

Accurate assessment of proton therapy treatments:

**Fast Monte Carlo dose engine
and extensive robustness tests**

Kevin Souris

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

October 21, 2017: Manuscript submission
November 21, 2017: Private defence
January 9, 2018: Public defence

During this thesis, Kevin Souris was supported by a research grant from Ion Beam Applications (IBA s.a.), by a Télévie grant from the Belgian Fonds National pour la Recherche Scientifique (F.R.S. - FNRS) with the project number 7.4591.15, and by a grant from Walloon Region under the project name InVivoIGT, convention number 1017266.



PhD committee

Thesis supervisors

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. Vincent Grégoire, PhD, MD

Radiation Oncology, Cliniques Universitaires Saint Luc
Molecular Imaging, Radiotherapy and Oncology, Université catholique de Louvain

Members

Prof. Eduardo Cortina Gil, PhD

Center for Cosmology, Particle Physics and Phenomenology, Université catholique de Louvain

Prof. Xavier Geets, PhD, MD

Radiation Oncology, Cliniques Universitaires Saint Luc
Molecular Imaging, Radiotherapy and Oncology, Université catholique de Louvain

Prof. Stefaan Vynckier, PhD

Radiation Oncology, Cliniques Universitaires Saint Luc

Invited members

Prof. Antje Knopf, PhD

Faculty of Medical Sciences, University of Groningen

University Medical Center Groningen

The Netherlands

Prof. Liyong Lin, PhD

Perelman School of Medicine, University of Pennsylvania

Department of Radiation Oncology, University of Pennsylvania

United States of America

Acknowledgments

This thesis represents the achievement of a long road toward the PhD. Overall, this was a great adventure for me. However, it was not always an obstacle-free way. Therefore, I would like to thank all people that contributed directly or indirectly to this thesis by supporting me and helping me during my journey.

It often requires strong motivations to engage in such long studies and enter the world of research. For me, this choice was inspired by a few persons that marked my training course. The story began in the physics class at secondary school. This was my very first introduction to modern physics (and even medical physics), transmitted with passion and pedagogy by our great teacher Marie-Cécile Lefebvre. Later, at the end of my bachelor and part of my master degree, I found a high interest in experimental physics with Eduardo Cortina. I really enjoyed designing and building exotic particle detectors with him. Thanks to his expertise and the time he invested, we successfully developed a spark chamber and a cosmic bench as extracurricular projects. Then, I continued playing with particle detectors during my master thesis, but this time oriented for medical applications. Indeed, I had the opportunity to work on PET image acquisition with Anne Bol and PET image reconstruction with John Lee. Realizing my master thesis with both of them allowed me to confirm my enthusiasm for research. Therefore, I would like to thank all these persons that inspired me and guided me toward the PhD.

The accomplishment of my thesis would certainly not be possible without my supervisors, John Lee and Edmond Sterpin. More than supervisors, they are also my mentors. They were always present to advice me and keep me in the right path. I would like to thank both of them for the time and the effort they dedicated to me. More specifically, thank you John for all your judicious and wise recommendations. It is always with a great pleasure and interest that I discuss with you. Thank you Edmond for your reflections and the real coaching that you provide with passion and a touch of humor.

I also would like to sincerely thank the other members of my PhD committee, Vincent Grégoire, Xavier Geets, Stefaan Vynckier, and Eduardo Cortina, who guided me during the past years. In addition, I thank the entire PhD jury, composed of the PhD committee and two external members, Antje Knopf and Liyong Lin. Thank you for the time you spent reading my manuscript, the interesting discussions during the defence, and your judicious comments.

During my PhD, I really enjoyed working in the MIRO lab, which regroups people from diverse scientific fields. I would like to thank all my colleagues from MIRO for the great atmosphere we constantly maintained, and for the sharing and mutual help that was always present. Especially, I appreciated very much working with all my office mates. Namely, Guillaume, Marie W., and Samuel who were already present when I started my PhD, but also Ana, Kinga, Simona, Marie C., and Umair who joined us later. A special thank you to Guillaume and Ana for the complicity and friendship we developed over the last years. I also would like to thank other colleagues with whom I had the chance to work, including John, Edmond, Jefferson, Anne, Séverine, Sarah, Dario, Geneviève, Karolin, and many more...

I had the opportunity to collaborate with several institutions around the world during my PhD. In particular, I stayed 9 months at the University of Pennsylvania, in the USA. I would like to thank all people I met there for their warm welcome. Especially, I thank Liyong Lin, for the time he invested in paperwork to make my stay in US possible and then for his supervision. Thank you Liyong for the clinical expertise that you shared with me. I also would like to thank my two colleagues and friends Sheng Huang and Minglei Kang from UPenn, who worked with me and still continue to collaborate today on so many projects. Many thanks to you guys for the hospitality you offered to me and for your friendship.

My work was partly funded by Ion Beam Applications (IBA s.a.). I would like to sincerely thank IBA, not only for their financial support, but also for all the resources and expertise they provided. Especially, I thank Guillaume Janssens for his supervision and all his scientific and technical contributions. I also thank Rudi Labarbe for the great management and support he provided. My research work was also financially supported by the Walloon region and a Télévie grant from FRS-FNRS.

Finally, this adventure would not be possible without the support of my family. Many thanks to my parents and my brother that constantly took care of me,

always encouraged me during my studies, and approved my choices. I could never thank my girlfriend, Roxane, enough for all she brings to me. Thank you for your presence, your patience and your support during the good moments, as well as during the difficulties. I also would like to thank Roxane's family for their encouragements.

Abstract

Radiation therapy is one of the main treatments for cancer care. It consists in irradiating the tumor, while limiting the toxicity associated with the exposure of healthy tissues. Proton therapy is an emerging radiation delivery modality, which has the potential to better spare healthy tissues than conventional radiotherapy treatments. However, this new modality is much more sensitive to treatment uncertainties, such as patient anatomy changes. To address the lack of robustness in proton therapy, this thesis provides accurate treatment preparation tools, such as fast Monte Carlo dose calculation and robust planning methods. Furthermore, a comprehensive and realistic treatment robustness verification tool was developed in order to assess the sensitivity of the treatment plan to uncertainties. By combining these tools, proton therapy could be delivered more safely, improving the treatment outcome for the patient. All tools developed during this thesis are released open source and are already used in several institutions for research and clinical purposes.

List of publications

Articles in peer-reviewed journals

Fast multipurpose Monte Carlo simulation for proton therapy using multi- and many-core CPU architectures

K. Souris, J.A. Lee, E. Sterpin

Medical physics 2016; 43 (4): 1700-1712

Monte Carlo methods to realistically evaluate the robustness of proton therapy plans

K. Souris, A. Barragan, G. Janssens, D. Di Perri, E. Sterpin, J.A. Lee

Under review in Medical physics

Evaluation of motion mitigation using abdominal compression in the clinical implementation of pencil beam scanning proton therapy of liver tumors

L. Lin and 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, E. Ben-Josef

Medical physics 2017; 44 (2): 703-712

Performance of a hybrid Monte Carlo-Pencil Beam dose algorithm for proton therapy inverse planning

A. Barragan, K. Souris, D. Sanchez, E. Sterpin, J.A. Lee

Medical physics 2017

Efficient implementation of 4D robust optimization for scanned proton therapy with Monte Carlo

A. Barragan, K. Souris, E. Sterpin, J.A. Lee

Manuscript in preparation

Combined influence of CT random noise and HU-RSP calibration curve nonlinearities on proton range systematic errors

S. Brousmiche, K. Souris, J. Orban, J.A. Lee, B. Macq, J. Seco
Physics in Medicine and Biology 2017; 62 (21): 8226-8245

Monte Carlo Dose Calculation is recommended for the pencil beam scanning proton therapy treatment of superficial tumors

S. Huang, K. Souris, S. Li, M. Kang, G. Janssens, A. Lin, E. Garver, C. Ainsley, P. Taylor, Y. Xiao, L. Lin
Manuscript in preparation

Analytical LET calculation algorithm, implemented on top of the Eclipse scripting API

P. Kimstrand, K. Souris, A. Carabe
Manuscript in preparation

A Fano cavity test for Monte Carlo proton transport algorithms

E. Sterpin, J. Sorriaux, K. Souris, S. Vynckier, H. Bouchard
Medical physics 2014; 41 (1): 11706

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

S. Michiels, A. Barragan, K. Souris, K. Poels, W. Crijns, J.A. Lee, E. Sterpin, S. Nuyts, K. Haustermans, T. Depuydt
Radiotherapy & Oncology (Green journal) 2017

Robust treatment planning in proton therapy

E. Sterpin, A. Barragan, K. Souris, J.A. Lee
Cancer radiotherapie 2016; 20 (6-7): 523-529

LET dependence of the response of a PTW-60019 microDiamond detector in a 62MeV proton beam

S. Rossomme, J-M. Denis, K. Souris, A. Delor, F. Bartier, D. Dumont, S. Vynckier, H. Palmans
Physica Medica 2016; 32 (9): 1135-1138

Validation and clinical implementation of a full Monte Carlo code for pencil beam scanning proton therapy

S. Huang, M. Kang, K. Souris, C. Ainsley, T.D. Solberg, J.E. McDonough, C.B. Simone, L. Lin
Under review in JACMP

Feasibility of online IMPT adaptation using fast, automatic and robust dose restoration

K. Bernatowicz, A. Barragan, G. Janssens, [K. Souris](#), X. Geets, E. Sterpin

Under review in Physics in Medicine and Biology

Experimental assessment of proton dose calculation accuracy in inhomogeneous media

J. Sorriaux, M. Testa, H. Paganetti, J. Orban, J.A. Lee, E. Traneus, [K. Souris](#), S. Vynckier, E. Sterpin

Physica Medica 2017; 38: 10-15

Administration of continuous positive airway pressure in the context of radiation therapy for lung tumours

D. Di Perri, A. Colot, A. Delor, R. Ghoul, G. Janssens, V. Lacroix, P. Matte, A. Robert, [K. Souris](#), X. Geets

Under review in Strahlentherapie und Onkologie

International conference communications

Simulation of Motion Amplitude Variations to Verify the Robustness of Treatment Plans

[K. Souris](#), A. Barragan, G. Janssens, E. Sterpin, D. Di Perri, J.A. Lee

[SU-H3-GePD-T-2](#) AAPM 2017, Denver, CO (USA)

Simulate baseline shift uncertainties to improve robustness of proton therapy treatments

[K. Souris](#), A. Barragan, D. Di Perri, X. Geets, E. Sterpin, J.A. Lee

[EP-1643](#) ESTRO 2017, Vienna (Austria)

A Model of Baseline Shift to Improve Robustness of Proton Therapy Treatments of Moving Tumors

[K. Souris](#), A. Barragan, E. Sterpin, D. Di Perri, X. Geets, J.A. Lee

[MO-FG-CAMPUS-TeP3-01](#) AAPM 2016, Washington, DC (USA)

Abdominal Compression Effectively Reduces the Interplay Effect and Enables Pencil Beam Scanning Proton Therapy of Liver Tumors

[K. Souris](#), A. Glick, M. Kang, G. Janssens, E. Sterpin, H. Lin, J. McDonough, C. Simone, T. Solberg, E. Ben-Josef, J.A. Lee, L. Lin

[SU-F-T-121](#) AAPM 2016, Washington, DC (USA)

4D Monte Carlo Robustness Test for Proton Therapy

K. Souris, E. Sterpin, J.A. Lee

[TU-EF-304-03](#) AAPM 2015, Anaheim, CA (USA)

Fast and realistic Monte Carlo evaluation of the robustness of proton therapy plans

K. Souris, J.A. Lee, E. Sterpin

[PO-0887](#) ESTRO 2015, Barcelona (Spain)

Intel Xeon Phi Implementation of a Fast Multi-Purpose Monte Carlo Simulation for Proton Therapy

K. Souris, J.A. Lee, E. Sterpin

[TH-A-19A-08](#) AAPM 2014, Austin, TX (USA)

Fast Monte Carlo simulation of proton therapy treatment using an Intel Xeon Phi coprocessor

K. Souris, J.A. Lee, E. Sterpin

[OC-0273](#) ESTRO 2014 Vienna (Austria)

On the Importance of Nuclear Models On the Accuracy of Fast Monte Carlo Methods for Proton Therapy

K. Souris, J.A. Lee, E. Sterpin

[WE-F-105-01](#) AAPM 2013, Indianapolis, IN (USA)

A Platform for Adaptive Radiation-Therapy Research

A. Pin, S. Puydupin, L. Yin, K. Souris, G. Janssens, B. Teo

[SU-I-GPD-J-2](#) AAPM 2017, Denver, CO (USA)

Performance of a Hybrid Monte Carlo-Pencil Beam Dose Algorithm for Proton Therapy

A. Barragan, K. Souris, D. Sanchez, J.A. Lee, E. Sterpin

[SU-K-108-13](#) AAPM 2017, Denver, CO (USA)

Evaluation of Motion Mitigation Using Abdominal Compression in the Clinical Implementation of Pencil Beam Scanning Proton Therapy of Liver Tumors

L. Lin, K. Souris, M. Kang, A. Glick, H. Lin, S. Huang, K. Stuetzer, G. Janssens, E. Sterpin, J. Lee, T. Solberg, J. McDonough, C. Simone, E. Ben-Josef

[MO-L-GePD-TT-2](#) AAPM 2017, Denver, CO (USA)

Validation and Clinical Implementation of a Full Monte Carlo Code for Scanned Proton Pencil Beams

S. Huang, M. Kang, [K. Souris](#), C. Ainsley, T. Solberg, J. McDonough, C. Simone, L. Lin

[TU-H-CAMPUS-TT-5](#) AAPM 2017, Denver, CO (USA)

CPAP ventilation might allow better sparing of normal lung tissue during lung cancer radiotherapy

D. Di Perri, A. Colot, A. Barragan, G. Janssens, V. Lacroix, P. Matte, [K. Souris](#), X. Geets

[PO-0670](#) ESTRO 2017, Vienna (Austria)

Can bolus range shifting improve plan quality in the IMPT of head and neck cancer?

S. Michiels, A. Barragan, [K. Souris](#), K. Poels, W. Crijns, J.A. Lee, E. Sterpin, S. Nuyts, K. Haustermans, T. Depuydt

[EP-1590](#) ESTRO 2017, Vienna (Austria)

Implementation of a hybrid superfast Monte Carlo-Pencil beam dose optimizer for proton therapy

A. Barragan, [K. Souris](#), D. Sanchez, A. Carabe, J.A. Lee, E. Sterpin

[EP-1519](#) ESTRO 2017, Vienna (Austria)

Multi-scenario sampling in robust proton therapy treatment planning

E. Sterpin, A. Barragan, [K. Souris](#), J.A. Lee

[PO-0825](#) ESTRO 2017, Vienna (Austria)

A Monte Carlo-Based Method to Include Random Errors in Robust Optimization

A. Barragan, [K. Souris](#), J.A. Lee, E. Sterpin

[WE-AB-209-01](#) AAPM 2016, Washington, DC (USA)

Efficient implementation of random errors in robust optimization for proton therapy with Monte Carlo

A. Barragan, [K. Souris](#), E. Sterpin, J.A. Lee

[OC-0265](#) ESTRO 2016, Turin (Italy)

Assessing How Stochastic CT Noise Can Lead to Systematic Proton Range Errors

S. Brousmiche, [K. Souris](#), J. Orban, J.A. Lee, B. Macq, J. Seco

[TH-CD-BRA-04](#) AAPM 2015, Anaheim, CA (USA)

Evaluation of 6 proton dose calculation algorithms using a critical experimental phantom

J. Sorriaux, M. Testa, K. Souris, J. Orban, J.A. Lee, D. Bertrand, E. Traneus, H. Paganetti, S. Vynckier

[PO-0888](#) ESTRO 2015, Barcelona (Spain)

Combined influence of CT noise and HU-RSP conversion curve discontinuities on proton range systematic errors

S. Brousmitche, K. Souris, J. Orban, B. Macq, J. Seco

[PO-0883](#) ESTRO 2015, Barcelona (Spain)

Fano Cavity Test of Proton Transport in Monte Carlo Codes Running On GPU and Xeon Phi

E. Sterpin, J. Sorriaux, J. Schuemann, K. Souris, J.A. Lee, S. Vynckier, X. Jia, S. Jiang, H. Paganetti

[SU-E-T-180](#) AAPM 2014, Austin, TX (USA)

Contents

1	Introduction	3
1.1	Motivation and aim of the thesis	4
1.2	Structure of the thesis	6
1.3	General overview of radiation therapy	7
1.3.1	Cancer	7
1.3.2	Treatment modalities	8
1.3.3	Rationale of radiotherapy	9
1.3.4	Delivery of proton therapy	13
1.4	Proton therapy workflow and uncertainties	17
1.4.1	Pre-treatment imaging	17
1.4.2	Contouring and margins	18
1.4.3	Dose prescription	20
1.4.4	Dose calculation	20
1.4.5	Treatment plan optimization	22
1.4.6	Treatment plan verification and QA	24
1.4.7	Treatment delivery	27
1.4.8	Summary of proton range uncertainties	28
1.5	Contributions	28
1.5.1	Fast Monte Carlo simulations	29
1.5.2	Improved planning tools	30
1.5.3	Realistic robustness verification	31
2	Fast Monte Carlo simulations	33
2.1	Introduction	34
2.2	Methods	36
2.2.1	Input data	37
2.2.2	Transport algorithm	38
2.2.3	Electromagnetic interactions	39
2.2.4	Nuclear interactions	41
2.2.5	Xeon Phi architecture	42

2.2.6	Code implementation	43
2.2.7	Benchmark	45
2.3	Results	48
2.3.1	Homogeneous water phantom	49
2.3.2	Homogeneous tissue phantom	50
2.3.3	Heterogeneous phantom	50
2.3.4	Computation speeds	52
2.4	Discussion	55
2.5	Conclusions	58
3	Robust treatment planning	59
3.1	Motion evaluation	60
3.2	Beam-specific PTV	62
3.3	Beam angle selection	67
3.4	Robust optimization	69
4	Realistic verification of the treatment plan robustness	75
4.1	Introduction	77
4.2	Methods	79
4.2.1	Monte Carlo dose calculation	79
4.2.2	Setup errors	80
4.2.3	Range errors	81
4.2.4	Organ motion	81
4.2.5	Motion amplitude errors	81
4.2.6	The interplay effect	83
4.2.7	Sampling of uncertainty scenarios	84
4.2.8	Robustness analysis	85
4.2.9	Patient case	86
4.3	Results	87
4.3.1	Validation of motion amplitude variation model	87
4.3.2	Treatment plan robustness evaluation	88
4.4	Discussion	92
4.5	Conclusion	94
5	Complement for treatment plan evaluation	95
5.1	A model of the baseline shift	96
5.1.1	Method	96
5.1.2	Validation	98
5.1.3	Discussion	100
5.2	LET calculation	100

CONTENTS	1
6 Discussion and conclusion	103
References	113
List of acronyms	127

Chapter 1

Introduction

1.1 Motivation and aim of the thesis

This first section provides a quick presentation of the motivations and objectives of the thesis. A more detailed introduction of radiotherapy concepts is presented in the Section 1.3.

Radiation therapy is one of the main treatments for cancer [1], along with surgery, and systemic therapy. It consists in irradiating the tumor, most of the time using an external radiation beam. For this particular type of treatment, the patient survival depends on the ability to control the tumor and limit the toxicity associated with the exposure of healthy tissues to the irradiation. To achieve this goal, various radiation delivery modalities can be employed. X-ray beam is currently the most widespread radiation delivery. This mature technology is well controlled nowadays. However, the maximum dose deposited by photon beams is located at shallow depth in the patient, which is not optimal to treat deep seated tumors. In order to reduce the radiation exposure of healthy tissues, proton therapy is slowly emerging [2, 3, 4]. It consists in delivering proton beams instead of conventional X-ray beams. Indeed, most of the proton dose is deposited in depth, at the end of the proton path, forming a peak of dose called Bragg peak. This peak of dose can be positioned inside the tumor in order to maximize its irradiation [5], while sparing healthy tissues. To do so, various parameters of the treatment delivery can be adjusted.

The list of delivery parameters, that constitutes the treatment plan, is optimized according to the patient geometry and dose prescription using a treatment planning system (TPS). The depth of the Bragg peak depends on the energy of the proton and the materials crossed along its path. Therefore, the proton energy must be calculated accurately according to the patient anatomy and tissue composition in order to position the Bragg peak at the desired depth. Multiple Bragg peaks with different proton energies may be combined to cover the entire tumor volume. Due to the high dose gradients in the Bragg peak, a wrong estimation of the proton range may lead to severe distortions of the delivered dose distribution [6, 7], with potential under-coverage of the target and over-dosage in the organs at risk. Therefore, the treatment plan optimization requires an extensive and accurate knowledge of the patient geometry and atomic compositions at the time of the treatment delivery. However, in the current clinical practice, the beam parameters are optimized several days before the beginning of the treatment, based on pre-treatment patient images. Due to the potential changes in the patient anatomy and in his positioning on the treatment couch, the treatment plan might not be optimal anymore.

In addition to the anatomical changes, multiple sources of proton range uncertainties have been identified [8]. For a static geometry, most of the uncertainties come from the TPS dose calculation. This process is composed of several operations, each introducing some uncertainties. First, the patient image is converted in stopping powers [9], which characterize the rate of proton energy loss in matter. As this conversion is not perfectly accurate, it may introduce a systematic error in the evaluation of the proton range. The second step consists in calculating the total dose distribution according to the treatment plan and the stopping powers map. Two types of algorithms are available for proton therapy dose calculation. Analytical algorithms rely on macroscopic models [10], yielding a fast computation, but with reduced accuracy in some conditions. Monte Carlo algorithms are based on microscopic simulations of proton interactions [11]. This provides a very good accuracy, but at the price of a long computation time. It has been shown that the uncertainty on the proton range can be significantly reduced using such Monte Carlo algorithms [8], especially for heterogeneous geometries. However, analytical algorithms are still preferred in clinic today for their fast calculation.

The sensitivity of the treatment against uncertainties can be improved with several strategies. A common technique consists in increasing the volume to be irradiated, and regions that must be spared, with safety margins. However, due to the proton range uncertainty, margin solution are not adequate for proton therapy [12]. Another solution consists in taking the uncertainties into account during the plan optimization, leading to the so-called robust plan optimization [13, 14]. A treatment plan is called robust when the resulting dose distribution remains within the clinical goals, even in presence of uncertainties.

While these planning strategies definitely help to increase the robustness of treatment plans, they are not perfect. Indeed, commercial tools used in clinic are often based on approximate dose calculation algorithms and only integrate limited models of treatment uncertainties. There is therefore no guarantee that the resulting treatment plan will actually be delivered as predicted by the TPS. The evaluation of the treatment plan robustness before delivery is therefore highly desirable. Unfortunately, the evaluation tools provided in the TPS are usually based on the same approximate algorithms and limited models. There is a need for independent tools to evaluate the treatment plan in a realistic and comprehensive way.

The main objectives of this thesis are to reduce and to control the impact of proton therapy treatment uncertainties. MCsquare [15], a fast Monte Carlo algorithm has been developed [16] to improve the accuracy of dose calculation, reducing the

range uncertainties. Its optimized implementation exploits all resources provided by modern processors and enables computation times compatible with clinical constraints. Improved robust planning tools are also proposed to control other treatment uncertainties. In particular, an adaptation of the safety margin concept typically used for conventional radiotherapy with photons was implemented to take into account proton range uncertainties. The Monte Carlo dose calculation was also adapted to feed a robust optimizer. Finally, an independent and accurate treatment plan evaluation tool has been developed based on the Monte Carlo code. Advanced uncertainty models have been implemented in MCsquare to provide a realistic and comprehensive verification of the robustness of proton therapy treatment plans. All developed tools are released open source [15, 17, 18].

While the implications and applications of this thesis are not limited to the treatment of mobile tumors, the methods and results presented are particularly focused on liver and lung tumor treatments. This is not an arbitrary choice. Indeed, for the physicist point of view, the breathing motion and tissue heterogeneity that characterize these particular anatomy sites, are part of the most challenging issues to be solved in proton therapy.

1.2 Structure of the thesis

This thesis is constructed around two main papers and is organized as following:

Chapter 1: Introduction

This is the present chapter. It aims at introducing the concepts of proton therapy that are necessary to read the next chapters. The motivations and objectives of the thesis are also presented. The main contributions of the author to address the current gaps in the research field are then described.

Chapter 2: Fast Monte Carlo simulations

The second chapter includes the first main publication entitled "Fast multipurpose Monte Carlo simulation for proton therapy using multi- and many-core CPU architectures", which is published in the Medical Physics journal. It describes the development of MCsquare, a fast Monte Carlo algorithm for proton therapy dose calculation. The use of such accurate dose calculation method in the clinic would enable to considerably reduce the treatment uncertainties.

Chapter 3: Robust treatment planning

In this chapter treatment planning methods dedicated to proton therapy are pre-

sented. Especially, a tool to evaluate the tumor motion was developed to help selecting the appropriate treatment modality. Another tool was also developed to guide the planner in the selection of optimal beam angles. Two methods are then presented to account for treatment uncertainties in proton therapy treatment planning.

Chapter 4: Realistic verification of the treatment plan robustness

This chapter is dedicated to the second main paper, which is entitled "Monte Carlo methods to realistically evaluate the robustness of proton therapy plans". This paper is currently under review in the Medical Physics journal. It presents a novel method to verify the robustness of proton therapy plans in a realistic way. In particular, motion effects are simulated using models of variable breathing amplitude and dynamic delivery.

Chapter 5: Complement for treatment plan evaluation

Additional plan evaluation topics are further discussed to complement the robustness verification method presented in Chapter 4. Especially, a model of the baseline shift is investigated. The simulation of LET distribution and its application in the evaluation of the treatment plan quality is then discussed.

Chapter 6: Discussion and conclusion

This final chapter provides a discussion of the implications of this thesis and exposes some of the remaining issues.

1.3 General overview of radiation therapy

This section provides an introduction of general concepts in radiation therapy. The reading of this section should answer the following questions. What is cancer and what are the typical treatments? Why radiation therapy is effective? What are the various radiation therapy modalities? For expert readers that already know the answers to these questions, the reading of this part might be optional and they may prefer to directly continue with the next section, which is more focus on proton therapy and its current limitations.

1.3.1 Cancer

Cancer is responsible for more than 22% of deaths worldwide, making it the second cause of mortality, after heart diseases [19]. In developed countries, about

one third of the male population and one fourth of the female population will be affected by cancer before the age of 75 years [20]. Cancer treatments are continuously being improved, thanks to research. However, in 2012, half of the patients suffering from cancer were still dying from their disease[20].

Cancer is a family of diseases involving uncontrolled cell divisions. Normally, human cells are renewed at a moderate rate by controlling their division and death with a complex system of biological signals. However, mutations of the DNA may occur from time to time due to environmental factors, viruses, or errors during the DNA replication [21]. These mutations may have no effect, or may lead to abnormal functioning of the cell. In particular, the mutation can affect the mechanism regulating the cell division and initiates a cancerous process. The cancerous cell can divide without stopping, multiplying itself and forming a cluster of cancerous cells called tumor. Some cancerous cells may spread and travel in the body, with the potential to initiate distant secondary tumors called metastasis.

1.3.2 Treatment modalities

Depending on the type of cancer and its stage, multiple treatment modalities are often combined to try to cure the disease. Nowadays, the three main treatment modalities are surgery, radiation therapy, and systemic therapy. Surgery can be performed to remove the tumor when possible. If the tumor is not removable, or when the patient could not tolerate a surgery, radiation therapy is an alternative. Radiation therapy is also often considered after surgery to eliminate cancerous cells that may have spread in surrounding tissues. These treatments can also be combined with systemic therapy to control distant metastasis and kill cancerous cells traveling in the body.

Half of the cancer treatment involve radiation therapy [1]. This treatment modality aims at eradicating cancerous cells by irradiation. The delivery of radiation therapy can be subdivided in multiple categories. For internal radiotherapy, the source of radiation is placed inside the patient by implanting radioactive seeds directly in contact with the tumor (brachytherapy), or by injecting radiopharmaceuticals that migrate to the tumor (nuclear medicine therapy). For external radiotherapy the target is irradiated using a beam of radiations generated outside the patient. Most of external radiotherapy treatments employ high energy photon beams. However, other particles can be used, offering different distributions of radiation damages.

The present thesis is dedicated to external radiation therapy with proton beams.

The choice of this specific delivery modality is motivated in the next subsection.

1.3.3 Rationale of radiotherapy

Depending on the nature of the radiation, various physical interactions with the patient tissues are possible [22, 23]. At the energy typically employed in radiotherapy, charged particles, such as protons, mainly interact with the patient tissues through ionization process. In contrast, neutral particles, like photons and neutrons, interact by generating secondary charged particles, which can then deposit their energy by ionization. In medical physics, the interaction of radiations with matter is usually quantified by the absorbed dose. This quantity is expressed in unit of Grays (Gy) and is defined by the energy (in Joules) deposited per unit of mass (in kilogram).

The DNA, which encodes the genetic information, is the part of the cell that is the most sensitive to the effect of radiations. When the DNA is damaged, reparation mechanisms try to reconstruct the missing information. However, if the damage is too important, with double-strand breaks for instance, the reparation may fail, causing mutations of the DNA and/or the death of the cell [21]. The challenge of radiation therapy is therefore to sufficiently irradiate cancerous cells to cause lethal damage, while sparing normal cells. Fortunately, the reparation mechanisms of healthy cells are more efficient than those of cancerous cells [22]. The tumor control probability (TCP) [24] is defined as the probability that a given dose, deposited in the tumor volume, will kill all cancerous cell. Similarly, the normal tissue complication probability (NTCP) is defined as the probability that the given dose will lead to complications due to the toxicity of radiations. As illustrated in Figure 6.1, these probabilities are typically characterized by a sigmoid function. In order to optimize the outcome of the patient treatment, the delivered dose should therefore maximize the TCP, while keeping the NTCP as low as possible. However, due to the proximity of TCP and NTCP curves, this is not easily achievable in practice.

Clinicians can influence the TCP and NTCP curves by working on various parameters. For instance, clinical studies have shown that the normal tissue complication rates can be reduced by fractionating the delivery in multiple irradiation sessions [22, 25]. This provides more time for the healthy cells to repair their DNA damages. Radiotherapy is therefore delivered in multiple fractions. The treatment is stretched over several weeks, with typically 5 delivery sessions per week.

TCP and NTCP are also dependent on the volume of the target or the volume

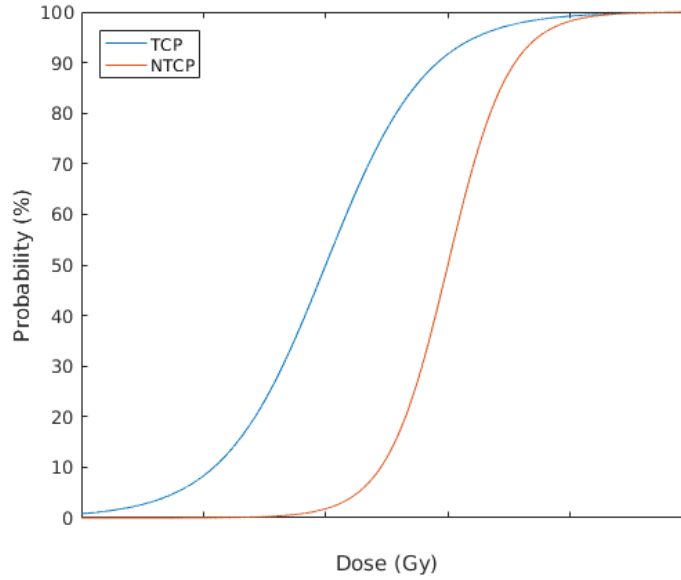


Figure 1.1: The tumor control probability (in blue) and normal tissue complication probability (in red) are represented by sigmoid functions. The optimal dose should maximize the TCP and minimize the NTCP.

of the organ at risk (OAR) that receive a given level of dose. The goal is therefore to deliver an homogeneous dose on the target volume, while minimizing, as much as possible, the dose deposited in OARs. Radiotherapy techniques have been improved over the years to deliver a dose distribution very conformal to the target volume, limiting the exposition of surrounding healthy tissues. Target coverage and OAR sparing are commonly analyzed in clinical practice using dose-volume histograms (DVH) [26]. As illustrated in Figure 1.2, these plots are actually cumulative histograms that enable an easy evaluation of the dose deposited in each volume of interest. The example illustrated in Figure 2 should be read as “the dose received by 95% of the target volume is at least 57.6 Gy”. This is usually shortened as the D95 is 57.6 Gy. Another way to read a DVH value is “the volume receiving at least 5 Gy is 20.6%”, which is shortened as the V5 is 20.6%. DVH metrics are highly correlated with the patient treatment outcome and are usually employed to express the dose prescription for a radiotherapy treatment.

The nature of the radiation plays an important role in the improvement of dose conformality to the target volume. Due to their difference in physical interactions with matter, different radiation types provide very different dose distributions. The

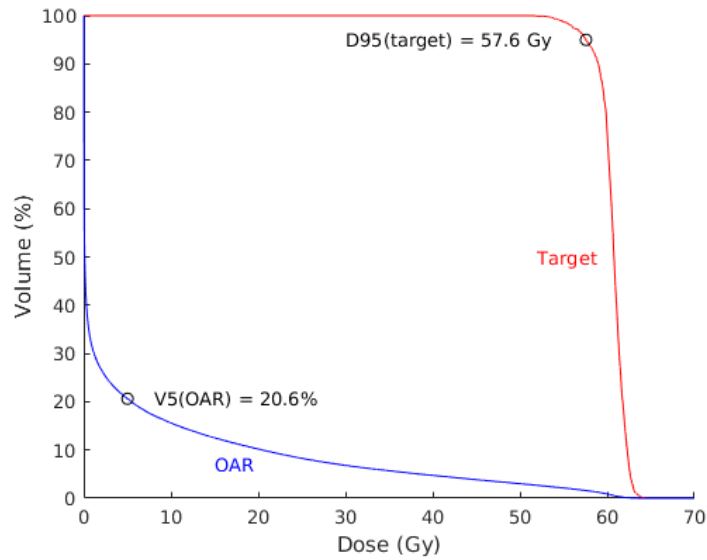


Figure 1.2: Example of dose-volume histogram (DVH) for the target (in red) and an organ at risk (OAR, in blue). The D95 and V5 values are marked to illustrate how to read a DVH.

doses deposited in water as a function of depth are compared in Figure 1.3 for photon, electron, proton, and carbon beams. One of the major differences is that charged particles are slowed down by interacting with matter and eventually stop at a certain depth [23]. In contrast, photons are not losing their energy continuously. Each photon interact at a random depth, exiting the primary beam. The intensity of photon beams is therefore attenuated exponentially with depth [22]. However, the beam is not stopped at a given depth as with charged particles and it continues to irradiate healthy tissues behind the tumor. Moreover, due to the exponentially decreasing depth-dose profile of photon beams, the maximum of dose is deposited at shallow depth, which is not optimal to treat deep tumors.

Nowadays, most of radiotherapy treatments are delivered with X-ray beams. Despite the inferiority of photon ballistics, it is currently the most mature radiotherapy technology. However, the use of heavy charged particle beams is emerging in clinic, thanks to the recent technological progress and the decrease of equipment costs [4, 5]. It offers several advantages compared to the conventional radiotherapy [28, 3]. First, as explained previously, charged particles stop at a given depth. This finite range enables the sparing of healthy tissues downstream the target. Moreover, most of the dose is deposited at the end of the particle path, forming

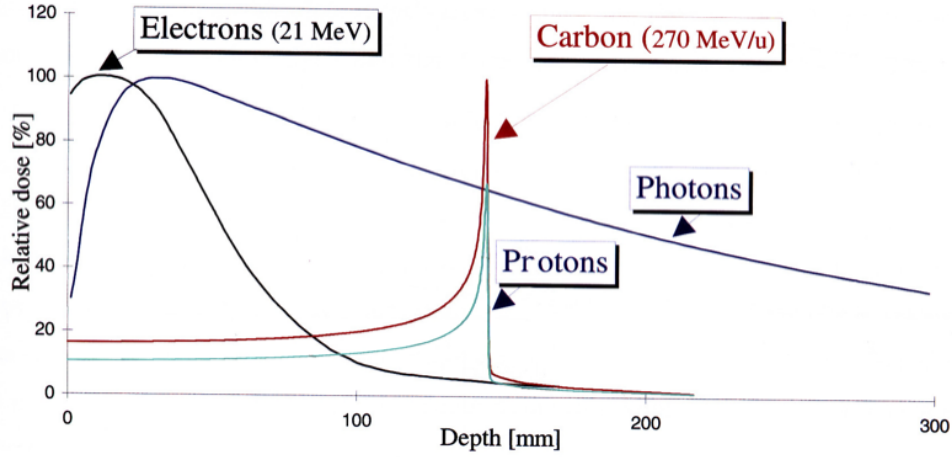


Figure 1.3: Comparison of the depth-dose profiles for electron, photon, proton, and carbon beams. The profile characterizing heavy charged particles starts with a low dose plateau at the entrance and end with a peak of deposited dose called Bragg peak. Source: [27]

the so called Bragg peak. Therefore, the low dose plateau preceding the peak permits to reduce the dose absorbed by healthy tissues upstream the target.

Besides the ballistic differences, some radiation types are biologically more efficient than others. Indeed, for a same delivered physical dose, the importance of biological damages differs from one radiation type to another. This effect can be quantified by the relative biological effectiveness (RBE) [29]. It is defined, for a reference radiation (typically photons) and another radiation type, as the ratio of doses that induce the same biological effect:

$$RBE = \frac{D_{\gamma}}{D_{\text{Ion}}} \Big|_{\text{Isoeffect}}$$

This leads to the definition of the biological dose, which is calculated by the multiplication of the physical dose and the RBE. The average RBE for proton is considered to be 1.1 [30], meaning that, on average, 10% additional dose is required with photons to achieve the same biological effect as with protons. The effect is even increased with heavier charged particles, such as carbon ions [31]. This biological effect can be explained by the distribution of the energy deposition along the particle path, which is very local for heavy charged particles and more dispersed for photons. The linear energy transfer (LET) is defined as the mean

energy lost by the particles per unit of distance. High LET radiations, such as protons and carbons, generate complex damages to the DNA, which are more difficult to repair. The LET is generally not constant with the depth and increases in the Bragg peak, which is normally located inside the target. This enables to raise the RBE inside the target, especially for very heavy particles such as in carbon therapy.

Heavy charged particles offer therefore physical and biological advantages over the conventional radiotherapy with photons. However, particle therapy is still a recent modality, with many challenges to be solved. Especially, there is an uncertainty on the actual proton range inside the patient. For these reasons this thesis is dedicated to proton therapy. Carbon therapy may offers even more theoretical advantages, but it is not as widespread as proton due to its high costs. Moreover, the solutions developed in this thesis to address some of the proton therapy issues can be applied almost directly to heavier charged particles.

1.3.4 Delivery of proton therapy

One of the reasons why proton therapy is not yet used in all radiotherapy centers is the cost of the equipment and of its maintenance [4, 32]. To achieve the delivery accuracy required for clinical applications, the proton beam is generated and adjusted through multiple steps [23] illustrated in Figure 1.4.

The first step consists in ionizing hydrogen atoms in order to extract the protons. The kinetic energy of protons is then increased using a particle accelerator (usually a cyclotron, less frequently a synchrotron or a linear accelerator). The depth of the Bragg peak can be adjusted by varying the proton energy. However, the proton energy at the output of the accelerator is often fixed. For IBA cyclotrons, protons are accelerated to around 230 MeV, corresponding to a proton range of 33 cm in water. The energy selection system is therefore responsible to degrade the proton energy at the exit of the accelerator in order to reach the desired range inside the patient. The size of the Bragg peak is generally too small to cover the entire thickness of the target. Multiple Bragg peaks of different energies are therefore combined to provide a larger coverage, forming a spread out Bragg peak (SOBP) as shown in Figure 1.5.

After the energy selection system, the proton beam is transported to the treatment rooms. The beam transport system is composed of several magnets to keep the beam focused inside the beamline. The beam eventually arrives in the treatment room and is directed to the patient through a fixed beamline or a rotating gantry. At the exit of the beamline, the final beam preparation is performed in

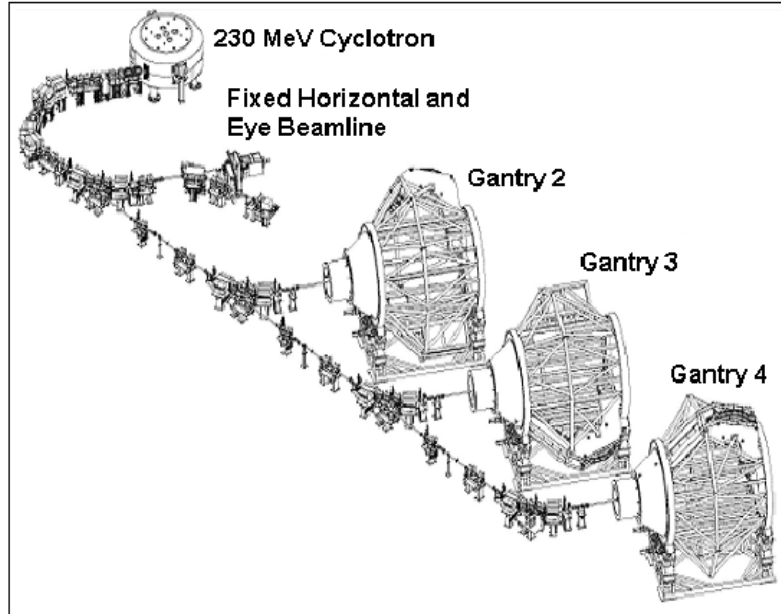


Figure 1.4: Example of proton therapy equipment composed of a cyclotron, a transport line, one fixed beamline and three gantries. Source: Essen Proton therapy center [33]

the nozzle, following the treatment plan specifications.

Two main delivery modes are employed in clinic [23, 34], namely the double scattering (DS) and the pencil beam scanning (PBS). As represented in Figure 1.6, the DS mode provides a broad beam shaped to cover the target in one shoot. The beam exiting the cyclotron first pass through the modulation wheel in the energy selection system. This rotating wheel offers multiple thickness of materials to degrade the mono-energetic beam to the various energies necessary to generate the SOBP. The pencil beam is then scattered two times to provide a flat broad beam. At the exit of the nozzle, patient specific collimators and compensators are employed to provide the final shape to the beam. The metallic collimator stops the peripheral part of the beam and determine the sectional shape of the beam. The range compensator is a plastic block milled with various depth to further degrade the beam energy locally. This way, the proton ranges are adjusted to the distal shape of the target.

Double scattering was the most common delivery mode in proton therapy. However, this technique leads to several issues. First, because the size of the SOBP is

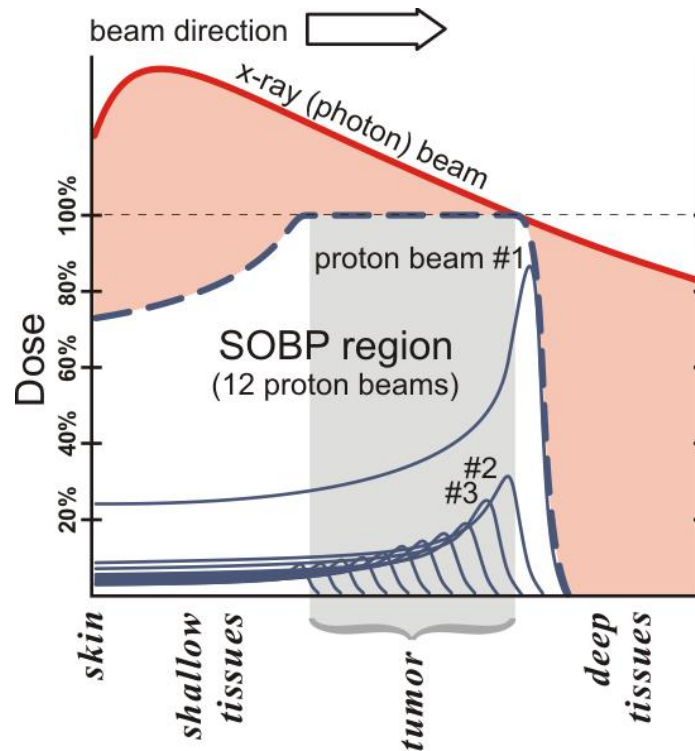


Figure 1.5: The SOBP (dashed blue curve) is formed by combining multiple Bragg peaks of different energies (solid blue curves). Source: [5]

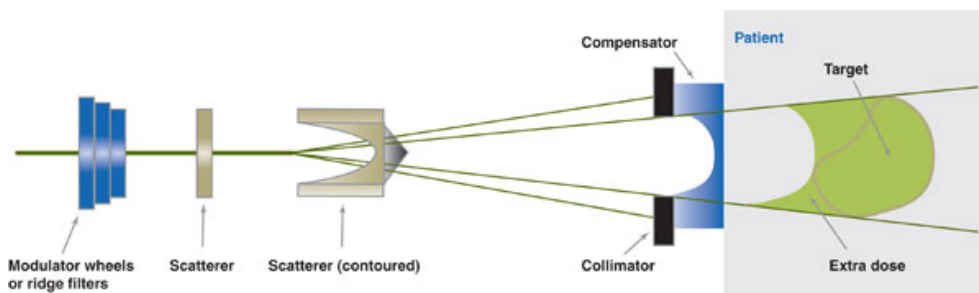


Figure 1.6: Representation of a proton therapy beamline with double scattering delivery. Source: [35]

fixed for the entire beam, the delivered dose distribution cannot be conformal to the proximal part of the target, giving unnecessary dose to the patient. Moreover, due to the proton interactions in the collimator and compensator, a neutron radiation is generated close to the patient. Finally, the cost, and manipulation time

of these devices is also problematic. Indeed, the collimators and compensators have to be fabricated and verified individually for each patient. It also requires an additional time during the treatment to manually set up the devices before the delivery of each beam. Moreover, these materials become activated with the repeated irradiation and must be stored properly before being recycled.

Due to the many issues encountered with DS delivery, it tends to be replaced by the PBS delivery mode. As suggested by the name, a thin beam is scanned, painting the dose sequentially spot by spot, and layer by layer over the target volume [36]. This provides a better flexibility, enabling the delivery of complex heterogeneous dose distributions, also called intensity modulated proton therapy (IPMT). As represented in Figure 1.7, scanning magnets are mounted inside the nozzle to bend the pencil proton beam. The magnetic field is adjusted between each spot in order to reach a specific location inside the patient. Once all spots of the layer are delivered, the energy is changed in order to "paint" the next layer at a different depth. The absence of collimators and range compensators highly reduces the neutron production and the manual operations. However, the delivery itself is generally slower due to the time required to switch the proton energy between each layer (typically one to two seconds).

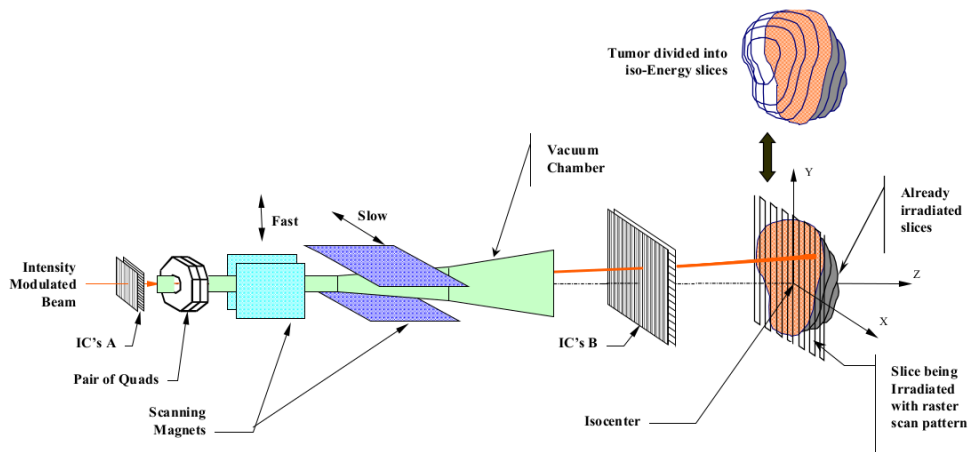


Figure 1.7: Representation of a proton therapy beamline with pencil beam scanning delivery. Source: [36]

1.4 Proton therapy workflow and uncertainties

In this section, the key steps of a typical proton therapy treatment are presented. The potential treatment uncertainties are also described for each step of the clinical workflow. The control of treatment uncertainties is very crucial because an error of a few percent in the delivered dose may lead to much larger variations of the TCP and NTCP.

Treatment uncertainties are generally subdivided into two categories. Systematic errors impact the dose delivery of all fractions and reduce the accuracy of the treatment. Random errors vary from one day to the other and reduce the precision of the treatment.

The impact of these treatment uncertainties on the dose distribution can also be divided in two groups. First, uncertainties may induce a rigid shift of the dose distribution with respect to the target. This is the main effect in conventional radiotherapy with photons. Safety margins employed in conventional radiotherapy are especially designed to cover such geometric misalignment of the dose with the target. The second possible effect of treatment uncertainties is a deformation of the dose distribution. In proton therapy, both effects are generally combined, due to the variation of the proton range. This dose distribution distortion is one of the major issue in proton therapy and represents the main topic of this thesis.

1.4.1 Pre-treatment imaging

The first step of the workflow consists in acquiring images of the patient anatomy. Typically, a 3D computed tomography (CT) image is acquired. This imaging modality provides a 3D map of the X-ray attenuation and is expressed in Hounsfield Units (HU). The CT scan generally provides a good quality representation of the patient anatomy at the moment of the acquisition. As the treatment planning is based on this image, the patient is imaged in the same conditions as for the treatment. An immobilization system is often employed to reduce the patient motions during imaging and future treatment delivery sessions and to increase the reproducibility of the patient positioning.

Due to organ motions, such as breathing or heart beats, the treated area is not always static. The use of 4D imaging provides information about the tumor and organs motions. Assuming that the motion is periodic, a 4DCT image can be reconstructed [37], providing multiple 3D CT images at different phases of the breathing motion. These images can be used during the treatment planning to

estimate the position of the tumor and organs along the motion cycle. Abdominal compression belt, or breath hold may be used to limit the organ motion when the patient can tolerate it.

Due to the limited soft tissue contrast of the CT scan, it may be not easy to distinguish the tumor when its density is close to the surrounding tissues. Other imaging modalities may be used in complement to the CT scan [38]. For instance, magnetic resonance imaging (MRI) provides anatomical images with better soft tissue contrast. Liver tumors are typically barely visible on CT images. A MRI image is often required for the contouring of these tumors. Another imaging modality also frequently employed for radiotherapy is the positron emission tomography (PET). This functional imaging technique does not provide anatomical information. However, it can provide physiological information such as the glucose uptake, which is usually abnormally high in tumor tissues. PET imaging is therefore useful to detect and follow the progression of tumors and metastasis.

This pre-treatment imaging process is crucial because the treatment preparation and clinical decisions are mainly based on this information. Therefore, the accuracy of the treatment is very dependent of the imaging quality. However, imaging devices feature limited spatial and temporal resolutions. Moreover, the image is often deteriorated by noise and artifacts [39]. These limitations will inevitably lead to uncertainties during the contouring and dose calculation steps. Especially, the estimation of the proton range inside the patient is based on physical properties extracted from the CT image [9].

Another major limitation of the current clinical practice is that the treatment planning is based on a pre-treatment image, which is a snapshot of the patient anatomy at the moment of the acquisition. However, the patient anatomy may change before and during the treatment [40, 41, 42]. Therefore, the dose actually delivered to the tumor and organs may not correspond to the expected one [7].

1.4.2 Contouring and margins

After imaging, physicians define the contours of organs at risk (OAR) and tumor volumes. The contours of organs and visible tumors are drawn over the CT image, referred as the planning CT. When the tumor is not identifiable on the CT image, it is contoured on another image (MRI or PET). This second image is then aligned with the CT image using registration tools. This enables the transfer of the contour to the planning CT.

The tumor volume contoured based on its visible extension on the images is called gross tumor volume (GTV). However, cancerous cells may have spread in the neighboring tissues. In order to irradiate all cancerous cells, the GTV is generally extended with an isotropic margin of a few millimeters, depending of the tumor type and location. The physician may also manually crop or increase the new contour according to the patient anatomy and his suspicion of cancer extension. Clinical target volumes (CTVs) include the expended GTV and other suspicious regions such as contaminated nodes.

The objective of the next steps of the treatment planning is to ensure a good coverage of the CTV. Indeed, if the treatment fails to kill all cancerous cells, the cancer may recur. However, the CTV coverage is not guaranteed due to all treatment uncertainties. The treatment is considered robust to a given level of uncertainty when the dose delivered in presence of such uncertainties still satisfies the clinical goals. The most common method to increase the treatment robustness against uncertainties is to expand the CTV with an additional safety margin. The expanded volume is called the planning target volume (PTV). Recipes have been proposed to compute the PTV margin according to the estimation of uncertainty values [43].

Motion effects can also be taken into account with additional margins. In proton therapy, the internal target volume (ITV) [44] is often employed. The ITV is the envelope of the CTV drawn on each phase of the 4DCT. By irradiating the ITV, the CTV is covered all along its motion. As shown in Figure 1.8, the PTV margin is then applied to the ITV instead of the CTV.

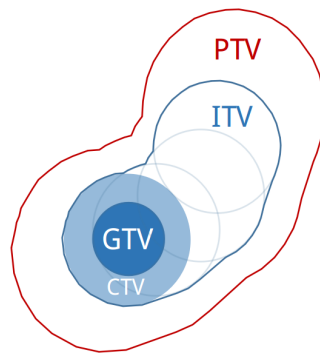


Figure 1.8: Illustration of the various level of target volumes.

The assumptions required to apply the PTV recipes usually employed in radiotherapy with photons are not always met with proton therapy due to the range uncertainties. This is especially true in presence of motion. Therefore the PTV formalism should be adapted for proton therapy. However, due to the lack of adapted tools in commercial software, many proton centers are still using the original PTV concept.

The contouring of GTV and CTV volumes is subject to the personal judgment of physicians. This variability [45] is also a treatment uncertainty and should be taken into account in the PTV margin or other robustness methods. Moreover, the contours are based on pre-treatment images and may not correspond to the actual patient anatomy during the treatment delivery.

1.4.3 Dose prescription

The dose to be delivered per fraction on the tumor and the number of treatment fractions are prescribed by the physician. Dose limits are also specified for organs at risk. These values are typically based on clinical experience and clinical trials. Dose prescriptions are usually reported in dose to water, which is the dose that would be deposited in patient tissues equivalent to water.

1.4.4 Dose calculation

In order to optimize the delivery parameters, we need a way to simulate the delivery and calculate the dose distribution in the patient geometry. This process starts with the conversion of the image units (HU for CT images) in physical quantities practical for the proton therapy dose calculation [9]. The most important quantity required for proton dose calculation is the mean energy deposited per unit of distance, called stopping power [23]. The Bethe-Bloch formula [46, 47, 48] is generally used to calculate the proton stopping power $-dE/dx$, which depends on the proton energy and the properties of the material crossed:

$$-\frac{dE}{dx} = K\rho \frac{Z}{A} \frac{z^2}{\beta^2} \left[\frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 T_{max}}{I^2} - \beta^2 - \frac{C}{Z} \right]$$

where:

$K = 4\pi N_A r_e^2 m_e c^2$ is a constant,

N_A is the Avogadro number,

m_e, r_e are the mass and classical radius of the electron respectively,

ρ is the mass density of the material,
 Z is the effective atomic number of the material,
 A is the effective atomic mass of the material,
 z is the charge of the particle (1 for protons),
 β, γ are the relativistic parameters of the particle,
 I is the mean excitation potential of the material,
 C is the shell correction for low energy effects,
 T_{max} is the maximum kinetic energy that can be transferred to an electron:,

$$T_{max} = \frac{2m_e c^2 \beta^2 \gamma^2}{1 + 2\gamma \frac{m_e}{M} + \left(\frac{m_e}{M}\right)^2}$$

with M the mass of the particle (938.3 MeV/ c^2 for protons).

The measurement or calculation of I values for human tissues is not easy. It is considered to be a significant source of uncertainty and can induce a 1.5% error on the proton range [8].

In proton therapy, the stopping power ratio (SPR) between the tissue and water materials is often employed:

$$SPR = \frac{dE}{dx} \Big|_{\text{tissue}} \Big/ \frac{dE}{dx} \Big|_{\text{water}}$$

As the composition of most human tissues is close to water and the effective atomic number is rather small, the SPR is often considered as a constant, neglecting its energy dependence. This approximation is valid for most human tissues, but may introduce range errors when certain materials, such as bone, are crossed.

The image HU are therefore converted in ρ , Z , A , or directly in SPR using a calibration curve specific to each CT scanner. The stoichiometric method proposed by Schneider[9, 49] is generally employed in proton therapy to generate such calibration curves. The HU conversion also induces large uncertainties on the calculation of proton range. Indeed, for one HU value, multiple combinations of ρ , Z , A exist. Dual energy CT have been recently introduced to decrease this ambiguity [50] by performing two, or more, CT acquisitions with different X-ray energy spectrum.

Based on the stopping powers extracted from the patient CT image and the machine parameters, the dose calculation algorithm is able to estimate a 3D map of the dose distribution inside the patient. The accuracy of the resulting dose

distribution depends on how close the implemented models are able to reproduce the real proton physics. In practice, there is usually a trade-off between accuracy and computation time. Two main categories of algorithms are generally employed for proton therapy, namely analytical and Monte Carlo algorithms.

Analytical algorithms are based on macroscopic models, which rely more on image processing techniques than actual modeling of the physics. For instance, the pencil beam algorithm (PBA) [10] uses multiple pencil beam dose kernels, measured or pre-computed in water, are stored in a database. The contribution of multiple kernels are summed to produce the full dose distribution. The position and intensity of each kernel are calculated according to the machine parameters. The proton range is, however, adapted by scaling each kernel in depth according to the SPR map. Some corrections may be applied to improve the scattering effect in the lateral directions. Analytical algorithms are typically used in clinical routine for their high computation speed. However, the accuracy of the dose calculation is limited [8, 51], especially in heterogeneous geometries as illustrated in Figure 1.9.

In contrast, Monte Carlo (MC) algorithms are based on microscopic simulations of proton interactions in the patient geometry [52, 11]. In general, MC algorithms enable to solve complex problems using random number generators. MC is therefore well suited to the simulation of proton interactions due to the stochastic nature of particle physics. Basically, protons are transported individually through the patient geometry. Instead of using SPR, like in analytical algorithms, physical interactions are randomly sampled according to the material stopping powers and interaction cross sections. The algorithm will be presented in detail in the second chapter of this thesis. As the simulation is based on random sampling, the dose results are subjected to a statistical uncertainty. This stochastic noise is reduced by simulating the interactions of a large number of protons. MC algorithms are very accurate, whatever the complexity of the geometry thanks to a better simulation of scattering effects. However, the MC dose calculation is much slower than analytical algorithms. For these reasons, MC simulations are generally performed for research and development purposes. However, multiple fast MC algorithms have been recently developed [16, 53, 54], enabling more applications for the clinical routine.

1.4.5 Treatment plan optimization

The treatment plan is composed of the list of machine parameters defining how the treatment must be delivered. For instance, in PBS delivery, the treatment plan contains the list of beams (each characterized by a gantry and a couch angle) and

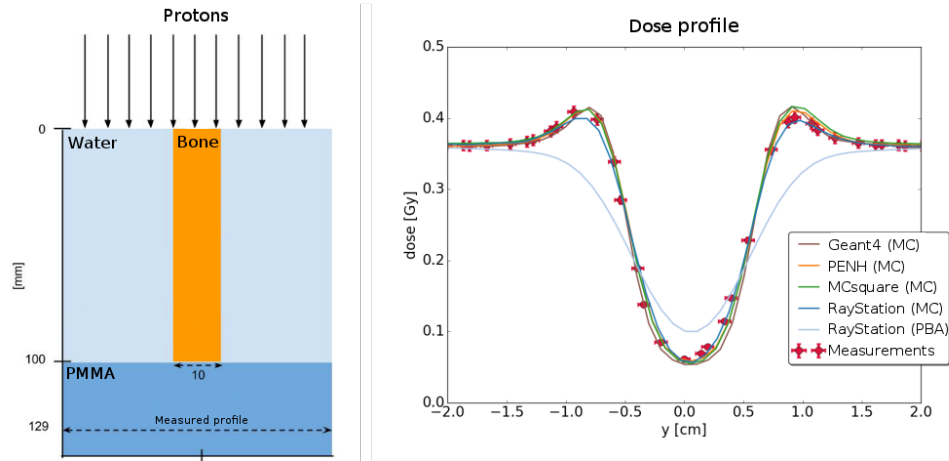


Figure 1.9: Comparison of analytical pencil beam algorithm (PBA) and Monte Carlo dose calculation results in a heterogeneous phantom. The dose profiles are compared at 129 mm depth and averaged over 4 mm for Geant4.9.5 (brown line), PENH (orange line), MCsquare (green line), RayStation MC (blue line), RayStation PBA (shadowed blue line) and experimental data (red circles points). This figure was adapted from the publication of Sorriaux et al. [51]

the list of spots for each beam. Each spot is defined by a position, an energy, and an intensity. All these parameters must be optimized so that the resulting delivered dose is in agreement with the physician prescriptions.

The treatment planning system (TPS) is the software employed to generate the treatment plan. The CT image and contoured volumes are first imported in the TPS. The gantry and couch angles are typically selected by the planner according to his clinical experience. The TPS computes then the spot positions to cover the target volume. The spot energies are calculated to reach the desired depth using the SPR extracted from the CT image. The dose engine, typically an analytical algorithm, is then employed to compute the contribution of each spot to the total dose distribution. The spot intensities are adjusted by an iterative optimizer [55] in order to deliver the prescribed dose on the target while keeping the dose to OAR within the clinical limits. Sometimes, there is no optimal solution satisfying both target coverage and OAR sparing. For such cases, the physician decides which objectives can be compromised.

As described in a previous sub-section, the PTV margin is usually employed to increase the robustness of the treatment plan against uncertainties. The treatment

plan is therefore optimized to cover the PTV volume with the prescribed dose. Safety margins are, however, not the ideal solution to take into account the proton range variation. Another method to take into account treatment uncertainties is called robust optimization [14, 56, 57, 58]. In this case, both robustness and plan optimization are performed in one step. The treatment plan is optimized to cover the CTV for multiple uncertainty scenarios. This procedure is more complex than margin solutions, but it enables to take into account proton range uncertainties properly. The dose distribution is indeed computed and optimized for multiple scenarios simulating various possible realizations of treatment uncertainties. Typically, systematic patient positioning errors and flat range uncertainties are modeled in commercial TPS. In some TPS, organ motion can also be taken into account by considering multiple 4DCT phases during the robust optimization process, instead of a single 3D CT image. The number of uncertainty scenarios are therefore multiplied by the number of 4DCT phases included in the optimization.

While robust optimization is recognized to provide better robustness against proton range uncertainties [12], the PTV solution is still often employed in clinical practice. This may be explained by an inertia of the clinical practice, but also because robust optimization requires more computational resources and is more complex to evaluate. Indeed, there is no PTV volume on which the resulting dose distribution can be reported and analyzed. The evaluation of treatment plans generated with a robust optimizer, requires to explicitly verify the impact of treatment uncertainties on the dose distribution. For this purpose, treatment uncertainties need to be simulated in a comprehensive and realistic way. This robustness evaluation process is described in the next sub-section.

1.4.6 Treatment plan verification and QA

Due to the limited accuracy of planning tools and other treatment uncertainties, it is not guaranteed that the optimized treatment plan will lead to the expected dose distribution when delivered to the patient. Whatever the planning strategy employed to take into account treatment uncertainties, the quality of the resulting plan should therefore be evaluated. Two conditions must be satisfied. First, the treatment plan has to meet the clinical constraints in term of tumor dose coverage and organ dose limits. Second, dose distribution computed by the TPS must be accurate and deliverable by the machine.

To verify the first condition, the DVH is often employed. Metrics of interest have to be within the clinical constraints. For instance, the dose delivered to OAR is analyzed with several metrics, such as the mean dose (D_{mean}), maximum dose

(Dmax), V20, etc ... The recommended metrics and their associated limits generally vary from one organ to another. When the treatment planning is performed with a PTV, the DVH metrics can be directly checked on the PTV. Typically, 95% of the PTV should receive at least 95% of the prescribed dose. The PTV margin is supposed to ensure the coverage of the CTV in presence of treatment uncertainties. However, when robust optimization is employed, there is no clear volume that takes into account treatment uncertainties. A simple evaluation of the dose distribution in the CTV would not reflect the actual delivered dose in presence of uncertainties. In order to properly evaluate the CTV coverage and OAR sparing, treatment uncertainties must be simulated explicitly. The dose distribution is calculated for multiple simulations of treatment scenarios, each representing a possible treatment realization that includes uncertainties. This concept is similar to the robust optimization, but the purpose is now to evaluate the treatment plan. To do so, the DVH is computed for each scenario. As illustrated in Figure 1.10, the envelope of all scenario DVHs is typically reported as a band [59], which represents the possible variations of the delivered dose. The width of the band is a good indicator of the plan sensitivity to uncertainties.

As the PTV margin does not fully compensate for proton range variations, such comprehensive robustness verification should be performed in addition to the simple PTV coverage verification. In practice, this is, however, infrequently performed due to the extensive computation resources necessary. Moreover, the tools provided by commercial applications are limited to simple models of patient positioning (setup) and range errors, which may not provide a realistic and comprehensive simulation of treatment errors. The lack of fast and comprehensive tools for plan verification is one of the reasons why the margin solution is still preferred to robust optimization in many proton centers.

To verify the accuracy of the TPS dose engine, an independent dose calculation software is desirable. The dose distribution is therefore recomputed and compared with the TPS result. Monte Carlo algorithms are particularly recommended for this task [60, 61] thanks to their improved accuracy. Fast Monte Carlo algorithms were recently released with some TPS. However, it is not yet widely employed in the clinical practice.

The quality assurance (QA) is an additional verification step that is generally performed for each treatment plan. This process aims at verifying the entire treatment workflow. It consists in delivering the treatment plan in water, or into a plastic cube, called phantom. The dose is measured [62, 63] at a certain depth using an array of ionization chambers, or Gafchromic films. The measured 2D

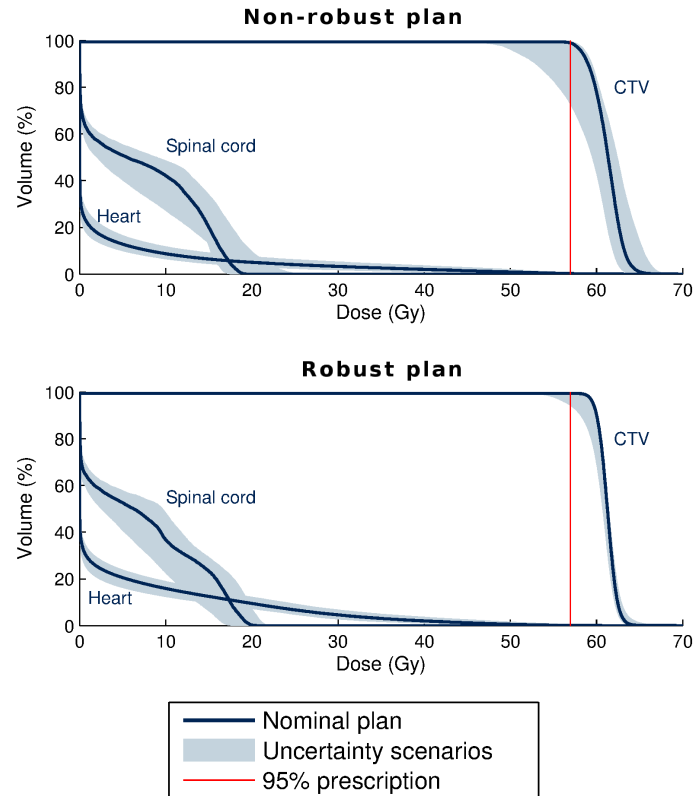


Figure 1.10: Example of DVH-bands for a non-robust plan (up) and a robust plan (down). The light blue bands represent the possible dose variation in presence of treatment uncertainties. The dark blue curves represent the nominal dose distribution as estimated by the TPS, without uncertainties. Typically, 95% of the target volume should receive at least 95% of the prescribed dose (red line), even in presence of treatment uncertainties.

dose distribution is then compared to the dose predicted by the TPS in same conditions (phantom geometry, measurements depth, ...). This measurement enables a verification of the treatment workflow, including the TPS dose calculation, the transfer of the plan to the delivery machine, the delivery accuracy. Machine quality control (QC) is also generally performed every morning before treating patients to verify the stability and accuracy of the delivery equipment.

1.4.7 Treatment delivery

For each delivery fraction, the patient needs to be set up on the treatment couch in the same position as during the pre-treatment imaging, which was employed to establish the treatment plan. Historically, markers on the patient skin or on the immobilization system were aligned with lasers. Nowadays, in-room imaging is often performed additionally to further guide the patient setup [64, 65]. Images can be acquired with two orthogonal radiographs, a cone-beam CT (CBCT), and/or in-room CT. The new images are aligned with the planning CT using rigid registration algorithms. The treatment couch is then translated and rotated using the registration information to align the patient correctly. This operation may be repeated multiple times to find the patient setup that best match the planning conditions. Daily imaging also enables to check when the patient anatomy has changed too much, indicating the need to adapt the treatment plan to the new anatomy.

Besides image guidance, residual setup errors are possible [66]. This may be caused by the imaging limits (noise, resolution, contrast), small errors in the alignment of the imaging device with the delivery system, registration uncertainties, and anatomical changes. The tumor position inside the patient, may also slightly varies from one day to the other [67, 68], especially in very soft tissues such as lung. This shift in the tumor position is called the baseline shift. In presence of organ motions, such as breathing, the baseline shift is defined as the shift of the mean tumor position. In conventional radiotherapy, with photons, this issue can be reduced by aligning the tumor positions during patient setup, instead of performing a global, or bony-structure based, image registration. In proton therapy, tumor based positioning may lead to range uncertainties and is therefore not recommended.

After the patient setup, the treatment fraction can be delivered. As described in the previous section, proton therapy treatments are typically delivered with DS or PBS. Double scattering is a passive delivery technique, while PBS requires an active scanning of the proton beam over the target.

The treatment of mobile tumors with proton therapy is still an active research topic. Indeed, organ motion may induce large variations of the proton range. Moreover, for dynamic delivery techniques, such as PBS, the delivery motion may interfere with the tumor motion. This motion interference leads to the so-called interplay effect[69, 70], that induces hot and cold spots in the dose distribution. Rescanning methods can be adopted to reduce this effect. The PBS plan is therefore splitted in several scans to deliver the dose more homogeneously over the

organ motion cycle.

The range uncertainty is one major limitation of proton therapy. To compensate for this uncertainty, the irradiated volume usually need to be enlarged, lowering the benefit of proton therapy in comparison to conventional radiotherapy. To decrease, the irradiated volume, the range uncertainty must first be reduced. In order to verify the actual proton range inside the patient and possibly adjust the treatment plan, new imaging techniques are being developed [71]. For instance, a gamma camera can be used during the delivery to detect prompt gammas emitted by nuclear interactions of protons with the patient materials [72, 73]. The prompt gamma profile can be correlated with the deposited dose using pre-computed simulations. This online imaging enables an in-vivo verification of the proton range during the treatment. This technology is not yet commercially available, but a few prototypes are in test. Alternatively, proton nuclear interactions also induce an activation of the patient materials. The activated nuclei eventually decay, releasing secondary particles, such as positrons that can be detected with a PET scan. PET imaging is already widely available, but the analysis of the PET signal for proton range monitoring is still a research topic.

1.4.8 Summary of proton range uncertainties

The sources of range uncertainties have been analyzed in the literature. For instance, the study of Paganetti [8], reported in Table 1.1, lists some of the main contributions to the range uncertainties. Some sources of uncertainties are dependent of the dose calculation. Paganetti estimated that using an analytical dose calculation algorithm, the total proton range uncertainty (1.5 sigma) is between $2.7\%+1.2\text{mm}$ and $4.6\%+1.2\text{mm}$, depending of the heterogeneity level in the geometry. However, using a Monte Carlo dose calculation algorithm, the range uncertainty estimation decreases to $2.4\%+1.2\text{mm}$, whatever the level of heterogeneity. Two important contributions are not considered in these numbers. First, the effect of inter- and intra-fraction anatomy changes should be taken into account. Second, image artifacts may also induce significant range uncertainties.

1.5 Contributions

The main objective of this PhD thesis is to address the issues related to treatment uncertainties in PBS proton therapy. First, some of the treatment uncertainties can be reduced using the fast and accurate Monte Carlo dose calculation software

Source of range uncertainty in the patient	Range uncertainty without Monte Carlo	Range uncertainty with Monte Carlo
Independent of dose calculation		
Measurement uncertainty in water for commissioning	± 0.3 mm	± 0.3 mm
Compensator design	± 0.2 mm	± 0.2 mm
Beam reproducibility	± 0.2 mm	± 0.2 mm
Patient setup	± 0.7 mm	± 0.7 mm
Dose calculation		
Biology (always positive) ^	$\pm \sim 0.8\%$	$\pm \sim 0.8\%$
CT imaging and calibration	$\pm 0.5\%$ ^a	$\pm 0.5\%$ ^a
CT conversion to tissue (excluding I-values)	$\pm 0.5\%$ ^b	$\pm 0.2\%$ ^g
CT grid size	$\pm 0.3\%$ ^c	$\pm 0.3\%$ ^c
Mean excitation energy (I-values) in tissues	$\pm 1.5\%$ ^d	$\pm 1.5\%$ ^d
Range degradation; complex inhomogeneities	-0.7% ^e	$\pm 0.1\%$
Range degradation; local lateral inhomogeneities *	$\pm 2.5\%$ ^f	$\pm 0.1\%$
Total (excluding *, ^)	$2.7\% + 1.2$ mm	$2.4\% + 1.2$ mm
Total (excluding ^)	$4.6\% + 1.2$ mm	$2.4\% + 1.2$ mm

The number are estimations based on finding by

^a Chvetsov and Paige (2010).

^b Schaffner and Pedroni (1998) and Matsufuji *et al* (1998).

^c Espana and Paganetti (2011).

^d ICRU (1993), Bichsel and Hiraoka (1992) and Kumazaki *et al* (2007).

^e Sawakuchi *et al* (2008), Bednarz *et al* (2010) and Urie *et al* (1986).

^f Bednarz *et al* (2010).

^g Espana and Paganetti (2010).

Table 1.1: Estimated proton range uncertainties and their sources and the potential of Monte Carlo for reducing the uncertainty. The estimations are average numbers based on 1.5 standard deviations. Extreme cases, such as lung treatments, might show bigger uncertainties. Source: Paganetti [8]

developed during the PhD. Multiple planning tools are also proposed to help to reduce the treatment plan sensitivity to uncertainties. Finally, a realistic robustness verification method is presented to verify the impact of remaining uncertainties.

1.5.1 Fast Monte Carlo simulations

Paganetti reported a significant difference in the proton range uncertainties depending on the type of algorithm employed for the dose calculation [8]. Monte Carlo algorithms are known to provide the best accuracy. However, their long computation times have historically prevented their use in the clinical routine. Recently, a few fast Monte Carlo algorithms have been implemented [53, 54]. MCsquare [15], the fast Monte Carlo code developed during this PhD [16], contributed to make accurate dose calculation accessible to the clinic. More than a dose engine, this multi-purpose Monte Carlo code enables more advanced simulations, such as the emission of prompt gammas, treatment uncertainties, or the

LET distribution. The implementation was optimized to fully exploit modern CPU architectures. In contrast to GPU implementations, MCsquare does not require specific hardware and can be run on dedicated calculation servers, or a simple laptop. This was the object of a publication during the PhD and is included in the chapter 2 of the present thesis. Today, MCsquare is used in multiple institutions over the world for research and clinical purposes. The source code is released open source on openmcsquare.org [15].

1.5.2 Improved planning tools

Due to the lack of planning tools adapted for proton therapy, the PTV margins, initially designed for conventional radiotherapy with photons, are still often employed in proton therapy, despite their inadequacy against proton range uncertainties. In order to fill this gap, several planning tools have been developed during the stay of the author at the proton center of the University of Pennsylvania (UPenn), in collaboration with UCL and IBA. First, motion evaluation tools have been developed to assess the impact of breathing motion on the proton range. Such tools, can help to choose the most adapted treatment modality for each patient. Moreover, the range variation can be studied for multiple beam orientations in order to select the beams that offer a reduced sensitivity to range uncertainties. Next, the PTV concept has been adapted for proton therapy. Indeed, proton range uncertainties are compensated by adding a customized margin in the proximal and distal part along the beam direction. This beam-specific PTV (BSPTV) takes into account a flat range uncertainty (typically 3%), but also the range variation caused by setup errors and breathing motion. These tools have been implemented in the open source application OpenReggui [18].

A second robust planning method has also been investigated at UCL. Indeed, a Monte Carlo based robust optimizer, MIROpt [17], was developed by Ana Baragan, in collaboration with the author of this thesis. MCsquare was adapted to provide the dose information and uncertainty models required by the robust optimizer. The current version of MIROpt integrates random and systematic setup errors, flat range errors, and breathing motion. More uncertainty models are available in MCsquare and may be integrated in MIROpt later. The source code of MIROpt is also released open source [17].

1.5.3 Realistic robustness verification

The tools provided by commercial TPS to verify the treatment plan are generally based on the same approximate dose calculation algorithms and limited uncertainty models as for the planning itself. Therefore, it does not yield an independent and comprehensive evaluation of the treatment plan. To address this issue, realistic simulations of treatment uncertainties have been developed to verify the robustness of proton therapy plans in a comprehensive way. In addition to the typical systematic setup errors and flat range uncertainties that are available in commercial solutions, several uncertainty models have been implemented directly into the Monte Carlo code MCsquare. Indeed, random setup errors, variable breathing amplitude, and interplay effects are also simulated by MCsquare. A model of the baseline shift is also being investigated, but not yet integrated in MCsquare. An original Monte Carlo approach has been proposed to randomly sample uncertainty scenarios and to provide a statistically sound robustness evaluation of the treatment plan. This novel method is fully described in another paper written during the PhD. The paper is entirely included in the chapter 4 of this thesis.

Chapter 2

Fast Monte Carlo simulations

The dose calculation process is used in the clinic to estimate the dose distribution inside the patient geometry based on a treatment plan. Dose calculation algorithms are generally classified in two categories. Analytical algorithms are based on macroscopic models of the dose. The computation is usually very fast, but the accuracy of analytical algorithms is often limited [8], especially in heterogeneous geometries. The second type of dose calculation algorithms is based on Monte Carlo techniques and provides an accurate simulation of particle interactions inside the patient geometry. The computation time is generally very long, but Monte Carlo simulations are recognized as the most accurate dose calculation for proton therapy. Today, analytical algorithms are still preferred in the clinical practice for their short computation time. In order to make Monte Carlo dose calculation acceptable in the clinic, their implementation had to be optimized. This represents the first challenge addressed in this thesis. The implementation of a fast Monte Carlo code is the basis for our subsequent developments like the robustness verification tool presented in chapter 4.

At the beginning of the development of our fast Monte Carlo code, another team was already working on a GPU (graphical processing unit) implementation [53]. We therefore decided to focus on CPU (central processing unit) architectures. Moreover, GPU computing is not optimal for Monte Carlo simulations of particle interactions. CPU architectures offer a better flexibility for the implementation of these algorithms. Afterwards, we can say that the decision to use CPU instead of GPU was judicious. Indeed, our CPU implementation is able to reach similar dose calculation speed as GPU, but does not require specific computer hardware. The developed fast Monte Carlo code, MCsquare is therefore compatible with any computers, ranging from simple laptops to dedicated calculation servers, allowing for easy use in research as well as optimal integration in clinics.

The rest of this chapter corresponds to the following paper:

Fast multipurpose Monte Carlo simulation for proton therapy using multi- and many-core CPU architectures

K. Souris, J.A. Lee, E. Sterpin

[Medical physics 2016; 43 \(4\): 1700-1712](#)

2.1 Introduction

Compared to photons, protons offer potentially decisive dosimetric advantages for cancer treatment. The Bragg peak and the finite range of charged particles allow

for better sparing of healthy tissues in treatment plans. However, the position of the Bragg peak may not be known with the required accuracy due to uncertainties in the treatment workflow (imaging and conversion to stopping powers, dose calculation, patient setup, particle beam delivery and organ motion). Safety margins grow with uncertainties and, consequently, reduce the potential benefit one can expect from protons. The sources of these range uncertainties have been studied by Paganetti [8]. In particular, it appears that the use of Monte Carlo based dose calculation instead of a typical analytical algorithm can reduce range uncertainties from $4.6\% + 1.2 \text{ mm}$ to $2.4\% + 1.2 \text{ mm}$ in the context of highly heterogeneous anatomies. However, their long computation time hinder their adoption in clinical routine.

Simplification of the physical models and optimization of their algorithmic implementation can reduce the running time of Monte Carlo simulations. For instance, the particle trajectories and interactions can be pre-computed for various materials and several energies. The dose is then calculated by repeating the appropriate tracks in the patient geometry [74]. It is also possible to increase computation efficiency by taking advantage of modern multi-core processors. Since the millions of simulated particles are statistically independent computation can easily be distributed among multiple parallel processing units. Graphical Processing Units (GPUs) are often used to speed up intensive computation [75, 76]. Their highly vectorized architecture is very powerful for signal and image processing or other calculations in which identical operations are applied to multiple data. However, the stochastic nature of particle interactions suggests that threads or instructions executed at a given time should be as independent as possible. Implementation on highly vectorized architectures like GPUs necessitates complex reformulation of the Monte Carlo algorithm with careful task distribution and scheduling. In particular, randomness in the emission of secondary particles during nuclear interactions usually requires the algorithm to be adapted and particles undergoing the same physical interaction can be regrouped to overcome the so-called thread divergence [76]. Nevertheless, the fast Monte Carlo package developed by Jia et al. [53], and other simplified algorithms [54, 77] have been successfully implemented on GPUs and are able to calculate dose distributions with 1% of statistical uncertainties in less than one minute. However, such a GPU implementation typically entails major redesign of existing code and algorithms in order to take into account hardware architecture specificities. As a result, most Monte Carlo tools that run on GPUs are often dedicated to a very specific task, like dose computation with the sole tracking of the protons.

GPUs are not the only alternative to carry out intensive computations very

fast [78]. Clusters of connected computers are the classical solution and still a very flexible and competitive one, except for its price, energy consumption, and maintenance. Intermediate approaches also exist, like the recently introduced Intel Xeon Phi coprocessors [79, 80, 81, 82]. They are available as affordable extension cards for workstations, just like GPUs. They combine the advantages of clusters, with many independent calculation units and those of GPUs, with access to a shared memory and vectorized calculation. This unique trade-off between cost, power, and flexibility led us investigating this new architecture, which allows for the implementation of more general Monte Carlo codes with no or little compromise.

A first contribution of this work consists in adapting the Monte Carlo algorithm presented by Fippel et al.[11] to modern CPU architectures. MCsquare, standing for many-core Monte Carlo, is our new Monte Carlo code. It has been designed for proton therapy simulations and can run on both regular multi-core Central Processing Unit (CPU) and Xeon Phi coprocessor architectures. Because accuracy of Monte Carlo methods is crucial, another contribution of this work aims at improving Fippel's algorithms. In our implementation, stopping powers and nuclear cross sections are computed from a database of multiple materials instead of applying correction to water data. Heavy charged secondary particles are fully simulated by scaling proton stopping powers according to their mass and charge. For comparison purposes, modularity in MCsquare allows the user to choose among several models and hence to reach his own trade-off between speed and accuracy. Moreover, MCsquare can simulate the emission of prompt gamma (PG), which can be used for in vivo range verification [83, 84, 85] before and/or during treatment delivery.

MCsquare performances, in terms of accuracy and speed, have been benchmarked against GATE [86, 87] as a gold standard. This application is an adaptation to medical uses of the Geant4 [88] Monte Carlo toolkit.

2.2 Methods

The algorithms and methods implemented in MCsquare are presented in this section. The first subsection introduces all data and parameters involved in MCsquare. The methods for transporting the charged particles are described in the second subsection and the physical models for electromagnetic and nuclear interactions follows in the next two subsections. The new Intel Xeon Phi coprocessor is then introduced and compared to the regular CPU architecture, followed by the

code optimizations that take full benefit of the computational resources offered by both architectures. All these subsections are summarized in Figure 2.1, which represents the flowchart of the implementation of MCsquare. Finally the different simulation setups used to benchmark MCsquare are described in the last subsection.

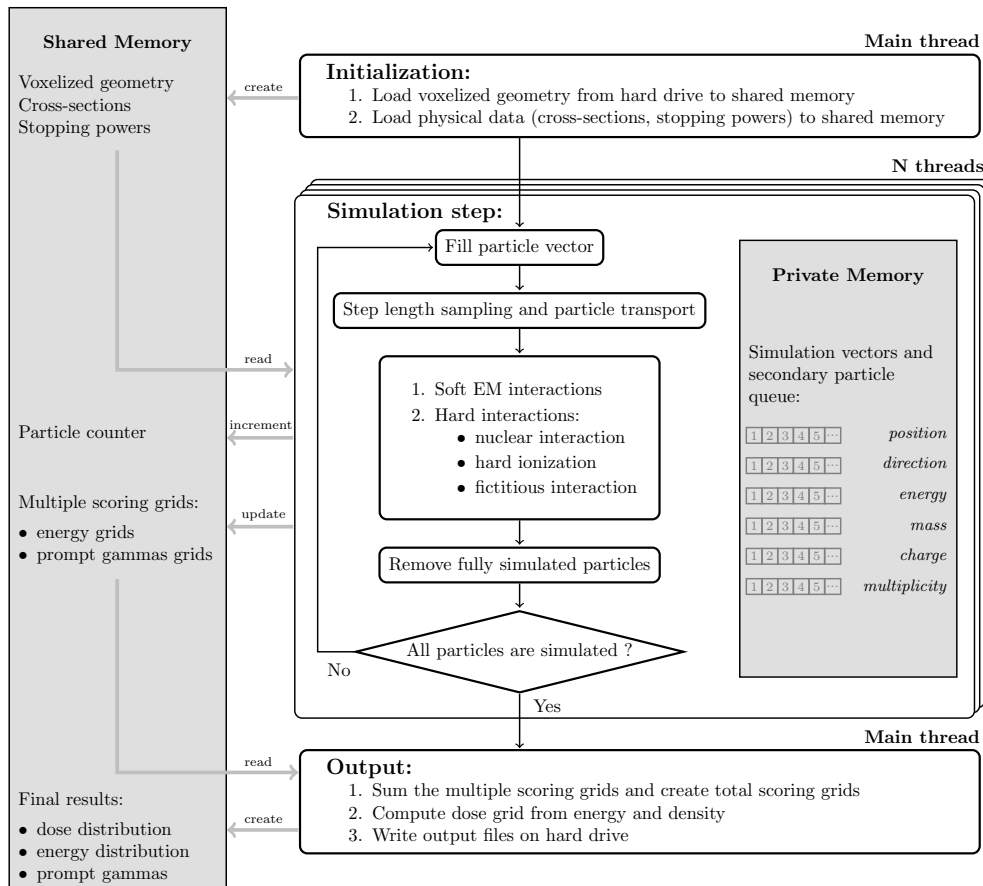


Figure 2.1: Flowchart representing the main steps of the algorithm implemented in MCsquare.

2.2.1 Input data

Dose calculation is performed in a voxelized geometry. This geometry can be provided to MCsquare using a simple ASCII text file or a MHD binary file. This MHD format is dedicated to medical images and is compatible with the GATE applica-

tion. The type of material is required by the Monte Carlo algorithm in order to compute the correct energy loss and select the corresponding cross sections. Each voxel is therefore labelled with the appropriate material according to its density. Multiple text files containing stopping powers and cross sections for each material are integrated in MCsquare, to form a database of all physical information required in the simulation. Users can easily modify, replace, or augment this database.

Various simulation parameters can be modified in the configuration file, in order to comply with specific constraints related to the intended application and to attain the best trade-off between computation speed and accuracy. For example beam parameters allow the energy, the position and the shape to be defined. Moreover, the transport of secondary particles (protons, deuterons, and alphas) can be enabled or disabled individually. Other parameters affecting the simulation accuracy will be presented in the next sections.

2.2.2 Transport algorithm

Charged particles undergo a huge number of electromagnetic (EM) interactions along their path. It is not possible to simulate all of them in a reasonable computation time. Particle interactions are therefore simulated in a condensed way using a class II mechanism [89, 90]. The proton trajectory is simulated step-by-step with random lengths. The energy loss is evaluated at each step, combining the contribution of multiple small inelastic EM interactions, also called soft interactions. Only hard interactions such as nuclear interactions or collisions with orbital electrons involving energy transfers higher than a user defined threshold T_{emin} are simulated individually at the end of each step. The process is continued until the kinetic energy falls below another user defined threshold $E_{cut_{pro}}$. At that point the remaining energy is locally absorbed.

The step length is defined as the distance between two hard interactions. The step length is therefore sampled at the beginning of each simulation step according to the total cross section of hard events, namely, hard ionizations and nuclear interactions. As cross sections can vary along the step, a third "fictitious" interaction is added. The cross section of this fictitious interaction varies along the step in order to keep the total cross section constant. This mechanism has been introduced in PENELOPE [89] in order to sample the step length correctly and to stabilize the simulation by constraining the step length. The user is able to define the maximum distance D_{max} travelled and the maximum fraction ϵ_{max} of energy loss during a step. Typically, these thresholds are fixed respectively to 2 mm and 25%. At the end of the step one physical or one fictitious interaction is simulated

according to their cross section.

Interaction probabilities vary from one voxel to another depending on the material and the density. Consequently, errors may be introduced if the particle crosses freely more than one voxel along its step. Two methods are implemented in MCsquare to handle interface crossing. In the first, "accurate", method the simulation step is inevitably stopped at each voxel interface using the fictitious interaction mechanism. In the second, "fast", method the simulation step is no longer stopped at each interface. Each time the particle enters a new voxel the step length is scaled by the local density and the mass stopping power ratios. The energy lost by the particle during the step is then scored in a random voxel along that path. This allows the different EM characteristics of each material to be taken into account without spending too much time restarting a new simulation step for each voxel. However this fast method leads to incorrect sampling of nuclear reactions near interfaces, which may slightly affect the simulation results.

2.2.3 Electromagnetic interactions

Protons lose their energy during multiple soft EM interactions, which are simulated in a condensed way, and hard inelastic interactions, which are simulated individually. The energy loss resulting from the condensed interactions is calculated using a database of stopping powers tabulated up to 300 MeV. The user can add his/her own table or opt for those found in the PSTAR or Geant4 databases, which are already integrated in MCsquare for each material recommended by Schneider [9]. Other heavy charged particles than protons can also be simulated by scaling the proton stopping powers [46]. The Bethe-Bloch formula can be approximately viewed as a function of the particle charge z and velocity β . Knowing the stopping powers $S_1 = -\frac{dE_1}{dx}$ of some particle with charge z_1 and mass M_1 , the stopping power S_2 of another particle (z_2, M_2) with kinetic energy E_2 can therefore be approximated by

$$S_2(E_2) = -\frac{dE_2}{dx}(E_2) = -\frac{z_2^2}{z_1^2} \frac{dE_1}{dx} \left(E_2 \frac{M_1}{M_2} \right) = \frac{z_2^2}{z_1^2} S_1 \left(E_2 \frac{M_1}{M_2} \right).$$

The rest of the EM algorithm essentially follows Fippel's method [11]. The contribution of hard ionizations (thus with energy transfers larger than Te_{min}) is first subtracted from the total stopping powers, leading to the restricted stopping powers $L_{Te_{min}}(E)$. The mean energy loss $\overline{\Delta E}$ along the step s is then calculated according to

$$\overline{\Delta E} = \int_0^s L_{T_{emin}} [E(s')] ds'.$$

The integral is evaluated numerically using formulas established by Kawrakow [91].

In inelastic EM interactions, energy transfers are random, leading to the so-called energy straggling phenomenon and the width of the Bragg peak. To simulate straggling, a Gaussian perturbation with standard deviation δE is added to the mean energy loss. This perturbation has been initially introduced by Bohr [92] and δE can be calculated from the step length and material characteristics using the method described in the Geant4 Physics reference manual [90] and ICRU 49 report [93].

The angular deflection resulting from multiple elastic Coulomb interactions is also sampled for each step from a Gaussian distribution with zero mean and variance given by the formula proposed initially by Rossi and Greisen [94]

$$\sigma_{MCS}^2 = \left(\frac{E_s}{\beta pc} \right)^2 \frac{s}{X_0},$$

where X_0 is the radiation length of the material, E_s is a coefficient tuned to reproduce other simulated or experimental results, β is the ratio of the particle velocity and the speed of light c , and p is the particle momentum.

Finally, the type of hard interaction (inelastic EM, nuclear or fictitious) is randomly sampled at the end of the step, depending on the cross section of each process. The hard EM cross section can be analytically calculated by integrating the ionization differential cross section as described in Fippel's method. When an ionization process is selected, the δ -electron kinetic energy is sampled from the differential cross section formula, using a rejection method. Secondary electrons are currently not transported in MCsquare and their energies are locally absorbed. This approximation does not impact the dose as long as the range of δ -electrons remains small in comparison with the voxel size (1-2 mm). Electron transport will be implemented in a future version of MCsquare, in order to further increase accuracy of the dose distribution in cases where the δ -electron range should not be neglected. For instance a 200 MeV proton beam can transfer a maximum energy of 0.5 MeV to δ -electrons which correspond to a range of 7 mm in lung tissues.

2.2.4 Nuclear interactions

The ICRU 44 report [95] shows that only a few elements are necessary to describe human tissues. All materials are therefore stored in the MCsquare database as a composition of H, C, N, O, P, and Ca, which are the only elements found in human tissues with concentrations higher than 1%. The ICRU 63 report [96] provides for each nuclear species, except hydrogen, all integral and differential cross sections required by MCsquare to simulate proton-nucleus elastic and inelastic nuclear interactions. Proton-nucleus elastic collisions are directly sampled from the ICRU 63 database. For each event, the kinetic energy transferred to the nucleus is locally absorbed.

To simulate inelastic interactions, we follow a similar methodology as Sterpin et al. [97]. Various secondary particles are produced in nuclear inelastic events, namely neutrons, protons, deuterons, alphas, prompt gammas and heavier recoil nuclei. In the present version, prompt gammas are produced but not transported. Neutrons are neglected, since the comparison with Geant4 results shows indeed that their contribution to the local dose distribution is less than 0.5% [11, 98, 99]. The kinetic energy transferred to heavier recoil nuclei is locally absorbed. All other charged particles are explicitly simulated. Their kinetic energy and emission angle are directly sampled from ICRU 63 data.

Nuclear elastic collisions between protons and hydrogen nuclei are treated separately. The total cross section for proton-proton elastic nuclear interaction is obtained with a curve that fits SAID data [100], such as proposed by Fippel and Soukup [11]. After colliding, the deviation angle is uniformly sampled in the reference system of the center of mass. The angular deflections and kinetic energies of both protons are then calculated in the reference system of the laboratory using kinematic transformations.

For each element constituting the material the elastic and inelastic nuclear cross sections are summed in order to obtain the interaction probability. These probabilities are employed to sample the target nucleus during a nuclear interaction. All these individual nuclear cross sections are also summed to the ionization and the fictitious cross sections to obtain the total hard interaction cross section from which the step length is sampled.

2.2.5 Xeon Phi architecture

The Intel Xeon Phi coprocessor can be seen as a massively parallel version of regular processors. It is provided as a PCI extension card that runs its own Linux embedded operating system. This new architecture integrates many small CPU cores in a single chip, which enables the simulation of several particles at the same time. Various approaches to parallelization can be followed to accelerate computation, each one being adapted to specific applications. On the one hand, task parallelization has been popularized in modern computer with the advent of multi core processors, which can execute multiple independent tasks simultaneously. On the other hand, GPUs take benefit of data parallelization, through the single instruction multiple data (SIMD) processing mode. Technical details on the GPU SIMD programming model can be found in recent surveys [75, 76]. This architecture can execute the same operation on multiple data simultaneously. This vectorial computation unit is very powerful for signal and image processing.

In order to optimize performance and power consumption, CPU architectures combine both types of parallelism and thus follows a multiple instruction on multiple data architecture (MIMD). Data parallelism is enabled by the small SIMD unit contained in each CPU core. The new coprocessor offers even more data parallelism thanks to its larger SIMD units capable of carrying out the same operation on 16 or 8 words of 32 or 64 bits like, e.g., single- or double-precision floating-point numbers. This brings the theoretical peak performance to 1 TFLOPS for double-precision operations and 2 TFLOPS for single-precision. This is lower than modern GPUs which reach 3.52 TFLOPS in single precision for the NVIDIA Tesla K20. However the flexibility offered by the multiple independent cores of the Xeon Phi architecture makes the exploitation of these maximum theoretical performances easier for complex applications such as Monte Carlo simulations. Technical specifications of the Xeon Phi 5110P coprocessor and both Xeon processors used in this study are compared in table 2.1.

Although the memory bandwidth of the Xeon Phi is larger than in many modern computation accelerators, it remains one of the main bottlenecks. All cores have to share the bandwidth of the external memory and many clock ticks may be spent just waiting for the required data to arrive. In order to limit this waste, two levels of high-speed cache memory are inserted between the external memory and calculation units. The second mechanism employed to indirectly overcome memory latency is multithreading. Four hardware threads are implemented in each core. Every time a thread has to wait for data, it is put on hold and another thread is loaded and run in the computation unit.

	Xeon E5-2670	Xeon Phi 5110P
Type	main CPUs	coprocessor
Number of cores	8	60
Threads per core	2	4
Frequency	2.6 - 3.0 GHz	1.053 GHz
Peak performance (single precision)	166.4 GFLOPS	2000 GFLOPS
SIMD units	256-bit	512-bit
RAM size	32 GB	8 GB
RAM Bandwidth	51.2 GB/s	320 GB/s
L1 Cache memory	8 × 32 KB	60 × 32 KB
L2 Cache memory	8 × 256 KB	60 × 512 KB
L3 Cache memory	20 MB shared	-

Table 2.1: Comparison of the specifications of the main CPU and the coprocessor employed in this study. Two CPUs Xeon E5-2670 and one coprocessor Xeon Phi 5110P are employed in this study. This accounts for a total of 16 CPU cores (32 threads) and 60 coprocessor cores (240 threads).

2.2.6 Code implementation

Programming and parallelization techniques for the Xeon Phi remain essentially the same as for usual CPUs. Any code already well optimized for multithreading and vector instructions on modern CPUs could be adapted to the coprocessor without difficulties. This also means that any optimization effort for one architecture may also benefit to the other one. This main advantage highly reduces the development and maintenance costs in comparison to other accelerators. MCsquare consists therefore of a single code which can be compiled and run on both architectures.

MCsquare has been implemented in regular C language and compiled with the Intel compiler. Several libraries and parallelization methods can be employed in order to exploit all the capabilities of the Xeon Phi [79, 81, 82, 101]. Intel's math kernel library (MKL) provides several mathematical functions that are already optimized for the Xeon Phi coprocessor. In particular, MCsquare uses the MCG59 random number generator of this library. This 59-bit multiplicative congruential generator offers a sequence of randomly distributed number with a period of $1.4 \cdot 10^{17}$ which suits very well the requirements of such Monte Carlo methods.

Modern compilers such as GCC or Intel solution are able to automatically optimize simple loops in the code to exploit data parallelism. To achieve better performances, the developer can guide the compiler using pragma instructions from OpenMP, Intel Cilk Plus, or other libraries. Intel Cilk Plus also provides an

explicit array notation [102] to guide the compiler in vectorizing more complex methods. In this work, the entire Monte Carlo algorithm has been written using that explicit vector notation. This C language extension allows us to define variables as vectors and explicitly write the code in a vectorized way. For all particle parameters (mass, charge, position, direction, energy, and multiplicity) one vector is created in each thread. These vectors regroup the status of 8 particles in double-precision mode or 16 particles in single-precision per thread. Completely simulated particles are replaced in the vectors after each step with new primary or secondary particles generated in previous steps.

To increase efficiency, simulation is performed simultaneously on all available threads. There are several possibilities to implement task parallelization such as MPI and OpenMP libraries [103, 104]. MPI (message passing interface) is a very generic framework allowing for point-to-point communication between processes running on distributed memory systems. In contrast, OpenMP is an API for shared-memory parallel programming. MCsquare has been implemented using the *pragma omp* directives from OpenMP, which are used to specify how code blocks can be run in parallel among the available threads, as shown in Figure 2.1. OpenMP allows the programmer to parallelize the code without explicitly implementing the management of multi threads and their communication. At runtime, MCsquare automatically adjusts the number of threads and task distribution according to the targeted architecture. Having 60 cores, 4 threads per core, and the capability to process 16 element vectors, the Xeon Phi coprocessor can track up to 3840 particles simultaneously.

Critical data such as simulation vectors are stored in a memory location that is private for each computation thread in order to guarantee that it can only be accessed by the associated thread. Data shared between multiple threads require more attention. Indeed, a common issue in parallel computing is memory conflicts. For instance, each thread increments a shared counter every time a primary particle has been completely simulated in order to keep track of the overall progression. However, when several threads attempt to write simultaneously in the same memory space, the final result may be corrupted. The use of atomic operations, implemented in OpenMP, can prevent this problem. When a thread initiates a memory writing operation in a shared variable, the atomic mechanism locks this memory space and forbids other threads to access it until completion of the first operation. However, serialization of memory writing may impact the computation efficiency for frequent operations such as the scoring of energy deposited during each interaction. To avoid this, the deposited energy is scored in multiple arrays which are finally summed at the end of the simulation. Depending on the beam

dimensions and voxel sizes, we found that using a single scoring grid can cause an increase of around 20% of the computation time on the Xeon Phi coprocessor in comparison with a simulation with one scoring grid per thread. The user has to find a tradeoff between calculation speed and memory size needed to store all scoring grids.

2.2.7 Benchmark

In order to validate MCsquare, its results are compared to those of the Geant4 Monte Carlo toolkit. Software such as GATE [86, 87] or TOPAS [105] may be very useful to simplify the configuration of Geant4 simulations. In this study the GATE platform version 6.2 is employed with Geant4 version 9.5 p2. GATE is often considered as a gold standard of Monte Carlo simulations for various medical applications [106]. The default configuration of the GATE *HadrontherapyStandardPhys* physics list was employed. In this physics list, the EM interactions are simulated with the generic HadronIonisation process from Geant4, while elastic nuclear interactions are handled by the G4LElastic process. All our Geant4 simulations have been performed using two nuclear inelastic interaction models, the Binary Cascade or the Precompound models. The simulation parameters have been optimized and validated by Grevillot et al. [107].

Various configurations of MCsquare are possible. Two configurations of MCsquare are compared for these validation tests. First, MCsquare is configured to reach the highest accuracy: the simulation of all secondary heavy charged particles is enabled, the transport algorithm stops the simulation step at each interface, and stopping powers are acquired from the Geant4 database. The second configuration is optimized for calculation speed. In this fast configuration, particles are transported through the voxelized geometry using the step scaling method. Moreover, secondary heavy charged particles (alphas and deuterons) are now locally absorbed. Only secondary protons are still fully transported. This last approximation is often used in fast Monte Carlo engines, due to the very short range of these heavy hadrons. In order to maximize SIMD parallelization all data are encoded in a single precision floating point format. The various configuration parameters are summarized in table 2.2.

The benchmarks are performed in different phantoms. A simple homogeneous water phantom is first employed in order to fully characterise the dose and prompt gamma distributions. Second, the accuracy of the proton range is verified for various tissues. The impact of heterogeneities on the dose calculation accuracy is then studied by simulating the proton beam in a non-homogeneous phantom.

Parameters	Possible values	MCsquare (accurate)	MCsquare (fast)
Computation parameters			
Number of CPU threads	from 0 to max CPU threads	32	32
Number of Xeon Phi threads	from 0 to max Xeon Phi threads	240	240
Floating point precision (data/compute/score)	single (S) or double (D)	S/S/S	S/S/S
Simulation parameters			
Number of primary protons	from 1 to 4 294 967 295	10^7	10^7
Hard ionization threshold (T_{emin})	from 0.001 to 10.0 MeV	0.05 MeV	0.05 MeV
Particle absorption threshold ($E_{cut_{pro}}$)	from 0.001 to 200 MeV	0.5 MeV	0.5 MeV
Maximum step length (D_{max})	from 0.001 to 100 cm	0.2 cm	0.2 cm
Maximum fraction of energy loss per step (ϵ_{max})	from 0.001 to 1.0	0.25	0.25
Nuclear interaction	enabled or disabled	enabled	enabled
Secondary particle simulation (p+/deuteron/ α/γ)	enabled (E) or disabled (D)	E/E/E/E	E/D/D/D
Transport algorithm	voxel interface or scaling method	voxel interface method	scaling method
Data parameters			
Stopping power database	PSTAR, Geant4 or user data	Geant4	Geant4
Nuclear model	ICRU63 or user data	ICRU63	ICRU63
Input file format	ASCII or MHD	ASCII	ASCII
Output file format	ASCII or MHD	ASCII	ASCII

Table 2.2: List of the various configuration parameters of MCsquare and the values employed in this study.

The simulation times of MCsquare and Geant4 are finally compared in the last subsection.

Homogeneous water phantom.

This first test involves a mono-energetic (200 MeV) proton pencil beam simulated in a $60 \times 60 \times 41$ cm³ homogeneous water phantom. The resulting physical quantities are scored in a grid of 1 mm³ cubic voxels. The depth dose curves of Geant4 and MCsquare are then compared after integrating 3D dose distributions over the entire 60×60 cm² cross section area, in order to take into account the dose deposited by secondary particles emitted with a large angle. To check the multiple Coulomb scattering implementation, transverse profiles acquired at depths of 20 cm and 25 cm are then compared.

In addition to dose distributions, prompt gamma (PG) production profiles are also analysed. While MCsquare directly provides this information by enabling the PG option in the configuration file, additional work is required to extract the profiles from Geant4 simulations. All photons are firstly recorded in a phase space file. The Root framework [108] is then employed to reconstruct a histogram of the PG emission. Our Root script identifies easily the prompt gammas, depending on the interaction that produced the photon. The production coordinates are then used to increment histogram bins. For the PG benchmark a third Monte Carlo code is compared. PENH [97, 109] is an extension of PENELOPE to protons based on the same ICRU 63 nuclear cross sections used for MCsquare.

Homogeneous tissue phantom.

MCsquare is also validated for various tissues. For the sake of conciseness, results for only two materials are reported. A 200 MeV proton pencil beam is simulated in average soft tissue and cortical bone, which correspond in Geant4 respectively to *G4_TISSUE_SOFT_ICRP* and *G4_BONE_CORTICAL_ICRP* materials. The same chemical composition is employed in MCsquare and the densities are set respectively to 1.03 g/cm^3 and 1.92 g/cm^3 . The phantom geometry and the scoring parameters are taken from the previous test.

Heterogeneous phantom.

The inhomogeneous phantom described by Fippel and Soukup [11] and illustrated in Figure 2.2 has been reproduced to assess the dose calculation accuracy in heterogeneous geometries. The 150 MeV proton pencil beam is simulated in a $4 \times 4 \times 30 \text{ cm}^3$ water phantom containing a $1 \times 2 \times 5 \text{ cm}^3$ cortical bone slab of density 1.92 g/cm^3 and a $1 \times 2 \times 5 \text{ cm}^3$ lung slab of density 0.26 g/cm^3 , forming an interface along the beam path. The depth-dose curve is integrated over the $4 \times 4 \text{ cm}^2$ section surface and transverse profiles are taken at depths of 12 cm and 19 cm. In order to evaluate the impact of heterogeneities on the calculation of nuclear interactions, the prompt gamma production profiles resulting from both MCsquare configurations are then compared.

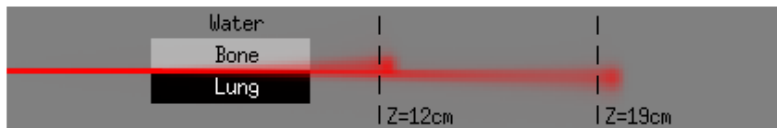


Figure 2.2: Dose deposited by a 150 MeV proton pencil beam into a heterogeneous phantom composed of lung, water, and bone materials.

Computation speeds.

The simulation speed of Geant4 and MCsquare are measured for the different benchmarked setups and various configurations. Currently, the GATE platform does not fully supports multithreaded computation and is not compatible with the Xeon Phi architecture. Unlike MCsquare, Geant4 is not optimized for speed and

therefore the reported speeds may not be compared directly. Nevertheless, they illustrate the gains in time made possible by specialized Monte Carlo code.

Simulations have been first executed on one single CPU thread. Moreover, Geant4 simulations have been repeated with a simplified physics list to reproduce MCsquare physics. The simulation of secondary electrons, photons, and neutrons are disabled. Instead, secondary electrons are locally absorbed. The test is also performed on one Xeon Phi thread using MCsquare in order to assess the performance of the code on the new architecture. For each benchmarked setup and configuration, the time required to simulate 10^7 primary protons is measured and converted in number of protons simulated per second. These measurements are repeated 5 times in order to obtain the average value and the standard deviation of the computation speed.

Since the Xeon Phi embeds its own Linux operating system, the *time* command is employed on both the host system and the coprocessor to measure the total execution time of Geant4 and MCsquare. Moreover the *omp_get_wtime* instruction of the OpenMP library is also used to internally check the computation time of the different stages in MCsquare whatever the adopted architecture.

To evaluate scalability of MCsquare, a second test is performed on both platforms. Computation speed is measured as a function of the number of computation threads employed for the simulation of 200 MeV protons into the water phantom.

In contrast with Geant4, MCsquare is able to benefit from all resources provided by the host CPUs (2 Intel Xeon E5-2670, 2x8 cores, 2.6 GHz) and the coprocessor (Intel Xeon Phi 5110P, 60 cores, 1.053 GHz). The entire computation power of the station is then exploited during a third test by performing the calculation with MCsquare simultaneously on the 240 Xeon Phi threads and the 32 host CPU threads.

2.3 Results

The simulation results are presented in this section. The accuracy of MCsquare is demonstrated by comparing integrated depth-dose curves, transverse profiles and prompt gamma production profiles with Geant4 results for different geometry and materials.

2.3.1 Homogeneous water phantom

The integrated depth-dose curves of a 200 MeV proton pencil beam has been simulated in water using MCsquare and compared with the results provided by the Geant4 Binary Cascade and Precompound nuclear models. The deviations caused by the different nuclear models are shown in Figure 2.3. MCsquare deviates less than 2% from Geant4 Binary Cascade at shallow depths (< 15 cm), whereas it comes closer ($< 2\%$ - 1mm deviations) to Geant4 Precompound deeper. The depth of the three Bragg peaks matches within 0.3 mm, which demonstrates the adequacy of the electromagnetic algorithm. The transverse profiles are also compared and, as shown in Figure 2.4, MCsquare is again in good agreement with Geant4 (1% root-mean-square deviations).

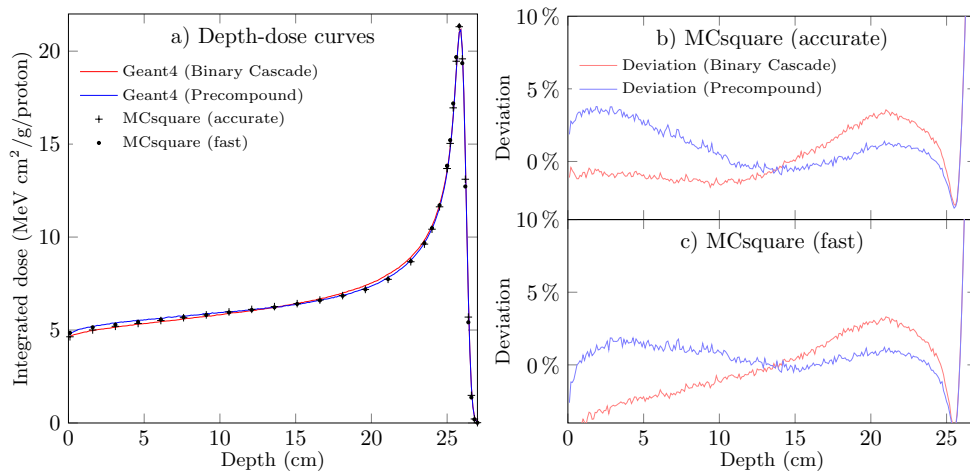


Figure 2.3: a) Integrated depth-dose curves of a 200 MeV proton pencil beam in water, simulated using two nuclear models of Geant4 (Binary Cascade, Precompound) and MCsquare. b,c) Deviations of MCsquare results in comparison with Geant4 simulations.

Finally the prompt gamma production profiles are compared in Figure 2.5. These results are much more sensitive to the nuclear model. Large differences (up to 66%) are observed between MCsquare and Geant4. However, the deviations with respect to PENH do not exceed 2% - 1mm.

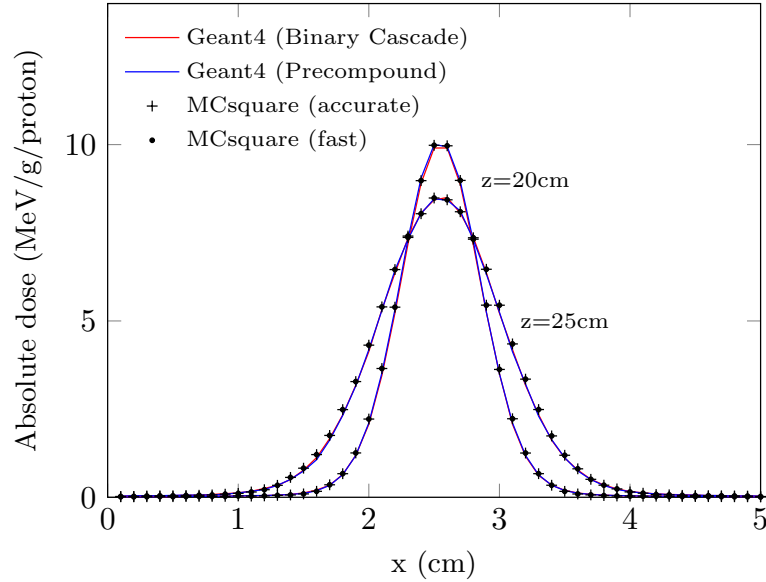


Figure 2.4: Transverse profiles of a 200 MeV proton pencil beam in depths of 20 cm and 25 cm into water, simulated using two nuclear models of Geant4 (Binary Cascade, Precompound) and MCsquare.

2.3.2 Homogeneous tissue phantom

The same proton pencil beam has been simulated in homogeneous phantoms of various tissues. MCsquare reproduced the proton range within 0.2% for every material. Figure 2.6 shows the integrated depth-dose results for the ICRP soft tissue and ICRP cortical bone. The differences between the results provided by the accurate configuration of MCsquare and both Geant4 models can be observed in the deviation plots. Once again MCsquare is closer to the Binary Cascade model at small depth (higher energy) and gets closer to the Precompound Model deeper (lower energy). The deviations between both MCsquare configurations are drawn in the right part of Figure 2.6. We can observe an overestimation of 4% of the entrance dose by the fast MCsquare simulation.

2.3.3 Heterogeneous phantom

The simulation is performed in the heterogeneous phantom represented in Figure 2.2. The resulting integrated depth-dose curves and transverse profiles are shown in Figures 2.7 and 2.8. MCsquare slightly overestimates the deposited dose in the first Bragg peak. The transverse profiles also reveal a small underestima-

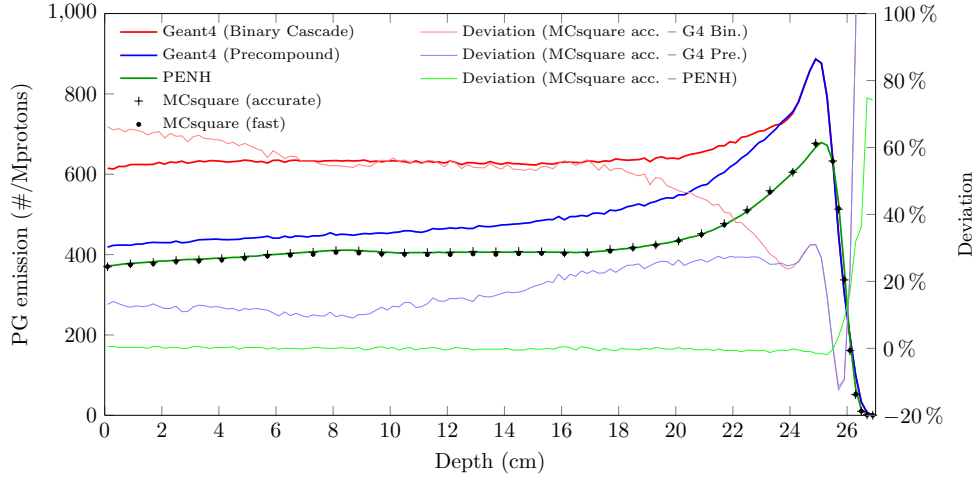


Figure 2.5: Dark curves represent the prompt gamma emission profiles for a 200 MeV proton pencil beam simulated in water using two nuclear models of Geant4 (Binary Cascade and Precompound), PENH and MCsquare. The deviations of MCsquare results (accurate configuration) in comparison with Geant4 and PENH are illustrated by the light curves.

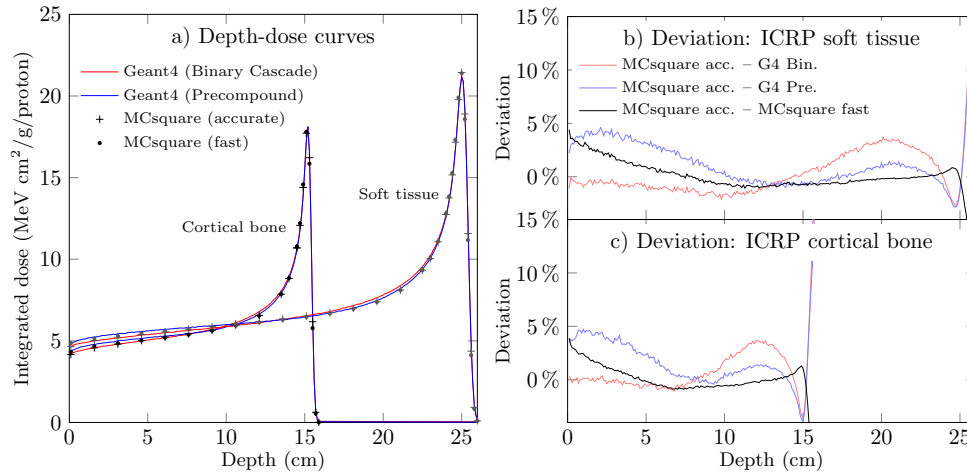


Figure 2.6: a) Integrated depth-dose curves of a 200 MeV proton pencil beam simulated in Soft tissue and Cortical bone using both nuclear models of Geant4 and both configurations of MCsquare. b,c) Deviations of the accurate simulation of MCsquare in comparison with Geant4 results and the fast simulation of MCsquare for both materials.

tion of the dose in the bone material and a slight overestimation in the lung. The prompt gamma production profiles generated by both configurations of MCsquare are compared in Figure 2.9. Both simulation modes provide similar results except at the interfaces where 16% deviations are induced by a smoother variation in the fast simulation results.

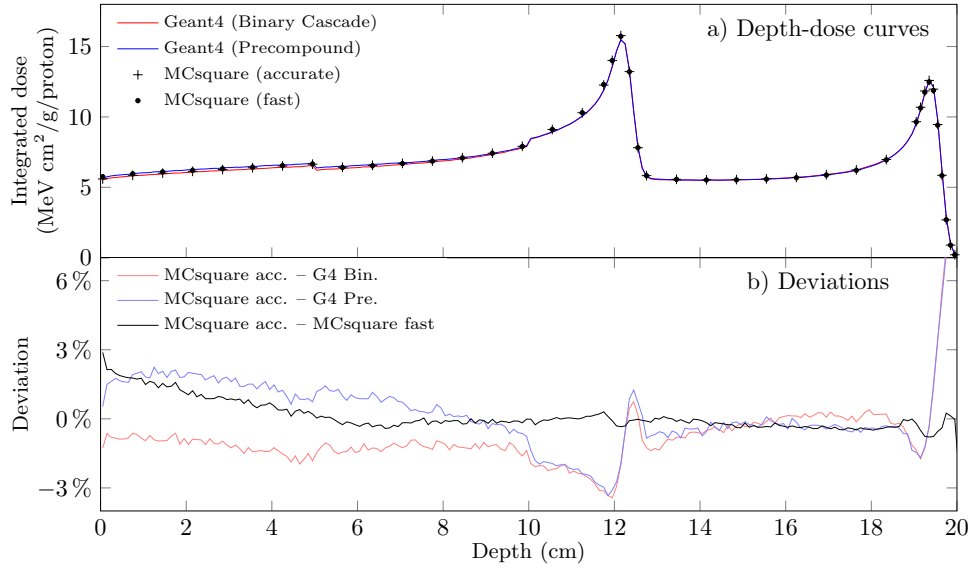


Figure 2.7: a) Integrated depth-dose curves for a 150 MeV proton pencil beam simulated in the heterogeneous phantom of Figure 2.2 using both nuclear models of Geant4 and both configurations of MCsquare. b) Deviations of the accurate simulation of MCsquare in comparison with Geant4 results and the fast simulation of MCsquare.

2.3.4 Computation speeds

The computation speeds for the four previous benchmark setups are listed in Table 2.3. Using equivalent computation resources, MCsquare runs 18 to 28 times faster than Geant4 for the accurate configuration and 22 to 40 times faster for the fast configuration. We also observe that MCsquare runs faster on a CPU thread than on a Xeon Phi thread.

The results of the scalability test performed with the accurate configuration of MCsquare on both architectures are plotted in Figure 2.10. For both architectures, the curve is linear until reaching their number of physical cores (16 CPU cores and

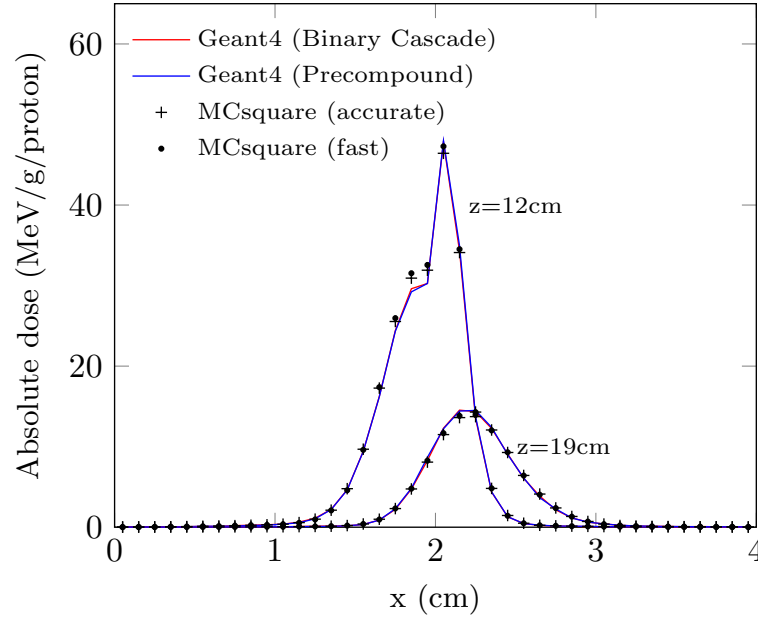


Figure 2.8: Transverse profiles of a 150 MeV proton pencil beam in depths of 12 cm and 19 cm into the heterogeneous phantom of Figure 2.2, simulated using both nuclear models of Geant4 and both configurations of MCsquare.

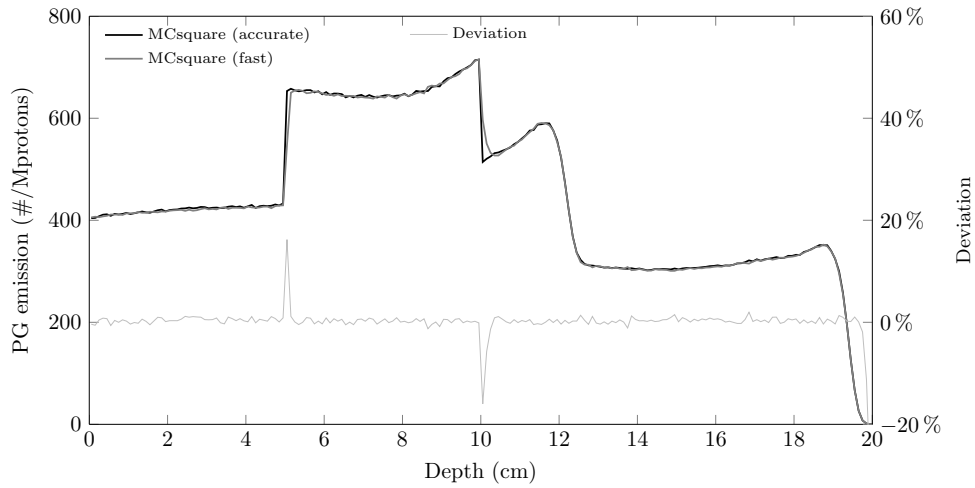


Figure 2.9: Prompt gamma emission profiles for a 150 MeV proton pencil beam simulated in the heterogeneous phantom of Figure 2.2 using MCsquare. The deviation between both configurations of MCsquare is illustrated by the light curve.

Monte Carlo code	Water phantom [protons / s / thread]	Soft tissue phantom [protons / s / thread]	Bone phantom [protons / s / thread]	Heterogeneous phantom [protons / s / thread]
GEANT4 (standard physics)	419.7 $\pm$ 5.7 (CPU)	296.6 $\pm$ 7.3 (CPU)	438.6 $\pm$ 7.2 (CPU)	486.0 $\pm$ 7.7 (CPU)
GEANT4 (simplified physics)	534.4 $\pm$ 7.9 (CPU)	355.7 $\pm$ 5.9 (CPU)	535.3 $\pm$ 6.5 (CPU)	628.9 $\pm$ 7.6 (CPU)
MCsquare (accurate config)	9803.7 $\pm$ 149.5 (CPU) 872.3 $\pm$ 4.3 (Xeon Phi)	9978.3 $\pm$ 190.4 (CPU) 890.6 $\pm$ 5.7 (Xeon Phi)	12663.1 $\pm$ 157.8 (CPU) 1127.2 $\pm$ 4.4 (Xeon Phi)	13411.4 $\pm$ 233.7 (CPU) 1126.4 $\pm$ 5.5 (Xeon Phi)
MCsquare (fast config)	14064.4 $\pm$ 280.9 (CPU) 1417.3 $\pm$ 4.9 (Xeon Phi)	14235.3 $\pm$ 439.1 (CPU) 1446.5 $\pm$ 7.1 (Xeon Phi)	16551.4 $\pm$ 418.3 (CPU) 1684.3 $\pm$ 7.0 (Xeon Phi)	13893.5 $\pm$ 257.8 (CPU) 1431.5 $\pm$ 5.4 (Xeon Phi)

Table 2.3: Number of protons simulated per second and per thread (CPU or Xeon Phi) for the different benchmark configurations and the different Monte Carlo engines.

60 Xeon Phi cores). After that point, the curves continue to rise linearly with a lower slope. The CPU curve stops when reaching its maximum 32 threads, while the Xeon Phi pursues to 240 threads. Despite its higher number of threads, the Xeon Phi maximum speed is lower than the CPUs.

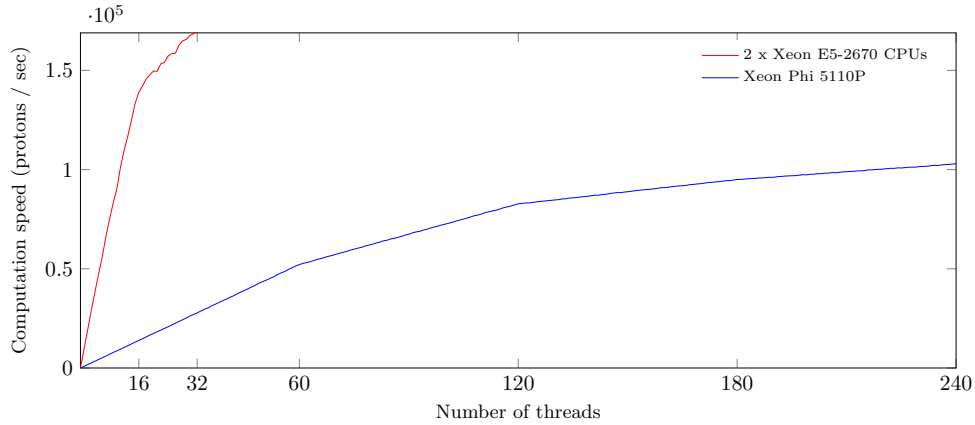


Figure 2.10: This scalability test describe the computation speed of MCsquare as a function of the number of calculation threads employed.

Running the accurate configuration of MCsquare simultaneously on the host CPUs and the coprocessor achieved simulation times of 38s, 37s, 29s, and 40s respectively for the water, soft tissue, bone, and heterogeneous phantom. The fast configuration of MCsquare reduces the computation times to 25s, 25s, 21s, and 35s respectively.

2.4 Discussion

This study has investigated acceleration techniques that considerably reduce the computation time of Monte Carlo simulations in proton therapy. Physical models and algorithms have been implemented to exploit task and data parallelization offered by modern CPU architectures. The dosimetric results and prompt gamma profiles provided by the new Monte Carlo software have been successfully validated with Geant4 simulations. Despite the simplifications implemented in MCsquare to accelerate the simulation of proton transport, the dosimetric results are in good agreement with Geant4. Using the same stopping powers as Geant4, the EM algorithm is able to reproduce the expected depth of the Bragg peak for every material.

Nuclear models employed in Geant4 are too complex and too slow to be implemented in our speed optimized code. Fippel and Soukup introduced a simplified iterative nuclear model tuned to reproduce Geant4 results. Alternatively, MCsquare directly samples nuclear events from ICRU-63 cross sections. Both solutions have been compared to the Binary Cascade model from Geant4 by simulating 200 MeV protons into water. Depth dose profiles are integrated over $4.1 \times 4.1 \text{ cm}^2$ and $60 \times 60 \text{ cm}^2$ surfaces and compared to Geant4 simulations. As shown in Figure 2.11, MCsquare and Fippel's methods reproduce Geant4 results within 2.0% for the small integration surface. However, Fippel's model is tuned to reproduce the $4.1 \times 4.1 \text{ cm}^2$ integrated results and larger deviations appear when the integration surface is increased. The disagreement between MCsquare and Fippel's results stems from the different sampling of the secondary proton emission angle. The deviations with Geant4 are compatible with the uncertainties in the cross sections, ranging from 5% to 10% according to the ICRU 63 report. MCsquare's method allows the user to easily modify the cross section database.

The deviations are larger for prompt gamma production profiles, which are very sensitive to the nuclear model. Deviations between prompt gamma profiles simulated with Geant4, MCNP6, TALYS and EMPIRE have already been reported by Verburg et al. [83], stressing the need for additional theoretical and experimental studies of nuclear cross sections. Moreover, MCsquare is able to reproduce results achieved with PENH, which has been programmed by different persons but still uses the same nuclear model based on ICRU-63 cross sections. By modifying the nuclear cross-sections, the user would be able to match the results of other Monte Carlo codes.

Some slight differences are also observed in transverse dose profiles, especially in the heterogeneous simulation. This can be explained by the simplified multiple

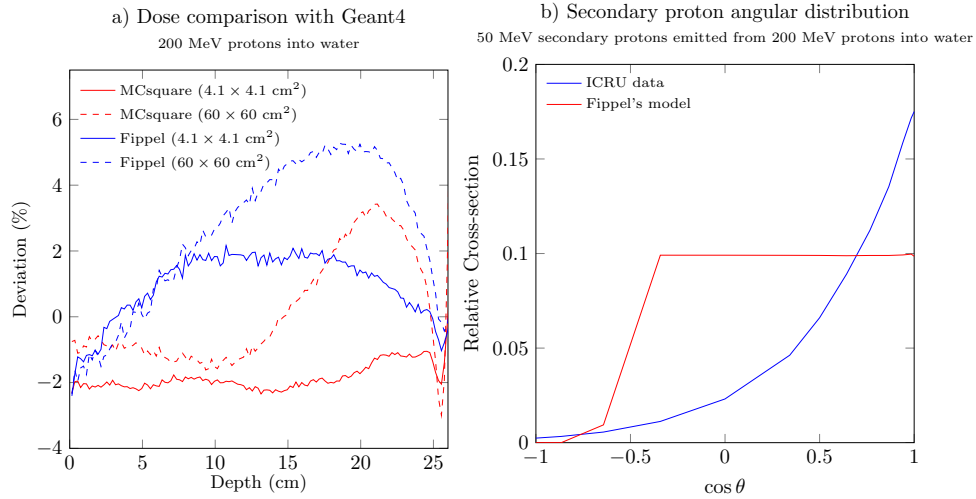


Figure 2.11: Comparison of Fippel's iterative nuclear model and MCsquare method based on ICRU-63 cross sections. a) Depth dose profiles are compared with Geant4 using two integration surfaces: $4.1 \times 4.1 \text{ cm}^2$ and $60 \times 60 \text{ cm}^2$. b) Comparison of emission angle distributions for 50 MeV secondary protons.

Coulomb scattering model, which depends on the ratio of the step length and the radiation length of the material. Geant4 formula includes higher-order terms and a correction factor that depends on the effective atomic number. However, larger deviations between experimental data and various general-purpose Monte Carlo engines have been reported in the literature [107]. The flexible method used in MCsquare to simulate the multiple Coulomb scattering allows the user to adjust the results according to his/her preferences. The lack of electron transport may also affect the dose distribution in low-density materials [53]. This feature will be implemented in future versions of MCsquare.

The modularity of MCsquare allows the user to make his own trade-off between speed and accuracy. The simulation results reveal that the fast configuration may lead to a slight overestimation of the entrance dose compared to the accurate configuration. This is explained by the local absorption of some secondary particles, which should have otherwise left the geometry or deposited their energy further in the phantom. In addition, the step scaling method that is employed to reduce the computation time in the fast configuration decreases the accuracy of the nuclear interaction sampling around interfaces. This approximation does not affect significantly the dosimetric results but may induce some deviations in the prompt gamma production profiles.

The new Intel Xeon Phi architecture has been investigated. This architecture benefits from a large number of cores and wide SIMD units. Since this technology is very close to regular CPUs, any CPU code can be operated on the Xeon Phi after recompiling. However the code has to be well optimized for task and data parallelism in order to benefit from this new technology. Currently, the Gate application does not fully support multithreading. Nevertheless, the last version of Geant4 has been optimized for task parallelism. The toolkit can therefore benefit from the large number of computation cores of the Xeon Phi coprocessor. However these small cores are slower than regular CPU cores. The computation power of the Xeon Phi actually comes from the wide vector units, which Geant4 will not gain much without a proper data parallelism optimization.

MCsquare implementation is optimized to exploit the task and data parallelism of modern computers. The same source code has been compiled for the host CPUs and the coprocessor. Scalability of the performances according to the number of thread was assessed. Speed increases linearly until reaching the number of physical cores. Since these architectures contains multiple threads per core, speed further increases after that point but with a smaller slope. This confirms the usefulness of multithreading to mitigate the memory latency. Despite its larger number of threads and its larger theoretical computation performances, the Monte Carlo simulation is slower on the Xeon Phi than the host CPUs. This slow-down can be explained by the limited memory bandwidth. Although the coprocessor memory bandwidth is larger than for regular CPUs, it is shared between many more computation threads. In this way, the CPU memory bandwidth is 6.4 GB/s per thread, while the Xeon Phi memory bandwidth is limited to 5.3 GB/s per thread. This small difference may not fully explain the slowdown of the Xeon Phi. The amount of cache memory per core is also very important (2.5 MB for the CPU and only 0.5 MB for the Xeon Phi). This memory issue is common for most of computation accelerators and requires specific optimization. We are currently studying various algorithm implementations to fit the memory limitation of the Xeon Phi. After this software optimization the performances of our Monte Carlo code on the coprocessor should be better than on the host CPU. We should also notice that the coprocessor employed in this study belongs to the first generation of a new technology. The main benefits of this type of architecture should be fully harnessed with the second generation, announced for the end of 2015, which let foresee a reoptimised core design with an increase of 300% of computation performance and a larger memory bandwidth [110, 111]. Still, the current version of the Xeon Phi coprocessor could help to increase the computation resources of your workstation at a lower price and energy consumption per TFLOPS than

high-end CPUs.

In spite of the coprocessor memory issues, the benchmarks reveal the capability of MCsquare to compute dose distributions very rapidly. Using the most accurate configuration of MCsquare, only 37 s are necessary to simulate the 10^7 primary protons into average soft tissues. This computation time is limited to 25 s using the fast configuration, which amounts to a gain of 32%. The very short computation time of MCsquare opens new perspectives, like online Monte Carlo simulations in clinical routine or the use of Monte Carlo calculation within the optimization loop of a TPS. Even without dedicated hardware, MCsquare can run simulations in less than 5 minutes on modern laptop computers. It also reveals that by investing as much optimization effort as for GPU programming, it is possible to reach, with CPU architectures, a comparable computation speed.

2.5 Conclusions

A new Monte Carlo code dedicated to proton therapy simulations has been implemented and optimized for modern multi- and many-core CPU architectures. In this sense, the new Intel Xeon processors and Xeon Phi coprocessors have been investigated. Different configurations of MCsquare allow users to reach their own trade-off between accuracy and speed. Simulation results have been successfully validated with Geant4 for homogeneous and heterogeneous geometries. The maximum deviations observed for the integrated depth dose curves are less than 2% - 1mm. These small discrepancies were attributed to the use of different nuclear models. In its fastest configuration, MCsquare takes no more than 25 seconds to simulate 10 million protons in average soft tissue.

Electron transport will be implemented in a future version of MCsquare to improve simulation accuracy in light tissues such as lung. Our benchmarks revealed the memory bandwidth limitation of the Xeon Phi technology. The next generation of Xeon Phi coprocessors and multiple algorithm solutions will be investigated to harness this limitation. MCsquare will be used in future studies, in order to realistically evaluate and optimize the robustness of proton treatment plans. Integration into existing proton therapy platforms is also planned, in order to benefit from their graphical user interface and DICOM support. Moreover the capability of MCsquare to simulate prompt gamma production profiles offers opportunities to verify the range of protons in-vivo. MCsquare accuracy is being validated for patient dose calculation. Its very short computation time would makes Monte Carlo simulations compatible with clinical routine.

Chapter 3

Robust treatment planning

This chapter describes the development of various tools that aim to improve and facilitate the treatment planning process in proton therapy. In the first section, the evaluation and quantification of the tumor motion is investigated. This evaluation can be performed before planning in order to help choosing an appropriate treatment delivery modality and treatment planning strategy. In the second section, the PTV margin concept is extended to proton therapy by considering proton range uncertainties during the generation of the PTV. A beam angle analysis method is then described in Section 3.3. Various metrics are analyzed as a function of the gantry angle in order to select optimal beam angles. Finally, the integration of Monte Carlo dose calculation in a robust optimizer is presented in the last section.

3.1 Motion evaluation

In proton therapy, organ motion, such as breathing, can lead to variations in the proton range and therefore potentially large distortions of the dose distribution. For this reason, the treatment of mobile tumors in proton therapy is still not very widespread and it remains an active research topic. Mobile tumors are often classified according to their motion amplitude. Tumors characterized by a large motion amplitude (larger than 10 mm, for instance) are typically not treated with proton therapy. For such cases, conventional radiotherapy with photons is preferred thanks to better stability of the photon dose distribution. Tumor with a smaller motion amplitude could be treated with proton therapy. The double scattering delivery mode is then preferably chosen. Indeed, the dynamic delivery of the PBS mode may interfere with breathing motion and further distort the dose distribution [69, 70]. The PBS mode is only used for very small tumor motion, with typically less than 5 mm of amplitude. This classification reflects the current clinical experience from the proton center of UPenn and may be different in other institutions. Improving planning tools may enable to treat larger motion cases with PBS proton therapy in the future.

Evaluation of the tumor motion is therefore important to choose the appropriate treatment modality. Motion information is generally provided by the acquisition of a 4DCT. This imaging technique yields a series of 3D images representing the patient anatomy at different phases of the breathing cycle. The motion amplitude is usually extracted by visually comparing the tumor position at the end-inhale and end-exhale phases. This manual operation is not practical and not very accurate.

To improve the evaluation of the tumor motion, an automatic method is proposed. The method is based on a deformable registrations algorithm [112], which

tries to align two images iteratively. As illustrated in Figure 3.1, a deformation vector field is therefore optimized to reproduce as well as possible the reference image by deforming the other. In the motion quantification method proposed here, the first phase of the 4DCT is arbitrarily chosen to be registered to all other phases using the deformable registration algorithm available in OpenReggui [18]. Once the deformation fields between phase 1 and all other phases are computed, they can be used to calculate motion statistics or deform other data. This registration operation is the basis of all other motion related tools presented in this thesis. The results computed at this stage could therefore be reused multiple times during the treatment planning.

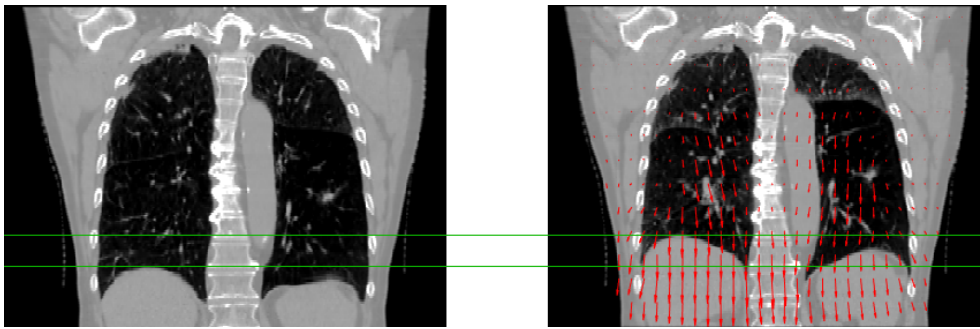


Figure 3.1: Illustration of two 4DCT phases representing the end-inhale (left) and end-exhale (right) of the breathing cycle. The deformation field illustrated with red arrows was computed with a deformable registration algorithm to reproduce the left image when applied to the right image.

To take into account the tumor motion during the treatment planning, the contour of the tumor is usually drawn on every 4DCT phases. This manual operation consumes much time. The deformation fields computed with the registration algorithm can be used to automatically deform the tumor contour drawn on the reference phase to all other phases, saving a considerable time for the physician. The tumor motion can then be evaluated by comparing the position of all tumor contours. Especially, the center of mass is calculated for each contour. The motion amplitude is therefore estimated as the largest difference between all centers of mass.

As patient tissues are not very rigid, a single number may not be sufficient to fully characterize the motion. In reality, motion of the tumor or organs is best described by a distribution of displacement vectors inside its volume. This is exactly what is calculated with deformable registration. The motion amplitude can

therefore be extracted for each voxel of the image using the pre-computed deformation fields. These fields only provide the displacement between phase 1 and all other phases. However, by combining multiple fields, it is possible to reconstruct the displacement fields between any other pairs of phases. The field showing the maximum displacement is then used to compute the motion amplitude distribution inside the target or organ volumes. Some statistics such as the median value, or other percentiles, can then be calculated to characterize the motion as illustrated in Figure 3.2.

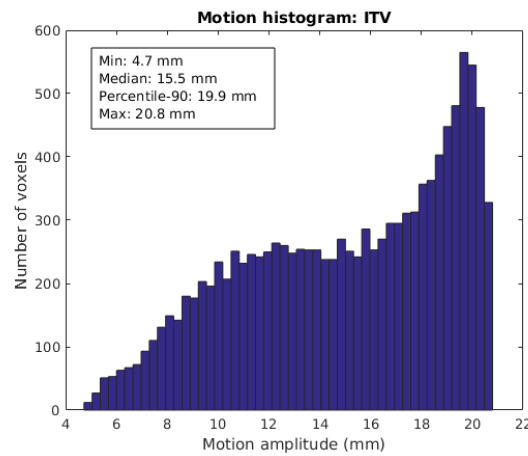


Figure 3.2: Histogram representing the distribution of the motion amplitude calculated using the deformation fields inside the ITV volume.

3.2 Beam-specific PTV

PTV margins are normally used in conventional radiotherapy with photons to make the plan more robust to treatment uncertainties. By enlarging the irradiated volume, it ensures coverage of the CTV even in presence of small setup errors, intra- and inter-fraction tumor motion, etc ... The PTV concept is based on the static dose cloud approximation. Therefore, the PTV margin is only valid under the assumption that the dose distribution is not deformed by treatment errors and is only rigidly shifted with respect to the tumor position. This assumption is not always met in proton therapy due to the finite range of protons and the high dose gradient of the Bragg peak. Indeed, the dose distribution may be substantially deformed in case of range variation. Robust optimization is considered as a better alternative method [12, 14, 56, 57, 58] for proton therapy. However, its implementation

in the clinic is more complex. It requires more computation, more evaluation of the resulting plan, and hence slows down the treatment planning process. For these reasons, the PTV concept is still often used in clinics for proton therapy, despite its limitations. This section aims at adapting the PTV margin to take into account the proton range uncertainty. The proposed method is probably not the best solution for proton therapy planning, but it may represent a temporary compromise for clinics that are not ready yet to move to alternative tools.

The deformation of the dose distribution by the variation of proton range is an unavoidable effect in proton therapy. In order to use PTV margins, the range variation should be compensated to ensure that the dose distribution remains uniform inside the target. This can be achieved by extending the PTV margin along the beam direction, leading to the definition of a beam-specific PTV (BSPTV). The BSPTV planning strategy consists therefore, first, to compute a specific margin for each considered beam direction. Then, the treatment plan is optimized so that each individual beam cover its own BSPTV. Moreover, to ensure that the dose distribution remains uniform inside the CTV, even in presence of range errors, each BSPTV must be covered with a uniform dose. This is also called single field uniform dose (SFUD) planning.

A few methods were already proposed in the literature to generate the BSPTV [113, 114, 115]. The method we propose here has been implemented in the open source application OpenReggui [18]. The BSPTV margin is constructed in two parts, namely, the lateral part that is transverse to the beam direction and the longitudinal part that is along the beam direction. The lateral part is equivalent to the original PTV margin and can be computed using the regular recipes proposed by Van Herk. This lateral margin is used to expand the CTV for static tumors, or the ITV for mobile tumors. The ITV can be generated by deforming the CTV drawn on the reference phase to all other phases of the 4DCT using the deformation fields generated at the motion evaluation step (see previous section). The union of all deformed CTV forms the ITV. The longitudinal part of the BSPTV margin can then be further subdivided in a proximal side and a distal side, with respect to the beam source. Both sides of the longitudinal margin need to be calculated to compensate for the potential range variation caused by setup errors, breathing motion, and other treatment uncertainties. Each of these three contributions are computed separately.

The physical proton range to reach a given voxel of the CT image is obtained by measuring the distance between this voxel and the edge of the image along the beam direction. To compute the proton energy necessary to reach this voxel,

the TPS needs to convert this physical range into a water-equivalent path length (WEPL). The WEPL can then be converted in proton energy using a look-up table as provided by the NIST [116], or using a fit of this table. Due to treatment uncertainties, the WEPL corresponding to this physical distance inside the patient may vary, inducing a variation in the physical range too. To compare the effect of various situations on the range, all results will therefore be computed in WEPL, which is the reference for TPS calculation.

The contribution of setup errors is computed by calculating, in each voxel of the target, what would be the WEPL variation if the entire CT image is shifted. The amplitude of the shift is provided by the user and is generally estimated based on population-based statistics. The planning CT image, typically the average CT for mobile tumors, is first converted in water-equivalent quantities, using the HU-to-SPR calibration curve. The WEPL corresponding to a given voxel is computed with a ray-tracing algorithm. To do so, the distance traveled in each crossed voxel is weighted by the voxel SPR and accumulated. As illustrated in Figure 3.3, this calculation is repeated for each voxel of the patient image to obtain a map of the WEPL, which is specific to the beam angle. The WEPL change that stems from a shift in the patient position can then be retrieved by shifting the WEPL map. Therefore, the WEPL variation for a given voxel of the target could be obtained by comparing the nominal WEPL of this voxel with the WEPL of its neighboring voxels. In practice, the morphological operations of dilation and erosion are applied to the WEPL map to provide respectively a map of the maximum and minimum WEPL found in the neighborhood of each voxel. The neighborhood is defined with a structuring element, which, in this case, is a disk perpendicular to the beam direction, with a radius equals to the maximum setup error. The contribution of setup errors to the proximal margin is therefore calculated in each voxel of the target with:

$$\text{Proximal margin}_{\text{setup}} = WEPL_{\text{nominal}} - WEPL_{\text{min}}$$

Similarly, the distal margin is calculated with:

$$\text{Distal margin}_{\text{setup}} = WEPL_{\text{max}} - WEPL_{\text{nominal}}$$

The WEPL variation caused by the breathing motion is extracted from the 4DCT. As for the calculation of the effect of setup errors, each phase of the 4DCT is converted in SPR and then in a WEPL map. The range variation can therefore be obtained by comparing the nominal WEPL values of the planning CT with the WEPL extracted from each phase of the 4DCT. The contribution of motion to the proximal and distal margins is calculated in each voxel of the

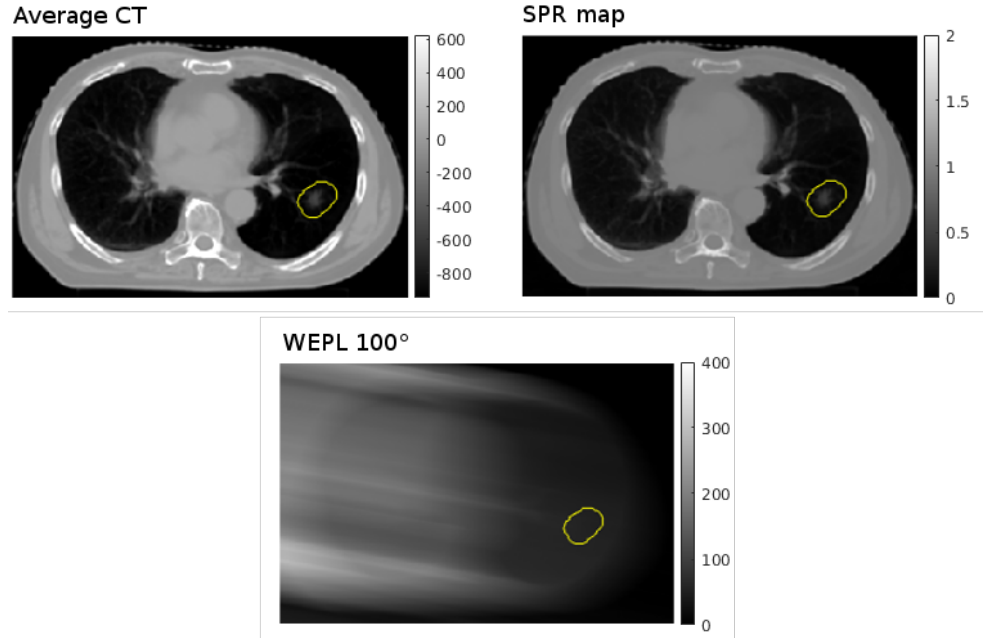


Figure 3.3: Example of average CT in Hounsfield units (top left). The CT image is converted in a SPR map (top right) using the CT calibration curve. A ray-tracing algorithm then computes the map of WEPL in mm (bottom) for a gantry angle of 100° . The ITV contour is also reported in yellow.

expanded ITV with:

$$\text{Proximal margin}_{\text{motion}} = WEPL_{\text{nominal}} - \min(WEPL_{\text{motion}})$$

$$\text{Distal margin}_{\text{motion}} = \max(WEPL_{\text{motion}}) - WEPL_{\text{nominal}}$$

A flat range uncertainty $R_{\%}$ of typically 3% of the nominal range is finally added in order to consider other sources of range variation, i.e. mainly the HU conversion errors. The contribution to both proximal and distal margins is given by:

$$\text{Margin}_{\text{others}} = WEPL_{\text{nominal}} \times R_{\%}$$

All these 3 contributions have then to be combined to provide the total proximal and distal margins. In the current implementation, the contributions are summed quadratically, assuming that each contribution can be considered as statistically independent. However, this approximation is biased. Moreover, the setup contribution was computed on the average CT, which is not a physical representation of the patient. To solve both issues, one possible improvement of our method would

be to compute the setup contribution on the 4DCT phases during the calculation of the motion contribution. This way, both setup and motion effects would be combined correctly and only the remaining flat range uncertainty would need to be added.

Finally, because all margin values are computed in WEPL, they have then be converted back to physical distances by ray-tracing the WEPL margin along the beam direction in the patient SPR image. The whole process is repeated for each beam angle. The resulting margins are illustrated in Figure 3.4 for a lung and a liver patient. Due to the low density of lung tissues, the margin gets very large after conversion from WEPL to physical lengths in the patient tissues.

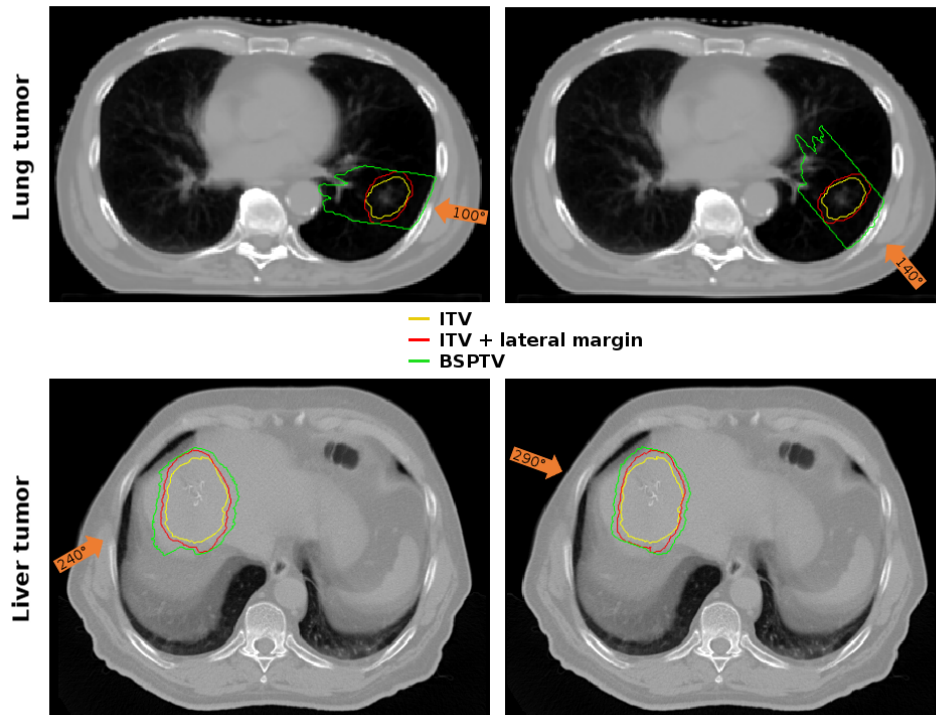


Figure 3.4: Illustration of target margins for a lung tumor (beam angles of 100° and 140°) and for a liver tumor (beam angles of 240° and 290°). The ITV (yellow contours) is first expanded laterally (red contours), then it is expanded along the beam direction to generate the BSPTV (green contours). The orange arrows represent the proton beam directions.

The concept of BSPTV is only valid for uniform dose distributions. Indeed, it is designed to restore the approximation of a static dose cloud inside the target. Therefore, the BSPTV margin can only be used for SFUD planning. The capability of PBS to deliver intensity modulated proton therapy (IMPT) cannot be fully exploited with the proposed margin solution. This is the reason why margins can only be a provisional tradeoff between improved accuracy, ease of implementation, and continuity in the current widespread use of margins in clinical practice. Robust optimization offers the potential to reach much better dose conformity by exploiting the flexibility of IMPT delivery.

3.3 Beam angle selection

A treatment plan is typically composed of two or more beams with different directions. There are some recommendations for the selection of beam angles. Obviously, organs at risk (OARs) should be avoided when possible. In proton therapy, it is also recommended not to use the finite range of proton to spare an OAR located next to the tumor. Indeed, in case of over-range caused by treatment uncertainties, Bragg peaks would move inside the organ. The beam angles are parameters that are generally not optimized by the TPS and require manual adjustment based on the planner's experience. Therefore, it is not always straightforward to find the optimal beam angles that minimize the irradiated volume and limit range sensitivity to uncertainties. This is especially true in presence of organ motion. In this section, we describe a method to assist the planner in the selection of the optimal beam angles by analyzing the variation of several metrics with respect to the gantry angle. Illustrative results will be reported for the lung patient shown in Figure 3.3.

For mobile tumors, some beam directions may lead to reduced sensitivity of the proton range to motion effects. This can be analyzed by computing the WEPL variation across all phases of the 4DCT inside the target. As for the motion analysis presented in Section 3.1, this method provides a distribution of the motion effect for all voxels of the target. Several statistics can then be computed to characterize this distribution. The average and the 95th percentile of the WEPL variation are computed as a function of the gantry angle and are reported in Figure 3.5. The sensitivity of the range to motion effects will be lower for the beam angles corresponding to a small value on this plot.

For treatment planning with BSPTV, the volume receiving the therapeutic dose can be estimated by computing the BSPTV volume. The analysis of the BSPTV

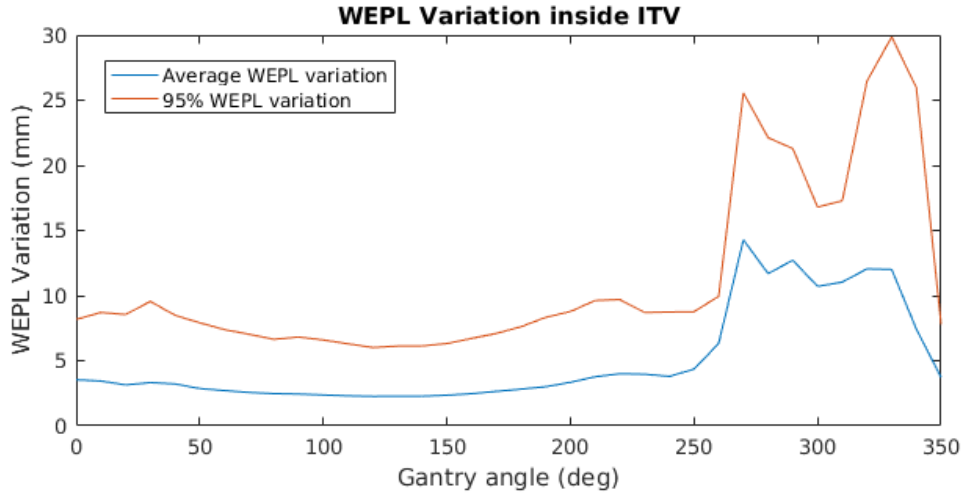


Figure 3.5: Analysis of the WEPL variation caused by the breathing motion as a function of the gantry angle. The blue curve represents the average, and the red curve represents the 95th percentile of the distribution of WEPL variations inside the ITV.

volume and each of the 3 individual contributions (setup errors, motion, and flat 3% range uncertainty) with respect to the gantry angle is presented in Figure 3.6. The intersection of organ structures with the BSPTV could also be useful information to optimize OAR sparing. This is illustrated in Figure 3.7 for various OARs. Similarly, the intersection of OARs with the beam-path is reported in Figure 3.8. For a given gantry angle, the volume of the beam-path is calculated by projecting the BSPTV contour along the beam direction to the edge of the CT image.

Calculating the dose distribution for all these angles would be time consuming. The proposed method only manipulate WEPLs, which is much faster than a full dose calculation. Therefore, our method provides useful information to the planner in order to select the optimal beam angles in a reasonable time. This analysis could also be used as prior information for better beam angle optimization based on a dose engine. The current implementation only supports co-planar beam directions, with a variable gantry angle and a couch angle fixed at 0° . However, the tool could be adapted to consider other couch angles too, offering a 4π angular analysis.

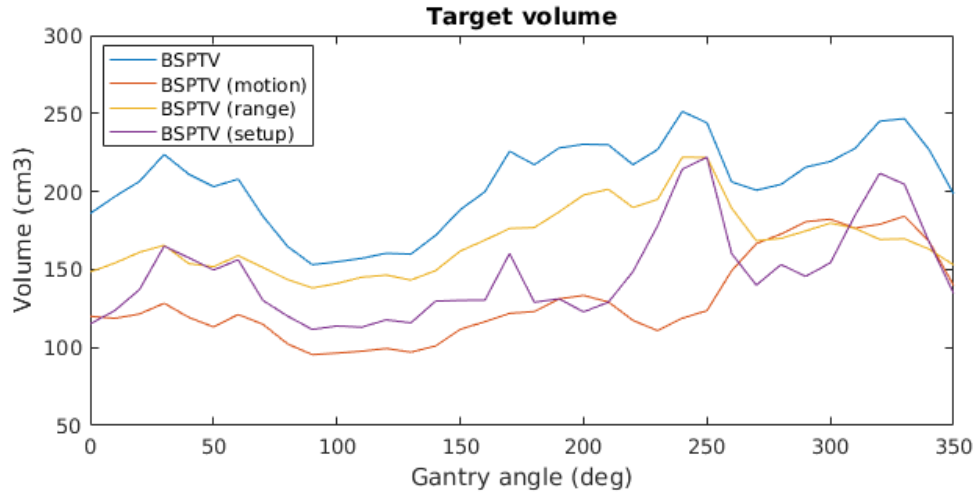


Figure 3.6: Variation of the BSPTV volume as a function of the gantry angle. The red, yellow, and purple curves represent the BSPTV margins for the 3 individual contributions only (motion, flat range error, and setup errors, respectively), and the blue line is the full BSPTV calculated with a quadratic sum of the 3 contributions.

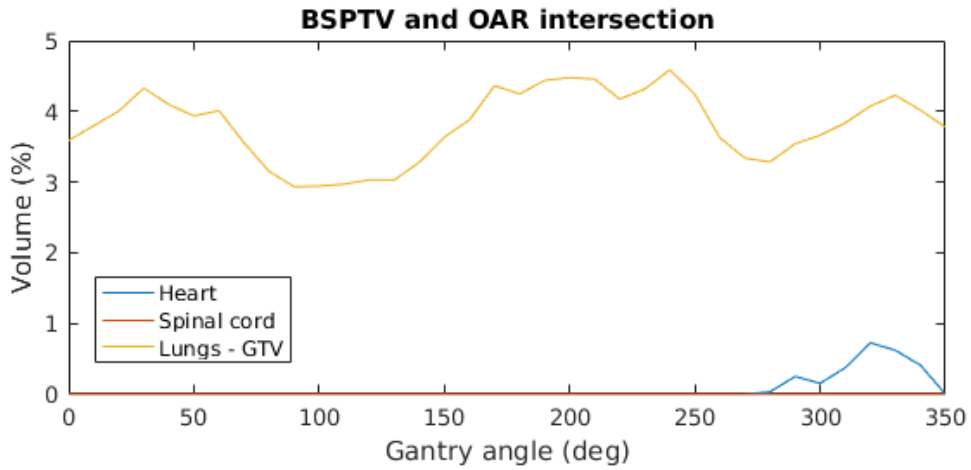


Figure 3.7: Analysis of the volume of organs at risk overlapped by the BSPTV as a function of the gantry angle.

3.4 Robust optimization

Robust optimization is a treatment planning method that integrates uncertainties directly in the optimization process. The treatment plan is optimized to reach the

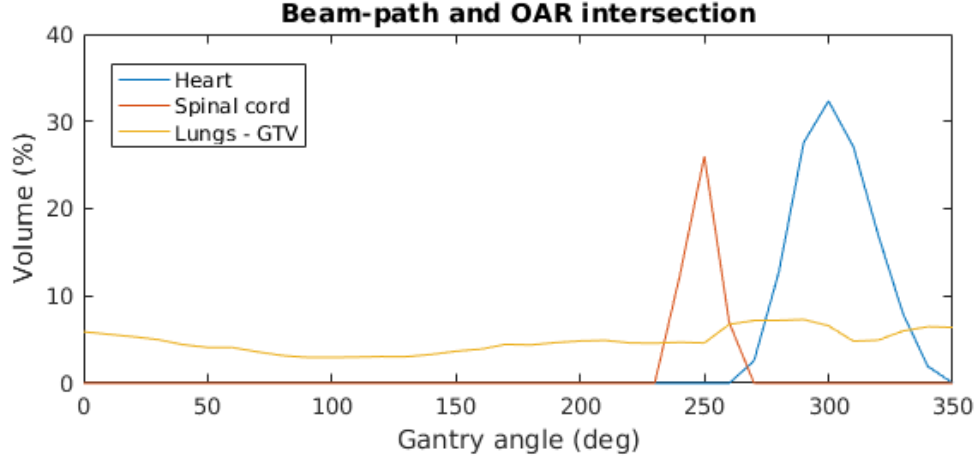


Figure 3.8: Analysis of the volume of organs at risk overlapped by the beam path as a function of the gantry angle. According to this plot, optimal beam angles are between 100° and 140°

clinical objectives for various scenarios of treatment uncertainties. This method allows for taking into account treatment errors, including range uncertainties, without any preliminary hypothesis on the dose distribution. In contrast to the BSPTV method, robust optimization does not require SFUD and enables the optimization of robust IMPT treatment plans.

MIROpt [17], the robust optimization algorithm developed in this project was the object of the PhD thesis of Ana Barragan. My contribution in this project was mainly the adaptation of MCsquare [15] to provide fast Monte Carlo simulations, 4D dose calculation, and various models of treatment uncertainties to the optimizer.

The algorithm presented here is dedicated to the optimization of proton PBS treatment plans. The proton beam is therefore scanned over the target, delivering the dose spot by spot. Each spot can be fully characterized by a set of 5 parameters: the beam direction, the proton energy, the intensity, and the (x,y) coordinates of the spot position in the beam-eye-view referential. Multiple spot layers corresponding to different proton energies are generally required to cover the entire target thickness. Before optimization, the layer positions are determined in depth for each beam with a constant spacing in the WEPL space in order to cover all target depths. These WEPLs are calculated with a ray-tracing algorithm, like

in the BSPTV method. Once each layer WEPL is determined, the corresponding proton energy can be calculated using a look-up table or a fitted formula. For each of these layers, the spot positions are also determined before optimization. The spots are positioned in a hexagonal grid to cover geometrically the target cross-section. The x and y coordinates are then calculated for each spot in the beam-eye-view referential.

The only remaining variable parameter is the spot intensity, which will be optimized by the algorithm to provide homogeneous dose coverage of the target, while sparing OARs. In order to optimize the intensity of each spot, the algorithm needs to know the contribution of each individual spot to the dose distribution. Therefore, a dose engine computes the dose distribution for each spot separately based on the pre-determined parameters (beam angle, energy, spot coordinates) and a unitary intensity. The resulting dose distributions are often called beamlets and form a dictionary of elementary functions that the optimizer can modulate and combine linearly. The optimizer can then find what are the optimal modulation weights for each beamlet, so that, after summation of all weighted beamlets, the total dose distribution meets the clinical objectives. The weights resulting from the optimization process provide the intensity of each spot, i.e. the number of protons, which then finalizes the parameterization of the treatment plan.

In our implementation, MCsquare is used to compute the dose distributions. The Monte Carlo simulation is therefore repeated for each individual spot. When beamlets are calculated with a Monte Carlo dose engine, it is important to simulate a sufficient number of particles to achieve a reasonable statistical noise level. Indeed, noisy beamlets may introduce a bias in the optimized plan. Therefore, each beamlet is typically simulated with 10^5 primary protons. Assuming that the treatment plan is composed of several thousands of spots, this leads to a much larger total number of simulated particles than for a regular forward dose calculation. This can be explained by the dose overlap of neighboring beamlets. For a full plan simulation, the dose reported in each voxel combines therefore the contribution of multiple beamlets, decreasing the number of simulated particles required per beamlet. As the computation time grows linearly with the number of particles, the simulation of each individual beamlet is therefore relatively slow and may represent the bottleneck of Monte Carlo TPS.

When a regular treatment plan optimizer is used, motion effects have to be manually taken into account in planning. Typically, an ITV margin is employed to ensure that the tumor remains inside the proton beam along its motion. Further adaptation, such as the BSPTV, should also be considered to compensate

for the motion-induced range variation. The use of a 4D optimizer would enable implicit modeling of the motion in the treatment plan optimization. Motion effect can be simulated based on the information of 4DCT images using a 4D dose calculation algorithm. This algorithm will be explained in Chapter 4. Briefly put, the dose distribution is computed on each phase of the 4DCT and accumulated on a reference phase. This results in a 3D distribution that represents the dose delivered into the mobile anatomy, typically a dose distribution blurred by motion. Therefore, motion information can be directly included in the beamlets, enabling 4D dose optimization with only few modifications in the optimization algorithm itself. In this project, the 4D dose calculation of MCsquare is used to compute each individual 4D beamlet.

In addition to motion, other treatment uncertainties are modeled in MCsquare, allowing for the optimization of the treatment plan robustness against various uncertainties. Several uncertainty models will be described extensively in Chapter 4. However, only setup errors and a flat range uncertainty are currently implemented in our robust optimizer. Setup errors can be simulated by shifting the patient image during the dose calculation. In MCsquare, setup errors are equivalently modeled by shifting the initial position of each particles in the opposite direction. A flat range uncertainty is also simulated with a global scaling of the stopping powers.

Treatment uncertainties can be classified in either systematic errors that remain the same for all delivery sessions, or random errors that vary from one fraction to the other. For a sufficient number of fractions, the effect of random errors can be seen as a blurring of the dose distribution. For instance, random setup errors can be approximated by randomly sampling a new position shift for every particle. The induced blurring effect is therefore directly reproduced in each beamlet. In contrast, systematic errors, such as systematic setup errors and range uncertainties, may lead to more important distortions of the dose distribution. The direction and amplitude of the error are not known a priori. Therefore, multiple scenarios of systematic errors are considered. All beamlets of the treatment plan are simulated for each scenario. In each step of the iterative optimization process, the algorithm is trying to optimize the treatment plan to improve the worst-case scenario [14, 57]. This way, the optimizer should find the optimal treatment plan parameters that satisfy the objectives for all scenarios.

Robust optimization enables more accurate modeling of treatment uncertainties and better flexibility in the treatment plan optimization. The resulting plan should be robust to the level of uncertainties considered during the optimization. The

use of Monte Carlo dose calculation increases the confidence in the robustness of the resulting plan. However, due to the computation limitations of such complex optimization problem, modeling of treatment uncertainties is still limited to a few scenarios of setup errors and range uncertainties. A more comprehensive and realistic verification of the robustness should therefore be performed to validate the plan. Such accurate and realistic robustness verification methods are presented in Chapter 4.

Chapter 4

Realistic verification of the treatment plan robustness

Various robust planning strategies are used to take into account treatment uncertainties during the generation of the plan. To keep the computation time practical, fast analytical dose calculation algorithms are typically used for plan optimization. These approximate algorithms also introduce some uncertainties that need to be taken into account during planning. Moreover, the models available in commercial TPS to compensate for treatment uncertainties are often limited to setup errors and a flat range uncertainty. Due to these limitations, the robustness of the resulting plan against actual treatment errors is not guaranteed. The treatment plan must therefore be evaluated properly before treating the patient. Especially, the impact of treatment uncertainties on the dose distribution should be verified.

A first study was conducted at the University of Pennsylvania to assess the impact of motion effects on the delivered dose for treatments of liver tumors with PBS proton therapy [69]. Organ motion and the interplay effect with PBS delivery were simulated. The dose calculation was performed by the TPS (Eclipse). The 4D dose distributions were then accumulated using the open source software OpenReggui [18], which provides image processing tools for radiation therapy applications. The implemented interplay verification tool allowed investigating the use of an abdominal compression belt for motion reduction during the treatment. However, we also found several limitations in its implementation. In particular, the dose calculation provided by the TPS does not yield an independent verification of the plan. Moreover, the use of various pieces of software is not practical. It involves multiple manual operations, which consume time and are prone to human errors. Finally, the developed tool only provides an evaluation of motion effects. All other treatment uncertainties still have to be evaluated separately.

To address these issues, a second robustness verification tool was developed, based on MCsquare [16], which was presented in chapter 2. The use of our fast Monte Carlo code provides an independent and accurate dose calculation method. Multiple uncertainty models were implemented directly in MCsquare in order to speed up the calculation and facilitate the robustness test procedure. The simulation of motion effects was combined with other uncertainties such as setup and range errors. This development was the object of a paper currently in revision in the Medical Physics journal. This paper is included in the following part of this chapter and is entitled:

Monte Carlo methods to realistically evaluate the robustness of proton therapy plans

K. Souris, A. Barragan, G. Janssens, D. Di Perri, E. Sterpin, J.A. Lee

4.1 Introduction

Proton therapy can potentially offer better dose conformity to the target and a reduced integral dose, compared with conventional radiotherapy. However, the high dose gradient in the Bragg peak also increases the sensitivity of the dose distribution to treatment uncertainties. Especially, the actual proton range in the patient may vary from the range calculated during treatment planning due to the inaccuracies in the treatment workflow [8], which include limited information conveyed by imaging, setup errors, image conversion into stopping powers, simplified dose calculation algorithms, inter- and intra-fraction motion... Range uncertainties hinder harnessing the full potential of the steep distal dose falloff to spare organs-at-risk and represent currently a major limitation for clinical indications of proton therapy.

The accuracy of the analytical dose calculation typically employed by the treatment planning system (TPS) deteriorates in highly heterogeneous geometries, such as lung tissues. In contrast, Monte Carlo dose calculation yields better accuracy thanks to realistic simulation of actual particle interactions using direct random sampling of the physics laws underlying particle transport. However, the gain in accuracy comes at the expense of nonzero precision: due to the stochastic nature of Monte Carlo process and limited number of simulated particles. A statistical uncertainty affects the resulting dose distributions and makes them noisy. A large number of particles must be simulated to reach an acceptable noise level, leading to long computation times that have historically prevented the use of Monte Carlo methods in clinical routine. Recently, several fast Monte Carlo codes have been developed to address this issue [75, 16] and some commercial TPSs feature a Monte Carlo dose engine performing dose computations at practical speed.

Motion is another challenge in proton therapy. In addition to the geometric effect related to target motion, the coverage may also be altered by the induced proton range variations. Moreover, dynamic delivery techniques, such as proton pencil beam scanning (PBS), may interfere with organ motion and distort the dose distributions, which is often referred to as interplay effect.

A treatment plan is considered robust to a given level of uncertainty when the resulting dose distribution still satisfies the clinical goals, even in presence of that uncertainty. Multiple methods have been developed to improve robustness of proton therapy. Safety margins recipes defining a Planned Target Volume (PTV) have been historically used in photon therapy to ensure target coverage in the

presence of a diverse set of geometrical uncertainties, namely setup errors, inter-fraction and intra-fraction motions (including breathing). The validity of these margin recipes relies on the static dose cloud approximation that is violated in proton therapy because of range uncertainties. Adaptation to the PTV has been suggested by integrating range uncertainties along the beam direction within the margin, leading to the so-called beam-specific PTV [115, 113, 114]. However, the latter gets overall inadequate in the context of multi-field intensity modulated proton therapy (IMPT), due to possibly high in-field dose gradients. On the contrary, robust optimization can ensure plan robustness for multi-field IMPT by including uncertainties in the definition of the objective function [58, 57, 56]. However, most proposed robust optimizers in the literature or in commercial software are limited to a few scenarios of systematic setup errors, a flat range uncertainty percentage, and a model of organ motion represented by a single 4D-CT. For the latter, it is important to mention that only robustness against breathing motion is ensured because the dynamics of the treatment is disregarded during optimization, thus neglecting the interplay effect. Moreover, the use of a single 4D-CT has well known limitations. This snapshot only provides an estimation of the motion at the time it was acquired. However, breathing motion may vary from one fraction to another. As this motion variation is not taken into account during planning, its effect on the delivered dose is unknown.

All these planning approximations and uncertainties in treatment delivery, and their limited integration in current robust optimizers, motivates the development of a tool to verify robustness beforehand in a comprehensive fashion. Such a robustness test should rely on realistic simulations of the entire space of relevant uncertainties. This is not the case for the uncertainty models provided by commercial applications, which are generally limited to systematic setup errors, a flat relative range error, and 4D dose calculation with incomplete or non-existent assessment of the interplay effect. Simulated setup and range errors are usually set at their extreme values for which robustness is sought. However, verifying the plan robustness for these extreme error values does not guarantee the plan robustness for intermediate errors. Thus, intermediate values of the errors should be simulated as well. Moreover, natural variations of the breathing motion should be incorporated together with a quantified assessment of the interplay effect. Effects of interplay have already been studied in the past [69, 70], but not in combination with all other treatment uncertainties. To assess the combined effect of these errors, all treatment uncertainty models should be mixed in a combinatorial and statistically consistent way. As a consequence, the number of scenarios to be simulated grows exponentially with the number of distinct uncertainties that are considered. For instance, by considering 3 error values (minimum, null, and

maximum), for n uncertainties, the total number of scenario to be simulated is 3^n . To be practical, such global assessment of the robustness would therefore require an adequate sampling of the space of the possible scenarios, which is not trivial.

In this work, we developed a comprehensive and accurate robustness test for proton therapy plans. Dose distributions are calculated for multiple uncertainty scenarios using MCsquare [15, 16], a fast Monte Carlo dose engine. Monte Carlo dose calculation improves dose computation accuracy in heterogeneous geometries. It also serves as an independent dose engine to verify the TPS results. Most clinically relevant uncertainties are modeled and implemented directly in the Monte Carlo code. Patient and delivery motions are simulated. Variations in breathing amplitude and period are taken into account in order to provide a more realistic treatment simulation. A Monte Carlo approach is also adopted to select and analyze the uncertainty scenarios. The resulting dose distributions are combined and reported in DVH-bands.

4.2 Methods

To assess the robustness of a treatment plan, the dose distribution is computed for multiple scenarios of treatment uncertainties. The dose calculation is performed by MCsquare [16], a fast Monte Carlo code presented in the following first subsection. As setup and range errors are already discussed in many publications [58], they will be only covered briefly in sections 4.2.2 and 4.2.3 respectively. Simulation of motion effects will then be described in more details (section 4.2.4), including the variation of breathing amplitude (section 4.2.5) and period (section 4.2.6). All error models are then combined to simulate realistic treatment realizations. A method of probabilistic scenario selection is presented in section 4.2.7. All types of uncertainties are characterized by a probability distribution and sampled individually for each scenario simulation. A Monte Carlo approach is employed to analyze the plan robustness in section 4.2.8. Lastly, the proposed solution is illustrated in section 4.2.9 with a clinical case.

4.2.1 Monte Carlo dose calculation

MCsquare [16] is a multi-purpose Monte Carlo code designed to offer high computation speed and accuracy. The code was benchmarked against Geant4. It is optimized to simulate proton PBS treatments in voxelized geometries. It was implemented to exploit both task and data parallelism of modern processors, ac-

celerating computation significantly. Simulation of electromagnetic interactions is based on a database of stopping powers, the formula of ionization cross section, and the multiple scattering model that was initially proposed by Rossi and Greisen [94]. Nuclear interactions are sampled from the ICRU 63 cross section database [96].

In order to enable realistic simulations, MCsquare was commissioned for a clinical beam line [117]. The transport algorithm was validated with a Fano test [109] and measurements in a heterogeneous phantom designed to highlight the multiple scattering process [51]. The beam model was further validated for clinical applications by comparing patient simulations calculated with MCsquare and Topas [105], which relies on the Monte Carlo code Geant4 [88].

Most TPSs usually compute and report the dose in water. The clinical experience is also based on dose prescriptions reported in water. In contrast, Monte Carlo dose engines calculate the dose with respect to tissue compositions. Indeed, a density and a material composition are assigned to each voxel of the patient CT image using the Schneider method [49]. To compare Monte Carlo results with the TPS and prescribed doses, the dose to medium is converted on the fly to dose to water using Paganetti's technique [118]. This conversion is directly implemented in the Monte Carlo code in order to take into account the energy and particle dependency of relative stopping powers.

In this study, the batch method was used to quantify the statistical precision of Monte Carlo results. The simulation is divided in at least 20 batches that provide each a 3D dose map. The final dose distribution is obtained by computing the mean value $\bar{x}$ in each voxel. Statistical noise on the mean value is then estimated by calculating the standard error ($SE_{\bar{x}}$) in each voxel. To do so, the standard deviation σ is computed across the 20 batches and then divided by the square root of the number $n = 20$ of batches.

$$SE_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

4.2.2 Setup errors

Setup errors can be simulated by translating the patient image before the Monte Carlo simulation. It is equivalent to translating the particle position in the opposite direction. In MCsquare, setup errors are modeled by translating particles at the beginning of the simulation, just after the sampling of particle initial states from the beam model.

4.2.3 Range errors

Range errors resulting from image conversion to stopping powers and other effects¹ are modeled in MCsquare by scaling the mass densities of the CT image for each scenario, before running the Monte Carlo simulation. This strategy is equivalent to that proposed in many publications⁵ and some commercial TPSs.

4.2.4 Organ motion

Breathing motion is typically represented by a 4DCT, composed of multiple 3D images of the patient at various phases of the motion cycle. The impact of the moving anatomy on dose distributions can be studied with a 4D dose calculation algorithm, which computes the doses for the multiple phases of the motion cycle and accumulates them at a reference position.

Prior to the 4D Monte Carlo dose calculation, a diffeomorphic deformable registration algorithm [112], called Morphon and implemented in OpenReggui [18], was used to register all 4DCT phases to the reference phase. One velocity field per phase is thereby pre-generated. The concept of velocity field facilitates their manipulation. Unlike the commonly used motion fields, velocity fields can be easily inverted, scaled, or combined, while maintaining physical consistency of deformations. Especially, rotations are preserved. After manipulation, velocity fields can be converted to regular motion fields for image deformation purpose as explained in Janssens et al [112].

After registration, velocity fields are imported in MCsquare for the 4D Monte Carlo calculation. The treatment plan is simulated on each phase of the 4DCT. As the image quality of the 4DCT is sometimes inadequate for dose calculation in proton therapy, a denoised 4DCT can be reconstructed as described in the next sub-section. The computed dose distributions are then deformed to the reference phase using the pre-generated velocity fields. Lastly, deformed doses are averaged in a 3D image representing the dose absorbed by the moving tissues.

4.2.5 Motion amplitude errors

Intra- and inter-fraction variations of the breathing motion have been reported in the literature [119, 120]. This sub-section describes how the organ motion

amplitude can be increased or decreased by manipulating the velocity fields. As illustrated in Figure 4.1, a new 4DCT series, with scaled motion, can be created by deforming the reference phase. This process starts by registering all phases of the original 4DCT as for the nominal organ motion simulation. In the case illustrated by Figure 4.1, all phases are registered to phase 1. Intermediate or fictitious phases can then be reconstructed from the original 4DCT. Here, the reference phase is set at the Mid-Position, which is defined by the average of all velocity fields. The first 4DCT phase can therefore be deformed to the Mid-Position using this average velocity field, represented by the vector from P1 to Mid-Position on Figure 4.1b. The velocity fields to deform other phases (P2, P3, ...) are generated by subtracting the original phase field from the average field. To generate the Mid-Position CT image (MidP-CT) [44], all 4DCT phases are deformed to the Mid-Position and the median HU value is computed in each voxel. This results in a CT image with less noise and fewer artifacts than the nominal 4D phases. Using the velocity field properties, the deformation can easily be inverted by taking the opposite of the velocity field. An improved 4DCT can therefore be reconstructed by deforming the newly created MidP-CT image back to each phase position using the inverted velocity fields. Moreover, these inverted velocity fields can also be scaled by multiplying them by a scalar factor. As illustrated in Figure 4.1d, the scaled fields can be used to increase or decrease the motion amplitude.

The model was validated using two 4DCT series that were acquired on the same lung patient over a one-week interval. A MidP-CT was reconstructed for both series and employed to contour the GTV. The GTV contours were then deformed to each phase of their respective 4DCT. A new 4DCT series, with improved image quality, was also generated by deforming the reference image (MidP-CT) back to each phase position. As a first validation test, deformed contours and images were visually inspected and compared to the original 4DCT images to assess the accuracy of the deformable registration. The motion amplitudes were evaluated by computing the 3D displacement of the GTV center of mass. The breathing amplitude variation between both 4DCT series was calculated with the ratio of their maximum GTV displacement. The proposed model was then used to reconstruct the second 4DCT from the first week images. To do so, the first week velocity fields were scaled with the breathing amplitude variation factor measured between both acquisitions. The first week MidP-CT is then deformed to generate the new 4DCT, with scaled motion. The GTV displacement of the generated 4DCT is compared with the actual second 4DCT series to validate the model. For this comparison, to remove any possible setup differences and baseline shifts, the images of week 1 and week 2 were rigidly registered to align the GTV position in the Mid-Position.

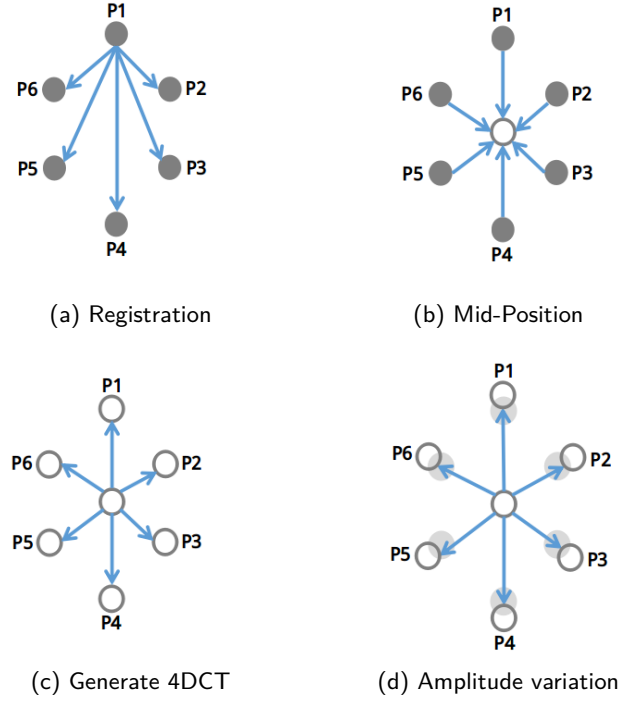


Figure 4.1: Reconstruction of 4DCT with scaled motion amplitude. Tumor position at each phase (P1 to P6) is represented by the circles. Velocity fields are illustrated by the blue vectors.

4.2.6 The interplay effect

To assess the dose degradation due to potential interplay effects, the 4D dose calculation should take into account the dynamic of the delivery equipment. The simulation of the interplay effect has already been described in the literature [69, 121, 122, 123], using various methods to reconstruct the delivery timing information. In our implementation, the IBA software ScanAlgo is employed to simulate the delivery process and extract the delivery time and duration for each spot of the PBS plan.

Knowing the delivery time t_i of the i^{th} spot and the organ motion period T_{period} , the 4D phase P_i in which the i^{th} spot would be delivered can be determined:

$$P_i = \frac{t_i + T_{\text{offset}}}{T_{\text{period}}/N} \mod N$$

where N is the number of 4DCT phases, and T_{offset} is the starting time of the delivery within the motion cycle, which may vary for each beam and fraction. This spot sorting method is employed to generate a partial plan for each phase. Instead of computing the full plan on all 4DCT phases, like for the regular 4D dose calculation, partial plans are simulated by MCsquare on their appropriate CT phase. The resulting partial doses are then deformed to the reference phase and accumulated, providing a 3D dose map that includes both organ and delivery motion effects.

4.2.7 Sampling of uncertainty scenarios

In commercial TPS, uncertainty scenarios are usually generated by mixing errors to reproduce typical PTV margins. However, limiting the scenario selection to these extreme error values does not guarantee the plan robustness for intermediate errors. A Monte Carlo approach is proposed in this study to overcome the challenges of scenario selection. Instead of mixing uncertainty models in a combinatorial way, the errors are randomly sampled for each uncertainty model. This permits to cover more values of uncertainties than the common practice that consists in simulating only the minimum, null, and maximum values for each error type. More specifically, in our implementation, the 3 spatial components of the systematic and random setup errors are randomly sampled individually from Gaussian distributions with zero mean and standard deviation set for each component by the user. Similarly, the global range uncertainty is randomly sampled from a Gaussian distribution, centered on the nominal range. The organ motion period and amplitude are sampled from a Gaussian distribution centered on their nominal value, with standard deviations defined by the user according to variations reported in the literature [8, 119, 120, 66]. The initial phase employed for the simulation of the interplay effect is randomly sampled using a continuous uniform distribution from the beginning to the end of the breathing cycle.

In our strategy, each scenario represents a full treatment simulation taking into account inter-fraction variability. Scenario simulation computes each fraction individually with MCsquare, taking into account systematic errors that are sampled once per treatment scenario, and random errors, re-sampled for each fraction. The resulting fraction doses are then summed together to obtain the scenario dose that represents one of the possible full treatment realizations. This process has to be

repeated many times to provide a representative set of full treatment scenarios.

4.2.8 Robustness analysis

The impact of treatment uncertainties on the dose distribution is usually reported by an uncertainty band around the nominal DVH. This band represents the envelope of all DVHs computed by taking into account multiple scenarios of treatment uncertainties. However, in our strategy, the sampling of treatment uncertainties can lead to very large errors due to the long tails of Gaussian distributions. Generally, safety margins recipes are applied in such a way that target coverage is ensured for a reasonable number of patients, typically 90% as suggested in van Herk et al [43]. To mimic this approach, we propose to generate the uncertainty bands by computing percentiles of the scenario DVHs. For instance, the 5th and 95th percentiles would provide the envelope of the dose degradation expected for 90% of the treatment errors.

In this dual Monte Carlo strategy, both dose calculation and scenario sampling lead to statistical uncertainties on the final result. This statistical variability typically decreases proportionally with the square root of the number of simulated events, whereas computation time grows linearly with this number. For Monte Carlo dose calculation, statistical uncertainty on the dose distributions is inversely proportional to the square root of the number of simulated particles. Statistical uncertainty can be assessed for each dose calculation using the batch method explained in section 4.2.1. This method can also predict the total number of particles to simulate for reaching a given statistical level. As each scenario is the sum of multiple treatment fractions, the number of particles simulated for each fraction can be divided by the number of fractions. The statistical level after the summation of all fraction doses should be very close to the desired one. The same principle is employed for the simulation of all phases during 4D dose calculation.

The robustness analysis resulting from the random scenario sampling is also subject to statistical uncertainties. If the robustness analysis is performed multiple times by randomly re-sampling the scenarios, the DVH-bands may vary from one run to another, even when the dose calculation statistical uncertainty is negligible. Moreover, the estimation of such 5th and 95th percentiles that constitute the band is very sensitive to the rare events in the tails of the scenario distribution. This means that the edges of the DVH-band converge slowly when increasing the number of simulated scenarios. It is, however, possible to calculate the probability that the considered percentiles fall between two given values. This confidence interval can be calculated after the simulation of each scenario to evaluate the progression

of the statistical uncertainty. The bootstrap method [124, 125] is employed in this study to efficiently estimate the percentile values and their confidence interval based on a limited number of scenario samples.

4.2.9 Patient case

The robustness test presented in this paper is applied to a patient case in order to illustrate the feasibility of this concept. The MIROpt 4D robust optimizer, developed in house and based on MCsquare dose calculation, was employed to generate a treatment plan for a lung tumor patient. Planning relies on a 4DCT series of this patient. The tumor volume of 3.5 cm^3 and tumor motion amplitude of 17 mm were measured in this 4DCT. A robust plan with 2 beams (150° and 180°) was optimized on the CTV (GTV + 5 mm) based on the 4DCT series. As listed in Table 4.1, the robustness parameters were selected to reproduce the values typically employed in the PTV margin recipe. The range uncertainty was set to $\pm 2.4\%$ as suggested in the table of Paganetti [8] for Monte Carlo dose calculation. The variation of motion amplitude was not considered as it is the case in most clinical treatment planning protocols. The dose prescription is 60 Gy in 35 fractions.

The generated treatment plan was then evaluated using the robustness test method proposed in this paper. Three configurations are compared. The first test simulates only setup errors, range errors, and the nominal motion based on the original 4DCT. In the second test, the motion amplitude variation is also simulated along with setup and range errors. In addition to all these errors, the interplay effect and motion period variations are simulated in the third test. A nominal breathing period of 4 seconds was considered. All robustness parameters in Table 4.1 were set according to typical values reported in the literature [8, 119, 66, 120].

Treatment uncertainties	Planning	Robustness verification
Systematic setup errors	$\pm 2.5 \times (0.4; 0.7; 0.6) \text{ mm}$	$\sigma = (0.4; 0.7; 0.6) \text{ mm}$ [66]
Random setup errors	$\sigma = (1.7; 1.4; 1.4) \text{ mm}$	$\sigma = (1.7; 1.4; 1.4) \text{ mm}$ [66]
Systematic range errors	$\pm 2.4 \%$	$\sigma = 1.6 \%$ [8]
Systematic motion amplitude errors	NA	$\sigma = 0 \%$ [120]
Random motion amplitude errors	NA	36 % [119]
Systematic motion period errors	NA	0 % [120]
Random motion period errors	NA	17 % [120]

Table 4.1: Summary of uncertainty parameters employed for the robust optimization and the robustness evaluation of the treatment plan.

In this study, the DVH-bands result from combining 1000 uncertainty scenarios for each of the 3 configurations. The band edges are computed using the bootstrap method²¹. After each scenario simulation, 1000 bootstrap samples are randomly sampled from the computed scenario DVHs. The 5th and 95th percentiles are estimated for each bootstrap sample. The DVH-band is then calculated by averaging the percentiles over all bootstrap samples. A 95% confidence interval can also be calculated around the band edges by estimating the 2.5th and 97.5th percentiles of the bootstrap distributions. The convergence of the band edges and their confidence intervals at the D95 level can therefore be evaluated after simulating each scenario.

Treatment uncertainties were directly sampled by MCsquare during the scenario simulations. The dose distributions were computed with a resolution of $2 \times 2 \times 2 \text{ mm}^3$. In order to reach a maximum noise level of 1% inside the ITV, 36 million primary particles are simulated per scenario. This number was estimated using the batch method described in sub-section 4.2.1. All computations were performed on dual-processor computer (Intel Xeon E5-2660 v4) running Linux. The computation times are analyzed in the results section.

4.3 Results

The results of the validation tests for the model described in section 4.2.5 are first presented. In the second sub-section, the robustness test algorithm presented in this paper is applied to a lung patient in order to demonstrate the feasibility of this concept.

4.3.1 Validation of motion amplitude variation model

The first validation experiment consists in reconstructing a new 4DCT series by deforming the reference image (MidP-CT) to each phase position. The reconstructed 4DCT is then visually compared with the original 4DCT set. As it can be observed in Figure 4.2, the method can reproduce the organ and target positions of the original 4DCT while reducing noise and artifacts in the generated 4DCT. Importantly, the visible tumor in the CT phase remains inside its GTV contour after reconstruction.

In the second validation experiment, the motion observed in the second 4DCT series is reproduced by deforming the first series using the proposed method. The motion hysteresis representing the position of the GTV center of mass at each

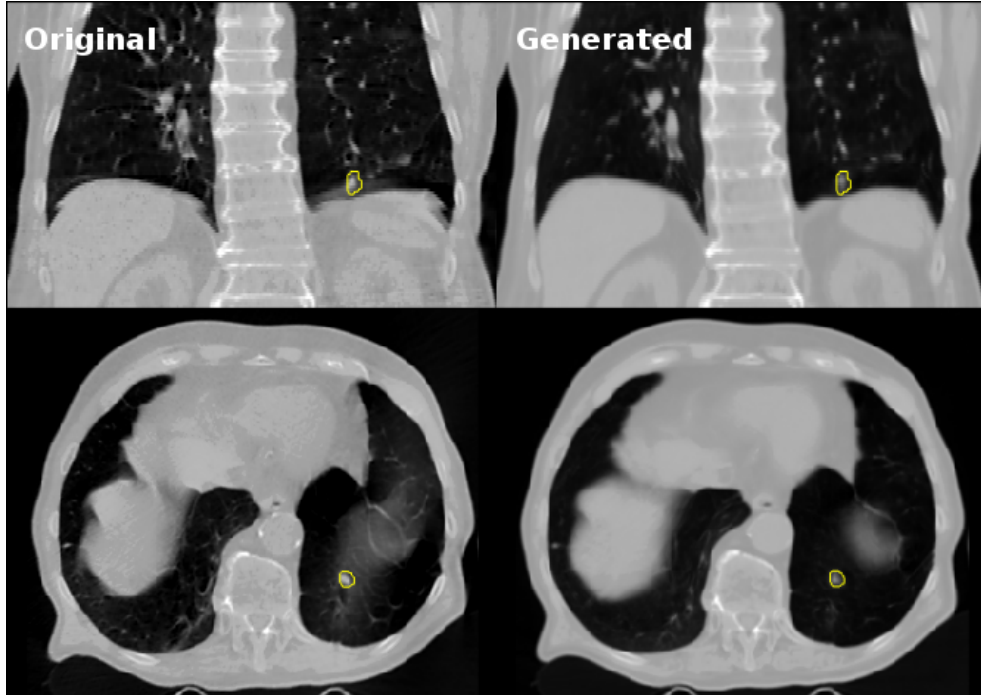


Figure 4.2: Comparison of the density maps calculated by MCsquare from the 8th phase of the original 4DCT (left) and generated 4DCT (right). The GTV drawn on the original CT phase is displayed in yellow on both images.

phase is displayed in Figure 4.3a for both 4DCT acquisitions. From these data, maximum GTV displacements of 33.59 mm and 38.08 mm were calculated for week 1 and week 2, respectively. This represents an increase of 13.4% of the target motion amplitude. The first-week MidP-CT and its associated GTV contour were deformed using the scaled velocity fields to reproduce the motion amplitude measured at week 2. The motion hysteresis of the adjusted week 1 images is compared to week 2 in Figure 4.3b. The motion amplitude of the new generated 4DCT is 37.82 mm. The difference between the modeled and the expected motion amplitude is therefore 0.68%.

4.3.2 Treatment plan robustness evaluation

The robustness of the lung treatment plan was evaluated for three different combinations of uncertainty types. The DVH-bands are displayed in Figure 4.4. DVH indexes are reported in Table 4.2 for the nominal plan resulting from a 4D dose

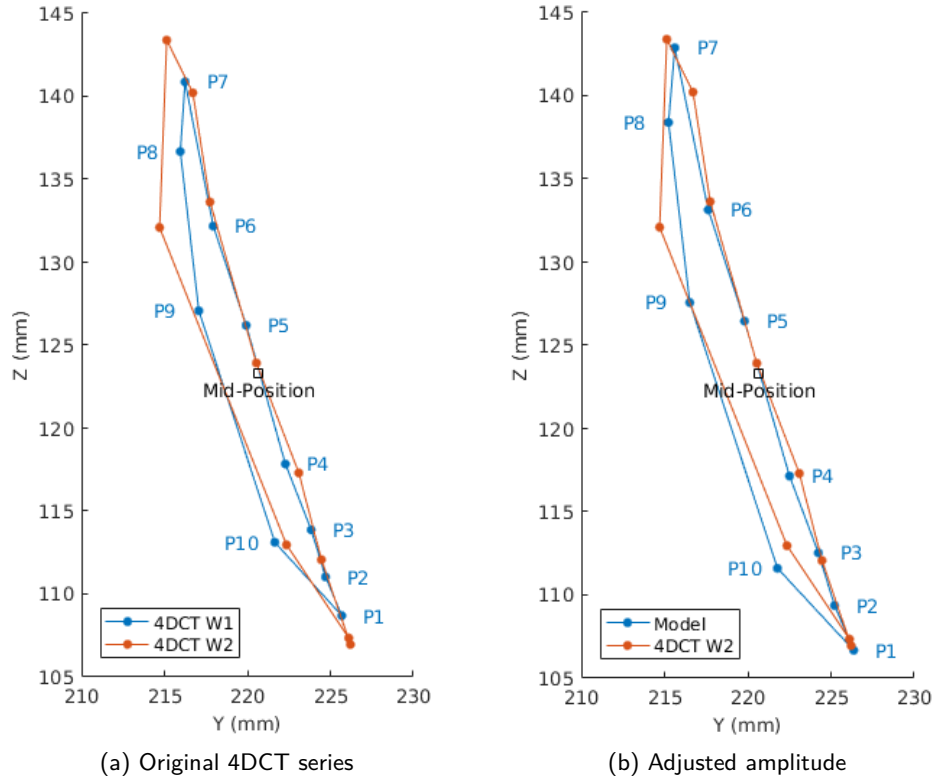


Figure 4.3: Motion hysteresis represented by the position of the GTV center of mass for each phase (P1 to P10).

- a) Comparison of the GTV motion extracted from original 4DCT series acquired at week 1 and week 2.
- b) The motion amplitude measured at week 2 is reproduced by the model using week 1 images.

calculation without treatment uncertainties and for the uncertainty band and its median curve.

For each of these 3 tests, 1000 uncertainty scenarios were simulated to generate the reported results. It is, however, possible to stop the simulation earlier, when the statistical uncertainties on the DVH-band reach an acceptable level. The convergence of the uncertainty band at the D95 level is shown in Figure 4.5 for the robustness test that includes all uncertainty models. The results look unstable below 300 scenarios, and converge slowly beyond. The statistical uncertainty is higher for the 5th percentile (left edge of the band). Its 95% confidence interval

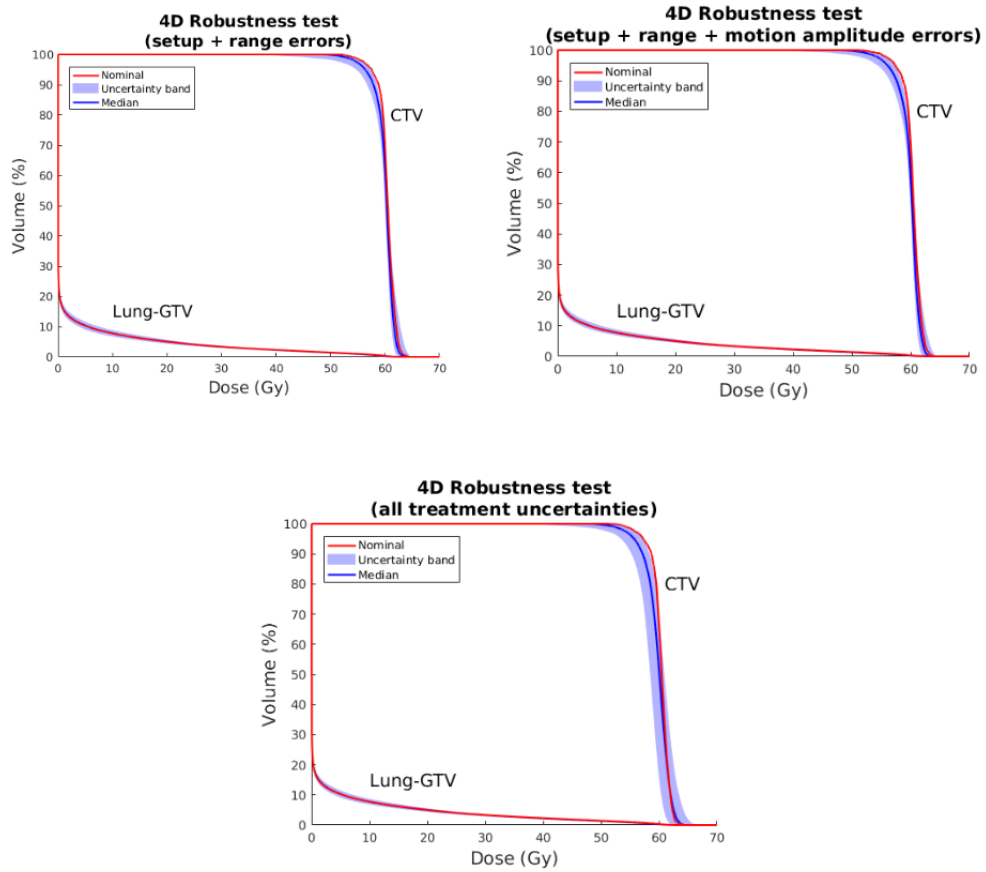


Figure 4.4: DVH-bands resulting from the robustness evaluation of the lung treatment plan. Top left: only setup and range errors are considered. Top right: the variation of motion amplitude is also simulated. Bottom: all treatment uncertainties are simulated, including the interplay effect and the variable breathing motion period. The red curve represents the nominal plan without uncertainties. The light blue band contains 90% of the uncertainty scenarios. The dark blue DVH is the median of all scenarios.

is 0.69 Gy wide for 300 scenarios, 0.49 Gy wide for 500 scenarios, and 0.37 Gy wide for 1000 scenarios.

The computation time was measured and analyzed. The pre-processing step, consisting in 4DCT phases registration, MidP-CT generation, and delivery time simulation is performed automatically by a workflow in the OpenReggui application and takes 12 minutes. The dose calculation and uncertainty models are

	CTV D95	CTV D5	Lung-GTV Dmean
Nominal DVH	57.6 Gy	62.4 Gy	2.7 Gy
Simulated uncertainties: range + setup			
Uncertainty band	54.7 - 57.4 Gy	61.5 - 63.4 Gy	2.4 - 3.0 Gy
Median	56.4 Gy	62.0 Gy	2.7 Gy
Simulated uncertainties: range + setup + variable amplitude			
Uncertainty band	54.4 - 57.3 Gy	61.3 - 63.1 Gy	2.5 - 3.0 Gy
Median	56.1 Gy	62.0 Gy	2.7 Gy
Simulated uncertainties: all			
Uncertainty band	53.4 - 57.3 Gy	60.9 - 64.2 Gy	2.4 - 3.0 Gy
Median	55.7 Gy	62.5 Gy	2.7 Gy

Table 4.2: DVH metrics resulting from the three robustness tests.

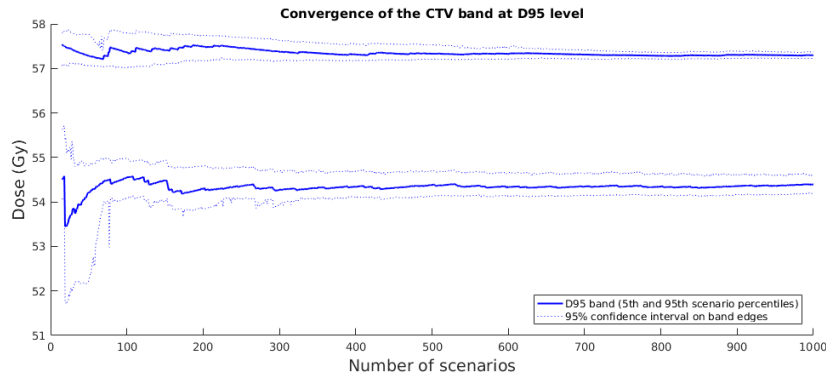


Figure 4.5: Variability of the uncertainty band edges and their 95% confidence interval with respect to the number of simulated scenarios.

performed with MCsquare. During an initialization time of 4 seconds, MCsquare is loading all data required for the Monte Carlo simulation. Then, the nominal dose distribution is computed and takes 97 seconds for this particular lung treatment plan. An average computation time of 385 seconds is required to simulate each uncertainty scenario. This time includes the generation of a new 4DCT (22% of computation time) with scaled motion amplitude for each fraction, other uncertainty models like setup errors (computation time negligible), Monte Carlo dose calculation (49%), and accumulations of 4D and fraction doses (29%). The total computation time necessary to perform the robustness test with 300 scenarios is 32 hours.

4.4 Discussion

In this study, a realistic and accurate robustness verification of proton therapy treatment plans is presented. A novel method for the simulation of breathing motion variation was implemented and combined with the conventional setup and range errors, as well as the simulation of the interplay effect. Systematic and random errors are randomly sampled for each treatment scenario, following a Monte Carlo approach. The variable motion amplitude model was first validated, and the robustness test was illustrated for a lung tumor treatment plan. The validation results demonstrated that the variable motion amplitude model is able to reproduce the motion measured on a second 4DCT acquisition by scaling the velocity fields and deforming the reference CT image. The difference between the motion amplitude measured on the generated 4DCT series and the expected one is 0.68%. This difference may be due to the inaccuracies during the registration or the motion amplitude measurements. The accuracy of the registration and deformation algorithms and their impact on this particular application still need to be studied.

As the reference phase is reconstructed from all 4DCT phases, it is usually less noisy and contains fewer artifacts than the original 4DCT phases. Therefore, the generated 4DCT series, with scaled motion, also benefit from this improved image quality. For this motion reduction/amplification model, the choice of the reference phase is important. For instance, using the Mid-Position phase, as presented in this paper, will lead to a symmetric motion amplification, extending both end-inhale and end-exhale positions. On the contrary, starting from the end-exhale phase, the motion will be amplified only in the inhale direction. This asymmetric motion variation may be preferred if a better stability of the end-exhale position is observed in the patient.

The uncertainty models presented in this paper are not specific to protons and can also be applied to photon treatments. However, density scaling, as used to simulate errors in the HU conversion, would obviously have less impact on photon dose distributions. In contrast, the new variable breathing amplitude model is certainly useful to verify the robustness of photon treatments, too. Moreover, the generated 4DCT series can be imported in a commercial TPS to easily simulate a variation of motion amplitude in a typical clinical workflow.

To illustrate our method, a 4D robust plan was optimized for a lung tumor treatment. The robustness parameters were selected from values typically reported in the literature. For instance, the setup error standard deviations were taken from the study of Brissonette et al [66] where they considered tumor-based positioning

with daily image guidance. This may not be the best choice for proton therapy due to the baseline shift [44]. The potential range variation due to such baseline shift was not considered in our robustness simulations. Another method would be to perform the patient positioning with image registration focused on the bony structures. In that case, the baseline shift effect should be explicitly simulated in the robustness test. Due to the lack of model for this uncertainty, the baseline shift is sometimes approximated as a setup error. The development of a realistic model for the simulation of such baseline shift effect is part of our future work.

Other treatment errors can be implemented in MCsquare. However, it is probably not possible to simulate everything. Also, trying to cover large errors in the treatment plan would lead to a very large irradiated volume. A replanning strategy may be adopted, but this procedure consumes more time. The desired plan robustness is probably a trade-off between the treatment uncertainties that can reasonably be covered and an acceptable number of replanned patients.

The variation of motion amplitude and the interplay effect were not taken into account during the plan optimization. We can observe in the results that, for the studied treatment plan, the interplay effect is the dominant contribution in the DVH band width. This indicates that repainting [70] may be necessary during the delivery to mitigate the interplay effect if clinical constraints are not fulfilled. Although it was not presented in this paper, repainting can be simulated by Scanalgo, and can therefore be taken into account during the robustness test.

Each simulated scenario combines the effect of multiple fractions, taking into account systematic errors, but also random errors that vary from one fraction to another. This allows us to analyze the robustness of the treatment plan for a full treatment. It may be also interesting to perform the simulation without the fraction combination in order to analyze the robustness at the fraction level. For the studied case, the computation time was measured at 385 seconds per full treatment scenario. Moreover, this time must be multiplied by a large number of scenarios to reduce the statistical uncertainty of the resulting DVH-band. The performances can be improved by working on two aspects. First, the computation time per scenario can be reduced by further working on the algorithm implementation. For instance, the generation of 4DCT and dose accumulations represent half of the computation load and mainly consist in image deformations. These operations could be performed by a GPU while the CPU already starts the next Monte Carlo simulation. A cluster of computers could also be used to simulate multiple scenarios simultaneously. The second possible improvement is to reduce the number of scenarios that are required to achieve an acceptable statistical uncertainty.

As for every Monte Carlo algorithm, variance reduction techniques can be applied. For instance, the uncertainty values are currently sampled from Gaussian distributions, favoring the sampling of small error values. However, the percentiles that define the DVH-band are more sensitive to extreme events in the tails of the scenario distribution. To improve the scenario sampling efficiency, uncertainty values could be sampled from uniform distributions with a corrected weighting factor in the percentile estimation. Another improvement could also be adopted for the simulation of the random setup error. Fredriksson [14] demonstrated that for a sufficient number of fractions, the random errors can be simulated by sampling a random shift for each primary particle during the Monte Carlo transport.

4.5 Conclusion

This paper describes a comprehensive, realistic, and statistically sound method of robustness verification for proton therapy treatment plans. A Monte Carlo approach is used to sample the treatment scenarios from multiple uncertainty models. Namely, the variability of the breathing motion, the interplay effect, and the conventional setup and range errors were modeled in the fast Monte Carlo dose engine MCsquare.

The variable breathing amplitude is simulated by generating new 4DCT series with scaled motion. These 4DCT images can be easily imported in a commercial TPS for robustness optimization or verification of photon or proton treatment plans.

Chapter 5

Complement for treatment plan evaluation

In this chapter, additional topics linked to treatment plan evaluation are further discussed to complement the robustness verification method presented in the previous chapter. Especially, a model of the baseline shift is investigated. The simulation of LET distribution and its application in the evaluation of the treatment plan quality is then discussed.

5.1 A model of the baseline shift

Several studies reported both systematic and random variations of the mean position of mobile tumors from fraction to fraction [67, 68]. This so-called baseline shift is a major source of uncertainties for mobile targets and can jeopardize treatment quality. Unlike conventional photon therapy, the inclusion of this error in a PTV margin is overall inadequate in proton therapy because of the range uncertainties. This effect could be taken into account in the BSPTV margin, but this would substantially enlarge the irradiated volume and limits the optimization problems to uniform dose distributions per field (see Section 3.2). Another solution also used in conventional radiotherapy to mitigate the effect of baseline shift is to perform a tumor-based patient positioning. Again, this can only work under the assumption of a static dose distribution. Therefore, this solution cannot guarantee a robust coverage of the target in proton therapy, due to the proton range variation induced by such anatomy shift. The only remaining viable solutions for proton therapy are either adapting the treatment plan, or accounting for this uncertainty in a robust optimization algorithm, using for instance population-based estimations of the shifts. In this section, a model of the baseline shift is proposed in order to simulate this effect during the robust optimization and, more importantly, during the verification of the treatment plan robustness.

5.1.1 Method

The aim of this model is to automatically generate a local deformation field that shift the tumor smoothly when applied on all phases of the planning 4D-CT. The model would therefore be able to generate new images that can be used as uncertainty scenarios for a robust optimization and evaluation of the treatment plan. This process was implemented in the open source application OpenReggui [18]. The entire workflow is illustrated in Figure 5.1.

As shown in Figure 5.1a, the first step of the algorithm consists in generating the initial deformation fields by copying a given shift vector in all voxels. This

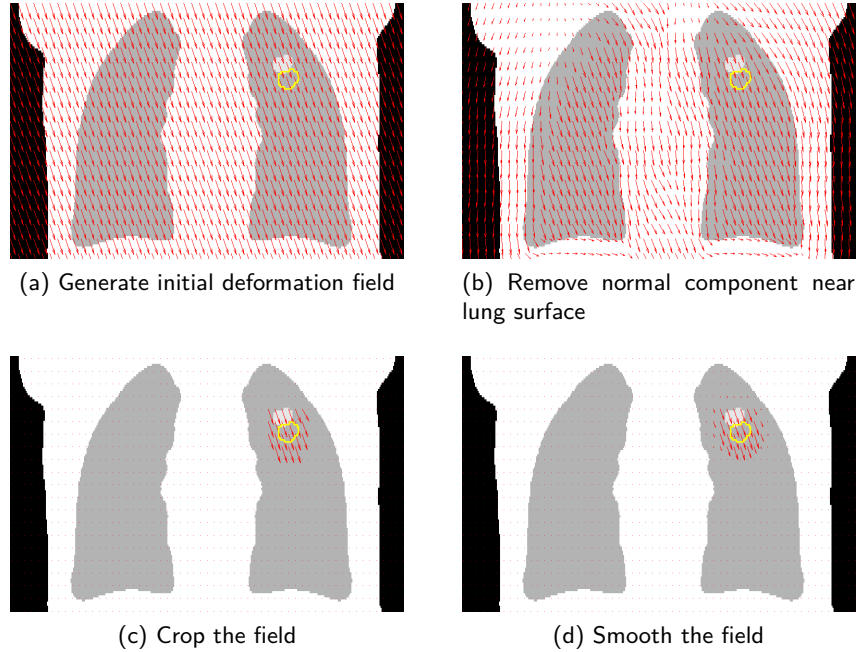


Figure 5.1: Illustration of each step of the baseline shift model. The resulting shifted tumor is represented by the yellow contour.

baseline shift can be either provided by the user, or generated randomly according to a population based estimation of its distribution in order to simulate various scenarios. The next step aims to regularize the deformation according to the patient anatomy so that the final result does not introduce any non-physical artifacts in the images. As represented in Figure 5.1b, the normal component is forced to zero near the lung wall in order to allow the tumor to slip within the lung instead of deforming the lung surface. The resulting deformation field is smoothed using a Gaussian filter to provide a progressive transition. In the third step (Figure 5.1c), the field is cropped to keep only a local deformation around the shifted tumor. Finally, the deformation field is smoothed with a Gaussian filter weighted by a certainty map to keep the correct translation inside the tumor and no deformation outside the lung (Figure 5.1d).

A deformation field is therefore created for each 4DCT phases using the described method. Each phase can then be deformed by the generated deformation field to obtain a new 4DCT with the shifted tumor. This new 4DCT series can then be utilized to compute the dose distribution resulting from the shift.

5.1.2 Validation

Two 4D-CT series acquired at one week interval were used to validate the proposed model. Average CT scans and Mid-Position CT scans (MidP-CT) have been generated from both 4D-CT series. By its definition, the MidP-CT represents an image of the mean position of the anatomy along the breathing cycle. The baseline shift can therefore be calculated by comparing both MidP-CTs. A rigid registration of both images is first performed to simulate an image guided patient positioning during the treatment. The remaining shift between both GTV contours represents the baseline shift. A 5.1 mm tumor shift was measured between the first and the second image acquisition.

Our baseline shift model was then used to attempt to reproduce the baseline shift observed in the second acquisition by deforming the first 4DCT series. A new average CT and MidP-CT were then computed from the generated 4DCT series. As we can see in Figure 5.2, the shift that was visible between week 1 and week 2 disappeared when comparing the simulated image with week 2 acquisition. To perform a more quantitative comparison, the water equivalent thickness (WET) of each voxel was computed for all 3 average CTs. The baseline shift between week 1 and week 2 led to a root mean square deviation (RMSD) of the WET of 0.34 mm inside the GTV. When comparing the simulated shift with the actual image, the WET RMSD is reduced to 0.08 mm. This demonstrate the capability of our model to realistically generate new images that reproduce a given baseline shift.

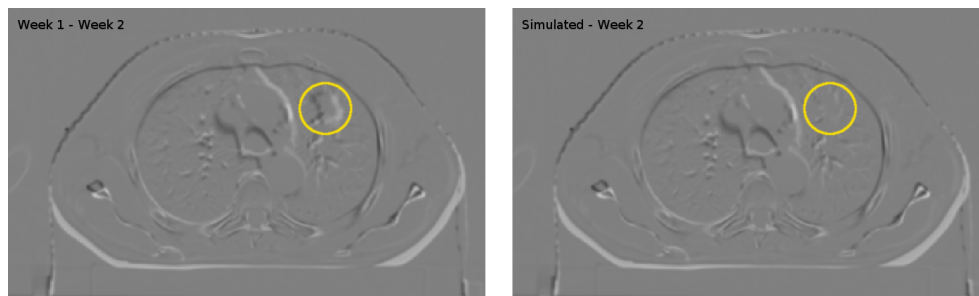


Figure 5.2: The difference between week 1 and week 2 average CTs is reported on the left side. The model is applied to the first week images to try to reproduce the baseline shift observed at the second week. The difference between the simulated shift and the actual image is presented on the right side.

To better evaluate the impact on the proton range, water equivalent path length (WEPL) maps were computed from the WET images for a beam with a gantry

angle of 0° . WEPL profiles are displayed in Figure 5.3. The difference between the WEPL profiles of week 1 and week 2 shows the anatomy changes between both fractions. Our model can reproduce the shift in the tumor position, but it is not able to simulate a full anatomy change. The overlap between WEPL profiles of week 1 and the model results demonstrates that the images are only locally deformed to shift the tumor. The rest of the anatomy remains the same as in week 1.

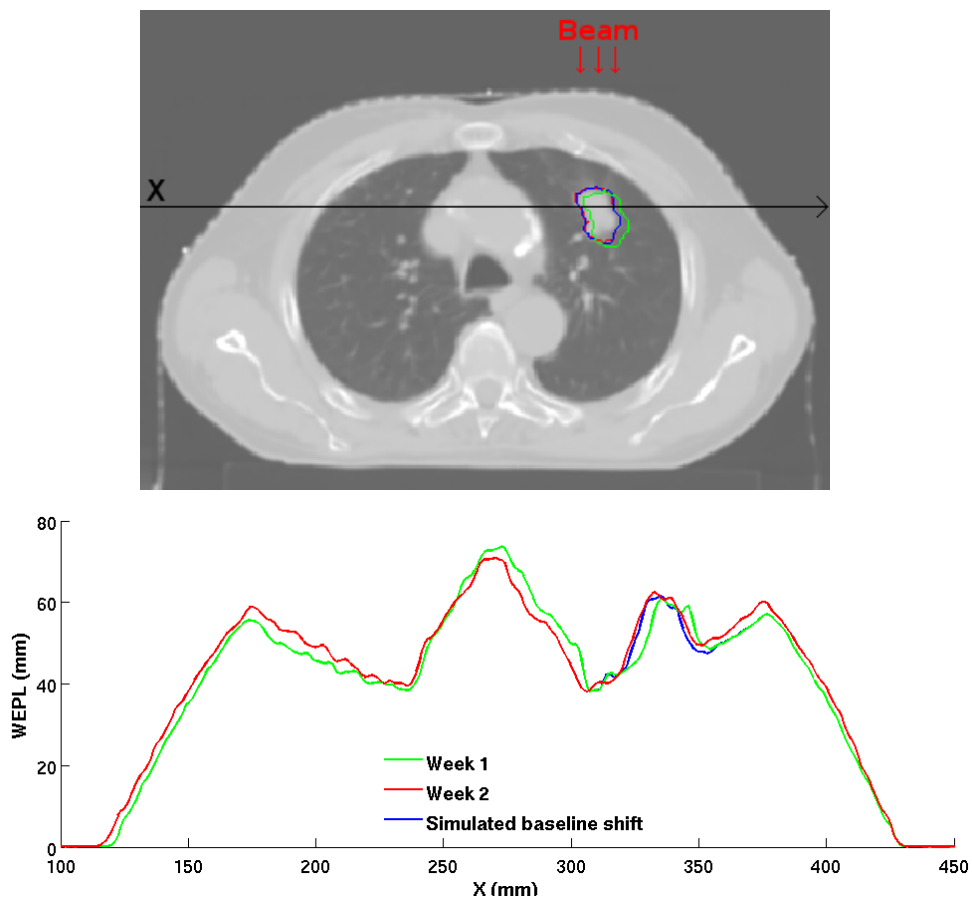


Figure 5.3: WEPL profiles computed along the X axis for a beam angle at 0° . Green and ref profiles are computed from two CT images acquired with 1 week interval. The blue curve is computed from the image generated by the baseline shift model.

5.1.3 Discussion

As for the variable breathing amplitude model, our baseline shift simulation algorithm is able to generate a new 4DCT series, with a shifted mean tumor position. These generated images can then be easily imported in any commercial software. For instance, the generated images could be imported in the 4D optimizer of the RayStation TPS to take into account the baseline shift effect during the treatment plan optimization. Moreover, the model could also be used to improve and verify the robustness of photon treatments.

While the model shows good results most of the time, we are not yet fully satisfied in some cases. Indeed, there is an issue when the tumor is just against the lung wall and the shift is perpendicular this surface. Because the deformation at the surface of the lung is forced to be null, the tumor cannot detach and is stretched by the deformation. Therefore, the work is still in progress. That explains why the model was not yet integrated in MCsquare for the robustness verification method presented in the Chapter 4.

5.2 LET calculation

The treatment plan robustness is not the only parameter that can be verified to evaluate its quality. Even if the plan passes the robustness verification test, it may not be optimal in term of biological effect. Indeed, due to the lack of biological model, the relative biological effect (RBE) is assumed to be a constant of 1.1 for proton therapy [30], thus a physical dose of 10 Gy with protons has the same effect than a physical dose of 11 Gy with photons. However, one of the factors that can affect the RBE is the linear energy transfer (LET). As shown in Figure 5.4, the LET is not constant along the Bragg curve. This means that for a same physical dose, the biological effect can vary along the beam direction. The LET is a physical value that can be easily estimated inside the patient geometry. Unfortunately, there is currently no accurate biological model to link the LET to the RBE value. However, it is known that the RBE is increasing with the LET and eventually reach a maximum [30]. Therefore, the LET distribution could still be verified to try to deliver high LET inside the tumor and low LET in OARs.

LET calculation has been implemented in MCsquare, following the method C described by Cortes-Giraldo and Carabe [126]. Basically, it consists in scoring the average stopping power of proton interacting inside each voxel. The reported LET is thus simply the average stopping power weighted by the energy deposited

by each interaction. The units of the resulting LET map are in $\text{keV} / \mu\text{m}$. This method is compared with micro-dosimetry theoretical results in Figure 5.4 for 160 MeV protons in water. The variance in LET calculation is higher than for dose calculation. Therefore, more simulated particles are required to decrease the statistical noise in LET results.

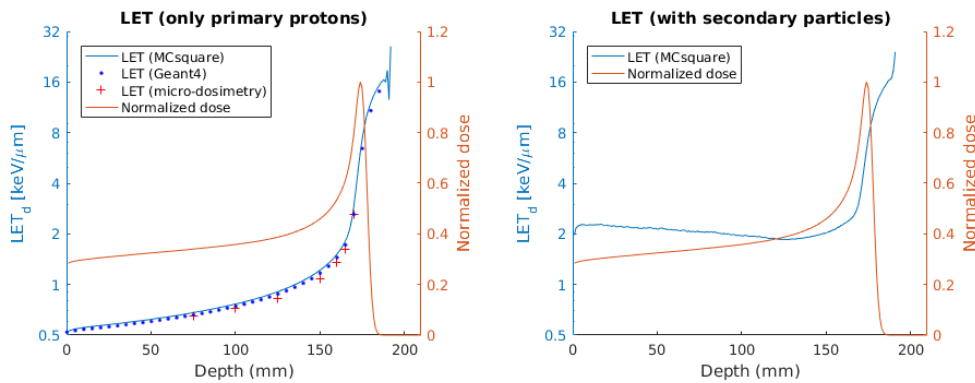


Figure 5.4: Comparison of the LET computed by Monte Carlo (MCsquare and Geant4) and the LET calculated with micro-dosimetry theory for a 160 MeV proton beam in water. On the left plot, only primary protons are taken into account in the LET scoring. On the right plot, all secondary particles generated during nuclear interactions are also included in the LET calculation. The LET axis is in log scale. The micro-dosimetry values are taken from the publication of Cortes-Giraldo and Carabe [126].

MCsquare was recently used at UMCG in Groningen (NL) to simulate the LET distribution for a brain tumor treatment plan. For this patient, the chosen proton beam configuration is untypical. Indeed, all beams come from the cranial side, which may be rather unsafe due to the possible dose increase in the brainstem. The LET results presented in Figure 5 confirm this risk. For this treatment plan, the dose weighted LET distribution, presented in Figure 5.5, shows its maximum value inside the brainstem. The biological dose may therefore be higher than expected inside the organ at risk. These results suggest that the treatment plan should be revised with another beam configuration.

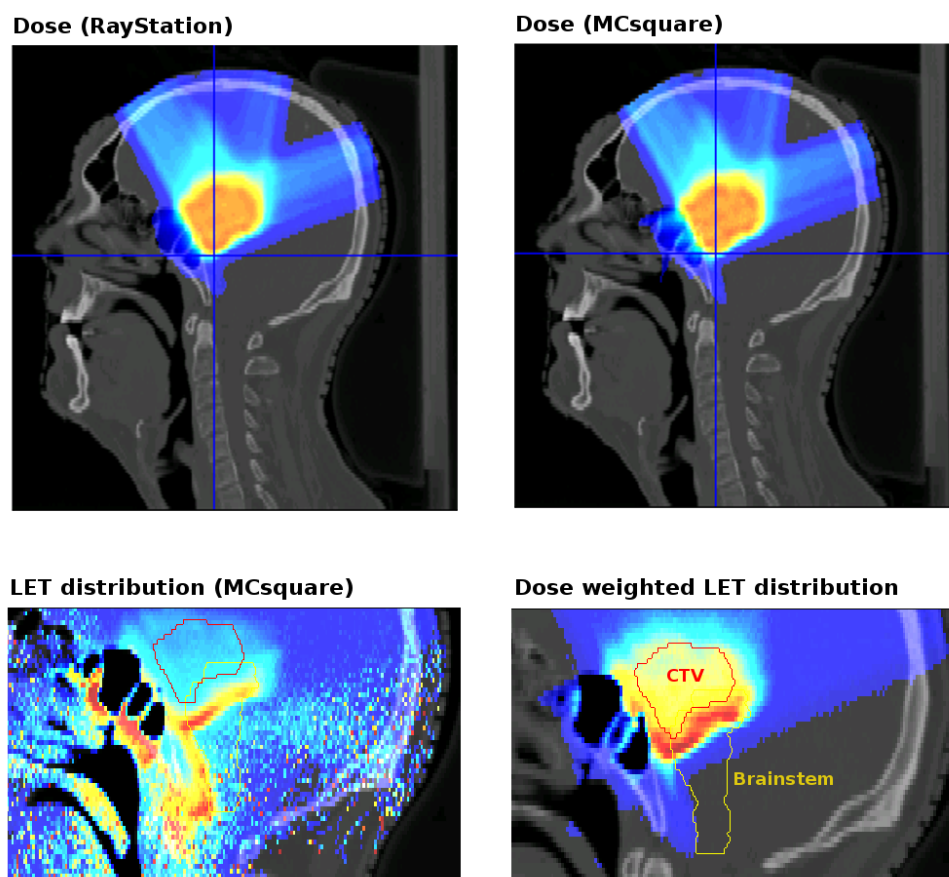


Figure 5.5: Dose and LET distributions for a brain tumor treatment plan. On top left, the dose distribution was calculated with the TPS (RayStation). On top right, the dose distribution was recomputed with MCsquare. On bottom left, the LET distribution was calculated with MCsquare. The statistical noise increases in very low dose regions. On bottom right, the LET distribution is weighted by the dose. High values are located inside the brainstem. Courtesy of Eric Korevaar from UMCg (NL).

Chapter 6

Discussion and conclusion

The Bragg peak is the main advantage of proton therapy, but it is also its main disadvantage. In contrast to the low sensitivity of photon treatments, protons do not forgive for any treatment uncertainties, due to the high dose gradients in the peak. To reach the theoretical potential of proton therapy, these uncertainties must be controlled first. This thesis aims at heading toward this objective. Multiple advanced tools have been developed for proton therapy in this work:

- In order to reduce the uncertainties related to the dose calculation, MC-square, a fast Monte Carlo code has been developed. Its optimized implementation enables accurate and fast dose calculation that is compatible with clinical practice. In addition to the dose calculation, prompt gamma emission, LET, and treatment uncertainties can also be simulated.
- Planning tools have been developed and adapted to proton therapy specificities. Motion evaluation tools have been implemented to provide an analysis of the proton range robustness with respect to the beam angle. Once the most robust beam angles are selected, a robust treatment plan can be generated using other developed tools, such as BSPTV margins for SFUD, or the robust optimizer MIROpt for IMPT.
- To evaluate the robustness of the treatment plan, treatment uncertainties are simulated in a realistic and comprehensive way using the Monte Carlo code MCsquare. These simulations include systematic and random setup errors, flat range uncertainties, variable breathing motion amplitude, and the interplay effect. Multiple scenarios of possible treatment realizations are randomly sampled. By combining the effect of individual fractions the robustness verification tool provides a statistically sound analysis of the effects of systematic and random treatment uncertainties.

The implications of this thesis, as well as some remaining issues are discussed in the next paragraphs.

Clinical benefits of proton therapy

Given the maturity and robustness of conventional radiotherapy with photons, one might wonder if proton therapy is worth its cost. Indeed, proton therapy equipment is expensive and cumbersome [4]. For these reasons, most hospitals do not offer proton therapy yet. Nevertheless, the finite proton range and the Bragg peak offer a great potential to improve the dose conformity to the target [28]. As shown in Figure 1, to achieve target coverage similar as proton therapy with photons, the treatment has to be delivered with many beams from all around the patient,

diluting the higher entrance dose of photons to a large volume. There is, however, an open question regarding the late effect caused by small doses [127]. Which of the two following cases leads to the smaller risk of radiation-induced cancer: a small dose delivered on a large volume, as in IMRT, or a slightly larger dose, but deposited over a smaller volume, as for IMPT? As a matter of fact, though, the integral dose deposited in the patient is generally lower with proton therapy than conventional radiotherapy [28, 4, 128]. The advantages of proton therapy can also be exploited in the other way. For a similar dose deposited to the OAR as what can be achieved with IMRT techniques, proton therapy allows a higher dose to be delivered in the target, therefore increasing the TCP.

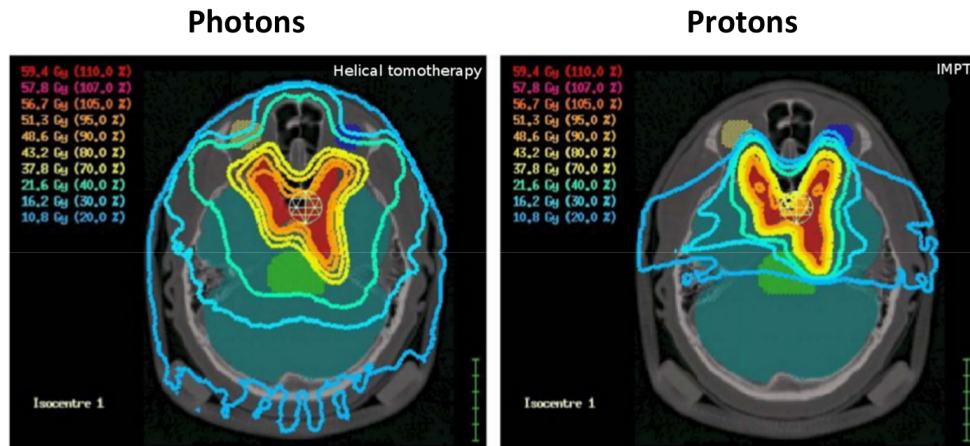


Figure 6.1: Comparison of the same patient planned for a state-of-the-art IMRT treatment (with photon) and for IMPT treatment (with proton). Source: Muzik et al. [28]

Monte Carlo simulations

Monte Carlo algorithms are recognized to provide the best accuracy [8] for proton therapy dose calculation by simulating physical interactions of a large number of particles inside the patient geometry. These simulations were initially too slow to be used for clinical applications and were rather dedicated to research. Fast Monte Carlo algorithms have been recently implemented [16, 53, 54] and released with some commercial TPS. Moreover, the continuous progress in computer technologies will enable even faster calculations in the future. The computation time of Monte Carlo dose engine is therefore not a major issue anymore. However, other issues need to be addressed before Monte Carlo calculation become a standard in clinic. Indeed, Monte Carlo dose engines are generally not easy to set up for a clinical use. One of the main issues is statistical noise, inherent to Monte Carlo

techniques. Efficient methods should be developed to quantify noise in the calculated dose distributions. The impact of noise on the clinical metrics and medical decisions should be studied carefully. It is not clear yet what maximum noise level can be tolerated in clinical applications. Some filters can be used to reduce the noise in the dose distribution [129], but they generally also distort the useful part of the signal. Dose filtering is therefore often a tradeoff between noise reduction and the bias introduced by the filter. In addition, the tools that are currently used in clinics to evaluate the dose are not always compatible with noisy data. In particular, the gamma-test often used to compare two dose distributions [130], is very sensitive to the presence of noise and may provide biased results. Another issue is the specific commissioning that is necessary to model the proton beam properly in the Monte Carlo simulations. This operation is often complex and time consuming due to the lack of automatic commissioning tools for Monte Carlo algorithms. Moreover, as an independent tool it is usually not well integrated in the clinical workflow. The Monte Carlo software should be interfaced with other commercial applications to avoid as much as possible the manual operations of data import/export that are otherwise typically subject to human errors. Working on all these aspects would further facilitate the use of MCsquare in the clinics.

Another open question related to Monte Carlo dose calculation is that it is not clear yet whether the dose calculated with the patient tissue compositions by Monte Carlo should be converted in dose-to-water. On one hand, the true physical dose deposited in the patient is the one reported in medium. On the other hand, the current clinical experience, and reference dosimetry are based on dose reported in water. Moving the clinical practice from dose-to-water to dose-to-medium would require to adapt the results of past clinical trials that define the recommendations in terms of dose prescriptions for tumors and dose limits for OARs. Moreover, Sterpin reported some limitations of the PTV planning strategy when dose-to-medium is used [131]. Robust optimization is therefore mandatory when the dose is reported in the medium. I personally think that moving the clinical practice to dose-to-medium would not improve considerably the accuracy of the treatment and it may not be worth the effort. Moreover, by converting the dose-to-medium calculated by Monte Carlo back to dose-to-water, it reduces the impact of the uncertainty on the material composition of each voxel.

While the computation time necessary to perform a forward dose calculation with Monte Carlo is now acceptable, it is still too slow for more complex applications that require many Monte Carlo calculations. Two examples of such applications were presented in this thesis. First the optimization of PBS treatment plans that requires the calculation of a 3D dose distribution for each individual

spot. The second application is the robustness verification tool. The Monte Carlo simulations has then to be repeated for multiple scenarios of treatment uncertainties. The parallelization scheme implemented in MCsquare consists in sharing the number of particles to be simulated between all computation cores available in the computer. To speed-up the calculation, a second level of parallelization could be considered. As these specific applications require to run multiple individual Monte Carlo calculations, several computers could be used to run multiple simulations simultaneously. Cluster of computers are already commonly used in radiotherapy departments to run commercial applications. Cloud computing is another interesting alternative. The computer resources can be rented in the cloud and scaled according to the work load. This avoids maintaining a large number of local computers that would not be fully exploited. However, this technology would also bring some new issues. As the computation is performed on distant servers, the time to transfer the data might lower the benefits. Moreover, transferring patient data outside the medical center may raise safety and privacy concerns.

Robust planning and robustness verification

Today, the solutions available for proton therapy planning are still limited and do not fully address the robustness issues related to proton range uncertainties. For instance, the BSPTV solution typically available in commercial TPS only take into account systematic setup errors and a flat range uncertainty. The BSPTV method described in this thesis also considers the range variation due to organ motion. Similarly, the uncertainty models available in commercial robust optimizers are limited to systematic range errors and a flat range uncertainty. Such robust plan optimization already demands lots of computational resources. Adding even more uncertainty models in the optimization process would exponentially increase the computation time and memory usage. Therefore, simplifications of the problem, such as the use of analytical algorithms for dose calculation, and limiting the number of uncertainty models, are necessary in the optimization process to keep the computation time acceptable. Due to these approximations, the quality of the resulting treatment plan is not guaranteed and should be verified carefully in the plan evaluation process. Especially, robustness of the treatment plan has to be verified in a comprehensive and realistic way. Unfortunately, the evaluation tools provided in the TPS are usually based on the same approximate algorithms and limited models as for plan optimization. In contrast, the robustness verification tool developed in this thesis offers an independent solution, which is based on accurate Monte Carlo dose calculation and more realistic treatment uncertainty models.

Despite its apparent superiority over margin solutions [12], robust optimization

is rarely employed in the current clinical practice. This can be explained by several issues:

- Robust optimization only considers the CTV, and generates no PTV. Therefore, without PTV contours, robust optimization does not provide a target volume on which the dose distribution could be easily evaluated and reported.
- In comparison to simple PTV-based optimization, robust optimization requires longer computation time and more computational resources to simulate the effect of several uncertainty scenarios.
- As the uncertainty models are limited to setup errors and a flat range variation, it is not clear how other uncertainty types could be taken into account.

The first issue can be addressed with a robustness verification tool. Regarding the second issue, the computation time of both robust optimization and robustness verification can be reduced by parallel computing and future progress in computation technologies. The last issue is, however, critical and will require more research. Some centers increase the setup and range error values in the robust optimization in order to compensate for the uncertainties that are not explicitly modeled. Some even use the value of the photon PTV margin as a setup error for robust optimization of proton treatment. However, the PTV margin includes uncertainties such as the delineation errors, or the baseline shift, that cannot be correctly represented by a setup error. There is no perfect solution given the tools available in the clinic. However, some improvements are probably possible. For instance, geometric errors, such as delineation errors and the geometric part of the baseline shift, could be taken into account with a margin, by expanding the CTV. Other errors that involves range variations can be included, in the model of setup and range errors. The final plan should then be verified with a more realistic simulation of treatment uncertainties using the robustness verification tool presented in this thesis.

Mobile tumors

The treatment of mobile tumors with proton therapy is still an active research topic. These treatments require specific strategies to take into account motion effects. There are currently no clear recommendations for 4D planning in proton therapy and various strategies are being studied independently. The use of a robustness verification tool is therefore very valuable and provides an objective comparison of all these strategies. Planning studies could then establish recommendations by comparing the robustness of several 4D strategies in terms of target coverage and OAR sparing for a large number of patients.

Among all 4D planning strategies, the ITV contour is often used in proton ther-

apy to ensure that the target remains inside the beam along its motion. However, the proton range is also affected by moving heterogeneous tissues. Historically, the CT values inside the ITV were replaced by soft tissue HU in order to improve the target coverage. Later, more sophisticated methods, such as BSPTV or robust optimization, were introduced. All these strategies offer a more or less effective coverage of the target. However, most of the time, the resulting treatment plan is evaluated using a static dose distribution computed on the average CT without real simulation of motion. There is therefore no guarantee of the actual target coverage and there is even less information regarding the dose deposited in organ at risks. This issue has been addressed in this work by implementing a 4D dose calculation algorithm in MCsquare. The dose is computed on each phase of the 4DCT individually and accumulated to a reference phase. Therefore, by combining the contribution of multiple static doses at various phases of the motion cycle, the resulting dose distribution reflects more accurately the dose actually deposited in the mobile tumor and OARs. Such 4D algorithms are now available in some commercial TPS. However, they are rarely employed in clinic due to their longer computation time. On the other hand, the Monte Carlo implementation of 4D algorithms does not increase the computation time. Indeed, the computation time of Monte Carlo dose calculation is linearly proportional to the number of simulated particles. During 4D dose calculation, the number of particles simulated for each phase can be divided by the number of phases. In such way, the statistical uncertainty of the final dose distribution, after accumulation of all phases, is similar as for a 3D dose calculation simulated with the same total number of particles.

Reduction of treatment uncertainties

Improving robustness of treatment plans against uncertainties is mandatory in proton therapy. However, we should also continue working on the reduction of such uncertainties in order to fully exploit the potential of proton therapy. Image guidance was a big improvement for patient positioning. For instance, Bissonnette et al [66] reported that without image guidance the setup error exceeded 5 mm for more than half of their lung patients. By using CBCT imaging to correct the patient position before each treatment, the setup error was lower than 3 mm for more than 80% of these patients. The range uncertainty caused by the conversion of image units (HU) into stopping powers can also be reduced. For instance the use of dual energy CT in planning can help. The development of proton radiography [132], or proton CT, is another ongoing research topic. This would enable a direct measurement of proton stopping powers inside the patient.

Other strategies may also be considered to remove or reduce the effect of tumor motion. Breath hold is sometimes employed when the patient can tolerate it.

This excludes the treatment of patients with limited respiratory conditions and treatments with PBS due to the longer delivery time. Gating is another solution that consists in delivering the treatment only during certain phases of the breathing cycle, pausing the irradiation when the tumor moves away from the beam. However, this necessarily increases the treatment time. Moreover, motion has to be monitored properly to start and stop the irradiation at the correct time. Unfortunately, the tumor motion, inside the patient, is not always well correlated with the motion monitoring, which is measured externally. Audio-coaching or mechanical ventilation can help to regularize the breathing. Mechanical ventilation is especially very promising for gating because it can help the patient to repeat multiple short and reproducible breath holds. In such condition, the delivery time could be reduced and become comparable with a normal PBS treatment.

Treatment uncertainties will probably continue to be reduced in the future. However, it will be very difficult to control or cancel them completely. For instance, even with image guidance, the setup error is not null due to imaging limitations. During the breath holds induced by mechanical ventilation, a small drift is likely to remain. Therefore, robust optimization and robustness verification will still remain necessary for a while. If residual uncertainties get small enough eventually, then the use of simplified models might be sufficient.

Adaptive therapy

To cover any possible treatment uncertainties, it would require irradiation of a very large volume. Therefore, a tradeoff is generally sought between the plan robustness and the conformity of the dose distribution to the tumor volume. Too robust a treatment plan is usually not optimal in terms of healthy tissue sparing. It is not easy to find the parameters for uncertainty models that will lead to a good compromise. For these reasons, it is not possible to cover all uncertainties and adaptation of the treatment plan is sometimes required. When the anatomy changed too much, a new treatment plan can be reoptimized, starting from a new planning image. Today, this operation is performed offline and the patient treatment may be delayed a few days, until the new plan is ready. A lot of developments are still necessary, but eventually the planning tools will probably become fast enough for online applications. An image of the patient will be acquired before and/or during the delivery and the treatment plan would be adapted automatically according to the new information. Thanks to this daily adaptation of the treatment plan, some uncertainties could be substantially reduced. For instance, the plan could be adapted for the anatomical changes, the patient positioning, and even the changes in breathing pattern. Some remaining uncertainties will still need to be considered in the robust optimization, but their

reduction would potentially enable a much better dose conformity to the tumor and better healthy tissue sparing. For instance, the delineation errors, the HU conversion uncertainties, or some residual setup errors could not be completely removed by the use of adaptive therapy strategies. Therefore, accurate models of uncertainties will remain necessary.

In addition to the computation time, several issues still need to be solved in order to achieve online adaptation of the treatment plan. Automating the entire planning procedure is necessary to minimize the time between the patient imaging and the treatment delivery. However, the current clinical practice requires the new plan to be accepted by a physician and validated through a QA procedure involving measurements of the delivered dose. The recalculation of the dose with an independent Monte Carlo algorithm is already one step toward an automatic QA process. Moreover, Monte Carlo simulations provides 3D dose distributions in the patient anatomy, while measurements are often limited to 2D dose distributions in a water phantom.

Implications for other radiotherapy modalities

Most of the tools developed during this PhD can be easily adapted to other delivery modalities. For instance, all uncertainty models can be directly employed for conventional radiotherapy with photons, or other particle therapy such as with carbon ions. The developed models to simulate a variable breathing amplitude and baseline shifts would probably be very useful for IMRT treatments too.

The Monte Carlo algorithm MCsquare can also be easily adapted for simulation of charged particles heavier than proton by tuning the multiple scattering model and by replacing the database of stopping powers and nuclear cross sections. The contribution of neutral particles to the total dose is negligible in proton therapy. Therefore, their transport was ignored in order to speed-up the simulation. However, in carbon therapy simulations, the contribution of neutrons to the total dose may be higher and their transport may be required to achieve a good accuracy.

Final conclusion

Proton therapy has the potential to better spare healthy tissues than conventional radiotherapy with photons. However, considerable uncertainty affects the proton range. To make the treatment plan robust against uncertainties, specific robust planning strategies must be adopted. In general, this leads paradoxically to a larger irradiated volume, which undermines substantially the potential benefits of proton therapy. The accurate dose calculation algorithm and realistic models of treatment uncertainties that are proposed in this thesis are key elements to

optimize and verify suitably the robustness of the treatment plan, and to avoid unnecessary irradiation of large volumes. More efforts are still necessary to further decrease the treatment uncertainties. New imaging techniques such as dual energy CT or proton CT can provide more information about the patient tissue composition. The introduction of prompt gamma imaging in the clinic would provide an in-vivo verification of proton range, which may be used to adapt the treatment plan accordingly. Online adaptation of the treatment plan using fresh CT or MR patient images has the potential to remove a large part of the treatment uncertainties. This will allow reducing the level of robustness that is necessary to deliver the treatment safely and, therefore, this leads also to better healthy tissue sparing.

Most advanced photon treatments permit to deliver intensity modulated beams while the gantry is rotating around the patient. Such volumetric modulated arc-therapy (VMAT) delivery is not available yet for proton therapy. However, preliminary studies demonstrated that proton therapy could benefit from arc-therapy, thanks to potential dosimetric advantages and a better treatment robustness[133, 134, 135]. To conclude, there are still many possible improvements, but eventually proton therapy will reach the same level of maturity as conventional radiotherapy. Online treatment plan adaptation combined with highly flexible delivery techniques may considerably change the future of proton therapy.

References

- [1] 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](https://doi.org/10.1016/j.radonc.2014.03.024)] [PubMed:[24833561](https://pubmed.ncbi.nlm.nih.gov/24833561/)].
- [2] X. Zhang, Y. Li, X. Pan, L. Xiaoqiang, R. Mohan, R. Komaki, J. D. Cox, and J. Y. Chang. Intensity-modulated proton therapy reduces the dose to normal tissue compared with intensity-modulated radiation therapy or passive scattering proton therapy and enables individualized radical radiotherapy for extensive stage IIIB non-small-cell lung cancer: a virtual clinical study. *Int. J. Radiat. Oncol. Biol. Phys.*, 77(2):357–366, Jun 2010. [PubMed Central:[PMC2868090](https://pubmed.ncbi.nlm.nih.gov/PMC2868090/)] [DOI:[10.1016/j.ijrobp.2009.04.028](https://doi.org/10.1016/j.ijrobp.2009.04.028)] [PubMed:[19660879](https://pubmed.ncbi.nlm.nih.gov/19660879/)].
- [3] T. A. van de Water, A. J. Lomax, H. P. Bijl, M. E. de Jong, C. Schilstra, E. B. Hug, and J. A. Langendijk. Potential benefits of scanned intensity-modulated proton therapy versus advanced photon therapy with regard to sparing of the salivary glands in oropharyngeal cancer. *Int. J. Radiat. Oncol. Biol. Phys.*, 79(4):1216–1224, Mar 2011. [DOI:[10.1016/j.ijrobp.2010.05.012](https://doi.org/10.1016/j.ijrobp.2010.05.012)] [PubMed:[20732761](https://pubmed.ncbi.nlm.nih.gov/20732761/)].
- [4] T. R. Bortfeld and J. S. Loeffler. Three ways to make proton therapy affordable. *Nature*, 549(7673):451–453, Sep 2017. [DOI:[10.1038/549451a](https://doi.org/10.1038/549451a)] [PubMed:[28959981](https://pubmed.ncbi.nlm.nih.gov/28959981/)].
- [5] Frederik Jozef Vernimmen and Kathy Rock. Technological progress in radiation therapy for brain tumors. *Journal of Cancer Therapy*, 5(1):38, 2013.
- [6] 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](https://doi.org/10.1088/0031-9155/53/4/014)] [PubMed:[18263956](https://pubmed.ncbi.nlm.nih.gov/18263956/)].
- [7] 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](https://doi.org/10.1088/0031-9155/53/4/015)] [PubMed:[18263957](https://pubmed.ncbi.nlm.nih.gov/18263957/)].
- [8] 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](https://pubmed.ncbi.nlm.nih.gov/PMC3374500/)] [PubMed:[22571913](https://pubmed.ncbi.nlm.nih.gov/22571913/)].
- [9] 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](https://pubmed.ncbi.nlm.nih.gov/8685250/)].
- [10] L. Hong, M. Goitein, M. Buccioli, 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](https://pubmed.ncbi.nlm.nih.gov/8858722/)].
- [11] 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](https://doi.org/10.1118/1.1769631)] [PubMed:[15377093](https://pubmed.ncbi.nlm.nih.gov/15377093/)].
- [12] 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](https://doi.org/10.1088/0031-9155/56/14/011)] [PubMed:[21709340](https://pubmed.ncbi.nlm.nih.gov/21709340/)].
- [13] 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](https://doi.org/10.1088/0031-9155/52/10/009)] [PubMed:[17473350](https://pubmed.ncbi.nlm.nih.gov/17473350/)].
- [14] 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](https://doi.org/10.1118/1.4883837)] [PubMed:[25086511](https://pubmed.ncbi.nlm.nih.gov/25086511/)].
- [15] Mcsquare: fast monte carlo simulations for proton therapy. <http://www.openmcsquare.org/>. Accessed: 2017-09-28.
- [16] K. Souris, J. A. Lee, and E. Sterpin. Fast multipurpose Monte Carlo simulation for proton therapy using multi- and many-core CPU architectures. *Med Phys*, 43(4):1700, Apr 2016. [DOI:[10.1118/1.4943377](https://doi.org/10.1118/1.4943377)] [PubMed:[27036568](https://pubmed.ncbi.nlm.nih.gov/27036568/)].
- [17] Miropt: 4d robust optimization for proton therapy. <http://www.openmiropt.org/>. Accessed: 2017-09-28.
- [18] Openreggui. <https://openreggui.org/>. Accessed: 2017-09-28.
- [19] K. D. Kochanek, S. L. Murphy, J. Xu, and B. Tejada-Vera. Deaths: Final Data for 2014. *Natl Vital Stat Rep*, 65(4):1–122, Jun 2016. [PubMed:[27378572](https://pubmed.ncbi.nlm.nih.gov/27378572/)].
- [20] 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](https://doi.org/10.3322/caac.21262)] [PubMed:[25651787](https://pubmed.ncbi.nlm.nih.gov/25651787/)].

- [21] Errol C Friedberg. Dna damage and repair. *Nature*, 421(6921):436–440, 2003.
- [22] Philip Mayles, Alan Nahum, and Jean-Claude Rosenwald. *Handbook of radiotherapy physics: theory and practice*. CRC Press, 2007.
- [23] Harald Paganetti. *Proton therapy physics*. CRC Press, 2016.
- [24] M. Baumann and C. Petersen. TCP and NTCP: a basic introduction. *Rays*, 30(2):99–104, 2005. [PubMed:16294901].
- [25] H. R. Withers. Biologic basis for altered fractionation schemes. *Cancer*, 55(9 Suppl):2086–2095, May 1985. [PubMed:3919923].
- [26] RE Drzymala, Radhe Mohan, L Brewster, J Chu, Michael Goitein, W Harms, and M Urie. Dose-volume histograms. *International Journal of Radiation Oncology* Biology* Physics*, 21(1):71–78, 1991.
- [27] The hidden face of cern. <http://www.quantumdiaries.org/2012/02/15/the-hidden-face-of-cern/>. Accessed: 2017-09-28.
- [28] J. Muzik, M. Soukup, and M. Alber. Comparison of fixed-beam IMRT, helical tomotherapy, and IMPT for selected cases. *Med Phys*, 35(4):1580–1592, Apr 2008. [DOI:10.1118/1.2890085] [PubMed:18491552].
- [29] H. I. Amols, B. Lagueux, and D. Cagna. Radiobiological effectiveness (RBE) of megavoltage X-ray and electron beams in radiotherapy. *Radiat. Res.*, 105(1):58–67, Jan 1986. [PubMed:3080801].
- [30] H. Paganetti, A. Niemierko, M. Ancukiewicz, L. E. Gerweck, M. Goitein, J. S. Loeffler, and H. D. Suit. Relative biological effectiveness (RBE) values for proton beam therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 53(2):407–421, Jun 2002. [PubMed:12023146].
- [31] W. K. Weyrather and G. Kraft. RBE of carbon ions: experimental data and the strategy of RBE calculation for treatment planning. *Radiother Oncol*, 73 Suppl 2:S161–169, Dec 2004.
- [32] A. Peeters, J. P. Grutters, M. Pijls-Johannesma, S. Reimoser, D. De Ruyscher, J. L. Severens, M. A. Joore, and P. Lambin. How costly is particle therapy? Cost analysis of external beam radiotherapy with carbon ions, protons and photons. *Radiother Oncol*, 95(1):45–53, Apr 2010. [DOI:10.1016/j.radonc.2009.12.002] [PubMed:20106540].
- [33] B Mukherjee, CL Fuentes, and B Timmermann. Operational health physics during the commissioning phase of west german proton therapy centre essen. In *Proceedings of International Conference on Radiation Protection in Medicine-Setting the Scene for the Next Decade*, pages 3–7, 2012.
- [34] 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].

- [35] Proton therapy: scattering versus scanning. <http://medicalphysicsweb.org/cws/article/opinion/42793>. Accessed: 2017-09-28.
- [36] B Marchand, D Prieels, B Bauvir, R Sépulchre, and M Gérard. Iba proton pencil beam scanning: an innovative solution for cancer treatment. In *Proceedings of EPAC*, pages 2539–2541, 2000.
- [37] J. Moorees and E. Bezak. Four dimensional CT imaging: a review of current technologies and modalities. *Australas Phys Eng Sci Med*, 35(1):9–23, Mar 2012. [DOI:[10.1007/s13246-012-0124-6](https://doi.org/10.1007/s13246-012-0124-6)] [PubMed:[22302463](https://pubmed.ncbi.nlm.nih.gov/22302463/)].
- [38] B. J. Pichler, A. Kolb, T. Nagele, and H. P. Schlemmer. PET/MRI: paving the way for the next generation of clinical multimodality imaging applications. *J. Nucl. Med.*, 51(3):333–336, Mar 2010. [DOI:[10.2967/jnumed.109.061853](https://doi.org/10.2967/jnumed.109.061853)] [PubMed:[20150252](https://pubmed.ncbi.nlm.nih.gov/20150252/)].
- [39] F Edward Boas and Dominik Fleischmann. Ct artifacts: causes and reduction techniques. *Imaging Med*, 4(2):229–240, 2012.
- [40] J. L. Barker, A. S. Garden, K. K. Ang, J. C. O’Daniel, H. Wang, L. E. Court, W. H. Morrison, D. I. Rosenthal, K. S. Chao, S. L. Tucker, R. Mohan, and L. Dong. Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. *Int. J. Radiat. Oncol. Biol. Phys.*, 59(4):960–970, Jul 2004. [DOI:[10.1016/j.ijrobp.2003.12.024](https://doi.org/10.1016/j.ijrobp.2003.12.024)] [PubMed:[15234029](https://pubmed.ncbi.nlm.nih.gov/15234029/)].
- [41] K. C. Bylund, J. E. Bayouth, M. C. Smith, A. C. Hass, S. K. Bhattia, and J. M. Buatti. Analysis of interfraction prostate motion using megavoltage cone beam computed tomography. *Int. J. Radiat. Oncol. Biol. Phys.*, 72(3):949–956, Nov 2008. [DOI:[10.1016/j.ijrobp.2008.07.002](https://doi.org/10.1016/j.ijrobp.2008.07.002)] [PubMed:[19014783](https://pubmed.ncbi.nlm.nih.gov/19014783/)].
- [42] J. G. Thomas, R. Kashani, J. M. Balter, D. Tatro, F. M. Kong, and C. C. Pan. Intra and interfraction mediastinal nodal region motion: implications for internal target volume expansions. *Med Dosim*, 34(2):133–139, 2009. [PubMed Central:[PMC2714530](https://pubmed.ncbi.nlm.nih.gov/PMC2714530/)] [DOI:[10.1016/j.meddos.2008.07.003](https://doi.org/10.1016/j.meddos.2008.07.003)] [PubMed:[19410142](https://pubmed.ncbi.nlm.nih.gov/19410142/)].
- [43] 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](https://pubmed.ncbi.nlm.nih.gov/10863086/)].
- [44] 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](https://doi.org/10.1016/j.ijrobp.2007.11.042)] [PubMed:[18313530](https://pubmed.ncbi.nlm.nih.gov/18313530/)].
- [45] C. Rasch, I. Barillot, P. Remeijer, A. Touw, M. van Herk, and J. V. Lebesque.

- Definition of the prostate in CT and MRI: a multi-observer study. *Int. J. Radiat. Oncol. Biol. Phys.*, 43(1):57–66, Jan 1999. [PubMed:[9989514](#)].
- [46] William R Leo. *Techniques for nuclear and particle physics experiments: a how-to approach*. Springer Science & Business Media, 2012.
- [47] Hans Bichsel. Passage of charged particles through matter. Technical report, California Univ., Berkeley (USA). Dept. of Physics, 1969.
- [48] HA Bethe. Moliere's theory of multiple scattering. *Physical Review*, 89(6):1256, 1953.
- [49] W. Schneider, T. Bortfeld, and W. Schlegel. Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions. *Phys Med Biol*, 45(2):459–478, Feb 2000. [PubMed:[10701515](#)].
- [50] M. Bazalova, J. F. Carrier, L. Beaulieu, and F. Verhaegen. Dual-energy CT-based material extraction for tissue segmentation in Monte Carlo dose calculations. *Phys Med Biol*, 53(9):2439–2456, May 2008. [DOI:[10.1088/0031-9155/53/9/015](#)] [PubMed:[18421124](#)].
- [51] 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](#)].
- [52] Joao Seco and Frank Verhaegen. *Monte Carlo techniques in radiation therapy*. CRC press, 2013.
- [53] 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](#)].
- [54] 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](#)].
- [55] U. Oelfke and T. Bortfeld. Inverse planning for photon and proton beams. *Med Dosim*, 26(2):113–124, 2001. [PubMed:[11444513](#)].
- [56] 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](#)].
- [57] 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](https://doi.org/10.1088/0031-9155/53/6/013)] [PubMed:[18367797](https://pubmed.ncbi.nlm.nih.gov/18367797/)].
- [58] 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](https://doi.org/10.1118/1.3556559)] [PubMed:[21520880](https://pubmed.ncbi.nlm.nih.gov/21520880/)].
- [59] A. Trofimov, J. Unkelbach, T. F. DeLaney, and T. Bortfeld. Visualization of a variety of possible dosimetric outcomes in radiation therapy using dose-volume histogram bands. *Pract Radiat Oncol*, 2(3):164–171, 2012. [PubMed Central:[PMC3388515](https://pubmed.ncbi.nlm.nih.gov/PMC3388515/)] [DOI:[10.1016/j.prro.2011.08.001](https://doi.org/10.1016/j.prro.2011.08.001)] [PubMed:[22773939](https://pubmed.ncbi.nlm.nih.gov/22773939/)].
- [60] H. Paganetti, H. Jiang, K. Parodi, R. Slopsema, and M. Engelsman. Clinical implementation of full Monte Carlo dose calculation in proton beam therapy. *Phys Med Biol*, 53(17):4825–4853, Sep 2008. [DOI:[10.1088/0031-9155/53/17/023](https://doi.org/10.1088/0031-9155/53/17/023)] [PubMed:[18701772](https://pubmed.ncbi.nlm.nih.gov/18701772/)].
- [61] H. Paganetti, H. Jiang, S. Y. Lee, and H. M. Kooy. Accurate Monte Carlo simulations for nozzle design, commissioning and quality assurance for a proton radiation therapy facility. *Med Phys*, 31(7):2107–2118, Jul 2004. [DOI:[10.1118/1.1762792](https://doi.org/10.1118/1.1762792)] [PubMed:[15305464](https://pubmed.ncbi.nlm.nih.gov/15305464/)].
- [62] B. Arjomandy, N. Sahoo, X. Ding, and M. Gillin. Use of a two-dimensional ionization chamber array for proton therapy beam quality assurance. *Med Phys*, 35(9):3889–3894, Sep 2008. [DOI:[10.1118/1.2963990](https://doi.org/10.1118/1.2963990)] [PubMed:[18841839](https://pubmed.ncbi.nlm.nih.gov/18841839/)].
- [63] L. Lin, M. Kang, T. D. Solberg, T. Mertens, C. Baeumer, C. G. Ainsley, and J. E. McDonough. Use of a novel two-dimensional ionization chamber array for pencil beam scanning proton therapy beam quality assurance. *J Appl Clin Med Phys*, 16(3):5323, May 2015. [PubMed:[26103492](https://pubmed.ncbi.nlm.nih.gov/26103492/)].
- [64] J. D. Lawson, T. Fox, E. Elder, A. Nowlan, L. Davis, J. Keller, and I. Crocker. Early clinical experience with kilovoltage image-guided radiation therapy for interfraction motion management. *Med Dosim*, 33(4):268–274, 2008. [DOI:[10.1016/j.meddos.2007.06.002](https://doi.org/10.1016/j.meddos.2007.06.002)] [PubMed:[18973853](https://pubmed.ncbi.nlm.nih.gov/18973853/)].
- [65] L. A. Dawson and M. B. Sharpe. Image-guided radiotherapy: rationale, benefits, and limitations. *Lancet Oncol.*, 7(10):848–858, Oct 2006. [DOI:[10.1016/S1470-2045\(06\)70904-4](https://doi.org/10.1016/S1470-2045(06)70904-4)] [PubMed:[17012047](https://pubmed.ncbi.nlm.nih.gov/17012047/)].
- [66] J. P. Bissonnette, T. G. Purdie, J. A. Higgins, W. Li, and A. Bezjak. Cone-beam computed tomographic image guidance for lung cancer radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 73(3):927–934, Mar 2009. [DOI:[10.1016/j.ijrobp.2008.08.059](https://doi.org/10.1016/j.ijrobp.2008.08.059)] [PubMed:[19095368](https://pubmed.ncbi.nlm.nih.gov/19095368/)].
- [67] J. J. Sonke, M. Rossi, J. Wolthaus, M. van Herk, E. Damen, and J. Belderbos. Frameless stereotactic body radiotherapy for lung cancer using four-dimensional cone beam CT guidance. *Int. J. Radiat. Oncol.*

- Biol. Phys.*, 74(2):567–574, Jun 2009. [DOI:[10.1016/j.ijrobp.2008.08.004](https://doi.org/10.1016/j.ijrobp.2008.08.004)] [PubMed:[19046825](https://pubmed.ncbi.nlm.nih.gov/19046825/)].
- [68] S. Takao, N. Miyamoto, T. Matsuura, R. Onimaru, N. Katoh, T. Inoue, K. L. Sutherland, R. Suzuki, H. Shirato, and S. Shimizu. Intrafractional Baseline Shift or Drift of Lung Tumor Motion During Gated Radiation Therapy With a Real-Time Tumor-Tracking System. *Int. J. Radiat. Oncol. Biol. Phys.*, 94(1):172–180, Jan 2016. [DOI:[10.1016/j.ijrobp.2015.09.024](https://doi.org/10.1016/j.ijrobp.2015.09.024)] [PubMed:[26700711](https://pubmed.ncbi.nlm.nih.gov/26700711/)].
- [69] 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](https://doi.org/10.1002/mp.12040)] [PubMed:[28133755](https://pubmed.ncbi.nlm.nih.gov/28133755/)].
- [70] A. C. Knopf, T. S. Hong, and A. Lomax. Scanned proton radiotherapy for mobile targets-the effectiveness of re-scanning in the context of different treatment planning approaches and for different motion characteristics. *Phys Med Biol*, 56(22):7257–7271, Nov 2011. [DOI:[10.1088/0031-9155/56/22/016](https://doi.org/10.1088/0031-9155/56/22/016)] [PubMed:[22037710](https://pubmed.ncbi.nlm.nih.gov/22037710/)].
- [71] A. C. Knopf and A. Lomax. In vivo proton range verification: a review. *Phys Med Biol*, 58(15):R131–160, Aug 2013. [DOI:[10.1088/0031-9155/58/15/R131](https://doi.org/10.1088/0031-9155/58/15/R131)] [PubMed:[23863203](https://pubmed.ncbi.nlm.nih.gov/23863203/)].
- [72] I. Perali, A. Celani, L. Bombelli, C. Fiorini, F. Camera, E. Clementel, S. Henrotin, G. Janssens, D. Prieels, F. Roellinghoff, J. Smeets, F. Stichelbaut, and F. Vander Stappen. Prompt gamma imaging of proton pencil beams at clinical dose rate. *Phys Med Biol*, 59(19):5849–5871, Oct 2014. [DOI:[10.1088/0031-9155/59/19/5849](https://doi.org/10.1088/0031-9155/59/19/5849)] [PubMed:[25207724](https://pubmed.ncbi.nlm.nih.gov/25207724/)].
- [73] C. Richter, G. Pausch, S. Barczyk, M. Priegnitz, I. Keitz, J. Thiele, J. Smeets, F. V. Stappen, L. Bombelli, C. Fiorini, L. Hotoiu, I. Perali, D. Prieels, W. Enghardt, and M. Baumann. First clinical application of a prompt gamma based in vivo proton range verification system. *Radiother Oncol*, 118(2):232–237, Feb 2016. [DOI:[10.1016/j.radonc.2016.01.004](https://doi.org/10.1016/j.radonc.2016.01.004)] [PubMed:[26774764](https://pubmed.ncbi.nlm.nih.gov/26774764/)].
- [74] K. Jabbari and J. Seuntjens. A fast Monte Carlo code for proton transport in radiation therapy based on MCNPX. *J Med Phys*, 39(3):156–163, Jul 2014. [PubMed Central:[PMC4154183](https://pubmed.ncbi.nlm.nih.gov/PMC4154183/)] [DOI:[10.4103/0971-6203.139004](https://doi.org/10.4103/0971-6203.139004)] [PubMed:[25190994](https://pubmed.ncbi.nlm.nih.gov/25190994/)].
- [75] X. Jia, P. Ziegenhein, and S. B. Jiang. GPU-based high-performance computing for radiation therapy. *Phys Med Biol*, 59(4):R151–182, Feb 2014. [PubMed Central:[PMC4003902](https://pubmed.ncbi.nlm.nih.gov/PMC4003902/)] [DOI:

- [9155/59/4/R151](#)] [[PubMed:24486639](#)].
- [76] G. Pratx and L. Xing. GPU computing in medical physics: a review. *Med Phys*, 38(5):2685–2697, May 2011. [DOI:[10.1118/1.3578605](#)] [[PubMed:21776805](#)].
 - [77] 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](#)].
 - [78] 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](#)].
 - [79] James Jeffers and James Reinders. *Intel Xeon Phi coprocessor high performance programming*. Newnes, 2013.
 - [80] Karl W Schulz, Rhys Ulerich, Nicholas Malaya, Paul T Bauman, Roy Stogner, and Chris Simmons. Early experiences porting scientific applications to the many integrated core (mic) platform. In *TACC-Intel Highly Parallel Computing Symposium, Austin, Texas*, 2012.
 - [81] Endong Wang, Qing Zhang, Bo Shen, Guangyong Zhang, Xiaowei Lu, Qing Wu, and Yajuan Wang. *High-Performance Computing on the Intel Xeon Phi*. Springer, 2014.
 - [82] R Ranman. Intel xeon phi coprocessor architecture and tools. *Apress open*, 2013.
 - [83] J. M. Verburg, H. A. Shih, and J. Seco. Simulation of prompt gamma-ray emission during proton radiotherapy. *Phys Med Biol*, 57(17):5459–5472, Sep 2012. [DOI:[10.1088/0031-9155/57/17/5459](#)] [[PubMed:22864267](#)].
 - [84] J. Smeets, F. Roellinghoff, D. Prieels, F. Stichelbaut, A. Benilov, P. Busca, C. Fiorini, R. Peloso, M. Basilavecchia, T. Frizzi, J. C. Dehaes, and A. Dubus. Prompt gamma imaging with a slit camera for real-time range control in proton therapy. *Phys Med Biol*, 57(11):3371–3405, Jun 2012. [DOI:[10.1088/0031-9155/57/11/3371](#)] [[PubMed:22572603](#)].
 - [85] E. Sterpin, G. Janssens, J. Smeets, F. Vander Stappen, D. Prieels, M. Priegnitz, I. Perali, and S. Vynckier. Analytical computation of prompt gamma ray emission and detection for proton range verification. *Phys Med Biol*, 60(12):4915–4946, Jun 2015. [DOI:[10.1088/0031-9155/60/12/4915](#)] [[PubMed:26057053](#)].
 - [86] S. Jan, G. Santin, D. Strul, S. Staelens, K. Assie, D. Autret, S. Avner, R. Barbier, M. Bardies, P. M. Bloomfield, D. Brasse, V. Breton, P. Bruyndonckx, I. Buvat, A. F. Chatzioannou, Y. Choi, Y. H. Chung, C. Comtat,

- D. Donnarieix, L. Ferrer, S. J. Glick, C. J. Groiselle, D. Guez, P. F. Honore, S. Kerhoas-Cavata, A. S. Kirov, V. Kohli, M. Koole, M. Krieguer, D. J. van der Laan, F. Lamare, G. Largeron, C. Lartizien, D. Lazaro, M. C. Maas, L. Maigne, F. Mayet, F. Melot, C. Merheb, E. Pennacchio, J. Perez, U. Pietrzyk, F. R. Rannou, M. Rey, D. R. Schaart, C. R. Schmidtlein, L. Simon, T. Y. Song, J. M. Vieira, D. Visvikis, R. Van de Walle, E. Wieers, and C. Morel. GATE: a simulation toolkit for PET and SPECT. *Phys Med Biol*, 49(19):4543–4561, Oct 2004. [PubMed Central:[PMC3267383](#)] [PubMed:[15552416](#)].
- [87] S. Jan, D. Benoit, E. Becheva, T. Carlier, F. Cassol, P. Descourt, T. Frisson, L. Grevillot, L. Guigues, L. Maigne, C. Morel, Y. Perrot, N. Rehfeld, D. Sarrut, D. R. Schaart, S. Stute, U. Pietrzyk, D. Visvikis, N. Zahra, and I. Buvat. GATE V6: a major enhancement of the GATE simulation platform enabling modelling of CT and radiotherapy. *Phys Med Biol*, 56(4):881–901, Feb 2011. [DOI:[10.1088/0031-9155/56/4/001](#)] [PubMed:[21248393](#)].
- [88] S Agostinelli, John Allison, K al Amako, J Apostolakis, H Araujo, P Arce, M Asai, D Axen, S Banerjee, G Barrand, et al. Geant4—a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.
- [89] Francesc Salvat, José M Fernández-Varea, and Josep Sempau. Penelope-2011: A code system for monte carlo simulation of electron and photon transport. In *Workshop Proceedings*, 2011.
- [90] Geant4 Collaboration et al. Geant4 physics reference manual, 2013. [Online].
- [91] I. Kawrakow. Accurate condensed history Monte Carlo simulation of electron transport. I. EGSnrc, the new EGS4 version. *Med Phys*, 27(3):485–498, Mar 2000. [DOI:[10.1118/1.598917](#)] [PubMed:[10757601](#)].
- [92] Niels Bohr. *The penetration of atomic particles through matter*, volume 18. I kommission hos E. Munksgaard, 1948.
- [93] International Commission on Radiation Units and Measurements. *ICRU Report 49, Stopping powers and ranges for protons and alpha particles*. ICRU, 1993.
- [94] Bruno Rossi and Kenneth Greisen. Cosmic-ray theory. *Reviews of Modern Physics*, 13(4):240–309, 1941.
- [95] International Commission on Radiation Units and Measurements. *ICRU Report 44, Tissue substitutes in radiation dosimetry and measurement*. ICRU, 1989.
- [96] International Commission on Radiation Units and Measurements. *ICRU Report 63, Nuclear data for neutron and proton radiotherapy and for radiation protection*. ICRU, 2000.

- [97] E. Sterpin, J. Sorriaux, and S. Vynckier. Extension of PENELOPE to protons: simulation of nuclear reactions and benchmark with Geant4. *Med Phys*, 40(11):111705, Nov 2013. [DOI:[10.1118/1.4823469](https://doi.org/10.1118/1.4823469)] [PubMed:[24320413](https://pubmed.ncbi.nlm.nih.gov/24320413/)].
- [98] Cross Sections Evaluation Working Group. Endf-6 formats manual. 2009.
- [99] H. Paganetti. Nuclear interactions in proton therapy: dose and relative biological effect distributions originating from primary and secondary particles. *Phys Med Biol*, 47(5):747–764, Mar 2002. [PubMed:[11931469](https://pubmed.ncbi.nlm.nih.gov/11931469/)].
- [100] Richard A Arndt, Igor I Strakovsky, and Ron L Workman. Nucleon-nucleon elastic scattering to 3 gev. *Physical Review C*, 62(3):034005, 2000.
- [101] James Reinders and James Jeffers. *High Performance Parallelism Pearls: Multicore and Many-core Programming Approaches*. Morgan Kaufmann, 2014.
- [102] LuisMiguel Sanchez, Javier Fernandez, Rafael Sotomayor, Soledad Escolar, and J.Daniel. Garcia. A comparative study and evaluation of parallel programming models for shared-memory parallel architectures. *New Generation Computing*, 31(3):139–161, 2013.
- [103] ARB OpenMP. Openmp application program interface, v. 3.0. *OpenMP Architecture Review Board*, 2008.
- [104] Tim Cramer, Dirk Schmidl, Michael Klemm, and Dieter an Mey. Openmp programming on intel xeon phi coprocessors: An early performance comparison. In *Proc. of the Many-core Applications Research Community Symposium*, pages 38–44, 2012.
- [105] 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](https://pubmed.ncbi.nlm.nih.gov/PMC3493036/)] [DOI:[10.1118/1.4758060](https://doi.org/10.1118/1.4758060)] [PubMed:[23127075](https://pubmed.ncbi.nlm.nih.gov/23127075/)].
- [106] D. Sarrut, M. Bardies, N. Bousson, N. Freud, S. Jan, J. M. Letang, G. Lourdos, L. Maigne, S. Marcatili, T. Mauxion, P. Papadimitroulas, Y. Perrot, U. Pietrzyk, C. Robert, D. R. Schaart, D. Visvikis, and I. Buvat. A review of the use and potential of the GATE Monte Carlo simulation code for radiation therapy and dosimetry applications. *Med Phys*, 41(6):064301, Jun 2014. [DOI:[10.1118/1.4871617](https://doi.org/10.1118/1.4871617)] [PubMed:[24877844](https://pubmed.ncbi.nlm.nih.gov/24877844/)].
- [107] 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.
- [108] Rene Brun and Fons Rademakers. Root—an object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research Sec-*

- tion A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389(1):81–86, 1997.
- [109] E. Sterpin, J. Sorriaux, K. Souris, S. Vynckier, and H. Bouchard. A Fano cavity test for Monte Carlo proton transport algorithms. *Med Phys*, 41(1):011706, Jan 2014. [DOI:[10.1118/1.4835475](https://doi.org/10.1118/1.4835475)] [PubMed:[24387498](https://pubmed.ncbi.nlm.nih.gov/24387498/)].
 - [110] Sebastian Anthony. Intel unveils 72-core x86 knights landing cpu for exascale supercomputing, 2013.
 - [111] D Kantor. Knights landing details, 2014.
 - [112] 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](https://pubmed.ncbi.nlm.nih.gov/PMC3005125/)] [DOI:[10.1155/2011/891585](https://doi.org/10.1155/2011/891585)] [PubMed:[21197460](https://pubmed.ncbi.nlm.nih.gov/21197460/)].
 - [113] 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](https://pubmed.ncbi.nlm.nih.gov/PMC3183263/)] [DOI:[10.1016/j.ijrobp.2011.05.011](https://doi.org/10.1016/j.ijrobp.2011.05.011)] [PubMed:[21703781](https://pubmed.ncbi.nlm.nih.gov/21703781/)].
 - [114] 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](https://doi.org/10.1088/0031-9155/58/17/6079)] [PubMed:[23939146](https://pubmed.ncbi.nlm.nih.gov/23939146/)].
 - [115] L. Lin, M. Kang, S. Huang, R. Mayer, A. Thomas, T. D. Solberg, J. E. McDonough, and C. B. Simone. Beam-specific planning target volumes incorporating 4D CT for pencil beam scanning proton therapy of thoracic tumors. *J Appl Clin Med Phys*, 16(6):5678, 11 2015. [DOI:[10.1120/jacmp.v16i6.5678](https://doi.org/10.1120/jacmp.v16i6.5678)] [PubMed:[26699580](https://pubmed.ncbi.nlm.nih.gov/26699580/)].
 - [116] Martin J Berger, JS Coursey, MA Zucker, and J Chang. *Stopping-power and range tables for electrons, protons, and helium ions*. NIST Physics Laboratory Gaithersburg, MD, 1998.
 - [117] 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](https://doi.org/10.1088/0031-9155/56/16/008)] [PubMed:[21791731](https://pubmed.ncbi.nlm.nih.gov/21791731/)].
 - [118] H. Paganetti. Dose to water versus dose to medium in proton beam therapy. *Phys Med Biol*, 54(14):4399–4421, Jul 2009. [DOI:[10.1088/0031-9155/54/14/004](https://doi.org/10.1088/0031-9155/54/14/004)] [PubMed:[19550004](https://pubmed.ncbi.nlm.nih.gov/19550004/)].
 - [119] J. P. Bissonnette, K. N. Franks, T. G. Purdie, D. J. Moseley, J. J. Sonke, D. A. Jaffray, L. A. Dawson, and A. Bezjak. Quantifying interfraction and intrafraction tumor motion in lung stereotactic body radiotherapy using respiration-correlated cone beam computed tomogra-

- phy. *Int. J. Radiat. Oncol. Biol. Phys.*, 75(3):688–695, Nov 2009. [DOI:[10.1016/j.ijrobp.2008.11.066](https://doi.org/10.1016/j.ijrobp.2008.11.066)] [PubMed:[19395200](https://pubmed.ncbi.nlm.nih.gov/19395200/)].
- [120] G. Bosmans, A. van Baardwijk, A. Dekker, M. Ollers, L. Boersma, A. Minken, P. Lambin, and D. De Ruyscher. Intra-patient variability of tumor volume and tumor motion during conventionally fractionated radiotherapy for locally advanced non-small-cell lung cancer: a prospective clinical study. *Int. J. Radiat. Oncol. Biol. Phys.*, 66(3):748–753, Nov 2006. [DOI:[10.1016/j.ijrobp.2006.05.022](https://doi.org/10.1016/j.ijrobp.2006.05.022)] [PubMed:[17011450](https://pubmed.ncbi.nlm.nih.gov/17011450/)].
- [121] S. Dowdell, C. Grassberger, G. C. Sharp, and H. Paganetti. Interplay effects in proton scanning for lung: a 4D Monte Carlo study assessing the impact of tumor and beam delivery parameters. *Phys Med Biol*, 58(12):4137–4156, Jun 2013. [PubMed Central:[PMC3752993](https://pubmed.ncbi.nlm.nih.gov/PMC3752993/)] [DOI:[10.1088/0031-9155/58/12/4137](https://doi.org/10.1088/0031-9155/58/12/4137)] [PubMed:[23689035](https://pubmed.ncbi.nlm.nih.gov/23689035/)].
- [122] C. Grassberger, S. Dowdell, A. Lomax, G. Sharp, J. Shackleford, N. Choi, H. Willers, and H. Paganetti. Motion interplay as a function of patient parameters and spot size in spot scanning proton therapy for lung cancer. *Int. J. Radiat. Oncol. Biol. Phys.*, 86(2):380–386, Jun 2013. [PubMed Central:[PMC3646997](https://pubmed.ncbi.nlm.nih.gov/PMC3646997/)] [DOI:[10.1016/j.ijrobp.2013.01.024](https://doi.org/10.1016/j.ijrobp.2013.01.024)] [PubMed:[23462423](https://pubmed.ncbi.nlm.nih.gov/23462423/)].
- [123] 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](https://pubmed.ncbi.nlm.nih.gov/PMC3843755/)] [DOI:[10.1118/1.4829500](https://doi.org/10.1118/1.4829500)] [PubMed:[24320499](https://pubmed.ncbi.nlm.nih.gov/24320499/)].
- [124] Bradley Efron and Robert Tibshirani. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical science*, pages 54–75, 1986.
- [125] Boris Vaisovich Vishnyakov and Andrei Ivanovich Kibzun. Application of the bootstrap method for estimation of the quantile function. *Automation and Remote Control*, 68(11):1931–1944, 2007.
- [126] M. A. Cortes-Giraldo and A. Carabe. A critical study of different Monte Carlo scoring methods of dose average linear-energy-transfer maps calculated in voxelized geometries irradiated with clinical proton beams. *Phys Med Biol*, 60(7):2645–2669, Apr 2015. [DOI:[10.1088/0031-9155/60/7/2645](https://doi.org/10.1088/0031-9155/60/7/2645)] [PubMed:[25768028](https://pubmed.ncbi.nlm.nih.gov/25768028/)].
- [127] E. J. Hall, C. S. Wu, and E. J. Hall. Radiation-induced second cancers: the impact of 3D-CRT and IMRT. *Int. J. Radiat. Oncol. Biol. Phys.*, 56(1):83–88, May 2003. [PubMed:[12694826](https://pubmed.ncbi.nlm.nih.gov/12694826/)].
- [128] B. R. Eaton, S. M. MacDonald, T. I. Yock, and N. J. Tarbell. Secondary Malignancy Risk Following Proton Radiation Therapy. *Front Oncol*, 5:261,

2015. [PubMed Central:[PMC4659915](#)] [DOI:[10.3389/fonc.2015.00261](#)] [PubMed:[26636040](#)].
- [129] I. El Naqa, I. Kawrakow, M. Fippel, J. V. Siebers, P. E. Lindsay, M. V. Wickerhauser, M. Vicic, K. Zakarian, N. Kauffmann, and J. O. Deasy. A comparison of Monte Carlo dose calculation denoising techniques. *Phys Med Biol*, 50(5):909–922, Mar 2005. [DOI:[10.1088/0031-9155/50/5/014](#)] [PubMed:[15798264](#)].
- [130] D. A. Low and J. F. Dempsey. Evaluation of the gamma dose distribution comparison method. *Med Phys*, 30(9):2455–2464, Sep 2003. [DOI:[10.1118/1.1598711](#)] [PubMed:[14528967](#)].
- [131] E Sterpin. Potential pitfalls of the ptv concept in dose-to-medium planning optimization. *Physica Medica*, 32(9):1103–1110, 2016.
- [132] Cinzia Talamonti, Valentina Reggioli, Mara Bruzzi, Marta Bucciolini, Carlo Civinini, Livia Marrazzo, David Menichelli, Stefania Pallotta, Nunzio Randazzo, Valeria Sipala, et al. Proton radiography for clinical applications. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 612(3):571–575, 2010.
- [133] 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](#)].
- [134] X. Ding, X. Li, A. Qin, J. Zhou, D. Yan, C. Stevens, D. Krauss, and P. Kabolizadeh. Have we reached proton beam therapy dosimetric limitations? - A novel robust, delivery-efficient and continuous spot-scanning proton arc (SPArc) therapy is to improve the dosimetric outcome in treating prostate cancer. *Acta Oncol*, pages 1–3, Aug 2017. [DOI:[10.1080/0284186X.2017.1358463](#)] [PubMed:[28774218](#)].
- [135] X. Ding, X. Li, J. M. Zhang, P. Kabolizadeh, C. Stevens, and D. Yan. Spot-Scanning Proton Arc (SPArc) Therapy: The First Robust and Delivery-Efficient Spot-Scanning Proton Arc Therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 96(5):1107–1116, Dec 2016. [DOI:[10.1016/j.ijrobp.2016.08.049](#)] [PubMed:[27869083](#)].

List of acronyms

2D	two dimensional
3D	three dimensional
4D	four dimensional
4DCT	four dimensional computed tomography
API	application programming interface
ASCII	American Standard Code for Information Interchange
BSPTV	beam-specific planning target volume
CBCT	cone beam computed tomography
CPAP	continuous positive airway pressure
CPU	central processing unit
CT	computed tomography
CTV	clinical target volume
D _#	minimum dose deposited in #% of the volume of interest
D _{mean}	mean dose deposited in the volume of interest
DNA	deoxyribonucleic acid

DS	double scattering
DVH	dose-volume histograms
EM	electromagnetic
FLOPS	floating-point operation per second
GPU	graphical processing unit
GTV	gross tumor volume
HU	Hounsfield units
IBA	Ion Beam Applications s.a.
IMPT	intensity modulated proton therapy
IMRT	intensity modulated radiotherapy
ITV	internal target volume
LET	linear energy transfer
MC	Monte Carlo
MCS	multiple Coulomb scattering
MidP	mid-position
MIMD	multiple instruction on multiple data
MPI	message passing interface
MRI	magnetic resonance imaging
NTCP	normal tissue complication probability

OAR	organ at risk
PBA	pencil beam algorithm
PBS	pencil beam scanning
PCI	peripheral component interconnect
PET	positron emission tomography
PG	prompt gamma
PTV	planning target volume
QA	quality assurance
QC	quality control
RAM	random access memory
RBE	relative biological effectiveness
RMSD	root mean square deviation
SE	standard error
SFUD	single field uniform dose
SIMD	single instruction multiple data
SOBP	spread out Bragg peak
SPR	stopping power ratio
TCP	tumor control probability
TPS	treatment planning system
UCL	Université catholique de Louvain

UMCG Universitair Medisch Centrum Groningen

UPenn University of Pennsylvania

V# volume receiving at least # Gy

VMAT volumetric modulated arc-therapy

WEPL water-equivalent path length

WET water-equivalent thickness