Robust, accurate, and patient-specific treatment planning for proton therapy

Ana María Barragán Montero



A thesis presented for the degree of
Doctor of Biomedical and Pharmaceutical Sciences





Molecular Imaging, Radiotherapy and Oncology (MIRO)
Institut de Recherche Expérimentale et Clinique (IREC)
Université catholique de Louvain (UCLouvain)
Louvain-La-Neuve, Belgium

Cover picture adapted by Ana Barragán, original by HildeErling (Own work) [CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0)], via Wikimedia Commons.

During this thesis, Ana Barragán was supported by a Télévie Grant from the Belgian 'Fonds National pour la Recherche Scientifique 'F.R.S.-FNRS with projet numbers 7.4566.13 (2013-2015) and 7.4581.15 (2015-2017).



Acknowledgements - Remerciements

J'ai quitté l'Espagne en septembre 2013 pour débuter cette thèse de doctorat au sein de l'unité MIRO. Je n'aurai jamais pu imaginer la richesse de cette expérience, tant sur le plan personnel que professionnel. Après quatre ans et de nombreuses heures de travail, après beaucoup de moments de stress et de joie, je n'ai pas assez de mots pour remercier, tout d'abord, les trois personnes qui ont rendu tout ça possible, tout en le rendant agréable: mes promoteurs John et Edmond et mon collègue (presqu'un troisième promoteur) et ami Kevin. Merci John pour ta patience sans limite, ta sagesse infinie et pour ta capacité à toujours transmettre un sentiment de sérénité à tes doctorants, même dans les situations compliquées. Merci aussi d'organiser les fameuses célébrations du vendredi après-midi. Merci Edmond de m'avoir transmis ta passion pour la science et la recherche, pour ton enthousiasme quotidien et d'avoir toujours partagé avec nous les idées qui te passent sans cesse par la tête. Merci aussi de m'avoir appris qu'à Bruxelles, il y a des peïs et des meïs, qu'il drache et qu'on peut, entre autres, faire des choses en stoemelings. Merci Kevin d'avoir toujours été là, littéralement, à deux mètres de moi, toujours prêt à discuter et à répondre à mes questions. Merci aussi d'avoir maintenu la synchronisation de l'horloge de nos estomacs à midi, et merci pour tes réflexions à la ≪Gandalf≫ qui nous ont tant fait rire.

Mes remerciements vont aussi au reste des membres de l'unité MIRO, qui ont toujours maintenu une ambiance agréable dans le laboratoire. Merci, notamment, pour toutes ces discussions intéressantes sur le temps de midi. Merci spécialement à Dario et Sarah pour votre aide concernant les questions médicales et les images des patients.

Merci aussi aux membres du comité d'encadrement qui ont suivi mon travail pendant ces quatre années (Jonathan, Dirk et Xavier) pour vos conseils et suggestions. Merci aux membres extérieurs d'avoir pris le temps de lire en détail mon manuscrit. Merci, en particulier, à Xavier d'avoir partagé avec nous ton expérience médicale et tes contributions pertinentes durant nos réunions entre physiciens.

Para terminar, gracias a mi familia (papá, mamá y Anto) por haber estado siempre ahí, apoyándome en los buenos y malos momentos, y en especial, por propiciarme un ambiente idóneo para la escritura final de la tesis, facilitando considerablemente estas últimas semanas de estrés y agobios. Y a Dani, por transmitirme siempre esa calma que te caracteriza y que tanto me ha ayudado en los peores momentos.

Abstract

The survival statistics for cancer patients treated with radiation therapy using photon beams show that many treatments fail due to poor tumour local control (TLC). A potential solution would be to increase the target dose, but this often entails high toxicity in the nearby healthy tissue. Unlike photons, protons release most of their energy at the end of their path (the so-called Bragg peak), reducing the dose to healthy tissue, which might be the key for safe dose escalation. Moreover, functional images can reveal spatial heterogeneity in the tumor radioresistance pattern and therefore allow for improved targeting (dose painting) and possibly enhanced TLC. In this context, special care must be taken to accurately model the possible uncertainties in the position of the Bragg peak, since they can strongly deteriorate treatment quality, especially in very heterogeneous dose distributions as in dose painting plans. In this thesis, we investigated the use of the most advanced techniques, such as Monte Carlo dose calculation and robust optimization, to ensure accurate and robust planning for (dose painted) proton therapy treatments.

Résumé

Les cancers traités par radiothérapie avec faisceaux de photons montrent un taux significatif de récidive locale. Pour y remédier, augmenter la dose d'irradiation à la cible est une approche intuitivement évidente, mais qui présente un risque de toxicité inacceptable. Contrairement aux photons, les protons libèrent un pic d'énergie au bout de leur parcours (le pic de Bragg), ce qui diminue les doses aux tissus sains et donc ouvre la voie vers une augmentation de dose à moindre risque. En outre, les tumeurs présentent une radiorésistance spatialement hétérogène dont le ciblage sur mesure grâce à l'imagerie fonctionnelle (dose painting) pourrait améliorer le contrôle local pour une toxicité acceptable. Dans ce cadre, les incertitudes relatives à la position du pic de Bragg doivent être adéquatement gérées car elles pourraient fortement perturber la qualité du traitement, a fortiori pour le dose painting. Dans cette thèse, nous avons étudié l'usage des techniques les plus avancées, telles que le calcul de dose par simulation Monte Carlo et l'optimisation robuste, pour assurer une planification de pointe des traitements de dose painting par protonthérapie.

PhD committee

Promoters

Prof. John A. Lee, PhD, Ir
Molecular Imaging, Radiotherapy and Oncology, Université catholique de Louvain
École polytechnique de Louvain, Université catholique de Louvain

Prof. Edmond Sterpin, PhD, Ir
Molecular Imaging, Radiotherapy and Oncology, Université catholique de Louvain
Laboratory of Experimental Radiotherapy, Katholieke Universiteit Leuven

President of the jury

Prof. Frédéric Lecouvet, PhD, MD
Radiology, Cliniques Universitaires Saint Luc
Université catholique de Louvain

Members

Prof. Xavier Geets, PhD, MD
Radiation Oncology, Cliniques Universitaires Saint Luc
Molecular Imaging, Radiotherapy and Oncology, Université catholique de Louvain

Dr. Jonathan Orban, PhD, Ir
Ion Beam Applications S.A. (IBA) - Louvain La-Neuve

Prof. Dirk Verellen, PhD
GZA Ziekenhuizen - Iridium Kankernetwerk

Faculty of Medicine and Pharmacy, Vrije Universiteit Brussel

Invited members

Prof. Uwe Oelkfe, PhD
Division of Radiotherapy and Imaging
Royal Marsden Hospital, The Institute of Cancer Research - ICR

Dr. Nick Reynaert, PhD
Service de Physique Médicale, Centre Oscar Lambret

List of publications

Peer reviewed publications

A. M. Barragán Montero, S. Differding, G. Janssens, J. A. Lee, and E. Sterpin, Feasibility and robustness of dose painting by numbers in proton therapy with contour-driven plan optimization, Med Phys. 2015;42(4):2006-17

A. M. Barragán Montero, K. Souris, D. Sánchez Parcerisa, E. Sterpin, and J. A. Lee, Performance of a hybrid Monte Carlo-Pencil Beam dose algorithm for proton therapy inverse planning, Under revision in Med Phys

A. M. Barragán Montero, K. Souris, J. A. Lee, and E. Sterpin, Efficient implementation of 4D robust optimization for scanned proton therapy with Monte Carlo, Manuscript in preparation

E. Sterpin, A. M. Barragán Montero, K. Souris, and J. A. Lee, *Robust treatment planning in proton therapy*, Cancer Radiother. 2016;20(6-7):523-9

S. Michiels, A. M. Barragán Montero, K. Souris, K. Poels, W. Crijns, J. A. Lee, E. Sterpin, S. Nuyts, K. Haustermans, and T. Depuydt, *Patient-specific bolus for range shifter air gap reduction in intensity-modulated proton therapy of head-and-neck cancer studied with Monte Carlo based plan optimization*, Radiother Oncol. 2017; pii: S0167-8140(17)32568-9

K. Souris, A. M. Barragán Montero, G. Janssens, D. Di Perri, E. Sterpin, and J. A. Lee, *Monte Carlo methods to realistically evaluate the robustness of proton therapy plans*, submitted to Med Phys (September 2017)

Conference communications

A. M. Barragán Montero, S. Differding, J. A. Lee, and E. Sterpin, *SU-E-T-182 Feasibility of dose painting by numbers in proton therapy by using contour-driven plan optimization*, AAPM 2014, Austin - TX (USA)

A. M. Barragán Montero, J. A. Lee, and E. Sterpin, *OC-0070 Automatic dose painting workflow: from tumor segmentation to optimization*, 3rd ESTRO Forum 2015, Barcelona (Spain)

A. M. Barragán Montero, E. Sterpin, and J. A. Lee, *SU-F-BRD-05 Robustness of dose painting by numbers in proton therapy*, AAPM 2015, Anaheim - CA (USA)

A. M. Barragán Montero, K. Souris, E. Sterpin, and J. A. Lee, *OC-0265 Efficient implementation of random errors in robust optimization for proton therapy with Monte Carlo*, ESTRO 2016, Turin (Italy)

A. M. Barragán Montero, K. Souris, E. Sterpin, and J. A. Lee, *WE-AB-209-01 A Monte Carlo-based method to include random errors in robust optimization*, AAPM 2016, Washington - DC (USA)

K. Souris, A. M. Barragán Montero, E. Sterpin, D. Di Perri, X. Geets, and J. Lee, *MO-FG- CAMPUS-TeP3-01 A model of baseline shift to improve robustness of proton therapy treatments of moving tumors*, AAPM 2016, Washington - DC (USA)

A. M. Barragán Montero, K. Souris, D. Sánchez Parcerisa, J. A. Lee, and E. Sterpin, *EP-1519 Implementation of a hybrid superfast Monte Carlo-Pencil Beam dose optimizer for proton therapy*, ESTRO 2017, Vienna (Austria)

A. M. Barragán Montero, K. Souris, D. Sánchez Parcerisa, J. A. Lee, and E. Sterpin, *SU-K-108-13 Performance of a hybrid Monte Carlo-Pencil beam dose algorighm for proton therapy*, AAPM 2017, Denver - CO (USA)

Contents

1 Introduction 3
1.1 General overview of the thesis . . . 4
1.1.1 General context and motivation . . . 4
1.1.2 Contributions . . . 7
1.1.3 Structure of the thesis . . . 9
1.2 State-of-the-art . . . 11
1.2.1 Proton therapy . . . 11
1.2.2 Personalized therapy through dose painting . . . 24

2 Treatment planning tools 31
2.1 Robust plan optimization with MIROpt . . . 32
2.1.1 Treatment design . . . 32
2.1.2 Optimization scheme . . . 34
2.1.3 Simulation of uncertainties . . . 38
2.1.4 Dose calculation . . . 42
2.2 Automatic dose painting workflow . . . 44
2.2.1 Target segmentation and dose prescription . . . 45
2.2.2 Plan optimization . . . 47
2.2.3 Robustness against treatment uncertainties . . . 48
2.2.4 Evaluation of the dose painting accuracy . . . 49

3 Accurate treatment planning for proton therapy 51
3.1 Introduction . . . 52
3.2 Materials and methods . . . 54
3.2.1 Dose engines . . . 54
3.2.2 Optimization scheme . . . 55
3.2.3 Feasibility study . . . 57
3.3 Results . . . 61
3.3.1 Test cases . . . 61
3.3.2 Patient cases . . . 62

3.4 Discussion . . . 72
3.5 Conclusion . . . 75

4 Robust treatment planning for proton therapy 77
4.1 Introduction . . . 78
4.2 Materials and methods . . . 80
4.2.1 Treatment planning system . . . 80
4.2.2 Monte Carlo 4D dose calculation . . . 81
4.2.3 Robustness against patient setup and range errors . . . 83
4.2.4 Robustness against motion . . . 83
4.2.5 Plan evaluation . . . 86
4.2.6 Clinical case . . . 86
4.3 Results . . . 87
4.3.1 Plan quality and robustness . . . 87
4.3.2 Computation time . . . 88
4.4 Discussion . . . 89
4.5 Conclusion . . . 93

5 Patient specific therapy through dose painting by numbers 95
5.1 Introduction . . . 96
5.2 Materials and methods . . . 97
5.2.1 Target definition . . . 97
5.2.2 DPBN prescription . . . 98
5.2.3 Plan robustness against setup and range errors . . . 99
5.2.4 Planning objectives . . . 100
5.2.5 Plan assessment . . . 101
5.2.6 Design of the study . . . 101
5.3 Results . . . 105
5.3.1 Planning parameters for DPBN plans . . . 105
5.3.2 Feasibility and robustness of DPBN plans . . . 107
5.3.3 Treatment time . . . 110
5.4 Discussion . . . 111
5.5 Conclusions . . . 113

6 Conclusions and outlook 119

A Convergence conditions of the hybrid algorithm 127

B Robustness against random setup errors: implementation and results 133
B.1 Background and purpose . . . 133
B.2 Materials and methods . . . 134

B.3 Results 135
B.4 Discussion 135
B.5 Conclusion 137

C Automatic dose painting module: implementation and results 139
C.1 Purpose 139
C.2 Materials and methods 139
C.3 Results 140
C.4 Conclusion 142

References 143

Chapter 1

Introduction

1.1 General overview of the thesis

1.1.1 General context and motivation

According to the World Health Organization (WHO), cancer is the second leading cause of mortality worldwide, and it was responsible for 8.8 million deaths in 2015 [118, 161, 162]. Cancer affects both developing and developed countries, and its incidence is expected to dramatically increase (70% over the next 20 years), due to several factors such as the growth and aging of the population, and the increasing prevalence of risk elements like tobacco and alcohol use, overweight, physical inactivity or viral infection [136], among others.

The main types of cancer treatment are surgery, radiation therapy, chemotherapy and other systemic therapies [123]. For radiation therapy, there are two ways of delivery: internal (placing radioactive material inside the body) and external (by means of a radiation beam), from which the external one constitutes the most used modality. In particular, radiation therapy is estimated to be needed in around 50% of the cases [11, 10, 69, 164]. Thus, an important weapon in the existing arsenal for fighting cancer.

Radiation therapy kills cancer cells by means of ionizing radiation such as X-rays, gamma rays or charged particles. When the ionizing radiation passes through the body, it causes damage in DNA of the cells, which makes cancer cells stop dividing or die. However, normal cells can also be damaged, leading to side effects. Therefore, the goal in radiation therapy is to reduce the exposure of healthy tissue as much as possible, while keeping a dose to the tumor sufficiently high to ensure the eradication of tumor cells.

External beam radiation therapy with photons, usually known as conventional radiotherapy, is the most used modality nowadays, since the modern machines achieve a very good treatment quality, with high dose conformity and healthy tissue sparing. However, the interest of the community in charged particles therapy, especially by proton therapy, has increased considerably in the last decade [129]. Indeed, protons deliver most of the energy at the end of their path, leading to the so-called *Bragg peak*, which entails a very low dose at the entrance and no dose after the tumor (Figure 1.1). This enables a more targeted therapy with increased tissue sparing and constitutes a huge advantage with respect to photons, where the maximum dose is located at the entrance followed by a smooth exponential decrease. This inherent physical advantage makes proton therapy a very appealing technique.

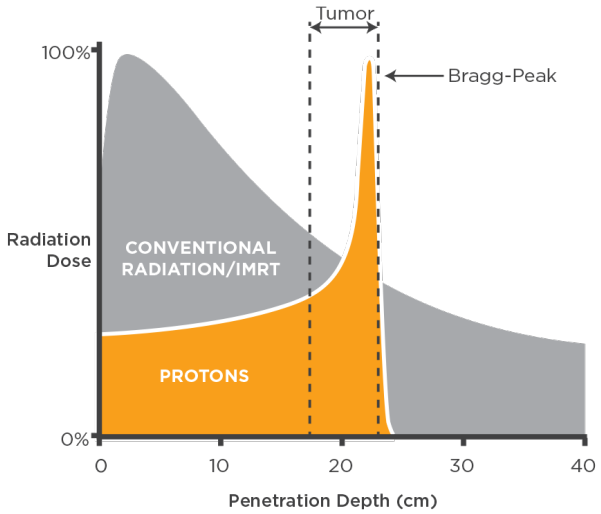


Figure 1.1: Dose deposition for conventional radiotherapy versus proton therapy [130].

However, in spite of the advances in radiation therapy, the long-term follow-up data (5 years survival) for patients treated with (chemo)radiation shows that an important part of the treatments fails due to unsuccessful tumor local control. For instance, the proportion of local control failure can reach up to 30 % for locally advanced head and neck cancer patients [79], and up to 70 % in lung cancer patients [8, 106]. Increasing the dose to the tumor could improve the tumor local control probability (TCP), but this would be at the expense of increased risk for the normal tissue (normal tissue complication, NTCP) [13]. A popular way to improve the therapeutic ratio (i.e, low NTCP vs high TCP) is to increase the dose only for the most resistant parts of the tumor [115, 167, 87, 139, 34, 22, 29]. Functional and molecular image can provide us with spatial information about the biological properties of the tumor provided, related with the radio-sensitivity pattern within the tumor. Performing a dose escalation based on this information is known as *dose painting* [17, 16].

Support for dose painting comes in part from the evidence of heterogeneous tumors (inter-population as well as intra-tumor volume) [109], Figure 1.2. On the basis of tumor heterogeneity, several mathematical modelling studies have shown that the optimal dose prescription for a heterogeneous radiosensitivity distribution is also heterogeneous, instead of the classical homogeneous dose prescription [25, 150, 157, 179].

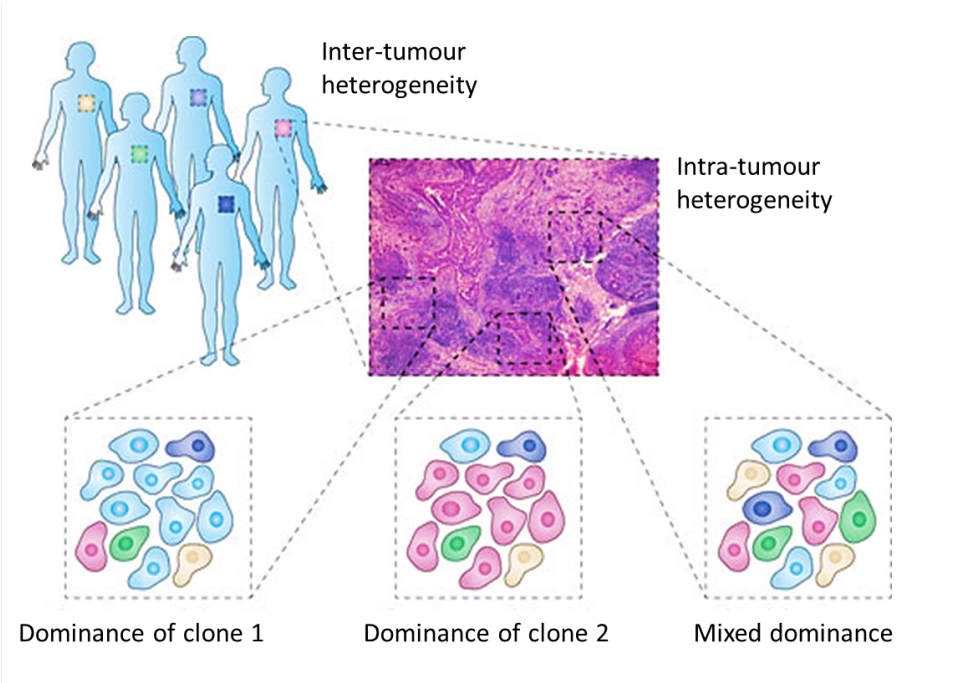


Figure 1.2: Inter-population tumor heterogeneity and intra-tumor heterogeneity. [109].

Therefore, dose painting could help us to achieve a patient specific treatment, in which the biological features of the tumor would be taken into account in the dose prescription. In addition, combining the potential of dose painting with proton therapy could enable a higher and more conformal dose escalation, which might significantly improve the therapeutic ratio for radiation treatments.

However, special care must be taken when planning proton therapy treatments, since the finite range of the protons can act like a double-edged sword: uncertainties in proton range estimation can lead to cases where the Bragg-peak is misplaced, potentially causing dramatic overdose in the surrounding healthy tissue and underdose in the tumor. Therefore, the proton range must be estimated as accurately as possible in order to achieve the full potential of proton therapy. In the case of dose painting treatments, where the heterogeneous prescription leads to steep dose gradients, proton range uncertainties may even have more pronounced effects on the final dose distribution, compared to a conventional plan with homogeneous prescription. Hence, addressing this type of uncertainties is crucial to enable the clinical implementation of *dose painted proton therapy*.

This thesis aims to combine proton therapy and dose painting to improve the outcome of radiation therapy treatments. In particular, we put especial focus on the development of tools that can be used clinically to create accurate and robust treatment plans for the most advanced state-of-the-art proton therapy and, in addition, are able to reproduce (dose-painting) heterogeneous prescriptions to achieve a patient specific therapy. Accurate dose calculation algorithms (i.e, Monte Carlo methods) have been available for a long time now, but their long computation times hamper their clinical use. The first objective of this thesis is to investigate dose calculation strategies to achieve accurate proton plans in a reduced computation time. Likewise, a wide range of robust optimization methods have been published in the literature, and they start to be available in commercial treatment planning systems (TPS). However, most of the research solutions are not open to the community, and the commercial ones are usually black-box systems that are hard to modify or interact with. In addition, these commercial solutions are very limited regarding the type of uncertainties they can simulate. The second objective of this thesis is to develop an open-source robust treatment planning system, with a versatile design so that it can be easily customized for research purposes and, in particular, that allows the simulation of other uncertainties than those currently implemented in modern TPSs. In the same line, the fact that dose painting could help to improve the therapeutic ratio for radiation therapy treatments has motivated the development of research solutions to create dose painting plans. However, the translation of these tools to the clinics is not straightforward, which is the third objective of this thesis. For this purpose, we developed an automatic workflow for dose painting, integrated in a commercial treatment planning system which includes robust optimization tools. This allows the clinicians to perform robust dose painting plans for proton therapy, with a single mouse-click.

1.1.2 Contributions

This section explains in detail the contributions from the author to address the gaps in this research field that were mentioned in the previous section. For non-expert readers, it might be useful to read the state-of-the-art (section 1.2) before going through this part.

Any radiation treatment can be divided in two phases: planning and delivery. This work is focused on the first part: treatment planning. As indicated in the title, our aim is to achieve treatment plans that are *accurate, robust and patient specific*, for the most advanced delivery technique in proton therapy, i.e., pencil beam scanning. For that, we developed 1) an in-house treatment planning system

able to achieve robust and accurate plans, called ***MIROpt*** [181], that will be available soon as open-source code for research purposes, and 2) an automatic dose painting workflow implemented in a commercial treatment planning system, which comprises all necessary steps from tumor segmentation to plan optimization. A complete description of these tools is given in Chapter 2.

Accuracy

In radiation therapy, the gold standard for accurate dose calculation are Monte Carlo methods (MC). The drawback of these methods is that usually they require very long computation times, which hampers their clinical application. Instead, most clinical treatment planning systems use analytical algorithms to compute the dose. But these algorithms may fail to predict proton range accurately in very heterogeneous media [120, 144], such as lungs [55]. In this work, we have implemented an hybrid dose calculation [142, 141] that combines the speed of analytical algorithms with the accuracy of MC to get accurate treatment plans in a reduced computation time, which enables its clinical use. For this purpose, we used a recently developed fast MC code (*MCsquare*, [147, 180]) and a research analytical algorithm (*FoCa*, [134]), both available open-source.

The in-house treatment planning system used to create the plans (*MIROpt*), as well as the dose engines, are described in Chapter 2. The hybrid dose calculation method and the obtained results are presented in Chapter 3.

Robustness

The classical way to improve robustness against treatment delivery errors consists in using margins around the target volume (PTV-based optimization). This works rather well for conventional radiotherapy, where one can assume that the dose distribution is invariant under spatial shifts (*static dose cloud approximation*). However, this hypothesis does not hold for proton therapy, due to the finite range of the protons [7]. Instead, robust optimization algorithms take the errors into account during the optimization process [48, 165, 98, 166, 128], and they are becoming the leading technique for intensity modulated proton therapy (IMPT) planning during the last years because of their excellent performance. In our case, we implemented a worst-case robust optimizer [48] that uses MC dose calculation and takes into account systematic and random setup errors, range uncertainties coming from image conversion errors, and breathing motion. Since current robust optimizers typically limit their scope to systematic setup and image conversion errors, our novelty lies in the modelling of random setup errors and breathing/organ motion. In addition, the use of hybrid methods to speed up the simulation of the

dosimetric effects caused by the latter is also investigated.

Detailed description of the robust optimizer and how the errors are simulated can be found in Chapter 2. Chapter 4 focuses on the modeling of the breathing motion and collects the results when the method is applied to a real patient case. Appendix B addresses exclusively the implementation of random setup errors.

Patient specific therapy through dose painting

The concept of dose painting sounds attractive for clinicians, since it might be a way to deliver tailored treatments that take into account patient specific biological features of the tumor, which could help to improve the therapeutic ratio. However, only a few research centers have developed the tools needed to apply this to patients using in-house software, and so far, there is no commercial solution supporting dose painting planning. In this work, we developed a complete and automatic workflow for dose painting, which supports the two main sub-modalities: dose painting by numbers (DPBN, [16]) and by contours (DPBC). The tool is integrated into a commercial treatment planning system (RayStation, RaySearch Laboratories, Sweden). It uses FDG-PET images to delineate the painted volume and to compute the dose prescription. We used this tool to test the feasibility and robustness of performing dose painting with proton pencil beam scanning. Regarding robustness, two methods were compared: classical PTV optimization versus robust optimization.

The automatic dose painting workflow is described in Chapter 2, and Appendix C collects the results of applying the workflow to a patient with head and neck cancer. The complete study about the feasibility and robustness of DPBN with scanned proton therapy is presented in Chapter 5

1.1.3 Structure of the thesis

This manuscript is organised in five main chapters and three appendices.

- The present chapter (**Chapter 1** - *Introduction*-) introduces the general context, motivation and contributions from the author, followed by a description of the state-of-the-art in proton therapy. This chapter helps the readers to have a general overview of the research field before losing themselves in the next chapters, where the author's contributions are presented in detail.

- **Chapter 2** -*Treatment planning tools*- describes the tools developed during this PhD thesis. It contains two separate sections. The first section presents the robust treatment planning system *MIROpt*. In particular, we describe the optimization libraries, the dose engines, and how the treatment delivery errors are simulated. The second section is focused on the automatic dose painting workflow and its implementation in the commercial treatment planning system RayStation (RaySearch Laboratories, Sweden). Several Conference communications have been published describing this dose painting workflow.

- **Chapter 3** -*Accurate treatment planning for proton therapy*- focuses on accurate dose calculation. It presents the implementation of the hybrid method for Monte Carlo based treatment planning, and collects the results for several phantom and real patient cases. This chapter is based on the paper: Performance of a hybrid Monte Carlo-Pencil Beam dose algorithm for proton therapy inverse planning, which is under revision in the journal *Medical Physics*.

- **Chapter 4** -*Robust treatment planning for proton therapy*- is focused on plan robustness. Since the novelty of this thesis is the simulation of breathing motion, the chapter is mainly focused on this aspect, but robustness against systematic setup and image conversion errors are also included. The use of hybrid methods to speed up the modelling of dosimetric effects caused by the breathing motion is also investigated. This chapter is based on a manuscript in preparation that will be submitted soon: Efficient implementation of 4D robust optimization for scanned proton therapy with Monte Carlo.

- **Chapter 5** -*Patient specific therapy through dose painting by numbers*- focuses on dose painting. It studies how to create robust dose painting plans for pencil beam scanning in a commercial treatment planning system. Several plan parameters are tested to find those which achieve the best plan quality. This chapter is based on the published paper: Feasibility and robustness of dose painting by numbers in proton therapy with contour-driven plan optimization.

- **Chapter 6** -*Conclusions and outlook*- concludes this thesis with a general discussion, presents the author's perspective about the evolution of this research field, and describes the role of the developed tools in current and future work.

- **Appendix A** -*Convergence conditions of the hybrid algorithm*- presents the limitations of the hybrid optimization method described in Chapter 3 through the mathematical formulation of some simplistic and illustrative examples.

- **Appendix B** -*Robustness against random setup errors: implementation and results*- describes the integration of random setup errors into a robust worst-case scenario optimization approach and collects the results for the application of the method to a clinical case.

- **Appendix C** -*Automatic dose painting module: implementation and results*- presents the results obtained after applying the automatic dose painting workflow to plan a dose painting by numbers treatment with scanned proton therapy for a patient with head and neck cancer.

1.2 State-of-the-art

The aim of this section is to provide the reader with a basic knowledge about the main concepts treated in this thesis, which comprises the state-of-the-art for proton therapy and dose painting. In both topics, the attention is focused in treatment planning and robustness against errors during treatment delivery, since this constitutes the central subject of this thesis.

1.2.1 Proton therapy

The idea of treating cancer patients with protons was first proposed by R. Wilson in 1946 [184]. The first patients were already treated in the 50's, in research centers at Uppsala and Berkeley. But due to the lack of technology and the high cost of particle accelerators, the first dedicated centers for proton therapy were only built in the early 90's. Since then, the number of proton therapy centers has experienced an exponential increase, with more than 60 centers worldiwide in 2017, and a few more under construction [131]. This exponential increase has been possible thanks to the intensive research performed during the last years, which has lead to progress in three main fields: 1) *treatment delivery*, with more compact solutions that have significantly decreased the cost of the proton therapy facilities, and enabled a more conformal and targeted therapy through modern delivery techniques as proton pencil beam scanning; 2) *treatment planning*, with faster and more accurate dose calculation, and increased robustness against treatment uncertainties; and 3) *quality assurance*, with advanced imaging tools that

allows for in-vivo proton range verification, among other tools that improve direct verification of the delivered dose, which increases the available information about the execution of the treatment and help us to gain more confidence and control on the technique itself. The following sections focus on the second point: treatment planning. But an extensive review of all these aspects and others is presented elsewhere [119, 122].

Treatment delivery: pencil beam scanning

In order to reach the tumor depth, protons need to be accelerated to increase their energy typically from 70 to 230 MeV, which corresponds approximately to a depth in water from 4 to 33 cm. Protons are accelerated inside a cyclotron (or synchrotron) and then transported into the treatment room through the beam line. A single mono-energetic proton beam (Bragg peak, Figure 1.1) is too narrow to cover a tumor three-dimensionally. Therefore, the beam needs to be spread somehow to create a uniform or other desired dose distribution both laterally (perpendicular to the beam), as well as in depth (beam direction). This is the role of the treatment head (or nozzle), where the proton beam is shaped to have a clinical 3D dose distribution.

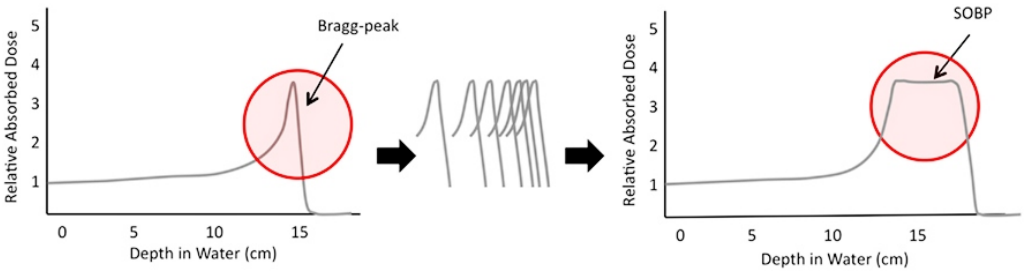


Figure 1.3: To cover a tumor in depth with an homogeneous dose, multiple Bragg peaks of different energies are superposed and lead to the so-called Spread out Bragg peak (SOBP). [62].

To cover the tumor in depth, several Bragg peaks of different energies are superposed to form the so-called Spread out Bragg peak (SOBP), Figure 1.3. To cover the tumor laterally, two methods are available: *passive scattering* (PS), where heavy materials scatter the proton beam to make it diverge to nominal dimensions large enough for most applications, and *pencil beam scanning* (PBS), where magnetic fields steer the proton beam over the desired area. The first proton centers were equipped with PS systems, but nowadays, nearly all new centers are provided with PBS [131], which is becoming the preferred modality. This is, in part, due to the fact that PS systems need patient specific beam compensators

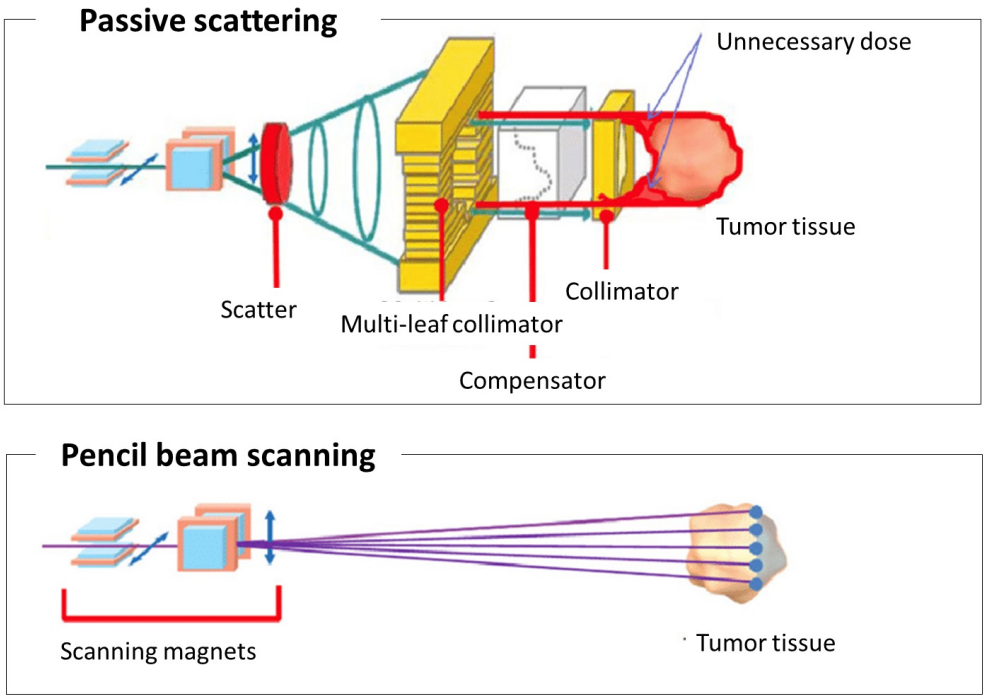


Figure 1.4: Proton delivery techniques: passive scattering (upper figure) versus pencil beam scanning (lower figure). Adapted from [37].

and collimators (Figure 1.4) to adapt the beam to the tumor shape, while in PBS, the tumor is scanned by the dynamic pencil beam, which avoids the tedious task of building patient specific beam shaping devices. In addition, PBS enables to combine individual Bragg-peaks from different beam angles to achieve the desired dose distribution. This is known as *intensity modulated proton therapy* (IMPT) [101, 85], which has demonstrated its superiority against PS and conventional radiotherapy in terms of dose conformity, healthy tissue sparing, and degrees of freedom to shape the dose.

In PBS, the tumor is divided in *energy layers*, in order to build the SOBP. These layers are scanned one after the other in the beam direction, starting at the most distant layer, i.e., the deepest tumor region. The spacing between energy layers is typically around 5 mm-WED [1]. The number of layers would depend on the tumor volume and the spacing between them, but it generally ranges from 10 to

[1] WED = Water Equivalent Depth

50 layers. The way to change the energy of the beam would depend on the accelerator system. In a synchrotron, the protons are accelerated until they achieve the desired energy. Therefore, the energy of the beam is directly the output energy given by the synchrotron. In a cyclotron, the protons are always accelerated to a maximum energy ($\approx$ 230 MeV), and modified afterwards with a degrader system, which allows decreasing the energy until a minimum value ($\approx$ 70-100 MeV). For energies below this minimum threshold, a piece of material (*range shifter*) is inserted at a certain point of the beam path. This device enables to decrease the energy of the beam down to a few MeV, which is very convenient for superficial tumor locations, such us tumors in head and neck close to the skin. However, any material in the beam path will scatter the beam, increasing the beam size.

To cover the tumor laterally (direction perpendicular to the beam), each energy layer is designed as a discrete scanning grid of dose delivery positions (*spots*), Figure 1.5. Usually, the dose delivery system is switched-off while the beam is being re-positioned between spots, which is known as *discrete scanning*. *Continuous or raster scanning*, where the beam is paused only between energy layers, has also been proposed [135], although this is so far used for research purposes mainly. The space between spots is typically around 3-5 mm, which lead to a total number of spots, including all energy layers, that can reach the several thousand for regular tumor sizes.

The time spent in treatment delivery depends on the machine itself. For IBA (Ion Beam Applications S.A., Louvain-La-Neuve, Belgium) systems, re-positioning of the beam to a different spot position takes usually less than 1 ms, while switching the energy layer requires around 1-2 s. The time to deliver a single spot depends on its intensity or *weight*, which is a magnitude proportional to the number of delivered protons, but generally it is around a few ms (2–3 ms). The most time-consuming part is to change the beam angle, which can take a few minutes (1-5 min). This is possible thanks to the *gantry*, a large rotating structure that allows the beam to spin 360 degrees around the patient.

The size of the proton beam that reaches the tumor (*spot size*) can be modelled as a Gaussian distribution with standard deviation σ. The choice of this parameter would influence the final dose distribution and robustness [182]. Small spot sizes can achieve steeper dose gradients, but they would be more sensible to treatment uncertainties. Most PBS systems use a spot size with σ around 3-5 mm at patient entrance [182].

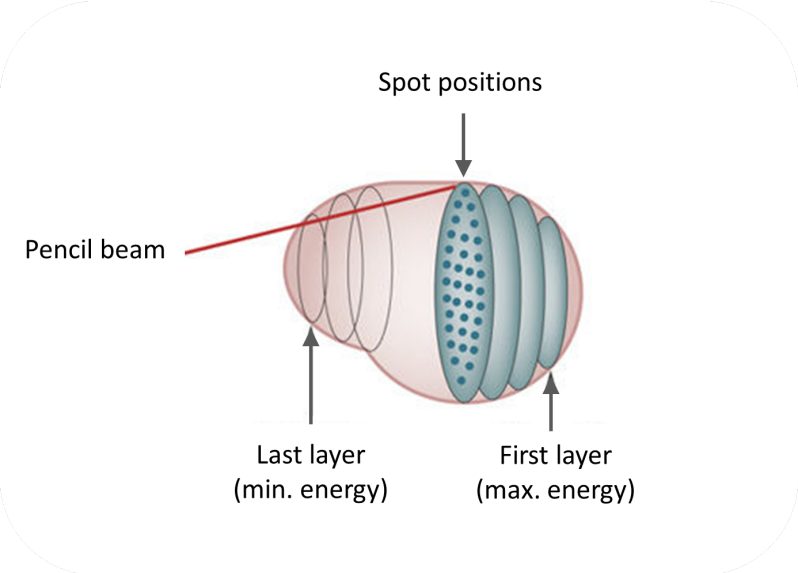


Figure 1.5: Illustration of the scanning grid energy layers and spot positions for PBS. Adapted from [36].

Treatment planning

Treatment planning for any radiation therapy modality comprises two main tasks: 1) *dose calculation* and 2) *plan optimization*. The latter consists in finding the optimal machine parameters that gives the best dose distribution. In the case of PBS plans, the variable subject to optimization are the spot weights. Therefore, both tasks are inter-connected, since one needs information about the dose contribution from each spot during plan optimization, in order to adjust the weight values to obtain a clinical dose distribution. In the following, we describe the main concepts for the state-of-the-art dose calculation and plan optimization algorithms for PBS treatments, putting especial focus on the strategies used in this thesis, i.e., Monte Carlo dose calculation and robust plan optimization.

Dose calculation

The dose for a given patient is computed by using information from a computed tomography (CT) image, acquired before treatment. This image contains three dimensional information about the radiation attenuation coefficient on each voxel, expressed in Hounsfield Units (HU). In order to compute the dose for a proton therapy plan, one must have information about the proton stopping powers (SPR). The relationship between HU and SPR is typically established by a

scanner-dependent calibration curve obtained empirically, from CT images of a phantom with several tissue-equivalent materials. Such curve allows us to extract the SPR on each voxel, and additional information about the tissue composition [137]. This information is then used to compute the dose.

The aim of any dose calculation algorithm is to provide accurate dose distributions in a computation time that is suitable for clinical use. The most simple and fast methods compute the dose analytically by means of *pencil beam* (PB) models [63]. In this model, elementary dose distributions of a proton pencil beam (also called *dose kernel*), computed in a water phantom, are modified to take into account the tissue density and the radiological depth, and integrated over the patient surface to get the dose distribution. The energy released in the medium is then quantified as dose to water. The drawback of PB algorithms is that they are limited by the lateral extent of the pencil beam itself, which makes them insensitive to lateral heterogeneities [120, 144]. A correct modelling of this phenomenon is crucial for heterogeneous tumor locations, such as lungs [55]. More accurate models have implemented multiple raytracing approaches that divide a single pencil beam into sub-spots, in order to better take into account the lateral heterogeneities [84, 145, 1]. However, for a complete and accurate modelling of the dose distribution in the patient, including a correct simulation of lateral heterogeneities, Monte Carlo methods are considered as the gold standard. An illustration of the differences between MC and PB algorithms in very heterogeneous media is presented in Figure 1.6.

Monte Carlo (MC) methods are based on the use of random number generators to sample the physical equations for radiation transport, being able to account precisely for all particle interactions along the beam path. In order to achieve a statistically acceptable result, MC methods need to simulate a large number of particle histories. The drawback of MC is that it requires long computation times, which hampers their application into clinical routine. However, the improving in computing power together with the development of dedicated MC dose engines for PBS [147, 127, 193, 83, 74, 40], is reducing the calculation time to an acceptable level for clinical application. These new fast MC codes exploit the potential of novel computing hardware, such as graphical processor units (GPU) [74] or Xeon Phi architectures [147], to achieve a full dose computation in the order of seconds or few minutes. In addition, most of them use simplified physics models, avoiding the unnecessary simulation of those interactions that have negligible contributions in the dose distribution [147, 40].

The increase in number and availability of these dedicated and fast MC dose en-

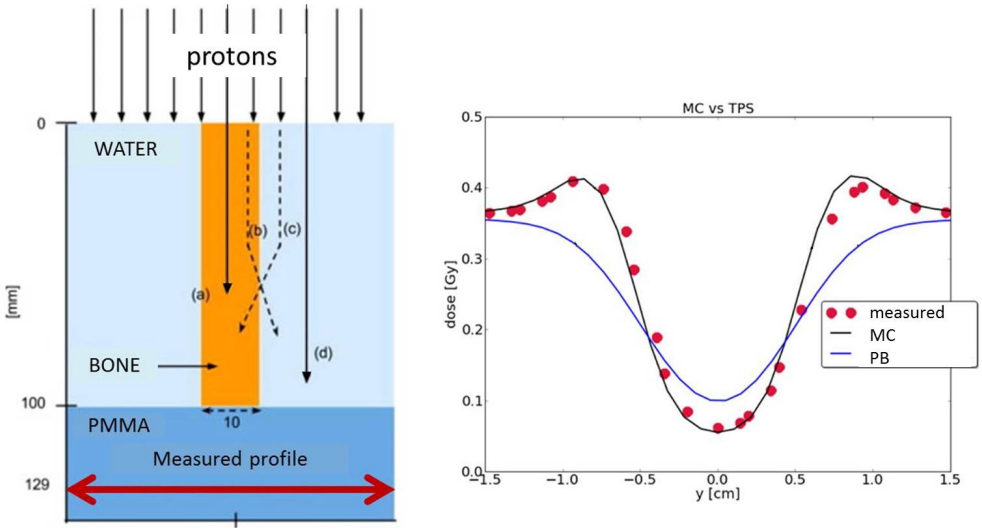


Figure 1.6: Illustration of the difference between Monte Carlo (MC) and analytical pencil beam (PB) algorithms in heterogeneous media. The plot on the left represents the phantom used for the measurements, which is a water tank with a bone slab inserted in the middle, and a layer of PMMA on the bottom. The dose is measured along the red line in the PMMA layer. The plot on the right collects the results from the measurements (red points), and the dose computed by PB (blue solid line) and MC (black solid line). The mismatch between the PB and the measured data is due to the limitations of this algorithm to model the bone-water interface (trayectories (b) and (c)). In contrast, MC is able to accurately simualte the interactions in this interface. Courtesy of J. Sorriaux [144].

gines motivates the development of MC based treatment planning systems (TPS) for PBS [105, 90, 91]. However, the TPS needs to have information about the dose contribution from all spots with a good statistical precision (See *Plan optimization* section), which implies the simulation of many particles per spot ($\approx 10^5$). Therefore, taking into account that the number of spots is typically around several thousand, MC based plan optimization requires to simulate a very large total number of particles. The most recent research in this field focuses on the development of strategies that allows to accomplish this task in the most efficient way possible [90, 91]. Chapter 3 discusses all possible ways of including MC in PBS treatment planning, and collects the results for the solution proposed in this thesis, which consists in a hybrid dose calculation method [142, 141] that combines PB and MC to get the best trade-off between speed, efficiency and accuracy .

Plan optimization

Traditionally, treatment planning was a trial and error process in which experienced medical physicists computed the dose distribution for a set of machine parameters, and repeated this task until finding the optimal configuration. This is known as *forward planning*, but nowadays, it has given the way to *inverse planning* [116], which uses mathematical optimization techniques to find the machine parameters that satisfy the dose objectives set by the medical physicist.

In inverse planning for PBS, the parameter subject to optimization is the spot weight (w). Previous to optimization, the user has to define the number of beam directions (usually from 1 to 4) and the scanning grid parameters (distance between spots and between energy layers). This information is used to place the spots to cover the target and to compute their dose contribution. The first step is to find out the energy range needed to cover the target in depth. This is done by performing a raytracing to estimate the water equivalent path length (WEPL) for each voxel, which is converted in energy units (MeV). Then, the energy interval is discretized in energy layers according to the layer spacing set by the user. The second step is to define the spot positions for each layer, which is usually done by following a square or hexagonal pattern to keep a constant space between spots. Once the spots positions are defined, the dose contributions for all spots at unitary fluence ($w = 1$) are computed and stored in the so-called *dose influence matrix*, P. This matrix will be used during optimization to compute the total dose distribution d at each iteration, for a given vector of spot weights,

$$d_i = \sum_j P_{ij} \cdot w_j \tag{1.1}$$

where i and j represent the voxel and spot indices, respectively. In order to guide the optimizer towards final spot weight values that give the desired clinical dose distribution, the medical physicist introduces several objectives (and/or constraints) on the dose itself. These objectives penalize deviations of the dose from a reference value (ξ) in a region of interest (ROI), and they are closely related with the clinical evaluation criteria. Some of the most common dose objectives use convex functions, such as the normalized sum of squared dose deviations for all voxels, to penalize the dose below ($minimum\ dose,\ f_{min}$) or above ($maximum\ dose,\ f_{max}$) the reference ξ [116]:

$$f_{min}(w) = \frac{1}{N_r} \sum_{i=1}^{N_r} ((\max\{0, \xi_r - d_i(w)\})^2) \tag{1.2}$$

$$f_{max}(w) = \frac{1}{N_r}\sum_{i=1}^{N_r}((\max\{0, d_i(w) - \xi_r\})^2) \tag{1.3}$$

where N_r represents the total number of voxels in ROI r, and d_i is the dose to voxel i given by equation 1.1. More sophisticated dose objectives include non-convex functions based on the dose given to a certain fraction of the ROI volume ($minimum$ or $maximum\ dose - volume$) [60], or based on biological considerations (EUD) [188]. A typical clinical example of the minimum dose-volume requires the 95 % of the target volume to receive at least 95 % of the prescribed dose ξ. Maximum dose-volume objectives penalize the overdose in a certain fraction of the total volume. Regarding the biologically based dose objectives, the term EUD stands for equivalent uniform dose, and it is defined as the uniform dose level that would have the same biological effect as the actual ROI dose distribution, under some biological model or fit to measured data.

In order to formulate the full optimization problem, one or several of the aforementioned dose objective functions $f(d(w))$ are associated to each ROI. Then, the most common approach is to weight each objective by a relative importance factor α, and add them up to build a composite of n functions that will be minimized by using mathematical optimization algorithms.

$$\begin{aligned} \underset{w}{\text{minimize}} \quad & \sum_{l=1}^{n} \alpha_l f_l(d(w)) \\ \text{subject to} \quad & w \geq 0 \end{aligned} \tag{1.4}$$

In this thesis, the only hard constraint is to have positive or zero weight values ($w \geq 0$), but other hard constraints based on the dose can also be used. As mentioned, the number of spots can amount up to several thousand, which leads to large scale optimization problems that need to be handled carefully. For this, one can build dedicated optimization algorithms or use existent multi-purpose libraries such as IPOPT [68], CPLEX [154] or SNOPT [148], among many others. In this thesis, we chose to use IPOPT because it is available open-source, and its MATLAB® interface makes it very convenient for research use. Recently, other groups have also chosen IPOPT libraries to build another open-source treatment planning system [183]. More information about the IPOPT settings used in our in-house treatment planning system -MIROpt- can be found in Chapter 2.

Concerning more practical details about the plan optimization for PBS treatments, two main approaches are distinguished: *single field optimization* (SFO) and *multi field optimization* (MFO). The first one requires to deliver a uniform

dose from each of the beam directions involved in the plan, which is also known as single field uniform dose (SFUD). In contrast, the second approach, popularly known as intensity modulated proton therapy (IMPT), optimizes simultaneously the weights of all spots, coming from different beam directions. This requires the total dose to be uniform, but not the individual beam contributions. Therefore, IMPT enables more degrees of freedom in plan optimization, which leads to more conformal dose distributions, but also a higher sensitivity to treatment uncertainties. This has motivated the development of robust optimization strategies that take into account the uncertainties during the optimization process. These techniques are presented in detail in the next section.

Treatment uncertainties and robust optimization

Uncertainties in conventional radiotherapy treatments have been divided in two groups: *treatment preparation* or *systematic errors*, and *treatment execution* or *random errors* [174, 173]. The treatment is given during a period of several days or weeks (depending on the modality), and the delivery is divided in *fractions*, with usually one fraction per day. Systematic errors will influence all treatment fractions in the same way. A typical example includes the uncertainty in the delineation of the target and organs in the planning CT. Random errors account for day-to-day variations, and they will influence each fraction differently, since their magnitude can change over the course of the treatment. Breathing induced organ motion is a clear example of random errors. Another important source of error is the patient positioning or *setup*, which has both systematic and random components.

In proton therapy, the finite range of the protons represents an extra challenge on top of the aforementioned sources of error. Uncertainties in proton range occur when the proton does not stop at the estimated depth, and they can have both systematic and random behaviour, depending on how they are originated. They are mainly caused by (a) anatomical modifications, (b) dose calculation approximations, (c) biological effects, (d) patient setup, (e) image conversion errors due to the presence of artefacts in the CT image and the conversion of HU to stopping powers, and (f) organ motion. Anatomical modifications over time (a) require the implementation of adaptive therapy [189], where the treatment plan is modified to take into account new information about the patient anatomy, such as daily CT images or in-vivo range verification. Uncertainties related to dose calculation approximations (b) can be reduced significantly by using Monte Carlo dose algorithms [120]. As mentioned in the previous section, this is especially crucial for heterogeneous media, where local lateral inhomogeneities will reduce

the accuracy of analytical algorithms. An extended analysis of dose calculation uncertainties and how MC could potentially reduce them is presented elsewhere [120]. Regarding the biological considerations (c), there is a consensus to use a constant value for the relative biological effectiveness (RBE $= 1.1$) of protons with respect to photons. However, it has been demonstrated that this value depends on factors like the biological endpoint, dose or linear energy transfer (LET). The latter quantifies the amount of energy transfer per unit distance. For a given dose and biological endpoint, elevated proton LET values are associated with an increase in RBE for most tissues. In particular, LET increases at the distal fall-off of the Bragg peak, leading to a shift in the biologically effective range of 1-2 mm [121]. Therefore, including the LET information into treatment plan optimization [51] can help to put the high LET regions away from critical structures, minimizing the impact of biological range uncertainties. The remaining sources of errors (d-f) can be also addressed during treatment planning, and the state-of-the-art strategies to deal with them are presented below.

The classical way to account for setup errors (d), and other geometric uncertainties, is to add a safety margin around the clinical target volume (CTV) to build the planning target volume (PTV), following well established margin recipes [174]. This method relies on the *static dose cloud approximation*, which assumes that the dose distribution is invariant under displacements of the target. This hypothesis is valid for conventional photon therapy, but it does not hold for proton therapy due to the finite range of the protons (Figure 1.7). Numerous studies has shown that PTV-based optimization is not appropriate for proton therapy. This is especially true for IMPT treatments, due to the high sensitivity of this modality to range uncertainties [7, 99, 100]. In IMPT, the doses from different beam directions are typically inhomogeneous and any variation can lead to a mismatch of the individual beam contributions, causing dramatic perturbations to the total dose distribution. This motivates the use of more sophisticated methods than PTV-based optimization, such as *robust optimization* [48, 98, 165], which takes into account the uncertainties during the optimization process, and has shown an excellent performance for IMPT plans.

In robust optimization, the planning is done directly on the CTV, and the dose depends on a set of uncertain parameters λ, $d(w, \lambda)$. Several approaches have been proposed, from which one can distinguish two main types: *probabilistic optimization* [165], which assigns a probability distribution $P(\lambda)$ to the set of uncertain parameter and optimizes the expected value of the objective function; and *worst-case optimization* [48, 98], where the uncertain parameters vary within a well defined interval, and the optimization is only done for those values that give the

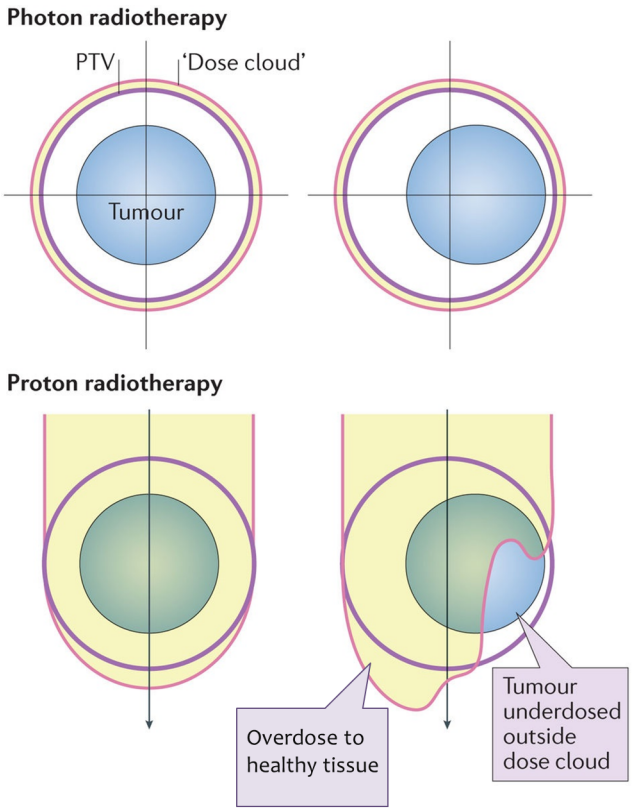


Figure 1.7: Conventional radiotherapy assumes that the dose distribution is invariant under displacements of the target (*static dose cloud approximation*), which enables to achieve robustness against uncertainties by simply adding a safety margin around the target (PTV). In proton therapy, this hypothesis does not hold due to the finite range of the protons, which leads to dramatic changes in the dose when the target moves. Adapted from [12].

highest objective function (worst-case). The latter can have some variants [46], such as worst-case dose distribution [128] or worst-case scenario (minimax) [48], among others. The worst-case dose distribution is a non-physical dose chart that contains, for each voxel, the worst of all possible values. Indeed, for every target voxel, the minimum dose of all possible scenarios is considered; while the rest of voxels takes the maximum. In contrast, the worst-case scenario approach considers only physically realizable scenarios, i.e., evaluates the full objective function in a number of scenarios and minimizes the worst of them. In this thesis, a worst-case scenario approach was used to create the plans, since it seems to be the preferred technique in the community, and it is also the one implemented in commercially available solutions, like RayStation (RaySearch Laboratories, Sweden). Chapter 2 gives detailed information about our robust treatment planning system, MIROpt.

So far, robust optimization has been mainly used to account for setup (d) and image conversion errors (e) [48, 98, 166, 165]. The reported results are very promising, since the plans created with robust optimization achieve robust target coverage without compromising the dose to the healthy tissue. This opens the way to the clinical implementation of these techniques, and to extend their applicability to other types of uncertainties.

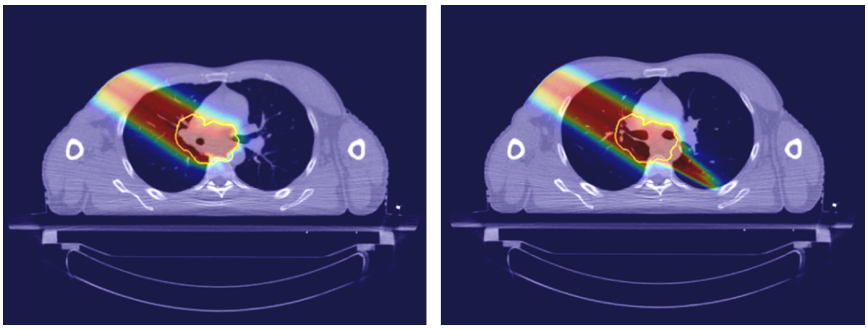


Figure 1.8: Dose distributions corresponding to a lung case in two different phases of the breathing cycle. The changes in density caused by the breathing motion lead to significant overdose of the healthy tissue. The tumor volume is contoured in yellow. Courtesy of G. Janssens (Ion Beam Applications SA)

Tumor and organ motion (f) represents another extra challenge for proton therapy treatments, since important changes in density along the beam path can lead to dramatic consequences for the patient, as illustrated in Figure 1.8. Again, the classical way to take into account tumor motion in radiotherapy, caused mainly by breathing, involves the use of safety margins. In this case, 4DCT images are used to create an internal target volume (ITV) that encloses the target while it moves over all respiratory phases. A more intelligent approach that reduces the exposed volume while maintaining target coverage, is to irradiate the geometric time-weighted mean tumor position (mid-position, MidP) [187, 186]. However, neither ITV nor MidP margins are robust to changes in the radiological depth of the tumor. Therefore, a common solution is to use range adapted ITVs for each beam direction [124, 81]. But this strategy only works if each field is optimized independently (SFUD), at the expense of conformity and lung exposure due to the large irradiated volumes; and extending this range adapted margins to IMPT still remains difficult [53]. More sophisticated methods like gating or tracking have also been proposed [54, 52, 38]. But the latter requires complex 4D delivery to synchronize the plan to the actual patient motion during treatment, which is difficult to implement clinically. In this thesis, robust worst-case optimization is

used to account for breathing motion, creating plans that can be delivered without the need of synchronization. Our approach is similar to the one published by Liu et al. [97], but we use 4D Monte Carlo dose calculation in order to have a correct modeling of tissue heterogeneities, which is crucial for lung tumors. In addition, we investigate the use of hybrid algorithms [142, 140] to speed up the 4D dose calculation process. The details about the complete solution proposed in this thesis is presented in Chapter 4.

It is important to note that besides the changes in the radiological depth of the tumor while it moves, there is another issue that needs to be addressed for PBS treatments, which regards the interference between the beam dynamic delivery and the motion itself, so-called *interplay effect*. In this thesis, interplay effects are disregarded, but this problem has been studied extensively and several publications have demonstrated that its dosimetric effects average out over the course of a conventionally fractionated treatment (i. e., 30 fractions) [92, 35], yielding a conformal overall dose distribution at the end of the treatment. For each individual fraction, interplay effects can be significantly reduced with simple techniques such as repainting [191, 75].

1.2.2 Personalized therapy through dose painting

Besides the accuracy and high dose conformality achieved with the most advanced radiation therapy delivery techniques, loco-regional recurrence still appears in important part of the treatments due to unsuccessful tumor local control. For instance, in patients with locally advanced lung cancer, 70% of the treatments fails because of that reason [8, 106]. The most straightforward way to improve the tumor control probability (TCP) would be to increase the dose. However, this would have the consequence of an increase in the normal tissue complications (Figure 1.9). A more intelligent way of improving TCP would be to increase the dose heterogeneously, instead of using the conventional homogeneous prescription. Thus, the more resistant parts of the tumor would receive a higher dose, while the usual dose level could be kept (or reduced) towards the more sensitive areas, which would enable to keep a low NTCP. Functional and molecular images can provide information to find these radio-resistant areas in the tumor, which are related with its biological characteristics. Indeed, tumor burden, cell density, hypoxia or proliferation, among others, are some of the features associated with a higher radioresistance, from which one could extract information by using functional imaging such as positron emission tomography (PET) or magnetic resonance (MR). This leads to the concept of *dose painting*, which consists in prescribing an heterogeneous dose according to spatial information about the biological properties of the tumor, given by functional or molecular images [93, 16].

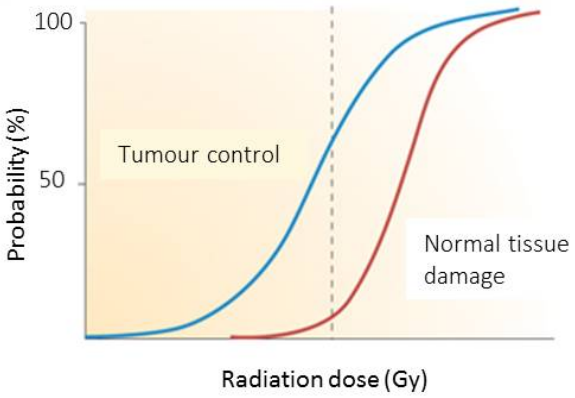


Figure 1.9: The tumor control probability (TCP, blue line) increases with increasing radiation dose. However, this also makes the normal tissue complications probability (NTCP, red line) to increase as a consequence, since the radiation passes unavoidably through part of the healthy tissue. [9].

The hypothesis of dose painting is supported by the fact that tumors have an heterogeneous composition, which is present at both intra-tumor and inter-population level [109]. This calls for a patient specific prescription, that could be achieved by means of dose painting. In addition, numerous mathematical studies using TCP and NTCP models have demonstrated that the optimal dose prescription for an heterogeneous radiosensitivity pattern is also heterogeneous, which again goes against the classical homogenous prescription [25, 150, 157, 179]. This makes the dose painting concept very appealing to the physicians, but clinical evidence of the gain that dose painting could bring for the patients is still needed [17, 72]. For this purpose, the work of the physicist is to provide all the necessary tools to plan and deliver dose painting treatments. This section describes the main concepts of dose painting from the physicist's point of view, paying special attention to treatment planning, which is the central subject of this thesis (See chapters 2 and 5).

Dose painting modalities: by contours VS by numbers

Two different approaches for dose painting have been considered in the literature: 1) *dose painting by contours (DPBC)*, which defines several sub-volumes in the tumor and prescribes a uniform dose to each of them, also known as *selective sub-volume boosting* if only two dose levels are used [93]; and 2) *dose painting by numbers (DPBN)* [16], which adjusts the dose prescription in each voxel in-

dependently, with respect to the corresponding voxel intensity in the registered functional image (Figure 1.10). Assuming that the biological properties of a tumor would be better modelled by a range of values extracted directly from the image information, rather than a few discrete dose levels, then DPBN would be a better option to target the molecular heterogeneities in the tumor.

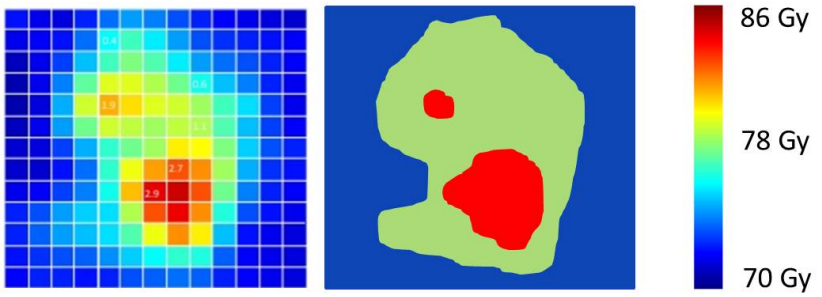


Figure 1.10: Illustration of the two different strategies used for dose painting: 1) dose painting by numbers (left), where each voxel is associated a specific dose ranging from 70 to 86 Gy, and 2) dose painting by contours (right), where two sub-volumes receive a boost of dose (green 78Gy, red 86Gy) on top of the tumor bed represented by the blue background (70Gy).

Imaging the biology of the tumor

Although some MRI techniques, such as diffusion weighted MRI or dynamic contrast-enhanced MRI, have been proposed to be used as functional image for dose painting [170], PET is by far the most used modality for this purpose [159, 17, 72].

The principle of a PET scan is to detect pairs of gamma rays coming from the positron-electron annihilation process inside the body. To that end, a positron emitting (β^+ decay) radionuclide is introduced into the body (usually injected into the blood circulation), on a biologically active molecule. As the radionuclide decays, it emits a positron that will travel in the tissue until it loses its kinetic energy and interacts with an electron. This causes their annihilation, with the consequent emission of two gamma photons in opposite directions, that will be detected by a array of scintillators. This signal is then used to reconstruct a three-dimensional image of the radionuclide concentration within the body.

The combination of the radionuclide and the biologically active molecule is often referred to as *tracer*. Different tracers can give information about different

features of the tumor. The most used one is the well-known fluorodeoxyglucose or FDG, which gives information about the tumor burden, but the list of possible tracers is very long [102]. Other examples are FLT, a surrogate of tumor cell proliferation, and hypoxia related tracers such as FMiso, Cu-ATSM or FAZA.

The main reasons for the FDG popularity are its wide availability and the accumulated clinical experience about its use, which includes the creation of standardised protocols. In addition, FDG has been suggested as a rational choice for dose painting [5]. This is supported by some studies reporting the relation between the relapse regions in the tumor with the pretreatment high FDG uptake areas [6, 4] and the stability of the uptake patterns during the treatment [3].

Dose prescription, planning and delivery

The dose prescriptions used nowadays are site-specific and have been derived from years of clinical experience, together with clinical trials where the dose was pushed as higher as possible, taking into account the toxicity and tolerance of the healthy tissue. In the new paradigm of dose painting, especially in DPBN, the optimal relationship between the information provided by the image and the dose prescription is still unknown , and probably would depend on factors like the type of tracer. Different options have been proposed in the literature [21], some of them trying to integrate TCP/NTCP models in the prescription function [158], others trying to keep the same integral dose as the conventional uniform prescription [43]. However, in the absence of evidence from any particular formulation, probably the most reasonable way of computing the prescription is to use a simple linear relationship between the tracer uptake and the dose. This has been one of the most used approach [175, 17, 34] and the one also chosen in this thesis (Chapters 2 and 5).

Once the prescription is defined, the next step is to communicate it somehow to the treatment planning system (TPS), so that the optimiser can find the best machine parameters to reproduce such prescription. Current TPSs use contour-driven optimization, that is, contours on each slice of the CT are used to define a 3D structure for which a certain dose objective is associated. Therefore, treatment planning for DBPC or sub-volume boosting can be easily done in any commercial TPS, by associating a dose objective with the desired prescription to each sub-volume. However, this framework is not appropriate for DPBN, since voxels cannot be constrained independently. To cope with this problem, some groups have developed in-house TPSs allowing for voxel-wise optimization [132, 160, 43, 42]. Other groups have investigated the use of several contours, with non-uniform dose objectives, to guide the optimiser towards the voxel-wise prescription [33, 86], which

has shown excellent results .

Following any of the aforementioned approaches for dose prescription and treatment planning, numerous studies have demonstrated the feasibility of performing dose painting with the current delivery techniques for intensity modulated photon therapy, achieving a very good target conformality. Even the agreement of the dose computed by the TPS and the measured one using alanine/EPR dosimetry have been verified, for dose painting plans in a thorax phantom [82]. Some groups have also compared the performance of photons versus protons, reporting a similar conformality for both techniques, but with significant reduced dose in the normal tissue for the protons plans [160, 43]. This motivates to further investigate the combination of dose painting and proton therapy, which is the topic of Chapter 5.

Uncertainties

There are several sources of uncertainties that affect all steps of the preparation and execution of a dose painting treatment.

On one hand, there are all those sources of errors that are present in any conventional radiotherapy treatment, such as variations in patient setup or organ motion, usually referred to as *geometric uncertainties* [174, 173]. These type of uncertainties have been reviewed in section 1.2.1, including those specific to proton therapy treatments, which arise as consequence of the finite range of the protons. For photon-based delivery modalities, previous studies have demonstrated that a minimal dose can be ensured in the presence of geometric uncertainties inside the dose painting volume by either using robust optimization [185] or incorporating directly the errors into a *robust prescription* [152]. The latter strategy extends the classical PTV margin approach for heterogeneous prescriptions. For scanned proton therapy, a proper analysis of the robustness against geometric errors is still missing, which is addressed in Chapter 5.

On the other hand, there are those uncertainties related to the dose painting concept and the capacity of targeting specific biological features in the tumor. This includes both biological factors (e.g. patient metabolism, tracer characteristics) and issues related to the molecular imaging (e.g., scanner calibration, reconstruction, spatial resolution). Unfortunately, this type of uncertainties cannot be included during treatment plan optimization, in contrast to the aforementioned geometric uncertainties. But some of them could be minimized or controlled through a properly designed workflow and quality assurance of the image acquisition procedure. A review of this topic has been published by Jeraj et al. [72]

Plan evaluation

For uniform prescriptions, treatment plan quality is generally assessed using dose volume histograms (DVH) and typical metrics like D_{95}, D_5 and D_{mean}. However, the spatial information is lost when creating a DVH, and therefore, such metrics no longer apply to non-uniform prescriptions as the ones used for dose painting. Instead, strategies to quantify the agreement between the computed dose and the prescription on each voxel must be applied. Vanderstraeten et al. [175] proposed the so-called *quality volume histogram* or QVH, which is the decumulative histogram of the ratio $Q_i = D_i^{planned}/D_i^{prescribed}$, where $D_i^{planned}$ and $D_i^{prescribed}$ are the planned and prescribed doses in the *i*-th voxel. The ideal QVH is a descending step function, where the relative volume V equals 100% for Q_i smaller than 1 and 0% for Q_i larger than 1. The same group also proposed the *quality factor* QF ($QF = \frac{1}{n}\Sigma_i |Q_i - 1|$). a quantity that averages the deviation of the ratios Q_i in the volume and therefore reflects the spread of the QVH curve around 1.

Witte et al. [185] introduced an alternative to the QVH: the ΔVH, defined as the volume histogram of the difference between the 3D planned and prescribed dose distributions. For the case of a uniform prescription, the ΔVH is identical to the conventional DVH, but with its origin moved to the prescription dose value.

Chapter 2

Treatment planning tools

Improving treatment planning of scanned proton therapy is the central topic of this thesis. For that purpose, we investigated two research lines that will be combined at the end with the objective of achieving a better therapeutic ratio for radiation treatments. The first research line focuses on improving the accuracy and robustness of current treatment planning systems, which results in the development of the software *MIROpt*. The second research line focuses on biologically based treatment planning, including information given by functional image, i.e., dose painting. In this case, our efforts are concentrated on developing an automatic workflow that enables the clinician to easily create dose painting plans with a commercial treatment planning system. These tools can be eventually combined to achieve a targeted therapy with reduced side-effects, thanks to the advantages of both proton therapy and dose painting strategies.

2.1 Robust plan optimization with MIROpt

Commercial treatment planning systems are usually black-box like systems that are difficult to modify or interact with. Some of them (i.e., RayStation, Eclipse) start to provide tools that can be very useful for research use, such as scripting environments or dedicated research interfaces. However, their internal structure is still mostly inaccessible, which represents a limitation for research users. This motivated us to develop our own treatment planning software: *MIROpt*, which was then used to test new ways of including treatment uncertainties during planning, as well as strategies to speed up the treatment planning process. Detailed information about the application of MIROpt in these areas can be found in Chapter 3 and 4.

MIROpt is conceived as a research tool for treatment planning of scanned proton therapy, with a versatile design that makes it very easy to modify or adapt for different applications. It is entirely written in MATLAB and the code will be released open-source at the end of 2017. This will include a stand-alone version and a graphical interface. In the following, we describe the main features of MIROpt software.

2.1.1 Treatment design

As every treatment planning system, MIROpt counts with a number of parameters that need to be defined by the user in order to design the treatment for a specific patient. In pencil beam scanning, those parameters will determine the dose delivery positions, i. e., the spots. Once the treatment plan is completed,

all parameters are encoded and exported in DICOM [66] format.

The volume used to place the spots, hereinafter referred to as optimization target volume (OTV), is computed as the target volume expanded by a user-defined margin that accounts for the dose fall-off (or penumbra) and treatment uncertainties. It is important to note that this margin has nothing to do with plan robustness, its only aim is to create a volume to place enough spots to cover the target in all possible situations. Once the OTV is computed, the first parameter to start with is the number of beam directions, which are defined by a gantry and couch angles, expressed in degrees (0-360^o). MIROpt enables to set as many beams as desired, but usually no more than four beams are needed to create a plan with PBS. For each beam direction, a raytracing is performed in the OTV in order to compute the maximum and minimum energy needed to cover the target in depth. This gives us information about the equivalent proton range in water (r) for each voxel, which can be easily converted into incident energy (E_0) by applying the following equation:

$$E_0 = e^{a+b(log(r))+c(log(r))^2+d(log(r))^3}, \tag{2.1}$$

where a, b, c, and d are obtained from a fitting to the NIST/ICRU database, published by the group of L. Grevillot et al.[58, 59]. The range was defined at the position where the dose has decreased to 80% of the maximum dose, i.e. in the distal dose falloff (r_{80}). For a mono-energetic proton beam with clinical energy, the r_{80} corresponds to the mean projected range of a proton, which is the range at which 50% of the protons have stopped. Therefore, r_{80} is considered as the most appropiate range definition, since it is independent of the initial energy spread of the beam [120].

For each beam direction, the target is then divided into n energy layers. The spacing between layers (ΔL) is specified by the user and expressed in water equivalent millimetres (mm - WED). The first layer to be irradiated (E_1) is the one with the highest energy. Each layer can be seen as a plane perpendicular to the beam direction, also called beam-eye-view (BEV) plane, where the spots are placed following a constant pattern. In particular, MIROpt uses an hexagonal pattern, with a constant user-defined distance between spots (ΔS) that corresponds to the side length of the hexagon (Figure 2.1). We chose this structure because the hexagon is the geometrical figure that best covers any surface. Indeed, in an hexagonal grid, each element boundary is as short as it can possibly be if a large area is to be filled with the fewest number of hexagons. For each layer, the spots positions are expressed in BEV coordinates (Figure 2.1), which origin is given by the beam isocenter. For each spot, four parameters are encoded in

the DICOM file: the position in BEV coordinates, the energy, and the intensity or *weight*, (x_{BEV}, y_{BEV}, E, w). The latter (w) being a variable proportional to the number of delivered protons. In the DICOM plan exported by MIROpt, the spots are ordered one after the other with increasing y_{BEV} firstly and increasing x_{BEV} secondly, starting at the deepest layer (E_1) until the most superficial one (E_n). The position (x_{BEV}, y_{BEV}) is expressed in mm, the energy (E) in MeV, and the spot weights (w) in Monitor Units (MU).

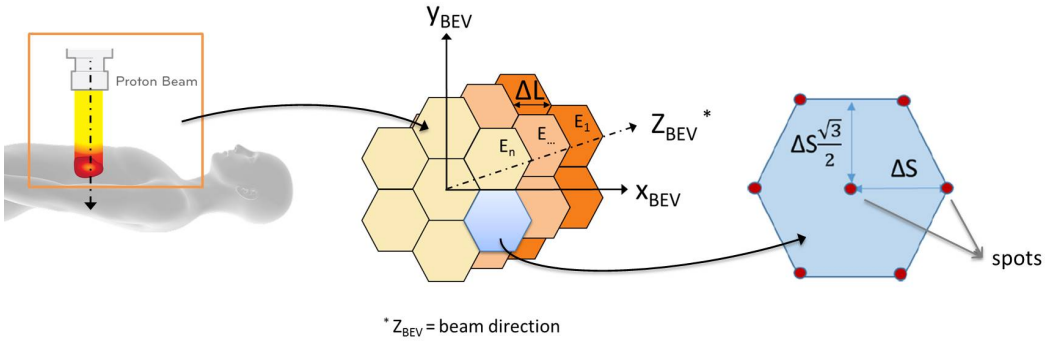


Figure 2.1: Illustration of the spot pattern and scanning grid for a given beam direction. The target is divided in n layers placed one after the other with constant spacing ΔL along the beam direction to cover the target in depth, starting at the deepest region (highest energy, E_1) until the most superficial one (lowest energy, E_n). For each layer, the spots positions are expressed in the Beam-Eye-View (BEV) coordinate system, following a regular hexagonal pattern with constant spacing ΔS.

The rest of user-defined parameters are those used to steer the optimizer towards a clinical and robust dose distribution. This includes 1) the dose objective related variables, such as target prescription and constraints for the organs at risk, and 2) the parameters for robust optimization, used to simulate the different error scenarios. The next sections (*Optimization scheme* and *Simulation of uncertainties*) give detailed information about how these parameters are interpreted and encoded in MIROpt.

2.1.2 Optimization scheme

IPOPT is an open-source software package for large-scale nonlinear optimization [68], which provides interfaces to several programming environments such as CUTEr, R, C++, Java or MATLAB. More information about the internal structure of IPOPT, as well as the interfaces can be found elsewhere [68, 177].

MIROpt uses the MATLAB interface of the IPOPT solver to find the optimal values for the spot weights in PBS plans (w), which are the variables subject to optimization. In order to steer the search direction in the variable space, IPOPT needs to be provided with an objective function (f) and its gradient (g). The objective function f formulates the problem according to the user requirements on the dose distribution; while the gradient g, i.e., the vector containing the first partial derivatives with respect to w, will help us to minimize/maximize f, since it points to the direction of fastest increase of f. If (nonlinear) constraints (c) are used, IPOPT also requires information about two more elements. The first of these elements is the Jacobian matrix (J), which contains all first partial derivatives of c with respect to w, and therefore, can be thought of as a set of direction vectors in the constraint space. The second element is the Hessian (H) of the Lagrangian (L), being the latter a mathematical tool to link the objective function f and the constraints c. The Lagrangian is formulated as $L = \sigma f + \lambda c$, where σ is the multiplier for the objective function f, and λ the vector of length m containing the Lagrange multipliers for the constraints c. The Hessian H is a matrix containing the second partial derivatives of L with respect to w, and provides information about the curvature of the Lagrange function. The Hessian is needed when nonlinear constraints are used, and it help us to find the local minima in the variable space. Detailed information about how these elements are used inside IPOPT can be found elsewhere [177]. In the following paragraphs, we describe how these terms (f, g, J and H) are computed in MIROpt, for the particular case of optimization for scanned proton therapy according to a robust worst-case scenario approach.

The objective function f is formulated as a weighted sum of quadratic terms that penalize the voxels with a dose D_i (being i the voxel index) that differs from the reference or prescription ξ_r. This is a very classical approach that has been widely used in inverse treatment planning [116], which can be stated as follows:

$$f(w) = \sum_{r \in T} \frac{\alpha_r}{N_r} \sum_{i=1}^{N_r} (\max\{0, \xi_r - D_i(w)\})^2 + \sum_{r \in T, OAR} \frac{\alpha_r}{N_t} \sum_{i=1}^{N_r} (\max\{0, D_i(w) - \xi_r\})^2 \tag{2.2}$$

where r denotes a region of interest (ROI) that is either a target ($\in T$) or an organ to spare ($\in OAR$). For each ROI, the penalty function is multiplied by a weighting factor α_r that accounts for its relative importance with respect to the other penalties, and divided by the total number of voxels N_r.

The dose in the i-th voxel is computed as

$$D_i = \sum_j P_{ij} \cdot w_j \tag{2.3}$$

where P represents the matrix containing the dose contribution for all spots (*dose influence matrix*) and j the spot index.

Robust optimization is implemented as a worst-case scenario (or minimax) problem [48]. In this approach, the uncertainties are simulated by using a set S of scenarios that models the error quantities within some interval. At each iteration, the dose is evaluated in all scenarios, and only the objective function corresponding to the worst of them (highest value of f) is minimized. The problem can then be formulated as:

$$\begin{aligned} \underset{w}{\text{minimize}} \quad & \max_{s \in S} \{f(D(w,s))\} \\ \text{subject to} \quad & w \geq 0 \end{aligned} \tag{2.4}$$

where the dose D for each scenario is computed by using equation 2.3. This means that one needs to provide a dose influence matrix for each of the considered scenarios $P(s)$, so that the dose can be expressed in matrix form as $D = P(s) \cdot w$.

The set S can be used to model any source of uncertainty, such as setup, range errors and organ motion (See section 2.1.3). However, the interval of error must be carefully selected in order to avoid too improbable scenarios that could compromise the plan quality. The advantage of this method is that only physically realizable scenarios are considered and the correlation between voxels is maintained, being less conservative than the worst-case dose distribution approach [128], where the voxels for all scenarios are handled independently to form a hypothetical worst dose distribution. In our approach, it is assumed that all scenarios have the same probability to occur, although one could eventually transform this formulation into a probability-weighted sum of all scenarios, where each of them could have a different probability to occur.

As proposed by Fredriksson et al. [48], an auxiliary variable t was introduced to reformulate the problem as a nonlinearly constrained system:

$$\begin{aligned} \underset{w,t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & t \geq f(D(w,s)) \quad \forall \quad s \in S \\ & w \geq 0 \end{aligned} \tag{2.5}$$

where the new set of optimization variables will be $x = \{t, w_1, ..., w_j, ..., w_n\}$.

The system 2.5 is convex whenever f is convex, and will have as many nonlinear constraints as the number of scenarios (m). MIROpt allows the user to select the dose objectives to be included in the max-function by using a binary term, i. e., equal to 1 when the user wants the objective to be robust, and 0 otherwise.

The form of the gradient for the objective function in 2.5 is straightforward:

$$g = \left(\frac{\partial t}{\partial t}, \frac{\partial t}{\partial w_1}, ..., \frac{\partial t}{\partial w_j}, ..., \frac{\partial t}{\partial w_n}\right) = (1, 0, ...0, ...0) \tag{2.6}$$

The Jacobian will be a matrix of size $(n+1) \times m$, being n the number of spots and m the number of nonlinear constraints (c). It can also be easily computed analytically as the partial derivatives of f (equation 2.2) evaluated for the set of scenarios $(s_1, ..., s_m)$:

$$J = \begin{bmatrix} \frac{\partial c_1}{\partial t} & \frac{\partial c_1}{\partial w_1} & \cdots & \frac{\partial c_1}{\partial w_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial c_m}{\partial t} & \frac{\partial c_m}{\partial w_1} & \cdots & \frac{\partial c_m}{\partial w_n} \end{bmatrix} = \begin{bmatrix} 0 & \frac{\partial f(w,s_1)}{\partial w_1} & \cdots & \frac{\partial f(w,s_1)}{\partial w_n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \frac{\partial f(w,s_m)}{\partial w_1} & \cdots & \frac{\partial f(w,s_m)}{\partial w_n} \end{bmatrix} \tag{2.7}$$

However, computing the Hessian matrix (H) of the Lagrangian [1], for which the matrix elements have the form $(H(L))_{ij} = \frac{\partial^2 L}{\partial x_i \partial x_j}$, requires much more computational effort. Therefore, we used the limited-memory Broyden – Fletcher – Goldfarb – Shanno algorithm (LM-BFGS), a quasi-Newton method that estimates the inverse H by using a reduced amount of computer memory. Indeed, where the classical BFGS would store a dense $(n+1) \times (n+1)$ approximation to the inverse H ($n+1$ being the number of variables), LM-BFGS stores only a few vectors that represent the approximation implicitly. Therefore, the LM-BFGS method is very convenient for optimization problems with a large number of variables, as it is the case of PBS plans, where the number of spots can reach the several thousands. This can be easily specified through the MATLAB interface of IPOPT by setting the parameter `options.ipopt.hessian_approximation` to `'limited-memory'` and `options.ipopt.limited_memory_update_type` to `'bfgs'`.

This information ($f, g, J,$ and H) is processed internally by IPOPT libraries, which will output the optimal values for the spot weights (w) as solution, after a

[1] $L = \sigma t + \lambda c$, where σ is the multiplier for the objective function t, and λ the vector of length m containing the Lagrange multipliers for the constraints c

certain number of iterations. Generally, a good solution is already achieved after 200-500 iterations. Instead of the number of iterations, one can also set stopping criteria based on the dose difference between two consecutive iterations, or any other quantity of interest.

2.1.3 Simulation of uncertainties

MIROpt can perform robust optimization including any source of uncertainty, provided that it can be simulated as a dose scenario. So far, we have included models to simulate 1) the patient positioning (setup) errors, with both their systematic and random components, 2) uncertainties in the proton range coming from image conversion errors, and 3) organ motion. In the following, we describe in detail how these uncertainties are simulated in MIROpt.

Systematic setup and image conversion errors are relatively easy to simulate, and several research and commercial implementations have been reported [166, 98, 165, 48]. Therefore, MIROpt simulates them by using existing models published in the literature. Our novelty relies mainly in the simulation of random setup errors and organ motion, which are seldom included in current robust optimizers. In this case, we take advantage of the stochastic nature of Monte Carlo (MC) methods to simulate both sources of error in a very efficient way. Regarding organ motion, what is described here refers to the change in the patient anatomy caused by the breathing motion. Although a model to simulate the change in the mean position of the tumor over time (*base-line shift*) was also investigated, and described elsewhere [146].

The concept of *nominal scenario* is used hereafter to represent the case where no errors are present, i. e., an ideal situation where the planning CT represents exactly the patient at the treatment stage.

Setup errors

Systematic setup errors

Systematic setup errors can be modelled as shifts of the patient position, or which is equivalent and easier to model, as shifts of the beam isocenters. Thus, they were simulated using the *virtual beamlets* technique [165, 48], i.e., by moving the spot weights to the neighbouring positions within the energy layers in the scanning grid and calculating the dose for the new weights using the same dose influence matrix as in the nominal scenario. Apart from the nominal, up to six scenarios can be considered, including shifts towards the positive and negative di-

rections of the R-L (right-left), A-P (anterior-posterior) and I-S (inferior-superior) axes, in the patient (DICOM) coordinates . The values for these shifts are defined by the user.

This method has the advantage of being very economical in terms of computation time and memory, since only one dose influence matrix is needed, instead of computing one for each shift. Nevertheless, using virtual beamlets could be an unsatisfactory approach when there are no spot positions corresponding exactly to the shift that needs to be performed. In this case, the nearest neighbours spots are used, which can be a more or less good approximation depending on the case. A better approach would be to compute a finer spot grid with much smaller spacing or even a dose influence matrix for each shift. Although none of these strategies are currently implemented in MIROpt, the code could be easily adapted to apply them if needed.

Random setup errors

The random component of setup errors may be simulated by sampling random shifts from a probability distribution over a patient population [45]. The latter requires computing a dose distribution for each shift if an analytical algorithm is used for dose calculation. This multiplies the computation time by the number of considered shifts, and it explains why commercial treatment planning systems have not included the modelling of this type of errors so far. However, MC methods have the capability to simulate random errors in a single dose calculation, with a free cost in terms of computation time [15, 149]. MC can simulate the effect of this type of errors in the dose distribution by shifting randomly the starting point of each particle, according to a given probability distribution. Since MIROpt is coupled with the MC dose engine *MCsquare*, we get advantage of this feature to simulate random errors as described, using a Gaussian probability distribution with a user-defined standard deviation (SD, in mm).

Applying this technique entails a blurring of the dose for each spot. Therefore, making the plan robust against a given SD may not guarantee an acceptable plan for the nominal case (SD = 0mm). Because SDs for random setup errors are typically population-based, we cannot foreseen the exact SD for a given patient. Thus, it is important to maintain an adequate plan quality for both the nominal case and the larger SD we want to be robust against. Detailed discussion about this subject, together with some results supporting these hypothesis, are presented in Appendix B.

It is important to note that this strategy assumes a infinite number of treatment fractions. However, it has been shown that using a sufficient high (finite) number, such as 20 or greater, produces equivalent results to the case with infinite fractions [45]. This supports the use of this method for conventional treatments, since typically the order of 30 fractions are used. However, using this method for hypo-fractionated treatments (1-5 fractions) still needs to be investigated.

Image conversion errors

Uncertainties in the proton range estimation may originate from artefacts in the CT image or the conversion of Hounsfield units (HU) to stopping powers, usually referred to as image conversion errors. One of the most common ways to simulate this type of errors consists in including a uniform density perturbation (ρ, expressed in %) in the CT calibration curve that gives the conversion from HU to stopping powers [48]. Therefore, all voxels in the CT image would suffer from a proportional escalation of their mass density. Typically, a value of $\rho = 3\%$ have been proposed in the literature [120], but different values can be specified in MIROpt by the user. From the user-defined ρ, three scenarios are created: the nominal case, undershoot ($\rho > 0$) and overshoot ($\rho < 0$) of all spots. This implies the computation of three dose influence matrix, one for each scenario.

This model assumes that all voxels have proportionally the same range uncertainty. However, one could refine the model by using a range uncertainty that depends on the voxel composition, i. e., function of the HU: $\rho_i(HU_i)$, being i the voxel index.

4D optimization

Including target motion in a robust worst-case scenario optimization approach can be easily done by simulating several scenarios where the target is located in different positions, with the corresponding changes in the anatomy that this might entail. In particular, the classical way of collecting information about the internal motion during the treatment is to acquire a 4D-CT image, which usually contains between 8 and 10 image captures (or *phases*) taken during the breathing cycle. Therefore, the most straightforward method to perform 4D robust optimization is to consider each phase as a dose scenario. However, this can considerably increase the computation time and memory, since the number of scenarios is multiplied by the number of phases. A more intelligent approach would be to compute the dose on each phase, to then accumulate them in a reference image using a deformable registration algorithm [97]. This would allow us to use a single scenario during

the optimization, reducing the computer memory needed. A similar approach is implemented in MIROpt, which in addition reduces the computation time thanks again to the stochastic nature of MC methods. The details of the method implemented in MIROpt are given below and the results from its application to a clinical case are presented in Chapter 4.

The number of particles N_p used by a MC dose engine to reach a good statistical precision is typically around 10^5 for a single spot dose calculation and 10^7 for a final dose computation. The 4D optimization approach implemented in MIROpt computes the dose on each phase by MC, using a reduced number of particles, i. e., N_p divided by the number of phases. Then, a 4D-Deformation Vector Field (4D-DVF) mapping each phase to a reference image is created, and used to deform the dose for each phase to the reference. At the end, all the doses are accumulated in the reference image to get the 4D dose distribution. MIROpt uses as reference the MidP-CT, which represents the tumor in its geometric time-weighted mean position [187, 186]. The MidP-CT image, as well as the 4D-DVFs are computed by using the research software REGGUI, which uses log-domain diffeomorphic Morphon algorithm for non-rigid registration [71].

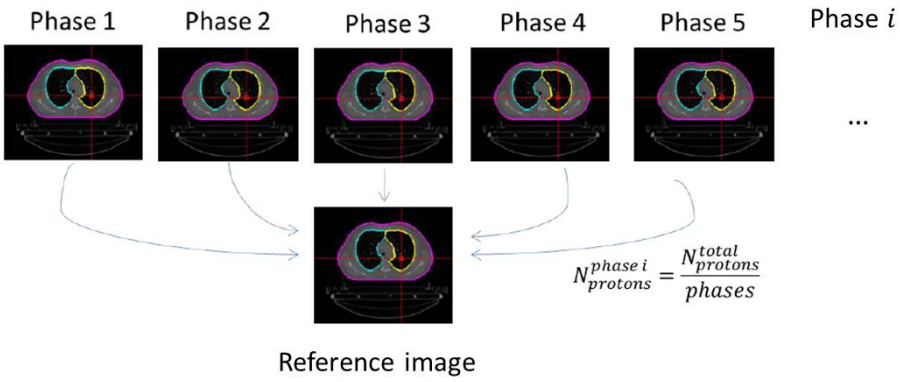


Figure 2.2: Workflow for 4D dose computation. The dose is computed on each phase of the 4DCT with a reduced number of particles, i.e., the number of particles for a regular 3D dose computation divided by the number of phases. Then, all doses are deformed and accumulated in a reference frame, which is the MidP-CT image in our case.

The advantage of this strategy is that simulating the motion is free in terms of computation time, since the total number of particles for 4D dose computation is the same as for a regular 3D case. The only extra time is coming from the final dose deformation and accumulation, which can be considered negligible with respect to the time employed in dose computation.

2.1.4 Dose calculation

MIROpt uses the fast MC code *MCsquare* [147] as main dose engine. However, the code can be easily adapted to use other dose calculation algorithms, as we did in Chapter 3 with the pencil beam algorithm provided in *FoCa* [134]. This section is focused on the use of the main dose calculation algorithm (MCsquare) inside MIROpt, putting special attention to the computation of the dose influence matrix. Although the features of MCsquare are described briefly, the details of its internal structure can be found elsewhere [147].

MCsquare is designed and optimized to simulate proton PBS treatments in voxelized geometries, such as a CT image. It exploits both task and data parallelisms of modern processors, being able to simulate 10^7 protons in less than one minute on a calculation server, which is usually enough to reduce the statistical noise to less than 2% average uncertainty for a 5×5 cm^2 beam with a 2× 2× 2 mm^3 dose grid. The proton interactions are simulated with a class-II condensed history algorithm. Stopping powers are extracted from Geant4 [117] and nuclear cross-sections are taken from ICRU 63 report [65]. Other heavy charged particles, like deuterons or alphas, are transported by scaling proton stopping powers with their charge and mass. Electrons are locally absorbed and neutral particles are neglected.

In order to compute the dose influence matrix P, MIROpt calls MCsquare repeatedly to compute the dose for each spot, with unitary weight ($w = 1$ MU). The number of particles simulated per spot (N_s) can be specified by the user, but we recommend to use at least $N_s = 5 \times 10^4$. This value is supported by the results obtained from several optimization runs on the same phantom performed with P matrices computed with different N_s, Figure 2.3. From these results, we can conclude that using $N_s = 5 \times 10^4$ or 10^5 provides a good trade-off between computation time and optimization bias, and it is consistent with the values proposed in the literature [105, 90]. Using a low precision dose influence matrix (smaller N_s) can entail errors that would be accumulated during the optimization and therefore, yield a non-optimal dose distribution (Figure 2.3).

The dose influence matrix P is stored in the MATLAB sparse matrix format [110], which is a two-dimensional matrix where each column contains the dose for each spot, i. e., P_{ij}, being i the voxel index and j the spot index. Since the main dose contribution for each spot is concentrated in a small area (around the pencil beam axis), encoding P as sparse matrix help us to have a compact storage and reduce computer memory. However, the dose computed by MC usually contains a big number of voxels with very low dose, located far from the main

dose contribution area, as a consequence of the scattering in different directions. This increases the number of entries in the sparse matrix, increasing the memory usage significantly. To cope with this issue, we implemented a threshold (t) below which the dose for a specific voxel is set to zero, and therefore, not encoded in the sparse P matrix. Again, this threshold must be carefully selected to reach a good trade-off between memory usage and optimization bias. Several thresholds were investigated, and the optimal t was found to be around 2 $\times 10^4$ eV/g/proton, Figure 2.4.

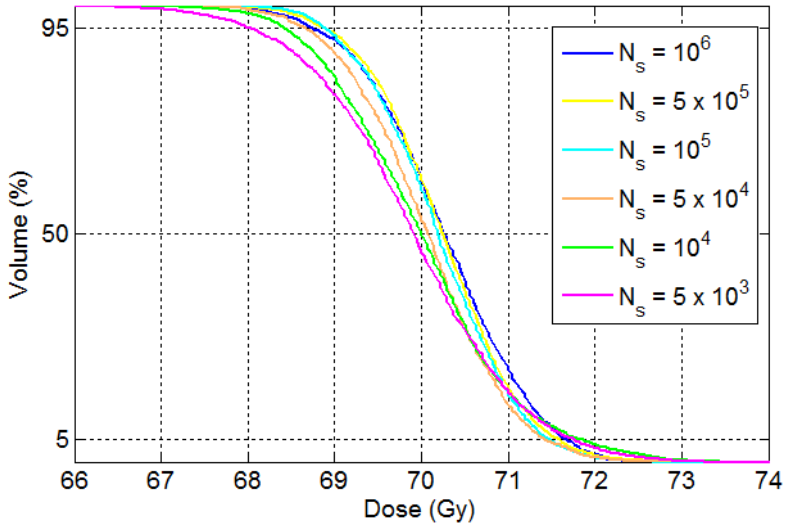


Figure 2.3: The dose influence matrix P was computed with different number of particles per spot (N_s), and used during optimization on the same phantom. The target dose prescription was set to 70Gy. The plots in the figure represent the target DVHs for the final dose distribution (computed with 10^7 particles), for the different optimization runs using P with decreasing N_s.

Other important feature regarding dose computation is the level of parallelization in MCsquare. Two different levels can be distinguished: parallelization on the number of particles or on the number of spots. By default, MCsquare divides the total number of particles between all available threads (parallelization on the number of particles), which is convenient when the number of particles is large, such as for final dose computation ($\approx 10^7$). However, when computing the dose influence matrix it might be more convenient to compute one spot dose by thread (parallelization on the number of spots), especially for 4D dose influence matrices, where the number of particles per spots is very small (divided by the number of phases) and many spots need to be computed (for all phases in the breathing cycle).

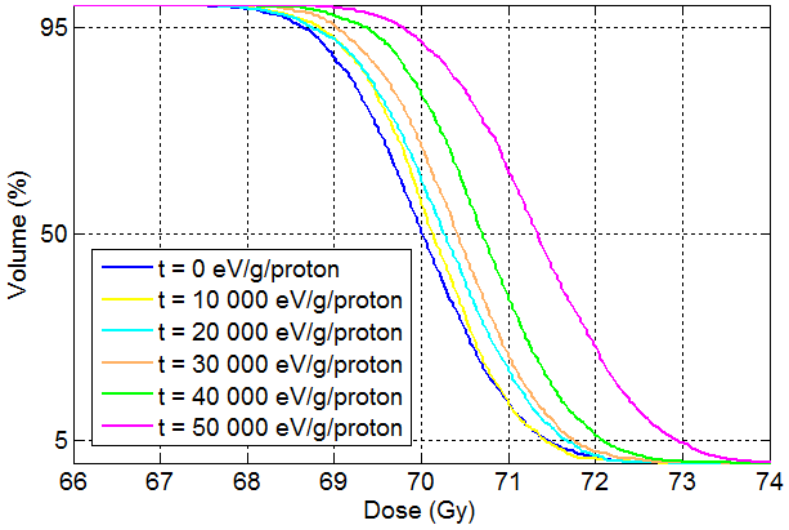


Figure 2.4: The dose influence matrix P was computed with MC and post-processed applying different thresholds (t) for all voxels. These matrices were used for optimization on the same phantom, with a target dose prescription equal to 70Gy. The plots in the figure represent the target DVHs for the final dose distribution (computed with 10^7 particles), for the different optimization runs using P with increasing t

2.2 Automatic dose painting workflow

Dose painting (See section 1.2.2) has become a hot topic during the last years, which is consequence of its potential to increase tumor local control, by using a heterogeneous dose prescription that would take into account the biological properties of the tumor. The information about these biological characteristics can be obtained from functional images, such as PET or MRI. Although the number of studies performing dose painting is very large, most of them use research tools developed in-house that are seldom open to the clinical community. Therefore, this Automatic dose painting workflow aims to bridge the gap between research and the clinics, providing a tool that can be easily used by the physician to create a dose painting plan within a few mouse clicks.

The workflow is integrated into the commercial treatment planning system RayStation (RaySearch Laboratories, Sweden) and covers all steps in treatment planning, from target segmentation to plan optimization. In this thesis, it was only used for scanned proton therapy plans, but the workflow is equally applicable to other treatment modalities available in RayStation, such as intensity modulated radiation therapy with photons. It combines several programming environments:

1) a set of MATLAB functions to segment the target volume, create the dose prescription, and evaluate the final plan, 2) the research RaySearch interface (written in C++) that allows customized optimization inside the RayStation, and 3) a script written in the Python programming language containing the main structure of the workflow, which links the MATLAB and research RaySearch interface environments with the RayStation itself, and vice-versa. The main features of this workflow and how these environments interact is described hereunder.

2.2.1 Target segmentation and dose prescription

The target segmentation and the computation of the dose painting prescription are performed by a set of MATLAB functions. These functions are called from the main script (written in Python programming language) and they are part of the software package for medical image processing IMREviewer, developed by J. A. Lee. Before starting, the workflow ask the user to import the CT and PET images.

Target segmentation

The method implemented inside IMREviewer to segment the primary tumor volume (GTV) is a gradient-based approach that uses 18-FDG-PET images as source of information. On one hand, the use of PET images is supported by studies where this modality was found to be the best for target segmentation [30], compared to other modalities like CT or MRI. On the other hand, the choice of FDG as tracer is also motivated by several publications affirming a relation between FDG uptake and radio-resistance [4], and therefore, supporting the dose escalation based on this information [5].

However, PET images are often very noisy and have poor resolution, which makes the application of gradient - based methods very difficult. Hence, the approach implemented in IMREviewer uses pre - processing tools such as denoising and deblurring to prepare the initial image for gradient-based segmentation. First, an (edge - preserving) filter is applied to the image so that the statistical noise is attenuated (image denoising). Second, the image is deconvoluted with the point spread function of the imaging device (image deblurring). Once the image is ready, a gradient - based method that combines watershed transforms and cluster analysis is applied to it. Detailed information about these three steps, as well as the validation of the algorithm is published elsewhere [50]. This algorithm has demonstrated to have a superior performance with respect to threshold-based methods.

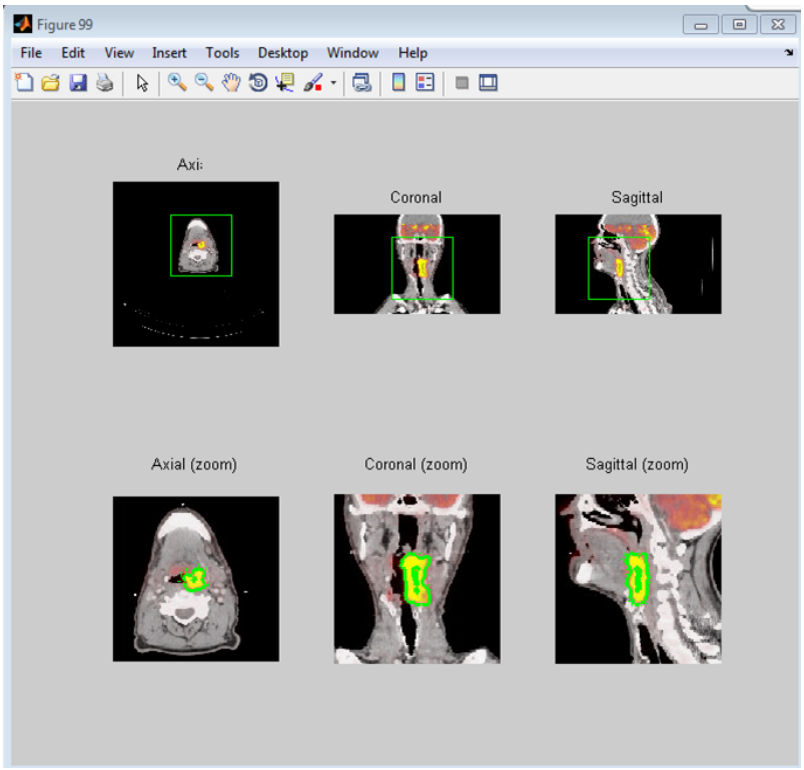


Figure 2.5: Screen-shot of the MATLAB pop-up window with the fusion of CT and PET images and the target segmentation (contour in green, figures on the bottom).

Dose prescription

Regarding the structure of the dose prescription, two different approaches are considered in the literature [17]: dose painting by contours (DPBC) and by numbers (DPBN). DPBC prescribes a uniform dose in a set of nested sub-volumes, while DPBN prescribes the dose in each voxel as a function of the uptake in the image. Both techniques are supported in this dose painting workflow.

The first step (for both DPBN and DPBC) is to compute a three - dimensional matrix with the same size and resolution as the PET image, containing the corresponding dose value for each voxel within the segmented GTV. For simplicity, we used a linear relationship between the dose and the standardized uptake value (SUV) on each voxel, although the code could be easily adapted to use another relationship. In particular, the linear SUV-dose conversion started from the median uptake value (SUV_{median}) until the maximum one (SUV_{max}), for which a user-defined minimum (D_{min}) and maximum (D_{max}) dose values were associated,

respectively. Below SUV_{median}, a constant dose equal to D_{min} was applied to all voxels, Figure 2.6. This choice aimed to reduce the partial volume effects [143], since the uptake values ranging from the minimum to the median can suffer from lower - uptake background contamination. Using the median as starting point could help to reduce the influence of cold and hot spots.

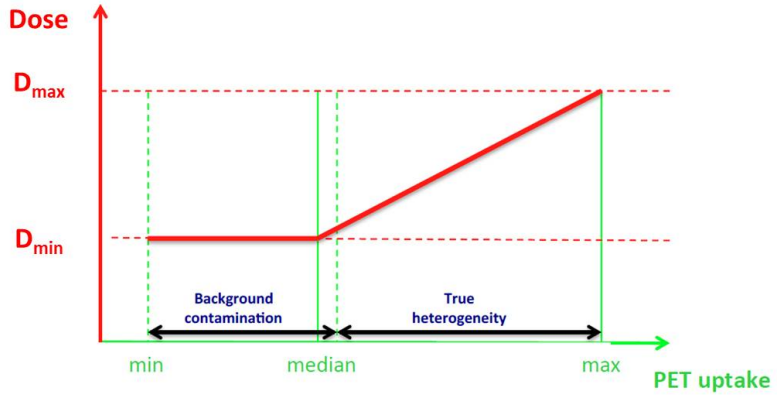


Figure 2.6: Piece-wise linear relationship between the uptake value and the dose on each voxel. The minimum and maximum dose values are set by the user. The linear conversion starts only at the median uptake value instead of the minimum one, since the latter often suffers from background contamination due to the limited resolution of PET images.

Once this three - dimensional dose matrix is computed, the user can select the desired dose painting option, either DPBN or DPBC. For DPBN, this matrix would be automatically resampled to have the size and resolution of the dose grid used inside RayStation for plan optimization. This enables to have a one-to-one correspondence between the dose prescription and the calculated dose. For DPBC, the number (N) of sub-contours to be used need to be specified by the user. This contours would be nearly equidistant and non-crossing, dividing the GTV into N sub-volumes. In both cases, either the 3D dose matrix or the N sub-contour are automatically imported into RayStation to be used during the optimization.

2.2.2 Plan optimization

The optimization, i. e. finding the optimal spot weights, is performed inside the RayStation treatment planning system. However, as most commercial systems, RayStation only supports contour - driven optimization. In this approach, the user applies dose objectives to different regions of interest (ROIs) representing the target and organs at risks. This type of optimization is very convenient for DPBC,

since one can just apply the desired dose objectives to the N sub-volumes, and the optimizer will do the job. For DPBN, where each voxel may have a different dose prescription, communicating this information to the optimizer needs some work around. Therefore, two possible methods were implemented in the automatic workflow to be able to reproduce a DPBN prescription in RayStation.

The first method for DPBN planning inside the workflow still uses contour - driven optimization. Indeed, for intensity modulated radiation therapy, it has been demonstrated that a DPBN prescription can be approximated by defining a sufficient number of non-uniform dose levels [33, 86]. This method applies a minimum and a maximum dose objective to each sub-volume, which creates a gradient within it. The minimum dose objective for the i-th sub-contour (being $i = 1$ the outermost one) is given by

$$D_{min} + (D_{max} - D_{min})(\frac{thr_i}{thr_N}), \tag{2.8}$$

where D_{max} and D_{min} are the user-defined maximum and minimum dose, and *thr* stands for the SUV thresholds: $thr_{1,...i,...N} = SUV_{median}$,...$SUV_i$,...$SUV_{max}$. The maximum dose objective for the *i*-th sub-volume equaled the minimum dose constraint for the next contour (*i*+1). During this thesis, this contour-driven optimization approach for DBPN was applied and validated for intensity modulated proton therapy, and the results are presented in Chapter 5.

The second available method to create DPBN plans enables optimization directly on the voxel scale, by means of a set of functions that enables customized optimization inside RayStation. These functions are written in C++ and are part of the RaySearch research interface. In this method, the 3D dose prescription matrix mentioned in the previous section is directly used to associate the right dose prescription to each voxel. This voxel-wise optimization approach was tested and compared against the contour-driven alternative strategy, for a set of DPBN plans. The results are presented in Appendix C.

2.2.3 Robustness against treatment uncertainties

In order to account for treatment uncertainties in the painted volume, such as setup or proton range errors, two different methods are supported in the automatic workflow. The first method extends the classical safety margin approach, which is valid only for homogeneous prescriptions, to be used for dose painting (heterogeneous) plans, by implicitly integrating the considered errors into a dilated prescription [152]. This method is available through the MATLAB functions from IMREviewer, which directly output the dilated prescription, in form of 3D matrix

or sub-contours. The second option is to use the robust optimization method implemented in RayStation, which is a worst-case scenario approach, where one can include errors coming from setup and proton range estimation. Unfortunately, this option is only available for contour - driven optimization, and it is still not supported when using the voxel-wise optimization implemented through the RaySearch research interface.

2.2.4 Evaluation of the dose painting accuracy

The conventional dose metrics used for homogeneous prescriptions, extracted from dose volume histograms (DVH), such as D_{95}, D_2 and D_{mean}, no longer apply to non-uniform prescriptions. Instead, quality volume histograms (QVH) [175] were used to assess the plan quality in the painted volume. QVHs for the painted volume can be automatically created by running a Python script after the optimization is finished and the plan is displayed.

The QVH compares the planned $D_i^{planned}$ and the prescribed dose $D_i^{prescribed}$ on each voxel i. More specifically, a QVH is computed as the decumulative histogram of the ratio $Q_i = D_i^{planned}/D_i^{prescribed}$. Therefore, ideal QVH is a descending step function, where the relative volume V equals 100% for Q_i smaller than 1 and 0% for Q_i larger than 1. From this, one can extract metrics such as $V_{Q=95\%}$ (%) and $V_{Q=105\%}$ (%), which reflect target coverage and overdosing, respectively. $V_{Q=XX\%}$ is the percentage of volume which receives at least $XX\%$ of the prescribed dose. Hence, ideal plans achieve $V_{Q=95\%} = 100\%$ and $V_{Q=105\%} = 0\%$. The quality factor QF ($QF = \frac{1}{n}\Sigma_i |Q_i - 1|$) averages the deviation of the ratios Q_i in the volume [175] and therefore reflects the spread of the QVH curve around 1.

In order to assess the robustness of the plan against treatment errors within the painted volume, the automatic workflow also provides a Python script that computes different scenarios where the dose is perturbed by the considered errors, and evaluates the QVH for each of them. All QVHs are then displayed within an encompassing 'band', which width represents the robustness of the plan, i.e., the narrower the band, the higher the robustness of the plan. Again, these tools were used to assess the plan quality for DPBN plans with scanned proton therapy, and the results are collected in Chapter 5 and Appendix C.

Chapter 3

Accurate treatment planning for proton therapy

Monte Carlo methods are considered as the *gold standard* for accurate dose calculation. However, their long computation times hampers their use in clinical routine. The first objective of this thesis is to provide treatment planning tools that allows to compute the dose accurately in a reduced computation time that is suitable for clinical implementation. Hence, this chapter investigates a hybrid optimization strategy to achieve the best trade-off between speed and accuracy, by combining analytical pencil beam (PB) and MC algorithms. The content presented here corresponds to the article: Performance of a hybrid Monte Carlo-Pencil Beam dose algorithm for proton therapy inverse planning, which is under revision in the American journal of *Medical Physics*.

3.1 Introduction

Proton therapy is becoming very popular due to its capability to focus the high dose region (i.e. Bragg peak) in the tumor volume, with a reduced dose to the surrounding healthy tissue in comparison with conventional radiotherapy. Among the different ways to deliver a proton treatment, Pencil Beam Scanning (PBS) is the most promising technique, since it enables intensity modulated proton therapy (IMPT). PBS uses narrow mono-energetic proton pencil beams, also called ≪beamlets≫ or ≪spots≫, which scan the tumor volume across the beam's transverse plane. The energy of the beamlets is then changed successively until the entire tumor is covered in depth. The intensity or weight of each beamlet can be individually adjusted to conform to the prescribed dose. This allows for high conformity of the dose distribution to the target volume and good sparing of organs at risk (OARs).

The optimal beamlets weights map results from solving an optimization problem that takes into account the dose constraints set by the user on the target and OARs. This process is known as inverse planning [116] and it has been widely used in radiation and particle therapy. For PBS inverse planning, the conventional approach consists in repeatedly calling the dose engine to compute the dose distribution of each beamlet beforehand, and store them in a dictionary usually called ≪dose influence matrix≫ used next in optimization [116, 194, 98].

In current treatment planning systems (TPS), analytical dose engines are typically used to compute the dose influence matrix because of their high speed. However, the use of analytical algorithms in treatment planning, such as pencil beam (PB), entails dose calculation errors that result from incorrect modelling of tissue inhomogeneities and scattering along density interfaces parallel to the beam [120, 144]. This can impact the dose distribution dramatically, especially

in very heterogeneous tumor sites, like lungs [55]. In contrast, Monte Carlo (MC) dose engines provide us with an accurate simulation of these phenomena [120]. Therefore, MC-based treatment planning seems to be the Holy Grail to achieve accurate proton therapy plans.

In the last years, the need for accurate treatment planning has boosted the development of fast MC engines dedicated to PBS [74, 190, 83, 193, 127, 147], which reach competitive times (less than 1 minute) for final dose calculation corresponding to a given spot weight map, hereinafter referred to as forward dose calculation. Running a forward dose calculation is an economic process in terms of number of particles simulated because the total number, typically around 107, is distributed among all the beamlets according to their previously optimized weights. However, MC engines are not efficient to compute the dose influence matrix that usually contains several thousands of spots, each of them requiring a sufficient high number of particles (around 104-105) to ensure adequate statistical precision [105]. This results in a large total number of particles simulated, including particles coming from spots with small or zero weights, which represents a waste of resources in computing beamlets that contribute very little or nothing to the dose distribution.

To cope with this issue, Y.Li et al. [90, 91], have recently developed an adaptive particle sampling (APS) scheme that iteratively performs MC dose computation and plan optimization. However, this approach needs a substantial change of the optimization workflow and perfect integration of the dose engine during the optimization process. This requires dedicated libraries to be directly applicable to current inverse planning systems [91]. Moreover, the average number of simulated particles per beamlet needed in the APS remains high (6x104 protons) in order to properly initialize the algorithm and maintain a good statistical accuracy that ensures convergence during the optimization.

In this work, we present an alternative strategy where the dose influence matrix is computed analytically and the MC is used only for forward dose calculation. Hence, this method combines the speed and efficiency of analytical algorithms with the advantages of MC planning, and can be easily applied without changing the existing optimization workflow used in current TPSs.

The concept of using two different dose engines during planning is known as hybrid optimization, which has demonstrated an excellent performance in conventional radiation therapy [142, 141] provided that the differences between the dose distributions computed by the two dose engines remain small [103]. So far, there

has been no attempt to apply PB-MC hybrid optimization to proton therapy, but it looks as an appealing alternative to full MC planning. However, due to range uncertainties, the differences between PB and MC doses in proton therapy can be much higher than those expected for photons. Therefore, proper performance evaluation of PB-MC hybrid optimization in proton therapy treatments is needed. This work addresses this issue by applying the method to different test cases in a water phantom, as well as to realistic clinical patient cases. The gain in computation time when using the hybrid method with respect to conventional MC optimization is also quantified.

3.2 Materials and methods

3.2.1 Dose engines

Monte Carlo

MCsquare [147] is a fast MC code developed in-house, designed and optimized to simulate proton PBS treatments in voxelized geometries, such as a CT image. It exploits both task and data parallelisms of modern processors, being able to simulate 10^7 protons in less than minute on a calculation server, or less than 5 minutes on a single desktop. This is usually enough to reduce the statistical noise to less than 2% for a 2x2x2 mm3 dose grid.

MCsquare simulates the proton interactions with a class-II condensed history algorithm. Stopping powers were extracted from Geant4 [117] and nuclear cross-sections are taken from ICRU 63 report [65]. Other heavy charged particles, like deuterons or alphas, are transported by scaling proton stopping powers with their charge and mass. Electrons are locally absorbed and neutral particles are neglected.

Pencil Beam

FoCa [134, 2] is in-house developed modular software that includes several tools for proton therapy plans, such as a forward dose calculation based on a convolution/superposition algorithm. It is a beam-centered approach, where a divergent calculation grid is created for each beam angle, matching the beam natural divergence, and the results are translated to the patient geometry in a final interpolation step. In its original form, FoCa computes the beamlets dose kernels analytically, but this module was modified to use lateral energy diffusion kernels in water precomputed by MCsquare, in order to perfectly align the proton ranges with the MC results.

3.2.2 Optimization scheme

Treatment planning system

The plans were created with MIROpt, our treatment planning system developed in-house. MIROpt is coupled with the two dose engines presented in the previous section: the analytical pencil beam algorithm embedded in FoCa [134, 2] and MC-square [147]. Optimization of the spot weights is performed with the large-scale non-linear solver IPOPT [177], through its MATLAB interface.

The optimization problem follows the classical approach where the objective function is a weighted sum of quadratic functions that penalize the voxels with a dose that differs from the prescription [116]. The problem is then stated as follows:

$$\min_{x} \left\{ \sum_{r \in T} \frac{\alpha_r}{N_r} \sum_{i=1}^{N_r} ((\max\{0, \xi_r - D_i(w)\})^2) + \sum_{r \in T, OAR} \frac{\alpha_r}{N_t} \sum_{i=1}^{N_r} ((\max\{0, D_i(w) - \xi_r\})^2) \right\}$$

$$\text{subject to } w \geq 0$$

where w is the vector containing the spot weights, and r denotes a region of interest (ROI) that is either a target ($\in T$) or an organ to spare ($\in OAR$). For each ROI, the penalty function is multiplied by weighting factor α_r and divided by number of voxels N_r. The penalty function depends on the difference between calculated dose D and prescription ξ_r. The dose in the i-th voxel is computed as $D_i = \sum_j P_{ij} \cdot w_j$, where P represents the dose influence matrix and j the spot index.

All the simulations were run on a machine with 2 Intel® Xeon® CPU E5-2670 of 8 cores each at 2.60 GHz, with 32 GB of RAM.

Hybrid optimization

The hybrid optimization method used in this study relies on the algorithm previously developed by Siebers et al. [141] and it is presented as a pseudo-code in Table 3.1. This method aims to combine two dose engines, one being accurate but slow and the other being fast but approximate, in the most efficient way. Let

$\ddot{\ddot{D}}$ denote the accurate dose and $\tilde{D}$ the approximate dose.

In our approach, the approximate algorithm is used to compute the dose influence matrix ($\tilde{D}$) and the accurate algorithm is only used for forward dose calculation ($\ddot{\ddot{D}}$). The forward $\ddot{\ddot{D}}$ is then used to compute a voxel-by-voxel correction term C_t (1) that will be added to the dose during the optimization process, D^c(2). Updates of C_t can be triggered as often as necessary by running the accurate dose engine, using as input the optimal values for the spot weights obtained from the last optimization loop (w_t). The index t denotes the number of updates performed so far ($t = 0, 1, 2, 3...$). The whole process can then be seen as two nested loops: 1) iterate until reaching the optimal w values for a given matrix C with index t (inner loop), and 2) update the C matrix until the desired $\ddot{\ddot{D}}$ is achieved (outer loop). Note that $C = 0$ in the first iteration of the outer loop ($t = 0$).

$$C_t = \ddot{\ddot{D}}(w_t) - \tilde{D}, \text{ where } \tilde{D} = \tilde{P} \cdot w_t \tag{3.1}$$

$$D^c = \tilde{D} + C_t \tag{3.2}$$

Before each update t, the optimized corrected dose $D^{c,o}$ and the accurate dose $\ddot{\ddot{D}}$ are properly evaluated according to the established stopping criteria. If they are satisfied, w_t is saved as the final optimal solution for the spot weights map. Otherwise, a new optimization process is run with w_t as initial weights and using the fresh-updated C_t matrix.

(1)	Initialize $C_t = 0$ $(t = 0)$
(2)	Optimize using D^c
(3)	If inner stopping criteria (*iSC*) achieved, save w_t and the optimized $D^{c,o}$
(4)	Compute $\ddot{\ddot{D}}(w_t)$ and C_t
(5)	If $\ddot{\ddot{D}}(w_t)$ and $D^{c,o}$ satisfies the outer stopping criteria (*oSC*),return w_t; else, $t = t + 1$ and go to (2)

Table 3.1: Hybrid optimization pseudo-code. Instruction 1 initializes the algorithm, Instructions 2-3 correspond to the inner optimization loop, while Instructions 4-5 represent the outer optimization loop.

Stopping criteria

Using a good termination criterion is crucial in any optimization algorithm since it will determine the quality of the final result as well as the number of iterations needed to reach it. Therefore, different stopping criteria (SC) were investigated

for the inner and outer loops, combined with a maximum number of iterations.

The first inner loop ($t = 0$) must achieve an optimal approximate dose in order to have a good starting point for the subsequent corrections ($t > 0$). Therefore, a convergence criterion ϵ 4.5 based on the relative dose difference between two consecutive iterations k and $k-1$ [90, 91] was always used for this initial run:

$$\epsilon = \frac{\sqrt{\frac{\sum_{i=1}^{N}(D_i^N - D_i^{k-1})^2}{N}}}{\xi_{PTV}} \times 100\%, \tag{3.3}$$

where N represents the total number of voxels summed over all ROIs involved in the objective function and ξ_{PTV} is the dose prescription for the target volume (PTV). Optimization was stopped when ϵ had not exceeded 0.005 for the last 5 iterations.

For $t > 0$, two different approaches were studied: 1) the same $\epsilon < 0.005$ criterion as for $t = 0$ (inner stopping criteria 1, *iSC1*); and 2) a set of early stopping criteria on the optimized corrected dose $D^{c,o}$ (2) to the target volume based on the usual clinical constraints $D_{95} \geq 0.95 D_{prescription}$, $D_5 \leq 1.05 D_{prescription}$ and a user-defined threshold for D_{5-95} (inner stopping criteria 2, *iSC2*). Moreover, a maximum of 500 iterations was set for all inner loops, which is enough to get a good dose distribution in our treatment planning system.

The outer loop computes which correction term C should be added to $\tilde{D}$ in order to get an optimal $\dddot{D}$. Ideally, the convergence of the outer loop would be reached when $D^{c,o} \approx \dddot{D}$ (outer stopping criteria 1, oSC1). This was quantified in Siebers et al. [141] via dose prediction errors (DPE) defined as $DPE = D^{c,o} - \dddot{D}$. However, one could already stop the algorithm if $\dddot{D}$ is within the clinical constraints (outer stopping criteria 2, oSC2), since the MC dose is considered as the gold standard and the aim is to get a good clinical $\dddot{D}$. This can be seen as an early stopping of the algorithm and could be directly evaluated by looking at the DVH metrics for $\dddot{D}$. Both oSC1 and oSC2 were combined with a maximum of 5 outer iterations (updates of C), which is considered as a good tradeoff between accurate final dose distribution and computation time.

3.2.3 Feasibility study

Previously published studies about the performance of hybrid optimization in conventional radiotherapy have reported excellent results when having small differences between the doses computed by the two considered algorithms [142, 141, 103]. In proton therapy, small range errors may have severe consequences in

the dose distributions, which might degrade performance of the hybrid algorithm. Therefore, the effect of different dose calculation errors in proton dose distributions must be carefully analyzed one by one. For this purpose, we first designed several test cases in a simplistic phantom and second, we applied the method to three realistic clinical cases.

Test cases

This section aims to analyze the effect of different types of dose calculation errors on the performance of the hybrid method. In order to analyze the errors in a controlled manner, a single dose engine was used and each error was simulated individually. MCsquare was chosen as reference dose engine since computation time was not an issue for this experiment. The approximate dose $\tilde{D}$ was then given by the product of the MC dose influence matrix ($\ddot{\tilde{P}} = P_{MC}$)and the spot weights. Accurate dose $\dddot{D}$ was also a MC dose but included the simulation of dose calculation errors (Table 3.2). The number of particles per beamlet used to compute $\tilde{P}$ was equal to 10^5, which is supposedly enough to achieve good statistical precision [105]; while 10^7 particles were used for forward dose calculation.

The four different dose calculation errors detailed below were simulated in $\dddot{D}$:

- Constant **dose scaling** was simulated by multiplying the dose in each voxel by a factor α. Two plans were created with $\alpha = 0.95$ and 0.90.

- **Dose blurring** was simulated by applying a convolution with a three-dimensional normalized Gaussian kernel (K) of standard deviation σ. The standard deviation for each dimension was proportional to the corresponding voxel size d, i.e.,$\sigma_x \propto d_x$, and the same for y and z coordinates. Two plans were created with σ equal to 0.3 d and 0.5 d.

- Errors in **range estimation** were simulated by scaling the mass density uniformly for all voxels of the CT image by a factor ρ. Range under- and over-estimation were simulated with positive and negative values of ρ, respectively. Four hybrid plans were created with density scaling of $\rho = +3, -3, +1.5$and -1.5%.

- **Lateral dose displacements** are simulated by simply shifting the dose distribution by x mm in the plane perpendicular to the beam direction. Two plans were generated for a shift x equal to 2.5 and 5 mm.

Only one stopping criterion was investigated for all inner loops ($t \geq 0$) in

Error type	$\dddot{D}$	Test cases
Dose scaling	$\dddot{D} = \tilde{D} \cdot \alpha$	$\alpha = 0.95$ and 0.90
Dose blurring	$\dddot{D} = \tilde{D} * K(\sigma)$	$\sigma = 0.3d$ and $0.5d$mm
Range under/over estimation	$\dddot{D} = \tilde{D}(\rho)$	$\rho = \pm 3$ and $\pm 1.5\%$
Lateral shift	$\dddot{D} = S_x(\tilde{D})$	$x = 2.5$ and 5mm

Table 3.2: Test cases generated for several dose calculation errors, and the corresponding form for $\dddot{D}$.

the hybrid optimization, namely, $\epsilon < 0.005$ during the last 5 iterations together with a maximum of 500 iterations (*iSC1*). For the outer loop, condition $D^{c,o} = \tilde{D} + C \approx \dddot{D}$ was used to stop the algorithm (*oSC1*), which was evaluated via $DPE = D^{c,o} - \dddot{D}$ on the target doses. If the DPE for DVH metrics D_5 and D_{95} were less than 0.25 Gy in absolute value, then condition $D^{c,o} \approx \dddot{D}$ was satisfied. Moreover, the possibility of early stopping based on clinical goals for D_{MC}^{sim} was also investigated (*oSC2*). These goals were set as $D_{95} \geq 57$ Gy and $D_5 \leq 63$ Gy for the target volume, given a dose prescription of 60Gy.

The designed phantom contained a cube with edges of 30 cm and Hounsfield units (HU) equal to 0 inside the cube (water-like) and -1000 outside the cube (air-like). The tumor was modelled as a sphere of 3 cm diameter located in the middle of the cube. The CT voxel size d was $2.5 \times 2.5 \times 2$ mm. A single beam at 0o was used to create all plans, and a dose of 60 Gy was supposed as target prescription.

Patient cases

The hybrid method was applied to three different clinical cases: prostate, brain, and lung. The dose influence matrix used to compute the approximated dose $\tilde{D}$ was calculated by the pencil beam algorithm in FoCa ($\tilde{P} = P_{PB}$) and the accurate forward dose was given by MCsquare ($\dddot{D} = D_{MC}$). The number of particles used to compute $\dddot{D}$ was estimated so that the mean relative uncertainty in the target volume was less than 1%, given a dose grid resolution of $2.5 \times 2.5 \times 2$ mm in the three cases. By using a batch method [74], this resulted in 5×10^7 particles for prostate (0.73%) and lung (0.7%), and 10^7 particles for the brain case (0.79%).

All hybrid plans were compared with the conventional full MC approach [105] – our reference case. This full MC-based approach uses a dose influence matrix

entirely computed by MC (P_{MC}) with 10^5 particles per beamlet, as in the previous section.

The stopping criteria used for the full MC-based plans equaled *iSC1*, i.e., $\epsilon < 0.005$ during the last 5 iterations, with a maximum of 500 iterations. For the hybrid method, two sets of plans were created, one for each of the two inner stopping criteria *iSC1* and *iSC2*, presented in section 3.2.2. The target dose objectives for *iSC2* ($D_{95} \geq 0.95 D_{prescription}, D_5 \leq 1.05 D_{prescription}$) were computed according to the dose prescription for each case: 74Gy for prostate, 59.4 Gy for brain and 60 Gy for lung. The D_{5-95} threshold was based on the results obtained from the full MC plans, since our aim was to prove that the hybrid method can reproduce the reference case as close as possible. For both sets of plans, the two stopping criteria for the outer loop presented in section 3.2.2, *oSC1* and *oSC2*, were investigated. The DPE threshold for *oSC1* was set to 0.25 Gy (absolute value) as in the previous section, while the target dose objectives for *oSC2* (D_{95}, D_5) were again computed according to the dose prescription for each patient. The doses for OARs were not taken into account when evaluating *oSC2*.

For simplicity, PTV-based optimization was used for all the cases. For the prostate case, the PTV was created by expanding the CTV by 5mm. For the brain case, an expansion of 2 mm was used. For the lung case, the ITV was expanded by 5mm to create the PTV and the average CT was used for planning. Range shifters of 4 and 5 cm water equivalent length (WED) were used for the brain and the lung case, respectively. The range shifters were simulated by delineating a rectangular structure of HU $= 0$ in the planning CT. Figure 3.1 illustrates the CT scans and the beam settings for each patient.

The HU outside the body contour were re-written to be -1000 (air) for all CTs, in order to minimize the differences between the MC and the PB algorithm due to possible artefacts outside the patient, as for instance the CT couch. The same CT calibration curve was used for both dose engines, in order to avoid discrepancies coming from converting HU into stopping powers.

For both methods, hybrid PB-MC and full MC, a final forward dose calculation D_{MC} was performed, which ensures a fair comparison with good statistical accuracy. The final doses were normalized to have $D_{50}^{P}TV$ equaling the dose prescription in order to facilitate plan comparison.

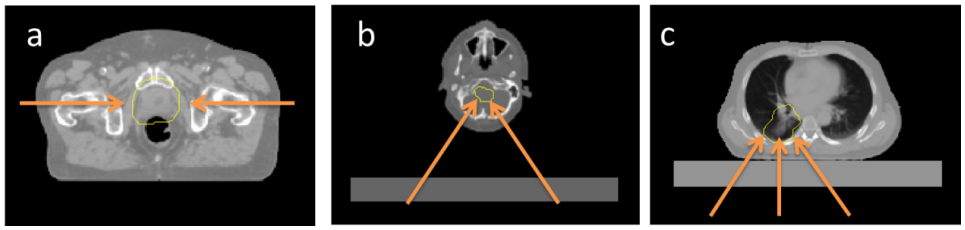


Figure 3.1: Axial CT slices of the three patients used for the study: a) prostate, b) brain and c) lung. The yellow contour represents the PTV volume, and the arrows indicate the beam angles used for treatment.

3.3 Results

3.3.1 Test cases

The first inner optimization loop ($C = 0$) resulted in an optimal dose distribution $D^{c,o}$ with $D_{95} = 59.41$ Gy and $D_5 = 60.5$ Gy for the target. Then, different dose errors were simulated in $\ddot{D}$ for the four test cases described in Section 3.2.3. The target DVH metrics and dose prediction errors (DPE) for $\dddot{D}$ after applying the hybrid method are summarized in Table 3.3.

For dose scaling errors (Table 3.3, first row), the hybrid method converged to excellent target dose distributions $\dddot{D}$, with $D_{95} \geq 59.2$ Gy and $D_5 \leq 60.5$ Gy. The DPE(D_{95}) ranged from 0.06 to 0.15 Gy and DPE(D_5) ranged from 0.05 to 0.18 Gy, which fully satisfies *oSC1* ($|DPE| < 0.25$ Gy). The *oSC1* and *oSC2* conditions were attained for the same number of outer iterations, which increased from 1 to 2 when decreasing the dose scaling factor (α) from 0.95 to 0.90.
For dose blurring errors (Table 3.3, second row), using a kernel with standard deviation $\sigma = 0.3$ required 2 outer iterations to satisfy *oSC1*, while $\sigma = 0.5$ required 4 iterations. In contrast, *oSC2* was fulfilled after only 1 iteration for both $\sigma = 0.3$ and 0.5. The dose distributions $\dddot{D}$ presented a very good target coverage in all cases ($D_{95} > 59.3$ Gy), regardless of the considered stopping criterion. The target overdose was also very low ($D_5 < 60.8$ Gy), except in the case $\sigma = 0.5$, where using the early stopping criteria (*oSC2*) led to $D_5 = 62.07$ Gy, which still fulfilled the clinical constraint ($D_5 < 63$ Gy).

For range errors (Table 3.3, third row), the dose distribution $\dddot{D}$ did not improve after applying the corrections in most of the cases. Indeed, the D_{5-95} after correction was larger than before correction ($C = 0$) in all cases except one ($\rho = +1.5\%$ for *oSC1*). For the smallest shift ($\rho = \pm 1.5\%$), D_{95} was slightly improved or maintained at the cost of increased overdose. However, increasing the density

perturbation to 3% had a dramatic impact on the final dose, degrading both coverage and overdose. For instance, in the case with $\rho = +3\%$, the optimization reached the maximum number of iterations without satisfying any of the stopping criteria (# iter = 5*), presenting a dose distribution $\ddot{\ddot{D}}$ with $D_{95} = 35.13$ Gy and $D_5 = 79.65$ Gy.

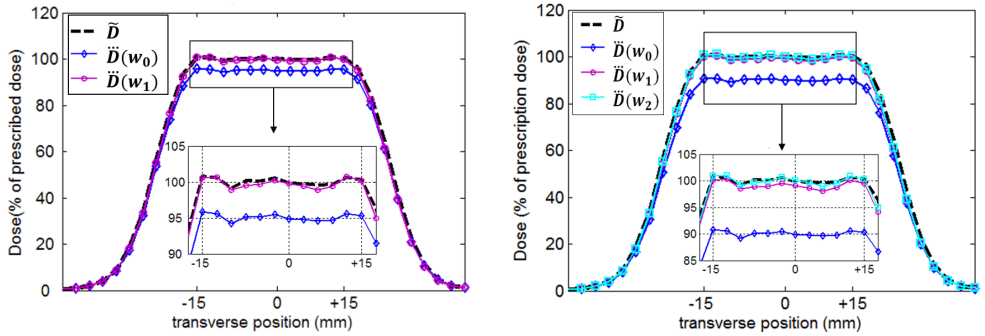
Simulation of lateral dose displacements (Table 3.3, fourth row) led to similar results than the cases with range errors. For the smallest shift ($x = 2.5$ mm), the target coverage was improved at the cost 310 of increasing overdose: D_{95} increased by 0.7 Gy for *oSC1*, and by 2.44 Gy for *oSC2*, with respect to the value before correction ($C = 0$). Again, increasing the magnitude of the shift ($x = 5$ mm) resulted in a totally degraded target dose, with a D_{5-95} that increased from 8.11 Gy to 24.29 Gy.

Dose profiles along the beam direction and perpendicular to it were also analyzed for all test cases. For dose scaling (Figure 3.2) and dose blurring cases (Figure 3.3), the hybrid method managed to add the required dose at the right place, recovering excellent target coverage. For cases involving shifts, range errors (Figure 3.4) and lateral displacements (Figure 3.5), the algorithm did not succeed to compensate the lack of dose in the shifted region. In this case, the first correction tried to compensate the lack of dose by adding a peak of high dose before the dose gap. The dose peak grew when increasing the magnitude of the shift, reaching doses up to 80 Gy for the cases with $\rho = +3\%$ (Figure 3.4c) and $x = 5$ mm (Figure 3.5b, right), which represents a 30% overdose with respect to the dose prescription (60Gy). The following corrections tried to decrease this overdose but this led to dose distributions that diverged from the optimal solution.

3.3.2 Patient cases

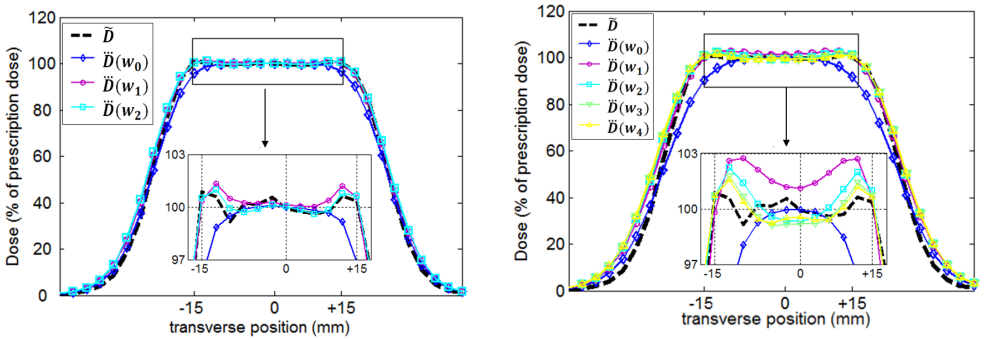
According to the results obtained from the full MC-based optimization ($\tilde{D}$ 4), and considering our reference case, the target D_{5-95} threshold for the early stopping criteria of the inner loop in the hybrid method (*iSC2*) was set to 3.5 Gy for the prostate case, 2.1 Gy for the brain and 2 Gy for the lung.

The approximate dose distribution ($\tilde{D} = D_{PB}$) obtained after the first inner optimization loop of the hybrid method ($C = 0$) achieved an excellent target coverage that fully satisfied the clinical constraints ($D_{95} \geq 0.95 D_{prescription}$, $D_5 \leq 1.05 D_{prescription}$) with a high homogeneity: D_{5-95} equaled 1.7 Gy for prostate, 0.31 Gy for brain, and 1 Gy for lung. However, the MC recomputed doses (Table 3.5, $C = 0$) showed important degradation of target coverage, with



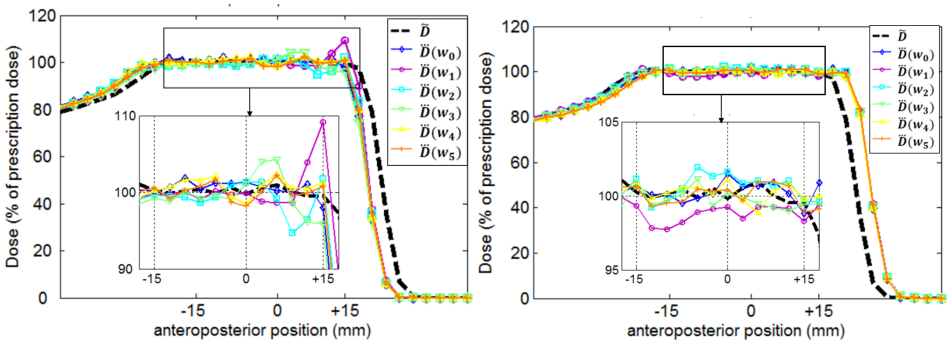
(a) Dose scaling $\alpha = 0.95$ (b) Dose scaling $\alpha = 0.90$

Figure 3.2: Dose profiles for the dose scaling test case, with the dose scaling factor (α) equal to 0.95 (a) and 0.90 (b). The profiles are taken in the plane perpendicular to the beam, along the line that divides the spherical target volume in two equal parts. The dashed line represents the approximate dose ($\tilde{D}$) whereas the solid lines with markers represents the accurate dose at the t-th correction ($\ddot{D}(w_t)$).

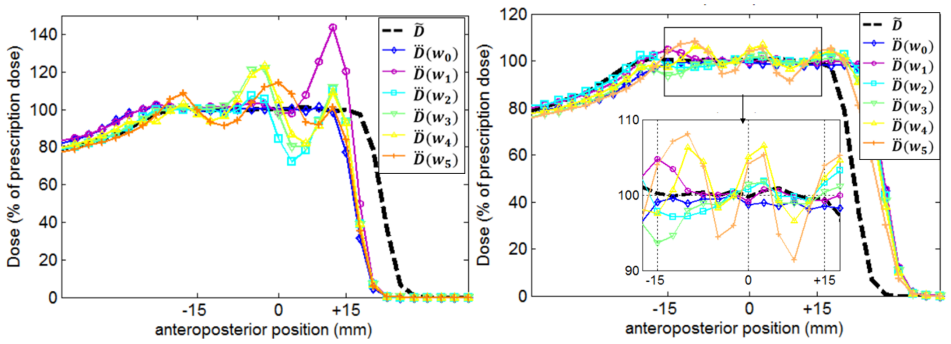


(a) Dose blurring $\sigma = 0.3d$ (b) Dose blurring $\sigma = 0.5d$

Figure 3.3: Dose profiles for the dose blurring test case, with the standard deviation (σ) for the Gaussian convolution kernel (K) equal to $0.3d$ (a) and $0.5d$ (b), being d the voxel size in milimeters. The profiles are taken in the plane perpendicular to the beam, along the line that divides the spherical target volume in two equal parts. The dashed line represents the approximate dose ($\tilde{D}$) whereas the solid lines with markers represents the accurate dose at the t-th correction ($\ddot{D}(w_t)$).



(a) Range undershoot $\rho = +1.5\%$ (b) Range overshoot $\rho = -1.5\%$

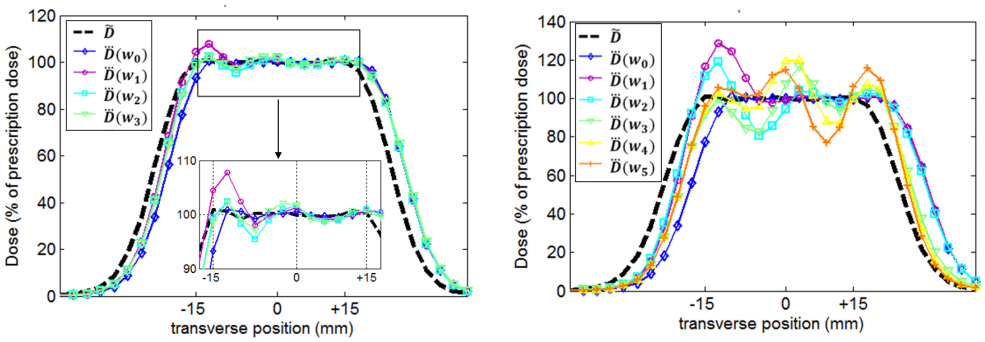


(c) Range undershoot $\rho = +3\%$ (d) Range overshoot $\rho = -3\%$

Figure 3.4: Dose profiles for the cases with range uncertainties (ρ) equal to 1.5% (a), -1.5% (b), 3% (c) and -3% (d). The profiles are taken along the beam direction, passing through the middle of the target. The dashed line represents the approximate dose ($\tilde{D}$) whereas the solid lines with markers represents the accurate dose at the t-th correction ($\ddot{D}(w_t)$).

Test case		Stopping criteria (SC)	# iter	D_{95} (Gy)	D_5 (Gy)	D_{5-95} (Gy)	DPE(D_{95}) (Gy)	DPE(D_5) (Gy)
Dose scaling	$\alpha = 0.95$	C=0	0	56.44	57.47	1.03	2.97	3.03
		oSC1& oSC2	1	59.19	60.41	1.22	0.15	0.18
	$\alpha = 0.90$	C=0	0	53.47	54.45	0.98	5.94	6.05
		oSC1 & oSC2	2	59.27	60.51	1.24	0.06	0.05
Dose blurring	$\sigma = 0.3$	C=0	0	57.03	60.00	2.97	2.38	0.50
		oSC1	2	59.74	60.52	0.78	-0.13	-0.17
		oSC2	1	59.79	60.79	1.00	-0.20	-0.42
	$\sigma = 0.5$	C=0	0	53.47	54.45	0.98	5.94	6.05
		oSC1	4	59.46	60.68	1.22	-0.04	-0.14
		oSC2	1	59.37	62.07	2.70	0.12	-1.64
Range under/over estimation	$\rho = +1.5\%$	C=0	0	58.08	61.15	3.07	1.33	-0.65
		oSC1	5*	58.78	61.33	2.55	0.29	-0.46
		oSC2	1	59.19	62.92	3.73	-0.16	-2.04
	$\rho = -1.5\%$	C=0	0	59.05	60.76	1.71	0.36	-0.26
		oSC1	2	58.99	60.91	1.92	0.23	-0.21
		oSC2	1	59.10	61.01	1.91	0.21	-0.36
	$\rho = -3\%$	C=0	0	51.37	61.37	10.00	8.04	-0.87
		oSC1 & oSC2	5*	35.13	79.65	44.52	12.45	-10.17
	$\rho = +3\%$	C=0	0	57.9	60.55	2.65	1.51	-0.05
		oSC1	5*	55.34	66.75	11.41	2.97	-4.58
		oSC2	1	59.12	62.71	3.59	0.21	-2.1
Lateral shift	$x = 2.5mm$	C=0	0	58.20	60.49	2.29	1.21	0.01
		oSC1	3	58.81	61.20	2.39	0.20	-0.21
		oSC2	1	59.30	62.93	3.63	0.02	-2.28
	$x = 5mm$	C=0	0	52.34	60.45	8.11	7.07	0.05
		oSC1 & oSC2	5*	48.52	72.81	24.29	7.26	-8.16

Table 3.3: Target volume DVH metrics (D_{95}, D_5 and D_{5-95}) and dose prediction errors (DPE) for the accurate dose including the simulated errors for each test case ($\ddot{D}$) before correction ($C = 0$) and after applying the hybrid method. The number of iterations (# iter) for the outer loop needed to satisfy the stopping criteria (SC) is presented in the 4th column. In cases where the maximum number of iterations (# iter = 5) is reached without satisfying the SC, an asterisk is presented next to the value (# iter = 5*). *oSC1* stands for the first outer stopping criteria, which requires DPE(D_{95}) and DPE(D_5) to be less than 0.25 Gy. *oSC2* stands for the second stopping criteria (early stopping), where only the clinical constraints (D_{95} $\geq$57Gy, D_5 $\leq$63Gy) for $\ddot{D}$ are evaluated. In some cases, both criteria are met for the same number of iterations (*oSC1 & oSC2*).



(a) Lateral displacement $x = 2.5$ mm (b) Lateral displacement $x = 5$ mm

Figure 3.5: Dose profiles for the cases including a lateral displacement of the dose (x) equal to 2.5 mm (a) and 5 mm (b). The profiles are taken in the plane perpendicular to the beam, along the line that divides the spherical target volume in two equal parts. The dashed line represents the approximate dose ($\tilde{D}$) whereas the solid lines with markers represents the accurate dose at the t-th correction ($\ddot{D}(w_t)$).

	D_{95} (Gy)	D_5 (Gy)	D_{5-95} (Gy)
Prostate	72.30	75.70	3.40
Brain	58.36	60.44	2.08
Lung	59.00	60.95	1.95

Table 3.4: Target dose metrics (D_5, D_{95}andD_{5-95}) for full MC-based optimization in the three clinical cases (prostate, brain and lung).

DPE(D_{95}) that reached 3.27 Gy (prostate), 4.44 Gy 379 (brain), and 2.55 Gy (lung). The difference in target overdose was also significant since DPE(D_5) exceeded 2.9 Gy in all cases (Table 3.5).

The hybrid method was able to correct for the differences between PB-MC, thereby providing plans with target coverage within the clinical constraints in all cases (Table 3.5). Plan quality varied according to the stopping criteria used in optimization. For all patients, the best target coverage was obtained by using DPE ¡0.25 Gy as stopping criterion for the outer loop (*oSC1*), with a minimum D_{95} of 97.5%, 97.6%, and 98.5% of the dose prescription for prostate, brain, and lung, respectively. However, the 386 combination [*iSC1, oSC1*] reached the maximum number of iterations for the outer loop without satisfying condition DPE¡0.25 Gy ($\#$ iter $= 5$*) for prostate and brain cases. In contrast, using the combination [*iSC2, oSC1*] led to plans with DPE ¡0.25 Gy after 2 iterations for prostate and

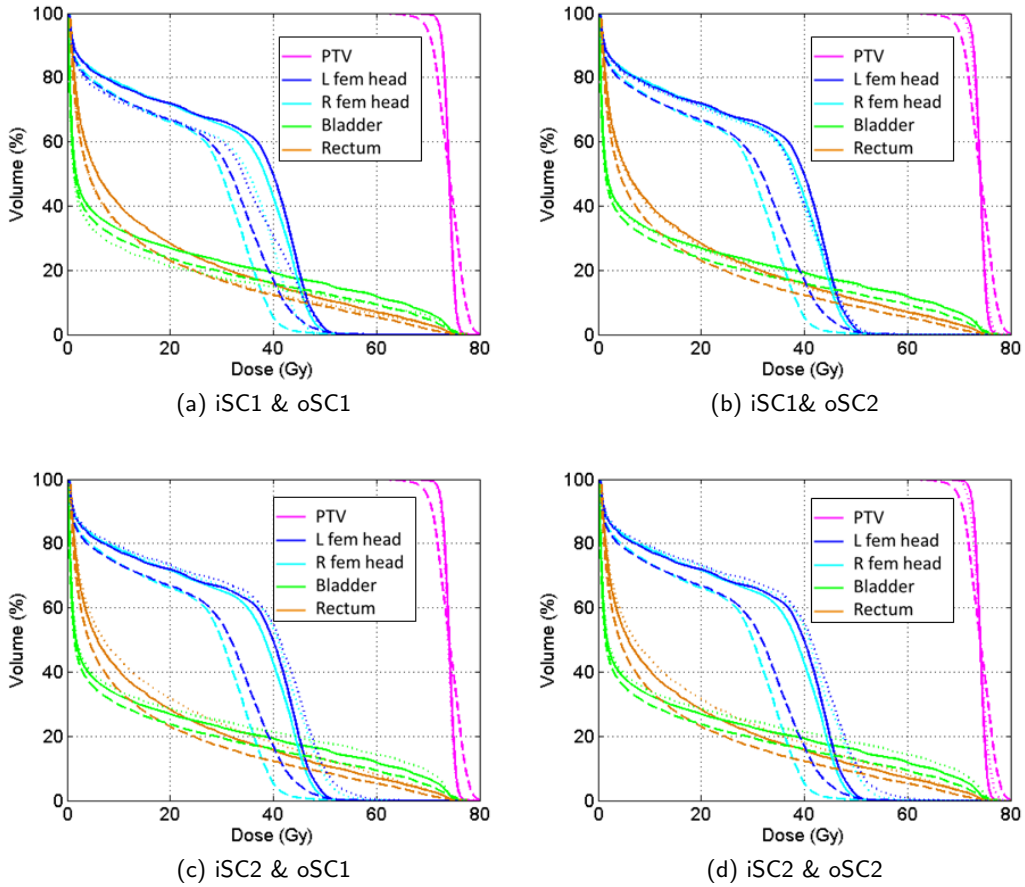
lung cases, and 3 iterations for the brain case (Table 3.5). Using the clinical constraints for D_{95} and D_5 as stopping criteria (*oSC2*) terminated the optimization after only 1 iteration of the outer loop for all patients. This led to slightly lower target coverage, compared to the results for *oSC1*, with minimum D_{95} equal to 96.8%, 96.6%, and 97.2% of the dose prescription for prostate, brain, and lung, respectively. The values of D_5 were always below 103.5% of the dose prescription, regardless of the used stopping criteria.

The dose to OARs in the hybrid plans was also influenced by the stopping criteria used for the inner and outer loops. This is illustrated with the DVHs presented in Figures 3.6 (prostate), 3.7 (brain), and 3.8 (lung). In general, the lowest doses were obtained when applying the convergence condition ($\epsilon < 0.005$) to the inner loop (*iSC1*). This resulted in DVHs that were equivalent to the reference full MC case (Figures 3.6,3.7,3.8 plots a-b). In the prostate case, the femoral heads volume that received a dose equal or bigger than 50Gy was less than 3% ($V_{50} <$ 3%), for combinations [*iSC1, oSC1*] and [*iSC1, oSC2*], which is very close to the reference case where $V_{50} < 2\%$. In the bladder, V_{60} equaled 10% for [*iSC1, oSC1*] and 12% for [*iSC1, oSC2*], which is equivalent to the reference value (V_{60} = 12%). The dose to the rectum was also the same as in the reference case ($V_{60} = 7\%$), with a $V_{60} = 6\%$ and 7% for combinations [*iSC1, oSC1*] and [*iSC1, oSC2*], respectively. For the brain case, the condition [*iSC1, oSC1*] led to a dose in the brainstem with D_2 equal to 59.6 Gy, which was identical to reference case value ($D_2 = 59.8$ Gy); and D2 for the spinal cord differed in no more than 1 Gy. The lung case presented a D_2 for the spinal cord that was reduced by 1.1 Gy with respect to the reference when 408 using the criteria [*iSC1, oSC1*] (Figure 3.8a), and by 2.2 Gy for [*iSC1, oSC2*] (Figure 3.8b). The rest of the OARs not mentioned here presented a DVH that totally overlapped those in the reference case, implying an identical dose distribution.

If the early stopping criteria for the inner loop (*iSC2*) were used instead, the doses to certain OARs for prostate and lung cases were slightly higher than the reference full MC plans. For both combinations [*iSC2, oSC1*] and [*iSC2, oSC2*], V_{50} to femoral heads in the prostate case was bigger than 7%; and the V_{60} for rectum and bladder were 1% and 2% higher than the reference case, respectively. For the lung case, these stopping criteria resulted in a higher mean dose to the spinal cord (Figure 3.8c, 3.8d), but D2 was as good as in the reference case since the maximum difference was 0.3 Gy. For the brain case, using *iSC2* as termination criteria led to dose distributions that were similar to those given by *iSC1* (Figure 3.7c,3.7d).

Patient	Stopping criteria (SC)		# iter	D95 (Gy)	D5 (Gy)	D5-95 (Gy)	DPE(D95) (Gy)	DPE(D5) (Gy)
Prostate	C=0		0	69.84	77.97	8.13	3.27	-3.13
	iSC1	oSC1	5*	72.29	75.54	3.25	0.66	-0.49
		oSC2	1	71.78	76.06	4.28	0.97	-0.78
	iSC2	oSC1	2	72.14	75.53	3.39	0.19	0.08
		oSC2	1	71.61	76.62	5.01	0.65	-0.95
Brain	C=0		0	54.80	62.85	8.05	4.44	-3.30
	iSC1	oSC1	5*	57.97	61.27	3.30	0.56	-1.04
		oSC2	1	57.54	61.42	3.88	1.16	-1.32
	iSC2	oSC1	3	58.16	60.6	2.44	0.18	-0.19
		oSC2	1	57.41	61.13	3.72	0.87	-0.77
Lung	C=0		0	57.16	63.71	6.55	2.55	-2.95
	iSC1	oSC1	3	59.24	61.35	2.11	0.23	-0.16
		oSC2	1	58.33	61.79	3.46	1.20	-0.66
	iSC2	oSC1	2	59.12	61.46	2.34	0.24	-0.11
		oSC2	1	58.49	61.59	3.10	0.93	-0.18

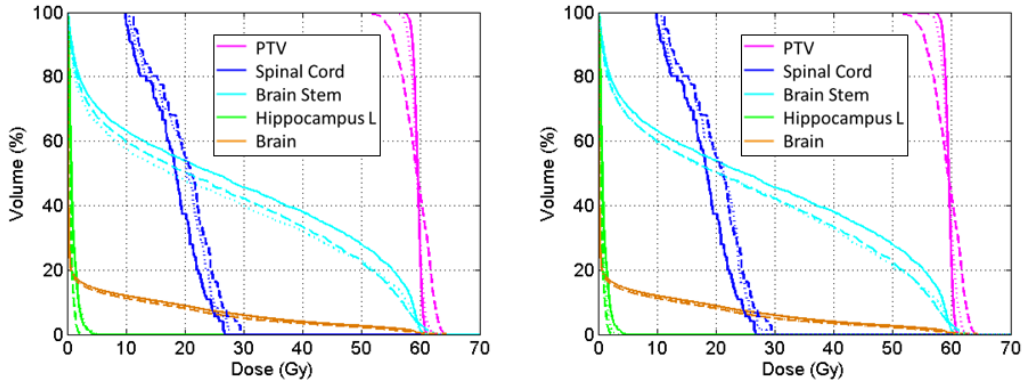
Table 3.5: Target volume DVH metrics (D_{95}, D_5andD_{5-95}) and dose prediction errors (DPE) for the MC doses obtained from hybrid optimization. For each patient, the first row contains the results before correction (C=0), while the second and third row correspond to the two sets of plans created for the inner stopping criteria iSC1 and iSC2. In both cases, the results after correction when applying the two different stopping criteria for the outer loop (oSC1 and oSC2) are presented. The number of iterations (# iter) needed to reach oSC1 and oSC2 is presented in the 4th column. In cases where the maximum number of iterations (# iter = 5) is reached without satisfying the SC, an asterisk is presented next to the value (# iter = 5*).



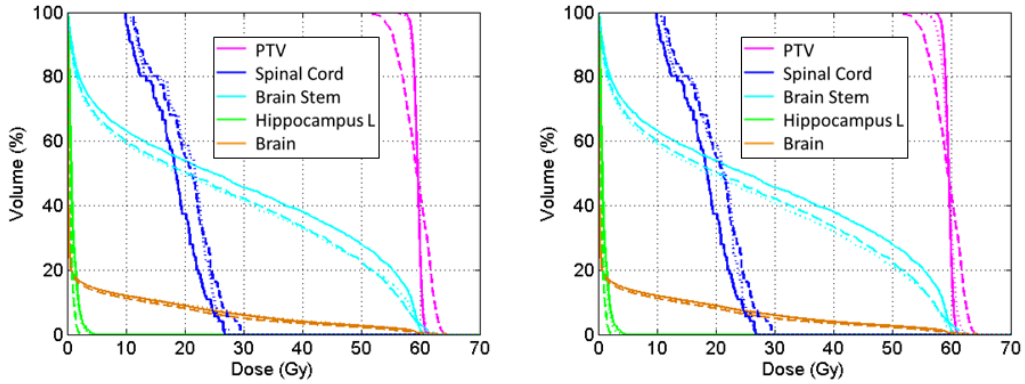
(a) iSC1 & oSC1 (b) iSC1& oSC2

(c) iSC2 & oSC1 (d) iSC2 & oSC2

Figure 3.6: Comparison of DVHs for MC doses for the prostate case, resulting from the reference full MC-based optimization (solid line) and hybrid optimization with different stopping criteria: the dashed lines represent the dose before correction (C=0), while the dotted lines correspond to the final corrected dose. The number of corrections needed for each set of stopping criteria is indicated in Table 5 (# iter).

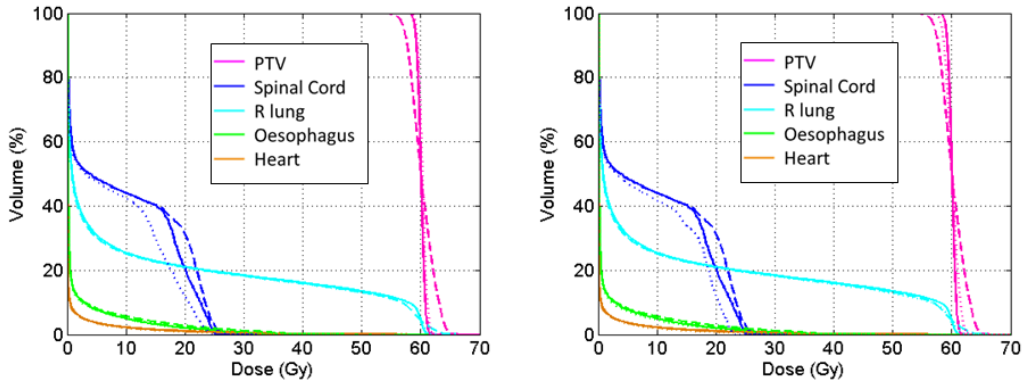


(a) iSC1 & oSC1 (b) iSC1& oSC2

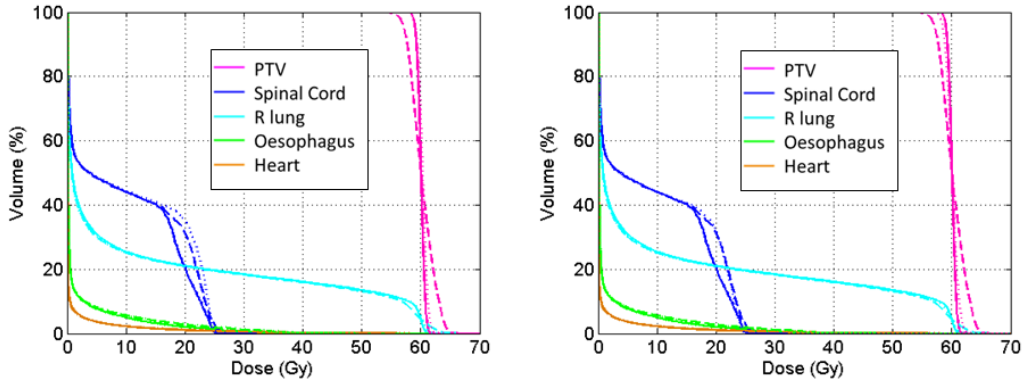


(c) iSC2 & oSC1 (d) iSC2 & oSC2

Figure 3.7: Comparison of DVHs for MC doses for the brain case, resulting from the reference full MC-based optimization (solid line) and hybrid optimization with different stopping criteria: the dashed lines represent the dose before correction (C=0), while the dotted lines correspond to the final corrected dose. The number of corrections needed for each set of stopping criteria is indicated in Table 5 (# iter).



(a) iSC1 & oSC1 (b) iSC1& oSC2



(c) iSC2 & oSC1 (d) iSC2 & oSC2

Figure 3.8: Comparison of DVHs for MC doses for the lung case, resulting from the reference full MC-based optimization (solid line) and hybrid optimization with different stopping criteria: the dashed lines represent the dose before correction (C=0), while the dotted lines correspond to the final corrected dose. The number of corrections needed for each set of stopping criteria is indicated in Table 5 (# iter).

Computation time

The time required to compute the accurate forward dose distribution ($t_{\ddot{D}}$) employed to calculate the correction matrix C was approximately 192 s for the prostate (5×10^7 particles), 46 s for the brain (10^7 particles), and 256 s for the lung (5×10^7 particles). The time employed to compute the accurate dose influence matrix ($\ddot{\ddot{P}} = P_{MC}$) with 10^5 particles per beamlet ($t_{\ddot{P}}$) was equal to 2346 s in the prostate case, 660 s in the brain case, and 4160 s in the lung case. This represents a reduction of a factor 14 on average ($t_{\ddot{P}}/t_{\ddot{D}}$) of the time spent in MC computation. This factor could be interpreted as the gain in computation time for the hybrid method with respect to the full MC optimization when only one correction ($\#$ iter $= 1$) is needed, provided that the time required to compute the approximate dose influence matrix ($t_{\tilde{P}}$) and the optimization time after correction are small enough. Arguments supporting these assumptions, together with some data summarizing the time employed in optimization and dose calculation for commercial TPS are presented in the Discussion section.

The computation time increases with the number of protons used in the MC simulation, as well as with the number of voxels traversed by the simulated protons. This explains the longer times obtained for prostate and lung cases; since the target was deeper than in the brain case, and more particles were simulated.

3.4 Discussion

The present study was divided in two parts. The first part consisted in analyzing the performance of the hybrid method when including different dose calculation errors. Each error was simulated individually in a controlled environment, using a single dose engine to avoid the dose discrepancies inherent to the use of two different algorithms, such as PB and MC, which could make the interpretation of results more difficult. The second part aimed to quantify the potential of using the hybrid method in clinical practice for IMPT treatments. Therefore, the hybrid method was applied to three clinical cases, simulating a realistic situation where two different dose engines were used: the MC code MCsquare and the PB algorithm in FoCa. The gain in computation time of the hybrid method with respect to the full MC-based optimization approach was also quantified.

The analysis of the four test cases revealed poor performance of the hybrid method for the cases involving shifts in the dose distributions, namely, range errors and lateral dose displacements. In these cases, the dose after correction were usually worse than those before correction ($C = 0$), even for small errors ($\rho = +1.5\%$,

x = 2.5 mm), which indicates a divergence of the algorithm. Dose degradation was especially strong when large shifts were introduced ($\rho = \pm 3\%, x = 5mm$). For instance, D_{95} decreased to 80% of the prescribed dose in the case with a lateral displacement of 5 mm (x = 5 mm), and to 59% of the prescribed dose in the case with a range undershoot of 3% ($\rho = +3\%$). In both cases, the target overdose led to D_5 that were larger than 120% of the dose prescription. This behavior can be explained by looking at the dose profiles in Figures 3.2, 3.3, 3.4, 3.5. The algorithm was unable to compensate for the lack of dose in the region of the shift. After the first correction ($C = 1$), a peak of high dose appeared just before the under-dosage region. In the following corrections ($C = 2-5$), the algorithm tried to get rid of this hot spot without success, which caused degradation of both target coverage and overdose. This is also illustrated in Figure A.1 (See appendix A).

On the other hand, the hybrid method performed very well in the test cases where the dose distribution did not suffer any shift, such as a constant scaling of the dose values or the convolution of the dose with a Gaussian kernel (dose blurring). In these cases, the hybrid method managed to perfectly correct for the differences between the approximate ($\tilde{D}$) and accurate dose ($\ddot{D}$), achieving $\ddot{D}$ with target doses of $D_{95} > 98.5\%$ and $D_5 < 103.5\%$ of the prescribed dose, regardless the magnitude of the error. Although larger errors ($\alpha = 0.90, \sigma = 0.5$) usually needed more corrections than smaller errors ($\alpha = 0.95, \sigma = 0.3$).

Previous studies about hybrid methods in conventional radiotherapy [142, 141, 103] concluded that good performance of the algorithm is ensured, provided that the differences between the two dose engines are small. However, our results demonstrated that on top of the magnitude, the type of error plays also an important role, since the algorithm fails completely in cases involving a shift of the dose distribution, even if the shift is small. This is an important issue for proton therapy plans, where the proton range estimation can easily differ from one dose engine to another. Therefore, proton range estimations must agree perfectly between the two dose engines used in the hybrid workflow, in order to ensure good performance. This means that the range of a proton beam passing through a homogeneous material of a given stopping power must be the same for both algorithms.

Applying the hybrid method to realistic patients gave us an insight about its potential use in clinical practice. The method exhibited outstanding performance for all patients, with final accurate dose distributions that always fulfilled the clinical constraints for the target volume and were very similar to the full MC plans used as reference. Note that in our workflow, the proton ranges from FoCa and

MCsquare were in perfect agreement for a given stopping power, satisfying the requirement stated above. Moreover, the dose kernels used by the PB algorithm to compute the approximate dose influence matrix ($\tilde{P}$) were computed with the same MC code (MCsquare) used for forward accurate dose calculation ($\ddot{\ddot{D}}$). This ensures a good performance of the algorithm, avoiding situations such as those presented in the test cases involving sharp range shifts.

The dose to OARs, as well as the number of corrections needed by the hybrid algorithm, depended strongly on the stopping criteria used. Since the goal is to use the hybrid method in clinical practice, a good trade-off between computation time (number of corrections) and quality of the plan must be found. According to our results, combining a convergence condition ($\epsilon < 0.005$) for the inner loop (*iSC1*) with an early stopping criteria ($D_{95}(PTV) \geq 0.95 D_{prescription}, D_5(PTV) \leq 1.05 D_{prescription}$) for the outer loop (*oSC2*), provided good quality plans after only 1 correction; with target D_{95} and D_5 that differed in less than 2% of the dose prescription with respect to the values for the full MC plan, and dose to OARs that were equivalent to the full MC plan. If the early stopping criteria based on the target volume metrics (D_{95}, D_5andD_{5-95}) were used for the inner loop (*iSC2*) instead of the convergence condition, the algorithm stopped also after one iteration but a higher dose to OARs was often obtained. This could be explained by the fact that our early stopping criteria concerned only the target dose. Therefore, it could be avoided by adding some extra dose constraints regarding the OARs in the stopping criteria.

If only one correction is needed to achieve a good plan with the hybrid method, as suggested by our results, the computation factor would decrease by a factor 14 on average with respect to the full MC approach. This is significantly bigger than the factor 5 achieved by Li et. al [90, 91] with their adaptive particle sampling scheme. Note that in our case, this factor was computed by dividing the time employed to compute the accurate dose influence matrix in the full MC approach ($t_{\ddot{\ddot{P}}}$) by the time needed to compute the accurate forward dose used in the correction matrix ($t_{\ddot{\ddot{D}}}$). Ideally, the time to compute the analytical dose influence matrix ($t_{\tilde{P}}$) and the time for subsequent optimization (t_{opt}) after correction (# iter > 0) should be added to $t_{\ddot{\ddot{D}}}$ to properly compute the gain factor. However, commercial TPSs usually achieve $t_{\tilde{P}} << 1$ min [49, 192], and total plan optimization can be also carried out in only a few seconds ($t_{opt} << 1$ min) [192], which can be neglected in comparison with the reported t_{opt}. Both contributions would depend on the total number of spots and the voxel size, but estimating as maximum $t_{\tilde{P}} + t_{opt} \approx 1$ min, the gain factor would be then 9, 6 and 13 for prostate, brain, and lung cases, respectively, which is still bigger than 5.

In this study, the number of particles that were simulated for accurate forward dose distribution ($\ddot{\ddot{D}}$) was chosen so that the target doses are computed with a mean relative statistical precision σ_{mean} of <0.8% and with a maximum σ_{max} below 1.5%, but additional time could be saved by relaxing the latter to $< 2\%$ [141] in combination with denoising techniques [76, 32]. Simulating fewer particles ($\sigma_{max} > 2\%$) is not recommended since it would reduce optimization precision, due to statistical noise caused by limited particle sampling [141, 73, 104, 140]. Lower precision would also compromise deterministic reproducibility of the results across different runs of the same plan.

For the sake of simplicity, optimization in this study relied on a PTV; although it is well-known that robust optimization is recommended in proton therapy in order to ensure robustness against treatment delivery errors such as patient setup or range errors, among others [98, 165, 128, 48]. One of the most popular strategies for robust optimization follows a worst-case approach [128, 48], where the objective function is evaluated for different error scenarios and each optimization steps considers only the currently highest objective value, corresponding to the worst-case scenario. The hybrid method could be adapted to such scheme of robust optimization, by simply computing a correction matrix for each scenario. The total computation time would be multiplied by the number of scenarios, but the relative gain factor with respect to full MC robust optimization would remain the same as reported in this study.

3.5 Conclusion

In this work, we analyzed the performance of PB-MC hybrid optimization for IMPT treatments, a simple strategy that combines the speed of analytical algorithms with the advantages of MC planning. Our results showed excellent performance of the hybrid method for the clinical cases under study, provided that the proton range estimations from the two dose engines involved are perfectly aligned.

The hybrid method can be easily implemented within most TPSs, with only minor changes to the existing optimization workflow, which makes it an appealing tool to be used together with commercial TPSs to get accurate plans without the integration of a MC dose engine within the TPS.

Chapter 4

Robust treatment planning for proton therapy

The second objective of this thesis is to develop software tools for robust treatment planning in scanned proton therapy, and in particular, that enables the simulation of other sources of uncertainty than those implemented in current commercial solutions. For that purpose, we developed a robust treatment planning system, MIROpt (Chapter 2) that is able to simulate systematic and random setup errors, proton range errors coming from image conversion issues, and breathing motion. This chapter is focused on the latter, breathing motion, although robustness against systematic setup and image conversion errors is also included. The use of a hybrid method, as the one presented in the previous chapter (Chapter 3), to speed up the modelling of dosimetric effects caused by the breathing motion is also investigated. The content is based on a manuscript in preparation that will be submitted soon: Efficient implementation of 4D robust optimization for scanned proton therapy with Monte Carlo.

4.1 Introduction

The finite range of protons enables the delivery of a conformal dose to the tumor. This is further enhanced by pencil beam scanning (PBS), where individually weighted proton beams ($\ll$ beamlets $\gg$ or $\ll$ spots $\gg$) are scanned to cover the tumor in all three dimensions. PBS can then perform intensity-modulated proton therapy (IMPT), which provides more degrees of freedom to shape the dose than conventional radiotherapy [101, 85]. IMPT plans result from an inverse optimization process where all beamlets from all fields are optimized simultaneously to find the weights that lead to an optimal dose distribution, also known as multi-field optimization (MFO).

IMPT treatments are especially sensitive to planning and treatment uncertainties, such as dose calculation errors, incorrect setup of the patient, and organ motion [99, 100]. Fortunately, the rapid advance of technology starts to provide tools to address these problems. On one hand, using Monte Carlo (MC) dose engines enable accurate dose calculation, especially for very heterogeneous tumor sites [120, 144]. On the other hand, applying robust optimization techniques ensures robustness against patient setup errors and proton range uncertainties coming from the conversion of the CT Hounsfield units to stopping powers, often referred to as image conversion errors [165, 128, 48, 98]. However, dealing appropriately with motion in IMPT treatments still remains a challenge [31, 18], due to two main issues: 1) the change in the radiological depth of the tumor due to motion, and 2) the interference between the beam dynamic delivery and the motion itself (*interplay effect*).The present study focus on the first issue, assuming that interplay effects can be compensated by using simple techniques such as

rescanning [18, 75, 191].

In external radiotherapy with X-rays, safety margins are typically used to ensure target coverage in the presence of motion. A popular approach consists in using 4DCT images to create an internal target volume (ITV) that encloses all possible tumor positions in the breathing cycle captured by the 4DCT. Another approach has been suggested which eventually reduces the irradiated volumes without compromising target coverage. This strategy consists in defining the margins around the geometric time-weighted mean tumor position, i.e. the so-called *midposition* (MidP) [187, 186]. In proton therapy, since neither ITV nor MidP margins are robust to changes in the radiological depth of the tumor, a common solution is to use range adapted ITVs for each beam direction [124, 81]. But this strategy only works if each field is optimized independently (*single field optimization, SFO*), at the expense of conformity and lung exposure due to the large irradiated volumes; and extending this range adapted margins to IMPT-MFO still remains difficult [53]. More sophisticated methods like gating or tracking has been also proposed [54, 52, 38]. But the latter requires complex 4D delivery to synchronize the plan to the actual patient motion during treatment, which is difficult to implement clinically.

Recently, Liu et al. [97] have developed an alternative 4D optimization strategy, without the need of synchronized delivery, that achieved robust plans against motion and reduced interplay effects. They computed the dose on each of the 4DCT phases and accumulated them in a reference phase (end-exhale) used during optimization. They also included patient setup and image conversion errors into the robust optimization approach. This 4D optimization strategy is very promising, but they used analytical algorithms to compute the dose, which still does not address the incorrect modeling of tissue heterogeneities [120]. This is especially crucial for mobile tumors located in regions with high density differences, such as lungs [55], where this study was performed. Therefore, MC-based 4D planning techniques are required in order to transfer all the advantages from IMPT to mobile tumors in the thoracic region.

4D MC codes have been available for many years in photons as well as in proton therapy. Many of them use the same strategy as in Liu et al. [97] of computing the dose on all phases and accumulating them in a reference image by means of deformable registration algorithms [77, 41, 64, 153]. Since the final result in MC dose calculation is the average after the simulation of many particles, it is associated with a standard deviation, which expresses the uncertainty due to the fact that the simulated number of particles is finite. In order to achieve a good

statistical precision (low standard deviation), MC simulations typically use a total number of particles (N) in the order of 10^7. Therefore, since the statistical precision of the result depends only on the total N, the advantage of this 4D MC method is that the dose on each phase can be computed with a reduced number of particles, provided that the accumulated dose would have a sufficient N [77]. More specifically, the statistical precision per voxel in the accumulated dose would depend on the number of energy deposition events on each voxel, which is independent on the number of phases accumulated. However, the statistical precision might be worse in regions with artefacts due to a large movement between the phases. In practice, we assume that this effect is small enough and can be neglected [77]. This involves that motion can be simulated for free in terms of computation time, since the total N is the same as for a regular 3D dose distribution. This advantage, together with the fact that the new generation of fast MC codes (dedicated to PBS) achieves very competitive times to be used clinically [74, 193, 127, 147, 105, 90, 91], motivates the development of a 4D MC-based robust optimizer. The first aim of this study is to develop and validate a robust 4D MC-based treatment planning system, which includes also setup and image conversion errors.

In addition, hybrid dose calculation methods have been proposed to reduce the computation time even further, for instance, by combining analytical and MC dose engines [141]. In this study, we hypothesize that the dosimetric effects of motion can also be compensated with an hybrid method that combines 3D and 4D dose calculations. Implementing and validating this hybrid method against full 4D treatment planning represents the second aim of this study.

4.2 Materials and methods

4.2.1 Treatment planning system

Optimization scheme

The plans were created with our in-house treatment planning system, MIROpt, which performs optimization of the spot weights by using libraries from the large-scale non-linear solver IPOPT [177], through its MATLAB interface.

The optimization problem follows the classical approach where the objective function is a weighted sum of quadratic terms that penalize the voxels with a dose that differs from the prescription [116]. Robust optimization is implemented as

a minimax problem, where the objective function for the worst-case scenario is minimized at each iteration [48]. The problem is then stated as follows:

$$\min_{x} \quad \max_{s \in S} \left\{ \sum_{r \in T} \frac{\alpha_r}{N_r} \sum_{i=1}^{N_r} ((\max\{0, \xi_r - D_i(w,s)\})^2) + \sum_{r \in T,OAR} \frac{\alpha_r}{N_t} \sum_{i=1}^{N_r} ((\max\{0, D_i(w,s) - \xi_r\})^2) \right\} \tag{4.1}$$

$$\text{subject to } w \geq 0$$

where w is the vector containing the spot weights, s denotes a certain scenario from set S and r denotes a region of interest (ROI) that is either a target ($\in T$) or an organ to spare ($\in OAR$). For each ROI, the penalty function is multiplied by a weighting factor α_r and divided by its number of voxels N_r, and it depends on the difference between the calculated dose D and prescription ξ_r.

The dose in the i-th voxel on scenario s is computed as

$$D_i(s,w) = \sum_j P_{ij}(s) \cdot w_j, \tag{4.2}$$

where $P(s)$ represents the *dose influence matrix* with the contribution from all spots in scenario s, being j the spot index.

Machine specifications

All the simulations were run on a computation server with 56 Intel® Xeon® CPU E5-2660 v4 of 14 cores each at 2.00 GHz, with 128 GB of RAM.

4.2.2 Monte Carlo 4D dose calculation

The fast MC code MCsquare [147], developed in-house, was used as dose engine inside MIROpt. MCsquare is designed and optimized to simulate proton PBS treatments in voxelized geometries, such as a CT image. It exploits both task and data parallelisms of modern processors, being able to simulate 10^7 protons in less than one minute on a calculation server, or less than 5 minutes on a single desktop. The details of the physics models, as well as the internal structure of the code have been previously published [147]. Here we concentrate on the description of

the 4D dose calculation method implemented in MCsquare and used for this study.

MCsquare performs a 4D dose calculation by computing the dose distribution on each phase of the 4DCT and accumulating them on a reference image by using deformable registration algorithms. In this study, the geometric time-weighted mean tumor position (midposition, MidP) [187, 186] was used as reference image (MidP-CT), and a log-domain diffeomorphic Morphon algorithm [71, 178] was used for deformable image registration. The whole process comprises the following steps:

1. In order to compute the MidP-CT, each phase ph_i of the 4DCT was non-rigidly registered to the first one ph_1 (end-inhale), which yields n inter-phase Deformation Vector Fields (DVF$(ph_i \rightarrow ph_1)$), being n the total number of phases. Then, the time-weighted average of all these inter-phase DVF was computed, which gives the displacement from ph_1 to the average anatomical position (i.e. mid-position, MidP), DVF$(ph_1 \rightarrow MidP)$. The DVF mapping each phase to the MidP (DVF $(ph_i \rightarrow MidP)$) is generated by subtracting DVF$(ph_i \rightarrow ph_1)$ from DVF$(ph_1 \rightarrow MidP)$. The resulting DVF$(ph_i \rightarrow MidP)$ is applied to each corresponding ph_i, and the MidP-CT is generated by taking the median value of each voxel from all deformed phases in the MidP position. This results in a MidP-CT image with a lower noise level and less artefacts than the 4DCT phases.

2. A regular 3D dose distribution was computed for each phase i (D^{ph_i}). The number of simulated particles per phase N_{ph_i} satisfied the following condition
$$N_{ph_i} \cdot n = N \tag{4.3}$$
where N is the accumulated number of particles from all n phases. N was computed so that the accumulated dose achieved a good statistical precision, which typically amounts to 10^5 for a single spot dose calculation [105, 90, 91] and 10^7 for a final dose distribution [74], in both cases using a common resolution dose calculation grid (i. e., 2×2×2 mm^3).

3. The same DVF$(ph_i \rightarrow MidP)$ computed in step 1) to generate the MidP-CT are applied to its corresponding 3D dose distribution D^{ph_i}. All deformed phase-specific 3D doses ($D^{ph_i}_{def}$) were accumulated, averaged by the number of phases n, to obtain the 4D dose distribution, i. e., $D^{4D} = \sum_i^n D^{ph_i}_{def}/n$ [89, 97]. This assumes that the patient spends equal time in each breathing phase, which was ensured by binning the 4D-CT datasets into n equally distributed temporal phases.

4.2.3 Robustness against patient setup and range errors

Systematic setup errors (SS) can be modelled as shifts of the beam isocenters. Thus, they were simulated by moving the spot weights to the neighbouring positions within the energy layers in the scanning grid and calculating the dose for the new weights using the same dose influence matrix as in the nominal scenario [165, 48]. Apart from the nominal (SS = 0 mm), six scenarios were considered, including shifts towards the positive and negative directions of the R-L (right-left), A-P (anterior-posterior) and I-S (inferior-superior) axes. The systematic shift was set to SS = 5mm for all patients, as often used in the literature [114, 48, 98, 96, 97, 67].

Proton range errors coming from converting Hounfield units (HU) into stopping powers, also known as image conversion errors, were simulated by including a uniform density perturbation (ρ) in the CT calibration curve [165, 48]. Three scenarios were taken into account: the nominal case, undershoot, and overshoot of all spots. The range perturbation was set to $\rho = +/-$ 3%, according with the proposed values in literature [120]. This amounts to a total of 21 scenarios, i. e., three scenarios to account for image conversion errors with their corresponding 7 systematic setup shifts.

4.2.4 Robustness against motion

Two Monte Carlo - based 4D optimization methods are presented and compared in this study. The first method, denoted as *full 4D optimization*, uses the 4D information during the optimization, by including the effects of motion directly on the dose influence matrix. The second method, denoted as *hybrid 3D-4D optimization*, uses a regular (3D) dose influence matrix during optimization and includes the 4D information a posteriori, by means of an iterative process. For both methods, a single, not phase-specific spot weights map was obtained after optimization. Patient setup and image conversion errors were included as described in sections 4.2.1 and 4.2.3. Thus, no planning target volume (PTV) margin was used, and the planning was done directly on the clinical target volume (CTV). The two methods are described below.

Full 4D optimization

For each spot j, the 3D dose distribution corresponding to each phase is calculated, deformed and accumulated on the reference image (MidP-CT), following the process described in section 4.2.2. This task is repeated for all spots, yielding a 4D dose influence matrix (P^{4D}). Since the total number of particles per spot is

the same as for a regular 3D dose influence matrix (P^{3D}), this strategy simulates the motion for free in terms of computation time (see equation 4.3). The only extra time with respect to computing P^{3D} is coming from the deformation and accumulation of the dose. Provided that the total number of spots can reach the several thousand, repeating this task for each spot could lead to an important difference of computation time that needs to be quantified.

During optimization, the dose on each scenario s will be then computed following equation 4.2, which in matrix form can expressed as the inner product of $P^{4D}(s)$ and the vector of weights w:

$$D(s, w) = P^{4D}(s) \cdot w \tag{4.4}$$

where s represents one of the 21 scenarios accounting for range and setup errors (Section 4.2.3). It is important to note that the P^{4D} matrix allows us to include the motion information in a single scenario, without increasing the number used to simulate range and setup errors.

A convergence criterion ϵ (eq. 4.5) based on the relative dose difference between two consecutive iterations k and $k-1$ [90, 91] was used:

$$\epsilon = \frac{\sqrt{\frac{\sum_{i=1}^{Nv}(D_i^{Nv} - D_i^{k-1})^2}{Nv}}}{\xi_{CTV}} \times 100\%, \tag{4.5}$$

where Nv represents the total number of voxels summed over all ROIs involved in the objective function and ξ_{CTV} is the dose prescription for the target volume (CTV). This criterion was only evaluated in the nominal scenario, and the optimization was stopped when ϵ had not exceeded 0.005 for the last 5 iterations.

Hybrid 3D-4D optimization

The hybrid optimization strategy applied in this study has been widely used in conventional radiotherapy to reduce computation time by combining two different dose algorithms, one being characterized by its accuracy and the other by its speed [142, 140]. Recently, our group demonstrated that this strategy provides also excellent results for PBS treatment planning (Chapter 3), and thus, correcting for dose calculation errors. In this work, we hypothesize that this hybrid optimization strategy can also be used to correct for the dosimetric effects caused by motion.

The proposed optimization workflow is presented as pseudo-code in table 4.1. Prior to optimization, a regular 3D dose influence matrix is computed on the

MidP-CT (P^{3D}), for each scenario s used to simulate range and setup errors (i.e., 21 scenarios). During optimization, this $P^{3D}(s)$ will be used to evaluate the 3D dose distribution ($D^{3D}(s)$) following equation 4.2, and a voxel-by-voxel correction term C_t that accounts for the 4D information will be added to get the corrected dose (D^c).

$$D^c(s) = D^{3D}(s) + C_t(s) \tag{4.6}$$

The correction term C_t for each scenario s is computed as the difference between a forward 4D dose computation and the 3D dose computed through P^{3D}:

$$C_t(s) = D^{4D}(s) - D^{3D}(s) \tag{4.7}$$

Note that the term *forward dose* denotes a MC simulation where the total number of particles N is properly distributed among all the spots according to their previously optimized weights (w). The D^{4D} is then obtained as explained in section 4.2.2, by computing a forward dose for each phase, deforming and accumulating all of them in the MidP-CT. Therefore, the task of deforming and accumulating the phase-specific doses needs to be performed only once for each scenario, i.e. 21 times. Updates of C_t can be triggered as often as necessary by running a forward D^{4D}, using as input the optimal values for the spot weights obtained from the last optimization loop (w_t). The index t denotes the number of updates performed so far ($t = 0, 1, 2, 3...$). The D^{3D} used in equations 4.6 and 4.7 is computed through the inner product of P^{3D} and w_t.

The whole process can then be seen as two nested loops: 1) iterate until reaching the optimal w values for a given matrix C with index t (inner loop), and 2) update the C matrix until the desired D^{4D} is achieved (outer loop). Note that C = 0 in the first iteration of the outer loop (t = 0).

Before each update t, the optimized corrected dose $D^{c,o}$ and the forward 4D dose D^{4D} are properly evaluated according to the established stopping criteria. If they are satisfied, w_t is saved as the final optimal solution for the spot weights map. Otherwise, a new optimization process is run with w_t as initial weights and using the fresh-updated C_t matrix.

According to the results reported in previous work (Chapter 3), a good trade-off between computation time and plan quality is obtained by combining a convergence condition for the inner loop (iSC) with an early stopping criteria, based on the target DVH for instance, for the outer loop (oSC). Therefore, for iSC we used the same convergence criterion ϵ as for the full 4D optimization method, given by equation 4.5. For oSC, an early stopping criteria based on the target dose metrics for the worst-case scenario was used, requiring $D_{95} \geq 0.95 D_{prescription}$, $D_5 \leq 1.05 D_{prescription}$.

(1)	Initialize $C_t = 0$ $(t = 0)$
(2)	Optimize using D^c
(3)	If inner stopping criteria (*iSC*) achieved, save w_t and the optimized $D^{c,o}$
(4)	Compute $D^{4D}(w_t)$ and C_t
(5)	If $D^{4D}(w_t)$ and $D^{c,o}$ satisfies the outer stopping criteria (*oSC*),return w_t; else, $t = t + 1$ and go to (2)

Table 4.1: Hybrid optimization pseudo-code. Instruction 1 initializes the algorithm, Instructions 2-3 correspond to the inner optimization loop, while Instructions 4-5 represent the outer optimization loop.

4.2.5 Plan evaluation

Robustness of the plan against systematic setup, image conversion and motion errors was evaluated by computing a final (forward) 4D dose distributions (D^{4D}) for 81 scenarios, i.e., the 21 scenarios used during optimization plus diagonal combinations of the systematic setup errors. The process to compute each D^{4D} was the same as described in section 4.2.2. Dose volume histograms (DVHs) were computed for each structures of interest in all scenarios, and plotted as a DVH-band, which represents the envelope enclosing all scenarios. Hence, the edges of such DVH band correspond to the worst case dose distribution.

Regarding target coverage, a plan was considered acceptable when the worst case D_{95} was bigger than 95% of the dose prescription, and the worst case D_5 was smaller than 105% of the dose prescription. For the organs at risk, we looked mainly at the V_{20} for the lungs and D_2 for the spinal cord.

4.2.6 Clinical case

The two 4D robust optimization methods described in section 4.2.4 were applied to the lung case shown in figure 4.1. The tumor is located in the lower part of the right lung, partly attached to the mediastinum. The CTV had a volume of 91.7 cm^3 and its average motion amplitude was 13.31 mm (R-L: 1.28 mm, A-P: 1.53 mmm, and I-S: 12.67 mm). A dose of 60 Gy was prescribed to the CTV and three beams were used to create the plan, with couch angles equal to zero and gantry angles equal to 145, 180 and 215 ^{o}. A range shifter of 5 cm water equivalent length (WED), mounted on the treatment couch, was used for the treatment. It was simulated by delineating a rectangular structure of HU $=$ 0 in the planning CT (Figure 4.1b). The dose grid resolution was equal to 2.5×2.5×2 mm in all cases.



(a) Coronal view (b) Axial view

Figure 4.1: Lung case used for the study, the CTV is shown is contoured in yellow in both coronal (a) and axial (b) views.

The 4D-CT comprised 10 respiratory phases and was acquired while patients were audio-coached to regularize their breathing, and thus, reduce 4D-CT image artefacts. The 4D-CT datasets were retrospectively binned into 10 equally distributed temporal phases, for motion management purposes [187].

4.3 Results

4.3.1 Plan quality and robustness

The two presented methods, *full 4D* and *hybrid 3D-4D optimization*, achieved excellent plan quality and robustness against the considered errors (systematic setup, image conversion, and motion). In both cases, the target worst-case D_{95} was above 95% of the prescribed dose (D_p) and the worst-case D_5 was below 105% D_p, which were the stopping criteria established for the outer loop (oSC) of the hybrid method (section 4.2.4), as well as the overall evaluation criteria (section 4.2.5). The specific values for these metrics (D_{95}, D_5), and those for the considered organs at risk (V_{20} for the right -ipsilateral- lung and D_2 for the spinal cord), are presented in table 4.2. The worst-case V_{20} for the right lung was always under 27%, which is far below than the recommended 30-35% [108], and the worst-case D_2 to spinal cord was always under 13Gy, which is much better than the reported maximum tolerance [80]. The DVH-bands for the two plans are presented in figures 4.2 and 4.3.

For the hybrid plan, the dose distribution before correction (C=0) did not satisfied the evaluation criteria for the worst-case target coverage, since D_{95} was only

93% of D_p. However, after applying the correction term C, the hybrid method was able to recover the worst-case target D_{95} up to 95.5%, which demonstrates the ability of this method to correct for the dosimetric effects caused by motion. Only one update of the correction term was needed, since the stopping criteria for the outer loop (worst-case $D_{95} \geq 95\%D_p$ and $D_5 \leq 105\%D_p$) were met at the first iteration. Surprisingly, the dose before correction presented very good target dose metrics in the nominal case ($D_{95} = 98.7\%D_p$ and $D_5 = 102\%D_p$), which indicates that including systematic setup and image conversion errors only, provides already some kind of robustness against the motion induced errors.

The results for the *full 4D optimization* were slightly better than those obtained for the *hybrid 3D-4D* method. The most important differences were found on 1) the worst-case D_{95}, which was 0.9Gy higher in the full 4D plan than in the hybrid one, and 2) the dose to the spinal cord, where the D_2 was about 3Gy lower in the full plan than in the hybrid, for both nominal and worst-case scenarios. The target overdose was very close in both cases (D_5 differed only in 0.3Gy), and the right lung V_{20} was also very similar in the two plans (only 0.7% difference).

4.3.2 Computation time

In the full 4D optimized plan, the 4D dose influence matrix (P^{4D}) was computed in 4h 42min, while the 3D dose influence matrix (P^{3D}) used for the hybrid 3D-4D plan took 3h 17min to be computed. This makes a difference of 1h 25min, which can be associated to the time spent in deforming and accumulating the phase-specific doses for all spots in the plan. Taking into account that the total number of spots (the same for the two methods) was 22 535, the time needed to perform this task only once can be estimated to be around 0.23s.

On the other hand, computing the forward 4D dose distributions (D^{4D}) used in the correction matrices C in the hybrid method, took around 45 min, including all 21 scenarios accounting for setup and range errors. This comprises the time spent in deforming and accumulating the phase-specifc doses. Therefore, the time needed to compute a single correction term for a given scenario s can be estimated to be around 2 min.

Method		Scenario	D_{95}(CTV) (Gy)	D_5(CTV) (Gy)	D_2(Spinal c.) (Gy)	V_{20}(R Lung) (%)
Full 4D		nominal	59.4	61.4	8.6	22.7
		worst	58.2	62.2	9.6	26.3
Hybrid 3D-4D	Before correction (C=0)	nominal	59.2	61.2	10.0	22.4
		worst	55.8	61.7	11.1	25.8
	After correction	nominal	59.0	61.4	12.3	23.3
		worst	57.3	62.5	12.7	27.0

Table 4.2: DVH metrics for the target ($D_{95}(CTV)$ and $D_5(CTV)$) and the main organs at risk (D_2 to spinal cord and V_{20} for the right lung), for both the *full 4D* and the *hybrid 3D-4D optimization* algorithms. In the latter case, the results before (C=0) and after correction are presented. Only one update of the matrix C was needed to satisfy the established stopping criteria. Note that the dose prescription to the CTV was equal to 60 Gy.

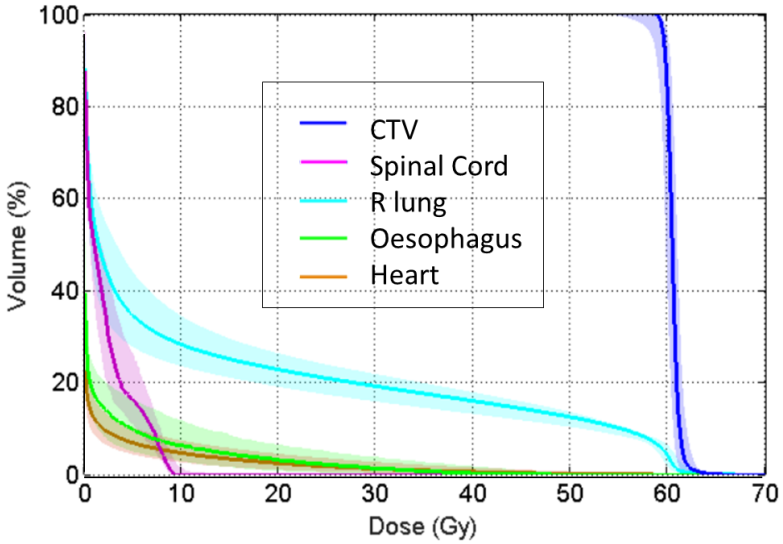


Figure 4.2: DVH-band for the target and organs at risk of interest, obtained for the full 4D optimized plan.

4.4 Discussion

In this work, two different methods (*full 4D optimization* and *hybrid 3D-4D optimization*) to achieve robustness against motion, were presented and validated for a lung case. Robustness against systematic setup and proton range errors was also included. Both methods make use of 4D Monte Carlo dose calculation, therefore ensuring accurate results, which is crucial for heterogeneous tumor sites such as

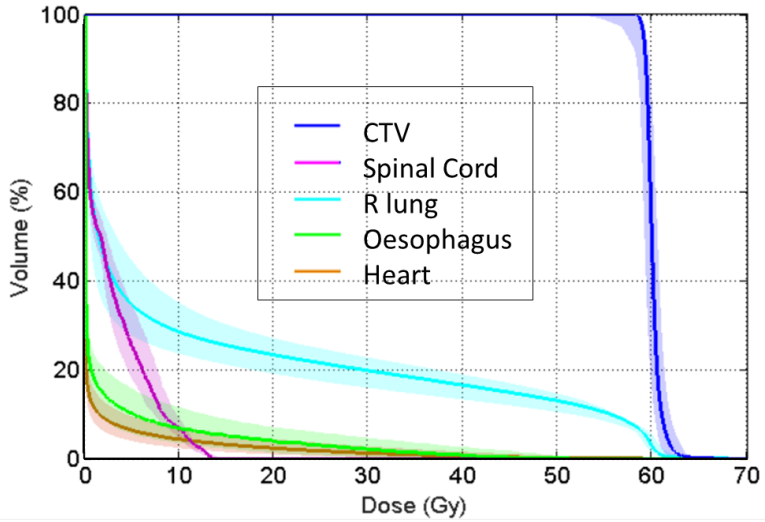


Figure 4.3: DVH-band for the target and organs at risk of interest, obtained for the hybrid 3D-4D optimized plan, after applying only one correction, i.e., a single iteration of the outer loop.

the lungs.

Each of the two methods includes the 4D information at different levels during the optimization process. The full 4D method uses this information already in the dose contributions from each spot, therefore, it requires to have a 4D MC code integrated in the treatment planning system (TPS). In contrast, the hybrid 3D-4D method optimizes using only 3D dose distributions, and incorporates the 4D information a posteriori. This method could be then used in combination with any robust TPS, as a post-processing tool to include robustness against motion. In addition, we have demonstrated in a previous work (Chapter 3) that the same hybrid algorithm can be used to correct for dose calculation errors, coming from the use of analytical algorithms such as pencil beam (PB). Thus, one could combine both features to create a hybrid 3D PB - 4D MC algorithm, to get advantage of the speed of PB, on the one hand, and the accuracy and efficiency of MC methods for 4D dose computation in heterogeneous media, on the other hand. Nevertheless, it is well-known that the hybrid method encounters some limitations when the size of the correction is too big [140, 103]. The performance and limitations of this type of hybrid 3D PB - 4D MC algorithm is under investigation in our group.

The results, however, showed small differences between the performance of the two methods. Indeed, although the plan quality and robustness was very good for

both plans (nominal and worst $D_{95}(CTV) \geq 95\%D_p$, $D_{95}(CTV) \leq 105\%D_p$) and the dose to the organs at risk was very low (nominal and worst $V_{20}(R.\,lung) \leq 27\%$, $D_2(Spinal\ c.) < 13Gy$), the results for the worst-case in the hybrid plan were sightly worse than those for the full one. In particular, the CTV D_{95} was 0.9Gy lower than the one for the full 4D method. This could be explained by the selected stopping criteria for the outer loop, since the optimization was terminated when the target dose satisfied the conditions $D_{95}(CTV) \geq 95\%D_p$ and $D_{95}(CTV) \leq 105\%D_p$. Better results, and thus closer to the full 4D plan, could be achieved by using a harder stopping criteria. However, this might entail the need of more than one correction, and as a consequence, the computation time would increase. We considered that the aforementioned stopping criteria provided a good balance between computation time and robustness, but if desired, the user could be easily change this criteria inside MIROpt. The second significant difference was found on the dose to the spinal cord, with a D_2 that was 3Gy higher in the hybrid plan than in the full one. Again, this could be a consequence of the selected dose objectives. In both plans, a dose objective of maximum dose (eq.4.1) equal to 25Gy was used, therefore, as long as the dose to the spinal cord was below this value, the corresponding penalty was invisible for the optimizer. One could reduce the dose even further by tuning the dose objectives in between corrections of the hybrid method.

Surprisingly, before applying the correction (C=0) in the hybrid plan, the nominal case presented a very robust target coverage. As mentioned, this indicates that including systematic setup and image conversion errors only, provides already some degree of robustness against the dose perturbation caused by the breathing motion. This was observed already by previous studies [96], where the fact of applying regular (not 4D) worst-case optimization produced motion-resistant treatment plans. Similar findings were also reported in Liu et al. [97] regarding the interplay effect. In this case, an identical approach to the full 4D optimization described in this study was applied, therefore, not accounting for interplay effects. However, their results showed that the full 4D robust optimization produced plans with reduced interplay effects. This suggests that the dosimetric effects coming from different sources of error merge during treatment planning, and therefore, one must treat the errors as a combination instead of doing it independently, as done in the current worst-case optimizers. Hence, strategies for scenario reduction and combination should be implemented to provide a more realistic and efficient way of robust optimization. In addition, this could also help to reduce the computation time and resources needed, in comparison with the current robust optimization techniques, which involves a set of typically 20 or more scenarios.

Our method is already an example of scenario reduction, since the 4D information is included in a single scenario, without increasing the number already used to simulate setup and range errors. This is very advantageous in comparison with the method implemented in the only commercial TPS including 4D optimization (RayStation, RaySearch Laboratories, Sweden). In their approach, each respiratory phase is considered as an independent scenario, which leads to large size optimization problems, where the number of scenarios is multiplied by the number of phases, requiring very long computation times and memory. Moreover, the fact that all phases are used independently might lead to very conservative results. However, the advantage of this method is that it does not depend on deformable registration algorithms.

Regarding more practical details, the two methods (*full 4D* and *hybrid 3D-4D*) rely on the hypothesis that breathing motion is perfectly represented by the 4DCT images taken at the planning stage, assuming no inter- or intra-fraction deviations. Therefore, special care should be taken to ensure as much as possible a reproducible breathing pattern, e.g., through coaching as in our case, or with more sophisticated techniques, such as mechanical ventilation [125]. Also, the 4D-CT datasets were retrospectively binned into equally distributed temporal phases, which allows to easily accumulate the phase-specific deformed doses by averaging the total result with the number of phases ($D^{4D} = \sum_i^n D_{def}^{ph_i}/n$).

The hybrid 3D-4D method was faster than the full 4D one. However, the difference was not significant, since taking into account the time needed to compute the 3D spot dose contributions (3h 17min), and the correction term containing the 4D information for the 21 scenarios (45 min), the final gain factor was only about 1.15. This is the best that one can get, since only one correction was performed. As mentioned, applying the hybrid method would be of notable interest if the 3D spot dose contribution matrix was computed by analytical algorithms, which can perform this task in a few minutes of even seconds [49, 192].

The type of clinical case also deserves some discussion. The patient used in this study presented a rather big tumor with median to large amplitude of motion (13 mm). However, the performance of the hybrid method for patients presenting much larger motions, as well as smaller target sizes, still need to be investigated. This study is a work in progress, together with the performance of the hybrid 3D PB - 4D MC algorithm.

4.5 Conclusion

This work presented two methods (*full 4D* and *hybrid 3D-4D*) to include MC based 4D optimization in a robust (worst-case scenario) optimizer, accounting also for setup and range errors. Both methods created a single, not phase-specific plan, avoiding the need for gating or synchronizing the proton beam dynamic delivery with the target motion. This makes its clinical implementation very easy, provided that more or less constant breathing motion can be assured. The computational resources needed for this 4D robust optimization were not bigger than for the current robust optimizers, since motion is simulated for free in terms of computation time when using MC and the 4D information was integrated in a single scenario.

Chapter 5

Patient specific therapy through dose painting by numbers

Dose painting is a very appealing concept that might enable a more personalized therapy through the delivery of a heterogeneous prescription that targets the patient-specific biological features of the tumor. Combining dose painting with proton therapy might allow for a very conformal dose escalation, with reduced normal tissue complications and increased tumor control. This is the third objective of this thesis, and the subject of this chapter. In particular, we study the feasibility and robustness of dose painting by numbers with scanned proton therapy, using a commercial treatment planning system. The content presented here comes from the published article: Feasibility and robustness of dose painting by numbers in proton therapy with contour-driven plan optimization

5.1 Introduction

Unsuccessful local tumor control is one of the main reasons for cancer treatment to fail. In order to improve the probability of local tumor control and/or reduce side effects, dose painting aims to deliver a non-uniform dose distribution to tumor volumes according to biological information provided by functional or molecular images [25, 157, 150, 24, 17]. IMRT can potentially deliver dose distributions that match complex dose prescriptions. But proton beams can be modulated even more finely than photons [132] and can spare organs at risk (OARs) more easily [160, 151, 112, 28, 169], which may allow for a higher dose intensification or escalation, like in dose painting . Combining both the potential of dose painting and proton therapy may help reach a better trade-off between disease control and toxicity.

The idea of using biological information conveyed by images to prescribe the dose [93] has led to two main approaches in the literature: selective sub-volume boosting or dose painting by contours (usually only one or two sub-volumes are selected), and dose painting by numbers (DPBN) [16]. Dose painting by contours prescribes a uniform dose "boost" within the image-defined contours, while DBPN adjusts the dose prescription in each voxel independently, with respect to the corresponding voxel intensity in the registered functional image. The feasibility of DBPN has been demonstrated for proton therapy in a few studies where in-house software was used to reproduce the non-uniform dose prescription [132, 160, 43, 42]. However, one of the remaining challenges of using DPBN is that most commercial treatment planning systems (TPS) do not allow optimization on the voxel scale and require contours or structures instead. More specifically, contours on each slice define a 3D structure (or volume). Each voxel in the structure is then subject to the same prescription and dose constraints. Such a framework naturally favors dose uniformity, because voxels cannot be constrained

independently. Nevertheless, for rotational IMRT treatments, several studies have demonstrated that DPBN prescriptions can be matched closely with an optimizer driven by contours, if an appropriate number of subcontours are defined to guide optimization towards the voxel-wise prescription [33, 86]. However, for proton therapy, no similar study has been conducted so far. Hence, this work aims to prove the feasibility of DPBN with proton therapy using optimization driven by contours in a recent TPS (RayStation, RaySearch Laboratories, Sweden). For this purpose, the ability of the TPS to reproduce a dose gradient typical for DPBN will be maximized by determining the TPS settings for protons (scanning pattern, beam configuration and number of subcontours), which best reproduce the dose prescription.

Similarly, the question of coping with delivery uncertainties (setup and range errors) in dose painting with protons remains mostly unaddressed. For photon-based modalities, previous studies have demonstrated that a minimal dose can be ensured in the presence of geometric errors inside the DPBN volume by either using robust optimization [185] or incorporating directly the errors into a "robust prescription" [152]. The latter strategy, like the classical approach with a PTV margin, relies on the "static dose cloud approximation". This approximation, generally valid for photons, does not hold for protons due to range uncertainties [7, 95]. Nevertheless, previous studies about dose painting with protons have still used a PTV margin for the external contour and even ignored setup and range errors inside the painted volume [132, 160, 43, 42]. In this work, a dilated prescription (PTV approach) [152] and robust "minimax" optimization [48] were applied separately in order to account for setup and range errors within the entire dose painting volume. The ability of both approaches to deal with range and setup uncertainties could then be assessed.

5.2 Materials and methods

5.2.1 Target definition

Two patients with locally advanced oropharyngeal squamous cell carcinoma (T4b N0 M0 for patient 1 and T4 N2c M0 for patient 2) were selected for the study. Based on the work of Daisne et al. [30], where PET was found to be the most accurate modality to delineate the primary tumor (GTV), pre-treatment ^{18}FDG-PET acquisitions provided functional images. Moreover, previous studies have suggested that FDG is a suitable tracer for dose painting [5].

PET images were acquired on a time-of-flight (TOF) Gemini TF PET-CT

(Philips Medical Systems, Cleveland, OH, USA). The performance of the machine is detailed by the manufacturers in a previous publication [156], as well as a spatial resolution of 4.8 mm (FWHM). But the FWHM measured in clinical conditions was about 6 mm. Both patients received 300 MBq of FDG and were scanned 1 hour post-injection on two bed positions of 150 seconds each one. The PET data were reconstructed with the 3D line of response-TOF blob-based OSEM algorithm from Philips, with three iterations and 33 subsets. The voxel size was 2 x 2 x 2 mm^3.

The external contour of the dose painting volume (primary tumor GTV_{PET}) was outlined in the ^{18}FDG-PET images using an automatic gradient-based method, which has demonstrated its superiority with respect to threshold-based methods [50]. Selection and delineation of clinical target volumes (CTV) and OARs were performed in the planning CT, following guidelines by Grégoire et al. [56, 57].

5.2.2 DPBN prescription

Standardized uptake values (SUV) within GTV_{PET} were converted to dose values using a linear relationship starting from the median uptake value (SUV_{median})(Figure 5.1). Due to partial volume effect [143], uptake values ranging from the minimum (SUV_{min}) to the median can be considered to result mainly from contamination of the lower-uptake background. Therefore, the median then appears as a reasonable estimator of the "baseline" uptake in the tumor, with limited influence of both the coldest and hottest spots. A non-uniform escalation from 70 Gy (for the median FDG uptake value SUV_{median}) to 86 Gy (for the maximal uptake value SUV_{max}) was prescribed, consistently with Vanderstraeten et al [175], Witte et al. [185], and Sterpin et al. [152]. The choice of a relative conversion function, with fixed bounds for dose escalation, aims to keep dose within safe clinical values. The lower bound corresponds to the dose prescribed homogeneously in current clinical protocols. The maximum is dictated by toxicity reports [107].

In order to guide the optimizer to approximate the voxel-by-voxel prescription, a number N of nearly equidistant and non-crossing subcontours were defined inside the GTV_{PET}, dividing it into N sub-volumes (Figure 5.2). SUV values from the ^{18}FDG-PET images were used to calculate the dose constraints: each sub-volume delimited by the *i*-th subcontour ($i = 1...N$) was prescribed a minimum dose equal to

$$70 + (D_{max} - D_{min})(\frac{thr_i}{thr_N}) \ ,$$

where D_{max} = 86 Gy, D_{min} = 70 Gy and *thr* stands for the SUV thresholds:

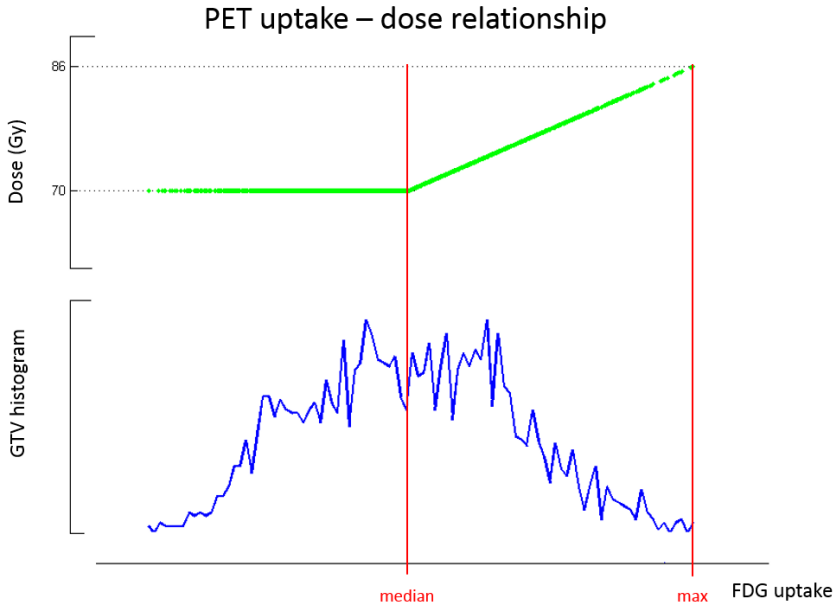


Figure 5.1: Histogram for GTV$_{PET}$ in patient 1 and linear relationship starting at the median uptake value used to calculate the dose in Gy.

thr$_{1,...i,...N}$ = SUV$_{median}$,...SUV$_i$,...SUV$_{max}$. The maximum dose constraint for the i-th sub-volume equaled the minimum dose constraint for the next contour (i+1). The whole process was carried out with in-house developed software.

5.2.3 Plan robustness against setup and range errors

Two different sets of plans were generated: the first were optimized on PTVs ("PTV-based") while the second entailed robust optimization on CTVs ("robust CTV-based").

In the PTV-based plans, uncertainties are taken into account with classical PTV margins. All target contours, including the ones in the GTV$_{PET}$ and all serial OARs were expanded to the corresponding PTVs and PRVs by 2.5 mm, using the usual formalism for homogeneous prescriptions [174, 111]. The dilation of the DPBN volume has been discussed previously for photons [152].

In the robust CTV-based plans, the "minimax" optimization method available in RayStation (research version 3.99) was used to account for uncertainties. Here,

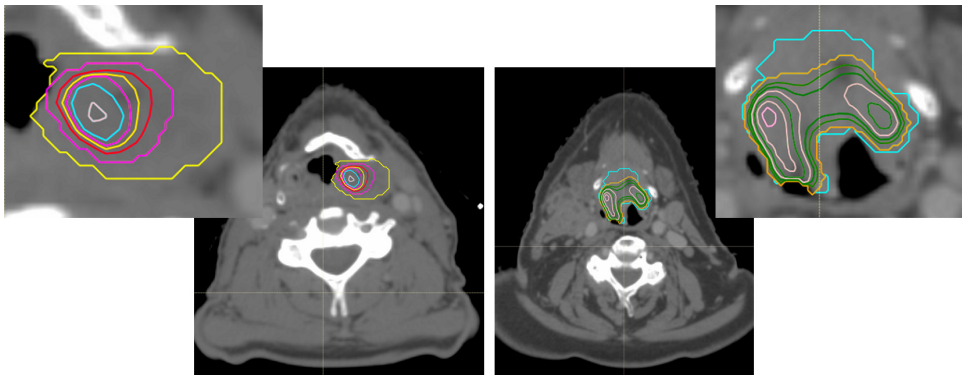


Figure 5.2: GTV_{PET} divided into seven subcontours for patient 1 (left) and 2 (right). Dose inside GTV_{PET} is expected to increase from contour to contour, from 70 to 86 Gy.

the PTV is no longer necessary and the dose is prescribed directly to the CTVs (and GTV_{PET} including subcontours). Doses to serial OARs were also taken into account in the robust optimization. The optimizer computes the objective function (the discrepancy between the prescribed and planned doses) for several error scenarios and minimizes it for the worst case [48]. Setup errors are simulated by shifting the beam isocenters, i.e., by moving the spot weights to adjacent positions within the isoenergy layers for each beam. Range uncertainties were simulated by a uniform change of the density. Setup shift was chosen to be 2.5 mm (equaling the margin used for the PTV-based plans), and the uncertainty on the density was set to 3% [7, 48, 166].

5.2.4 Planning objectives

The total treatment was divided in 35 fractions. For the high-dose CTVs (therapeutic CTVs, CTV_{70}), a uniform prescription of 70 Gy was applied, whereas a uniform 56 Gy dose was prescribed for the low-dose CTVs (prophylactic volumes, CTV_{56}). Regarding the heterogeneous (DPBN) prescription, an escalation from 70 to 86 Gy was prescribed to the GTV_{PET} volume according to the formalism explained in section 5.2.2.

Plans were considered to be acceptable if target coverage satisfied the established clinical constraints: 95% of the target volumes (D_{95}) received at least 95% of the prescribed dose. For the OARs, traditional constraints used in our center for radiotherapy plans were applied as maximum limit: $D_{max} < 48$ Gy for spinal cord and brain stem, $D_{mean} < 30$ Gy for the oral cavity and $D_{mean} < 26$ and 30 Gy for contralateral and ipsilateral parotid glands, respectively. For critical OARs in HNC (spinal cord and parotid glands), the latter dose constraints were

pushed down as much as possible, while keeping target coverage within clinical constraints.

5.2.5 Plan assessment

For uniform prescriptions, treatment plan quality is generally assessed using dose volume histograms (DVH) and typical metrics like D_{95}, D_2 and D_{mean}. Such metrics no longer apply to non-uniform prescriptions. In order to address this issue, quality volume histograms (QVHs) [175] were computed in the GTV_{PET} and PTV_{PET} for robust CTV-based and PTV-based plans, respectively. A QVH is the decumulative histogram of the ratio $Q_i = D_i^{planned}/D_i^{prescribed}$, where $D_i^{planned}$ and $D_i^{prescribed}$ are the planned and prescribed doses in the *i*-th voxel. The ideal QVH is a descending step function, where the relative volume V equals 100% for Q_i smaller than 1 and 0% for Q_i larger than 1.

Plan quality was assessed with the QVH metrics $V_{Q=95\%}$ (%) and $V_{Q=105\%}$ (%), which reflect target coverage and overdosing, respectively. $V_{Q=XX\%}$ is the percentage of volume which receives at least $XX\%$ of the prescribed dose. Hence, ideal plans achieve $V_{Q=95\%} = 100\%$ and $V_{Q=105\%} = 0\%$. The quality factor QF ($QF = \frac{1}{n}\Sigma_i |Q_i - 1|$) averages the deviation of the ratios Q_i in the volume [175] and therefore reflects the spread of the QVH curve around 1.

5.2.6 Design of the study

After sub-segmentation of the dose painting volume (GTV_{PET} for robust CTV-based plans and PTV_{PET} for PTV-based plans), subcontours were translated into a standard prescription, with DICOM-RT structures that are compatible with the chosen TPS (RayStation from RaySearch Laboratories, Sweden).

Proton pencil beam scanning was the modality used to create the treatment plans. The beam model used for the optimization is based on beam data, which is representative of an IBA dedicated pencil-beam nozzle (Ion Beam Applications, Louvain-la-Neuve, Belgium). A 2 x 2 x 2 mm^3 resolution was used for dose calculation.

The study was divided in two parts. The first part aims at determining the optimal parameter values to reproduce accurately and efficiently a DPBN prescription inside a given volume. The second part applies these values to the actual clinical plans, where all the targets and OARs are considered for optimization and geometric uncertainties are taken into account. Treatment delivery time was also quantified for some plans.

Testing of planning parameters for DPBN plans

In order to simplify the problem, an ideal case was considered where geometric uncertainties were neglected and only the primary tumor (GTV_{PET}) and non-dilated subcontours were used for the optimization. OARs were not considered for the optimization in this part.

Number of beams and number of subcontours

According to previous studies for photons [33, 86], one of the parameters that influence the most the final DPBN prescription is the number of subcontours involved in the optimization. Therefore, the effect of this parameter in the final dose distribution is tested by increasing the number of subcontours from 3 to 11.

Regarding the beam configuration, the number of beams used clinically in a classical (homogeneous prescription) IMPT treatment for HNC patients is usually only two or three [169, 44, 94]. Although studies with five, seven, and even nine beams have been published [99, 155, 151], it has been demonstrated that for homogeneous IMPT treatments, there is little improvement in plan quality when increasing the number of beams [151]. However, since DPBN prescriptions are highly heterogeneous, increasing the number of beams incidences could provide the optimizer with more degrees of freedom to better match the DPBN prescription. Indeed, Seco et al. [138] reported an improved conformality when increasing the number of beams. Therefore, the number of beams was set to 3, 5, or 7 (Figure 5.3 a, c, d). Coplanar and equally spaced gantry angles were used, starting from a posterior beam (180^o), rather than an anterior one (0^o) to avoid crossing the oral cavity [151, 44, 155, 182]. Range shifters were not used in this part.

For all possible combinations of beam incidences (n_b = 3, 5, 7) and number of subcontours (n_{sc} = 3, 5, 7, 9, 11), conventional (non-robust) optimization was performed in RayStation and the corresponding QVHs in the GTV_{PET} volume were calculated.

Spot scanning grid

Other important parameters are those that determine the spot pattern: spot size, spot spacing, and layer separation. Scanning patterns with smaller distances between spots and smaller spot sizes could lead to tighter dose gradients and more accurate "painting." The effect of some of these features in the dose distribution has already been investigated in previous studies [151, 182, 14]. For classical

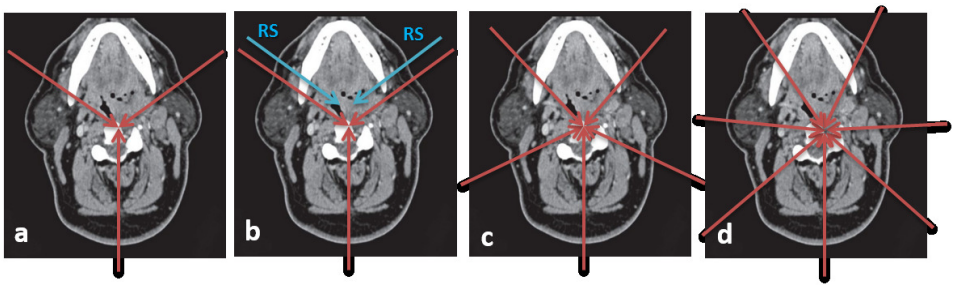


Figure 5.3: Beam configurations: 3, 5, and 7 equally spaced fields without range shifter (RS) (a, c, d) and 3 beam plus two additional with RS in the frontal beams (b).

(homogeneous) HNC treatments, Widesott et al. [182] found that the maximum acceptable sigma (σ) of the initial Gaussian size of the proton pencil beam was $\sigma \geq 4$ mm, while a spacing of 5 mm between spots was the best. Regarding the distance between layers, a constant spacing of 4 [160] or 5 mm^{WED} [43, 155, 182] has been used in the literature.

For the IBA machine used in our study, σ equaled 3 mm (in air at isocenter at 230 MeV). This is our sole spot size commissioned, which should be small enough according to previous studies [160, 43]. Concerning the spot pattern settings, a "dynamic" approach versus different constant configurations [160, 151, 43, 155] were investigated.

The dynamic ("automatic") option in RayStation enables dynamic selection of distances between spots and layers at the beginning of the optimization in order to have a smooth dose with as few spots as possible. This means that in layers where the spots are large, the spot spacing will be higher. The layer separation is also non-constant in depth. In the "constant" option, a uniform spot spacing ($\Delta x = \Delta y = SS$) within each layer and constant layer separation (LS) in depth is chosen by the user.

We assume that the number of beams and subcontours, as well as the spot pattern, are variables that influence independently in the final DPBN gradient. Therefore, for a beam configuration fixed to three beams (Figure 5.3.a) and seven subcontours, six constant patterns (SS = 3.5, 4 and 5 mm, for LS = 4 and 5 mm^{WED}) versus the automatic approach were analyzed.

Feasibility and robustness of DPBN plans

The "automatic" option for the scanning pattern parameters (dynamic SS and LS) was used for all the plans. The number of subcontours was fixed to 7.

All the targets and OARs were considered for optimization, trying to reach the planning objectives described in section 5.2.4, and minimizing the dose to the spinal cord and to the parotid glands as much as possible. As a consequence, high in-field dose gradients may appear close to these regions, thus increasing the sensitivity to uncertainties [7].

The number of beams could influence plan robustness, since range uncertainties would be spread out over many beams and beam angles [138, 61]. In order to analyze this feature, different beam configurations with 3, 5 and 7 beam incidences were considered.

For both patients, the CTV_{56} was close to the surface in the anterior-oblique directions. Therefore, the effect of a range shifter was investigated by creating another plan (with 5 beams) where range shifters (Lucite 40 mm^{WED}) were added to two of the frontal beams (Figure 5.3.b).

For all four different beam configurations (Figure 5.3), the two sets of plans (PTV-based and robust CTV-based) were generated for each patient as described in section 5.2.3, in order to test robustness in both strategies. For robust CTV-based plans, robustness was enabled for the objective functions associated with all targets (CTVs), subcontours, and serial OARs (spinal cord and brain stem). For other OARs, like parotid glands or oral cavity, the robust option was disabled.

For each plan, robustness was evaluated by computing the dose in different error scenarios. Positive (+2.5 mm) and negative (-2.5 mm) shifts of the beam isocenters were simulated in the anterior-posterior (A-P), superior-inferior (S-I), and lateral (R-L) directions. For each shifted dose, 3% overshoot and undershoot of all Bragg peaks (-3 and +3% density perturbation, respectively) were calculated. All these cases plus the nominal ones (no shift and/or 0% density perturbation) amount to a total of 21 scenarios.

Treatment time

The delivery time for each beam was evaluated in six plans of patient 2: two plans for the patient planned with OARs (section 5.2.6, PTV-based with 3 and 7 beams) and four plans for the GTV_{PET} only (section 5.2.6, dynamic scanning grid for 3

and 7 beams; and constant SS = 4mm & LS = 4mm^{WED}, and SS = 4mm & LS = 5mm^{WED}). It is noteworthy that it estimates only the time required by the proton therapy system to deliver the fraction, excluding gantry rotation.

5.3 Results

5.3.1 Planning parameters for DPBN plans

Dose painting volume segmentation and FDG distribution

The segmented volume for the primary tumor (GTV$_{PET}$) obtained from the automatic gradient-based method was 21.27 cm^3 for patient 1, and 12.82 cm^3 for patient 2. Figure 5.2 shows a cranio-caudal slice for each patient where the GTV$_{PET}$ is divided in 7 subcontours. For patient 2, two dose escalation peaks occur in the lateral direction (right-left). This can also be observed in Figure 5.4, where the corresponding PET image with the FDG uptake distribution is shown. For patient 1, the FDG uptake is more centered in the lateral direction but also two peaks are observed in the cranio-caudal direction.

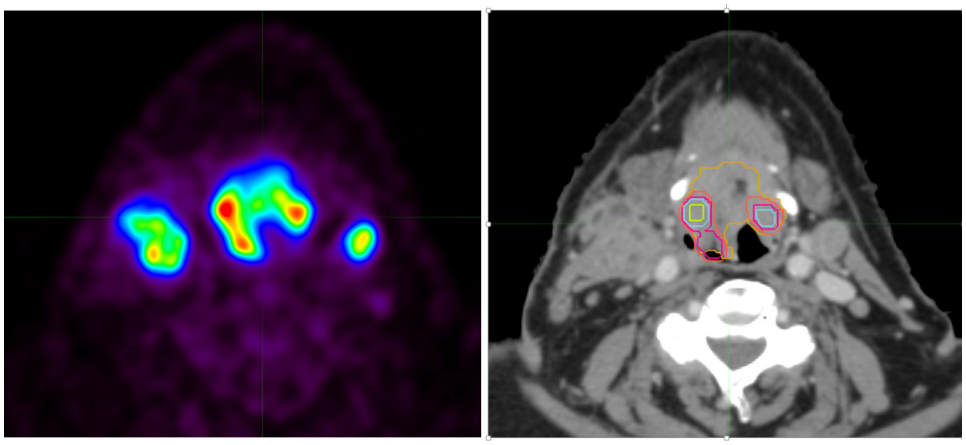


Figure 5.4: PET and CT images with 7 subcontours for patient 2. FDG uptake distribution is not uniform and presents two peaks for dose escalation.

Number of subcontours and number of beams

Table 5.1 shows the QVH metrics V$_{Q=95\%}$, V$_{Q=105\%}$ and QF values, for increasing number of subcontours and beams. The differences between QFs values for 3, 5, and 7 beams, given a fixed number of subcontours, are within 0.2%. This

hardly affects on the final QVH curve, as shown in Figure 5.5.

On the other hand, increasing the number of subcontours clearly impacted the QVHs. In general, $V_{Q=95\%}$ remained almost the same ($V_{Q=95\%} > 99\%$) while target overdose was strongly reduced. Going from 3 to 11 subcontours, $V_{Q=105\%}$ decreased from 11-13% to less than 0.5% for patient 1 and from 7-8% to less than 0.3% for patient 2. The largest difference was observed when going from 3 to 5 subcontours (Figure 5.5), with an improvement in the QF of about 1%. Beyond 5 or 7 subcontours, the QF increases, which means that painting quality starts deteriorating when too many sub-volumes are used, though not significantly (within 0.1%). According to Table 5.1, the optimal number of subcontours is almost equivalently 5 or 7 for patient 1, and 7 for patient 2. For both patients, the best QF is obtained for a combination of 7 subcontours and 7 beams.

Scanning parameters

For all considered scanning patterns, the differences in QF values were less than 0.05% in most cases, which means that all of them allow for accurate painting (Table 5.2.

The dynamic approach led to layer spacings that decreased continuously from 8 mm^{WED} for the high-energy layers, to 2 mm^{WED} for low-energy layers. The spot spacing varied from 5 to 7 mm, depending on the spot size. For this strategy, the number of spots (4997 and 3661 for patient 1 and 2, respectively) was reduced with respect to the constant grids; except for the case with SS = 5 mm and LS = 5 mm^{WED}, which used even less spots (4702 and 3546, for patient 1 and 2). In general, the number of spots for patient 2 is smaller than for patient 1 because the GTV_{PET} volume is smaller.

Regarding the number of energy layers, the dynamic option used more than 60 layers, whereas 56 and 46 were used on average by the LS = 4 mm^{WED} and LS = 5 mm^{WED}, respectively.

For both patients, the best QF was obtained for a grid with constant SS = 4 mm and a LS = 4mm^{WED}. But the total number of spots for this configuration (9045 and 6864 for patient 1 and 2, respectively) almost doubled compared to the dynamic approach.

PATIENT 1						
No. of subcontours		3	5	7	9	11
3 beams	QF (%)	2.26	1.10	1.13	1.22	1.18
	$V_{Q=95\%}$ (%)	99.81	99.98	99.94	99.93	99.93
	$V_{Q=105\%}$ (%)	13.28	1.57	0.66	0.90	0.47
5 beams	QF (%)	2.12	1.00	1.03	1.03	1.02
	$V_{Q=95\%}$ (%)	99.92	99.98	99.98	99.99	99.98
	$V_{Q=105\%}$ (%)	12.90	1.18	0.55	0.45	0.39
7 beams	QF (%)	2.09	0.98	0.92	1.02	1.07
	$V_{Q=95\%}$ (%)	99.87	99.99	100.00	99.99	99.96
	$V_{Q=105\%}$ (%)	11.25	1.15	0.29	0.41	0.39
Averaged QF		2.15	1.03	1.03	1.09	1.09
PATIENT 2						
No. of subcontours		3	5	7	9	11
3 beams	QF (%)	1.97	1.23	1.01	1.05	1.07
	$V_{Q=95\%}$ (%)	99.95	99.95	99.95	99.96	99.95
	$V_{Q=105\%}$ (%)	8.43	1.48	0.45	0.33	0.29
5 beams	QF (%)	1.89	1.13	0.96	0.93	1.02
	$V_{Q=95\%}$ (%)	99.99	99.99	99.99	99.99	99.99
	$V_{Q=105\%}$ (%)	7.40	1.31	0.58	0.31	0.27
7 beams	QF (%)	1.81	1.11	0.86	0.91	1.01
	$V_{Q=95\%}$ (%)	99.99	100.00	99.99	100.00	100.00
	$V_{Q=105\%}$ (%)	6.96	1.53	0.53	0.41	0.28
Averaged QF		1.89	1.17	0.94	0.96	1.03

Table 5.1: QVH metrics for all plans generated with increasing number of subcontours (3 to 11) and increasing number of beams (3, 5 and 7). The averaged QF for each number of subcontours is also presented in the last row.

5.3.2 Feasibility and robustness of DPBN plans

All nominal plans satisfied the initially established clinical constraints in section 5.2.4, i.e. 95% of the targets volumes ($PTV_{56,70}$ for PTV-based plans and $CTV_{56,70}$ for robust CTV-based plans) received more than 95% of the prescribed

PATIENT 1							
Layer spacing - LS (mm)	4	4	4	5	5	5	dynamic
Spot spacing - SS (mm)	3.5	4	5	3.5	4	5	dynamic
QF (%)	1.04	1.00	1.06	1.06	1.04	1.11	1.13
No. of spots	11665	9045	5702	9624	7438	4702	4997
No. of layers	55	55	55	45	45	46	63
PATIENT 2							
Layer spacing - LS (mm)	4	4	4	5	5	5	dynamic
Spot spacing - SS (mm)	3.5	4	5	3.5	4	5	dynamic
QF (%)	0.96	0.95	0.96	0.98	0.98	1.02	1.01
No. of spots	8780	6864	4289	7274	5678	3546	3661
No. of layers	59	57	58	48	46	48	62

Table 5.2: Total number of spots and energy layers for all different scanning grids used. All plans were created with a configuration of 3 beams and 7 subcontours. QVH were calculated for each plan and the QF values are also displayed in the table.

dose: D_{95}^{PTV56}, D_{95}^{CTV56} >54 Gy and D_{95}^{PTV70}, D_{95}^{CTV70} > 67 Gy.

Data for D_{95} to CTV_{56} and CTV_{70}-GTV_{PET}, D_{mean} to parotid glands, and the dose to 2% of the spinal cord volume (D_2) is shown in Table 5.3, for nominal (regular/roman font) and worst case (italicized font).

Regarding the OARs, nominal doses are the same for PTV-based and robust CTV-based plans (differences in dose lower than 0.5 Gy) and they could be reduced significantly while maintaining target coverage (Table 5.3). D_2 to spinal cord was reduced to 12 Gy for patient 1 and 14 Gy for patient 2, which represent $1/4$ of the initial constraint (D_{max} <48 Gy). Concerning the parotid glands, the dose could be much more reduced for patient 1 than for patient 2, because the targets in the former overlapped a big part of the parotids. In patient 1, D_{mean} equaled 14 and 8 Gy for the left (L) and right (R) parotid, respectively. In patient 2, D_{mean} for L and R parotids was 14 and 24 Gy, respectively.

No significant improvement was observed when changing the beam configuration, neither in terms of OAR sparing nor in target coverage (differences in dose lower than 0.1 Gy in most cases).

Robustness test

Robustness evaluation revealed the degradation of target coverage when including range and setup errors in the set of scenarios. As an example, Figure 5.6 shows the DVH bands for CTVs and OARs for the plan with 3 beams and 7 subcontours, for conventional (left) and robust optimization (right) in patient 1.

For PTV-based plans, the difference between nominal and worst case in D_{95} for CTV_{56} and CTV_{70}-GTV_{PET} grew up to 5 Gy (Table 5.3). As a consequence, target coverage for the worst case was below the clinical constraints [1] for all plans generated: $D_{95}^{CTV56} \approx$ 50.5-53 Gy and $D_{95}^{CTV70-GTVPET} \approx$ 63-66.3 Gy.

In contrast, for robust CTV-based plans, the worst-case dose to target volumes was always above or very close to the clinical constraint: $D_{95}^{CTV56} \approx$ 53.5-54 Gy and $D_{95}^{CTV70-GTVPET} \approx$ 65.5-66.7 Gy. In cases it was below, the difference with respect to the constraint did not exceed 1 Gy, which can be considered still as robust coverage. No improvement in robustness with respect to target coverage was observed when increasing the number of beams.

Concerning the worst case for the dose to OARs, they still satisfied the clinical constraints largely, for both PTV and robust CTV-based plans. Only in the case of the R parotid gland for patient 2, the dose ($D_{mean} \approx$ 28 Gy) was close to the clinical constraint (D_{mean} <30 Gy).

As shown in Figure 5.6, robustness of OARs for PTV-based plans and robust CTV-based plans was very similar in both cases: differences in worst-case doses were below 1 Gy, for a given beam configuration (Table 5.3). Only in this part we see a minor trend with the number of beams, since the smallest worst-case dose to OARs was always for the plans with 5 or 7 beams. However the differences with respect to the plan with only 3 beams were about 1 Gy or smaller, which can be considered as negligible when doses to OARs are largely below the clinical constraint.

[1] Clinical constraints for target volumes are $D_{95}^{CTV56} >$ 53.2 Gy, and $D_{95}^{CTV70} >$ 66.5 (See Section 5.2.4)

Dose painting volume

No clear relationship between the QF and the number of beams was observed (differences in nominal QF of maximum 0.3%, Table 5.4).

The nominal QVH for PTV-based plans on PTV$_{PET}$ was obviously better than the nominal QVH for robust CTV-based plans on GTV$_{PET}$, since we are comparing directly with the dilated prescription (Figure 5.7 , Table 5.4). The difference concerned overdosing mainly, which leads to a difference of 0.6-0.8% in QF. However, in order to compensate for setup and range errors, the dilated prescription (PTV-based plans) generated a high overdose on GTV$_{PET}$: V$_{Q=105\%}$ = 35-38% for patient 1 and V$_{Q=105\%}$ = 29-31% for patient 2. On the opposite, for robust CTV-based plans, overdosing was reduced to V$_{Q=105\%} \approx$ 7% at maximum, while maintaining target coverage (V$_{Q=95\%} \approx 100\%$). Overdosing was interestingly reduced even further for plans with 5 and 7 beams. The reduction was bigger for patient 1 (V$_{Q=105\%}$ ≈4%) than for patient 2 (V$_{Q=105\%} \approx$ 6.5%).

Looking at the spread of the QVH bands shown in Figure 5.7, it is clear that deterioration of target coverage in PTV-based plans is much more pronounced than for the homogeneous CTVs, since heterogeneous prescriptions are expected to be more sensitive to uncertainties. QVH for worst cases in PTV-based plans presented a V$_{Q=95\%}$ that decreased to 94% for patient 1, and to 90% for patient 2. Moreover, they presented an important increase of overdosing since values up to V$_{Q=105\%}$ = 62% and 73% were obtained in patients 1 and 2, respectively. However, for robust CTV-based plans, worst case achieved V$_{Q=95\%} >$ 96% while keeping V$_{Q=105\%}$ lower than 15%.

For PTV-based plans, differences between the nominal and worst QVH on GTV$_{PET}$ grew up to 3% in QF for patient 1, and up to 4% for patient 2. For CTV-based robust plans, the difference never exceeded 1% for both patients.

5.3.3 Treatment time

Switching the energy layer required 2 seconds (s), while changing the spot position took less than 1 ms. The time to deliver a single spot depends on its weight (i.e, the MUs associated for the spot), but generally it was around a few ms (2-3 ms). The number of layers per beam varied between 20 and 40, as shown in Table 5.2.

For plans with only GTV$_{PET}$, the computed total time for 3 beams was 2.3 and 1.9 min for a constant configuration of SS = 4mm & LS = 4mm^{WED} and SS = 4mm & LS = 5mm^{WED}, respectively. For the dynamic configuration, the time for

3 beams was also 2.3 min, which increased up to 6.4 min when 7 beams were used.

In plans with all target and OAR structures (PTV-based, dynamic scanning grid), treatment time was 5.4 min for 3 beams and 12.7 min for 7 beams.

Therefore, the factor that influences the most treatment time is the number of layers, and consequently, the number of beams since a new beam will add between 20 and 40 layers to the total.

5.4 Discussion

To our knowledge, this study is the first attempt to test feasibility of dose painting by numbers with protons in a commercial treatment planning system, considering setup and range errors within the entire patient, including the painted volume. The patients in the study had complex non-uniform FDG uptake distributions. Two peaks for dose escalation in the cranio-caudal and lateral directions were present for patient 1 and 2, respectively.

The first part of the study investigated the influence of several parameters on the capability to accurately reproduce DPBN prescriptions. These were 1) the number of beams, 2) the number of subcontours used to approximate the DPBN prescription, and 3) the selected spot pattern (spot spacing and layer separation).

Only minor differences were observed for the beam number and spot pattern. Concerning the latter, it was not necessary to reduce SS and LS below 5 mm although previous studies of dose painting used smaller spot spacing or energy layer separation (Flynn et al. used SS = 3.5mm and LS = 5mm [43], and Thorwarth et al. employed SS=5mm and LS = 4mm [160]). For the two considered patients, the spot pattern associated with the fastest treatment time could thus be selected. The factor that influences the most the final treatment time is the number of energy layers, since the number of spots increases when adding layers. In addition, the switching time between two layers is about 2 s. Consequently, varying the number of beams from 3 to 7 more than doubled treatment time.

Concerning quality, the number of subcontours impacted the final QVH. The optimal number of subcontours was found to be around 5 and 7 for both patients. For 3 contours, target overdosing increased because the number of contours was not sufficient to guide accurately the optimizer to reproduce the DPBN prescription. Above 7 subcontours, plan quality slightly decreased. This could be caused by conflicting assignments of several prescription levels to the same voxel, when

the number of subcontours is such that their distance goes below the voxel size. Smaller voxels could improve the QF values, at the expense of longer optimization times. The optimal number of subcontours may also depend on the gradient, since the steeper the gradient in a region the more contours are needed to reproduce accurately the prescription. However, the fact that we use a fixed dose range from 70 to 86 Gy makes the gradient dependent on the volume, i.e., in small volumes we will have less space to escalate from 70 to 86 Gy, and consequently the gradient will be more pronounced. Therefore, the initial number of 7 subcontours could also be refined for each individual patient, after preliminary tests taking into account the target volume and uptake gradient.

The second part of the study tested the feasibility and robustness of dose painting in a realistic situation, including all targets and OARs in the optimization, and simulating both setup and range uncertainties in two strategies, either with a PTV or robust optimization on the CTV. The influence of the number of beams was also investigated.

No important differences in target coverage were observed when increasing the number of beams. Although all nominal plans easily met the clinical constraints, robustness evaluation showed that target coverage could be seriously degraded in PTV-based plans. Indeed, the worst-case D_{95} to CTV_{56} and CTV_{70} were up to 3 Gy below the clinical constraint for all PTV-based plans, whereas for robust CTV-based plans, D_{95} fulfilled it in most cases.

Regarding the OARs, both nominal- and worst-case received doses that were largely below the initial established constraints. Contrarily to target coverage, where CTV-based plans were much more robust than PTV-plans, the dose to OARs was equally robust for both set of plans. However, only the spinal cord was optimized robustly. Because the spinal cord lies relatively far from target structures in both patients, range errors will not impact significantly this OAR. The conclusions could be different if an OAR was in the vicinity of the tumor volume.

As regards to the dose painting volume, in order to compensate for setup and range error and keep the coverage to 100%, nominal QVHs on GTV_{PET} for dilated prescriptions (PTV-based plans) had a very high overdose ($V_{Q=105\%} > 50\%$). In contrast, the nominal QVH for robust CTV-based plans led to rather low overdosing ($V_{Q=105\%} < 8\%$). The differences in robustness between the PTV- and CTV-based plans were bigger than for the homogeneous CTVs. The worst case for PTV-based plans yielded QVHs with poor target coverage ($V_{Q=95\%} < 95\%$) and extremely high overdosing ($V_{Q=105\%} > 50\%$). Consequently, the use in proton

therapy of a dilated prescription, following a PTV approach, cannot cope appropriately with errors, because it entails non-robust and excessive dose distributions. On the opposite, robust CTV-based plans led to worst cases with fairly good GTV_{PET} coverage ($V_{Q=95\%}$ > 96-98%) and moderate overdose ($V_{Q=105\%} \leq$ 15%). Hence, robust optimization allows dose escalations to be robust with respect to both target coverage and overdosing.

The fact that PTV approach is not appropriate to account for errors in proton therapy, especially with pencil beam scanning, has been demonstrated by recent studies for classical (homogeneous) treatments [7, 95, 48, 94, 99, 98]. Our results showed that this effect is even stronger in dose painting since a high heterogeneous dose with high in-field dose gradients is delivered, making the painted volume more sensitive to uncertainties.

5.5 Conclusions

This work demonstrated that DPBN in proton therapy is feasible with a commercially available TPS (RayStation from RaySearch Laboratories in this case). It just requires the definition of a proper number of subcontours (around 7), typical scanning parameters, and only 3 beam incidences. Such settings still allow for relatively short treatment times (around 1-2 min per beam).

The use of a PTV fails to account appropriately for setup and range errors in dose painting since it leads to dose distributions that are not robust and present excessive overdose in the painted volume. In contrast, robust optimization yields dose distributions that are robust for both target coverage and overdosing.

As to the OARs, both approaches, with either a PTV or a robust CTV, generated plans with doses to OARs that were equally robust and far below the clinical constraints. This means that low doses to OARs can be achieved with proton therapy even in the case of high dose escalations, like in DPBN. The proposed methodology aims at bridging the gap between research in the field of DPBN, especially with proton therapy, and its implementation in clinical practice, thereby facilitating the execution of future clinical trials. It is generic enough to apply to any contour-driven TPS that comes equipped with robust optimization.

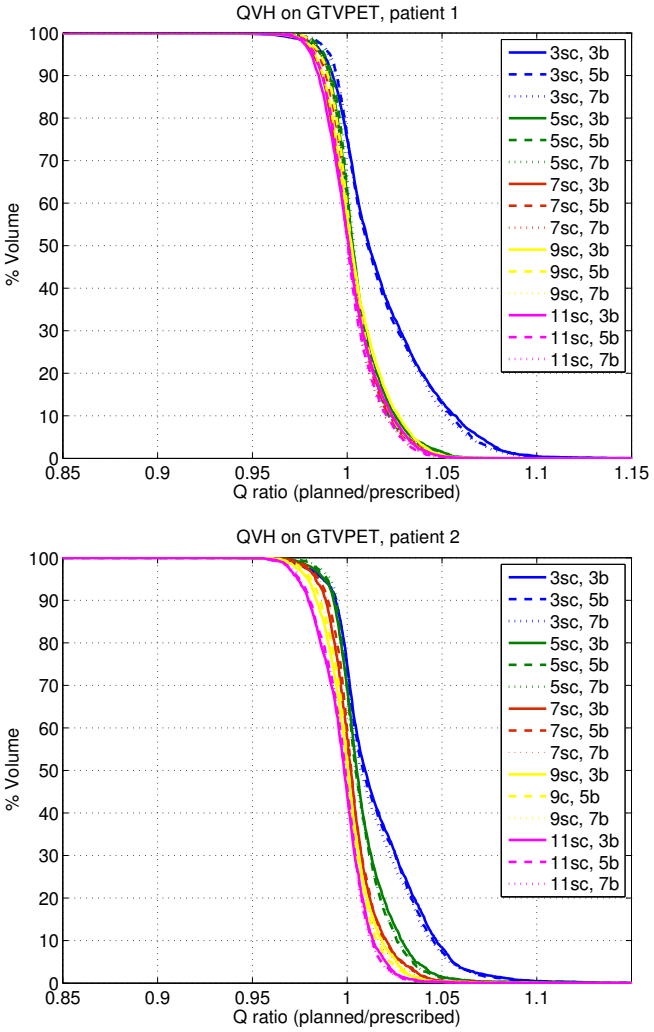


Figure 5.5: QVH curves for all plans generated with increasing number of subcontours (3 to 11) and increasing number of beams (3, 5 and 7), for patient 1 (left) and 2 (right).

	PATIENT 1				PATIENT 2			
Beam setting	3	5	7	3 + 2RS	3	5	7	3 + 2RS
PTV-based								
CTV_{56}	55.33	55.47	55.48	55.37	54.83	54.92	55.01	54.90
D_{95} (Gy)	*50.47*	*50.88*	*51.58*	*51.55*	*52.92*	*51.85*	*52.10*	*53.30*
CTV_{70}-GTV_{PET}	69.35	69.46	69.47	69.37	69.34	69.36	69.46	69.30
D_{95} (Gy)	*64.66*	*63.34*	*64.47*	*64.20*	*66.14*	*63.57*	*65.46*	*66.27*
Spinal cord	11.64	12.55	11.70	11.93	14.76	18.35	14.31	14.36
D_2 (Gy)	*14.30*	*13.72*	*13.28*	*13.62*	*15.61*	*14.76*	*15.33*	*15.29*
Parotid L	14.31	14.27	14.27	14.32	14.17	14.10	14.12	14.16
D_{mean} (Gy)	*19.14*	*18.54*	*18.45*	*18.86*	*18.44*	*17.51*	*17.29*	*17.84*
Parotid R	8.15	8.14	8.13	8.16	24.60	24.38	24.25	24.48
D_{mean} (Gy)	*11.08*	*10.94*	*10.81*	*10.93*	*28.73*	*28.34*	*27.78*	*28.26*
robust CTV-based								
CTV_{56}	55.05	55.23	55.23	55.19	54.28	54.50	54.42	54.23
D_{95} (Gy)	*53.68*	*54.14*	*54.03*	*54.27*	*53.43*	*53.90*	*53.58*	*53.56*
CTV_{70}-GTV_{PET}	67.69	68.17	68.28	67.73	67.57	67.76	67.59	67.57
D_{95} (Gy)	*66.41*	*66.66*	*66.72*	*66.38*	*65.36*	*66.47*	*66.14*	*65.56*
Spinal cord	12.82	12.14	12.17	12.39	15.02	15.22	13.53	13.14
D_2 (Gy)	*15.85*	*14.45*	*14.33*	*15.08*	*16.01*	*16.71*	*15.27*	*14.92*
Parotid L	14.24	14.21	14.19	14.21	14.16	14.12	14.14	14.12
D_{mean} (Gy)	*18.67*	*17.82*	*17.83*	*18.48*	*17.92*	*17.23*	*16.91*	*17.57*
Parotid R	8.13	8.13	8.10	8.12	24.60	24.37	24.49	24.53
D_{mean} (Gy)	*10.95*	*10.69*	*10.52*	*10.87*	*28.39*	*28.02*	*27.76*	*28.05*

Table 5.3: Dose data for nominal (regular font) and worst case (italicized font) for target volumes (CTV_{56}, and CTV_{70} where the volume for dose escalation - GTV_{PET}- was subtracted) and OARs (mean dose to parotid glands and the dose to 2% of the spinal cord volume (D_2)).

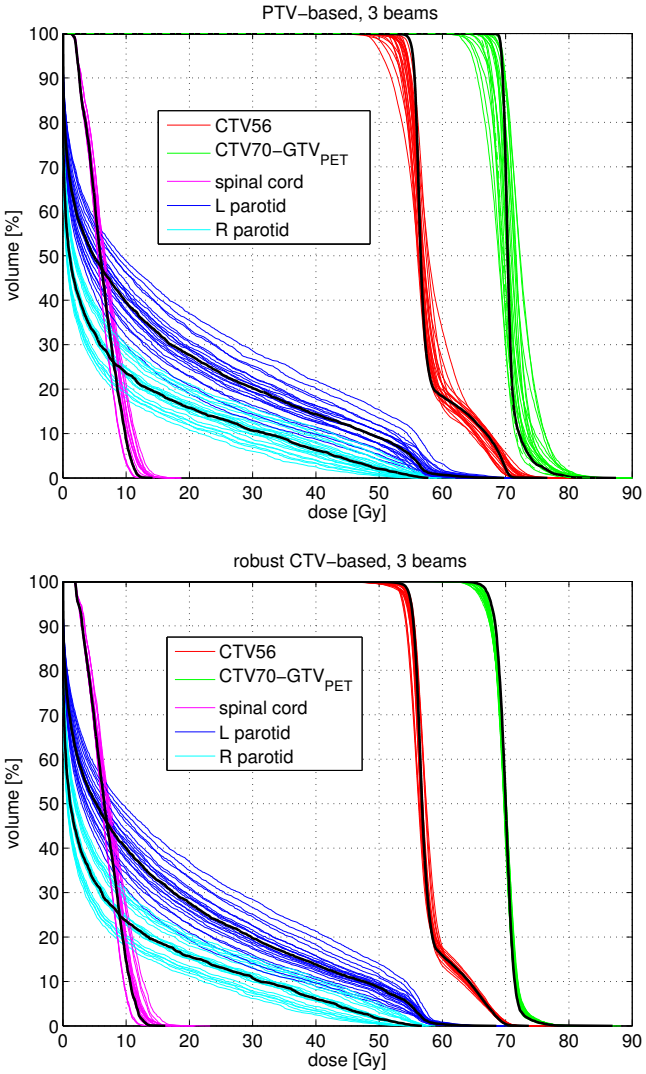


Figure 5.6: DVH bands for a configuration of 3 beams and 7 subcontours for patient 1, for PTV-based conventional optimization (left) and robust optimization (right).

		PTV-based							
		PATIENT 1				PATIENT 2			
Beam setting		3	5	7	3 + 2RS	3	5	7	3 + 2RS
nominal (on PTV)	QF (%)	1.13	0.92	0.89	1.11	1.22	1.14	1.10	1.27
	$V_{Q=95\%}$ (%)	99.98	100	100	99.98	100	100	100	100
	$V_{Q=105\%}$ (%)	0.49	0.22	0.16	0.56	0.68	0.43	0.25	0.71
nominal (on GTV)	QF (%)	4.27	4.39	4.38	4.27	3.68	3.75	3.86	3.65
	$V_{Q=95\%}$ (%)	100	100	100	100	100	100	100	100
	$V_{Q=105\%}$ (%)	35.96	37.76	38.01	36.19	28.88	30.73	31.70	29.24
worst (on GTV)	QF (%)	6.02	7.08	7.03	6.01	5.60	7.95	6.25	6.23
	$V_{Q=95\%}$ (%)	97.24	94.44	93.82	97.20	95.29	88.99	92.39	93.78
	$V_{Q=105\%}$ (%)	51.75	59.80	62.81	53.79	50.47	73.78	56.27	57.79
		robust CTV-based							
nominal (on GTV)	QF (%)	1.90	1.50	1.46	1.89	1.95	1.76	1.85	1.95
	$V_{Q=95\%}$ (%)	100	100	100	100	99.94	100	100	99.88
	$V_{Q=105\%}$ (%)	7.12	3.87	3.81	7.09	7.43	6.29	6.58	7.25
worst (on GTV)	QF (%)	2.51	2.20	2.16	2.49	2.85	2.33	2.52	2.93
	$V_{Q=95\%}$ (%)	98.31	98.61	98.35	98.50	96.67	97.27	98.13	96.62
	$V_{Q=105\%}$ (%)	11.13	8.00	7.95	10.86	14.75	9.07	11.68	15.96

Table 5.4: Nominal and worst case QVH metrics for the different beam configurations, for conventionally optimized PTV-based plans and robust CTV-based plans.

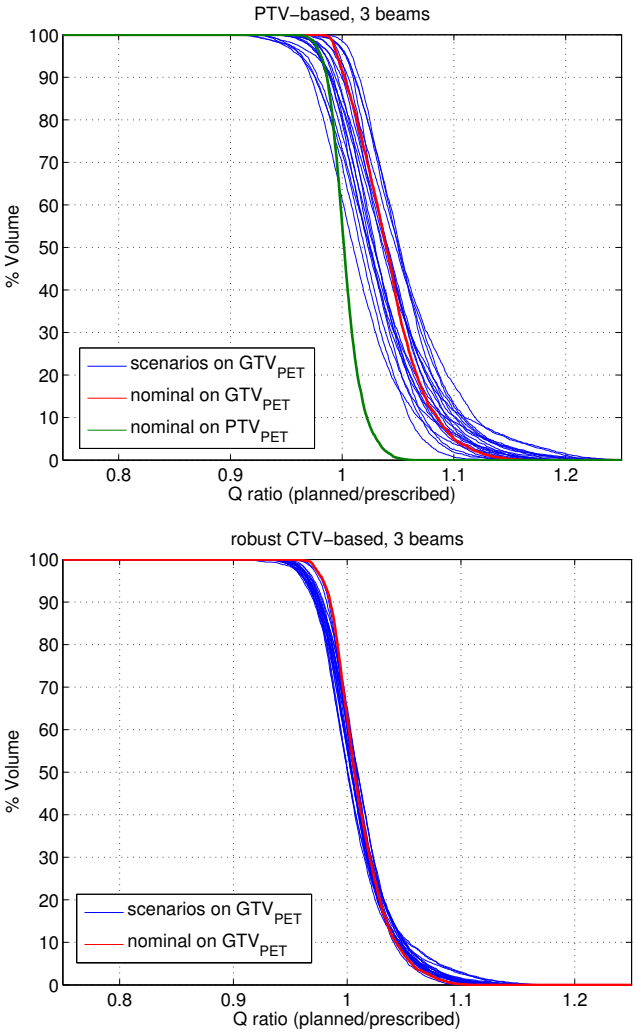


Figure 5.7: QVH bands for GTV_{PET} volume in patient 1, with a configuration of 3 beams and 7 subcontours, for both PTV-based conventional optimization (left) and robust optimization (right). Green and red solid line represents the nominal QVH on $\text{PTV}_{PET}$and GTV_{PET}, respectively; while the blue lines correspond to the different scenarios.

Chapter 6

Conclusions and outlook

This thesis aimed at improving treatment planning for scanned proton therapy. The contributions from the author can be summarized with the following three points:

- An efficient method to achieve **accurate** dose calculation in reduced computation time was implemented and validated for several clinical cases. The method combined the speed of analytical pencil beam algorithms (PB) with the accuracy of Monte Carlo (MC) in a hybrid PB-MC optimization approach (Chapter 3). It could be easily integrated in commercial treatment planning systems based on PB dose calculation, to overcome the limitations the latter might have to ensure accurate dose distributions in heterogeneous media.

- We developed a valuable research tool for **robust** treatment planning of scanned proton therapy (Chapter 2), and used it to investigate novel strategies for the simulation of different sources of errors. We completed the current spectrum of the uncertainty models included in robust optimizers (systematic setup and proton range errors related to image conversion) by providing efficient methods to simulate random setup errors (Appendix B) and motion induced by breathing (Chapter 4).

- The new treatment planning paradigm of dose painting, which aims at targeting the **patient-specific** biological heterogeneities within the tumor, has been integrated through an automatic workflow into a commercial treatment planning system (Chapter 2 and appendix C). The workflow was used to prove the feasibility of generating robust dose painting plans with scanned proton therapy (Chapter 5).

The combination of these three points may potentially benefit cancer patients through the delivery of more *robust, accurate, and patient-specific proton therapy* treatment plans, which was the objective of this thesis. However, there is still room for improving the presented methods, as well as there are other related topics not treated here that should be further investigated. This is discussed in the following paragraphs.

The hybrid PB - MC optimization approach for accurate proton plans was implemented in our in-house treatment planning system (TPS), and tested for PTV-based optimization (Chapter 2). On one hand, it should be ideally integrated into a commercial TPS so that the clinical community can benefit from this tool directly. In addition, analysing the performance of the hybrid method with other

(clinical) PB algorithms would give us more insight into the efficacy and possible limitations of the method. On the other hand, although the extension of the hybrid strategy for robust optimization is straightforward and has been discussed at the end of Chapter 3, we did not report relevant results on this subject yet. This needs to be carefully investigated, which is already ongoing.

The 4D optimization methods presented in Chapter 4 to include the breathing motion into a robust worst-case optimization approach contain several features that make them good candidates to be used in clinical routine. First, they are based on MC dose calculation, providing more accurate treatment plans than the method proposed by Liu et al. [97]. Second, the output treatment plan does not need the synchronization of the beam dynamic delivery with the tumor motion, as it was required in most of the studies published so far [54, 52, 38]. And third, the effects of the breathing motion are included in a single scenario, without increasing the computation resources with respect to those needed for a regular 3D optimization. The only requirement is that motion needs to be regularized as much as possible, since the method relies on the 4DCT as a perfect representation of sucessive breathing cycles during treatment, i.e., no inter-fraction or intra-fraction variations. This can be partially achieved by using audio-coaching, as we currently do in our hospital. However, the group of Parkes et al. have recently published a innovative article affirming that the breathing motion can be considerably regularized by using a mechanical ventilator during treatment [125]. We are confident that combining our 4D robust optimization method with the use of a mechanical ventilator to ensure a fully regular motion could be the key to treat mobile tumors in the thorax region (e. g. lung, liver) with scanned proton therapy. Indeed, our group has started a pioneering study to investigate the use of this device and its implementation into clinical routine. Then, the remaining issue for treating moving tumors would be to mitigate the interplay effects. In this line, numerous studies have already reported that interplay effects average out over the course of a conventionally fractionated treatment (i.e., 30 fractions) [35, 92], yielding a conformal overall dose distribution, and for each individual fraction they can be compensated by using simple techniques such as *rescanning* [18, 75, 191].

The developed treatment planning system, MIROpt, can be potentially used as a research tool to investigate new ways to integrate different sources of errors into a robust worst-case scenario optimization approach, as we did for the breathing motion or the random setup errors. Another example would be to include a few scenarios to simulate the change in the mean position of the tumor over time, known as base-line shift, which is one of the main sources of errors for lung tumors. A model for the latter was also developed by our group and pre-

sented elsewhere [146]. However, the results collected in Chapter 4 suggest that all the sources of errors combine and that treating them as independent scenarios might not be the best option. Moreover, the more scenarios we use, the higher the computation time and resources we need to be able to solve the optimization problem. Taking into account that the number of scenarios multiplies when adding new ones, including several sources of uncertainty in a robust worst-case scenario approach would lead to an exponential increase of this number. For instance, setup errors were simulated inside MIROpt by using seven scenarios (Chapter 2), which were multiplied by three when including those used to model the range uncertainty, giving a total of 21 scenarios. If a few more scenarios are added to model a new source of error like the baseline shift, the number of scenarios would be again multiplied by the latter and would easily rise above the 100 scenarios. Therefore, strategies to reduce the number of scenarios and the time employed in their simulation must be investigated. A good example on strategies to reduce the computation time for scenario simulation have been suggested by Perko et al., using polynomial chaos expansion techniques [126]. These techniques can create a meta-model of the dose engine by approximating the dose in every voxel with multidimensional polynomials, which is then used to compute the expected dose, the dose variance, and other evaluation dose metrics. These models and the obtained information could be then used to feed robust optimizers. However, the questions regarding how to combine different uncertainties in the same scenario and how to reduce their total number remain still unaddressed.

We chose to use a worst-case scenario approach to implement robust optimization, which is the leading technique because it is the one provided in some of the most known commercial TPSs, such as RayStation (RaySearch Laboratories, Sweden) or Eclipse (Varian Medical Systems, USA). However, other techniques have been proposed in the literature, like expected value optimization, also referred to as probabilistic optimization [166, 165], and voxelwise worst case optimization [166, 128]. See section 1.2.1 for a review of these techniques. All these techniques, including the worst-case scenario one, have been studied in detail by several groups [27, 45, 46]. Although all of them can provide robust plans against the considered error scenarios, they count with some limitations. For instance, worst-case scenario approach risks neglecting scenarios with low objective function, expected value optimization might increase integral dose in order to ensure robustness against systematic errors, and voxelwise worst case might be overly conservative since it uses non-physical dose distributions. New methods to overcome these limitations must be investigated, as it is been done by the group of Fredriksson et al., with a scenario-based generalization of the conventional margin-based planning [47, 20]. In this line, MIROpt could be used as a research platform

to explore different methods, since its internal code can be easily modified by the user. In addition, the code will be released open source soon, which is a good opportunity to attract new research users and combine their inputs to improve MIROpt, so that it can be used for further research in other emerging areas in proton therapy treatment planning (e.g., automatic planning [26] or spatio temporal optimization [78]).

Regardless of the method used for robust optimization, the values for the uncertainty parameters must be carefully selected. Paganetti et al. [120] did an extensive quantification of the contributions from different factors to the proton range uncertainty, obtaining a final value around 3%, which is the one used in most publications. However, this estimation should be reviewed by each center, and if necessary, updated to incorporate the specificities related to their clinical practice, such us their patient positioning accuracy or their CT calibration protocols. For systematic errors, the current practice is to use the same value as the margin used to create the PTV (e.g., 5 mm for lung tumors [114, 48, 98]), which might not be a good approach. Indeed, the PTV margin, if properly calculated, is derived from recipes that combine several population-based parameters accounting from different sources of errors, such as target delineation, systematic and random setup errors, baseline shift or organ motion [174]. Therefore, associating the PTV margin to a single robust parameter accounting for setup errors, has definetely no clinical meaning. Thus, a proper quantification of the clinical impact of the parameters used in robust optimization must be done, and if possible, derive population-based parameters that can be used to derive recipes or guidelines for robust optimization, mimicking the well known van Herk's formalism [174] for margins. This will allow us to transfer the knowledge accumulated in conventional radiotherapy using safety margins into robust optimization, by still keeping an intelligent and clinically meaningful definition of the uncertainty parameters. Some studies on this subject have been already published recently, quantifying the impact of different robustness parameters for head and neck cancer patients [168, 171]. But this analysis should be also done for other treatment sites and include higher number of patients. Again, MIROpt could be used as a research tool to facilitate this process.

Robust optimization is a very valuable tool that has improved a lot the treatment planning for scanned proton therapy. However, it does not ensure a complete robustness in the presence of important anatomical modifications. In this case, adaptive therapy [189] must be used to modify the treatment according to the new anatomy. During the last years, significant progress has been done on this topic, in order to automate and speed up the treatment adaptation process. But

further research is needed to reach a fully automated and online adaptive therapy. Our group is also working on the subject very intensively, using the existing robust optimization tools. However, the role of robust optimization as it is known today, might need to be reviewed if we move towards a daily adaptation framework. Since robust optimization is usually very costly in terms of computation time, a daily adaptation workflow might not afford its use unless very fast replanning techniques will be available. Again, this converges to the need of strategies to speed up and/or reduce the number of scenario simulated, as previously discussed.

Regarding treatment planning in the particular case of dose painting, we have developed and integrated all the necessary tools into a commercial TPS to generate plans in an easy and automatic way. However, although the concept of dose painting is very appealing, it is not yet validated clinically as an effective treatment option. So far, indirect evidence for dose painting comes from 1) studies reporting a correlation between the (non-uniform) pattern of response to radiation and the spatial distribution of the tracer uptake [6, 4, 113, 23], and 2) mathematical models demonstrating that an optimal dose escalation according to the heterogeneities in the tumor improves the tumor control probability [25, 150, 157, 179]. But the ultimate proof should ideally come from clinical trials. In order to have a fair comparison, Bentzen et al. [17] suggested that such trials must compare two arms, the first one delivering a heterogeneous dose escalation according to the tracer uptake spatial distribution, and the second one applying a homogeneous dose with the same integral dose. Alternatively, a third arm delivering a random heterogeneous dose, or even a inverse dose escalation (i.e., higher dose to the low SUV values [163]), would help us to better explore the space of possibilities. Recently, a preclinical study with rats used this approach to evaluate the therapeutic efficacy of a dose escalation based on FDG uptake from PET images [163]. Unfortunately, the results for the heterogeneous and the homogeneous arms were not significantly different, failing to prove that dose painting is superior to an homogeneous boost. However, some methodological factors could have probably contributed to this negative results, such as the low spatial resolution of PET imaging or partial volume effects, which might jeopardize the final uptake pattern shown by the PET image, with respect to the real one. Several clinical studies are currently ongoing for lung [172, 133, 176] and head and neck patients [39, 19], and their final results are eagerly awaited. Nevertheless, this might not completely confirm the dose painting hypothesis yet due to several factors. First, some of these studies do not perform a fair comparison, since they have a lower integral dose in the homogeneous arm. Therefore, not being able to distinguish if the therapeutic gain comes from the increased dose or from the specific targeting of the most radioresistant areas. Second, several methodological issues might bias the

results, such as the low spatial resolution of PET images and the related partial volume effects [143], the mathematical function that convert the image information into dose prescription, the tracer signal-to-noise ratio, or even microscopical features that cannot be targeted with the current image technology. This makes the analysis of the results very difficult, and therefore, confirming or refuting the hypothesis of dose painting might require several years of clinical experience and numerous trials. For this purpose, understanding the limitations of the selected imaging modality is one of the key elements to properly design the treatment and analyse the outcome. In any case, from the treatment planning point of view, our work is accomplished and the tools are ready to generate robust dose painting plans in a commercial TPS, which can indeed facilitate the implementation of multicenter clinical trials with a large number of patients.

Several groups are also working on the use of functional image information to re-distribute the dose to the organs at risk in the safest way possible, avoiding the parts that might have a higher risk for toxicity [88]. The automatic dose painting workflow can also be easily adapted for this kind of *functional avoidance dose painting* plans. In addition, the voxel-wise optimization feature can be very useful in other fields beyond dose painting, such as adaptive planning. Using voxel-wise optimization one could reproduce the optimal three-dimensional dose distribution generated at the planning stage and validated by the physician, adapting the proton range and the spot weights for the new anatomy [70]. This would help to automate the adaptive replanning process and eventually bypass the physician approval, to reach a full online adaptive therapy.

So far, all the clinical studies about dose painting have been performed with intensity modulated photon therapy, but using proton therapy could be advantageous in term of the dose to the normal tissue. In fact, several planning studies have shown that, for a given dose escalation, proton plans present a reduced dose to the normal tissue with respect to the photon plans [160, 43]. Therefore, using proton therapy could be a decisive factor to enable a high dose escalation, with reduced normal tissue complications. This will avoid high toxicity cases as reported in some of the abovementioned clinical trials [107].

Appendix A

Convergence conditions of the hybrid algorithm

The following formulation supports theoretically the results obtained in Chapter 3 for the test cases under study, and exposes the limitations of the hybrid method when there are significant differences between the two considered dose engines. The notation $\ddot{x}$ will be used for the accurate dose, whereas $\tilde{x}$ will design the approximate dose.

Let E be a simplified version of the objective function presented in Section 3.2.2 of Chapter 3. E is expressed as the squared difference between the prescription ξ and the calculated dose d, summed over all voxels, designed with index i. Then we have

$$E = \sum_i (\xi_i - d_i)^2 \tag{A.1}$$

$$d_i = P_{ij} \cdot w_j \tag{A.2}$$

where the dose for a certain voxel i is given by equation A.2, being P the dose influence matrix and w_j the weight for spot j. The ensemble of voxels (N) leads to a matrix denoted by $D = [d_i]_{1 \leq i \leq N}$.

For the hybrid method, the dose d used during optimization was given (in the matrix form) by

$$D^c = \tilde{D} + C_t \tag{A.3}$$

where C_t is the correction matrix for the outer loop t,

$$C_t = \dddot{D}(w^t) - \tilde{D} \tag{A.4}$$

According to equation A.1, then for the hybrid method we have

$$E^{hybrid} = (\xi - (\tilde{D} + C_t))^2 \tag{A.5}$$

Thus the gradient of the hybrid objective function will be

$$\frac{\partial E^{hybrid}}{\partial w} = \tilde{P}^T(\xi - C_t - \tilde{P}w) \underbrace{=}_{\text{eq.}A.4} \tilde{P}^T(\xi - \dddot{P}w^t + \tilde{P}w^t - \tilde{P}w) \tag{A.6}$$

where $\tilde{P}^T$ denotes the transpose of $\tilde{P}$.

Using $w^t = w^t + w - w$ and reorganizing,

$$\frac{\partial E^{hybrid}}{\partial w} = \tilde{P}^T(\xi - \dddot{P}w + (\tilde{P} - \dddot{P})(w^t - w)) \tag{A.7}$$

To ensure that the hybrid method converges to the same or similar result as full MC optimization, the gradients of both methods should have the same direction. We have formulated this by requiring the scalar product to be positive,

$$\left(\frac{\partial E^{hybrid}}{\partial w}\right)^T \frac{\partial \dddot{E}}{\partial w} > 0 \tag{A.8}$$

which results into the expression

$$\left(\dddot{P}^T(\xi - \dddot{P}w)\right)^T \left(\tilde{P}^T(\xi - \dddot{P}w + (\tilde{P} - \dddot{P})(w^t - w))\right) > 0 \tag{A.9}$$

$$(\xi - \dddot{P}w)^T \dddot{P}\tilde{P}^T \left(\xi - \dddot{P}w + (\tilde{P} - \dddot{P})(w^t - w)\right) > 0 \tag{A.10}$$

Assuming that the difference between the current weights and the ones for the outer loop t is small enough [103], $(w^t - w) << \epsilon$, we get

$$(\xi - \dddot{P}w)^T \tilde{P}\dddot{P}(\xi - \dddot{P}w) > 0 \tag{A.11}$$

which requires $\dddot{P}\tilde{P}^T$ to be positive definite in order to satisfy the condition A.8

Note that a symmetric $n \times n$ real matrix M is positive definite if the scalar $z^T M z$ is positive for every non-zero column vector z of n real numbers. Any positive definite matrix M has positive eigenvalues.

In the following we present three theoretical cases (scaling, convolution and shift) where the approximate dose influence matrix ($\tilde{P}$) is modelled in matrix form as a perturbation of the accurate dose influence matrix ($\ddot{\ddot{P}}$). $\ddot{\ddot{P}}$ is simulated as the identity matrix I for the shake of simplicity.

- **Scaling**

  Let assume a scaling factor α that will multiply the accurate dose influence matrix $\ddot{\ddot{P}}$,

  $$\tilde{P} = \alpha \ddot{\ddot{P}} = \alpha \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

  Then we have $\ddot{\ddot{P}}\tilde{P}^T = \ddot{\ddot{P}}(\alpha\ddot{\ddot{P}})^T = \alpha\ddot{\ddot{P}}\ddot{\ddot{P}}^T$, which is trivially positive since for any real invertible matrix M, the product $M^T M$ is a positive definite matrix. The eigenvalues of αI are evidently positive too.

- **Convolution**

  Let C denote the convolution matrix that will multiply the accurate dose influence matrix $\ddot{\ddot{P}}$,

  $$\tilde{P} = C\ddot{\ddot{P}} = \begin{bmatrix} 1 & \epsilon & \epsilon & \epsilon \\ \epsilon & 1 & \epsilon & \epsilon \\ \epsilon & \epsilon & 1 & \epsilon \\ \epsilon & \epsilon & \epsilon & 1 \end{bmatrix}$$

  The eigenvalues of $\ddot{\ddot{P}}\tilde{P}^T$ are always positive, regardless of the value of ϵ, provided that $\epsilon > 0$.

- **Shift**

  Let S_x be the permutation operator that will introduce a shift in the accurate

dose influence matrix $\dddot{P}$,

$$\tilde{P} = S_x \dddot{P} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

The eigenvalues of $\dddot{P}\tilde{P}^T$ comprise negative and complex numbers, which tell us that $\ddot{P}\tilde{P}^T$ is not positive definite.

The three cases presented above demonstrated that the convergence of the hybrid method to a result close to the full MC optimization is ensured for cases that do not involve any shift of the dose distribution (scaling, convolution). However, when a shift is present the convergence of the algorithm is not guaranteed.

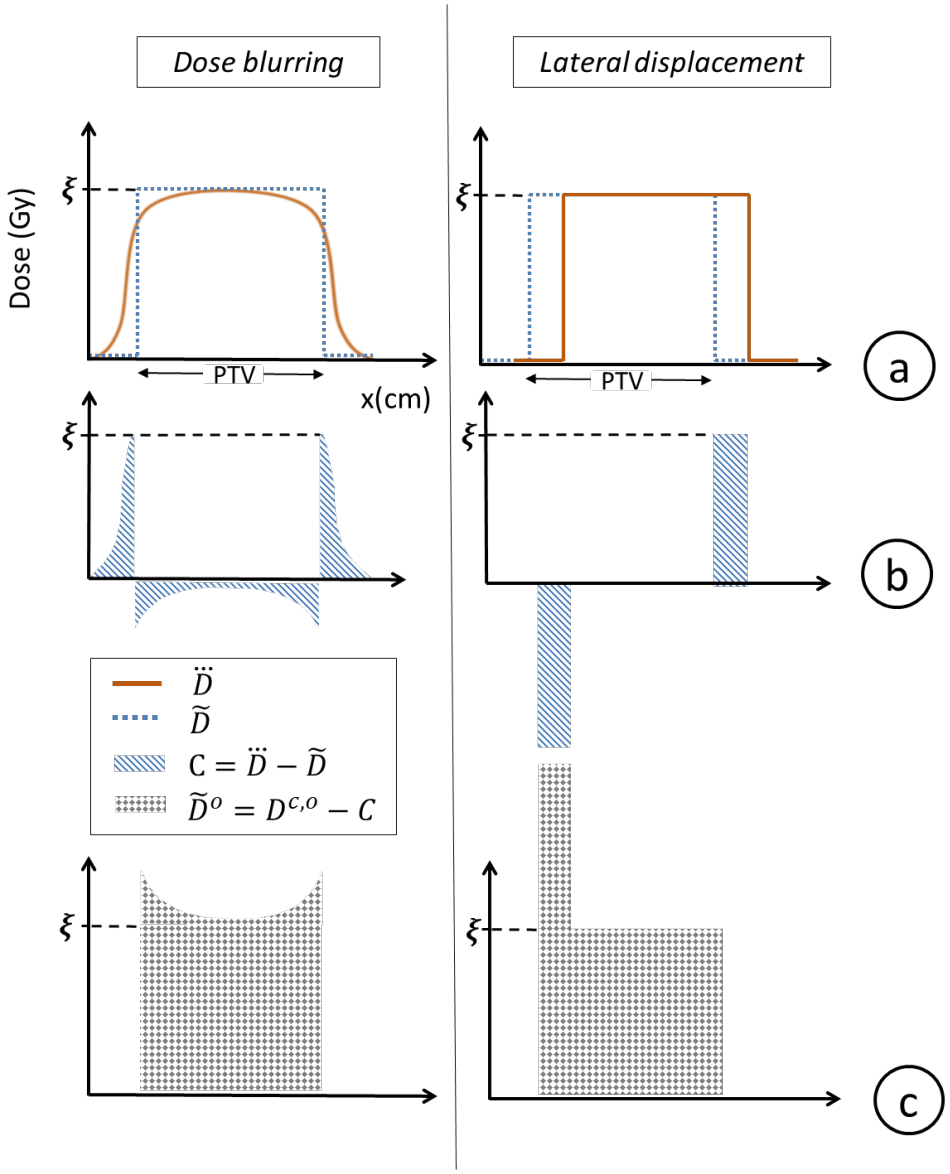


Figure A.1: Representative scheme of the dose distributions involved in the hybrid workflow for a test case involving a lateral shift of the dose (right column), and a case where the dose is blurred by using a convolution with a Gaussian kernel (left column). The plots in the first row (a) represent the dose profiles for the accurate ($\ddot{\ddot{D}}$) and approximated dose ($\tilde{D}$), which is ideally equal to the prescribed dose (ξ). The second row (b) shows the shape of the correction term (C) for each case; while the third row (c) presents the approximate dose after optimization ($\tilde{D}^o$), which results from optimization of $D^c = \tilde{D} + C$.

Appendix B

Robustness against random setup errors: implementation and results

This appendix describes an efficient way of simulating random setup errors by using Monte Carlo methods, and its integration in a robust worst-case scenario optimization approach. The plans were created with our in-house treatment planning system MIROpt, presented in the first part of Chapter 2. The following content is based on the abstracts accepted as oral communication at the annual meetings of ESTRO (Turin, Italy) and AAPM (Washington, USA) in 2016. See Efficient implementation of random errors in robust optimization for proton therapy with Monte Carlo and A Monte Carlo-based method to include random errors in robust optimization in the *List of Publications*.

B.1 Background and purpose

In treatment planning for proton therapy, robust optimizers typically limit their scope to systematic setup and proton range errors. Treatment execution errors (patient and organ motion) are seldom included, mainly because their simulation is very costly in terms of computation time and resources. Indeed, the only way to integrate this type of errors in a robust worst-case scenario approach, if an analytical dose engine is used, is to sample random positions from a probability distribution and compute the dose distribution for each new position of the patient (organs). This increases the computation time since many dose calculations need to be performed to have a proper representation of the motion effects. However, the stochastic nature of Monte Carlo (MC) methods allows treatment execution errors to be simulated in a single dose calculation, and therefore, without increas-

ing the computation time. This work focuses on the implementation of patient positioning (setup) random errors.

B.2 Materials and methods

The plans were created with our in-house treatment planning system MIROpt, which is based on worst-case scenario optimization and coupled with the superfast MC dose engine MCsquare [147, 180] (See Chapter 2 for a detailed description). MCsquare enables computing the spot dose contributions (*dose influence matrix,* P) used during the optimization of the spot weights, as well as final dose distributions (less than 1 minute for final dose).

Random setup errors were simulated with MCsquare, by shifting randomly the starting point of each particle, according to a Gaussian probability distribution with standard deviation σ. This strategy assumes an infinite number of treatment fractions, but it has been demonstrated that the results converge if a sufficient number, e. g. 30 fractions, is used [45]. Therefore, the method can be used for conventionally fractionated treatments.

Two strategies are presented to account for random errors: 1) *Full robust optimization* with a dose influence matrix that already includes the effect of random errors (P_σ) and 2) *Hybrid robust optimization*, where the nominal dose influence matrix ($P_{nominal}$, $\sigma = 0$) is involved, but a correction term C adds the contribution of the random setup errors. Starting from $C = 0$, the method alternates optimization of the spot weights (w) with $P_{nominal}$ and updates of C, with $C = D_\sigma - D_{nominal}$ and where D_σ results from a forward MC computation that includes the random setup errors and $D_{nominal} = P_{nominal} \cdot w$. The term forward MC computation refers to a dose calculation where the total number of simulated particles is properly divided among the spots according to their optimal weight values (w). Therefore, this type of simulation is only possible after optimization, when the information about the optimal w is available. Updates of C can be triggered as often as necessary by running forward dose MC simulations with the last corrected values for w as input.

The two methods were applied to a patient with lung cancer. The standard deviation for random setup errors (σ) was set to 5 mm. Proton range errors (ρ) and systematic setup errors (SS) were also included in all the three plans, with values equal to 3% and 5mm, respectively. This amounted to a total of 21 scenarios, including 3 scenarios for range errors ($\rho = +3,0,-3$ %) and 7 scenarios for systematic setup errors (SS = 0, and SS = 5mm for the positive and negative

directions of the DICOM axis). Three plans were generated: 1) full robust plan with $\sigma = 5mm$ (21 scenarios), 2) full robust plan with $\sigma = 0mm$ and $\sigma = 5mm$ (42 scenarios), and 3) hybrid robust plan with $\sigma = 5mm$ (21 scenarios).

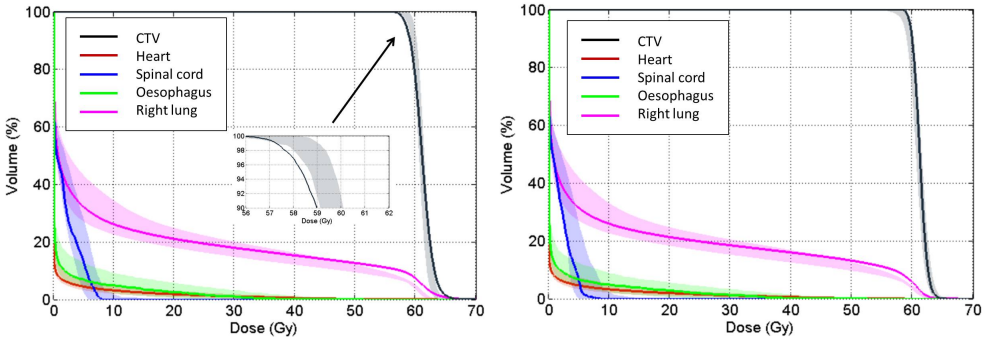
The robustness of the plans were evaluated by computing the final dose for the aforementioned 21 scenarios, plus the diagonal combinations of the DICOM axis for the systematic errors. This amounted to a total of 81 scenarios, all of them including $\sigma = 5mm$. In addition, the nominal scenario ($\sigma = 0mm, \rho = 0\%$, and $SS = 0mm$) was simulated to verify that, in the absence of errors, the patient would also receive a correct dose. The plan was considered acceptable when the CTV dose in the worst-case scenario satisfied the conditions $D_{95} \geq 95\%D_p$ and $D_5 \leq 105\%D_p$.

B.3 Results

The three plans satisfied the evaluation criteria for the target volume, since $D_{95} \geq 95\%D_p$ and $D_5 \leq 105\%D_p$ (Table B.1), and the dose to the organs at risk was similar in the three cases (Figure B.1). For plan 3) (hybrid opt. plan), the target coverage before correction in the worst case was only 91% of D_p, therefore not satisfying the established evaluation criterion. But after applying a single correction, the hybrid method was able to recover an excellent target dose distribution, equivalent to plan 2) (full opt. plan with $\sigma = 0$ and $5mm$) in both the nominal (difference in D_{95} and D_5 of 0.3% and 0.8%, respectively) and wost case (difference in D_{95} and D_5 of 0% and 0.2%, respectively). In contrast, plan 1) (full opt. plan with $\sigma = 5mm$) presented a particularly deteriorated nominal dose distribution. Indeed, the nominal case (solid lines in Figure B.1a) was the worst one ($D_{95} = 95.6\%D_p$ and $D_5 = 104.4\%D_p$).

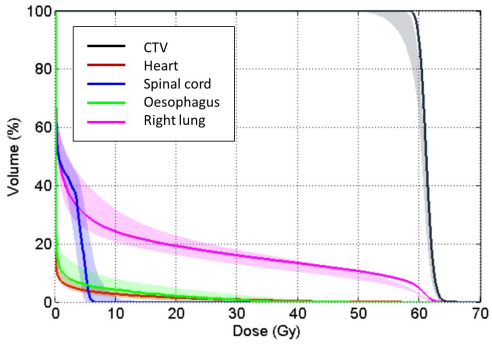
B.4 Discussion

Using full robust optimization including $\sigma = 5mm$ for random errors (plan 1 $\sigma = 5mm$) leads to a low quality dose distribution in the the nominal case, where no errors are considered. This can be explained by the fact that no nominal scenario in terms of random errors ($\sigma = 0mm$) was included in the optimization, and therefore, the optimizer was unable to foreseen a robust dose for this case. In order to guarantee an acceptable nominal dose distribution, while being robust against a given σ, one should use two sets of scenarios taking into account both $\sigma = 0$ and $5mm$ (plan 2). This is important because the standard deviations for the probability distributions of random setup errors are typically population-based, and thus, we cannot predict the exact σ for a given patient. In this case, being

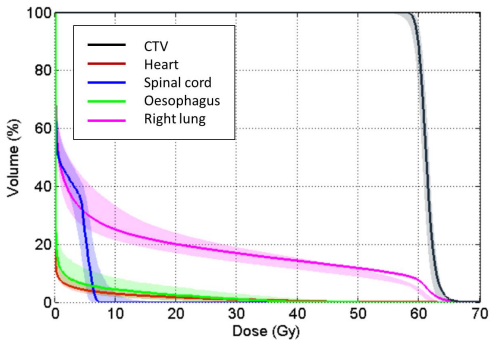


(a) Full robust plan, $\sigma = 5mm$

(b) Full robust plan, $\sigma = 0$ and 5 mm



(c) Hybrid robust plan before correction, $C = 0$



(d) Hybrid robust plan after correction, $C = D_{\sigma=5mm} - D_{nominal}$

Figure B.1: DVH bands corresponding to the target volume (CTV) and organs at risk for the three cases under study: 1) full robust optimization using $\sigma = 5mm$ (21 scenarios) (Figure B.1a), 2) full robust optimization using $\sigma = 0$ and $5mm$ (42 scenarios) (Figure B.1b), and 3) hybrid robust optimization (21 scenarios). For the latter, two plots are presented corresponding to the dose before applying the correction (Figure B.1c) and after applying the correction (Figure B.1d). The solid lines represent the nominal case, where no errors are included ($\sigma = 0mm, \rho = 0\%$, and $SS = 0mm$).

Method		Scenario	D_{95}(CTV) (% of D_p)	D_5(CTV) (% of D_p)
1) Full robust opt., $\sigma = 5mm$		nominal	95.6	104.4
		worst	95.6	104.4
2) Full robust opt., $\sigma = 0$ and $5mm$		nominal	97.5	103.1
		worst	95.7	104.1
3) Hybrid robust opt.	Before correction (C=0)	nominal	97.4	102.8
		worst	91.0	103.1
	After correction ($C = D_{\sigma=5mm} - D_{nominal}$)	nominal	97.2	103.9
		worst	95.7	104.3

Table B.1: DVH metrics for the target volume ($D_{95}(CTV)$ and $D_5(CTV)$)), expressed in % of the dose prescription (D_p), for the three cases under study: 1) full robust optimization with $\sigma = 5mm$, 2) full robust optimization with $\sigma = 0$ and $5mm$, and 3) hybrid robust optimization. In the latter case, the results before (C=0) and after correction are presented. Note that the established evaluation criteria with respect to the target volume were the following: $D_{95} \geq 95\%D_p$ and $D_5 \leq 105\%D_p$.

robust to the largest possible error, as well as to the case with no errors, might be the safest approach. The hybrid optimization approach (plan 3) achieved an excellent target dose distribution for both nominal and worst case, after a single update of the correction term, which validates the method and converts it in an appealing alternative to the full robust optimization. In addition, it could be used in combination with an analytical algorithm such as pencil beam (PB), using the latter to compute the dose influence matrix used for optimization, and correcting for random errors by running the MC including the considered σ (Chapter 3). This would allow to reduce the computation time considerably, while including robustness against random errors.

B.5 Conclusion

The proposed strategies achieved robust plans in term of target coverage without increasing the dose to the CTV nor to the organs at risk. For full robust optimization, using two sets of scenarios accounting for a given value of random setup errors, as well as no random errors, is needed to ensure a good dose distribution. The hybrid strategy is able to correct for the effect of random errors and provides plans that are equivalent to full robust optimization.

Appendix C

Automatic dose painting module: implementation and results

This appendix presents the results for the application of the automatic dose painting workflow (second part of Chapter 2) to a patient with head and neck (H& N) cancer. The content is based on the abstract accepted as oral communication at ESTRO annual meeting in 2015 (Barcelona, Spain). See Automatic dose painting workflow in the *List of Publications*.

C.1 Purpose

Dose painting becomes a popular strategy to increase tumor local control in radiotherapy. However, only a few research centers have developed the tools to apply it to patients. Our aim is to develop an automatic workflow for dose painting, which is integrated in a commercial treatment planning system (TPS).

C.2 Materials and methods

A set of MATLAB functions (GTVPETCT) are called through scripting from RayStation (RaySearch Laboratories, research version 3.99) to segment the primary tumor (GTV_{PET}) in FDG-PET images, using an automatic gradient-based method [50]. The user selects the lower and upper limits (Gy) for dose escalation. The FDG uptake in each voxel is linearly converted to dose (Gy), starting from the median uptake to avoid background contamination. Optimization can be performed using either a number N of sub-volumes (N selected by the user)

or directly on the voxel scale thanks to a customized function developed with the RaySearch research package (C++). The N contours can be used to drive the optimizer towards a dose painting by number prescription (DPBN) or either to perform sub-volume boosting or dose painting by contours (DPBC) if uniform dose is prescribed inside each contour.

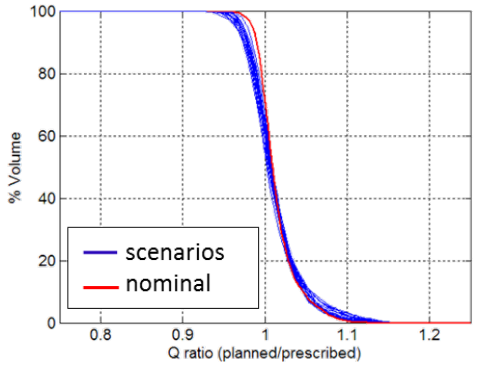
The workflow can be applied for radiation therapy with photons and protons, but in this work we considered only the latter, and in particular, being delivered with proton pencil beam scanning (PBS). To ensure robustness against setup (and proton range) errors, two strategies are implemented: 1) integration of margins in a dilated prescription (PTV-based plan) [152]; and 2) robust optimization (robust plan) [48], but only for contour-driven optimization. Work is in progress to extend this feature to voxel-wise optimization.

Plan quality was assessed with Quality Volume Histograms (QVH) [175], where quality (Q) is the ratio of the planned dose to the prescribed dose in each voxel. Robustness is evaluated by calculating the dose perturbed by setup errors (and range errors for protons) gathered in a set of scenarios.

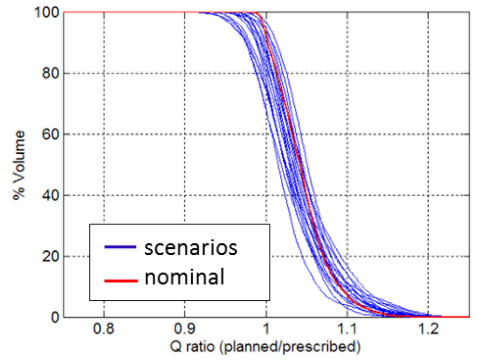
As an illustration, the workflow was applied to a H&N patient, aiming at reproducing a DPBN prescription. Direct voxel-wise optimization and 7 non-uniform dose levels were used for scanned proton therapy. PTV-based (dilated prescription) and robust CTV-based optimization were compared.

C.3 Results

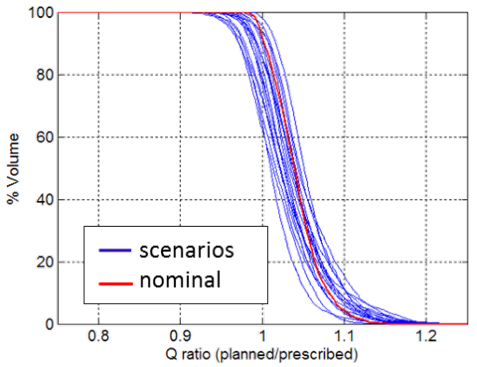
The DPBN prescription was reproduced successfully for both contour-driven and voxel-wise optimizations: more than 99% of the PTV_{PET} received at least 95% of the prescribed dose ($V_{Q=95\%} > 99\%$) and less than 1% received more than 105% ($V_{Q=105\%} < 1\%$). For the robust plan (Figure 1, a), the nominal QVH on GTV_{PET} achieved perfect target coverage ($V_{Q=95\%} = 100\%$) and very little overdosing ($V_{Q=105\%} = 9\%$). However, for dilated prescriptions (Figure 1, b and c), overdosing in GTV_{PET} was very high ($V_{Q=105\%} > 35\%$ and 40% for voxel-wise and contour-driven optimizations, respectively) , which means that this strategy is not sufficient to be robust against setup errors in proton treatments. Voxel-wise optimization reduced by a factor 2 the number of iterations to find the optimal solution, being twice faster than optimization with contours.



(a) robust (contour-driven) optimization



(b) PTV-based (contour-driven) optimization



(c) PTV-based (voxel-wise) optimization

Figure C.1: Robustness evaluation by means of QVHs plots for 20 error scenarios (blue lines) including 2.5mm of setup error (six shifts in positive and negative directions on right-left, anterior-posterior, and inferior-superior) and $\pm 3\%$ density perturbation for range errors. The red line corresponds to the nominal case, with no setup errors and zero density perturbation. All QVHs were computed in the GTV_{PET}

C.4 Conclusion

We have developed an automatic workflow for dose painting that can be directly implemented in clinics since it is integrated in a commercial TPS. Voxel-wise optimization leads to a simplified and faster optimization. The results for the robust plan indicated that robust dose escalation in proton therapy can be performed without having significant target overdosing. On the other hand, a dilated prescription entails too much overdosing and, furthermore, is a non-robust method to account for errors in proton therapy.

References

[1] RaySearch Laboratories, white paper on Dose calculation.

[2] http://nuclear.fis.ucm.es/foca.

[3] H. J. Aerts, G. Bosmans, A. A. van Baardwijk, A. L. Dekker, M. C. Oellers, P. Lambin, and D. De Ruysscher. Stability of 18F-deoxyglucose uptake locations within tumor during radiotherapy for NSCLC: a prospective study. *Int. J. Radiat. Oncol. Biol. Phys.*, 71(5):1402–1407, Aug 2008. [DOI:10.1016/j.ijrobp.2007.11.049] [PubMed:18234432].

[4] H. J. Aerts, J. Bussink, W. J. Oyen, W. van Elmpt, A. M. Folgering, D. Emans, M. Velders, P. Lambin, and D. De Ruysscher. Identification of residual metabolic-active areas within NSCLC tumours using a pre-radiotherapy FDG-PET-CT scan: a prospective validation. *Lung Cancer*, 75(1):73–76, Jan 2012. [DOI:10.1016/j.lungcan.2011.06.003] [PubMed:21782272].

[5] H. J. Aerts, P. Lambin, and D. D. Ruysscher. FDG for dose painting: a rational choice. *Radiother Oncol*, 97(2):163–164, Nov 2010. [DOI:10.1016/j.radonc.2010.05.001] [PubMed:20561697].

[6] H. J. Aerts, A. A. van Baardwijk, S. F. Petit, C. Offermann, J. v. Loon, R. Houben, A. M. Dingemans, R. Wanders, L. Boersma, J. Borger, G. Bootsma, W. Geraedts, C. Pitz, J. Simons, B. G. Wouters, M. Oellers, P. Lambin, G. Bosmans, A. L. Dekker, and D. De Ruysscher. Identification of residual metabolic-active areas within individual NSCLC tumours using a pre-radiotherapy (18)Fluorodeoxyglucose-PET-CT scan. *Radiother Oncol*, 91(3):386–392, Jun 2009. [PubMed Central:PMC4693609] [DOI:10.1016/j.radonc.2009.03.006] [PubMed:19329207].

[7] F. Albertini, E. B. Hug, and A. J. Lomax. Is it necessary to plan with safety margins for actively scanned proton therapy? *Phys Med Biol*, 56(14):4399–4413, Jul 2011. [DOI:10.1088/0031-9155/56/14/011] [PubMed:21709340].

[8] A. Auperin, C. Le Pechoux, E. Rolland, W. J. Curran, K. Furuse, P. Fournel, J. Belderbos, G. Clamon, H. C. Ulutin, R. Paulus, T. Yamanaka, M. C. Bozonnat, A. Uitterhoeve, X. Wang, L. Stewart, R. Arriagada, S. Bur-

dett, and J. P. Pignon. Meta-analysis of concomitant versus sequential radiochemotherapy in locally advanced non-small-cell lung cancer. *J. Clin. Oncol.*, 28(13):2181–2190, May 2010. [DOI:10.1200/JCO.2009.26.2543] [PubMed:20351327].

[9] G. C. Barnett, C. M. West, A. M. Dunning, R. M. Elliott, C. E. Coles, P. D. Pharoah, and N. G. Burnet. Normal tissue reactions to radiotherapy: towards tailoring treatment dose by genotype. *Nat. Rev. Cancer*, 9(2):134–142, 02 2009. [PubMed Central:PMC2670578] [DOI:10.1038/nrc2587] [PubMed:19148183].

[10] M. B. Barton, M. Frommer, and J. Shafiq. Role of radiotherapy in cancer control in low-income and middle-income countries. *Lancet Oncol.*, 7(7):584–595, Jul 2006. [DOI:10.1016/S1470-2045(06)70759-8] [PubMed:16814210].

[11] M. B. Barton, S. Jacob, J. Shafiq, K. Wong, S. R. Thompson, T. P. Hanna, and G. P. Delaney. Estimating the demand for radiotherapy from the evidence: a review of changes from 2003 to 2012. *Radiother Oncol*, 112(1):140–144, Jul 2014. [DOI:10.1016/j.radonc.2014.03.024] [PubMed:24833561].

[12] M. Baumann, M. Krause, J. Overgaard, J. Debus, S. M. Bentzen, J. Daartz, C. Richter, D. Zips, and T. Bortfeld. Radiation oncology in the era of precision medicine. *Nat. Rev. Cancer*, 16(4):234–249, Apr 2016. [DOI:10.1038/nrc.2016.18] [PubMed:27009394].

[13] M. Baumann and C. Petersen. TCP and NTCP: a basic introduction. *Rays*, 30(2):99–104, 2005. [PubMed:16294901].

[14] C. Baumer and J. B. Farr. Lateral dose profile characterization in scanning particle therapy. *Med Phys*, 38(6):2904–2913, Jun 2011. [DOI:10.1118/1.3578603] [PubMed:21815365].

[15] W. A. Beckham, P. J. Keall, and J. V. Siebers. A fluence-convolution method to calculate radiation therapy dose distributions that incorporate random set-up error. *Phys Med Biol*, 47(19):3465–3473, Oct 2002. [PubMed:12408475].

[16] S. M. Bentzen. Theragnostic imaging for radiation oncology: dose-painting by numbers. *Lancet Oncol.*, 6(2):112–117, Feb 2005. [DOI:10.1016/S1470-2045(05)01737-7] [PubMed:15683820].

[17] S. M. Bentzen and V. Gregoire. Molecular imaging-based dose painting: a novel paradigm for radiation therapy prescription. *Semin Radiat Oncol*, 21(2):101–110, Apr 2011. [PubMed Central:PMC3052283] [DOI:10.1016/j.semradonc.2010.10.001] [PubMed:21356478].

[18] C. Bert and M. Durante. Motion in radiotherapy: particle therapy. *Phys Med Biol*, 56(16):R113–144, Aug 2011. [DOI:10.1088/0031-9155/56/16/R01]

[PubMed:21775795].

[19] D. Berwouts, L. A. Olteanu, F. Duprez, T. Vercauteren, W. De Gersem, W. De Neve, C. Van de Wiele, and I. Madani. Three-phase adaptive dose-painting-by-numbers for head-and-neck cancer: initial results of the phase I clinical trial. *Radiother Oncol*, 107(3):310–316, Jun 2013. [DOI:10.1016/j.radonc.2013.04.002] [PubMed:23647760].

[20] R. Bokrantz and A. Fredriksson. Scenario-based radiation therapy margins for patient setup, organ motion, and particle range uncertainty. *Phys Med Biol*, 62(4):1342–1357, Feb 2017. [DOI:10.1088/1361-6560/aa524d] [PubMed:28114114].

[21] S. R. Bowen, R. T. Flynn, S. M. Bentzen, and R. Jeraj. On the sensitivity of IMRT dose optimization to the mathematical form of a biological imaging-based prescription function. *Phys Med Biol*, 54(6):1483–1501, Mar 2009. [PubMed Central:PMC2858011] [DOI:10.1088/0031-9155/54/6/007] [PubMed:19218733].

[22] J. Bradley. A review of radiation dose escalation trials for non-small cell lung cancer within the Radiation Therapy Oncology Group. *Semin. Oncol.*, 32(2 Suppl 3):S111–113, Apr 2005. [PubMed:16015546].

[23] T. J. Bradshaw, S. R. Bowen, M. A. Deveau, L. Kubicek, P. White, S. M. Bentzen, R. J. Chappell, L. J. Forrest, and R. Jeraj. Molecular imaging biomarkers of resistance to radiation therapy for spontaneous nasal tumors in canines. *Int. J. Radiat. Oncol. Biol. Phys.*, 91(4):787–795, Mar 2015. [PubMed Central:PMC4355478] [DOI:10.1016/j.ijrobp.2014.12.011] [PubMed:25752393].

[24] A. Brahme. Biologically optimized 3-dimensional in vivo predictive assay-based radiation therapy using positron emission tomography-computerized tomography imaging. *Acta Oncol*, 42(2):123–136, 2003. [PubMed:12801131].

[25] A. Brahme and A. K. Agren. Optimal dose distribution for eradication of heterogeneous tumours. *Acta Oncol*, 26(5):377–385, 1987. [PubMed:3426851].

[26] S. Breedveld, P. R. Storchi, P. W. Voet, and B. J. Heijmen. iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. *Med Phys*, 39(2):951–963, Feb 2012. [DOI:10.1118/1.3676689] [PubMed:22320804].

[27] M. Casiraghi, F. Albertini, and A. J. Lomax. Advantages and limitations of the 'worst case scenario' approach in IMPT treatment planning. *Phys Med Biol*, 58(5):1323–1339, Mar 2013. [DOI:10.1088/0031-9155/58/5/1323] [PubMed:23391569].

[28] A. W. Chan and N. J. Liebsch. Proton radiation therapy for head and neck cancer. *J Surg Oncol*, 97(8):697–700, Jun 2008. [DOI:10.1002/jso.21013]

[PubMed:18493920].

[29] A. Chi, N. P. Nguyen, J. S. Welsh, W. Tse, M. Monga, O. Oduntan, M. Almubarak, J. Rogers, S. C. Remick, and D. Gius. Strategies of dose escalation in the treatment of locally advanced non-small cell lung cancer: image guidance and beyond. *Front Oncol*, 4:156, 2014. [PubMed Central:PMC4064255] [DOI:10.3389/fonc.2014.00156] [PubMed:24999451].

[30] J. F. Daisne, T. Duprez, B. Weynand, M. Lonneux, M. Hamoir, H. Reychler, and V. Gregoire. Tumor volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. *Radiology*, 233(1):93–100, Oct 2004. [DOI:10.1148/radiol.2331030660] [PubMed:15317953].

[31] D. De Ruysscher, E. Sterpin, K. Haustermans, and T. Depuydt. Tumour Movement in Proton Therapy: Solutions and Remaining Questions: A Review. *Cancers (Basel)*, 7(3):1143–1153, Jun 2015. [PubMed Central:PMC4586762] [DOI:10.3390/cancers7030829] [PubMed:26132317].

[32] B. De Smedt, M. Fippel, N. Reynaert, and H. Thierens. Denoising of Monte Carlo dose calculations: smoothing capabilities versus introduction of systematic bias. *Med Phys*, 33(6):1678–1687, Jun 2006. [DOI:10.1118/1.2198188] [PubMed:16872075].

[33] M. A. Deveau, S. R. Bowen, D. C. Westerly, and R. Jeraj. Feasibility and sensitivity study of helical tomotherapy for dose painting plans. *Acta Oncol*, 49(7):991–996, Oct 2010. [PubMed Central:PMC3040036] [DOI:10.3109/0284186X.2010.500302] [PubMed:20831487].

[34] S. Differding, E. Sterpin, N. Hermand, B. Vanstraelen, S. Nuyts, N. de Patoul, J. M. Denis, J. A. Lee, and V. Gregoire. Radiation dose escalation based on FDG-PET driven dose painting by numbers in oropharyngeal squamous cell carcinoma: a dosimetric comparison between TomoTherapy-HA and RapidArc. *Radiat Oncol*, 12(1):59, Mar 2017. [PubMed Central:PMC5364636] [DOI:10.1186/s13014-017-0793-0] [PubMed:28335778].

[35] S. Dowdell, C. Grassberger, and H. Paganetti. Four-dimensional Monte Carlo simulations demonstrating how the extent of intensity-modulation impacts motion effects in proton therapy lung treatments. *Med Phys*, 40(12):121713, Dec 2013. [PubMed Central:PMC3843755] [DOI:10.1118/1.4829500] [PubMed:24320499].

[36] M. Durante and J. S. Loeffler. Charged particles in radiation oncology. *Nat Rev Clin Oncol*, 7(1):37–43, Jan 2010. [DOI:10.1038/nrclinonc.2009.183] [PubMed:19949433].

[37] M. Durante and H. Paganetti. Nuclear physics in particle therapy: a review. *Rep Prog Phys*, 79(9):096702, Sep 2016. [DOI:10.1088/0034-4885/79/9/096702] [PubMed:27540827].

[38] J. G. Eley, W. D. Newhauser, R. Luchtenborg, C. Graeff, and C. Bert. 4D optimization of scanned ion beam tracking therapy for moving tumors. *Phys Med Biol*, 59(13):3431–3452, Jul 2014. [PubMed Central:PMC4139294] [DOI:10.1088/0031-9155/59/13/3431] [PubMed:24889215].

[39] FiGaRO. https://clinicaltrials.gov/ct2/show/NCT02953197.

[40] M. Fippel and M. Soukup. A Monte Carlo dose calculation algorithm for proton therapy. *Med Phys*, 31(8):2263–2273, Aug 2004. [DOI:10.1118/1.1769631] [PubMed:15377093].

[41] S. Flampouri, S. B. Jiang, G. C. Sharp, J. Wolfgang, A. A. Patel, and N. C. Choi. Estimation of the delivered patient dose in lung IMRT treatment based on deformable registration of 4D-CT data and Monte Carlo simulations. *Phys Med Biol*, 51(11):2763–2779, Jun 2006. [DOI:10.1088/0031-9155/51/11/006] [PubMed:16723765].

[42] R. T. Flynn, D. L. Barbee, T. R. Mackie, and R. Jeraj. Comparison of intensity modulated x-ray therapy and intensity modulated proton therapy for selective subvolume boosting: a phantom study. *Phys Med Biol*, 52(20):6073–6091, Oct 2007. [PubMed Central:PMC2712448] [DOI:10.1088/0031-9155/52/20/001] [PubMed:17921573].

[43] R. T. Flynn, S. R. Bowen, S. M. Bentzen, T. Rockwell Mackie, and R. Jeraj. Intensity-modulated x-ray (IMXT) versus proton (IMPT) therapy for theragnostic hypoxia-based dose painting. *Phys Med Biol*, 53(15):4153–4167, Aug 2008. [PubMed Central:PMC2695924] [DOI:10.1088/0031-9155/53/15/010] [PubMed:18635895].

[44] S. J. Frank, J. D. Cox, M. Gillin, R. Mohan, A. S. Garden, D. I. Rosenthal, G. B. Gunn, R. S. Weber, M. S. Kies, J. S. Lewin, M. F. Munsell, M. B. Palmer, N. Sahoo, X. Zhang, W. Liu, and X. R. Zhu. Multifield optimization intensity modulated proton therapy for head and neck tumors: a translation to practice. *Int. J. Radiat. Oncol. Biol. Phys.*, 89(4):846–853, Jul 2014. [PubMed Central:PMC4171724] [DOI:10.1016/j.ijrobp.2014.04.019] [PubMed:24867532].

[45] A. Fredriksson. A characterization of robust radiation therapy treatment planning methods-from expected value to worst case optimization. *Med Phys*, 39(8):5169–5181, Aug 2012. [DOI:10.1118/1.4737113] [PubMed:22894442].

[46] A. Fredriksson and R. Bokrantz. A critical evaluation of worst case optimization methods for robust intensity-modulated proton therapy planning. *Med Phys*, 41(8):081701, Aug 2014. [DOI:10.1118/1.4883837] [PubMed:25086511].

[47] A. Fredriksson and R. Bokrantz. The scenario-based generalization of radiation therapy margins. *Phys Med Biol*, 61(5):2067–2082, Mar 2016.

[DOI:10.1088/0031-9155/61/5/2067] [PubMed:26895381].

[48] A. Fredriksson, A. Forsgren, and B. Hardemark. Minimax optimization for handling range and setup uncertainties in proton therapy. *Med Phys*, 38(3):1672–1684, Mar 2011. [DOI:10.1118/1.3556559] [PubMed:21520880].

[49] R. Fujimoto, T. Kurihara, and Y. Nagamine. GPU-based fast pencil beam algorithm for proton therapy. *Phys Med Biol*, 56(5):1319–1328, Mar 2011. [DOI:10.1088/0031-9155/56/5/006] [PubMed:21297243].

[50] X. Geets, J. A. Lee, A. Bol, M. Lonneux, and V. Gregoire. A gradient-based method for segmenting FDG-PET images: methodology and validation. *Eur. J. Nucl. Med. Mol. Imaging*, 34(9):1427–1438, Sep 2007. [DOI:10.1007/s00259-006-0363-4] [PubMed:17431616].

[51] D. Giantsoudi, C. Grassberger, D. Craft, A. Niemierko, A. Trofimov, and H. Paganetti. Linear energy transfer-guided optimization in intensity modulated proton therapy: feasibility study and clinical potential. *Int. J. Radiat. Oncol. Biol. Phys.*, 87(1):216–222, Sep 2013. [PubMed Central:PMC4520412] [DOI:10.1016/j.ijrobp.2013.05.013] [PubMed:23790771].

[52] C. Graeff, A. Constantinescu, R. Luchtenborg, M. Durante, and C. Bert. Multigating, a 4D optimized beam tracking in scanned ion beam therapy. *Technol. Cancer Res. Treat.*, 13(6):497–504, Dec 2014. [PubMed Central:PMC4527435] [DOI:10.7785/tcrtexpress.2013.600277] [PubMed:24354752].

[53] C. Graeff, M. Durante, and C. Bert. Motion mitigation in intensity modulated particle therapy by internal target volumes covering range changes. *Med Phys*, 39(10):6004–6013, Oct 2012. [DOI:10.1118/1.4749964] [PubMed:23039638].

[54] C. Graeff, R. Luchtenborg, J. G. Eley, M. Durante, and C. Bert. A 4D-optimization concept for scanned ion beam therapy. *Radiother Oncol*, 109(3):419–424, Dec 2013. [DOI:10.1016/j.radonc.2013.09.018] [PubMed:24183865].

[55] C. Grassberger, J. Daartz, S. Dowdell, T. Ruggieri, G. Sharp, and H. Paganetti. Quantification of proton dose calculation accuracy in the lung. *Int. J. Radiat. Oncol. Biol. Phys.*, 89(2):424–430, Jun 2014. [PubMed Central:PMC4028367] [DOI:10.1016/j.ijrobp.2014.02.023] [PubMed:24726289].

[56] V. Gregoire, E. Coche, G. Cosnard, M. Hamoir, and H. Reychler. Selection and delineation of lymph node target volumes in head and neck conformal radiotherapy. Proposal for standardizing terminology and procedure based on the surgical experience. *Radiother Oncol*, 56(2):135–150, Aug 2000.

[PubMed:10927132].

[57] V. Gregoire, A. Eisbruch, M. Hamoir, and P. Levendag. Proposal for the delineation of the nodal CTV in the node-positive and the post-operative neck. *Radiother Oncol*, 79(1):15–20, Apr 2006. [DOI:10.1016/j.radonc.2006.03.009] [PubMed:16616387].

[58] L. Grevillot, D. Bertrand, F. Dessy, N. Freud, and D. Sarrut. A Monte Carlo pencil beam scanning model for proton treatment plan simulation using GATE/GEANT4. *Phys Med Biol*, 56(16):5203–5219, Aug 2011. [DOI:10.1088/0031-9155/56/16/008] [PubMed:21791731].

[59] Loïc Grevillot, Thibault Frisson, Nabil Zahra, Damien Bertrand, Frédéric Stichelbaut, Nicolas Freud, and David Sarrut. Optimization of geant4 settings for proton pencil beam scanning simulations using gate. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 268(20):3295 – 3305, 2010.

[60] T. Halabi, D. Craft, and T. Bortfeld. Dose-volume objectives in multi-criteria optimization. *Phys Med Biol*, 51(15):3809–3818, Aug 2006. [DOI:10.1088/0031-9155/51/15/014] [PubMed:16861782].

[61] P. Hill, D. Westerly, and T. Mackie. Fan-beam intensity modulated proton therapy. *Med Phys*, 40(11):111704, Nov 2013. [PubMed Central:PMC3799727] [DOI:10.1118/1.4822485] [PubMed:24320412].

[62] C. E. Hill-Kayser, S. Both, and Z. Tochner. Proton Therapy: Ever Shifting Sands and the Opportunities and Obligations within. *Front Oncol*, 1:24, 2011. [PubMed Central:PMC3355860] [DOI:10.3389/fonc.2011.00024] [PubMed:22655235].

[63] L. Hong, M. Goitein, M. Bucciolini, R. Comiskey, B. Gottschalk, S. Rosenthal, C. Serago, and M. Urie. A pencil beam algorithm for proton dose calculations. *Phys Med Biol*, 41(8):1305–1330, Aug 1996. [PubMed:8858722].

[64] T. C. Huang, J. A. Liang, T. Dilling, T. H. Wu, and G. Zhang. Four-dimensional dosimetry validation and study in lung radiotherapy using deformable image registration and Monte Carlo techniques. *Radiat Oncol*, 5:45, May 2010. [PubMed Central:PMC2890615] [DOI:10.1186/1748-717X-5-45] [PubMed:20509955].

[65] ICRU. Nuclear data for neutron and proton radiotherapy and for radiation protection. Report 63, 2000.

[66] Digital Imaging and Communications in Medicine (DICOM). http://medical.nema.org.

[67] T. Inoue, J. Widder, L. V. van Dijk, H. Takegawa, M. Koizumi, M. Takashina, K. Usui, C. Kurokawa, S. Sugimoto, A. I. Saito, K. Sasai, A. A. Van't Veld, J. A. Langendijk, and E. W. Korevaar. Limited Impact of Setup and Range Uncertainties, Breathing Motion, and Interplay Effects in

Robustly Optimized Intensity Modulated Proton Therapy for Stage III Non-small Cell Lung Cancer. *Int. J. Radiat. Oncol. Biol. Phys.*, 96(3):661–669, Nov 2016. [DOI:10.1016/j.ijrobp.2016.06.2454] [PubMed:27681763].

[68] IPOPT Interior Point OPTimizer. https://projects.coin-or.org/Ipopt.

[69] D. A. Jaffray and M. K. Gospodarowicz. Chapter 14 Radiation Therapy for Cancer. *Cancer: Disease Control Priorities, Third Edition*, 3:239–247, Nov 2015. [DOI:10.1596/978-1-4648-0349-9_ch14] [PubMed:26913334].

[70] T. Jagt, S. Breedveld, S. van de Water, B. Heijmen, and M. Hoogeman. Near real-time automated dose restoration in IMPT to compensate for daily tissue density variations in prostate cancer. *Phys Med Biol*, 62(11):4254–4272, Jun 2017. [DOI:10.1088/1361-6560/aa5c12] [PubMed:28140380].

[71] G. Janssens, L. Jacques, J. Orban de Xivry, X. Geets, and B. Macq. Diffeomorphic registration of images with variable contrast enhancement. *Int J Biomed Imaging*, 2011:891585, 2011. [PubMed Central:PMC3005125] [DOI:10.1155/2011/891585] [PubMed:21197460].

[72] R. Jeraj, T. Bradshaw, and U. Simon?i? Molecular Imaging to Plan Radiotherapy and Evaluate Its Efficacy. *J. Nucl. Med.*, 56(11):1752–1765, Nov 2015. [DOI:10.2967/jnumed.114.141424] [PubMed:26383148].

[73] R. Jeraj and P. Keall. The effect of statistical uncertainty on inverse treatment planning based on Monte Carlo dose calculation. *Phys Med Biol*, 45(12):3601–3613, Dec 2000. [PubMed:11131187].

[74] X. Jia, J. Schumann, H. Paganetti, and S. B. Jiang. GPU-based fast Monte Carlo dose calculation for proton therapy. *Phys Med Biol*, 57(23):7783–7797, Dec 2012. [PubMed Central:PMC4474737] [DOI:10.1088/0031-9155/57/23/7783] [PubMed:23128424].

[75] L. Kardar, Y. Li, X. Li, H. Li, W. Cao, J. Y. Chang, L. Liao, R. X. Zhu, N. Sahoo, M. Gillin, Z. Liao, R. Komaki, J. D. Cox, G. Lim, and X. Zhang. Evaluation and mitigation of the interplay effects of intensity modulated proton therapy for lung cancer in a clinical setting. *Pract Radiat Oncol*, 4(6):e259–268, 2014. [PubMed Central:PMC4399385] [DOI:10.1016/j.prro.2014.06.010] [PubMed:25407877].

[76] I. Kawrakow. On the de-noising of Monte Carlo calculated dose distributions. *Phys Med Biol*, 47(17):3087–3103, Sep 2002. [PubMed:12361212].

[77] P. J. Keall, J. V. Siebers, S. Joshi, and R. Mohan. Monte Carlo as a four-dimensional radiotherapy treatment-planning tool to account for respiratory motion. *Phys Med Biol*, 49(16):3639–3648, Aug 2004. [PubMed:15446794].

[78] M. Kim, R. D. Stewart, and M. H. Phillips. A feasibility study: Selection of a personalized radiotherapy fractionation schedule using spatiotemporal optimization. *Med Phys*, 42(11):6671–6678, Nov 2015. [DOI:10.1118/1.4934369] [PubMed:26520757].

[79] Y. S. Kim. Reirradiation of head and neck cancer in the era of intensity-modulated radiotherapy: patient selection, practical aspects, and current evidence. *Radiat Oncol J*, 35(1):1–15, Mar 2017. [PubMed Central:PMC5398346] [DOI:10.3857/roj.2017.00122] [PubMed:28395502].

[80] J. P. Kirkpatrick, A. J. van der Kogel, and T. E. Schultheiss. Radiation dose-volume effects in the spinal cord. *Int. J. Radiat. Oncol. Biol. Phys.*, 76(3 Suppl):S42–49, Mar 2010. [DOI:10.1016/j.ijrobp.2009.04.095] [PubMed:20171517].

[81] A. C. Knopf, D. Boye, A. Lomax, and S. Mori. Adequate margin definition for scanned particle therapy in the incidence of intrafractional motion. *Phys Med Biol*, 58(17):6079–6094, Sep 2013. [DOI:10.1088/0031-9155/58/17/6079] [PubMed:23939146].

[82] I. S. Knudtsen, J. G. Svestad, E. P. Skaug Sande, B. L. Rekstad, J. R?dal, W. van Elmpt, M. Ollers, E. O. Hole, and E. Malinen. Validation of dose painting of lung tumours using alanine/EPR dosimetry. *Phys Med Biol*, 61(6):2243–2254, Mar 2016. [DOI:10.1088/0031-9155/61/6/2243] [PubMed:26913997].

[83] R. Kohno, K. Hotta, S. Nishioka, K. Matsubara, R. Tansho, and T. Suzuki. Clinical implementation of a GPU-based simplified Monte Carlo method for a treatment planning system of proton beam therapy. *Phys Med Biol*, 56(22):N287–294, Nov 2011. [DOI:10.1088/0031-9155/56/22/N03] [PubMed:22036894].

[84] H. M. Kooy, B. M. Clasie, H. M. Lu, T. M. Madden, H. Bentefour, N. Depauw, J. A. Adams, A. V. Trofimov, D. Demaret, T. F. Delaney, and J. B. Flanz. A case study in proton pencil-beam scanning delivery. *Int. J. Radiat. Oncol. Biol. Phys.*, 76(2):624–630, Feb 2010. [DOI:10.1016/j.ijrobp.2009.06.065] [PubMed:20117294].

[85] H. M. Kooy and C. Grassberger. Intensity modulated proton therapy. *Br J Radiol*, 88(1051):20150195, Jul 2015. [PubMed Central:PMC4628542] [DOI:10.1259/bjr.20150195] [PubMed:26084352].

[86] S. S. Korreman, S. Ulrich, S. Bowen, M. Deveau, S. M. Bentzen, and R. Jeraj. Feasibility of dose painting using volumetric modulated arc optimization and delivery. *Acta Oncol*, 49(7):964–971, Oct 2010. [PubMed Central:PMC3055798] [DOI:10.3109/0284186X.2010.498440] [PubMed:20831483].

[87] M. Leclerc, P. Maingon, M. Hamoir, C. Dalban, G. Calais, S. Nuyts, A. Serre, and V. Gregoire. A dose escalation study with intensity modulated radiation therapy (IMRT) in T2N0, T2N1, T3N0 squamous cell carcinomas (SCC) of the oropharynx, larynx and hypopharynx using a simultaneous integrated boost (SIB) approach. *Radiother Oncol*, 106(3):333–340, Mar

2013. [DOI:10.1016/j.radonc.2013.03.002] [PubMed:23541643].

[88] E. Lee, J. Zeng, R. S. Miyaoka, J. Saini, P. E. Kinahan, G. A. Sandison, T. Wong, H. J. Vesselle, R. Rengan, and S. R. Bowen. Functional lung avoidance and response-adaptive escalation (FLARE) RT: Multimodality plan dosimetry of a precision radiation oncology strategy. *Med Phys*, 44(7):3418–3429, Jul 2017. [PubMed Central:PMC5546910] [DOI:10.1002/mp.12308] [PubMed:28453861].

[89] H. Li, Y. Li, X. Zhang, X. Li, W. Liu, M. T. Gillin, and X. R. Zhu. Dynamically accumulated dose and 4D accumulated dose for moving tumors. *Med Phys*, 39(12):7359–7367, Dec 2012. [PubMed Central:PMC3523466] [DOI:10.1118/1.4766434] [PubMed:23231285].

[90] Y. Li, Z. Tian, T. Song, Z. Wu, Y. Liu, S. Jiang, and X. Jia. Monte Carlo-based Inverse Treatment Plan Optimization for Intensity Modulated Proton Therapy. *World Congress on Medical Physics and Biomedical Engineering, IFMBE Proceedings*, 51, Jun 2015. [PubMed Central:PMC5206994] [DOI:10.1007/978-3-319-19387-8_116].

[91] Y. Li, Z. Tian, T. Song, Z. Wu, Y. Liu, S. Jiang, and X. Jia. A new approach to integrate GPU-based Monte Carlo simulation into inverse treatment plan optimization for proton therapy. *Phys Med Biol*, 62(1):289–305, Jan 2017. [PubMed Central:PMC5206994] [DOI:10.1088/1361-6560/62/1/289] [PubMed:27991456].

[92] L. Lin, K. Souris, M. Kang, A. Glick, H. Lin, S. Huang, K. Stutzer, G. Janssens, E. Sterpin, J. A. Lee, T. D. Solberg, J. E. McDonough, C. B. Simone, and E. Ben-Josef. Evaluation of motion mitigation using abdominal compression in the clinical implementation of pencil beam scanning proton therapy of liver tumors. *Med Phys*, 44(2):703–712, Feb 2017. [DOI:10.1002/mp.12040] [PubMed:28133755].

[93] C. C. Ling, J. Humm, S. Larson, H. Amols, Z. Fuks, S. Leibel, and J. A. Koutcher. Towards multidimensional radiotherapy (MD-CRT): biological imaging and biological conformality. *Int. J. Radiat. Oncol. Biol. Phys.*, 47(3):551–560, Jun 2000. [PubMed:10837935].

[94] W. Liu, S. J. Frank, X. Li, Y. Li, P. C. Park, L. Dong, X. Ronald Zhu, and R. Mohan. Effectiveness of robust optimization in intensity-modulated proton therapy planning for head and neck cancers. *Med Phys*, 40(5):051711, May 2013. [PubMed Central:PMC3651255] [DOI:10.1118/1.4801899] [PubMed:23635259].

[95] W. Liu, S. J. Frank, X. Li, Y. Li, R. X. Zhu, and R. Mohan. PTV-based IMPT optimization incorporating planning risk volumes vs robust optimization. *Med Phys*, 40(2):021709, Feb 2013. [PubMed Central:PMC3562272] [DOI:10.1118/1.4774363] [PubMed:23387732].

[96] W. Liu, Z. Liao, S. E. Schild, Z. Liu, H. Li, Y. Li, P. C. Park, X. Li, J. Stoker, J. Shen, S. Keole, A. Anand, M. Fatyga, L. Dong, N. Sahoo, S. Vora, W. Wong, X. R. Zhu, M. Bues, and R. Mohan. Impact of respiratory motion on worst-case scenario optimized intensity modulated proton therapy for lung cancers. *Pract Radiat Oncol*, 5(2):77–86, 2015. [PubMed Central:PMC4355168] [DOI:10.1016/j.prro.2014.08.002] [PubMed:25413400].

[97] W. Liu, S. E. Schild, J. Y. Chang, Z. Liao, Y. H. Chang, Z. Wen, J. Shen, J. B. Stoker, X. Ding, Y. Hu, N. Sahoo, M. G. Herman, C. Vargas, S. Keole, W. Wong, and M. Bues. Exploratory Study of 4D versus 3D Robust Optimization in Intensity Modulated Proton Therapy for Lung Cancer. *Int. J. Radiat. Oncol. Biol. Phys.*, 95(1):523–533, May 2016. [PubMed Central:PMC4834263] [DOI:10.1016/j.ijrobp.2015.11.002] [PubMed:26725727].

[98] W. Liu, X. Zhang, Y. Li, and R. Mohan. Robust optimization of intensity modulated proton therapy. *Med Phys*, 39(2):1079–1091, Feb 2012. [PubMed Central:PMC3281975] [DOI:10.1118/1.3679340] [PubMed:22320818].

[99] A. J. Lomax. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 1: the potential effects of calculational uncertainties. *Phys Med Biol*, 53(4):1027–1042, Feb 2008. [DOI:10.1088/0031-9155/53/4/014] [PubMed:18263956].

[100] A. J. Lomax. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: the potential effects of inter-fraction and inter-field motions. *Phys Med Biol*, 53(4):1043–1056, Feb 2008. [DOI:10.1088/0031-9155/53/4/015] [PubMed:18263957].

[101] A. J. Lomax, E. Pedroni, H. Rutz, and G. Goitein. The clinical potential of intensity modulated proton therapy. *Z Med Phys*, 14(3):147–152, 2004. [PubMed:15462415].

[102] E. Lopci, I. Grassi, A. Chiti, C. Nanni, G. Cicoria, L. Toschi, C. Fonti, F. Lodi, S. Mattioli, and S. Fanti. PET radiopharmaceuticals for imaging of tumor hypoxia: a review of the evidence. *Am J Nucl Med Mol Imaging*, 4(4):365–384, 2014. [PubMed Central:PMC4074502] [PubMed:24982822].

[103] W. Lu. A non-voxel-based broad-beam (NVBB) framework for IMRT treatment planning. *Phys Med Biol*, 55(23):7175–7210, Dec 2010. [DOI:10.1088/0031-9155/55/23/002] [PubMed:21081819].

[104] C. M. Ma, J. S. Li, S. B. Jiang, T. Pawlicki, W. Xiong, L. H. Qin, and J. Yang. Effect of statistical uncertainties on Monte Carlo treatment planning. *Phys Med Biol*, 50(5):891–907, Mar 2005. [DOI:10.1088/0031-9155/50/5/013] [PubMed:15798263].

[105] J. Ma, C. Beltran, H. Seum Wan Chan Tseung, and M. G. Herman.

A GPU-accelerated and Monte Carlo-based intensity modulated proton therapy optimization system. *Med Phys*, 41(12):121707, Dec 2014. [DOI:10.1118/1.4901522] [PubMed:25471954].

[106] M. Machtay, R. Paulus, J. Moughan, R. Komaki, J. E. Bradley, H. Choy, K. Albain, B. Movsas, W. T. Sause, and W. J. Curran. Defining local-regional control and its importance in locally advanced non-small cell lung carcinoma. *J Thorac Oncol*, 7(4):716–722, Apr 2012. [PubMed Central:PMC3335930] [DOI:10.1097/JTO.0b013e3182429682] [PubMed:22425920].

[107] I. Madani, F. Duprez, T. Boterberg, C. Van de Wiele, K. Bonte, P. Deron, W. De Gersem, M. Coghe, and W. De Neve. Maximum tolerated dose in a phase I trial on adaptive dose painting by numbers for head and neck cancer. *Radiother Oncol*, 101(3):351–355, Dec 2011. [DOI:10.1016/j.radonc.2011.06.020] [PubMed:21742392].

[108] L. B. Marks, S. M. Bentzen, J. O. Deasy, F. M. Kong, J. D. Bradley, I. S. Vogelius, I. El Naqa, J. L. Hubbs, J. V. Lebesque, R. D. Timmerman, M. K. Martel, and A. Jackson. Radiation dose-volume effects in the lung. *Int. J. Radiat. Oncol. Biol. Phys.*, 76(3 Suppl):S70–76, Mar 2010. [PubMed Central:PMC3576042] [DOI:10.1016/j.ijrobp.2009.06.091] [PubMed:20171521].

[109] A. Marusyk, V. Almendro, and K. Polyak. Intra-tumour heterogeneity: a looking glass for cancer? *Nat. Rev. Cancer*, 12(5):323–334, Apr 2012. [DOI:10.1038/nrc3261] [PubMed:22513401].

[110] MathWorks documentation MATLAB sparse matrix format. https://nl.mathworks.com/help/matlab/ref/sparse.html.

[111] A. McKenzie, M. van Herk, and B. Mijnheer. Margins for geometric uncertainty around organs at risk in radiotherapy. *Radiother Oncol*, 62(3):299–307, Mar 2002. [PubMed:12175561].

[112] N. P. Mendenhall, R. S. Malyapa, Z. Su, D. Yeung, W. M. Mendenhall, and Z. Li. Proton therapy for head and neck cancer: rationale, potential indications, practical considerations, and current clinical evidence. *Acta Oncol*, 50(6):763–771, Aug 2011. [DOI:10.3109/0284186X.2011.590147] [PubMed:21767172].

[113] L. S. Mortensen, J. Johansen, J. Kallehauge, H. Primdahl, M. Busk, P. Lassen, J. Alsner, B. S. S?rensen, K. Toustrup, S. Jakobsen, J. Petersen, H. Petersen, J. Theil, M. Nordsmark, and J. Overgaard. FAZA PET/CT hypoxia imaging in patients with squamous cell carcinoma of the head and neck treated with radiotherapy: results from the DAHANCA 24 trial. *Radiother Oncol*, 105(1):14–20, Oct 2012. [DOI:10.1016/j.radonc.2012.09.015] [PubMed:23083497].

[114] M. F. Moyers, D. W. Miller, D. A. Bush, and J. D. Slater. Methodologies and tools for proton beam design for lung tumors. *Int. J. Radiat. Oncol. Biol. Phys.*, 49(5):1429–1438, Apr 2001. [PubMed:11286851].

[115] C. M. Nutting, C. M. Corbishley, B. Sanchez-Nieto, V. P. Cosgrove, S. Webb, and D. P. Dearnaley. Potential improvements in the therapeutic ratio of prostate cancer irradiation: dose escalation of pathologically identified tumour nodules using intensity modulated radiotherapy. *Br J Radiol*, 75(890):151–161, Feb 2002. [DOI:10.1259/bjr.75.890.750151] [PubMed:11893639].

[116] U. Oelfke and T. Bortfeld. Inverse planning for photon and proton beams. *Med Dosim*, 26(2):113–124, 2001. [PubMed:11444513].

[117] World Health Organization. Geant4 - a simulation toolkit. *Nucl Instrum Methods Phys Res*, 506(3):250–303, Dec 2003. 10.1016/S0168-9002(03)01368-8.

[118] World Health Organization. Cancer, key facts. Feb 2017. http://www.who.int/mediacentre/factsheets/fs297/en/.

[119] H. "Paganetti. *Proton therapy physics*. 2012. CRC Press, ISBN 9781439836446.

[120] H. Paganetti. Range uncertainties in proton therapy and the role of Monte Carlo simulations. *Phys Med Biol*, 57(11):99–117, Jun 2012. [PubMed Central:PMC3374500] [PubMed:22571913].

[121] H. Paganetti. Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer. *Phys Med Biol*, 59(22):R419–472, Nov 2014. [DOI:10.1088/0031-9155/59/22/R419] [PubMed:25361443].

[122] Harald Paganetti. *Proton Beam Therapy*. 2399-2891. IOP Publishing, 2017.

[123] M. O. Palumbo, P. Kavan, W. H. Miller, L. Panasci, S. Assouline, N. Johnson, V. Cohen, F. Patenaude, M. Pollak, R. T. Jagoe, and G. Batist. Systemic cancer therapy: achievements and challenges that lie ahead. *Front Pharmacol*, 4:57, 2013. [PubMed Central:PMC3646247] [DOI:10.3389/fphar.2013.00057] [PubMed:23675348].

[124] P. C. Park, X. R. Zhu, A. K. Lee, N. Sahoo, A. D. Melancon, L. Zhang, and L. Dong. A beam-specific planning target volume (PTV) design for proton therapy to account for setup and range uncertainties. *Int. J. Radiat. Oncol. Biol. Phys.*, 82(2):e329–336, Feb 2012. [PubMed Central:PMC3183263] [DOI:10.1016/j.ijrobp.2011.05.011] [PubMed:21703781].

[125] M. J. Parkes, S. Green, A. M. Stevens, S. Parveen, R. Stephens, and T. H. Clutton-Brock. Reducing the within-patient variability of breathing for radiotherapy delivery in conscious, unsedated cancer patients using a mechanical ventilator. *Br J Radiol*, 89(1062):20150741, Jun 2016. [PubMed

Central:PMC5258146] [DOI:10.1259/bjr.20150741] [PubMed:26959610].

[126] Z. Perko, S. R. van der Voort, S. van de Water, C. M. Hartman, M. Hoogeman, and D. Lathouwers. Fast and accurate sensitivity analysis of IMPT treatment plans using Polynomial Chaos Expansion. *Phys Med Biol*, 61(12):4646–4664, Jun 2016. [DOI:10.1088/0031-9155/61/12/4646] [PubMed:27227661].

[127] J. Perl, J. Shin, J. Schumann, B. Faddegon, and H. Paganetti. TOPAS: an innovative proton Monte Carlo platform for research and clinical applications. *Med Phys*, 39(11):6818–6837, Nov 2012. [PubMed Central:PMC3493036] [DOI:10.1118/1.4758060] [PubMed:23127075].

[128] D. Pflugfelder, J. J. Wilkens, and U. Oelfke. Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy. *Phys Med Biol*, 53(6):1689–1700, Mar 2008. [DOI:10.1088/0031-9155/53/6/013] [PubMed:18367797].

[129] The Wall Street Journal Proton-Beam Therapy for Cancer Gets Renewed Attention. https://www.wsj.com/articles/proton-beam-therapy-for-cancer-gets-renewed-attention-1463428188.

[130] Knoxville Provision Cares, Proton therapy center. https://www.provisionproton.com/ about-proton-therapy/advantages-of-proton/.

[131] Particle Therapy Co-Operative Group (PTCOG). Particle therapy facilities in operation. Jul 2017. https://www.ptcog.ch/index.php/facilities-in-operation.

[132] M. Rickhey, Z. Moravek, C. Eilles, O. Koelbl, and L. Bogner. 18F-FET-PET-based dose painting by numbers with protons. *Strahlenther Onkol*, 186(6):320–326, Jun 2010. [DOI:10.1007/s00066-010-2014-8] [PubMed:20559789].

[133] RTOG1106. https://www.rtog.org/ClinicalTrials/ProtocolTable/ StudyDetails.aspx?study=1106.

[134] D. Sanchez-Parcerisa, M. Kondrla, A. Shaindlin, and A. Carabe. FoCa: a modular treatment planning system for proton radiotherapy with research and educational purposes. *Phys Med Biol*, 59(23):7341–7360, Dec 2014. [DOI:10.1088/0031-9155/59/23/7341] [PubMed:25387249].

[135] A. Schatti, D. Meer, and A. J. Lomax. First experimental results of motion mitigation by continuous line scanning of protons. *Phys Med Biol*, 59(19):5707–5723, Oct 2014. [DOI:10.1088/0031-9155/59/19/5707] [PubMed:25197938].

[136] J. T. Schiller and D. R. Lowy. Virus infection and human cancer: an overview. *Recent Results Cancer Res.*, 193:1–10, 2014. [DOI:10.1007/978-3-642-38965-8_1] [PubMed:24008290].

[137] U. Schneider, E. Pedroni, and A. Lomax. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Phys Med Biol*, 41(1):111–124, Jan 1996. [PubMed:8685250].

[138] J. Seco, G. Gu, T. Marcelos, H. Kooy, and H. Willers. Proton arc reduces range uncertainty effects and improves conformality compared with photon volumetric modulated arc therapy in stereotactic body radiation therapy for non-small cell lung cancer. *Int. J. Radiat. Oncol. Biol. Phys.*, 87(1):188–194, Sep 2013. [DOI:10.1016/j.ijrobp.2013.04.048] [PubMed:23920395].

[139] S. Servagi-Vernat, S. Differding, E. Sterpin, F. X. Hanin, D. Labar, A. Bol, J. A. Lee, and V. Gregoire. Hypoxia-guided adaptive radiation dose escalation in head and neck carcinoma: a planning study. *Acta Oncol*, 54(7):1008–1016, Jul 2015. [DOI:10.3109/0284186X.2014.990109] [PubMed:25562382].

[140] J. V. Siebers. The effect of statistical noise on IMRT plan quality and convergence for MC-based and MC-correction-based optimized treatment plans. *J Phys Conf Ser*, 102:12020, Apr 2008. [PubMed Central:PMC2818598] [DOI:10.1088/1742-6596/102/1/012020] [PubMed:20148126].

[141] J. V. Siebers, I. Kawrakow, and V. Ramakrishnan. Performance of a hybrid MC dose algorithm for IMRT optimization dose evaluation. *Med Phys*, 34(7):2853–2863, Jul 2007. [DOI:10.1118/1.2745236] [PubMed:17821993].

[142] J. V. Siebers, M. Lauterbach, S. Tong, Q. Wu, and R. Mohan. Reducing dose calculation time for accurate iterative IMRT planning. *Med Phys*, 29(2):231–237, Feb 2002. [DOI:10.1118/1.1446112] [PubMed:11865994].

[143] M. Soret, S. L. Bacharach, and I. Buvat. Partial-volume effect in PET tumor imaging. *J. Nucl. Med.*, 48(6):932–945, Jun 2007. [DOI:10.2967/jnumed.106.035774] [PubMed:17504879].

[144] J. Sorriaux, M. Testa, H. Paganetti, J. Orban de Xivry, J. A. Lee, E. Traneus, K. Souris, S. Vynckier, and E. Sterpin. Experimental assessment of proton dose calculation accuracy in inhomogeneous media. *Phys Med*, 38:10–15, Jun 2017. [DOI:10.1016/j.ejmp.2017.04.020] [PubMed:28610689].

[145] M. Soukup, M. Fippel, and M. Alber. A pencil beam algorithm for intensity modulated proton therapy derived from Monte Carlo simulations. *Phys Med Biol*, 50(21):5089–5104, Nov 2005. [DOI:10.1088/0031-9155/50/21/010] [PubMed:16237243].

[146] K Souris, A Barragan Montero, E Sterpin, D Di Perri, X Geets, and J Lee. Mo-fg-campus-tep3-01: A model of baseline shift to improve robustness of proton therapy treatments of moving tumors. *Medical Physics*, 43(6Part32):3727–3727, 2016.

[147] K. Souris, J. A. Lee, and E. Sterpin. Fast multipurpose Monte Carlo simulation for proton therapy using multi- and many-core CPU archi-

tectures. *Med Phys*, 43(4):1700, Apr 2016. [DOI:10.1118/1.4943377] [PubMed:27036568].

[148] SNOPT Standford Business Software Inc. http://www.sbsi -sol-optimize.com/asp/sol_product_snopt.htm.

[149] S. Stapleton, S. Zavgorodni, I. A. Popescu, and W. A. Beckham. Implementation of random set-up errors in Monte Carlo calculated dynamic IMRT treatment plans. *Phys Med Biol*, 50(3):429–439, Feb 2005. [PubMed:15773721].

[150] N. A. Stavreva, P. V. Stavrev, and W. H. Round. A mathematical approach to optimizing the radiation dose distribution in heterogeneous tumours. *Acta Oncol*, 35(6):727–732, 1996. [PubMed:8938221].

[151] M. Steneker, A. Lomax, and U. Schneider. Intensity modulated photon and proton therapy for the treatment of head and neck tumors. *Radiother Oncol*, 80(2):263–267, Aug 2006. [DOI:10.1016/j.radonc.2006.07.025] [PubMed:16916557].

[152] E. Sterpin, S. Differding, G. Janssens, X. Geets, V. Gregoire, and J. A. Lee. Generation of prescriptions robust against geometric uncertainties in dose painting by numbers. *Acta Oncol*, 54(2):253–260, Feb 2015. [DOI:10.3109/0284186X.2014.930171] [PubMed:24991892].

[153] E. Sterpin, G. Janssens, J. Orban de Xivry, S. Goossens, M. Wanet, J. A. Lee, A. Delor, V. Bol, S. Vynckier, V. Gregoire, and X. Geets. Helical tomotherapy for SIB and hypo-fractionated treatments in lung carcinomas: a 4D Monte Carlo treatment planning study. *Radiother Oncol*, 104(2):173–180, Aug 2012. [DOI:10.1016/j.radonc.2012.06.005] [PubMed:22841518].

[154] IBM ILOG CPLEX Optimization Studio. https://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/.

[155] M. Stuschke, A. Kaiser, J. Abu-Jawad, C. Pottgen, S. Levegrun, and J. Farr. Re-irradiation of recurrent head and neck carcinomas: comparison of robust intensity modulated proton therapy treatment plans with helical tomotherapy. *Radiat Oncol*, 8:93, Apr 2013. [PubMed Central:PMC3648492] [DOI:10.1186/1748-717X-8-93] [PubMed:23601204].

[156] S. Surti, A. Kuhn, M. E. Werner, A. E. Perkins, J. Kolthammer, and J. S. Karp. Performance of Philips Gemini TF PET/CT scanner with special consideration for its time-of-flight imaging capabilities. *J. Nucl. Med.*, 48(3):471–480, Mar 2007. [PubMed:17332626].

[157] A. S?vik, E. Malinen, ?. S. Bruland, S. M. Bentzen, and D. R. Olsen. Optimization of tumour control probability in hypoxic tumours by radiation dose redistribution: a modelling study. *Phys Med Biol*, 52(2):499–513, Jan 2007. [DOI:10.1088/0031-9155/52/2/013] [PubMed:17202629].

[158] D. Thorwarth, S. M. Eschmann, F. Paulsen, and M. Alber. Hypoxia

dose painting by numbers: a planning study. *Int. J. Radiat. Oncol. Biol. Phys.*, 68(1):291–300, May 2007. [DOI:10.1016/j.ijrobp.2006.11.061] [PubMed:17448882].

[159] D. Thorwarth, X. Geets, and M. Paiusco. Physical radiotherapy treatment planning based on functional PET/CT data. *Radiother Oncol*, 96(3):317–324, Sep 2010. [DOI:10.1016/j.radonc.2010.07.012] [PubMed:20673689].

[160] D. Thorwarth, M. Soukup, and M. Alber. Dose painting with IMPT, helical tomotherapy and IMXT: a dosimetric comparison. *Radiother Oncol*, 86(1):30–34, Jan 2008. [DOI:10.1016/j.radonc.2007.11.003] [PubMed:18061294].

[161] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal. Global cancer statistics, 2012. *CA Cancer J Clin*, 65(2):87–108, Mar 2015. [DOI:10.3322/caac.21262] [PubMed:25651787].

[162] L. A. Torre, R. L. Siegel, E. M. Ward, and A. Jemal. Global Cancer Incidence and Mortality Rates and Trends–An Update. *Cancer Epidemiol. Biomarkers Prev.*, 25(1):16–27, Jan 2016. [DOI:10.1158/1055-9965.EPI-15-0578] [PubMed:26667886].

[163] D. Trani, A. Yaromina, L. Dubois, M. Granzier, S. G. Peeters, R. Biemans, G. Nalbantov, N. Lieuwes, B. Reniers, E. E. Troost, F. Verhaegen, and P. Lambin. Preclinical Assessment of Efficacy of Radiation Dose Painting Based on Intratumoral FDG-PET Uptake. *Clin. Cancer Res.*, 21(24):5511–5518, Dec 2015. [DOI:10.1158/1078-0432.CCR-15-0290] [PubMed:26276892].

[164] S. Tyldesley, G. Delaney, F. Foroudi, L. Barbera, M. Kerba, and W. Mackillop. Estimating the need for radiotherapy for patients with prostate, breast, and lung cancers: verification of model estimates of need with radiotherapy utilization data from British Columbia. *Int. J. Radiat. Oncol. Biol. Phys.*, 79(5):1507–1515, Apr 2011. [DOI:10.1016/j.ijrobp.2009.12.070] [PubMed:20605338].

[165] J. Unkelbach, T. Bortfeld, B. C. Martin, and M. Soukup. Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning. *Med Phys*, 36(1):149–163, Jan 2009. [PubMed Central:PMC2673668] [DOI:10.1118/1.3021139] [PubMed:19235384].

[166] J. Unkelbach, T. C. Chan, and T. Bortfeld. Accounting for range uncertainties in the optimization of intensity modulated proton therapy. *Phys Med Biol*, 52(10):2755–2773, May 2007. [DOI:10.1088/0031-9155/52/10/009] [PubMed:17473350].

[167] J. Uzan, A. E. Nahum, and I. Syndikus. Prostate Dose-painting Radiotherapy and Radiobiological Guided Optimisation Enhances the Ther-

apeutic Ratio. *Clin Oncol (R Coll Radiol)*, 28(3):165–170, Mar 2016. [DOI:10.1016/j.clon.2015.09.006] [PubMed:26482453].

[168] S. van de Water, I. van Dam, D. R. Schaart, A. Al-Mamgani, B. J. Heijmen, and M. S. Hoogeman. The price of robustness; impact of worst-case optimization on organ-at-risk dose and complication probability in intensity-modulated proton therapy for oropharyngeal cancer patients. *Radiother Oncol*, 120(1):56–62, Jul 2016. [DOI:10.1016/j.radonc.2016.04.038] [PubMed:27178142].

[169] T. A. van de Water, H. P. Bijl, C. Schilstra, M. Pijls-Johannesma, and J. A. Langendijk. The potential benefit of radiotherapy with protons in head and neck cancer with respect to normal tissue sparing: a systematic review of literature. *Oncologist*, 16(3):366–377, 2011. [PubMed Central:PMC3228110] [DOI:10.1634/theoncologist.2010-0171] [PubMed:21349950].

[170] U. A. van der Heide, A. C. Houweling, G. Groenendaal, R. G. Beets-Tan, and P. Lambin. Functional MRI for radiotherapy dose painting. *Magn Reson Imaging*, 30(9):1216–1223, Nov 2012. [PubMed Central:PMC5134673] [DOI:10.1016/j.mri.2012.04.010] [PubMed:22770686].

[171] S. van der Voort, S. van de Water, Z. Perko, B. Heijmen, D. Lathouwers, and M. Hoogeman. Robustness Recipes for Minimax Robust Optimization in Intensity Modulated Proton Therapy for Oropharyngeal Cancer Patients. *Int. J. Radiat. Oncol. Biol. Phys.*, 95(1):163–170, May 2016. [DOI:10.1016/j.ijrobp.2016.02.035] [PubMed:27084639].

[172] W. van Elmpt, D. De Ruysscher, A. van der Salm, A. Lakeman, J. van der Stoep, D. Emans, E. Damen, M. Ollers, J. J. Sonke, and J. Belderbos. The PET-boost randomised phase II dose-escalation trial in non-small cell lung cancer. *Radiother Oncol*, 104(1):67–71, Jul 2012. [DOI:10.1016/j.radonc.2012.03.005] [PubMed:22483675].

[173] M. van Herk. Errors and margins in radiotherapy. *Semin Radiat Oncol*, 14(1):52–64, Jan 2004. [DOI:10.1053/j.semradonc.2003.10.003] [PubMed:14752733].

[174] M. van Herk, P. Remeijer, C. Rasch, and J. V. Lebesque. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 47(4):1121–1135, Jul 2000. [PubMed:10863086].

[175] B. Vanderstraeten, W. Duthoy, W. De Gersem, W. De Neve, and H. Thierens. [18F]fluoro-deoxy-glucose positron emission tomography ([18F]FDG-PET) voxel intensity-based intensity-modulated radiation therapy (IMRT) for head and neck cancer. *Radiother Oncol*, 79(3):249–258, Jun 2006. [DOI:10.1016/j.radonc.2006.03.003] [PubMed:16564588].

[176] P. Vera, S. Thureau, P. Chaumet-Riffaud, R. Modzelewski, P. Bohn, M. Ver-

mandel, S. Hapdey, A. Pallardy, M. A. Mahe, M. Lacombe, P. Boisselier, S. Guillemard, P. Olivier, V. Beckendorf, N. Salem, N. Charrier, E. Chajon, A. Devillers, N. Aide, S. Danhier, F. Denis, J. P. Muratet, E. Martin, A. B. Riedinger, H. Kolesnikov-Gauthier, E. Dansin, C. Massabeau, F. Courbon, M. P. Farcy Jacquet, P. O. Kotzki, C. Houzard, F. Mornex, L. Vervueren, A. Paumier, P. Fernandez, M. Salaun, and B. Dubray. Phase II Study of a Radiotherapy Total Dose Increase in Hypoxic Lesions Identified by (18)F-Misonidazole PET/CT in Patients with Non-Small Cell Lung Carcinoma (RTEP5 Study). *J. Nucl. Med.*, 58(7):1045–1053, Jul 2017. [DOI:10.2967/jnumed.116.188367] [PubMed:28254869].

[177] A. Wachter and L. T. Biegler. On the Implementation of a Primal-Dual Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming. *Mathematical Programming*, 106(1):25–57, Mar 2006. [[DOI:10.1007/s10107-004-0559-y].

[178] M. Wanet, E. Sterpin, G. Janssens, A. Delor, J. A. Lee, and X. Geets. Validation of the mid-position strategy for lung tumors in helical TomoTherapy. *Radiother Oncol*, 110(3):529–537, Mar 2014. [DOI:10.1016/j.radonc.2013.10.025] [PubMed:24424385].

[179] S. Webb and A. E. Nahum. A model for calculating tumour control probability in radiotherapy including the effects of inhomogeneous distributions of dose and clonogenic cell density. *Phys Med Biol*, 38(6):653–666, Jun 1993. [PubMed:8346278].

[180] MCsquare website. http://www.openmcsquare.org/.

[181] MIROpt website. http://www.openmiropt.org/.

[182] L. Widesott, A. J. Lomax, and M. Schwarz. Is there a single spot size and grid for intensity modulated proton therapy? Simulation of head and neck, prostate and mesothelioma cases. *Med Phys*, 39(3):1298–1308, Mar 2012. [DOI:10.1118/1.3683640] [PubMed:22380362].

[183] H. P. Wieser, E. Cisternas, N. Wahl, S. Ulrich, A. Stadler, H. Mescher, L. R. Muller, T. Klinge, H. Gabrys, L. Burigo, A. Mairani, S. Ecker, B. Ackermann, M. Ellerbrock, K. Parodi, O. Jakel, and M. Bangert. Development of the open-source dose calculation and optimization toolkit matRad. *Med Phys*, 44(6):2556–2568, Jun 2017. [DOI:10.1002/mp.12251] [PubMed:28370020].

[184] R. R. WILSON. Radiological use of fast protons. *Radiology*, 47(5):487–491, Nov 1946. [DOI:10.1148/47.5.487] [PubMed:20274616].

[185] M. Witte, G. Shakirin, A. Houweling, H. Peulen, and M. van Herk. Dealing with geometric uncertainties in dose painting by numbers: introducing the Î"VH. *Radiother Oncol*, 100(3):402–406, Sep 2011. [DOI:10.1016/j.radonc.2011.08.028] [PubMed:21937139].

[186] J. W. Wolthaus, J. J. Sonke, M. van Herk, J. S. Belderbos, M. M. Rossi, J. V. Lebesque, and E. M. Damen. Comparison of different strategies to use four-dimensional computed tomography in treatment planning for lung cancer patients. *Int. J. Radiat. Oncol. Biol. Phys.*, 70(4):1229–1238, Mar 2008. [DOI:10.1016/j.ijrobp.2007.11.042] [PubMed:18313530].

[187] J. W. Wolthaus, J. J. Sonke, M. van Herk, and E. M. Damen. Reconstruction of a time-averaged midposition CT scan for radiotherapy planning of lung cancer patients using deformable registration. *Med Phys*, 35(9):3998–4011, Sep 2008. [DOI:10.1118/1.2966347] [PubMed:18841851].

[188] Q. Wu, R. Mohan, A. Niemierko, and R. Schmidt-Ullrich. Optimization of intensity-modulated radiotherapy plans based on the equivalent uniform dose. *Int. J. Radiat. Oncol. Biol. Phys.*, 52(1):224–235, Jan 2002. [PubMed:11777642].

[189] D. Yan, F. Vicini, J. Wong, and A. Martinez. Adaptive radiation therapy. *Phys Med Biol*, 42(1):123–132, Jan 1997. [PubMed:9015813].

[190] P. P. Yepes, D. Mirkovic, and P. J. Taddei. A GPU implementation of a track-repeating algorithm for proton radiotherapy dose calculations. *Phys Med Biol*, 55(23):7107–7120, Dec 2010. [PubMed Central:PMC3001323] [DOI:10.1088/0031-9155/55/23/S11] [PubMed:21076192].

[191] Y. Zhang, I. Huth, M. Wegner, D. C. Weber, and A. J. Lomax. An evaluation of rescanning technique for liver tumour treatments using a commercial PBS proton therapy system. *Radiother Oncol*, 121(2):281–287, Nov 2016. [DOI:10.1016/j.radonc.2016.09.011] [PubMed:27726957].

[192] P. Ziegenhein, C. P. Kamerling, M. Bangert, J. Kunkel, and U. Oelfke. Performance-optimized clinical IMRT planning on modern CPUs. *Phys Med Biol*, 58(11):3705–3715, Jun 2013. [DOI:10.1088/0031-9155/58/11/3705] [PubMed:23656861].

[193] P. Ziegenhein, S. Pirner, C. Ph Kamerling, and U. Oelfke. Fast CPU-based Monte Carlo simulation for radiotherapy dose calculation. *Phys Med Biol*, 60(15):6097–6111, Aug 2015. [DOI:10.1088/0031-9155/60/15/6097] [PubMed:26216484].

[194] P. Ziegenhein, J. J. Wilkens, S. Nill, T. Ludwig, and U. Oelfke. Speed optimized influence matrix processing in inverse treatment planning tools. *Phys Med Biol*, 53(9):N157–164, May 2008. [DOI:10.1088/0031-9155/53/9/N02] [PubMed:18401066].