is needed. In this section, the physics of protons is quickly described, then a brief overview of the existing dose calculation methods is presented and finally, more details about the specific method used in MIROpt are given.

3.1.1 Protons physics

The interesting property of protons is that they lose more energy as they slow down, with a large dose enhancement just before they stop, resulting in the Bragg peak. They slow down mainly through multiple head-on collisions with atomic electrons [1]. Therefore, the rate at which protons slow down depends on the type of material they are going through, which can thus be represented by a stopping power.

However, electromagnetic (EM) interactions have other consequences as well.

First, protons are deflected due to multiple collisions with atomic electrons and nuclei. The effect of a single collision is very small, but the contribution of multiple collisions results in a non-negligible deflection angle. This is known as multiple Coulomb scattering, which can be described by a statistical distribution [1].

Moreover, monoenergetic protons may also stop at different depths, given the statistical nature of EM interactions. This effect is called range straggling, and causes the narrow Bragg peak predicted by stopping theory to have increased width [20].

The second type of interactions is the nuclear interaction : protons experience a head-on collision with a nucleus. These interactions occur seldom and can be elastic or non-elastic. In an elastic collision, the proton is deflected by a non-negligible angle. In an inelastic collision, the nuclei is destroyed and several secondary particles are emitted. Those secondary particules might be protons, neutrons, light nucleon clusters or photons called prompt gammas [1]. Prompt gammas are interesting since they can be used to measure the range of protons inside the patient during treatment. This possibility is still being explored but could be promising in order to have a feedback on the range error [21].

3.1.2 Dose calculation methods

Dose calculation methods used in a clinical environement must be a trade-off between speed and accuracy. Among existing methods are the analytical methods and Monte Carlo methods. Both types will be briefly overviewed here.

Analytical methods

The first, simplest analytical method is the ray tracing or broad beam method. It can be used for dose calculations involving little anatomical heterogeneities. A pre-calculated dose distribution in water is scaled by the water equivalent path depth of each spot along the beam direction. This method relates to what is used in MIROpt in order to place the spots. It is very fast, but it does not account for the scattering and only models energy losses [20].

A more accurate analytical method is the pencil beam algorithm (PBA). The beam is divided into smaller beam elements that are individually transported through the patient. The final dose is the sum of the resulting beam dose distributions [20]. The dose distribution due to one beam element is computed as a set of interactions between the beam and the patient and may include Multiple-Coulomb scattering effects, which is another improvement from the simple ray tracing method [20]. However, PBA is less sensitive to complex geometries and density variations than Monte Carlo methods [1].

Monte Carlo methods

Instead of computing the dose for a beam element composed of many thousands protons, Monte Carlo (MC) methods track each proton individually through the material, which is why they are more accurate but also slower.

The basic idea of MC methods is to estimate a value based on its known probability distribution by sampling it a sufficient amount of times. As an example, consider a rectangular zone with known area A containing a lake of unknown area. Then, consider n balls being randomly and uniformly thrown all over the rectangular area. The number of balls on the ground is counted, x, thus the number of balls in the lake should be $n - x$. The area of the lake can be estimated as $A \times \frac{n-x}{n}$ [22]. With this example, it is easy to understand that the more balls are thrown, the more accurate the estimation may be. The probability distribution in this case is uniform.

Thus, MC simulations are based on numerous random samplings from a known distribution. Therefore, to estimate the trajectory of a proton, an interaction probability distribution is defined, based on its properties (as described in the previous section) and the traversed material properties (such as electronic density).

Originally, MC simulations define the trajectory of a particle by simulating its energy, position and direction after each collision. Then, the simulation randomly samples from the distribution in order to choose how the particle might interact next. Since every collision is taken into account, this is called a detailed history MC simulation and it is very slow [20].

Instead of simulating every single collision, condensed history MC simulations divide the particle trajectory into small steps, based on cross sections per unit distance [1]. At each step, EM interactions involving little energy loss are condensed and the resulting deflecting angle is predicted by multiple-Coulomb scattering theory. The energy lost in one step is estimated by the stopping theory. EM interactions involving a sufficient energy loss are individually simulated. Since nuclear interations are more unusual, they are also individually simulated [21]. This method is faster than detailed history MC simulations but still slow [20].

Finally, track-repeating algorithms have been developped. Trajectories are pre-computed in water and scaled to other materials by their SPR. This further improves the calculation speed, however it requires to store a large amount of trajectories which uses a lot of computer ressources [20].

	Analytical methods	
Method	Ray-Tracing	PBA
Calculation speed	Very fast	Fast
Accuracy	Accurate (except for heterogeneities)	Accurate

	Monte Carlo methods		
Method	Detailed history	Condensed history	Track repeating
Calculation speed	Very slow	Slow	Fast
Accuracy	Very accurate	Very accurate	Very accurate

Table 3.1: Summary of analytical and MC methods.

While very accurate, Monte Carlo methods either require a long computation time or use a lot of memory, and therefore are currently not used in the clinical environment. Indeed, when using sufficient margins, analytical methods are accurate enough to perform treatment planning. However, MC methods present the advantage of precisely computing the dose in very heterogeneous regions, and may simulate random events (for instance random setup errors, breathing) with no additionnal cost in treatment time. Table 3.1 shows a summary of the presented methods.

3.1.3 MCsquare

Kevin Souris has designed a new Monte Carlo code called MCsquare, which stands for multi-core Monte Carlo [21]. This is the code used for the dose contribution matrix computation in MIROpt. For better visualization, a diagram is presented in Figure 3.1.

The purpose of this new implementation was to significantly reduce the computation time of a condensed-history MC simulation, while keeping a reasonnable accuracy. In order to achieve this goal, Kevin Souris proposes a simplification of the physical models and an optimization of their algorithmic implementation. As a result, comparisons with a simulation program called GATE [23] showed reasonable accuracy as well as a reduced computation time [21]. MCsquare can also simulate the emission of prompt gammas.

The simplification of the physical models involves several changes compared to GATE. Instead of being computed at each step, physical models are approximated using a database or a fit. Moreover, same physical models are used to compute the trajectories of protons and secondary particles. This allows a parallelization of the code. Finally, the simulation is made inside a voxelized geometry, in which each voxel is labeled with the appropriate material according to its density [21].

The algorithmic implementation was optimized by using Intel Xeon Phi coprocessors. Those processors have a parallell architecture and possess many independent calculation units, which enables the simulation of several particles at a time and the distribution of tasks among CPUs (Central Processing Units). They are a unique trade-off between cost, power and flexibility [21].

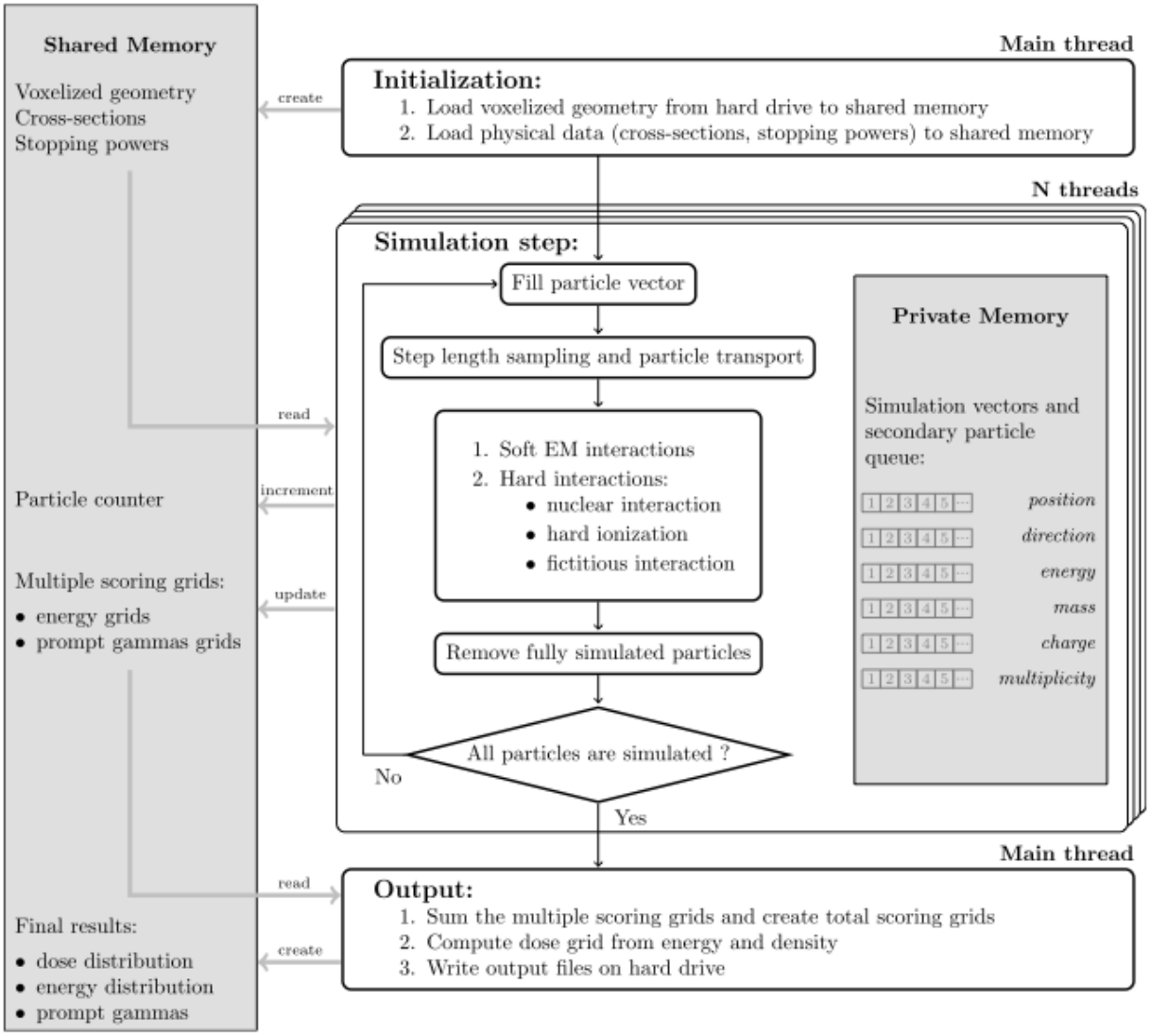


Figure 3.1: Diagram representing the main steps of the MCsquare algorithm. Credit : [21].

3.2 Optimization

Once the dose contribution matrix has been computed as explained in section 3.1, the weight of each beamlet has to be optimized so the final dose can respect the treatment constraints. This results in quite a large optimization problem given that the number of spots can be up to several thousands, especially if we include robustness.

In this section, the general problem of optimization is first introduced in order to set a few basic concepts. Then, the specific dose optimization problem is further described, followed by the minimax problem, which allows to include robustness. Finally, the specificities of the algorithm used in MIROpt (an interior-point filter line search algorithm) are presented.

3.2.1 General problem of optimization

The general formulation of an optimization problem is given by

$$\min_{x \in S} f_0(x) \tag{3.1}$$
$$\text{subject to } f_j(x) \geq 0 \; \forall \; j = 1, \ldots, m$$

where x is the vector of variables subject to optimization (in our case, the weights of each beamlet), S is the set over which those variables are defined, f_0 is the function to minimize, which we call the objective function, and f_j are the functionnal constraints [24]. Some problems are not subject to functionnal constraints, then they are called unconstrained.

There exists a numerous amount of methods to solve this problem which can most of the time be classified according to the type of objective function (i.e. differentiable or not), the type of set S (convex, closed, etc), or the type of constraints (if they exist) [24].

Most algorithms for solving optimization are based on an iterative process which tries to bring an iterate x_k towards an optimal value x^*. A new iterate is computed according to a step size parameter, which will determine how close two subsequent iterates are from each other, and a search direction, which is the direction supposed to lead the algorithm towards x^*. A typical iterative process for an unconstrained problem can be described as follows :

1. Choose $x_0 \in S$, a stopping criterion and compute $f(x_0)$.
2. While $k \geq 0$ and the stopping criterion is not satisfied, iterate :
   (a) Compute a new search direction p_k.
   (b) Compute a new step size parameter h_k or use a constant step size $h_k = h$.
   (c) Compute the next iterate as $x_{k+1} = x_k + h_k p_k$.
   (d) Compute $f(x_{k+1})$.

The computation of the search direction is crucial. The simplest and most intuitive method is the gradient method, in which $p_k = -\nabla f(x_k)$. Indeed, the anti-gradient is the direction of the fastest local decrease of f at that point. This method is also called gradient descent.

3.2.2 Dose optimization problem

An optimization problem is thus defined by its objective function, its constraints and its definition set.

The function to minimize must ensure that the prescribed dose is delivered to the CTV and that the dose in each OAR does not exceed a given dose. Therefore, it will measure the difference between the prescribed dose and the computed dose at each iteration of the optimization process.

The importance of respecting the dose in each ROI can be variable, therefore, the objective function f is a weighted combination of m optimization functions f_i [17] :

$$f(x) = \sum_{i=1}^{m} \text{impw}_i f_i(x) \tag{3.2}$$

where x is the vector of weights to optimize, and impw_i is the importance accorded to the ith function.

Each optimization function is either used to minimize the dose or to maximize it according to a given dose parameter d_i. For the CTV, two functions are needed since we want exactly d_i in the whole volume. Therefore, penalty is given for voxels with lower dose but also for voxels with higher dose. For the OARs, only a max-type function is used. Therefore the functions $f_i(x)$

compute the normalized square dose deviation from d_i for all voxels in the ith ROI [17] :

$$f_{i,min}(x) = \frac{1}{\text{nvoxels}_i} \sum_{j=1}^{\text{nvoxels}_i} (\max(0, d_i - D(j))^2) \tag{3.3}$$
$$f_{i,max}(x) = \frac{1}{\text{nvoxels}_i} \sum_{j=1}^{\text{nvoxels}_i} (\max(0, D(j) - d_i)^2)$$

where nvoxels_i is the number of voxels in the ith ROI, and $D(j)$ is the dose computed in the jth voxels of the ROI such as $D(j) = P_j x$, where P_j is the jth line of P, the beamlets matrix.

The definition set is defined by the physical meaning of the variables, hence they must be non-negative : $x \geq 0$.

The constraints are only introduced by the minimax method, which is presented next.

3.2.3 Minimax problem

As discussed in section 2.4, a robust optimization takes into account uncertainties which have been defined as random set up errors, systematic set up errors and range errors. This section explains how this is done in MIROpt, as introduced by Fredriksson [3].

As mentionned before, the purpose of robust optimization is to minimize the consequences of uncertainties. Scenarios are thus defined in order to include uncertainties. In the optimizer's point of view, the worst case scenario corresponds to the scenario for which the value of the objective function is the largest. Thus, let $\mathcal{S}$ be the set of scenarios, $P(s)$ be the dose contribution matrix for scenario $s \in \mathcal{S}$ and let $D(x, s) = P(s)x$ be the corresponding dose distribution. The objective function thus becomes also dependent on s and the problem can be defined as

$$\min_x \max_{s \in \mathcal{S}} \{f(x, s)\} \tag{3.4}$$
$$\text{subject to } x \geq 0 \ .$$

When introducing the variable $t \in \mathbb{R}$, corresponding to the objective function of the worst case scenario, the problem becomes

$$\min_{x,t} t \tag{3.5}$$
$$\text{subject to } t \geq f(x, s) \quad \forall s \in \mathcal{S}$$
$$x \geq 0 \ .$$

Therefore the problem becomes constrained. This is known as the minimax formulation since we are trying to minimize the maximum value of the objective function.

3.2.4 Interior-point filter line-search algorithm

The optimization of the weights vector x is done in MIROpt by means of an open-source software for large scale non-linear optimization: IPOPT. It uses an interior point algorithm with a filter-line search method [25]. This section will explain the main characteristics of the algorithm.

The problem to solve is of the form [1]

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} f(x) \\ &\text{subject to } c(x) = 0 \\ &x \geq 0 \end{aligned} \tag{3.6}$$

The objective function $f : \mathbb{R}^n \to \mathbb{R}$ and the equality constraint $c : \mathbb{R}^n \to \mathbb{R}^m$ with $m \leq n$ are assumed to be twice continuously differentiable on $\mathbb{R}^n$ [25].

Interior-point method

The interior-point method or barrier method consists in reformulating the problem as

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} f(x) - \mu \sum_{i=1}^{n} \ln x_i \\ &\text{subject to } c(x) = 0 \end{aligned} \tag{3.7}$$

where μ is a positive parameter. This new formulation gets rid of the inequality because it creates a barrier for non positive variables. Indeed, let us note that [26]

$$\lim_{x_i \to 0} \ln(x_i) = -\infty \tag{3.8}$$

Filter method

The filter method also consists in reformulating the problem. One could reformulate the objective function by using a penalty function which is a linear combination of the objective function f and some measure of the constraint violation [27]. By doing this, we get rid of the constraints, since they are taken into account in the objective function, which leaves us with an uncontrained problem. Using a penalty function is similar to what is done in the barrier method for the inequalities. However, introducing such a term requires to define an penalty parameter, which must be chosen wisely.

As an alternative to penalty functions, the filter method does not require the introduction of a penalty parameter. Instead, it defines two objective functions, one being the actual function f and the other being the constraints function c. Between those two objective functions, the filter method gives priority to c since we must find a point for which $c(x) = 0$ [27].

Limited-memory BFGS method

In IPOPT, the search direction is computed using limited-memory BFGS method.

As mentionned before, the simplest way to optimize a function is the gradient method. Another simple and well known method is the Newton method. This one uses more information about the objective function f since it also uses the second derivative of the function (the Hessian), which gives the topology of the function: the search direction is thus given by $p_k = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k)$ [28].

The BFGS method is a quasi-Newton method, in which the search direction p_k is computed such as $p_k = -H_k \nabla f(x_k)$, where H_k is an approximation of $[\nabla^2 f(x_k)]^{-1}$. Indeed, computing the Hessian as it is done in the Newton method can be time consuming and approximating its inverse allows to still use its information while reducing computation time [28].

The limited-memory BFGS method aims to save memory and computation time when the dimension n of the problem is large, as it does not store the entire approximated matrix, but only stores information from the $m << n$ past iterations [29].

[1] The method can be applied to the one defined in Eq. 3.5, for instance inequality constraints are converted in equality constraints by introducing slack variables.

Line search method

The line search method refers to the computation of the step size parameter. It allows to ensure that $f(x_{k+1}) = f(x_k + h_k p_k) < f(x_k)$ [30]. This is a way of compensating for approximating the Hessian.

Stopping criterion

The last important characteristic in IPOPT is the stopping criterion. Basically, an optimality error E is computed : when it becomes lower than a certain value, the algorithm stops. In order to explain how the optimality error is defined, let us briefly introduce the method of Lagrange multipliers, which, for a general problem as in Eq. 3.6, consists in defining the Lagrangian :

$$\mathcal{L}(x, \lambda) = f(x) - \lambda c(x) \tag{3.9}$$

where $\lambda \in \mathbb{R}^m$ is the vector of the Lagrange multipliers associated with each constraint $c_i : \mathbb{R}^n \to \mathbb{R}$. The Lagrange multipliers have the following property [28] :

Property 1. *If x^* is a local extremum of f and if the set formed by the gradients of each constraint $\nabla c_i(x^*)$ is linearly independent, then there is a vector λ^* such as (x^*, λ^*) is a critical point of $\mathcal{L}(x, \lambda)$* [2].

When applying this to the barrier problem (Eq. 3.7), we get

$$\mathcal{L}(x, \lambda) = f(x) - \mu \sum_{i=1}^{n} \ln x_i - \lambda c(x)$$
$$\nabla_x \mathcal{L}(x, \lambda) = \nabla f(x) - \mu X^{-1} e - \lambda \nabla c(x) = 0$$

where X is the diagonal matrix whose elements along the diagonal are the elements of the vector x and $e = (1, \ldots, 1)^T$ [26]. By introducing the vector $z = \mu X^{-1} e$ and its corresponding diagonal matrix Z, three equations can be derived that are equivalent to the problem 3.7 :

$$\nabla f(x) - \lambda \nabla c(x) - z = 0 \tag{3.10}$$
$$c(x) = 0 \tag{3.11}$$
$$XZe - \mu e = 0 \tag{3.12}$$

The optimality error E is thus given by :

$$E_\mu = \max \left\{ \frac{||\nabla f(x) - \lambda \nabla c(x) - z||_\infty}{s_d}, ||c(x)||_\infty, \frac{||XZe - \mu e||_\infty}{s_c} \right\} \tag{3.13}$$

where s_d and s_c are scaling parameters. This error is computed for a given μ which is driven towards zero [25].

3.3 Standards for dose characterization and quantification methods

In this section, the standards used to define the parameters of the dose optimization (maximum or minimum dose in each ROI) will be presented as well as the methods used to characterize the dose computed by MIROpt. These quantification methods will determine if the newly planned treatment is similar to the original one.

[2] meaning that the gradient $\nabla \mathcal{L}(x^*, \lambda^*)$ is nul.

3.3.1 Dose-Volume histograms

Dose-volume histograms (DVH) are a graphical representation of the dose in a given volume. The x-axis corresponds to the dose while the y-axis corresponds to the percentage of volume. It allows to describe the dose delivered in the CTV and in OARs. Of course it depends on the quality of the delineation of each ROI [31].

The DVH allows for some characteristic values to be determined. In the CTV, d_{95} is the minimum dose delivered to 95% of the CTV and should be higher than 95 % of the prescribed dose; d_5 is the maximum dose delivered to 5% of the CTV and should be lower than 105% of the prescribed dose. This ensures that the dose delivered to the tumor respects the clinical specifications. For organs at risk, the important values given by the DVH will depend on the organ's structure, as will be explained in the next subsection.

3.3.2 Standards for dose prescription

Organs at risk As said before in section 2.3, before the optimization, organs at risk (OARs) are delineated and a maximum dose value is assigned to each of them. This value depends on their structure and composition. Indeed, each organ can be defined according to the way the tissue is organized : serial or parallel structure [32]. Depending on the type of organs, the way we define the maximum dose will be different.

An organ with a serial structure means that its function depends on all its subunits : if one of them does not work, the whole function is lost [32]. Typically, the spinal cord is a serial organ, since a lesion at one point would disrupt all connections involving nerves passing through that point. The analogy can be made with a series circuit. For serial organs there is therefore a critical dose d_{max} not to exceed anywhere in the whole organ volume [32].

A parallel organ is made of subunits that are more or less independent from one another. Therefore, the loss of several subunits is necessary to lead to function loss, which means that a high dose in a small volume can still be tolerable [32]. The lung is typically a parallel organ. For parallel organs, the tolerable dose is thus exprimed as $Vx \leq Y\%$, meaning that the dose x Gy should not be delivered in more than $Y\%$ of the volume of the OAR [3] [31]. Therefore, the maximum value assigned to such an OAR must ensure this constraint. Typical values for serial or parallel OARs are given in Table 3.2.

OAR	Tolerance dose
Brain stem	$d_{max} = 50$ Gy
Spinal cord	$d_{max} = 45$ Gy
Larynx	$d_{max} = 20$ Gy
Lung	$V20 \leq 35\%$, $V30 \leq 20\%$
Liver	$V30 \leq 50\%$
Heart	$d_{max} = 35$ Gy
Kidney	$d_{max} = 20$ Gy

Table 3.2: Table showing tolerance dose value for a few OAR. Credit : [31]

CTV The dose prescribed in the tumor should ensure that in 90% of the case [31], the tumor will be completely eradicated. The necessary dose depends on the type of tissue in which the tumor is and its volume. The typical doses range between 60 and 80 Gy [31].

[3] This is thus easily ensured by looking at the DVH.

3.3.3 Gamma analysis

A gamma analysis allows to compare two dose distributions [33]. It is based on two other methods: dose difference and distance-to-agreement (DTA).

For a point a_1 in the reference distribution D_{ref} and its corresponding point a_2 in the target distribution D_{tar}, the dose difference is simply $D_{tar}(a_2) - D_{ref}(a_1)$. Each point is tested: the passing criterion checks that the absolute dose difference should not exceed a certain percentage. The drawback of this method is its sensitivity to steep dose gradient, as points in such a region may not pass the test if the dose is even slightly shifted [34].

The DTA method computes the smallest vector r such as $D_{tar}(a_2 + r) = D_{ref}(a_1)$. Each point is tested: the passing criterion checks that the vector r should not exceed a certain radius. While this method is robust to high gradient, it is not in low gradient regions where differences between close points are small but the radius needed to find the matching dose can be very large [34].

Therefore, DTA and dose difference seem to be complementary, which is why the gamma analysis is a combination of both. It defines γ as an index which combines two passing criteria: one accounting for DTA, and one accounting for dose difference. Those terms are scaled so that if $\gamma \leq 1$, the point passes the test [33].

In order to perform such an analysis, we have used a free software developped in the MATLAB environment by Mark Geurts and available on GitHub [35].

Practically, for each voxel position (x, y, z), a Gamma function is computed as a function of a position offset $r = (i, j, k)$ between the target dose D_{tar} and the reference dose D_{ref} [35]

$$\Gamma_{(x,y,z)}(i,j,k) = \left(\frac{D_{tar} - D_{ref}}{d_{max} \cdot \frac{p}{100}}\right)^2 + \left(\frac{i}{s_{dta}}\right)^2 + \left(\frac{j}{s_{dta}}\right)^2 + \left(\frac{k}{s_{dta}}\right)^2 \tag{3.14}$$

where d_{max} is the maximum value in D_{ref}, p is the percentage of error accepted for dose difference and s_{dta} is the distance-to-agreement criterion (i.e. the maximum radius). Then the gamma index for the voxel (x, y, z) is the minimum of the Gamma function for all possible shifts $r = (i, j, k)$

$$\gamma_{(x,y,z)} = \min_{i,j,k} \Gamma_{(x,y,z)}(i,j,k) \tag{3.15}$$

3.4 Image registration

Image registration consists in matching two images together. It will therefore allow to compute the transformation between the original CT scan (used for the original treatment planning) onto the new one. In this work, registration is done through an image processing software called REGGUI (REGistration Graphical User Interface), developped by Guillaume Janssens (IBA) [36]. This section aims to describe the methods behind this software.

3.4.1 General problem of image registration

The objective of image registration is to find the transformation that maps an image, the moving image, onto the other, fixed image. In our case, the moving image corresponds to the original CT and the fixed to the new one. The typical registration workflow is illustrated in Figure 3.2 and can be described as follows.

A digital 3D image is a 3D array in which an intensity value is assigned to each voxel.

In order to find the right transformation, a measure of the difference between the two images is needed. This measure is called the metric and can take various forms depending on the images characteristics. The most commonly used metric is the sum of square difference (SSD), when images are of the same modalities. When it is not the case, the mutual information (MI) is more useful, since it is based on intensity histograms of both images [37]. Those are two examples but other types of metrics can be used, like the local phase which will be introduced later.

The purpose of image registration is thus to find the transformation that minimizes the metric between two images. That is the role of the optimizer. An iterative process allows to reduce the metric at each step until it is close to zero.

Figure 3.2 illustrates one more step. After a geometric transformation, intensity values are shifted towards new positions that may not correspond to the 3D array, therefore the intensity at each voxel position needs to be interpolated [37].

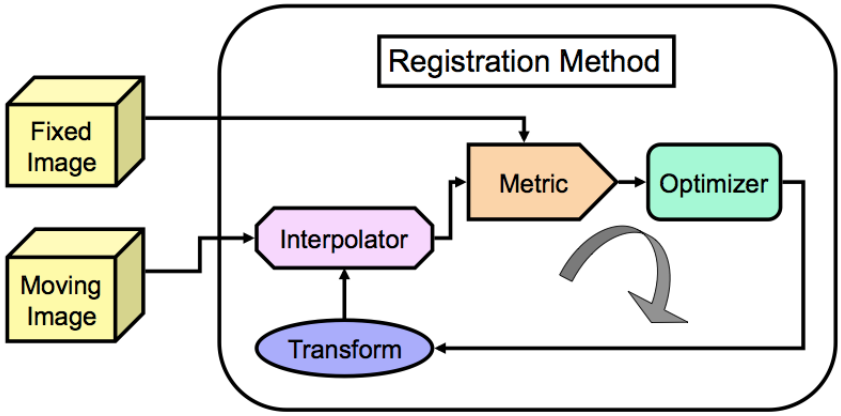


Figure 3.2: Typical registration workflow. The optimizer minimizes the metric and returns a transformation. The transformed image is then interpolated on the discrete voxel grid and the metric is computed between the warped image and the fixed image. The process is iterated until a stopping criterion on the metric is reached. Credit : [37].

Different methods for registration exist which depend on the type of transformation expected. The two main types are rigid and non-rigid registration.

A rigid registration matches two images with a rigid transformation, i.e. a translation, a rotation or a combination of both. It implies that sizes and shapes are preserved. Any rigid transformation can be modelized by some parameters (a translation vector, a rotation matrix) [37].

A non-rigid registration matches two images with any other type of transformation. The focus here is on free form deformation and methods that define a deformation field for every voxel [37]. In opposition with a rigid transformation where a few parameters are needed, the number of parameters in a non rigid registration is proportional to the number of voxels and the number of space dimensions (i.e. a 3D vector per voxel). So many degrees of freedom can lead to non-physical or non-smooth deformation fields, which is why non-rigid registration requires additionnal steps like field regularization. This is also why the concept of diffeomorphism should be introduced. A diffeomorphism is a geometrical deformation that is invertible, differentiable, and its inverse is differentiable. Therefore, a diffeomorphic registration allows to compute an invertible deformation field, which is a necessary condition for it to be physical. Indeed, organs cannot undergo non-invertible spatial transformations, such as mirror effects [38].

In REGGUI, both rigid and non-rigid registration use a multi-scale approach, which allows for more stability [38]. The basic concept consists in repeating the computation at different scales. We define the voxel grid at scale 0 as $\mathcal{G} = \mathcal{G}_0 = \{(x_1, x_2, x_3) \in \mathbb{Z}^3\}$. Then, each scale is defined by a down-sampled version of the initial grid. Therefore, at scale j (with $j = 0, \ldots, J$), the grid is $\mathcal{G}_j$ such as $\mathcal{G}_j = \kappa^j \mathcal{G}$. The approach is illustrated in Figure 3.3. The iteration process starts with the coarser grid $\mathcal{G}_J$. At each scale, the fixed image f and moving image m are down-sampled accordingly, and the displacement field is estimated. The transfer to the next, finer grid, is made by an up-sampling of the deformation field which serves as the initial estimate to start the process at the next scale [38].

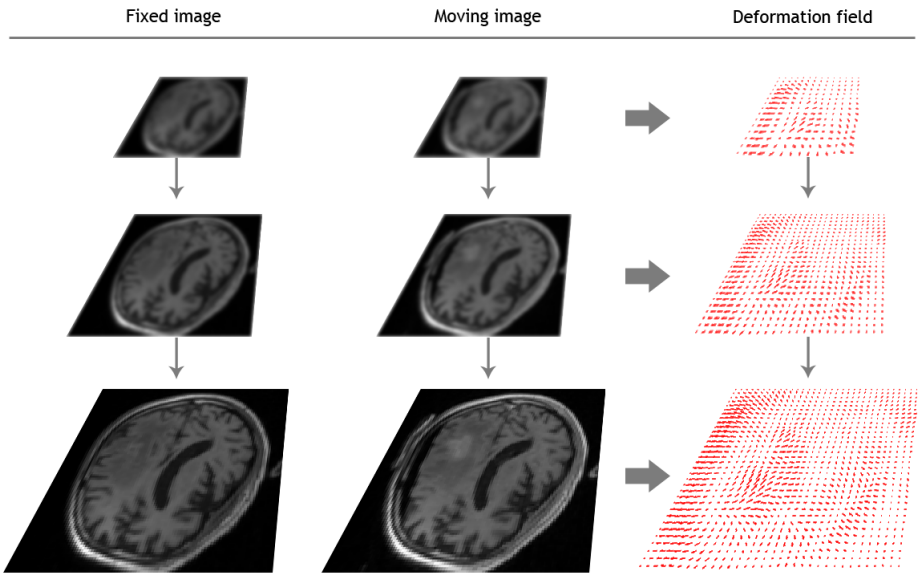


Figure 3.3: Illustration of the multi-resolution approach with a free deformation field computed at each scale. Credit : [36].

3.4.2 Rigid registration

The method presented here is the mono-modal automatic translation registration proposed in REGGUI [36]. Such a registration consists in finding the translation vector. Since both CTs have the same modality, the metric used is SSD. The problem can thus be solved by

$$\min_T \sum_x (f(x) - m(Tx))^2 \tag{3.16}$$

where f and m are respectively the fixed and moving images, represented at each voxel $x = (x_1, x_2, x_3)$ by a scalar function of three variables, and T is the 3-dimensionnal translation vector to be optimized.

Such a problem can be solved by a simple optimization method introduced in the section 3.2: the gradient method [24]. Starting from an initial estimate T_0, it computes at each iteration the direction of the steepest descent which corresponds to the negative of the gradient at that point :

$$T_{k+1} = T_k - hg(T_k) \tag{3.17}$$

where $g(T_k)$ is the gradient of the objective function and h the step size.

3.4.3 Non rigid registration

The method presented here is the diffeomorphic multi-scale non-parametric registration with certainty map proposed in REGGUI [36], based on the method presented by Guillaume Janssens *et al* [38]. It is a registration method that computes a deformation field at each voxel.

The iterative process for computing a free form deformation is more complex than the one presented in Figure 3.2. It is based on three main steps : field computation, field accumulation, and field regularization [38]. Field computation corresponds to the optimizer. It is made by a method called the Morphons [39]. Field accumulation and field regularization are supplementary steps, needed here to reduce the many degrees of freedom (as mentionned previously) by introducing physical constraints on the operations between fields and on the deformation field itself. Those three steps allow for improving the estimation at each iteration until a stopping criterion is reached. The process is illustrated in Figure 3.4.

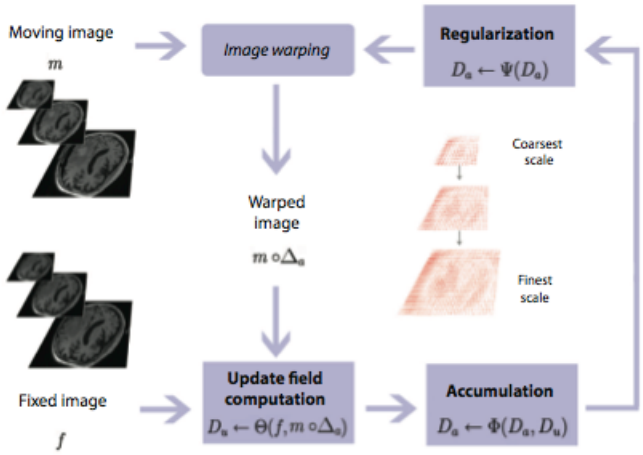


Figure 3.4: Workflow of the process of non-rigid registration described. The three main steps are iterated at each scale until a stopping criterion is reached. Image warping corresponds to transformation and interpolation. Credit : [38].

Field computation - the Morphons method

Let D be a deformation field. It is associated with the deformation operation Δ by the relation $\Delta = D + I$, where I is the identity deformation. Therefore $\Delta(x) = x + D(x)$. The warping of m by the optimized deformation field D should give $f \simeq m \circ \Delta$, where $\circ$ is the common function composition [38].

The updated displacement field D_u is computed as a function Θ of the fixed image f and the warped image, w, i.e. m warped by the accumulated field D_a: [38]

$$D_u = \Theta(f, m \circ \Delta_a) = \Theta(f, w) \tag{3.18}$$

where Δ_a is the operator associated with the field D_a. The field computation is also associated with a certainty map, which is a measure of confidence in the estimation of the update field D_u [38]. As a consequence, the field displacement will be damped at voxels for which the certainty is lower, i.e. in regions that are less reliable. This certainty map also needs to be updated, which split the transformation into two quantities, Θ_D and Θ_c [38].

The Morphons method tries to match transitions rather than intensities [38]. Therefore the metric used is not the SSD but the local phase. Indeed, as local structures are associated with a specific phase, knowing the local phase difference between both images allows to estimate the spatial shift between them. Since the phase is one-dimensional and our image tri-dimensional, we need to compute the phase in different directions. The Morphons method uses a set of quadrature filters h_k associated with an angular vector ω and a direction η_k. A discrete number of directions can be used to cover the whole frequency plane [39]. The local displacement is thus estimated by solving a weighted least squares optimization [38] :

$$\Theta_D(f, w) = \arg\min_{D_u} \sum_k [c_k(\hat{\eta}_k^T D_u - \Delta\phi_k)]^2 \tag{3.19}$$

where $\Delta\phi_k$ is the local dephasing between f and w in the direction η_k, and c_k is the certainty measure associated with the quadrature filter h_k, derived from the magnitude of the phase difference [39].

The certainty map is also updated according to [38]

$$\Theta_c(f, w) = \sum_k c_k(x) \tag{3.20}$$

Field accumulation

Since the update field is computed based on the warped image and not the moving image, it needs to be accumulated after each iteration to the total field displacement D_a :

$$D_a \leftarrow \Phi(D_a, D_u) \tag{3.21}$$

This is done by a specific operation called composition [38]. This means that two displacement fields D_1 and D_2 are accumulated by the compositive operation "$\oplus$" :

$$D_1 \oplus D_2 \triangleq \Delta_1 \circ \Delta_2 - Id = D_2 + D_2 \circ \Delta_1 \tag{3.22}$$

Thus the field D_2, instead of being simply added to D_1, is added to D_1 warped by D_2, which corresponds to the image m being first warped by D_1, then by D_2 [38]. The compositive operation is interesting since it allows to keep properties such as diffeomorphism. [38].

In this case too, the accumulation step is divided into two operations :Φ_d and Φ_c [38]. The field accumulation is done by computing the exponential of the update field which allows smoother fields than the traditional addition or composition :

$$\Phi_D = D_a \oplus \left(\exp\left(\frac{c_u}{c_a + c_u} D_u \right) - Id \right) \tag{3.23}$$

The accumulation of the certainty map is done by :

$$\Phi_c = \frac{c_a^2 + c_u^2}{c_a + c_u} \tag{3.24}$$

Field regularization

Field regularization allows for a smoother transformation and a reduced noise. It is simply performed by applying a Gaussian smoothing on each component of the accumulated field. As the accumulative field D_a is regularized, it corresponds to an elastic solid modeling. If the increment field D_u was regularized, it would be a viscous fluid modeling [38]. Figure 3.5 illustrates the effect of smoothing on the deformation field.

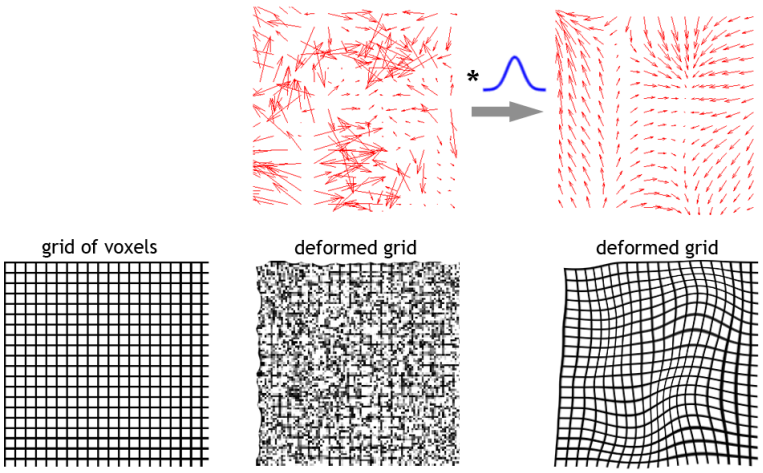


Figure 3.5: The effect of smoothing by convolving the deformation field with a gaussian kernel. Credit : [36].

In order to take into account the certainty map in the regularization, a normalized convolution is done :

$$\Psi_D = D_a *_{c_a} g \tag{3.25}$$

$$\Psi_D = c_a *_{c_a} g \tag{3.26}$$

where g is the gaussian kernel and the operator $*_h$ is such as $s *_h g \triangleq \frac{(hs)*g}{h*g}$ [4].

[4] with * begin the convolution operator.

4 | Methods

This chapter describes the method designed for adapting the treatment. After planning an initial dose distribution based on an initial CT scan, registration methods allows to compute the mapping of the CT used for initial treatment planning onto the CT used for real-time replanning, i.e. acquired just before the treatment session. The deformation field produced is used to compute the reference dose distribution. Optimization then aims at finding the planning parameters that will replicate the reference dose distribution. A quantification of each dose distribution as well as the difference between the adapted dose distribution and the reference dose distribution will finally determine whether the adapted plan has kept the quality and robustness of the initial plan.

Initial treatment planning was done as described in the previous chapter using MIROpt. Registration was performed by the tools offered in REGGUI. Some changes were brought to MIROpt in order to perform adaptive treatment planning.

4.1 Computation of the initial dose

Two phantom cases are considered : a lung case and a paraspinal case, in which the dose prescribed to the target is 70 Gy. The voxel size is set to $2 \times 2 \times 2$ mm for the lung case and $1.074 \times 1.074 \times 2$ mm for the paraspinal case. Sizes of the patient cases are summarized in Table 4.1.

Phantom case	Number of contributing voxels	Spot spacing (mm)	Number of robust spots	Number of nominal spots
Lung	131,946,402	5	2633	901
Paraspinal	380,703,327	5	16,406	6718

Table 4.1: Sizes of the phantom cases.

For the lung case, the phantom is a cube of low radiodensity (-500 HU) representing the lung in which the CTV is defined as a sphere of radius 15 mm (0 HU), as illustrated in Figure 4.1. We have introduced a bone (500 HU) in shape of a parallelipiped and an OAR (40 HU). The disposition of organs in the CT is not completely realistic but what is interesting is that the structures represent obstacles of different density in the beampaths. Two beams are used, located at 0° and 35°.

The paraspinal case is illustrated in Figure 4.2. The phantom is a cylinder in which radiodensity is equal to 0 HU . The spinal cord is represented as a smaller cylinder (OAR) and the target (CTV) is a U-shape structure surrounding it. Three beams are used, located at 0°, 45° and 315°.

In both cases, the physical range error is assumed to be ±3%, and the systematic set up scenarios are going from -5mm to 5mm in the three directions. Random set up errors are not taken into account. For each case, this results in 21 robustness scenarios (7 set ups times 3 density scal-

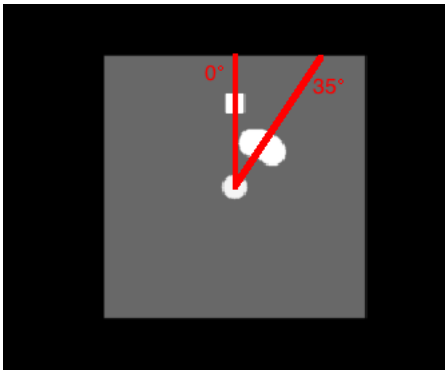


Figure 4.1: An axial slice of the lung case. The beam directions are indicated by the red lines. Black voxels correspond to a radiodensity of -1000 HU (air). Brighter color means higher radiodensity.

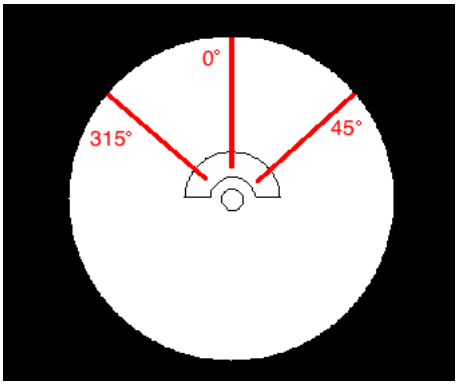


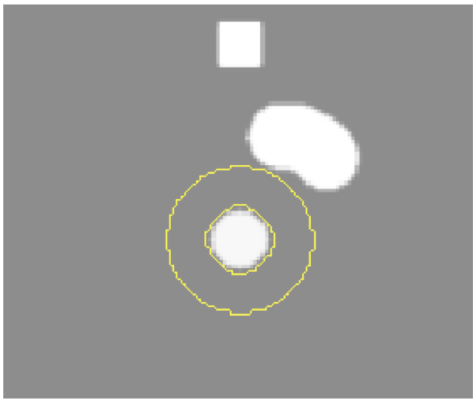
Figure 4.2: An axial slice of the paraspinal case. The beam directions are indicated by the red lines. Black voxels correspond to a radiodensity of -1000 HU (air), white ones to 0HU (water).

ings) defined for robust optimization. Thus the optimization problem has 21 constraints function.

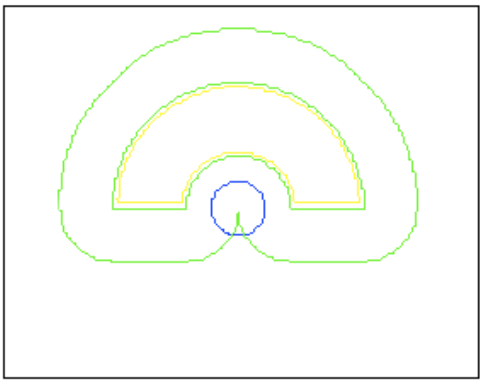
The objective function is made of several optimization functions $f_i(x)$ as explained in section 3.2 (see Equation 3.3). Five are used for the lung cases and four for the paraspinal case. Parameters are presented in Table 4.2. Additional structures were defined in order to limit the dose just outside the target. Indeed, dose restriction in an area surrounding the target can help improve target conformance. Therefore, "CTV+3mm" was defined by adding a margin of 3 millimeters to the CTV. "CTV+2.3cm" was defined by adding a margin of 2.3 centimeters to the CTV. "CTV+2.3cm-CTV+3mm" is a hollow structure made of "CTV+2.3mm" from which we have taken "CTV+3mm". Our choice was guided by previous papers [19]. Those additionnal structures are shown in Figures 4.3a and 4.3b on an axial slice. The choice of importance weight parameters was guided by the assumption that the main purpose is to deliver the prescribed dose to the target, then exact values were found by trial and error. Maximum number of iterations was set to 500 for both cases.

	ROI (i)	Type	Importance weight (impw$_i$)	Reference dose (d_i)
Lung	CTV	max	2	70
	CTV	min	0,5	72
	bone	max	0,1	45
	OAR	max	0,1	45
	CTV2.3cm-CTV+3mm	max	0,3	60
Paraspinal	CTV	max	0,6	70
	CTV	min	0,2	72
	OAR	max	0,3	38
	CTV2.3cm-CTV+3mm	max	0,2	60

Table 4.2: Parameters of the optimization functions for both cases.



(a) Lung case. The yellow contour corresponds to "CTV2.3cm-CTV+3mm". Thus the inner circle is the "CTV+3mm".



(b) Paraspinal case. The green contour corresponds to "CTV2.3cm-CTV+3mm". Thus the inner circle is the "CTV+3mm". The CTV is in yellow and the OAR in blue.

Figure 4.3: Illustration of the additionnal structures used for optimization.

4.2 Registration

The morphological change considered is a translation of the target. For the lung case, the target is translated by 2 cm laterally, as shown in Figure 4.4. For the paraspinal case, the target is moved by 3mm towards the OAR. It is a slight translation but as both structures are already so close, it will be relevant for the optimization. The same additional structures as for the initial plan are defined.

As the target only is translated, the global image deformation must be computed by a non-rigid registration. However, we consider also the rigid translation applied to the target, of which we know the translation vector. Thus we will further refer to two transformations : a translation and a free form deformation.

When performing the non-rigid registration, we had to deal with the problem of using phantoms, that are made of structures of homogeneous density. In consequence, matching two voxels together can lead to completely distorted deformation fields since one voxel can match numerous others, as it is shown in Figure 4.5.

In order to deal with this, the CT images are not used as such but a distance map is com-

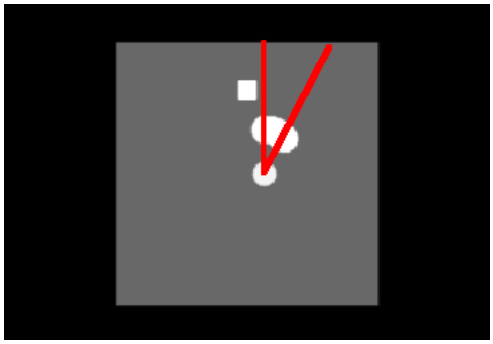


Figure 4.4: An axial slice of the lung case after translation of the target. The beam directions are indicated by the red lines.

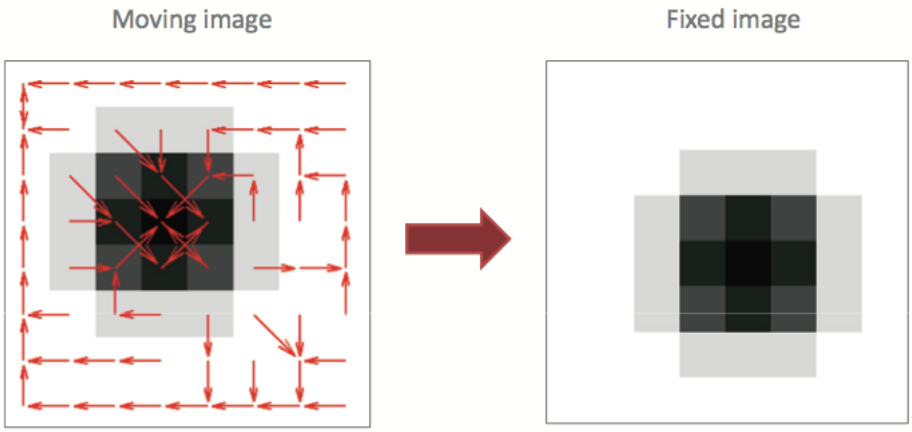


Figure 4.5: An example of deformation field computed on an image with homogenous voxel intensity and corresponding to a perfect matching while it clearly isn't. Credit : [36].

puted. In each structure (ROI), voxels were assigned a new intensity according to their distance from the edges of the structure : the closer to the edge, the closer to zero. It creates isointensity lines inside the structures as shown in Figure 4.6. Distance maps thus allow for the registration method to match isointensity voxels with one another. Thanks to field regularization and diffeomorphism that smooth the deformation field, this is enough to ensure that the computed deformation field is relevant.

Non rigid registration was performed on REGGUI at 8 scales, with 10 iterations per scale, with a weighted compositive field accumulation and a solid field regularization by means of a normalized Gaussian of which variance was set to 1.5 voxels. At the end of the process, the difference between the registered image and the fixed image is measured in terms of SSD for information.

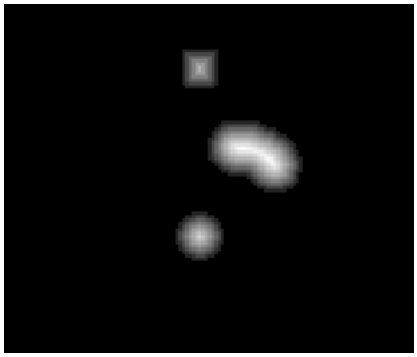


Figure 4.6: An axial view of the distance map for the lung case.

4.3 Computation of the reference dose

A dose distribution can be visualized as a 3D image in which the intensity of each voxel corresponds the absorbed dose (in Gy). Therefore the two transformations can be applied to this image. The reference dose (D_{ref}) is a combination of two images: the rigid reference dose, which is the initial dose on which the translation is applied, and the non-rigid reference dose, which is the initial dose on which the free form deformation is applied. The diagram displayed in Figure 4.7 summarizes the construction of the reference dose.

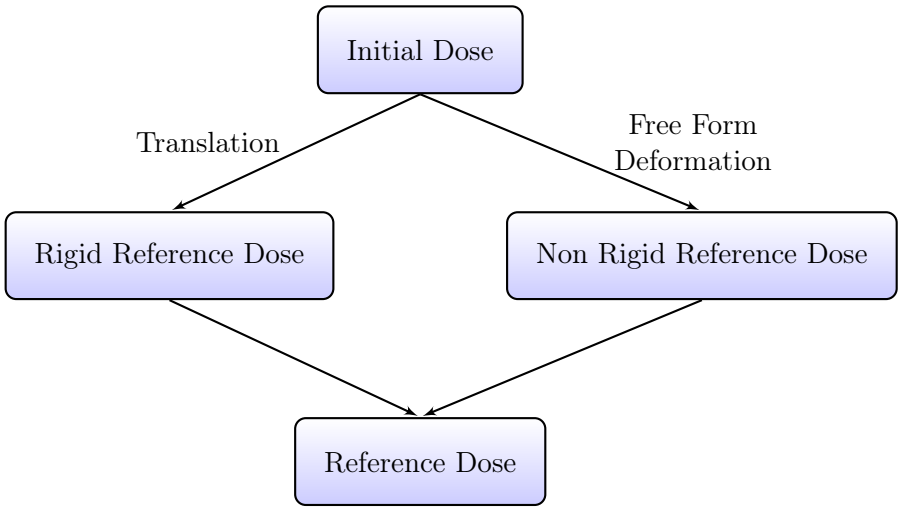
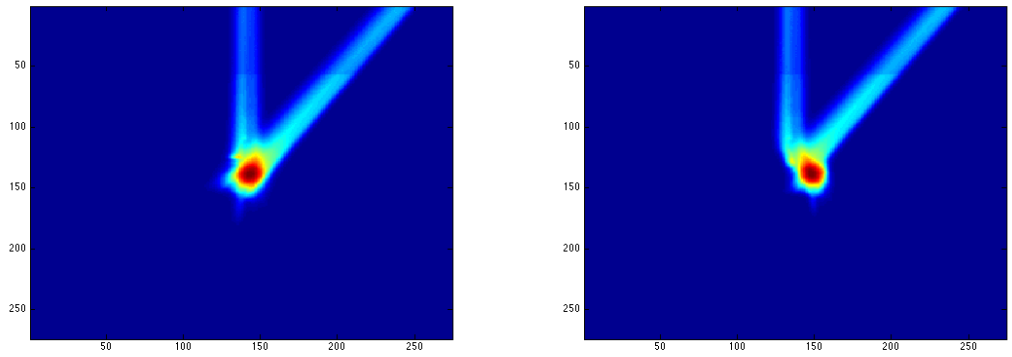


Figure 4.7: Diagram showing the construction of the reference dose.

The reference dose must be as close as possible to the initial dose (D_1). While the non rigid reference dose keeps all properties of the initial dose since the dose distribution in each structure is reproduced, it also yields a dose that is impossible to deliver physically since most of the beampath will be distorted. This is especially true in the lung case, since the shift is lateral, as illustrated in Figure 4.8b. On the other hand, the translation performs a simple shift of the dose distribution, as illustrated in Figure 4.8a, which results in a dose that is possible to deliver but may not keep all properties of the initial distribution. Therefore, a sensible combination of both is needed.

As the method should be as systematic as possible so it does not involve any decision making and does not call for a physician's approval, the combination follows a simple rule : in the target and in serial OARs, the non-rigid reference dose is replicated, while in other structures, the rigid reference dose is replicated.



(a) Rigid reference dose. (b) Non rigid reference dose.

Figure 4.8: Illustration of the deformation of the initial dose distribution for the lung case by applying (a) the translation and (b) the non-rigid deformation.

The reason for this choice comes from the properties of serial and parallel organs. As parallel organs are generally big and can tolerate a certain amount of dose in any part as long as it does not exceed a certain volume, the shape of the dose distribution has little effects on them. Therefore, the rigid reference dose is the best choice since it will reproduce the initial shape of the beam. On the other hand, serial organs, as well as the target, can suffer from a local dose extremum, which is why it is very important to keep the properties of the dose distribution in those regions.

In the lung case, we consider that the bone and OAR are like parallel organs. Thus, the reference dose is built by reproducing the non-rigid reference dose in the shifted CTV and the rigid reference dose in "Ext-CTV+3mm", which is another additional hollow structure made of the whole CT from which we have taken "CTV+3mm" (thus containing the bone and the OAR). In the paraspinal case, the OAR is the spinal cord, therefore it is a serial organ. Thus, the reference dose is built by reproducing the non-rigid reference dose in the shifted CTV, in "CTV+2.3cm-CTV+3mm", and in the OAR. The rigid reference dose was replicated in the remaining parts of "Ext-CTV+3mm".

4.4 Computation of the adapted dose

The adapted treatment is planned by using a modified version of MIROpt. The modifications involve two steps of the planning : the computation of the spot positions and the optimization.

4.4.1 Spot positions

A different way of computing the spot positions allows for a smart initialization of the weights to optimize. This is done based on the work of Geoffrey Schepmans.

A smart initialization of the weights means than the optimization process for the adapted plan starts with an initial estimate x_0 equal to the weights optimized for the initial plan. In order to do this, the spots must be placed in the translated target in the same relative positions as in the initial plan.

As explained in section 2.3, spot positions for the initial plan are defined for each beam direction by transverse coordinates ($x_{BEV,1}$ and $y_{BEV,1}$) and an energy E_1. Thus we need to find

the corresponding set $\{x_{BEV,2}, y_{BEV,2}, E_2\}$ in the new geometry.

The pair $(x_{BEV,2}, y_{BEV,2})$ is not difficult to find. Since BEV coordinates are centered on the beam isocenter, which in our case is the center of target, and the target is only translated, then the new isocenter is computed and we have $(x_{BEV,2}, y_{BEV,2}) = (x_{BEV,1}, y_{BEV,1})$.

The inital energy E_1 of one spot depends on its depth and on $(x_{BEV,1}, y_{BEV,1})$. The first step is thus to compute the depth of the spot in BEV coordinates : $z_{BEV,1}$. Indeed, once this is done, we have again $z_{BEV,2} = z_{BEV,1}$ for the same reason as above. The second step is to retrieve E_2 from $(x_{BEV,2}, y_{BEV,2}, z_{BEV,2})$.

The second step is quite straightforward: it is done by the same ray tracing algorithm used in MIROpt (see Figure 2.9 again). The difference in this case is that the ray tracing is done for each spot instead of for each voxel. Furthermore, the ray tracing was used solely to determine the minimum and maximum energy needed by the beam to cover the target. After that, the energy range was discretized into energy layers. However, in this case, each spot is associated with a particular energy E_2. Therefore we could have a different energy for each spot, which could be practically impossible to deliver. In further work, an optimization of the number of energy layers will thus have to be considered. However, since we focus here on the feasibility of the prescription strategy, this issue is considered as secondary.

The first step is trickier. Basically, a ray tracing is done to find all the points in the beam direction associated with a pair $(x_{BEV,1}, y_{BEV,1})$. Those points are associated with an energy. $z_{BEV,1}$ is defined as the point whose energy is closest to E_1, the energy associated with the pair $(x_{BEV,1}, y_{BEV,1})$. As this step includes approximations, errors may arise. As a consequence, the target coverage can be slighty modified, which may lower the benefits of smart initialization. A comparison between a plan adapted with smart initialization and without is thus performed.

4.4.2 Optimization

The idea is to reproduce exactly the initial dose in the new geometry, which is given by the reference dose. However the weights associated with the reference dose are unknown and must be computed. Therefore, a new optimization that will replicate the reference dose in each ROI is needed and will be defined by a new objective function. The latter is still a weighted combination of optimization functions $f_i(x)$:

$$f(x) = \sum_{i=1}^{m} \text{impw}_i f_i(x) \tag{4.1}$$

with the importance weights being the same as for the initial treatment plan. This ensures the replicability of the method.

In the target, the reference dose must be exactly reproduced. The optimization function $f_i(x)$ associated with the target must thus compute the normalized square dose deviation from the reference dose :

$$f_i(x) = \sum_{j=1}^{\text{nvoxels}_i} (D_{ref}(j) - D(j))^2) \tag{4.2}$$

where $D_{ref}(j)$ is our reference dose computed in the jth voxel of the ith ROI and $D(j)$ is the dose computed in the jth voxels of the ith ROI such as $D(j) = P_j x$, where P_j is the jth line of P, the beamlets matrix.

In other structures, we could allow the adapted dose to be smaller than the reference dose. Therefore, we have a different optimization function for those ROI :

$$f_i(x) = \sum_{j=1}^{\text{nvoxels}_i} (\max(0, (D(j) - D_{ref}(j))^2) \tag{4.3}$$

4.4.3 Parameters

The adapted dose is planned in a non-robust manner. This allows for a reduced computation time and is justified by the fact that our reference dose should already be robust against uncertainties. Therefore, only one dose contribution matrix is computed and the spots used are the nominal spots only. The number of beams and their angle is of course the same as for the initial plan. Maximum number of iterations was set to 250 iterations (instead of 500) for both cases.

4.5 Quantitative evaluation

The three doses (initial, reference and adapted) are first individually quantified in each ROI i for the nominal case, by computing :

- The maximum dose :
$$d_{max} = \max_j D(j),\ j = 1, ...\text{nvoxels}_i \tag{4.4}$$
- The mean :
$$\bar{d} = \frac{1}{\text{nvoxels}_i} \sum_{j=1}^{\text{nvoxels}_i} D(j) \tag{4.5}$$
- The quantiles d_{95}, d_{50} and d_5 from the DVH.

The initial plan undergoes a robustness test that will determine how robust it is against uncertainties and assess the quality of the robust optimization. While the adapted plan is not optimized in a robust way, it also undergoes a robustness test in order to assess how well it has kept the properties of the initial plan. This test consist in computing the DVH in 81 scenarios, thus 60 more scenarios than the ones defined for robust optimization. This is because the test takes also into account diagonal shifts, thus leading to 27 set ups times 3 density scalings. The detail of each scenario can be found in the appendix A.2.

From those DVHs, more characteristic values are determined for the initial and adapted dose. We define $\hat{d}_x$ in a ROI as the dose delivered to $x\%$ of its volume in the worst case scenario, i.e. $\hat{d}_{95}$ is the minimum d_{95} among all scenarios. In our evaluation, $\hat{d}_{95}$ and $\hat{d}_5$ are computed in the CTV, and for the paraspinal case $\hat{d}_{max}$ is computed in the OAR, according to the standards stated in section 3.3.2.

The difference between the reference dose D_{ref} and the adapted dose D_2 in the nominal case is computed voxel by voxel in each ROI i in three ways :

- The root mean square error (RMSE) :
$$d_{\text{RMSE}} = \sqrt{\frac{1}{\text{nvoxels}_i} \sum_{j=1}^{\text{nvoxels}_i} (D_{ref}(j) - D_2(j))^2} \tag{4.6}$$

- The mean absolute difference (AD) :

$$d_{\mathrm{AD}} = \frac{1}{\mathrm{nvoxels}_i} \sum_{j=1}^{\mathrm{nvoxels}_i} |D_{ref}(j) - D_2(j)| \tag{4.7}$$

- The gamma analysis, as described in the previous chapter (see Equation 3.14), where D_{tar} is replaced by D_2. The percentage of voxels that pass the test in the ith ROI is computed. The parameters used were $s_{dta} = 0.5$ mm for DTA and $p = 3\%$ for dose difference.

5 | Results

In this chapter, the quantitative evaluation of the initial and adapted plans for the lung case and the paraspinal case is presented. It is done as described in the previous section. Moreover, we propose a visualization of the dose by plotting isodose curves and line doses. An isodose curve is a line passing through voxels of equal dose, while the line dose shows the evolution of the dose on a particular line of voxels.

All simulations were performed in the MATLAB environment, on a server using 16 Intel(R) Xeon(R) (3.33GHz) processors with 23GB of memory under 64-bit Linux.

5.1 Lung case

5.1.1 Initial dose

Dose contribution matrices computation with MCsquare lasted 36 hours and robust optimization with IPOPT took 6 hours. It stopped after 500 iterations with an optimality error given by 2.959 $\times 10^{-7}$ (see section 3.2 for the definition of the optimality error).

Isodose curves for the nominal scenario and a perturbed scenario in an axial slice are shown on Figure 5.1a. Line doses from the same slice are shown on Figures 5.1b. The nominal 95% isodose curve encloses most of the CTV but leaves a gap in the middle of it. Extra dose is delivered posteriorly and to the right of the CTV. The perturbed 95% isodose curve is similar but also does not enclose a left anterior part of the CTV.

DVHs are displayed on Figure 5.2. Mean and quantiles of the dose are given in Table 5.1. Values in bold exceed the standards stated in section 3.3.2. Indeed, $\hat{d}_{95}$ is lower than 66.5 Gy, and d_5 and $\hat{d}_5$ are greater than 73.5 Gy.

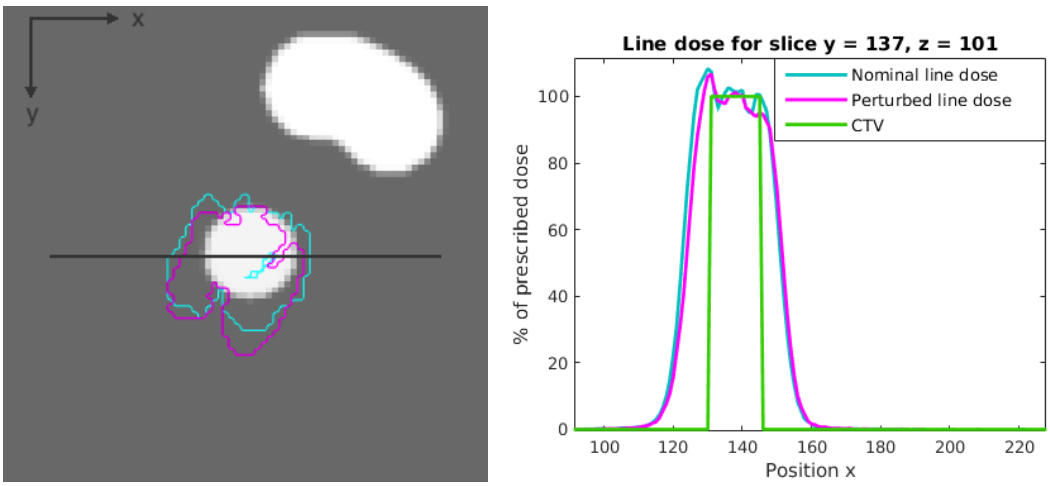
	$\bar{d}$	d_{95}	d_{50}	d_5	d_{max}	$\hat{d}_{95}$	$\hat{d}_5$
CTV	69.94	66.64	69.9	**73.65**	80.37	**62.49**	**76.01**
bone	19.34	9.865	19.63	27.31	32.7	/	29.18
OAR	12.59	0.1153	4.96	35.85	55.17	/	39.11

Table 5.1: Initial dose statistics for the lung case. $\bar{d}$ denotes the mean dose, d_x denotes minimum dose to x % of the ROI volume in the nominal case, d_{max} denotes the maximum dose in the nominal case, and $\hat{d}_x$ denotes dose to x % of the ROI volume in the worst scenario (minimum for $x = 95$ and maximum otherwise).

5.1.2 Reference dose

Non rigid registration with REGGUI ended with a remaining SSD of 6.356 $\times 10^{-4}$. It lasted 16 minutes.

DVHs (only for the nominal case) are displayed on Figure 5.3 and show a behavior similar to the initial dose distribution in the nominal case.



(a) (b)

Figure 5.1: Isodose curves and line doses of the initial dose distribution in the nominal case and a perturbed scenario in a transversal slice of the lung case. (a) The isodose curves show 95% of the prescribed dose : the blue curve shows the nominal case and the pink one a perturbed scenario. In the perturbed scenario the density is 3% lower than planned and the isocenter is shifted 2.89 mm diagonally in the directions anterior, inferior and right. The spherical CTV and the OAR are visible. The grey line displays the trace of the line doses. (b) Line doses : the height of the CTV is the prescribed dose level. x,y and z refers to DICOM coordinates.

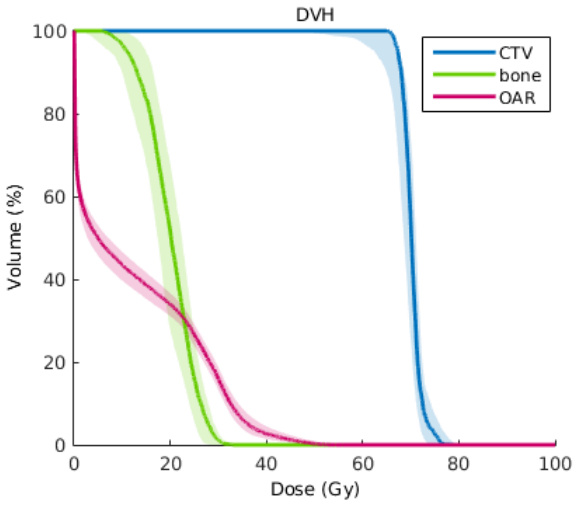


Figure 5.2: DVHs of the initial dose distribution for the lung case in the nominal case (darker line) and in 81 scenarios (27 set ups times 3 density scalings).

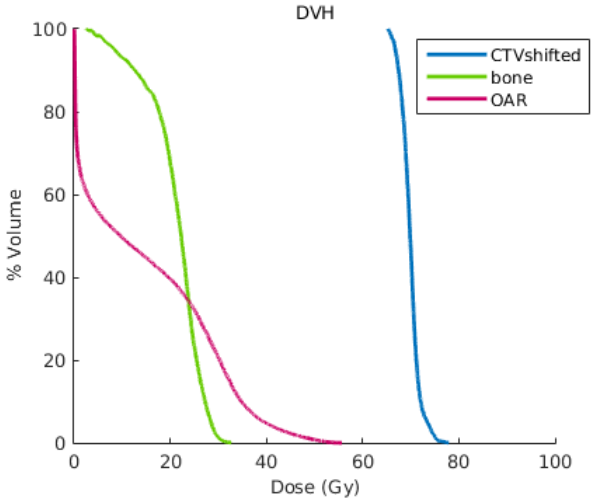


Figure 5.3: DVHs of the reference dose distribution for the lung case in the nominal scenario.

Mean and quantiles of the dose are given in Table 5.2. Note that values do not exceed the standards stated in section 3.3.2. The difference between those values and the one of the initial dose distribution does not exceed 0.01% of the nominal initial dose distribution value for the CTV. For the bone and the OAR the difference is bigger but the percentage is not higher than 1%.

	$\bar{d}$	d_{95}	d_{50}	d_5	d_{max}
CTV shifted	69.88	66.8	69.86	73.42	77.93
bone	21.11	8.475	22.29	28.5	32.7
OAR	14.76	0.1476	9.66	39.49	55.71

Table 5.2: Reference dose statistics for the lung case. $\bar{d}$ denotes the mean dose, d_x denotes minimum dose to x % of the ROI volume in the nominal case, and d_{max} denotes the maximum dose.

5.1.3 Adapted dose

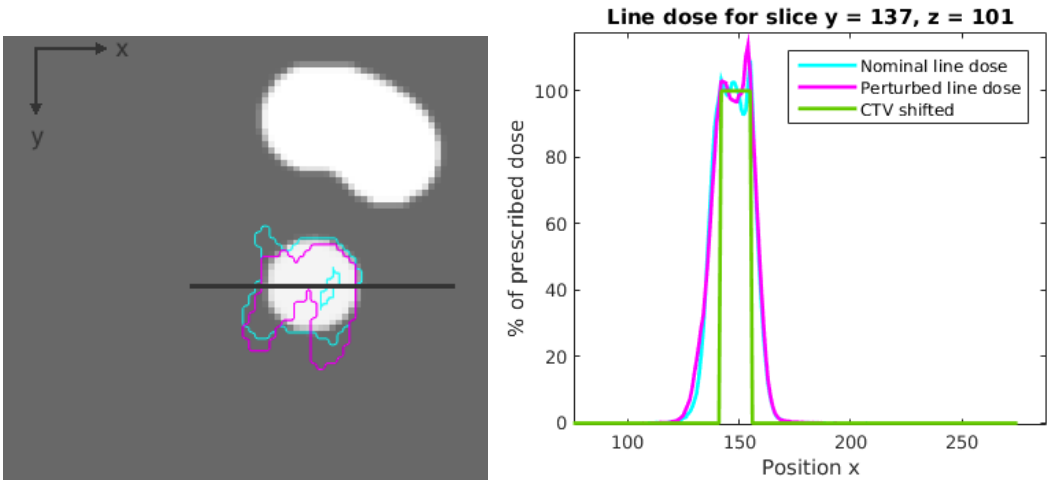
Without smart initialization

Dose contribution matrix computation with MCsquare lasted 2 hours, and optimization with IPOPT took 16 minutes. It stopped after 250 iterations with an optimality error of 5.935×10^{-4}.

Isodose curves for the nominal scenario and a perturbed scenario in an axial slice are shown on Figure 5.4a. Line doses from the same slice are shown on Figures 5.4b. The nominal 95% isodose curve encloses most of the CTV except for a gap in the middle of it, just like in the initial case. Also, extra dose is delivered to the right. The perturbed 95% isodose curve does not enclose some parts of the CTV, and extra dose is delivered posteriorly.

DVHs are displayed on Figure 5.5. The adapted plan is less robust than the initial plan. In the worst case scenario, the dose to the CTV declines earlier than in the initial plan. Mean and quantiles of the dose are given in Table 5.3. Values in bold exceed the standards stated in section 3.3.2. Indeed, $\hat{d}_{95}$ is way lower than 66.5 Gy, and $\hat{d}_5$ is greater than 73.5 Gy.

Differences between the reference dose and the adapted dose are given by the quantities computed in Table 5.4. The adapted dose distribution in the CTV is very close to the reference dose distribution, with a percentage of success in the gamma test of almost 99%. However, the difference is larger in the OAR, and even larger in the bone.



(a) (b)

Figure 5.4: Isodose curves and line doses for the adapted dose distribution without smart initialization in the nominal case and a perturbed scenario in a transversal slice of the lung case. (a) The isodose curves show 95% of the prescribed dose: the blue curve shows the nominal case and the pink one a perturbed scenario. In the perturbed scenario the density is 3% lower than planned and the isocenter is shifted 2.89 mm diagonally in the directions anterior, inferior and right. The spherical CTV and the OAR are visible. The grey line displays the trace of the line doses. (b) Line doses : the height of the CTV is the prescribed dose level. x,y and z refers to DICOM coordinates.

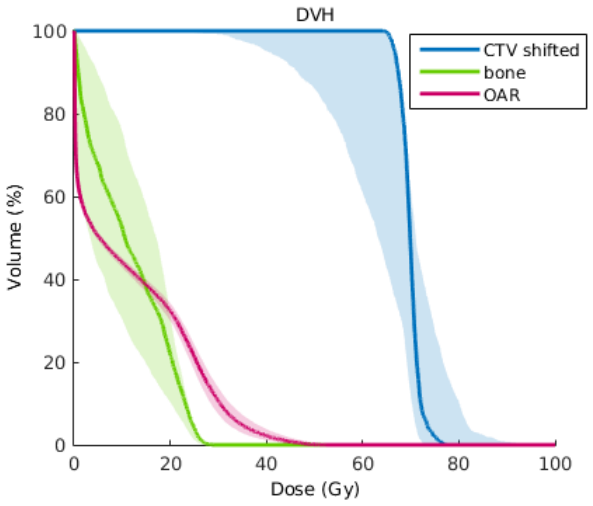


Figure 5.5: DVHs of the adapted dose without smart initialization in the nominal case (darker line) and in 81 scenarios (27 set ups times 3 density scalings).

	$\bar{d}$	d_{95}	d_{50}	d_5	d_{max}	$\hat{d}_{95}$	$\hat{d}_5$
CTV shifted	69.92	66.75	69.86	73.46	77.71	**40.32**	**81.72**
bone	11.44	0.4246	10.57	24.63	28.89	/	24.98
OAR	11.89	0.1097	5.39	34.7	52.65	/	37.88

Table 5.3: Adapted dose statistics for the lung case without smart initialization. d_x denotes minimum dose to x % of the ROI volume in the nominal case, $\hat{d}_x$ denotes dose to x % of the ROI volume in the worst scenario (minimum for $x = 95$ and maximum otherwise), $\bar{d}$ denotes the mean dose, and d_{max} denotes the maximum dose.

	d_{RMSE}	d_{AD}	$\gamma_{\%}$
CTV shifted	1.171	0.8887	98.82
bone	11.79	9.819	21.83
OAR	5.217	3.195	61.53

Table 5.4: Statistics between reference dose and adapted dose without smart initialization for the lung case. d_{AD} denotes the mean absolute dose difference, d_{RMSE} denotes the root mean square dose deviation, and $\gamma_{\%}$ denotes the percentage of voxels that has passed the gamma test.

With smart initialization

Dose contribution matrix computation with MCsquare lasted 2 hours, and optimization with IPOPT took 15 minutes. It stopped after 250 iterations with an optimality error of 6.053×10^{-5}.

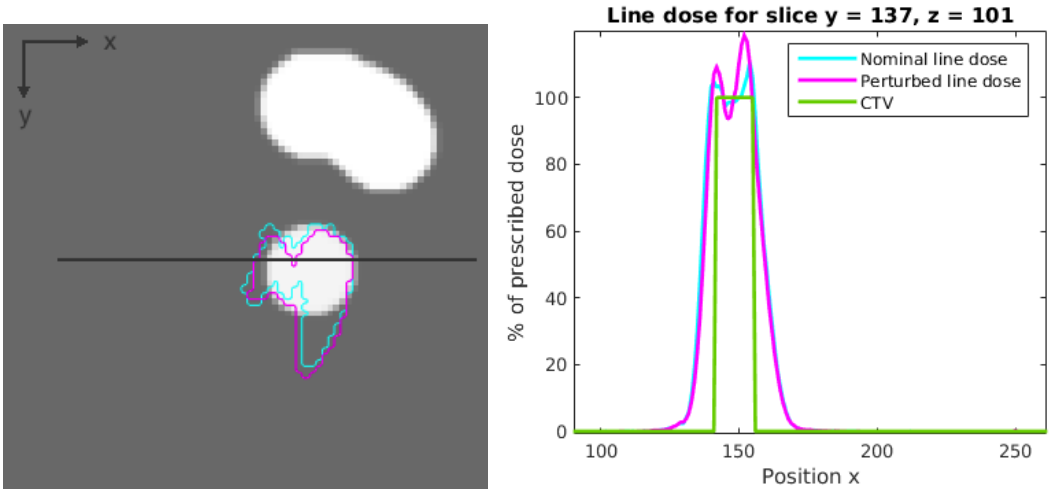
Isodose curves for the nominal scenario and a perturbed scenario in an axial slice are shown on Figure 5.6a. Line doses from the same slice are shown on Figure 5.6b. The nominal 95% isodose curve does not enclose some parts of the CTV and extra dose is delivered posteriorly. The perturbed 95% isodose curve is very similar.

DVHs are displayed on Figure 5.7. The adapted plan is less robust than the initial plan. In the worst case scenario, the dose to the CTV declines earlier than in the initial plan. Mean and quantiles of the dose are given in Table 5.5. Values in bold exceed the standards stated in section 3.3.2. Indeed, d_{95} and $\hat{d}_{95}$ are lower than 66.5 Gy, and d_5 and $\hat{d}_5$ are greater than 73.5 Gy.

Differences between the reference dose and the adapted dose aer given by the quantities computed in Table 5.6. The dose distribution in the CTV is not reproduced as well as without smart initialization, as the percentage of the success of the gamma test only reaches 70%. However values in the OAR and the bone are similar as without smart initialization.

	$\bar{d}$	d_{95}	d_{50}	d_5	d_{max}	$\hat{d}_{95}$	$\hat{d}_5$
CTV shifted	69.49	**65.09**	69.19	**75.55**	81.97	**46.2**	**85.2**
bone	10.53	0.2913	8.755	25.4	31.37	/	25.7
OAR	11.43	0.1096	4.921	33.57	53.42	/	36.68

Table 5.5: Adapted dose statistics for the lung case with smart initialization. d_x denotes minimum dose to x % of the ROI volume in the nominal case, $\hat{d}_x$ denotes dose to x % of the ROI volume in the worst scenario (minimum for $x = 95$ and maximum otherwise), $\bar{d}$ denotes the mean dose, and d_{max} denotes the maximum dose.



(a) (b)

Figure 5.6: Isodose curves and line doses for the adapted dose distribution with smart initialization in the nominal case and a perturbed scenario in a transversal slice of the lung case. (a) The isodose curves show 95% of the prescribed dose: the blue curve shows the nominal case and the pink one a perturbed scenario. In the perturbed scenario the density is 3% lower than planned and the isocenter is shifted 2.89 mm diagonally in the directions anterior, inferior and right. The spherical CTV and the OAR are visible. The grey line displays the trace of the line doses. (b) Line doses : the height of the CTV is the prescribed dose level. x,y and z refers to DICOM coordinates.

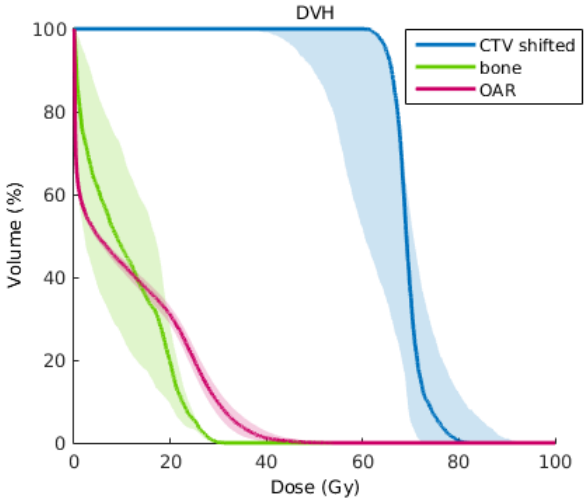


Figure 5.7: DVHs of the adapted dose with smart initialization for the lung case in the nominal case (darker line) and 81 scenarios (27 set ups times 3 density scalings).

	d_{RMSE}	d_{AD}	$\gamma_{\%}$
CTV shifted	3.256	2.469	70.85
bone	12.84	10.7	24.29
OAR	5.822	3.68	59.2

Table 5.6: Statistics between reference dose and adapted dose with smart initialization for the lung case. d_{AD} denotes the mean absolute dose difference, d_{RMSE} denotes the root mean square dose deviation, and $\gamma_{\%}$ denotes the percentage of voxels that has passed the gamma test.

5.2 Paraspinal case

5.2.1 Initial dose

For this case, the initial plan was computed on a faster machine than the other computations, using 32 Intel(R) Xeon(R) (2.60GHz) processors with 32 GB of memory. Dose contribution matrices computation with MCsquare lasted a few hours, and robust optimization with IPOPT took 20 minutes. It stopped after 500 iterations with an optimality error given by 8.367 $\times 10^{-5}$.

Isodose curves for the dose distribution in the nominal scenario and a perturbed scenario in an axial slice are shown on Figure 5.8a. Line doses from the same slice are shown on Figures 5.8b and 5.8c. The nominal 95% isodose curve encloses all parts of the CTV. The perturbed 95% isodose curve encloses most of the CTV but is shifted in the directions anterior and right compared to the nominal isodose curve.

DVHs are displayed on Figure 5.9. The plan is less robust around the OAR than around the target. Mean and quantiles of the dose are given in Table 5.7. Values in bold exceed the standards stated in section 3.3.2. Indeed, $\hat{d}_{95}$ is lower than 66.5 Gy, $\hat{d}_5$ is greater than 73.5 Gy and $\hat{d}_{max}$ in the OAR is greater than 45 Gy.

	$\bar{d}$	d_{95}	d_{50}	d_5	d_{max}	$\hat{d}_{95}$	$\hat{d}_5$	$\hat{d}_{max}$
CTV	70.89	68.88	70.87	72.96	75.49	**64.79**	**74.09**	77.5
OAR	13.89	2.716	14.14	24.82	27.79	/	40.55	**52.7**

Table 5.7: Initial dose distribution statistics for the paraspinal case. $\bar{d}$ denotes the mean dose, d_x denotes the dose delivered to x % of the ROI volume in the nominal case, d_{max} denotes the maximum dose in the nominal case, $\hat{d}_x$ denotes the dose to x % of the ROI volume in the worst case scenario, and $\hat{d}_{max}$ denotes the maximum dose in the worst case scenario.

5.2.2 Reference dose

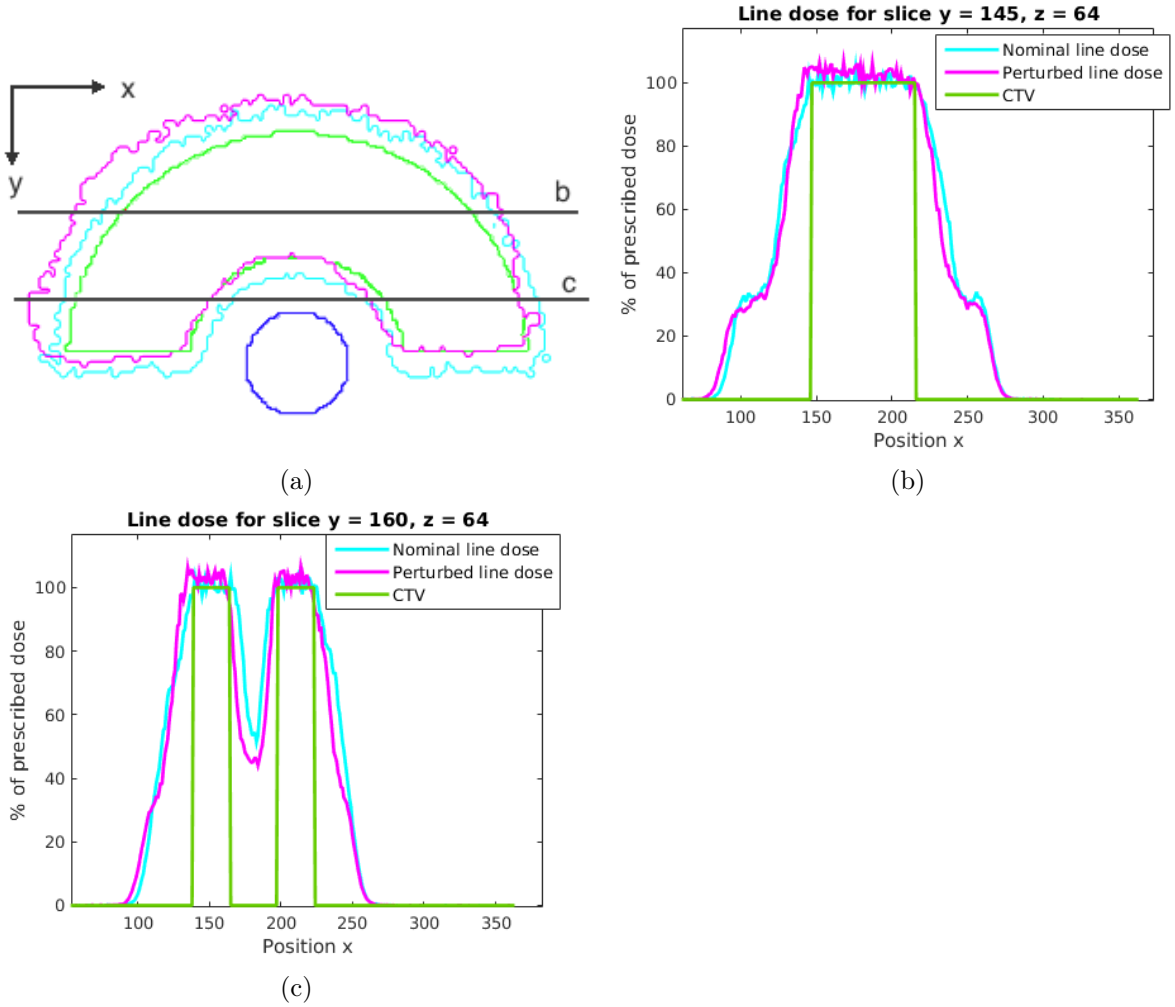
Non rigid registration ended with a remaining SSD of 2.6327 $\times 10^{-4}$. It lasted 20 minutes.

DVHs are displayed on Figure 5.10 and display a behavior similar to the initial dose distribution in the nominal case.

Mean and quantiles of the dose in each ROI are given in Table 5.8. The difference between those values and the one of the initial dose distribution does not exceed 0.01% of the initial nominal distribution value.

	$\bar{d}$	d_{95}	d_{50}	d_5	d_{max}
CTV shifted	70.88	69.29	70.86	72.52	74.97
OAR	14.16	2.803	14.51	25.06	27.97

Table 5.8: Reference dose statistics for the paraspinal case in the nominal case. $\bar{d}$ denotes the mean dose, d_x denotes minimum dose to x % of the ROI volume, and d_{max} denotes the maximum dose.



(a) (b) (c)

Figure 5.8: Isodose curves and line doses of the initial dose distribution in the nominal case and a perturbed scenario in a transversal slice of the paraspinal case. (a) The isodose curves show 95% of the prescribed dose: the light blue curve shows the nominal case and the pink one a perturbed scenario. In the perturbed scenario the density is 3% higher than planned and the isocenter is shifted is shifted 2.89 mm diagonally in the directions posterior, superior and left. The green contour delineates the CTV and the blue contour delineates the OAR. The grey line displays the trace of the line doses. (b) and (c) Line doses : the height of the CTV is the prescribed dose level. x,y and z refers to DICOM coordinates.

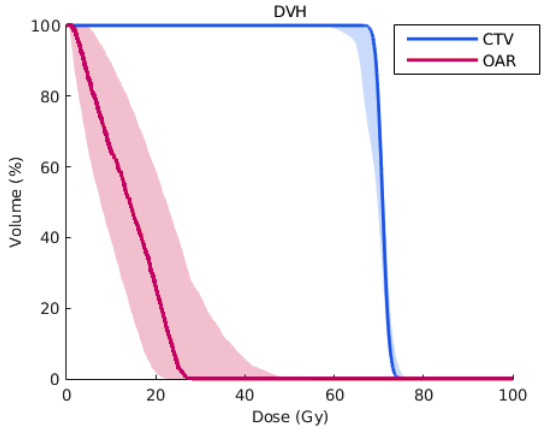


Figure 5.9: DVHs of the initial dose distribution for the paraspinal case in the nominal scenario (darker line) and in 81 scenarios (27 set ups times 3 density scalings).

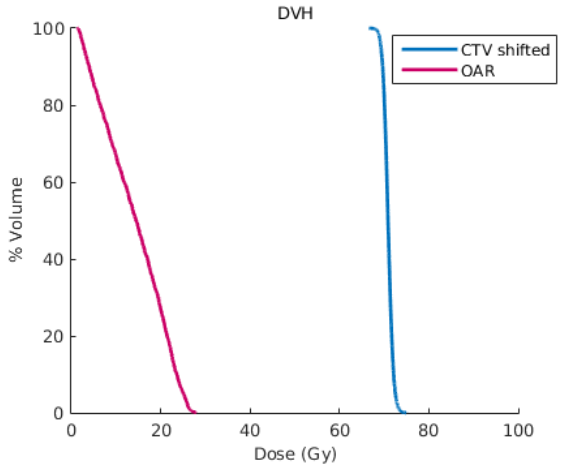


Figure 5.10: DVHs of the reference dose distribution for the paraspinal case in the nominal scenario.

5.2.3 Adapted dose

Without smart initialization

Dose contribution matrix computation with MCSquare lasted 3 hours, and optimization with IPOPT took 14 minutes. It stopped after 250 iterations with an optimality error given by 6.28×10^{-2}.

Isodose curves for the dose distribution in the nominal case and a perturbed scenario in an axial slice are shown on Figure 5.11a. Line doses from the same slice are shown on Figures 5.11b and 5.11c. The nominal 95% isodose curve encloses all parts of the CTV. The perturbed 95% isodose curve leaves gaps in the CTV and does not conform well to the more posterior part. The dose delivered in between the branches of the CTV, closer to the OAR, is lower than in the initial plan.

DVHs are displayed on Figure 5.12. The adapted plan is less robust than the initial plan. In the worst case scenario, the dose to the CTV declines earlier than in the initial plan and the OAR receives higher maximum dose.

Mean and quantiles of the dose in each ROI are given in Table 5.9. Values in bold exceed the standards stated in section 3.3.2. Indeed, $\hat{d}_{95}$ is lower than 66.5 Gy, $\hat{d}_5$ is greater than 73.5 Gy and $\hat{d}_{max}$ in the OAR is greater than 45 Gy.

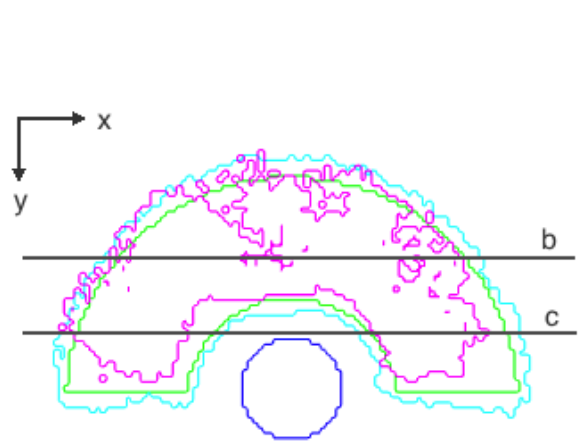
Differences between the reference dose and the adapted dose are given by the quantities computed in Table 5.10. The adapted dose distribution is much closer to that of the initial plan in the CTV than in the OAR. Especially, the percentage of success of the gamma analysis in the OAR is very low.

	$\bar{d}$	d_{95}	d_{50}	d_5	d_{max}	$\hat{d}_{95}$	$\hat{d}_5$	$\hat{d}_{max}$
CTV shifted	71.04	69.04	71.02	73.03	75.19	**57.4**	**77.7**	80.5
OAR	8.311	0.581	6.339	21.36	26.86	/	**46.5**	**59.3**

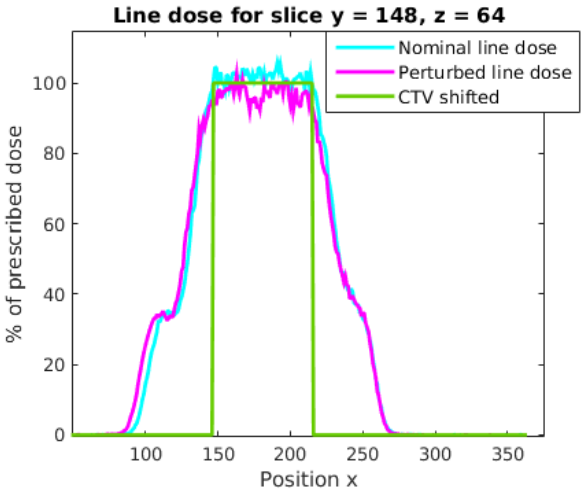
Table 5.9: Adapted dose statistics for the paraspinal case without smart initialization. $\bar{d}$ denotes the mean dose, d_x denotes minimum dose to x % of the ROI volume in the nominal case, d_{max} denotes the maximum dose in the nominal case, $\hat{d}_x$ denotes dose to x % of the ROI volume in the worst scenario (minimum for $x = 95$ and maximum otherwise), and $\hat{d}_{max}$ denotes the maximum dose in the worst case scenario.

	d_{RMSE}	d_{AD}	$\gamma_{\%}$
CTV shifted	1.434	1.148	93.6
OAR	6.291	5.849	9.625

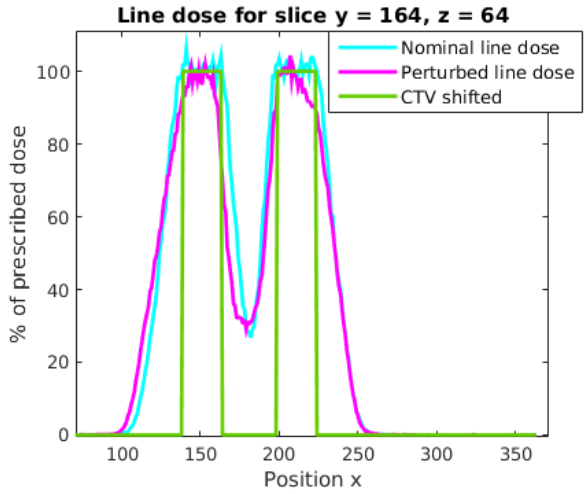
Table 5.10: Statistics between reference dose and adapted dose without smart initialization for the paraspinal case. d_{AD} denotes the mean absolute dose difference, d_{RMSE} denotes the root mean square dose deviation, $\gamma_{\%}$ denotes the percentage of voxels that has passed the gamma test.



(a)



(b)



(c)

Figure 5.11: Isodose curves and line doses of the adapted dose distribution without smart initialization in the nominal case and a perturbed scenario in a transversal slice of the paraspinal case. (a) The isodose curves show 95% of the prescribed dose : the light blue curve shows the nominal case and the pink one a perturbed scenario. In the perturbed scenario the density is 3% higher than planned and the isocenter is shifted is shifted 2.89 mm diagonally in the directions posterior, superior and left. The green contour delineates the CTV and the blue contour delineates the OAR. The grey line displays the trace of the line doses. (b) and (c) Line doses : the height of the CTV is the prescribed dose level. x,y and z refers to DICOM coordinates.

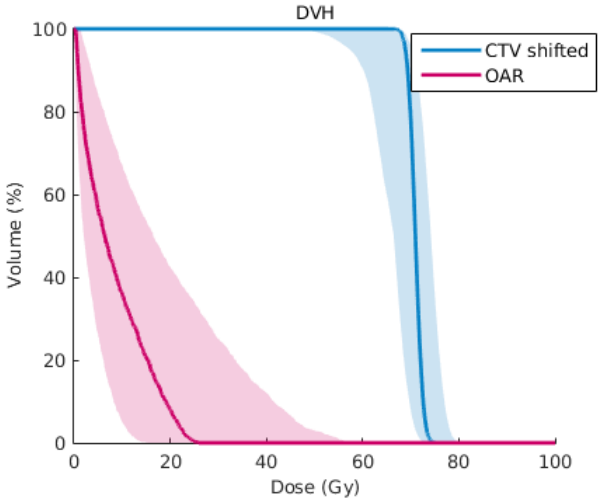


Figure 5.12: DVHs of the adapted dose without smart initialization for the paraspinal case in the nominal scenario (darker line) and in 81 scenarios (27 set ups times 3 density scalings).

With smart initialization

Dose contribution matrix computation with MCSquare lasted 3 hours, and optimization with IPOPT took 12 minutes. It stopped after 250 iterations with an optimality error given by 7.675×10^{-3}.

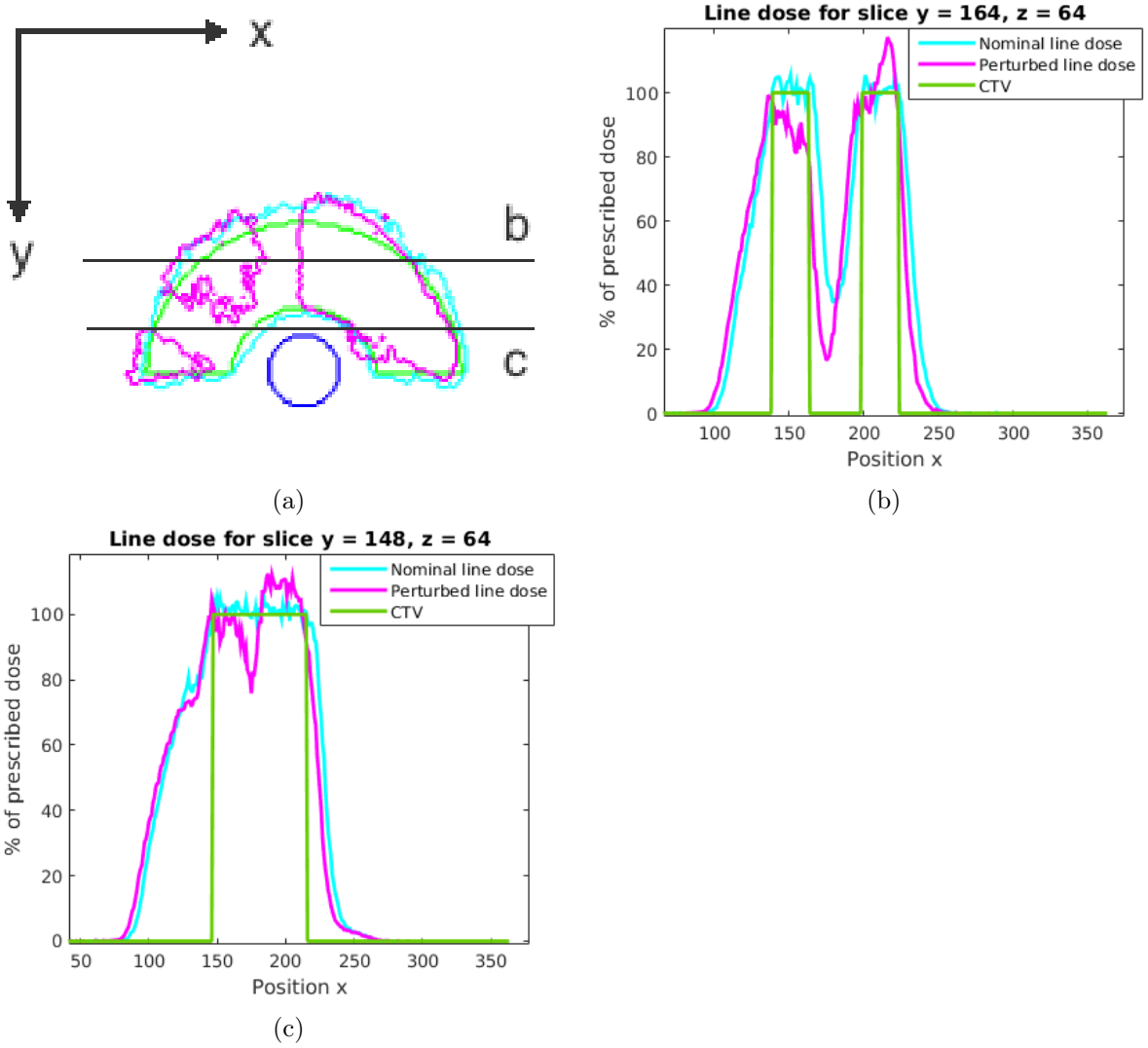
Isodose curves for the dose distribution in the nominal scenario and a perturbed scenario in an axial slice are shown on Figure 5.13a. Line dose from the same slice is shown on Figures 5.13b and 5.13c. The nominal 95% isodose curve encloses all the CTV. The perturbed 95% isodose curve does not enclose large parts of the CTV.

DVHs are displayed on Figure 5.14. The adapted plan is less robust than the initial plan. In the worst case scenario, the dose to the CTV declines earlier than in the initial plan and the OAR receives higher maximum dose. Mean and quantiles of the dose in each ROI are given in Table 5.11. Values in bold exceed the standards stated in section 3.3.2. Indeed, $\hat{d}_{95}$ is lower than 66.5 Gy, $\hat{d}_5$ is greater than 73.5 Gy and $\hat{d}_{max}$ in the OAR is greater than 45 Gy.

Differences between the reference dose and the adapted dose are given by the quantities computed in Table 5.6. The adapted dose distribution in the CTV is close to the reference dose distribution and compared to the plan without smart initialization, the dose delivered to the OAR is also much more like the reference dose distribution.

	$\bar{d}$	d_{95}	d_{50}	d_5	d_{max}	$\hat{d}_{95}$	$\hat{d}_5$	$\hat{d}_{max}$
CTV shifted	70.82	68.4	70.83	73.21	76.55	**52.26**	**83.85**	89.55
OAR	12.37	2.913	12.25	23.36	27.23	/	**56.48**	**67.05**

Table 5.11: Adapted dose statistics for the paraspinal case with smart initialization. $\bar{d}$ denotes the mean dose, d_x denotes minimum dose to x % of the ROI volume in the nominal case, d_{max} denotes the maximum dose in the nominal case, $\hat{d}_x$ denotes dose to x % of the ROI volume in the worst scenario (minimum for $x = 95$ and maximum otherwise), and $\hat{d}_{max}$ denotes the maximum dose in the worst case scenario.



(a) (b) (c)

Figure 5.13: Isodose curves and line doses of the adapted dose distribution with smart initialization in the nominal case and a perturbed scenario in a transversal slice of the paraspinal case. (a) The isodose curves show 95% of the prescribed dose : the light blue curve shows the nominal case and the pink one a perturbed scenario. In the perturbed scenario the density is 3% higher than planned and the isocenter is shifted is shifted 2.89 mm diagonally in the directions posterior, superior and left. The green contour delineates the CTV and the blue contour delineates the OAR. The grey line displays the trace of the line doses. (b) and (c) Line doses : the height of the CTV is the prescribed dose level. x,y and z refers to DICOM coordinates.

	d_{RMSE}	d_{AD}	$\gamma_{\%}$
CTV shifted	1.685	1.345	91.54
OAR	2.379	1.929	69.15

Table 5.12: Statistics between reference dose and adapted dose with smart initialization for the paraspinal case. d_{AD} denotes the absolute dose difference, d_{RMSE} denotes the root mean square dose deviation, $\gamma_{\%}$ denotes the percentage of voxels that has passed the gamma test.

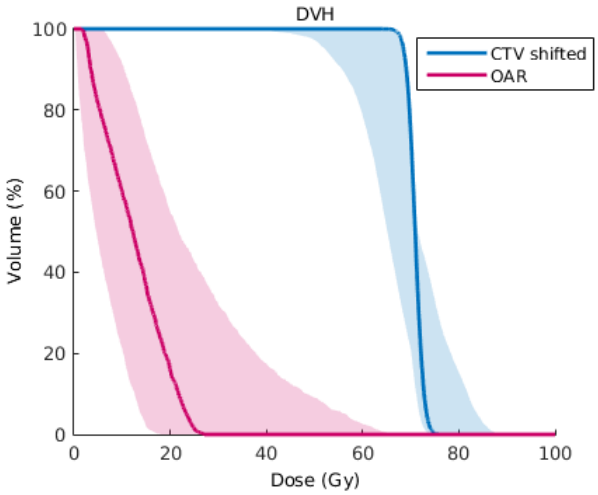


Figure 5.14: DVHs of the adapted dose distribution with smart initialization for the paraspinal case in the nominal scenario (darker line) and in 81 scenarios (27 set ups times 3 density scalings).

6 | Discussion

In this chapter, the results are discussed according to the objectives stated before. Our first purpose is to determine whether our reference dose distribution is close enough to the initial one. This depends on the quality of the non-rigid registration and on the combining strategy for constructing the reference dose. The second one is to determine if the optimization allowed to find planning parameters that keep the previous quality and robustness of the plan. Moreover, an estimation of how much time we have saved compared to a fresh planning can be made.

Reference dose

The quality of the registration can be assessed by looking at the remaining SSD after non-rigid registration of the distance maps, and the difference between characteristic values of the initial and reference dose distributions in the nominal case ($\bar{d}, d_{95}, d_{50}, d_5, d_{max}$).

In both cases (lung and paraspinal), registration ended up with a remaining SSD lower than 0.001. Also, the DVH of the reference dose distribution is very similar to the one of the initial dose distribution in the nominal case. When comparing characteristic values, the difference was lower than 0.01 % in ROIs in which the non-rigid reference dose was replicated. In other structures (the bone and the OAR for the lung case), where the rigid reference dose was replicated, the difference was a bit higher (<1%), but it was expected since their positions relative to the beam paths changed. Indeed, as mentionned before, the use of the rigid registration has a mainly geometrical purpose since it keeps the shape of the beam.

Interestingly, some characteristic values were improved after registration. For instance, in the lung case, d_5 in the CTV for the initial dose distribution was higher than 73.5 Gy, i.e. out of the standard range stated in section 3.3.2. However, d_5 in the CTV for the reference dose is below 73.5 Gy. This could be due to a general decrease of the dose to the CTV, since the mean is also a little lower, which would mean that the registration was not performed well enough, but d_{95} on the other hand is higher for the reference dose, which is also an improvement. Therefore, the non-rigid registration of the initial dose in the CTV resulted in a deformation that yielded a little better dose distribution. The same observation can be made for the CTV in the paraspinal case. However, this is not systematic. Indeed, the mean reference dose to the OAR in the paraspinal case (also constructed from the non-rigid registration of the initial dose distribution) is a little higher than the mean initial dose. In conclusion, we can say that the non-rigid registration was performed well but still changes the dose distribution. While it has very little effects here (<0.01 %), we have seen that it can lead to different conclusions regarding the standards stated in section 3.3.2.

Optimization

The quality of the optimization can be assessed by looking at the difference between the reference and adapted dose distributions, which is given in the nominal case by the root mean square dose

deviation, the mean absolute dose difference and the percentage of success of the gamma test. For robustness, the DVHs and values $\hat{d}$ are compared.

Dose distribution in the CTV In the lung case and without smart initialization, the dose distribution in the CTV was very well reproduced, with a percentage of success for the gamma test of almost 99% and around 1 Gy of mean dose difference between two corresponding voxels. With smart initialization, the percentage of success only reached 70%, and the mean dose difference between two corresponding voxels went up to about 3 Gy. Therefore it seems that smart initialization in this case was not the best option. This is further confirmed by the values d_{95} and d_5 of the adapted dose, which are out of the standard range.

In order to explain this, we have two possibilities. On one hand, the target coverage can be slightly altered by the way spot positions are computed in the case of smart initialization, since the computation of new energies uses approximations (see section 4.4.1). On the other hand, choosing a specific starting point may sometimes be counterproductive. Indeed, since our problem may have several local minima, the choice of starting point can have a strong influence on the solution. For instance, the optimization may not find the best solution because of its proximity to another local minima. As our problem has changed, the weights that were optimal for the initial plan might be too far from the optimal solution of the adapted plan. Moreover, in IPOPT, the starting point can be chosen through a warm start, meaning that we provide not only an initial weights vector, but also the corresponding parameters such as λ and μ (see Equation 3.10). They can also further reduce the flexibility of the optimization method.

In the paraspinal case, the dose distribution in the CTV is similar when using smart initialization or not. It is quite well reproduced, with a percentage of success for the gamma test higher than 90% and less than 2 Gy of root mean square dose difference between two corresponding voxels. Those results are lower than in the lung case without smart initialization, which could be explained by the specificity of this test case since the CTV is brought closer to the OAR. However the dose to the CTV in the adapted plan is generally higher than the reference dose distribution, which seems quite contradictory. Indeed, we would have expected that the dose to the CTV would be generally lower in order to respect specifications of the dose to the OAR. Nevertheless, the dose delivered to the CTV in the nominal case is satisfying.

Dose distribution in other ROIs In both cases, with and without smart initialization, the dose distribution is not reproduced so well. Indeed, the dose delivered in the nominal case of the adapted plan is lower than in the reference dose distribution. However, this was actually allowed by our objective function. Therefore, this difference is acceptable.

Robustness In the lung case, we have noticed that the value $\hat{d}_{95}$ in the CTV has dramatically dropped to about 45 Gy. Although the initial plan already had a discutable target coverage robustness, the adapted plan is thus way worse. Indeed, the DVHs in the CTV are spread over a larger dose range. The same thing can be observed in a lesser extent for the DVHs in the bone, however, OAR sparing robustness seems to be preserved.

In the paraspinal case, without smart initialization, the drop of $\hat{d}_{95}$ in the CTV is clearly not as dramatic (-7 Gy) as in the lung case (about -20 Gy), but the DVHs in the CTV are still spread over a larger dose range, and we cannot talk about a preserved target coverage robustness. Also, OAR sparing robustness is not preserved since $\hat{d}_{max}$ has increased (+ 7 Gy). The loss of robustness in this case seems stronger when using smart initialization.

Compared to the lung case, the loss of robustness here is general and could therefore have occurred as a result of the fact that we allow a lower dose in OARs. Therefore we have performed an additional test which can be found in appendix A.3. The results are very similar to those obtained with the first optimization strategy. Therefore, we can conclude that the loss of robustness is not related to the fact that our objective function allows a lower dose in the OAR. The reason for the loss of robustness should therefore be further investigated.

Computational time

The most time-consuming parts of treatment planning are the dose contribution matrix computation and the optimization. They were both greatly reduced by not introducing robustness into computations. Indeed, for the lung case the time needed for dose contribution matrix was reduced by a factor 18, and the optimization time was reduced by a factor 12 per iterations.

Nevertheless, planning still takes a few hours which is not acceptable for real-time replanning. Indeed, the objective was to perform re-planning while the patient is on the table, thus computation should not exceed a few minutes. However, more efficient machines could greatly further reduce the computation time, as we have already proven when computing the initial plan for the paraspinal case on a faster machine. Also, recall that in the clinical environment, MC computations (for dose contribution matrix) are usually not used since they are slower than other methods like PBA.

7 | Conclusion

The goal of this thesis was to propose and test a strategy for replanning in real time a proton therapy treatment plan. Ideally, the adapted plan would preserve the quality and robustness of the initial plan. Such replanning can be of great use in the future in view of the conjonction of two facts. First, regular treatment planning is cumbersome and is subject to tedious quality assurance protocols in order to be accepted by the physician and physicist in charge. Then, the treatment can last up to 7-8 weeks when it is fractionated. As a consequence, on top of uncertainties inherent to the delivery of protons, we have to deal with inter-fractionnal motion: the patient's tumor and organs may have different relative positions from one treatment session to another.

The strategy proposed includes different steps : image registration, reference dose construction and optimization.

Image registration was performed with an image processing software called REGGUI, using a diffeomorphic Morphons method. It allowed for computing the non-rigid deformation mapping the CT used for initial treatment planning onto the CT used for real-time replanning, i.e. acquired just before the treatment session. A second deformation, rigid this time, was considered, which only took into account the motion of the tumor, regardless of its position relative to other structures.

Reference dose construction was based on two dose distributions constructed by applying both deformations described above on the initial dose distribution. Then, those two dose distributions were combined with each other according to the structure of each region of interest. Typically, the dose distribution computed with the rigid deformation was used to replicate the beam geometry while the other one was used to keep the specific properties of the dose distribution within a specific region of interest. This combination resulted in a so-called reference dose distribution.

Optimization was used to find the planning parameters that would allow to replicate the reference dose distribution. In order to do this, the normalized square dose deviation from the reference dose was minimized, but we allowed lower dose in OARs. Time was gained by using non-robust optimization. The assumption allowing us to do so was that the reference dose would keep the robustness against uncertainties included in the initial plan even after deformation. We also investigated a way to improve our optimization by using a smart initialization of the planning parameters, i.e. using the parameters optimized for the initial plan as the starting point of the new optimization.

Our strategy was tested on two phantom cases : a lung case and a paraspinal case. The inter-fraction motion was represented by a translation of the target. We have shown that our strategy allowed to keep the properties of the initial dose distribution in the nominal case, thanks to a reference dose distribution that was accurately computed and combined, as well as a rela-

tively efficient objective function. However, robustness was hardly preserved. Moreover, smart initialization did not improve the results. Computational time was greatly reduced, but in order to reach the real-time level, further efforts need to be made.

In future work, different improvements can be made to enhance and developp results. First of all, further investigation should be pursued in other to developp additionnal constraints that would help preserve robustness without going into a new robust optimization, which requires the computation of three dose contribution matrices.

Then, it could be interesting to determine cases in which a smart initialization could really improve results. Also, we have mentionned before that using smart initialization induces a very large number of energy layers, which results in a treatment plan that is most likely impossible to deliver. Therefore, smart initialization cannot be performed realistically without optimizing the number of energy layers. The optimization strategy should however keep the spot positions as close as possible to the initial ones in order not to loose the benefits that could be achieved by using smart initialization.

One line of inquiry to reduce re-planning time could be to avoid a new computation of the dose contribution matrix by rescaling the initial one according to the change in density computed in each voxel.

Finally, other types of tumor motion should be considered, such as shrinkage, or any free form deformation. Then the construction of the reference dose would be more complex. The rigid deformation should probably consist in a local free deformation concerning the target only, coupled with the translation of the target isocenter.

A | Appendix

A.1 List of acronyms and abbreviations

- AD : Absolute Difference
- BEV : Beam Eye View
- CT : Computed Tomography
- CTV : Clinical Target Volume
- DICOM : Digital Imaging and COmunication in Medecine
- DTA : Distance-To-Agreement
- DVH : Dose-Volume Histogram
- EM : ElectroMagnetic
- GTV : Gross Target Volume
- HU : Hounsfield Units
- IMPT : Intensity-Modulated Proton Therapy
- IPOPT : Interior Point Optimization
- MC : Monte Carlo
- MCsquare : Mutiple Core Monte Carlo
- MIROpt : MIxed Robust Optimization
- OAR : Organ at risk
- PBA : Pencil Beam Algorithm
- PBS : Pencil Beam Scanning
- PTV : Planning Target Volume
- REGGUI : REGistration Graphical User Interface
- RMSE : Root Mean Square Error
- ROI : Region Of Interest
- SOBP : Spread-out Bragg Peak
- SPR : Stopping Power Ratio

- SSD : Sum of Square Differences
- TPS : Treatment Planning System
- WEPL : Water Equivalent Path Length

A.2 Robustness scenarios

```
Robustness Parameters:
----------------------
Systematic Setup Error = 0.500 0.500 0.500 cm
Random Setup Error = 0.000 0.000 0.000 cm
Systematic Range Error = 3.00 %

Uncertainty scenarios:
----------------------
Nominal: Systematic_Setup(0.000 0.000 0.000 cm) Systematic_Range(0.00 %)
Scenario (1/81): Systematic_Setup(-0.289 -0.289 -0.289 cm) Systematic_Range(-3.00 %)
Scenario (2/81): Systematic_Setup(-0.289 -0.289 -0.289 cm) Systematic_Range(0.00 %)
Scenario (3/81): Systematic_Setup(-0.289 -0.289 -0.289 cm) Systematic_Range(3.00 %)
Scenario (4/81): Systematic_Setup(-0.354 -0.354 0.000 cm) Systematic_Range(-3.00 %)
Scenario (5/81): Systematic_Setup(-0.354 -0.354 0.000 cm) Systematic_Range(0.00 %)
Scenario (6/81): Systematic_Setup(-0.354 -0.354 0.000 cm) Systematic_Range(3.00 %)
Scenario (7/81): Systematic_Setup(-0.289 -0.289 0.289 cm) Systematic_Range(-3.00 %)
Scenario (8/81): Systematic_Setup(-0.289 -0.289 0.289 cm) Systematic_Range(0.00 %)
Scenario (9/81): Systematic_Setup(-0.289 -0.289 0.289 cm) Systematic_Range(3.00 %)
Scenario (10/81): Systematic_Setup(-0.354 0.000 -0.354 cm) Systematic_Range(-3.00 %)
Scenario (11/81): Systematic_Setup(-0.354 0.000 -0.354 cm) Systematic_Range(0.00 %)
Scenario (12/81): Systematic_Setup(-0.354 0.000 -0.354 cm) Systematic_Range(3.00 %)
Scenario (13/81): Systematic_Setup(-0.500 0.000 0.000 cm) Systematic_Range(-3.00 %)
Scenario (14/81): Systematic_Setup(-0.500 0.000 0.000 cm) Systematic_Range(0.00 %)
Scenario (15/81): Systematic_Setup(-0.500 0.000 0.000 cm) Systematic_Range(3.00 %)
Scenario (16/81): Systematic_Setup(-0.354 0.000 0.354 cm) Systematic_Range(-3.00 %)
Scenario (17/81): Systematic_Setup(-0.354 0.000 0.354 cm) Systematic_Range(0.00 %)
Scenario (18/81): Systematic_Setup(-0.354 0.000 0.354 cm) Systematic_Range(3.00 %)
Scenario (19/81): Systematic_Setup(-0.289 0.289 -0.289 cm) Systematic_Range(-3.00 %)
Scenario (20/81): Systematic_Setup(-0.289 0.289 -0.289 cm) Systematic_Range(0.00 %)
Scenario (21/81): Systematic_Setup(-0.289 0.289 -0.289 cm) Systematic_Range(3.00 %)
Scenario (22/81): Systematic_Setup(-0.354 0.354 0.000 cm) Systematic_Range(-3.00 %)
Scenario (23/81): Systematic_Setup(-0.354 0.354 0.000 cm) Systematic_Range(0.00 %)
Scenario (24/81): Systematic_Setup(-0.354 0.354 0.000 cm) Systematic_Range(3.00 %)
Scenario (25/81): Systematic_Setup(-0.289 0.289 0.289 cm) Systematic_Range(-3.00 %)
Scenario (26/81): Systematic_Setup(-0.289 0.289 0.289 cm) Systematic_Range(0.00 %)
Scenario (27/81): Systematic_Setup(-0.289 0.289 0.289 cm) Systematic_Range(3.00 %)
Scenario (28/81): Systematic_Setup(0.000 -0.354 -0.354 cm) Systematic_Range(-3.00 %)
Scenario (29/81): Systematic_Setup(0.000 -0.354 -0.354 cm) Systematic_Range(0.00 %)
Scenario (30/81): Systematic_Setup(0.000 -0.354 -0.354 cm) Systematic_Range(3.00 %)
Scenario (31/81): Systematic_Setup(0.000 -0.500 0.000 cm) Systematic_Range(-3.00 %)
Scenario (32/81): Systematic_Setup(0.000 -0.500 0.000 cm) Systematic_Range(0.00 %)
Scenario (33/81): Systematic_Setup(0.000 -0.500 0.000 cm) Systematic_Range(3.00 %)
Scenario (34/81): Systematic_Setup(0.000 -0.354 0.354 cm) Systematic_Range(-3.00 %)
Scenario (35/81): Systematic_Setup(0.000 -0.354 0.354 cm) Systematic_Range(0.00 %)
```

```
Scenario (36/81): Systematic_Setup(0.000 -0.354 0.354 cm) Systematic_Range(3.00 %)
Scenario (37/81): Systematic_Setup(0.000 0.000 -0.500 cm) Systematic_Range(-3.00 %)
Scenario (38/81): Systematic_Setup(0.000 0.000 -0.500 cm) Systematic_Range(0.00 %)
Scenario (39/81): Systematic_Setup(0.000 0.000 -0.500 cm) Systematic_Range(3.00 %)
Scenario (40/81): Systematic_Setup(0.000 0.000 0.000 cm) Systematic_Range(-3.00 %)
Scenario (41/81): Systematic_Setup(0.000 0.000 0.000 cm) Systematic_Range(0.00 %)
Scenario (42/81): Systematic_Setup(0.000 0.000 0.000 cm) Systematic_Range(3.00 %)
Scenario (43/81): Systematic_Setup(0.000 0.000 0.500 cm) Systematic_Range(-3.00 %)
Scenario (44/81): Systematic_Setup(0.000 0.000 0.500 cm) Systematic_Range(0.00 %)
Scenario (45/81): Systematic_Setup(0.000 0.000 0.500 cm) Systematic_Range(3.00 %)
Scenario (46/81): Systematic_Setup(0.000 0.354 -0.354 cm) Systematic_Range(-3.00 %)
Scenario (47/81): Systematic_Setup(0.000 0.354 -0.354 cm) Systematic_Range(0.00 %)
Scenario (48/81): Systematic_Setup(0.000 0.354 -0.354 cm) Systematic_Range(3.00 %)
Scenario (49/81): Systematic_Setup(0.000 0.500 0.000 cm) Systematic_Range(-3.00 %)
Scenario (50/81): Systematic_Setup(0.000 0.500 0.000 cm) Systematic_Range(0.00 %)
Scenario (51/81): Systematic_Setup(0.000 0.500 0.000 cm) Systematic_Range(3.00 %)
Scenario (52/81): Systematic_Setup(0.000 0.354 0.354 cm) Systematic_Range(-3.00 %)
Scenario (53/81): Systematic_Setup(0.000 0.354 0.354 cm) Systematic_Range(0.00 %)
Scenario (54/81): Systematic_Setup(0.000 0.354 0.354 cm) Systematic_Range(3.00 %)
Scenario (55/81): Systematic_Setup(0.289 -0.289 -0.289 cm) Systematic_Range(-3.00 %)
Scenario (56/81): Systematic_Setup(0.289 -0.289 -0.289 cm) Systematic_Range(0.00 %)
Scenario (57/81): Systematic_Setup(0.289 -0.289 -0.289 cm) Systematic_Range(3.00 %)
Scenario (58/81): Systematic_Setup(0.354 -0.354 0.000 cm) Systematic_Range(-3.00 %)
Scenario (59/81): Systematic_Setup(0.354 -0.354 0.000 cm) Systematic_Range(0.00 %)
Scenario (60/81): Systematic_Setup(0.354 -0.354 0.000 cm) Systematic_Range(3.00 %)
Scenario (61/81): Systematic_Setup(0.289 -0.289 0.289 cm) Systematic_Range(-3.00 %)
Scenario (62/81): Systematic_Setup(0.289 -0.289 0.289 cm) Systematic_Range(0.00 %)
Scenario (63/81): Systematic_Setup(0.289 -0.289 0.289 cm) Systematic_Range(3.00 %)
Scenario (64/81): Systematic_Setup(0.354 0.000 -0.354 cm) Systematic_Range(-3.00 %)
Scenario (65/81): Systematic_Setup(0.354 0.000 -0.354 cm) Systematic_Range(0.00 %)
Scenario (66/81): Systematic_Setup(0.354 0.000 -0.354 cm) Systematic_Range(3.00 %)
Scenario (67/81): Systematic_Setup(0.500 0.000 0.000 cm) Systematic_Range(-3.00 %)
Scenario (68/81): Systematic_Setup(0.500 0.000 0.000 cm) Systematic_Range(0.00 %)
Scenario (69/81): Systematic_Setup(0.500 0.000 0.000 cm) Systematic_Range(3.00 %)
Scenario (70/81): Systematic_Setup(0.354 0.000 0.354 cm) Systematic_Range(-3.00 %)
Scenario (71/81): Systematic_Setup(0.354 0.000 0.354 cm) Systematic_Range(0.00 %)
Scenario (72/81): Systematic_Setup(0.354 0.000 0.354 cm) Systematic_Range(3.00 %)
Scenario (73/81): Systematic_Setup(0.289 0.289 -0.289 cm) Systematic_Range(-3.00 %)
Scenario (74/81): Systematic_Setup(0.289 0.289 -0.289 cm) Systematic_Range(0.00 %)
Scenario (75/81): Systematic_Setup(0.289 0.289 -0.289 cm) Systematic_Range(3.00 %)
Scenario (76/81): Systematic_Setup(0.354 0.354 0.000 cm) Systematic_Range(-3.00 %)
Scenario (77/81): Systematic_Setup(0.354 0.354 0.000 cm) Systematic_Range(0.00 %)
Scenario (78/81): Systematic_Setup(0.354 0.354 0.000 cm) Systematic_Range(3.00 %)
Scenario (79/81): Systematic_Setup(0.289 0.289 0.289 cm) Systematic_Range(-3.00 %)
Scenario (80/81): Systematic_Setup(0.289 0.289 0.289 cm) Systematic_Range(0.00 %)
Scenario (81/81): Systematic_Setup(0.289 0.289 0.289 cm) Systematic_Range(3.00 %)
```

A.3 Additional test

This additional test was performed in order to determine whether the choice of allowing the adapted dose to be lower than the reference dose was legitimate. For this test (in the paraspinal

case and without smart initialization), the optimization function in the OAR was the same as the one in the CTV.

Dose contribution matrix computation with MCSquare lasted 3 hours, and optimization with IPOPT took 13 minutes. It stopped after 250 iterations with an optimality error given by 8.89×10^{-3}.

Isodose curves for the dose distribution in the nominal scenario and a perturbed scenario in an axial slice are shown on Figure A.1a. Line dose from the same slice is shown on Figures A.1b and A.1c. The nominal 95% isodose curve encloses all parts of the CTV, but extra dose is delivered anterior to the CTV. The perturbed 95% dose does not enclose all of the CTV, and leaves some gaps at several positions.

DVHs are displayed on Figure A.2. As for other adapted plans, the dose to the CTV declines earlier than in the initial dose distribution. Mean and quantiles of the dose in each ROI are given in Table A.1. Values in bold exceed the standards stated in section 3.3.2. Indeed, $\hat{d}_{95}$ is lower than 66.5 Gy, $\hat{d}_5$ is greater than 73.5 Gy and $\hat{d}_{max}$ in the OAR is greater than 45 Gy.

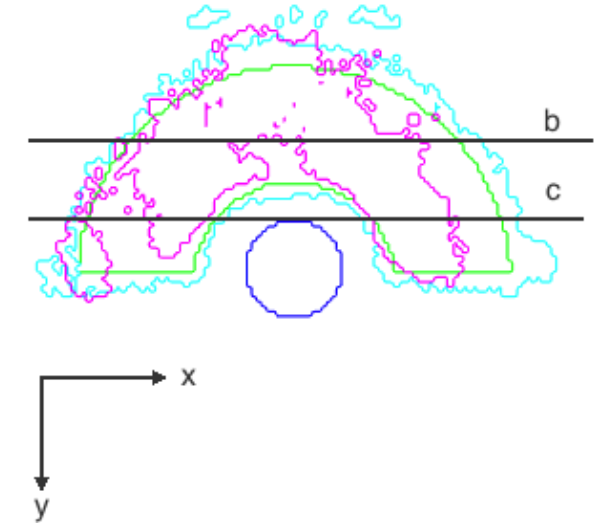
Difference between the reference dose and the adapted dose is given by the quantities computed in Table A.2. The percentage of success of the gamma test in the CTV does not reach 50%.

	$\bar{d}$	d_{95}	d_{50}	d_5	d_{max}	$\hat{d}_{95}$	$\hat{d}_5$	$\hat{d}_{max}$
CTV shifted	70.9	68.64	70.9	73.13	75.9	**58.07**	**79.45**	86
OAR	13.92	2.056	14.44	25.82	29.5	/	**49.45**	**57.7**

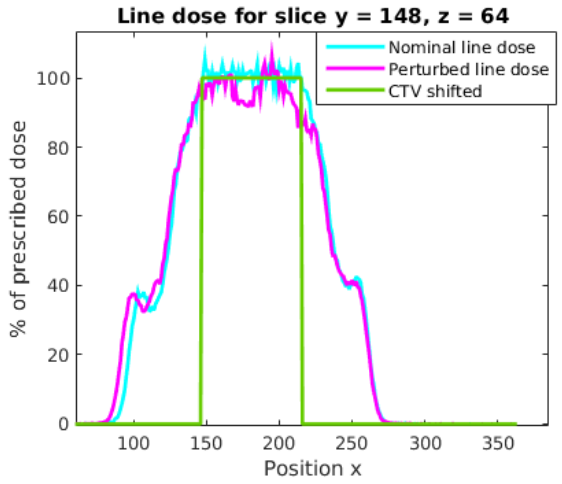
Table A.1: Adapted dose statistics for the paraspinal case with a new optimization strategy 2. $\bar{d}$ denotes the mean dose, d_x denotes minimum dose to x % of the ROI volume in the nominal case, d_{max} denotes the maximum dose in the nominal case, $\hat{d}_x$ denotes dose to x % of the ROI volume in the worst scenario (minimum for $x = 95$ and maximum otherwise), and $\hat{d}_{max}$ denotes the maximum dose in the worst case scenario.

	d_{RMSE}	d_{AD}	$\gamma_{\%}$
CTV shifted	54.77	49.27	45.5
OAR	25.67	17.27	29.69

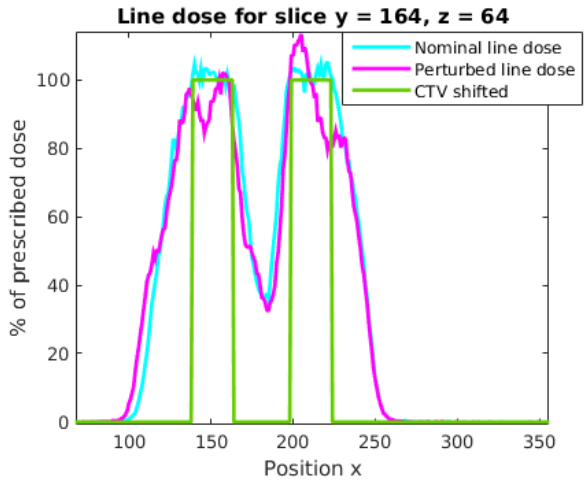
Table A.2: Statistics between reference dose and adapted dose with the new optimization strategy for the paraspinal case. d_{AD} denotes the absolute dose difference, d_{RMSE} denotes the root mean square dose deviation, $\gamma_{\%}$ denotes the percentage of voxels that has passed the gamma test.



(a)



(b)



(c)

Figure A.1: Isodose curves and line doses of the adapted dose distribution with the new optimization strategy and without smart initialization in the nominal case and a perturbed scenario in a transversal slice of the paraspinal case. (a) The isodose curves show 95% of the prescribed dose : the light blue curve shows the nominal case and the pink one a perturbed scenario. In the perturbed scenario the density is 3% higher than planned and the isocenter is shifted is shifted 2.89 mm diagonally in the directions posterior, superior and left. The green contour delineates the CTV and the blue contour delineates the OAR. The grey line displays the trace of the line doses. (b) and (c) Line doses : the height of the CTV is the prescribed dose level. x,y and z refers to DICOM coordinates.

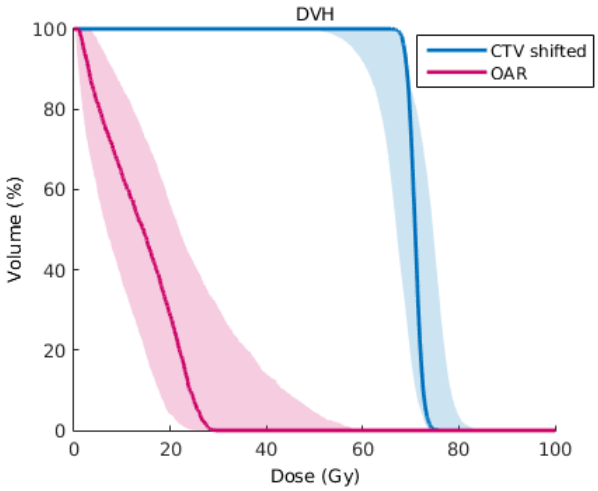


Figure A.2: DVHs of the adapted dose distribution with the new optimization strategy for the paraspinal case in the nominal scenario (darker line) and in 81 scenarios (27 set ups times 3 density scalings).

Bibliography

[1] H. Paganetti. *Proton therapy physics.* CRC Press, 2011.

[2] A. Barragan Montero. Implementation of delivery uncertainties in robust optimization for proton therapy. Université Catholique de Louvain - PhD Day, 2016.

[3] A. Fredriksson. *Robust optimization of radiation therapy accounting for geometric uncertainty.* Phd thesis, Royal Institute of Technology, Stockholm, Sweden, 2013.

[4] S. Jiang, X. Gu, C. Men, X. Jia, O. Fluck, D.J. Choi, and A. Majumdar. Real-time replanning for online adaptive radiotherapy. Slideshow.

[5] A. Bol. Medical imaging, GBIO2050 course. Université Catholique de Louvain, 2014.

[6] J. Malo, C. Luraschi-Monjagatta, D. M. Wolk, R. Thompson, C. A. Hage, and K. S. Knox. Update on the diagnosis of pulmonary coccidioidomycosis. *Annals of the American Thoracic Society*, 11:243–53, 2014.

[7] Dicom. URL = http://dicom.neam.org.

[8] A. M. Barragan Montero. MIROpt user guide, 2016.

[9] N.G. Burnet, S.J. Thomas, K.E. Burton, and S.J. Jefferies. Defining the tumor and target volumes for radiotherapy. *Cancer Imaging*, 4:153–161, 2004.

[10] National cancer institute at the national institutes of health, u.s. department of health and human services. URL = http://www.cancer.gov.

[11] N. Saito and C. Bert. Scanned ion beam therapy of moving targets with beam tracking. *Modern Practices in Radiation Therapy*, pages 70–71, 2012.

[12] M. Durante and J. S. Loeffler. Charged particles in radiation oncology. *Nature Reviews Clinical Oncology*, 7:37–43, 2010.

[13] Stopping power (n.d.), 2011. Retrieved May 11 2016 from http://medical-dictionary.thefreedictionary.com/stopping+power.

[14] Loic Grevillot. Procedure to extract the BDL parameters using GATE, 2011. Report.

[15] J. Amanatides and A. Woo. A fast voxel traversal algorithm for ray tracing.

[16] Fluence (n.d.), 2012. Retrieved May 11 2016 from http://medical-dictionary.thefreedictionary.com/fluence.

[17] U. Oelfke and T. Bortfeld. Inverse planning for photons and proton beams. *Medical Dosimetry*, 26:113–124, 2001.

[18] F. Albertini, E. B. Hug, and A.J. Lomax. Is it necessary to plan with safety margins for actively scanned proton therapy? *Physics in Medicine and Biology*, 56, 2011.

[19] A. Fredriksson, A. Forsgren, and B. Hardemark. Minimax otpimization for handling range and setup uncertainties in proton therapy. *Medical Physics*, 38:1672–1684, 2011.

[20] J. W. Chapman. *Photon dose calculation in homogenous media.* PhD thesis, Louisiana State University, 2012.

[21] K. Souris, J. Lee, and E. Sterpin. Fast multipurpose Monte Carlo simulation for proton therapy using multi- and many- core CPU architecture. *Medical Physics*, 43:1700–1712, 2016.

[22] Wikipedia. Methode de Monte Carlo. Site web.

[23] S. Jan, G. Santin, D. Strul, S. Staelens, K. Assi, D. Autret, S. Avner, R. Barbier, M. Bardi, P. S. Bloomfield, D. Brasse, V. Breton, P. Bruyndonckx, I. Buvat, A. F. Chatziioannou, Y. Choi, Y. H. Chung, C. Comtat, D. Donnarieix, L. Ferrer, S. J. Glick, C. J. Groiselle, D. Guez, P-F. Honore, S. Kerhoas-Cavata, A. S. Kirov, V. Kohli, M. Koole, M. Krieguer, D. J. van der Laan, F. Lamare, G. Largeron, C. Lartizien, Lazaro C., M. C. Maas, L. Maigne, F. Mayet, F. Melot, C. Merheb, E. Pennacchio, J. Perez, U. Pietrzyk, F. R. Rannou, M. Rey, D. R. Schaart, C. R. Schmidtlein, L. Simon, T. Y. Song, J-M. Vieira, D. Visvikis, R. Van de Walle, R. Wie, and C. Morel. GATE : a simulation toolkit for pet and spect. *Physics in Medecine and Biology*, 48:4543–4561, 2004.

[24] Y. Nesterov. Non linear optimization, LINMA2460 course. Université Catholique de Louvain, 2016.

[25] A. Wachter and L. T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 2005.

[26] M. Glavic and Wehenkel L. Interior point method : A survey, short survey of applications to power systems, and research opportunities. Technical report, 2004.

[27] R. Fletcher and S. Leyffer. Non linear programming without a penalty function. *Mathematical Programming*, 91:239–269, 2002.

[28] F. Glineur. Modèles et méthode d'optimisation, LINMA1702 course. Université Catholique de Louvain, 2013.

[29] A. Skajaa. *Limited Memory BFGS for Nonsmooth Optimization.* PhD thesis, Institute of Mathematical Science new York University, 2010.

[30] R. Hauser. Numerical linear algebra and optimization. Oxford University Computing Laboratory, 2007.

[31] Société Française de Radiothérapie Oncologique and Société Française de Physique Médicale. Guide des procédures de radiothérapie externe 2007, 2007.

[32] P. Maingon, V. Mammar, K. Peignaux, G. Truc, and I. Barillot. Les contraintes aux organes à risque en radiothérapie par modulation d'intensité des cancers orl. *Cancer/Radiothérapie*, 8:234–247, 2004.

[33] D. A. Low, W. B. Harms, S. Mutic, and J. A. Purdy. A technique for the quantitative evaluation of dose distribution. *Medical Physics*, 25:656, 1998.

[34] Roy. Comparing dose distributions : Dta and dose-difference, 2012. Blog : 365 Days of Medical Physics. URL = http://medphys365.blogspot.be/2012/05/comparing-dose-distributions-dta-and.html.

[35] Mark Geurts, 2015. Software available on GitHub : https://github.com/mwgeurts/gamma.

[36] G. Janssens. REGGUI course, 2011.

[37] J. Lee. Medical imaging, GBIO2050 course. Université Catholique de Louvain, 2014.

[38] G. Janssens, L. Jacques, J. Orban de Xivry, X. Geets, and B. Macq. Diffeomorphic registration of images with variable contrast enhancement. *International Journal of Biomedical Imaging*, 2011.

[39] A. Wrangsjo, J. Pettersson, and H. Knutsson. Non-rigid registration using morphons. *Proceedings of the 14th Scandinavian conference on image analysis*, pages 1226–1229, 2005.