

Working towards a fast prompt gamma emission simulator based on the Boltzmann Equation

W.M.A. Pols



Working towards a fast prompt gamma emission simulator based on the Boltzmann Equation

by

W.M.A. Pols

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Tuesday October 23, 2018 at 13:00 AM.

Student number:	4171969
Project duration:	January 8, 2018 – October 23, 2018
Thesis committee:	Dr. ir. D.R. Schaart TU Delft, chair
	Dr. ir. D. Lathouwers TU Delft, supervisor
	Dr. Z. Perkó TU Delft
	Dr. E. Lens TU Delft, supervisor

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

Abstract

Range verification based on prompt gamma detection is an important step to improve dose control for proton therapy. To deduce the proton range from the detected prompt gamma emission, a prediction of the measured profile is required. This study introduces the Boltzmann solver as a faster alternative to the Monte Carlo simulations to produce dose distributions and prompt gamma source terms from proton therapy treatment plans.

The Boltzmann solver traces mono-energetic pencil beams of finite width through the CT volume. The code only simulates scatterless propagation. The pencil beams are divided into mathematical pencil beams that are individually forward traced using numerical integration to keep track of the fluence and proton energy. The code employs cross section libraries to evaluate stopping power and prompt gamma emission and has an option to include absorption interactions.

To verify the Boltzmann solver, the simulated distributions in simple geometries were compared to data produced by Topas MC. The calculation time of the Boltzmann solver was in the order of seconds, while the simulations in Topas MC took >4 h. It was found that the solver calculated shorter proton ranges than the Monte Carlo code. This is caused by the difference of cross sections between Topas MC and the library employed by the Boltzmann solver. Furthermore, the prompt gamma yield of the Boltzmann solver was found to be significantly higher than that of Topas MC.

Furthermore, the dose deposition of a full treatment plan was compared to the planned dose included in the plan. The simulation time of a full treatment plan was 0.43 h. A γ -index analysis showed that severe deviations occur in high-dose areas, especially when few pencil beams of high intensity deposit the dose.

The Boltzmann solver is mathematically reliable. Additional research must be done to provide it with accurate cross sections. Furthermore, the balance between calculation time and accuracy is to be determined based on the range verification system the code serves. Ultimately, an extensive validation study must be performed.

Acknowledgements

I would like to express my gratitude towards a group of people who have been indispensable over the course of this thesis. Firstly, I would like to thank my supervisors. Danny Lathouwers, who, with his persistent - and I suspect very conscious - ignorance of my insecurities regarding my mathematical skills and programming capabilities, made me push beyond my comfort zone. Furthermore his ideas whenever we met a dead end were vital for the development of the code. It has been an honour to contribute to his exciting simulation tool.

I would like to thank Eelco Lens for taking up the daily supervision of a student whose daily activities largely lied outside his field of expertise. His insights were indispensable and his enthusiasm about our tool worked contagious.

Furthermore I must thank Jules. Over the many walks outside the RID building and coffee breaks we were able to exchange mental support, experience with our simulation tools, and Matlab codes, which allowed me to fast forward through my research.

Lastly I would like to thank my parents, Robert and Hanneke, my boyfriend Sam and house mates Sophie and Daan for their mental support, especially over the last few weeks of the project when I needed them most.

Contents

1	Introduction	1
1.1	Radiotherapy	1
1.1.1	Introduction to radiotherapy.	1
1.1.2	Proton therapy.	1
1.2	Proton interactions with matter.	3
1.2.1	Electronic interactions.	3
1.2.2	Nuclear interactions	4
1.2.3	Widening of the Bragg peak	4
1.3	Proton range verification	4
1.4	Objective	5
2	Methods to simulate prompt gamma profiles for proton therapy: a review	7
2.1	Introduction	7
2.2	Range verification system requirements	8
2.3	Monte Carlo method	8
2.3.1	Geant4	9
2.3.2	MCNP/MCNPX	11
2.3.3	FLUKA	12
2.4	Dedicated nuclear interaction codes	13
2.5	Convolution-based methods	13
2.6	Conclusion	14
3	Derivation of the proton transport equation	15
3.1	Linear Boltzmann equation.	15
3.2	Application to proton therapy.	17
3.3	Fokker-Planck approximation.	17
3.4	Method of characteristics	18
4	Boltzmann solver code	21
4.1	Initialization	21
4.1.1	Tracing volume	21
4.1.2	Pencil beams.	23
4.2	Interpretation of ODEs	23
4.3	Numerical integration methods.	23
4.3.1	Euler explicit method.	23
4.3.2	Heun's method.	24
4.4	Tracing algorithms	26
4.4.1	Gaussian quadrature.	26
4.4.2	Needle beam method	28
4.4.3	Comparison between quadrature and needle beam methods	30
4.5	Reduction of computation time.	30
4.5.1	Euler versus Heun	31
4.5.2	Cut-off radius	31
4.5.3	Binary search	31
5	Simulation set-up	33
5.1	Single beam on homogeneous phantom	33
5.2	Single beam on inhomogeneous phantom	34
5.3	Full treatment plan on a planning CT	34

6	Results	37
6.1	Single beam on homogeneous phantom	37
6.1.1	Prompt gamma emission	40
6.2	Single beam on inhomogeneous phantom	40
6.3	Full treatment plan on a planning CT	43
6.3.1	Gaussian blur	43
6.3.2	Cut-off radius	45
7	Discussion	49
	Bibliography	53

Introduction

1.1. Radiotherapy

1.1.1. Introduction to radiotherapy

Radiotherapy is one of the three main methods to treat cancer, along with chemotherapy and surgery. The aim of this type of treatment is to kill tumor cells with ionizing radiation.

The energy administered during radiation therapy is low: a treatment plan typically prescribes a homogeneous dose of 60-90Gy to a tumor volume, which corresponds to the amount of heat to raise its temperature by 0.02°C. The effectiveness of radiation dose lies in the ionizing interactions with DNA, which can cause single and double strand breaks and eventually cell death.

Healthy cells are able to repair DNA damage to some extent. However, tumor cell functions are disrupted, causing them to be less successful at self-repair. This difference is exploited by administering the dose in multiple fractions of a few Grays. In between the fractions healthy tissues have the opportunity to recover. This way tumor cells are killed while surrounding tissues are spared.

While healthy cells are moderately spared by the fractionation of the treatment plan, the overall dose to healthy tissues, so-called organs at risk (OAR), is minimized. Some OAR are particularly sensitive to radiation, cause discomfort to the patient or cannot regenerate. A maximum dose constraint may be applied for these tissues. Treatment plans carefully balance dose prescriptions to the target volume and OAR while minimizing integral dose. Deviations from the plan may cause sparing of tumor cells or unnecessary healthy cell death.

Correct dose delivery is complicated because of a variety of uncertainties, including deviations in patient alignment, organ motion and anatomical changes of the patient. This is accounted for by application of a safety margin around the clinical target volume (CTV), thereby creating the planning target volume (PTV). This margin ensures that the prescribed dose is delivered to the CTV under these circumstances, but introduces a high dose to tumor-adjacent tissues.

The aim of radiotherapy development is to kill all tumor cells while the dose to healthy tissues is as low as possible. The characteristics of particle therapy allow for significant improvement relative to conventional photon treatment. However, the particle modality is specifically sensitive for uncertainty in dose delivery. Therefore, to fully exploit the advantages of particle therapy, the development of dose control is essential.

1.1.2. Proton therapy

Over the past years, proton irradiation has gained ground in the field of radiation oncology. The main benefit of proton therapy relative to conventional treatment is the concentrated dose delivery at the distal end of the particle range, which allows for more efficient targeting of the tumor volume. The difference between the depth-dependent dose delivery photon and proton treatment is illustrated in figure 1.1.

The dose from photon beams is characterized by a skin sparing effect in the first ~1.5 cm, during which the dose rises to its maximum. After the peak both the photon fluence and radiation dose decrease exponentially. Due to this behaviour high dose is delivered to healthy tissues located before the tumor volume.

In contrast to the photon dose, the dose delivered by proton beams is characterized by a low entrance plateau. The rate of energy-transferring interactions is inversely correlated with the incident proton energy, except for proton energies below 0.1 MeV. Therefore, the dose eventually rises steeply into the so-called Bragg

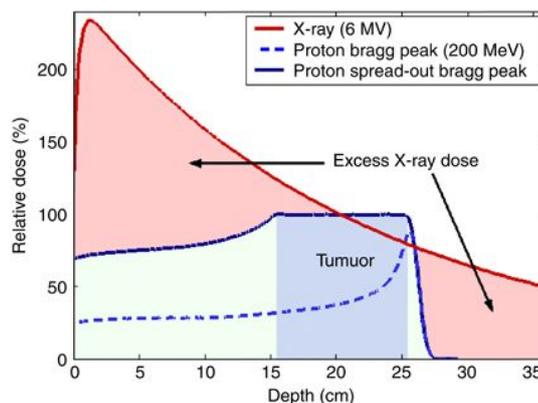


Figure 1.1: Integrated depth dose in a homogeneous medium. The goal is to administer a 100% dose to a tumor volume highlighted in grey. The dose distribution delivered by a photon beam is indicated in red; by a proton beam in blue. A pristine Bragg peak is displayed in dashed blue, the SOBP in solid blue. Image from Efstathiou et al., 2013 [1].

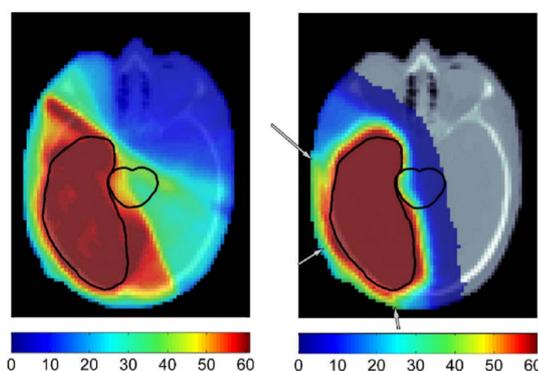


Figure 1.2: Left image: Dose in Gray as calculated by a treatment planning system for a brain tumor, using an intensity modulated radiotherapy (IMRT) planning system. The PTV is the delineated volume to the left. The OAR are the brain stem, optic nerve and chiasm which represented by the central delineation. Right image: Dose calculated by an intensity modulated proton therapy (IMPT) planning system for the same patient. The arrows indicate the directions of the proton beams. Images from Wedenberg et al., 2014 [2].

peak. Behind the Bragg peak protons have lost all momentum and undergo no more interactions, so the dose quickly falls to zero.

The proton range is defined as the depth at which the dose is at 80% of the maximum in the fall-off region. The range depends on the initial proton energy and the material characteristics of the target. To cover a tumor volume with a homogeneous dose, a set of pristine Bragg peaks is accumulated to create a plateau of constant dose, the so-called spread-out Bragg peak (SOBP).

The benefits of the depth dose curve of the SOBP relative to that of a photon beam is thus threefold. Firstly, the dose proximal to the Bragg peak is relatively low. Furthermore, the compounded Bragg peaks allow for homogeneous dose coverage and finally, the tissue distal to the Bragg peak is spared because of the finite range of protons. Consequently, lower integral dose to the patient can be achieved with proton therapy than when using photon therapy (figure 1.2).

On the other hand, the dose delivery profile of proton beams entails the risk of range uncertainty. The proton interaction rates depend linearly on the mass density of the material they interact with. Therefore density deviations between the map used in the treatment plan and the body during irradiation may lead to beam over- or undershoot, causing high dose delivery to healthy tissues and underdosage of the target volume. Sources of uncertainty include errors in CT conversion into mass density, deviations of daily setup and anatomical changes. The latter are illustrated in figure 1.3.

Because of the risk of range uncertainty, proton treatment requires different PTV margins than photon therapy. Multiple margin recipes circulate [3–5], prescribing generous margins (e.g. 6.5mm for 150 MeV proton beams [3]). To reduce these margins and thereby the integral dose, on-line range verification systems

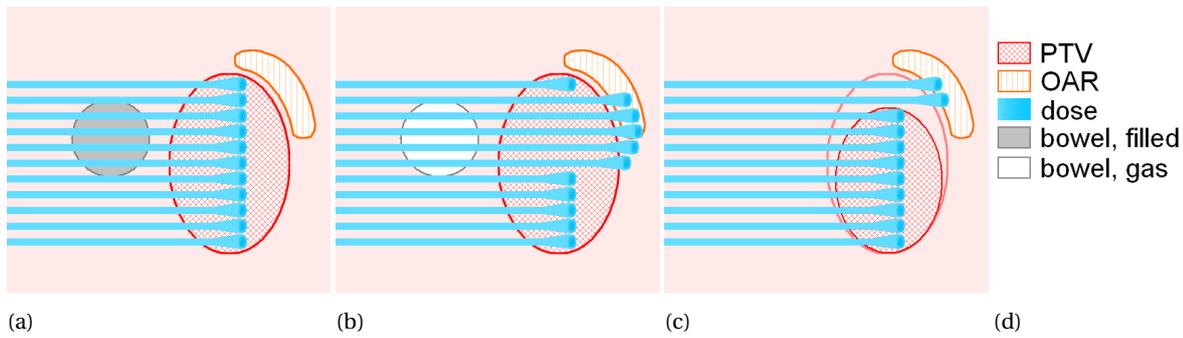


Figure 1.3: Effects due to anatomical changes. Simplified representation of dose deposited by a set of proton pencil beams. **a.** Treatment plan; **b.** Proton overshoot into an organ at risk (OAR) due to a gas bubble in the bowel; **c.** Proton overshoot into OAR due to tumor shrinkage, leaving a volume with lower density than the tumor tissue (e.g. a dense tumor in the lungs).

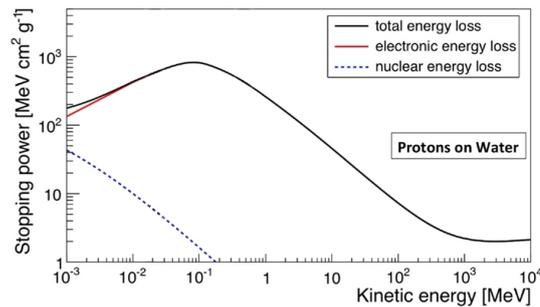


Figure 1.4: Underlying interactions of stopping power of protons in water. Nuclear interactions account for little stopping power relative to electrical interactions. Image from Kraan et al., 2015 [6], based on data from National Institute of Standards and Technology (NIST).

are currently being developed. Using these instruments, deviations in the proton range can be identified early during administering a treatment fraction instead of accounted for using large margins.

1.2. Proton interactions with matter

Proton range verification systems are based upon the physical properties of interactions between the incident particles and matter. Proton interactions can be divided into electrical and nuclear interactions. As figure 1.4 shows, the electrical interactions are the main source of energy loss for the protons.

1.2.1. Electronic interactions

Firstly, protons interact inelastically with atomic electrons. They either excite the electron or ionize the atom.

Because $m_p \gg m_e$, the deflection of the proton due to these interactions is negligible. The mass difference also causes the maximum energy transfer per interaction to be low. However, these interactions occur at high rates: for a 100 MeV proton in water, the mean free path before an inelastic electronic interaction is 0.5 nm [7]. Therefore, the accumulated energy transfer is significant and the inelastic proton-electron interactions are the main source of energy transfer to the medium.

Because the interactions occur very frequently, the statistical mean energy loss per unit path length is commonly employed to describe proton energy loss. This concept is called the continuous slowing down (CSD) approximation. It describes the stopping power (S) that depends on the proton energy (E) and the atomic number of the particle it interacts with (Z):

$$S(E, Z) = -\frac{dE}{dx} \quad (1.1)$$

The rate at which energy loss occurs is described by the Bethe-Bloch equation. From this equation follows that the relation between energy loss dE/dx and the proton velocity v is

$$\frac{dE}{dx} \propto \frac{1}{v^2}. \quad (1.2)$$

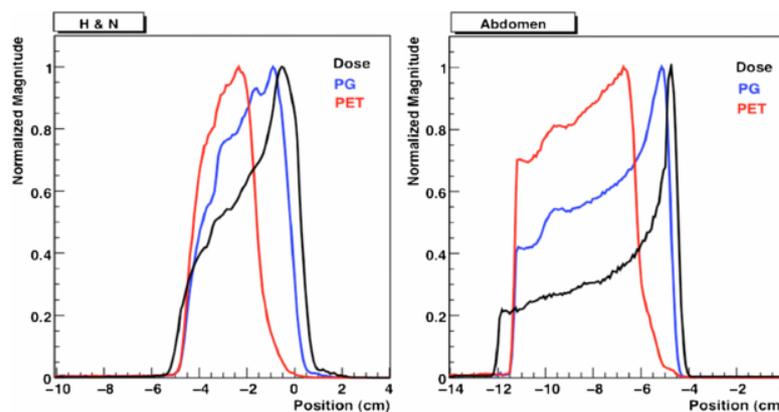


Figure 1.5: Integral depth profiles for dose, PG and PET following from patient irradiation with a pencil beam. The data are simulated using Geant4. Left image: Head and neck patient. Right image: Abdominal patient. Image from Moteabbed et al., 2011 [12].

The fact that energy loss increases as the proton loses velocity, explains the shape of the Bragg peak.

Secondly, protons interact elastically with atomic nuclei. During these reactions energy transfer is negligible, but the deflection of the proton can be significant. As these interactions occur frequently, multiple Coulomb scattering occurs. This process causes the fluence of an initially uni-directional beam spread laterally into a near-Gaussian spatial distribution.

1.2.2. Nuclear interactions

Thirdly, protons undergo inelastic interactions with atomic nuclei. These reactions cause generation of secondary particles, including secondary protons, neutrons, and gamma rays, either instantaneously or with a delay. Secondaries of special interest are positrons and prompt gammas (PG), as those provide the possibility to determine the proton range external of the patient.

Nuclear interaction rates are expressed by cross sections: the probability that an interaction occurs per unit distance propagated. The cross section depends on the type of incident particle, its energy, the target particle and the type of interaction itself.

1.2.3. Widening of the Bragg peak

Proton interactions are a stochastic process. Therefore the path length of individual protons can deviate from the calculated range. This so-called range straggling causes the Bragg peak to have a finite width.

The peak is further widened by energy straggling. The latter occurs because clinical proton beams are never strictly mono-energetic upon production, so incoming protons already have a finite energy range.

1.3. Proton range verification

Proton range verification is a method to confirm if the dose delivered to a patient corresponds to the treatment plan. If a range verification system detects an over- or undershoot, the irradiation can be stopped before dose is faultily administrated. Based on the circumstances the patient may be moved, the fractionation can be postponed or a new treatment plan must be designed.

Proton range can be determined based on direct or indirect measurement of dose. Direct in-vivo measurement of dose using implanted dosimeters was investigated [8, 9]. Though the results are promising, this method is invasive and has a problematic trade-off between the amount of units implanted and resolution.

Therefore proton range is deemed to be monitored externally, based on indirect measurements. The secondary particles produced by nuclear reactions provide an opportunity to do so. Prompt gamma imaging and positron emission tomography (PET) systems are being developed. Both methods are based on the fact that the photon profiles are (indirectly) proportional to the dose distribution (Figure 1.5). Encouraging results have been shown for both methods [10, 11].

Though the main focus of range verification research lay with PET over the past years, it has serious downsides. First of all, PET scans require acquisition times in the order of 20 minutes, making on-line PET scanning inconvenient. Both low positron emitter-related cross sections and low decay rates lead to long measurement

times; Parodi et al. report tissue-effective half-lives of positron emitters in the order of 1000s while describing PET acquisition times of 30 minutes [13].

A second downside of PET scans is the low image resolution. The positron path length in tissue is inherent to PET and causes a blur in the order of 3mm FWHM [14]. Additionally, the long acquisition times cause further decrease in image resolution through patient motion (up to 3cm) and biological washout (4mm), though these features are can at least be partially overcome [15]. However, the development of PET detectors is further developed than PG detectors.

On the other hand, prompt gamma emission occurs instantaneously and is consequently more appropriate for on-line range verification. Furthermore, interactions yielding prompt gammas have higher cross sections than PET-related reactions, which creates higher count rates. Moteabbed et al. report a PG/positron emission ratio of 60 to 70 when including time delay and wash-out in the latter case [12]. However, it must be taken into account that this comparison was based on Monte Carlo simulations which may have depended on cross sections with significant uncertainty for nuclear reactions.

As shown in figure 1.5, PG emission is not directly proportional to delivered dose. Therefore, a range verification system relies on a PG profile prediction. Numerical methods to simulate PG source terms already exist. However, the long computation time and lack of optimal settings make clinical application of full-numerical methods unlikely.

1.4. Objective

The aim of this study is to develop an alternative to full numerical methods for PG profile calculation. The current state of prompt PG simulation tools is reviewed in a literature study, which is included in this report as chapter 2.

In this study a Fortran90 code was designed to simulate PG emission of a treatment plan with good reliability, but with shorter computation times than existing methods. This code employs the Boltzmann Transport Equation to trace pencil beams individually through a target volume. The code is therefore named the Boltzmann solver (BS). The mathematical foundation of the BS is explained in chapter 3. Next the features of the Boltzmann solver code are described in chapter 4.

To validate the performance of the BS it is compared to simulations by Topas MC and the planned dose from a treatment plan, produced by the iCycle treatment planning system. The set-ups are described in chapter 5 and the results are shown in chapter 6. Lastly, the current state of the BS is discussed in chapter 7

The scope of this research includes the unscattered part of the fluence only. The BS extracts cross sections from cross section libraries. It currently employs cross sections for stopping power and PG emission; selection of libraries for absorption was considered out of scope. Therefore the validation of the BS in chapter 6 is based on dose profiles and a PG source term.

2

Methods to simulate prompt gamma profiles for proton therapy: a review

Abstract

In order to improve dose control for proton therapy, range verification systems based on prompt gamma activity are under development. These systems depend on a predicted profile of prompt gamma emission. Simulation thereof is difficult due to the lack of an integral theory to describe nuclear proton interactions. Prompt gamma profiles from the literature, produced by Monte Carlo-based simulation toolboxes like Geant4, MCNP and FLUKA, are compared and found to depend heavily on the selected physics lists and cross section libraries. Furthermore, these simulations require long computation times that make them clinically inviable. To comply with the time constraint, the calculation algorithms should include analytical solutions instead of the full numerical approach.

2.1. Introduction

Over the past years hadron therapy, especially proton therapy, has gained ground in the field of radiation oncology. The main benefits of particle irradiation as compared to photons lie in the characteristics of the dose deposition curve: the Bragg peak and finite range allow for relatively low integral dose to the patient. However, the concentrated dose deposition of the Bragg peak requires more accurate targeting of the clinical target volume (CTV). Deviations in treatment circumstances such as anatomical changes and setup variations may cause beam overshoot or undershoot, leaving parts of the CTV undertreated while adjacent healthy tissues suffer from a radiation dose that was not assigned.

To secure the prescribed radiation dose to the CTV, a safety margin is applied around it to which the same dose is assigned, thus creating the planning target volume (PTV). Depending on the applied proton range and the treatment site, this margin can be several millimeters up to over a centimeter [5]. Large margins introduce high dose to adjacent healthy tissues and may interfere with maximal dose constraints of sensitive organs at risk (OAR). Therefore, to make full clinical use of particle treatment, the margin and thus range uncertainty should be minimized.

A range verification system traces deviations from the planned dose. It would preferably be executed early in the irradiation procedure. Early detection of deviations from the treatment plan allows for alteration or termination of the current fraction, before the entire dose is administered.

Multiple approaches of range verification are investigated. Lu et al. focus on a method to determine dose directly from dosimeters implanted in the patient [8, 9]. The direct measurement is a distinct advantage of this method, but its invasive character and the limitation of resolution due to the finite amount of point measurements are serious drawbacks of this method. Range monitoring based on external measurement of secondary radiation, including positron emission tomography (PET) imaging and prompt gamma (PG) imaging, are therefore more suitable.

A Monte Carlo-based comparison between PET and PG imaging for proton range verification shows that PG can overcome fundamental resolution limitations of PET imaging and is especially preferable in cases of small treatment volumes and heterogeneous anatomies [12]. Furthermore, PG imaging can be performed on-

line, while PET has a time delay due to the half-life of β^+ -emitting nuclei, which requires long measurement times.

PG emission is not directly proportional to delivered dose. Therefore, to determine if the measured PG profile corresponds to the planned dose, the range verification system requires a prediction of the PG profile as produced by the treatment plan. The reliability of the range verification is closely related to the accuracy of the calculated profile.

To date, no default method to predict PG emission has been found to meet the requirements of a clinically feasible range verification system. The modelling of the nuclear interactions that produce PG can currently not be described in one integral theory.

This literature study aims to provide an overview of the available approaches to estimate PG profiles from treatment plans. Section 2.2 discusses the requirements of a PG profile calculation code, and sections 2.3-2.4 and 2.5 describe validation studies of numerical approaches and convolution-based methods, respectively.

2.2. Range verification system requirements

The range verification system is comprised of (1) a detector to measure the PG profile, (2) a code that predicts the PG profile based on the treatment plan and (3) a code that compares the measured and predicted profiles and states if deviations of the treatment plan occur. To be clinically viable, the range verification system should meet certain requirements as described in the following sections.

Resolution

In order to detect small deviations of the planned range, the proton beam integral verification system, including both software and hardware, should be accurate on the single or sub-millimeter scale. Current detector designs are estimated to have resolution down to 1 mm [16]. The simulation tool should therefore predict the PG profile with an accuracy up to 1mm: lower accuracy compromises the performance of the verification system, while the gain of higher accuracy is nullified by the detector resolution.

PG emission

PG imaging systems currently designed rely on energy-integrated profiles [16–18]. However, to assess existing PG simulation codes, a limited collection of energy lines are analyzed. The elements C-12, N-14, O-16 and, depending on the region of interest, Ca-40 are the most prominent PG producing elements in the body [12, 19]. The 6.13MeV gamma line produced by O-16 is of special interest as it holds the potential to be a measure of the oxygen concentration in tissues [20].

As described in section 2.3, the accuracy of the computational models may be high for certain energy ranges. Depending on the availability of detectors that have sufficient spatial and energy resolution, range verification systems that use a limited energy range may be viable. In that case, calculation methods that are accurate only in certain energy ranges would suffice, which relieves the process of validating simulation codes.

Detector response

The aim of this study is to review the reliability of PG emission of various simulation codes. However, to relate simulated PG yields and measured gamma profiles, the former must be converted into a simulated detector response. To model the detector response, other secondary particle emissions must be included, as do interactions of all secondary particles with the body, collimator and detector.

Time frame

Lastly, for the PG-based range verification to be clinically relevant, the calculation time is limited. Each treatment plan requires a tailored PG profile prediction, meaning that one or, in case of re-evaluation of the treatment plan, multiple runs must be executed per patient. Therefore, the computation time should be in the order of seconds or minutes.

2.3. Monte Carlo method

Interactions of protons in matter can be divided into electrical and nuclear interactions. Electrical interactions are well-understood and can be easily modelled using the Bethe-Bloch equation, whereas the physics of nuclear interactions has not been captured in one complete computational model to date. The challenge of creating a PG profile prediction code therefore is to model nuclear interactions well.

Because of its versatility, the Monte Carlo (MC) method holds the potential to produce reliable profiles for secondary particles produced by proton beams. MC toolboxes are comprised of various modules that each model certain particle interaction types in distinct proton energy ranges. This structure is suitable for the simulation of the interaction of proton beams in an inhomogeneous volume. MC is however a time-consuming method, especially when a large amount of beams are simulated. A trade-off must always be made between minimizing calculation time and count statistics, as the latter is related to accuracy.

Various MC packages have been employed to model patient irradiation. While MC codes are already applied for testing and optimizing medical device designs, the reliability of the emission yields is still the subject of research. Because there is no validated optimal set of modules to produce PG profiles, the range of possible combinations makes MC toolbox verification a cumbersome process.

Though various sets of modules are used, the general backbone of nuclear interaction handling is consistent. The probability of a nuclear reaction occurring is handled first. The probability of processes is determined either looking-up, based on tabulated cross sections that can be parametrized to account for varying characteristics of the interaction, or on-the-fly, i.e. based on physical models [6]. The modelling of the interaction itself can again be based on tabulated lists or physical modelling.

Cross section libraries

A range of cross section libraries which are in MC code-compatible format is available. The libraries vary not only in terms of interactions or elements included and energy ranges covered, but the values of the cross sections themselves may differ from library to library.

The general purpose library ENDF is the default alternative to physics models for toolboxes, like MCNP. While ENDF is of American origin, other countries and international collaborations have published alternative general purpose libraries.

As the tables included in these libraries are thoroughly tested, the libraries should provide reliable data. However, inter- or extrapolation may be required to find cross sections at certain energies, and because the PG yields behave volatily over the energy spectrum, this may lead to unreliable cross sections. Therefore not only the choice of library, but also the interplay between application of physical models and interpolation of cross sections from libraries affect the outcome of a numerical simulation toolkit.

Physical models

When a MC toolbox is used, the applied modules compose the physics list. Standard 'reference' physics lists are available, but alternative lists can be built from the available modules. This yields various methods to model nuclear interactions, but the general modelling of interactions is based on three consecutive mechanisms [6]:

- **Intra-nuclear cascade** The intra-nuclear cascade (INC) is associated with high-energy incoming particles, down to a few tens of MeV. In this model the medium is modelled as quasi-free nucleons, which interact with the incoming particles via a series of two-body interactions. [6]. The secondary particles produced by these interactions are traced until they reach a lower energy limit.
- **Pre-equilibrium** The second mechanism takes over calculations from the INC model after the particles reach the lower energy limit, or handles incident particles with energies below the lower limit of the INC model. For the equilibration process interactions are commonly represented based on the exciton model, which describes the formation of nucleon-hole pairs. Through emission of nucleons the nucleus ultimately reaches equilibrium, but it remains in excited state [21].
- **De-excitation** In this stage the nuclei lose their remaining energy surplus via gamma emission, Fermi break-up or, for heavier nuclei, nuclear evaporation.

Multiple MC packages are used to simulate PG profiles for proton therapy. In the following sections 2.3.1-2.3.3 the performance of Geant4, MCNP and FLUKA will be discussed.

2.3.1. Geant4

The Geant4 package was developed by CERN and its original purpose was the simulation of high-energy physics. However, its general approach of particle interaction evaluation makes it an interesting program for proton therapy calculations. Geant4 can be applied directly, but Geant4 overlays dedicated to medical radiation purposes are also available, like GATE [22] and TOPAS [23].

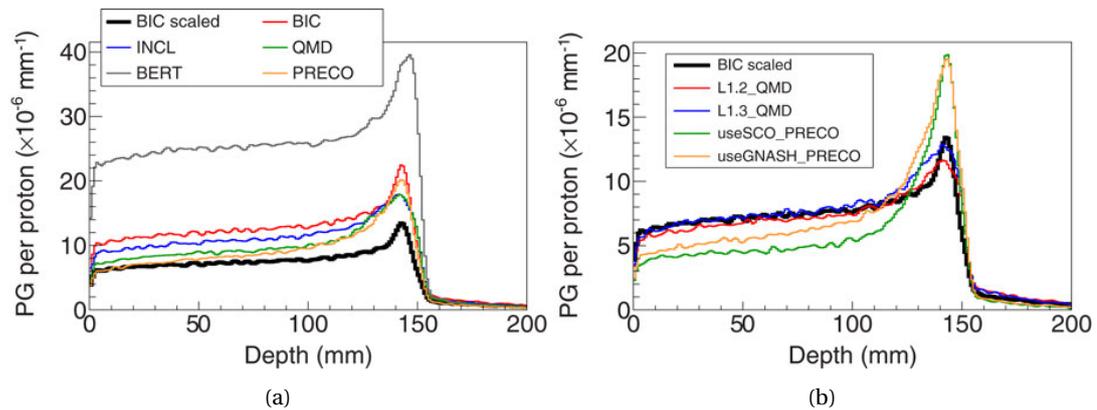


Figure 2.1: **a.** Proton yield produced by various hadronic inelastic models in Geant4 version 10.01.p02, for 160MeV proton beams with angular acceptance 1.5° . The line 'BIC scaled' is the normal BIC dataset scaled by an empirically found factor relating the cross sections from the BIC model and experimental data. Abbreviations: BIC: binary cascade, BERT: Bertini cascade, PRECO: pre-compound, INCL: Liège intranuclear cascade and QMD: quantum molecular dynamics. **b.** Changes of settings in the QMD and PRECO models cause better agreement between simulated and measured data. Both images from Pinto et al., 2016 [27].

The dose profile calculation performance of Geant4 and related programs has been thoroughly tested and validated [24, 25]. Meanwhile, the wide range of modules associated with nuclear reactions and settings confound the evaluation of PG production reliability.

Early PG verification research concluded that Geant4 severely overestimated PG yields [26]. Le Foulher et al. found an overestimation by a factor of 12 using Geant4 v9.1. More recent versions of the Geant4 toolkit include improved models, but the simulation reliability still relies heavily on the chosen parameters. This was illustrated by Pinto et al., who simulated the PG yield in a homogeneous PMMA phantom using five different proton inelastic models, using the default settings of each module (figure 2.1a) [27]. The PG yields vary in magnitude, but they all overestimate the reference data. However, each model produced the same range of PG emission. Therefore, whether the produced profiles are useful for range verification depends on which characteristic is used to compare the predicted and measured PG profiles: in terms of absolute emission rates, the models do not suffice, but the range or shape of the profile are well-predicted.

To improve the absolute emission rates, the settings of two modules were changed (figure 2.1b). These yields resembled the reference data better, showing that both choice of model and subsequent settings determine the simulation reliability.

In some cases the complexity and multitude of available settings inspire to vary variables at random until a simulation's outcome agrees with reference data. Lestand et al. demonstrated this when selecting a tolerance value [28]. The tolerance value applies on looking-up calculations: only when the difference between excitation energy and an energy level from a nuclear database is below the tolerance value, a radiative transition can take place. Because there was no physical foundation for this variable, a range of values was tested to investigate the impact of the tolerance value. PG depth profiles produced by carbon beams were simulated and compared to a measured profile. At 100keV the emission yields matched best, so this value was adopted in further simulations to investigate the feasibility of imaging methods.

Both Pinto et al. and Lestand et al. published energy-integrated PG yields, without reporting on energy dependence of their results. The fact that the reliability of PG profile simulation may vary over photon energies is shown by Schumann et al. [29]. They performed an energy spectrum analysis using simulated (Geant4 v10.00.p01, using a binary cascade (BIC) based set-up) and experimental data of a PMMA phantom with a germanium detector (Figure 2.2). Most major peaks in the experimental cross sections were reproduced in the simulations. However, in terms of absolute PG yields, the simulation overestimated the PG production by 37%. The majority of the overestimation originated from PG with energies below 4.5MeV. For higher energies, the cross sections matched experimental data better.

The homogeneous PMMA phantom is a recurrent target for validation runs [27, 28]. Schumann et al. provided some variation by analyzing the performance of Geant4 on polyethylene, graphite and a cortical bone equivalent material, SB3, benchmarking the MC code against experimental data [29]. The simulations showed overestimations of 38%, 48% and 38%, respectively, for the gamma yield for $3\text{MeV} \leq E \leq 7\text{MeV}$.

The wide range of modules and subsequent setting possibilities establish the versatility that is the strength of Geant4, but it confounds the design of a reliable medical planning tool. By adjusting the settings until sim-

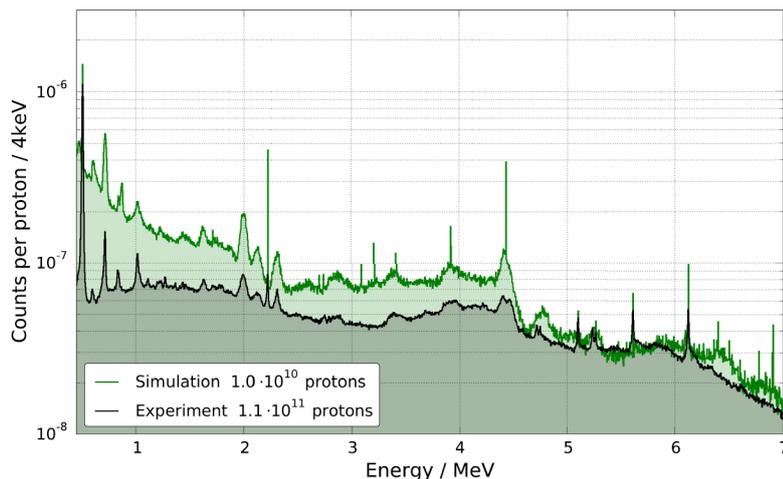


Figure 2.2: Comparison of PG spectra from simulated and measured data; 160MeV beam in PMMA. The yields for γ energies over 4.5MeV correspond better than lower energies. Image from Schumann et al., 2015 [29].

ulations match experimental data, quasi-reliability may be established [27, 28]. Therefore, attempts have been made to produce toolkits of which the defaults produce reliable outcomes for medical irradiation calculations. These have resulted in GATE and TOPAS.

GATE is a community-driven toolkit developed specifically for medical imaging and radiation. Like the underlying Geant4 code, the dose and range accuracy were already found to be simulated accurately based on depth-dose analysis. However, lateral spreading was found to be underestimated [30].

GATE was originally designed to model PET and SPECT [22]. Hence, few reports of applications on PG imaging exist. However, Gueth et al. performed a series of GATE simulations to identify valuable classifiers for a machine-learning approach to range verification [17]. The simulated data were not evaluated using any benchmark so this publication did not contribute to the validation of PG emission by GATE.

The publication of Gueth et al. is unique in the sense that the computation time was reported: a single spot consisting of $50 \cdot 10^6$ protons took 5.2h [17]. Given the time constraint from section 2.2, this simulation is no viable component for a range verification system. However, Huisman et al. found significant time reduction can be achieved [31]. They used GATE to produce a spatial and spectral distribution of PG yields based on the low-statistical simulation of a proton beam. The emission profile is normalized for a single incident proton can be scaled up to the desired beam intensity, thus producing a full-intensity emission profile at low computational cost. The propagation of the produced PG is handled in a separated step.

The creators of TOPAS aimed to produce a user-friendly tool with a default physics list that is tailored to simulate proton therapy. Like GATE, TOPAS is built as an overlay of Geant4. Upon presentation, TOPAS was validated for dose deposition only [23].

Testa et al. performed a validation study of PG energy profiles in a series of set-ups that contained different collimator configurations [32]. They used the default TOPAS physics list based on Geant4 v9.6.p02 to simulate a 160MeV pencil beam in a PMMA phantom. The PG profile was scored using simulated NaI detectors. The simulated PG spectra were benchmarked against experimental data. The simulation underestimated the PG emission in a geometry without collimator by 10%, while addition of collimators increased the underestimation up to 50%. Testa et al. state that the underestimation is partly due to lack of objects in the simulation, while in experimental measurements background radiation and scatter from surrounding objects was scored as well. However, when related to results produced by the underlying Geant4 code as produced by e.g. Schumann et al., a 10% deviation from measured data is a significant improvement. [29].

2.3.2. MCNP/MCNPX

The MCNP code is a general purpose MC code. It is specifically designed for neutron, photon and electron transport. MCNPX, in which the "X" stands for "eXtended", was created by adding the cross sections other particle interactions, and on-the-fly calculations were included to be applied when no cross section libraries are available [33].

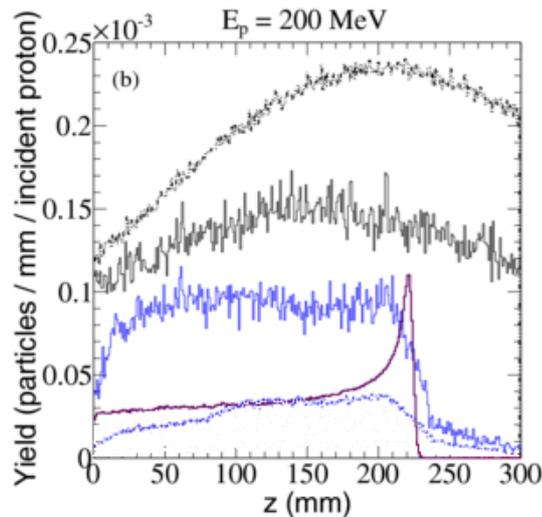


Figure 2.3: Lateral profile of yields for γ emission with angular acceptance $90^\circ \pm 3^\circ$ (blue) and neutrons (black), as produced from a 200 MeV proton beam in a homogeneous PMMA phantom. Data were obtained from Geant4 (line) and MCNPX (scatter). The integral depth dose is added in purple. Image from Biegun et al., 2012 [34].

The first validation for a MCNPX code in the field of medical PG emission was performed by Smeets et al. [16]. During their research on optimal slit camera characteristics, MCNPX version 2.5.0 was used to produce PG profiles from a 160 MeV proton beam in a PMMA phantom. The simulations employed a looking-up structure using the la150n and la150h cross section libraries for proton and neutron transport at energies below 150 MeV, and the BIC model otherwise. The PG were scored on a perfect, cylindrical scintillator. The simulated counts were compared to measured data from a similar geometry.

It was found that in an open environment without collimator, the MCNPX toolbox underestimated the PG emission by 19.3% and 26.9% for distances of 50 cm and 100 cm between the beam axis and scintillator, respectively. The largest count differences were measured in the energy range of 0-3 MeV. When including only PG in the energy range of 3-6 MeV, the underestimation was reduced to 8.9% and 16.5%, respectively.

Performance of both Geant4 and MCNPX were analyzed by Biegun et al. [34]. In their paper on the application of time-of-flight restriction, PG depth profiles simulated by MCNPX 2.7.D and Geant4 v9.2.p02 were compared [34]. The setup contained a PMMA phantom, which was irradiated with 100 MeV and 200 MeV protons. The MCNPX code employed cross section libraries for proton and neutron energies below 150 MeV (la150), and physics models for other particle types and energy ranges. The Geant4 code was tested with two physics lists which were found to yield equal results. Despite the effort to select a reliable set of modules, the PG yields found by the Geant4 simulation were a factor 2 to 5 higher than those simulated by MCNPX (Figure 2.3). No experimental data were included in this research, so the simulations were only compared among themselves.

2.3.3. FLUKA

Like MCNP, FLUKA is a particle tracking toolkit for general purpose. Evaluation of dose delivery simulations were performed [35], but validation of the quality of PG profiles is ongoing.

Battistoni et al. tested FLUKA's performance for PG emission using multiple light ions and PMMA targets, using measured data as benchmark [36]. Simulations and measurements were done with an open geometry, with a collimator and with a wall, i.e. a closed collimator, between the target and detector. PG data for 160 MeV proton beams were reported as "opening difference" and "wall difference" energy spectra only, representing the difference in measured PG between the collimator and the wall, and the open geometry and the wall, respectively. For both cases the lower energies showed the largest discrepancies. However, when only PG with $E > 2$ MeV are involved, the simulated yields lie within 10% of the experimental data, which Battistoni et al. were content with. The simulated spectral distributions are shown in figure 2.4. A systematic energy shift of the yields is visible in the figure, but remains undiscussed by Battistoni et al.

Robert et al. performed a series of comparisons between FLUKA and GATE for proton and carbon irradiation of a PMMA target without experimental benchmark [37]. For a 134 MeV proton beam the yield of γ with $E > 1$ MeV produced by GATE was higher than FLUKA by a factor 1.92. A following analysis of the energy spec-

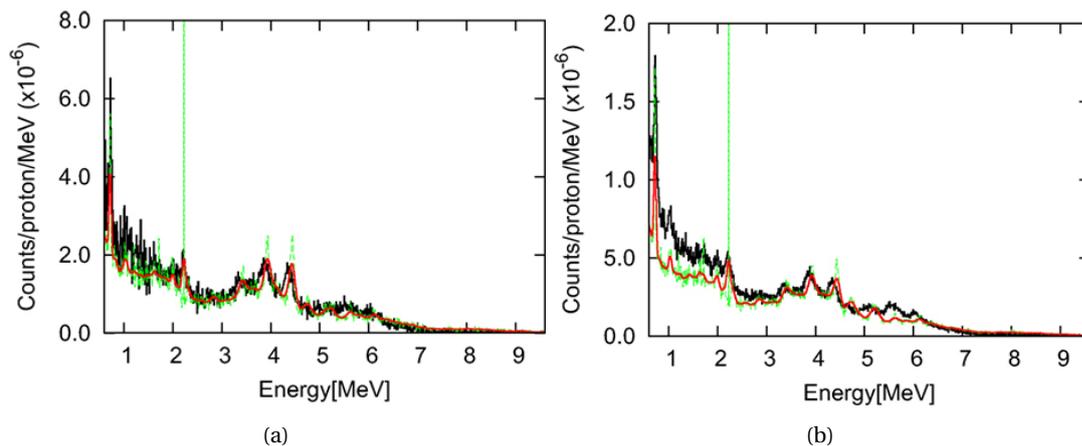


Figure 2.4: Background-subtracted photon energy spectra, produced by 160MeV protons in PMMA, simulated by FLUKA. Experimental data in black, simulated data in red (with intrinsic detector resolution) and green (without intrinsic detector resolution). **a.** Opening difference and **b.** wall difference. Image from Battistoni et al., 2016 [36].

trum shows that GATE systematically estimates higher yields, though the gamma lines near 4.44MeV (C-12), 5.21 (O-15) and 6.13MeV (O-16) show smaller deviations.

2.4. Dedicated nuclear interaction codes

An alternative for MC codes are dedicated nuclear interaction codes. These codes calculate cross sections over wide energy regions, based on both physics models and data from cross section libraries. They however do not contain the spatial tracing algorithm of MC codes. In this section the dedicated nuclear reaction codes TALYS and EMPIRE are discussed.

TALYS aims to provide reliable simulations of nuclear reactions for incident particles in the 1keV-200MeV region, for target nuclei with atomic mass number 12 and up [38]. The incident energy matches that of proton therapy, and the limitation of atomic numbers allows for the most prominent PG lines, so TALYS may be a good fit to simulate proton therapy.

EMPIRE has, like TALYS, a wide incident particle energy range that reaches from 1keV to hundreds of MeV [39]. Because no lower limit of the atomic mass numbers for target nuclei is reported, EMPIRE may be advantageous relative to TALYS.

The fact that the cross sections measured from experiments, tabulated in cross section libraries and simulated by MC codes and dedicated nuclear reaction codes vary significantly is demonstrated by the work of Verburg et al. [19]. The cross sections for eight gamma lines from C-12, N-14 and O-16 as produced by Geant4 v9.5, MCNP6 beta2, TALYS 1.4 and EMPIRE 3.1 were compared to experimentally determined cross sections reported in the literature and the ENDF/B-VII cross section library. Figure 2.5 shows the comparison of all cross sections for two gamma lines, illustrating how each method yields different results.

2.5. Convolution-based methods

As the aforementioned simulation methods require long calculation times, the time-related prerequisite mentioned in Section 2.2 is not met. Alternatively, quicker methods based on convolution have been applied for dose deposition calculations and PET modelling.

For example, Parodi et al. relied on pre-calculated dose profiles [40]. They developed a kernel function based on a Gaussian distribution and a power function, which was convolved with a dose profile in order to generate a PG profile. The method was benchmarked with both FLUKA simulations (for an inhomogeneous target) and measured data (for a homogeneous target). The convolution method was able to identify the distal positron emission fall-off location with sub-millimetre precision based on emission from C-11. However, the total positron intensity had a 10% deviation from the measured intensity.

The performance of the method developed by Parodi et al. depends on the quality of the dose profile. Miyatake et al. proposed an alternative by developing a kernel that expressed the yields of the most prominent positron emitters in the body, C, O and Ca, for a proton pencil beam [41]. The kernels were based on measured activity from various homogeneous phantoms and three different proton energies. The following

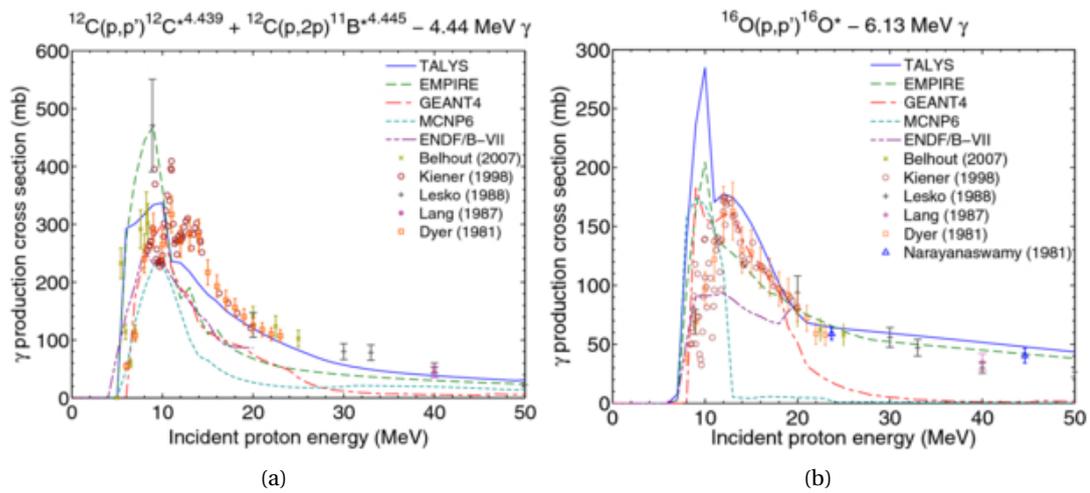


Figure 2.5: PG cross sections for two gamma lines, from dedicated nuclear codes, MC codes, the ENDF/B-VII cross section library and experimental data from literature. **a.** Cross sections for 4.44MeV PG production from C-12 **b.** Cross sections for 6.13MeV PG production from O-16. Images from Verburg et al., 2012 [19].

convolution method would be very similar to dose calculations performed by treatment planning systems, and requires roughly equal calculation time, i.e. in the order of minutes. The aim of the study was the mere development of kernels, so no comparison with experimental profiles was made.

Applications of similar convolution methods for PG imaging have not been reported on. However, given the results of PET simulations, convolution methods are a promising approach.

2.6. Conclusion

Proton range verification systems rely on a tool which predicts PG emission accurately with computation time in the order of minutes. Due to the lack of an integral theory, nuclear interactions require a patchwork of models.

MC-based simulation codes are versatile because of the availability of various toolboxes, modules, cross section libraries and settings. They therefore hold the potential to produce reliable PG profiles.

However, a thorough validation study of MC toolboxes has not been performed to date. The literature contains many studies in which the results from a certain physics list are benchmarked against experimental data, but these comparisons are usually limited to one proton beam energy and a homogeneous target. Therefore, no reliable default combination of settings for PG production has been determined.

A second problem lies with the calculation time, which is usually not reported on. A full treatment plan consists of roughly 10^{12} protons, which requires calculation time in the order of days. For PG imaging to be clinically feasible, a method to drastically reduce the amount of traced particles must be applied. Methods to reduce calculation time could entail reducing the amount of pencil beams used for range verification or replacing part of the simulation with analytical solutions.

An alternative to the full numerical approach is the convolution-based method. The development of convolution kernels has already been applied for dose and PET simulations, but no evidence of application for PG imaging has been reported to date.

3

Derivation of the proton transport equation

The Boltzmann equation originally described dynamics of particles in a gas. However, it can be applied to calculate the statistical distribution of a type of particle in a fluid in general. The equation describes the rate of change in the number of particles in a volume.

The linear Boltzmann equation (LBE) is derived for general particles in section 3.1. The adjustments necessary to apply the LBE to proton therapy are discussed in section 3.2. Next, the equation that describes the unscattered flux is isolated in section 3.3. Lastly, the partial differential equation derived in the previous sections is converted into a set of ordinary differential equations in section 3.4. These equations can be evaluated by the Boltzmann solver code.

3.1. Linear Boltzmann equation

The LBE describes the behaviour of $n(\mathbf{r}, E, \hat{\Omega}, t) dV dE d\hat{\Omega}$. This is the amount particles in the infinitesimal volume dV , with energy dE about E and direction $d\hat{\Omega}$ about $\hat{\Omega}$, at time t . Energy levels that do not lie within dE about E are indicated with E' and directions not in $d\hat{\Omega}$ about $\hat{\Omega}$ are indicated with $\hat{\Omega}'$.

The value of n varies over time due to particle production, streaming into and out of the volume and interactions that change E and $\hat{\Omega}$. The LBE was derived from these processes by Duderstadt and Hamilton [42]. The derivation below follows their method.

The gain and loss of n in an arbitrary, finite volume V , which is enclosed by surface \mathbf{S} , is described using five mechanisms. The mathematical description of these mechanisms form the foundation of the LBE.

$$\frac{\partial}{\partial t} \int_V n(\mathbf{r}, E, \hat{\Omega}, t) dV dE d\hat{\Omega} = a + b + c - d - e; \quad (3.1)$$

- Gain
 - a Amount of particles produced by sources within V .
 - b Amount of particles streaming through \mathbf{S} into V .
 - c Amount of particles in V with E' and $\hat{\Omega}'$, undergoing an interaction so the new energy level lies within dE about E and $d\hat{\Omega}$ about $\hat{\Omega}'$. This process is described as $E' \rightarrow E$; $\hat{\Omega}' \rightarrow \hat{\Omega}$.
- Loss
 - d Amount of particles streaming through \mathbf{S} out of V .
 - e Amount of particles in V with $E, \hat{\Omega}$ undergoing a collision, so $E \rightarrow E'$; $\hat{\Omega} \rightarrow \hat{\Omega}'$.

To express mechanism a , the source density s is introduced. The quantity s is defined as the production rate of particles with applicable $E, \hat{\Omega}$ from a source that lies within volume dV . The gain of n in V is then

$$a = \left[\int_V s(\mathbf{r}, E, \hat{\Omega}, t) dV \right] dE d\hat{\Omega}. \quad (3.2)$$

Secondly, the mechanisms that describe streaming into and out of V are considered. The inflow of mechanism b and outflow of d are collected into one net outflow term: $d - b$. To describe streaming, the scalar

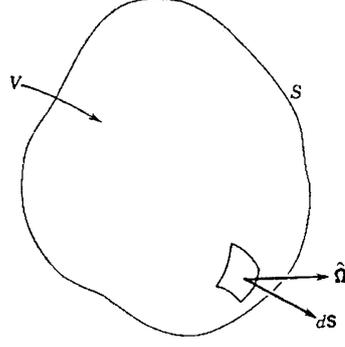


Figure 3.1: An arbitrary volume V with surface area \mathbf{S} . A differential element is indicated with its unit normal vector $d\mathbf{S}$ and a net outflow direction $\hat{\mathbf{\Omega}}$. Image by Duderstadt and Hamilton, 1976 [42].

angular particle flux ϕ is introduced:

$$\phi(\mathbf{r}, E, \hat{\mathbf{\Omega}}, t) \equiv v n(\mathbf{r}, E, \hat{\mathbf{\Omega}}, t), \quad (3.3)$$

where v is the particle speed. When flowing into or out of V , surface \mathbf{S} is passed. In this derivation, the convention is that $d\mathbf{S}$ is the unit vector normal to the surface (figure 3.1). The total streaming term is the angular particle flux ϕ with direction $\hat{\mathbf{\Omega}}$ passing through surface \mathbf{S} :

$$d - b = \left[\int_{\mathbf{S}} d\mathbf{S} \cdot \hat{\mathbf{\Omega}} \phi(\mathbf{r}, E, \hat{\mathbf{\Omega}}, t) \right] dE d\hat{\mathbf{\Omega}}.$$

The surface integral is rewritten into a volume integral using Gauss theorem, which states that $\int_{\mathbf{S}} \mathbf{F} d\mathbf{S} \equiv \int_V (\nabla \cdot \mathbf{F}) dV$. The net outflow term now becomes

$$d - b = \left[\int_V dV \nabla \cdot \hat{\mathbf{\Omega}} \phi(\mathbf{r}, E, \hat{\mathbf{\Omega}}, t) \right] dE d\hat{\mathbf{\Omega}}. \quad (3.4)$$

The third gain mechanism c describes the amount of particles undergoing scatter interactions that cause both $E' \rightarrow E$ and $\hat{\mathbf{\Omega}}' \rightarrow \hat{\mathbf{\Omega}}$. The probability of such interactions per unit length is given by the differential cross section Σ_s , and the interaction rate is $\Sigma_s \phi dE d\hat{\mathbf{\Omega}}$. This term includes particles with any energy E' and angle $\hat{\mathbf{\Omega}}'$. Therefore Σ_s is integrated over both E' and $\hat{\mathbf{\Omega}}'$:

$$c = \left[\int_V dV \int_{4\pi} d\hat{\mathbf{\Omega}}' \int_0^\infty dE' \Sigma_s(E' \rightarrow E, \hat{\mathbf{\Omega}}' \rightarrow \hat{\mathbf{\Omega}}) \phi(\mathbf{r}, E', \hat{\mathbf{\Omega}}', t) \right] dE d\hat{\mathbf{\Omega}}. \quad (3.5)$$

For the final loss term, e , the amount of particles that undergo interactions that cause $E \rightarrow E'$ and $\hat{\mathbf{\Omega}} \rightarrow \hat{\mathbf{\Omega}}'$ are determined. The collision rate is described using the cross section for collision, Σ_t :

$$e = \left[\int_V \Sigma_t(\mathbf{r}, E) \phi(\mathbf{r}, E, \hat{\mathbf{\Omega}}, t) \right] dE d\hat{\mathbf{\Omega}}. \quad (3.6)$$

Putting the mathematical descriptions of mechanisms $a - e$ together, the equation reads

$$\begin{aligned} \int_V \frac{\partial n(\mathbf{r}, E, \hat{\mathbf{\Omega}}, t)}{\partial t} dV dE d\hat{\mathbf{\Omega}} = & \int_V dV \left[\{s(\mathbf{r}, E, \hat{\mathbf{\Omega}}, t)\} - \{\nabla \cdot \hat{\mathbf{\Omega}} \phi(\mathbf{r}, E, \hat{\mathbf{\Omega}}, t)\} + \right. \\ & \left. \left\{ \int_{4\pi} d\hat{\mathbf{\Omega}}' \int_0^\infty dE' \Sigma_s(E' \rightarrow E, \hat{\mathbf{\Omega}}' \rightarrow \hat{\mathbf{\Omega}}) \phi(\mathbf{r}, E', \hat{\mathbf{\Omega}}', t) \right\} - \right. \\ & \left. \left\{ \Sigma_t(\mathbf{r}, E) \phi(\mathbf{r}, E, \hat{\mathbf{\Omega}}, t) \right\} \right] dE d\hat{\mathbf{\Omega}}, \end{aligned}$$

and rearranging all terms to the left hand of the equation gives

$$\begin{aligned} \int_V dV \left[\frac{\partial n(\mathbf{r}, E, \hat{\mathbf{\Omega}}, t)}{\partial t} - \{s(\mathbf{r}, E, \hat{\mathbf{\Omega}}, t)\} + \{\nabla \cdot \hat{\mathbf{\Omega}} \phi(\mathbf{r}, E, \hat{\mathbf{\Omega}}, t)\} - \right. \\ \left. \left\{ \int_{4\pi} d\hat{\mathbf{\Omega}}' \int_0^\infty dE' \Sigma_s(E' \rightarrow E, \hat{\mathbf{\Omega}}' \rightarrow \hat{\mathbf{\Omega}}) \phi(\mathbf{r}, E', \hat{\mathbf{\Omega}}', t) \right\} + \left\{ \Sigma_t(\mathbf{r}, E) \phi(\mathbf{r}, E, \hat{\mathbf{\Omega}}, t) \right\} \right] dE d\hat{\mathbf{\Omega}} = 0. \end{aligned}$$

As V was arbitrarily chosen and the equation must hold for any volume, the integrand over V must be 0. By rearranging the integrand and applying equation 3.3 on the first term, the LBE appears:

$$\frac{1}{v} \frac{\partial \phi(\mathbf{r}, E, \hat{\Omega}, t)}{\partial t} + \left\{ \nabla \cdot \hat{\Omega} \phi(\mathbf{r}, E, \hat{\Omega}, t) \right\} + \left\{ \Sigma_t(\mathbf{r}, E) \phi(\mathbf{r}, E, \hat{\Omega}, t) \right\} = \left\{ s(\mathbf{r}, E, \hat{\Omega}, t) \right\} + \left\{ \int_{4\pi} d\hat{\Omega}' \int_0^\infty dE' \Sigma_s(E' \rightarrow E, \hat{\Omega}' \rightarrow \hat{\Omega}) \phi(\mathbf{r}, E', \hat{\Omega}', t) \right\}. \quad (3.7)$$

3.2. Application to proton therapy

The general LBE can be applied to evaluate proton transport for proton therapy after minor adjustments. Firstly, no proton sources exist in the target volume, so term a can be eliminated. Secondly, as the proton flux reaches steady state almost instantaneously during irradiation, dn/dt can be considered to be 0 and all terms become time-independent. Lastly, the energy-changing interactions in term c will only reduce energy, so $E' > E$. Therefore the integral over dE' is performed on the interval $[E, \infty]$. The remaining equation reads

$$\left\{ \nabla \cdot \hat{\Omega} \phi(\mathbf{r}, E, \hat{\Omega}) \right\} + \left\{ \Sigma_t(\mathbf{r}, E) \phi(\mathbf{r}, E, \hat{\Omega}) \right\} = \int_{4\pi} d\hat{\Omega}' \int_E^\infty dE' \Sigma_s(E' \rightarrow E, \hat{\Omega}' \rightarrow \hat{\Omega}) \phi(\mathbf{r}, E', \hat{\Omega}'). \quad (3.8)$$

The next step is to separate the cross sections of equation 3.8 into three components each. The first component Σ_{in} , accounts for inelastic scatter, which causes slowing down of protons. Secondly, elastic scatter, which causes angular deflection, is described by Σ_{el} . The last term, Σ_{nu} , accounts for nuclear interactions, which cause both energy reduction and angular deflection. This split yields the following:

$$\begin{aligned} \nabla \cdot \hat{\Omega} \phi(\mathbf{r}, E, \hat{\Omega}) = & \int_E^\infty dE' \Sigma_{s,in}(E' \rightarrow E, \hat{\Omega}) \phi(\mathbf{r}, E', \hat{\Omega}) - \Sigma_{t,in} \phi(\mathbf{r}, E, \hat{\Omega}) + \\ & \int_{4\pi} d\hat{\Omega}' \Sigma_{s,el}(E, \hat{\Omega}' \rightarrow \hat{\Omega}) \phi(\mathbf{r}, E, \hat{\Omega}') - \Sigma_{t,el} \phi(\mathbf{r}, E, \hat{\Omega}) + \\ & \int_{4\pi} d\hat{\Omega}' \int_E^\infty dE' \Sigma_{s,nu}(E' \rightarrow E, \hat{\Omega}' \rightarrow \hat{\Omega}) \phi(\mathbf{r}, E', \hat{\Omega}') - \Sigma_{t,nu} \phi(\mathbf{r}, E, \hat{\Omega}). \end{aligned} \quad (3.9)$$

3.3. Fokker-Planck approximation

Solving equation 3.9 is complicated because of the behaviour of the cross sections for elastic and inelastic interactions. The former has a high maximum as $\Delta\Omega \rightarrow 0$ (figure 3.2), while the latter peaks as $\Delta E \rightarrow 0$. To avoid the complexity of describing these cross sections, the Fokker-Planck approximation is applied. The approximation addresses the terms describing small energy transfers (Σ_{in}) and small angular deflections (Σ_{el}) by two operators.

The two terms involving inelastic scatter are simplified using the fact that the differential cross section for energy transfer increases as the incident proton energy decreases. Therefore the energy dependence can be estimated by a Taylor expansion around $E = 0$:

$$\int_E^\infty dE' \Sigma_{s,in}(\mathbf{r}, E' \rightarrow E) \phi(\mathbf{r}, E', \hat{\Omega}) - \Sigma_{t,in} \phi(\mathbf{r}, E, \hat{\Omega}) \approx \frac{\partial S(\mathbf{r}, E) \phi(\mathbf{r}, E, \hat{\Omega})}{\partial E} + \frac{1}{2} \frac{\partial^2 T(\mathbf{r}, E) \phi(\mathbf{r}, E, \hat{\Omega})}{\partial E^2}. \quad (3.10)$$

Here the first term of the right hand side represents the slowing down of incident protons, using the stopping power $S(\mathbf{r}, E)$. The second term can be interpreted physically as energy straggling, with $T(\mathbf{r}, E)$ a diffusion coefficient that represents the measure of straggling.

Secondly, the terms involving elastic scatter are approximated by the continuous scatter operator. For the full derivation hereof the reader is referred to the work of Leakeas and Larson [44]. In this equation Σ_{tr} is the momentum transfer cross section and ∇_Ω^2 represents the angular part of the Laplace operator.

$$\int_{4\pi} d\hat{\Omega}' \Sigma_{s,el}(E, \hat{\Omega}' \rightarrow \hat{\Omega}) \phi(\mathbf{r}, E, \hat{\Omega}') - \Sigma_{t,el} \phi(\mathbf{r}, E, \hat{\Omega}) \approx \frac{1}{2} \Sigma_{tr}(\mathbf{r}, E) \nabla_\Omega^2 \phi(\mathbf{r}, E, \hat{\Omega}). \quad (3.11)$$

When put together, the PTE for proton therapy reads

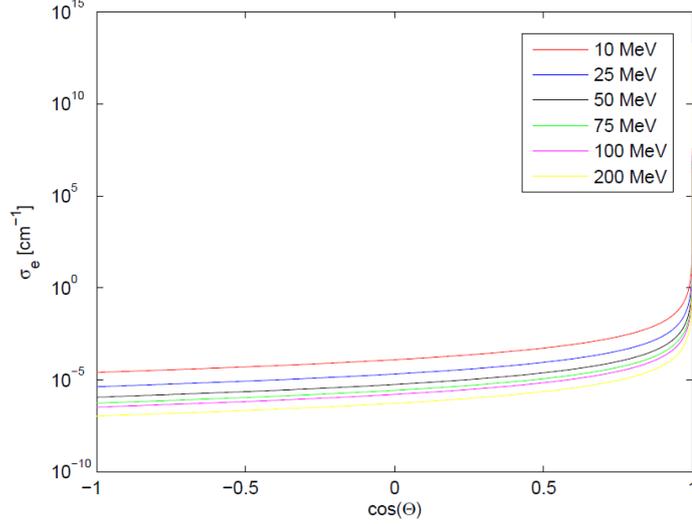


Figure 3.2: Angular dependence of the cross section for elastic electrical interactions of protons in water. The cross section is nearly singular for $\cos\theta$, i.e. $\Delta\mathbf{\Omega}$. Image by Uilkema, 2012 [43].

$$\begin{aligned} \nabla \cdot \hat{\mathbf{\Omega}}\phi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) \approx & \frac{\partial S(\mathbf{r}, E)\phi(\mathbf{r}, E, \hat{\mathbf{\Omega}})}{\partial E} + \frac{1}{2} \frac{\partial^2 T(\mathbf{r}, E)\phi(\mathbf{r}, E, \hat{\mathbf{\Omega}})}{\partial E^2} + \\ & \frac{1}{2} \Sigma_{tr}(\mathbf{r}, E) \nabla_{\Omega}^2 \phi(\mathbf{r}, E, \hat{\mathbf{\Omega}}) + \\ & \int_{4\pi} d\hat{\mathbf{\Omega}}' \int_E^{\infty} dE' \Sigma_{s,nu}(E' \rightarrow E, \hat{\mathbf{\Omega}}' \rightarrow \hat{\mathbf{\Omega}}) \phi(\mathbf{r}, E', \hat{\mathbf{\Omega}}') - \Sigma_{t,nu} \phi(\mathbf{r}, E, \hat{\mathbf{\Omega}}). \end{aligned} \quad (3.12)$$

To solve ϕ , equation 3.12 is split into an unscattered and a scattered part. The unscattered flux ϕ_u can be solved using Σ_t and Σ_s :

$$\nabla \cdot \hat{\mathbf{\Omega}}\phi_u(\mathbf{r}, E, \hat{\mathbf{\Omega}}) + \Sigma_t \phi_u(\mathbf{r}, E, \hat{\mathbf{\Omega}}) = \frac{\partial S(\mathbf{r}, E)\phi_u(\mathbf{r}, E)}{\partial E}. \quad (3.13)$$

Then, using ϕ_u , ϕ_s can be solved. The total flux is $\phi_{total} = \phi_u + \phi_s$.

3.4. Method of characteristics

In this section the method of characteristics is applied to equation 3.13 to create a set of ordinary differential equations (ODE) that can be solved by the Boltzmann solver. To simplify this process, the equation for unscattered flux is reduced to one spatial dimension. This is possible because the Boltzmann solver only uses uni-directional pencil beams. In 1D equation 3.13 becomes

$$\frac{\partial \phi_u(x, E)}{\partial x} + \Sigma_t \phi_u(x, E) = \frac{\partial S(x, E)\phi_u(x, E)}{\partial E}. \quad (3.14)$$

The solutions to the partial differential equation (PDE) that is equation 3.14 lie on a surface graph $\phi(x, E)$. The method of characteristics presents a curve, called the characteristic, on which the PDE becomes an ODE. This curve develops along the characteristic variable q , so that $x(q)$ and $E(q)$.

To evaluate equation 3.14 over the characteristic, it is first expanded using the chain rule:

$$\frac{\partial \phi_u(x, E)}{\partial x} - \frac{\partial \phi_u(x, E)}{\partial E} S(x, E) = \left(\frac{\partial S(x, E)}{\partial E} - \Sigma_t \right) \phi_u(x, E). \quad (3.15)$$

Then, using equation 3.15 and

$$\frac{d\phi_u}{dq} = \frac{d\phi_u}{dx} \frac{dx}{dq} + \frac{d\phi_u}{dE} \frac{dE}{dq},$$

the Lagrange-Charpit equations emerge:

$$\frac{dx}{dq} = 1 \quad (3.16)$$

$$\frac{dE}{dq} = -S(x, E) \quad (3.17)$$

$$\frac{d\phi_u}{dq} = \left(\frac{\partial S(x, E)}{\partial E} - \Sigma_t \right) \phi_u(x, E) \quad (3.18)$$

The last equation, 3.18, seems counter-intuitive, because it states that the number of particles decreases due to interactions that only underlie slowing down. However, the mono-energeticity of the pencil beams offers a solution. The energy dependence of ϕ_u is a single delta peak at energy level E . $\int \phi_u dE$ is thus constant, and stopping power only causes a shift of the delta peak over E . Therefore, equation 3.18 can be re-written as

$$\frac{d\phi_u(x)}{dq} = -\Sigma_t \phi_u(x) \quad (3.19)$$

4

Boltzmann solver code

This chapter describes the components of the BS code. The BS commences with the set-up of tracing volume and beams, which is described in section 4.1. Then the beams are traced through the volume. Over the tracing, the flux is evaluated using the ODEs as derived in section 3.4. The implementation of these ODEs into the BS is described in section 4.2. Next, two possible integration methods are discussed in section 4.3. Section 4.4 describes two options for tracing algorithms, from which 3D distributions for dose and PG emission can be calculated. Lastly, methods to reduce calculation time are elaborated on in section 4.5.

4.1. Initialization

The BS initiates a simulation by the interpretation of its input data. The input consists of the planning CT and a list of proton beam characteristics. The conversion of the planning CT into the simulation grid is described in section 4.1.1. The handling of the beams is elaborated on in section 4.1.2.

4.1.1. Tracing volume

The BS traces proton beams through a 3D grid, which is based on the planning CT scan. The code translates the Hounsfield Units (HU) into two tissue characteristics that dictate interaction features: mass density ρ and a component vector \mathbf{f} . \mathbf{f} is an array of elemental weight fractions (f_i) for the elements that are abundant in human tissues.

In the BS the stoichiometric calibration as proposed by Schneider et al. is applied [45]. They describe the handling of HU in the interval [-1000,1600]. If the CT scan contains values outside this region they are overwritten by -1000 and 1600 for lower and higher HU, respectively. Furthermore, this method creates \mathbf{f} of 12 elements: H, C, N, O, Na, Mg, P, S, Cl, Ar, K, Ca. Argon is not present in human tissues, but it is a component of air.

When protons travel through the volume, the stopping power, absorption and PG emission depend on the proton energy E , ρ and \mathbf{f} . These processes are handled by a set of cross section libraries. The current version of the BS includes cross sections for stopping power and PG yields.

Because the stopping power differs for each element, the BS employs stopping power lists for a range of energies for each of the 12 elements described in \mathbf{f} . The Bragg-Kleeman rule states that the stopping power of compounds in a mixture is additive [46], so

$$\frac{1}{M} \left(\frac{dE}{dx} \right) = \sum_{i=1}^n \frac{f_i}{M_i} \left(\frac{dE}{dx} \right)_i, \quad (4.1)$$

where M is the molar mass of the mixture and M_i the molar mass of the element corresponding with mass weight fraction f_i . Note that $(f_i M)/M_i$ is the mole fraction of element i in the mixture.

The proton stopping power libraries from the National Institute Standards Technology (NIST) PSTAR [47] and Stopping and Range of Ions in Matter (SRIM) 2013 [48] were considered for implementation. The SRIM database was used because NIST does not cover all 12 of the elements included by Schneider et al. The cross sections are shown in figure 4.1a.

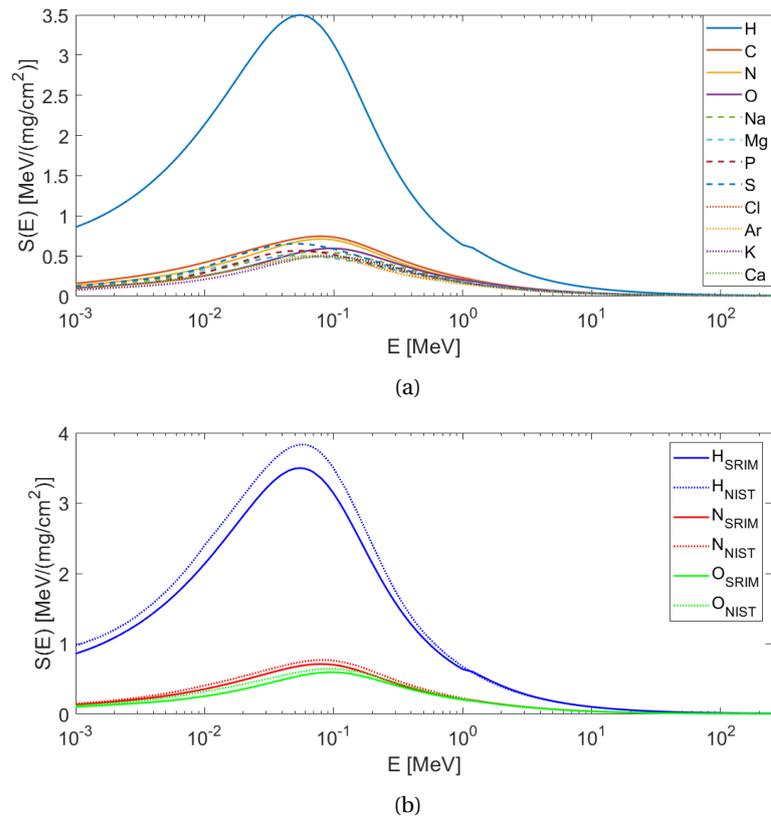


Figure 4.1: **a.** The cross sections for stopping power from the SRIM database for the 12 components required by the stoichiometric conversion as proposed by Schneider et al. [45, 48]. **b.** Comparison between the stopping powers of H, N and O between SRIM (non-gaseous state of matter) and NIST (gaseous state) [47, 48].

Another advantage of the SRIM database is that it distinguishes compounds in gaseous and non-gaseous state. The stopping power [MeV cm²/g] of an element in a gas differs significantly from other states of matter, as shown in figure 4.1b. The NIST PSTAR database includes stopping power for hydrogen, nitrogen and oxygen in gaseous state only. However, these elements occur mostly in non-gaseous state in the body. Therefore, the stopping power tables for the non-gaseous state were included in the BS solver. This means that the stopping power in the air surrounding the body is mishandled. Still, because of the low density of air, the deviations the dose distributions are insignificant.

The stopping power lists are implemented with an energy range of [0.999 keV, 250 MeV]. Energy intervals are proportional to the energy level: from small steps in between low energies ($\Delta E \leq 1$ keV for energies below 18 keV) and larger for high energies (25 MeV steps beyond 200 MeV). Lastly, SRIM provides stopping powers in [MeV/(mg/cm²)]. To apply them in the BS, they are converted into [MeV/cm] by multiplication by the mass density ρ .

Furthermore, the BS contains microscopic cross sections for PG production. The PG library developed by Tolboom was used [49]. These cross sections are based on simulations in Topas MC.

4.1.2. Pencil beams

The pencil beams in the BS are mono-energetic, uni-directional beams of finite width, with a lateral intensity profile described by a 2D Gaussian distribution. Because the Gaussian distribution is non-zero over the entire real domain, a cut-off radius is defined, beyond which the intensity is deemed 0. The consideration of this radius is described in section 4.5.2. The treatment plan characterizes each pencil beam using five features:

- Initial energy E_0
- Initial intensity I_0 , which represents the number of protons in the beam
- Beam width σ , which is the standard deviation of the lateral intensity profile
- Origin \mathbf{r}_0 , which is the gantry location
- Direction $\hat{\Omega}$.

4.2. Interpretation of ODEs

As described in section 3.4, the method of characteristics yields three ODEs that are used to evaluate $\phi_u(\mathbf{r}, E)$ for each pencil beam over its path through the tracing volume. Integrating the first equation, $dx/dq = 1$, yields $x(q) = q + x_0$. This means that the solutions for $\phi_u(\mathbf{r}, E)$ lie on the uni-directional path of the beam, which is a straightforward conclusion. The beams in the BS are forward traced in steps of step size h , so $\mathbf{r}_{n+1} = \mathbf{r}_n + h\hat{\Omega}$

For the second ODE, $dE/dq = -S(x, E)$, the equation for stopping power emerges as equation 4.2. To find E , the equation is numerically integrated along the beam path, so $E_{n+1} = f(E_n)$.

$$\frac{dE}{dq} = \frac{dE}{dx} \frac{dx}{dq} = \frac{dE}{dx} = -S(x, E) \quad (4.2)$$

Lastly, the third ODE describes absorption of protons. Like the second ODE, it can be re-written: $d\phi_u/dx = -\Sigma_t(x, E)\phi_u(x)$. To find ϕ_u , the equation will be numerically integrated: $\phi_{u,n+1} = f(\phi_{u,n})$. The numerical integration methods used to evaluate E and ϕ_u are described in section 4.3.

4.3. Numerical integration methods

4.3.1. Euler explicit method

The Euler explicit method is a first-order routine to numerically approximate ODEs. To estimate an equation given by

$$y'(t) = f(t, y(t)); \quad y(t_0) = y_0 \quad (4.3)$$

the Euler method prescribes, in steps of size h ,

$$y_{n+1} = y_n + hf(t_n, y_n). \quad (4.4)$$

When the Euler method is applied to integrate the proton beam energy, it translates into

$$E_{n+1} = E_n - hS(E_n) \quad (4.5)$$

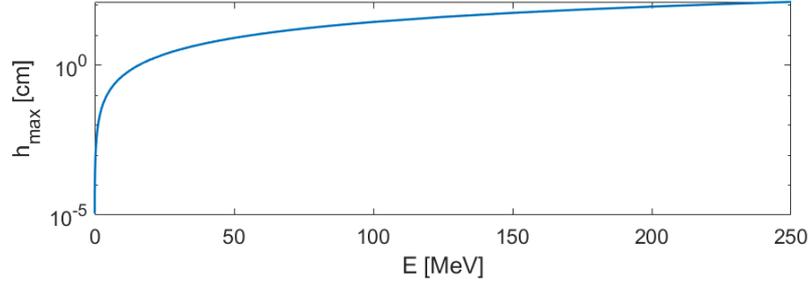


Figure 4.2: Relation between E and the maximum step size h_{max} for the Euler explicit integration method in water.

In order to be numerically stable, the solution should be exponentially decaying, and therefore obey

$$\left| \frac{E_n - hS(E_n)}{E_n} \right| \leq 1. \quad (4.6)$$

This yields a condition for h : $0 \leq hS(E_n)/E_n \leq 2$. Figure 4.2 shows the relation between E and the maximum value of h for water. If h is to be constant over a whole simulation in BS, it should be in the order of 10^{-5} cm. This requires long calculation times.

However, the BS can still produce accurate outcomes if the Euler method is not stable over the whole energy range. If instability occurs only when E is so low that the beam will finish in the current voxel, all remaining energy will be deposited in that voxel and the instability will have no effect on the outcome.

The effect of various h on the integration accuracy was further illustrated by a set of simulations in the BS, summarized in figure 4.3. Here a 100 MeV proton beam was traced through a volume of muscular tissue. The voxels were 0.05 cm in each dimension. The voxels were crossed in one step ('Euler 1'; $h = 0.05$ cm), 10 steps ('Euler 10'; $h = 5 \cdot 10^{-3}$ cm), and so on. To provide a reference data set, the more accurate Heun method, described in section 4.3.2, was run with $h = 5 \cdot 10^{-4}$ cm.

Figure 4.3a shows that for low E the Euler method is accurate, but significant deviations from the reference data occur near the Bragg peak, where E is low. For Euler 1 and Euler 10, the range of the protons increased due to inaccurate integration. For smaller steps, the Euler method calculates the right range, but the height of the Bragg peak still shows deviations: even for $h = 5 \cdot 10^{-4}$ mm, the Euler method shows a 1% deviation.

The deviations from the reference data set are shown in figure 4.3b. This figure too shows that the deviations are largest for high h and low E .

Varying h on the course of the beam, for example based on E upon entering each voxel, could improve the trade-off between computation time and accuracy. However, this method would still require very small h near the range. Therefore, the Heun integration method was applied.

4.3.2. Heun's method

The Heun method is an extension of the Euler explicit method. Being a second order integration procedure, the Heun method has a smaller error. It consists of two steps: first, an estimate \tilde{y}_{n+1} , based on the slope at point y_n , is calculated, as done in equation 4.4. Therefore, the slope at point \tilde{y}_{n+1} is determined: $f(t_n, \tilde{y}_{n+1})$. This second slope is then used to correct the first slope. The solution, as dictated by Heun's method, is then

$$y_{n+1} = y_n + \frac{h}{2} (f(t_n, y_n) + f(t_n, \tilde{y}_{n+1})) \quad (4.7)$$

where

$$\tilde{y}_{n+1} = y_n + hf(t_n, y_n).$$

This, in terms of E and S , is

$$E_{n+1} = E_n + \frac{h}{2} (S(E_n) + S(\tilde{E}_{n+1})). \quad (4.8)$$

Heun's method converges faster than Euler's method. Figure 4.4 shows that the energy distribution produced with Euler's method and $h = 5 \cdot 10^{-4}$ cm still changes, while the results from Heun's method hardly change when h is varied from 0.05 cm to $5 \cdot 10^{-4}$ cm. Heun's method can therefore be applied with relatively large step sizes.

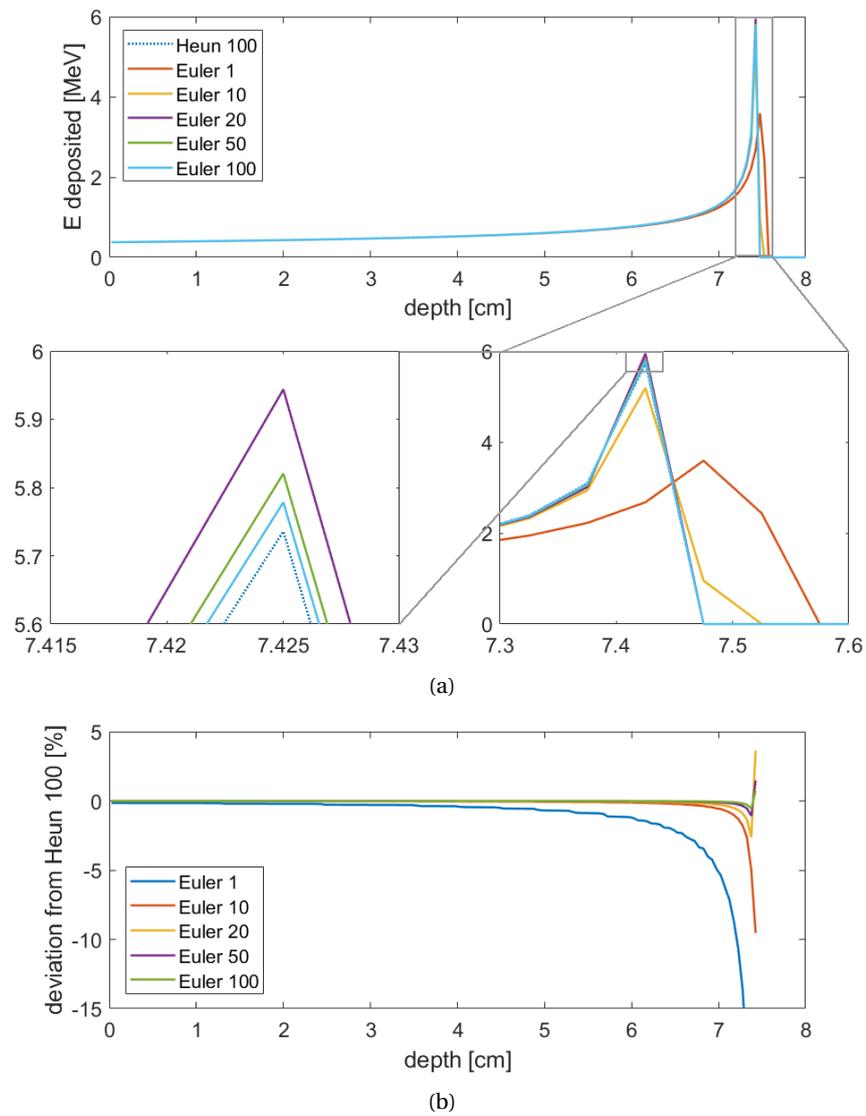


Figure 4.3: **a.** Integrated depth dose for a 100 MeV proton in a homogeneous muscle tissue target, simulated with BS in a grid of voxels of 0.05 cm in each dimension. The lines represent different integration methods. 'Name X' in the legend means the integration method of Name was used, with step size $h = 0.05/X$ cm. **b.** Deviations in the integrated depth dose from figure a. The 'Euler 100' data set was used as the reference data set.

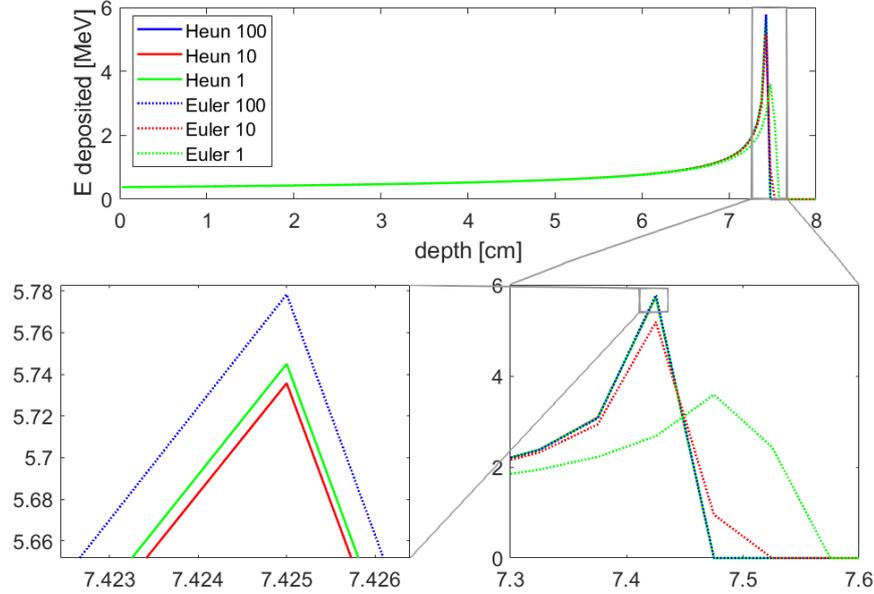


Figure 4.4: Integrated depth dose for a 100 MeV proton in a homogeneous muscle tissue target, simulated with BS in a grid of voxels of 0.05 cm in each dimension. The lines represent different integration methods. 'Name X' in the legend means the integration method of Name was used, with step size $h = 0.05/X$ cm. The 'Heun 100' line is not visible because it is overlapped by 'Heun 10'.

4.4. Tracing algorithms

The Boltzmann solver produces 3D distributions for dose deposition and PG emission. Its general procedure is to trace individual mono-energetic pencil beams of finite width through the treatment volume, evaluating the fluence and beam energy at each step of the way. The tracing continues either until the beam energy reaches below the cut-off energy of 1 keV or until it exits the volume, depending on which event occurs first.

The BS was initially built using a Gaussian quadrature method and determine dose deposition or PG emission. This method is described in section 4.4.1. However, long calculation time and a lack of energy conservation impaired this approach. Therefore, an alternative method based on mathematical sub-beams was implemented, which is elaborated on in section 4.4.2.

In these sections, the tracing methods to calculate dose distributions $D(\mathbf{r})$ are described. The same methods can be employed for PG activity distribution (PG_i for individual gamma lines i) or any other cross section. Lastly, because the cross sections for absorption are currently excluded, ϕ_u remains constant over the path until E reaches a lower limit; then $\phi_u = 0$.

4.4.1. Gaussian quadrature

The original tracing method is referred to as the (Gaussian) quadrature method because the BS scores dose only on a set of points that are described by the Gaussian quadrature rules. Using these points, the dose distribution can be determined for the whole volume.

Following the discontinuous Galerkin method, the BS determines a polynomial that describes the dose distribution $D(\mathbf{r})$ within an integration domain [50]. In the BS these integration domains are voxels on the same grid as the planning CT. It was assumed that the behaviour of the dose distribution is approximately linear on voxel-scale, so a 1st order polynomial was implemented:

$$D(\mathbf{r}) = \sum_{i=1}^N D_i h_i(\mathbf{r}) \quad (4.9)$$

where D_i are unknown coefficients for the basis functions h_i . For a 1st order polynomial, $N = 4$ and $h_i = [1; x; y; z]$.

To solve D_i , the value of $D(\mathbf{r})$ at a set of P specifically located integration points in each integration domain, must be known. The Gaussian quadrature rules dictate the locations of a set of quadrature points (QP), which lie at optimal positions to approximate definite integrals (figure 4.5). To evaluate a polynomial of order

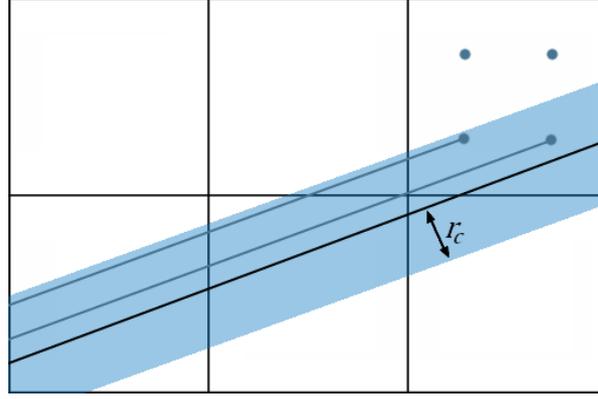


Figure 4.5: Proton beam (in blue) on a 2D grid. To determine a 1^{st} order polynomial for $D(\mathbf{r})$ in the upper right pixel, the dose in the four QP is scored. The upper two QP lie beyond cut-off radius r_c , so the dose is considered 0. The lower two are traced towards over the paths indicated.

1, 2 points in each dimension are necessary, thus requiring $P = 2^3$ points per voxel. The BS calculates the dose at these (QP) by tracing towards each point individually. Hereto the following procedure is used:

1. Using a pencil beam's origin \mathbf{r}_0 and direction Ω , the distance r between the beam axis and each QP, measured perpendicularly to the axis, is determined. The algorithm assumes an infinite range of the beam, because the range is unknown at this point. A cut-off radius r_c is used, after which the dose intensity is considered 0. Only nodes with $r \leq r_c$ are included in step 2.
2. The selected QP are traced towards individually. Each QP is weighted according to the 2D normal distribution which defines the lateral beam intensity profile:

$$W_{QP} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right). \quad (4.10)$$

3. The location at which the path from the source towards the QP enters the planning CT is determined. The voxel adjacent to the entry point, is the first voxel that will be traced through.
 - (a) The path through the voxel is determined. For most voxels, the path reaches from the entry point of the current voxel to the exit point. However, if the voxel contains the QP that is traced towards, the path goes from the entry point to the QP. The path length is l .
 - (b) The path is divided into m intervals, with step size $h = l/m$.
 - (c) For each interval, $S(\mathbf{r}, E)$ is determined based on E , the voxel's ρ and \mathbf{f} , and the cross section library for stopping power, using equation 4.1.
 - (d) The beam propagates $h\hat{\Omega}$. E is re-evaluated using $S(\mathbf{r}, E)$, based on an integration method from section 4.3. When absorption cross sections are included, ϕ_u is updated too.
 - (e) Steps c-d are repeated until the path is crossed, or until $E < 1$ keV, in which case the tracing for this QP terminates.
 - (f) When a QP is reached, the dose is scored:

$$D_{QP} = W_{QP} \frac{(\phi_u(\mathbf{r}_{QP} - h\hat{\Omega})E(\mathbf{r}_{QP} - h\hat{\Omega}) - \phi_u(\mathbf{r}_{QP})E(\mathbf{r}_{QP}))}{h} \quad (4.11)$$

4. Steps 1-3 are repeated for each pencil beam. The D are accumulated per voxel.
5. Using the discontinuous Galerkin method, the dose to each voxel is calculated.

To solve the coefficients D_i from D_{QP} , the discontinuous Galerkin method is employed. Hereto both sides of equation 4.9 are multiplied by a set of indices, h_j , and then integrated over the voxel volume V_e :

$$\int_{V_e} D(\mathbf{r}) h_j(\mathbf{r}) d\mathbf{r} = \sum_{i=1}^N D_i \int_{V_e} h_i(\mathbf{r}) h_j(\mathbf{r}) d\mathbf{r}. \quad (4.12)$$

The left-hand side of equation 4.12 is approximated by the QP:

$$\int_{V_e} D(\mathbf{r}) h_j(\mathbf{r}) d\mathbf{r} \approx \sum_{QP} D_{QP} h_j(\mathbf{r}_{QP}) w_{QP} |J|, \quad (4.13)$$

where w_{QP} is the quadrature weight, prescribed by the quadrature rules, and $|J|$ is the quadrature determinant. On the right-hand side, $\int_{V_e} h_i(\mathbf{r}) h_j(\mathbf{r}) d\mathbf{r}$ is the so-called mass matrix $M_{i,j}$. This yields a linear system:

$$M_{i,j} D_i = LHS, \quad (4.14)$$

where LHS represents the approximation for the left-hand side from equation 4.13.

This tracing method entails limitations. First of all, for the dose deposition, the dose gradients of the Bragg peak are so steep that severe discontinuities and negative values are likely to occur in the final voxel. To handle this, a gradient limiter is applied. This procedure overwrites the coefficients for basis functions $h_2 - h_4$ with 0 when $D(\mathbf{r}) < 0$ in any vertex, thus yielding $D(\mathbf{r}) = D_1 h_1$.

Another method to handle non-linear behaviour and steep gradients would be to use more nodes per voxel, so more complex spatial dependencies can be described within the voxels. Increasing the number of nodes from 2 to 3 per dimension would however increase the computation time by a factor ($3^3/2^3$). Using the described algorithm with 8 nodes and 20 intervals per voxel required multiple hours of computation per pencil beam. Such calculation times do not produce a clinically viable tool, so increasing it further by adding nodes is not an option. It must be noted here that at that time, the code had other sub-optimal features at that further slowed it down. However, section 4.4.3 points out that the quadrature algorithm is slow relative to its alternative described in section 4.4.2.

The steep gradients introduced a second problem: energy conservation. The 8 quadrature points did not capture the behaviour of energy deposition within the used voxel sizes, as it is not nearly linear. Laterally, the energy deposition can be described as a Gaussian distribution, for which a linear approximation causes relatively minor deviations in the integral dose. However, distally, the energy deposition near the Bragg peak evolves in an extreme manner, causing $D(\mathbf{r})$ to become very sensitive to the location of the QP relative to the maximum of the Bragg peak. This is illustrated in figure 4.6.

Due to the high dose gradients, these final voxels often are subject to the gradient limiter, meaning that the whole voxel is assigned with a constant value of the high dose measured in the Bragg peak (figure 4.6c). This severely overestimates the dose to the voxel, and consequently a significant difference between the total energy of the beams and the integrated dose deposition was observed.

4.4.2. Needle beam method

The procedure that was implemented secondly, divides each pencil beam into mathematical pencil beams: needle beams. Each needle beam is assigned the weight of part of the lateral beam intensity distribution:

$$W_{nb} = \int_{\theta}^{\theta+\Delta\theta} \int_r^{r+\Delta r} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{y^2+z^2}{2\sigma^2}\right) \quad (4.15)$$

for a pencil beam in the x-direction.

The values Δr and $\Delta\theta$ are determined based on the requirement that each voxel crossed by the pencil beam should be crossed by at least 10 needle beams, except for the voxels on the edges of the pencil beam. This condition is satisfied when $10\Delta r\Delta\theta \leq A_v$, where A_v is the area of the face of the voxel crossed by the beam. The number of needle beams is a trade-off between calculation time and lateral deviations, which are shown in figure 4.7.

The needle beam method consists of the following steps:

1. A pencil beam is divided into needle beams.
2. Using the pencil beam direction's and the needle beam's translation relative to the pencil beam axis, the entry point of the needle beam in the planning CT is determined. The voxel adjacent to the entry point is the first-evaluated voxel.
 - (a) The path through the voxel is determined. The path reaches from the entry point of the current voxel, \mathbf{r}_{in} to the exit point \mathbf{r}_{out} . The path length is l .

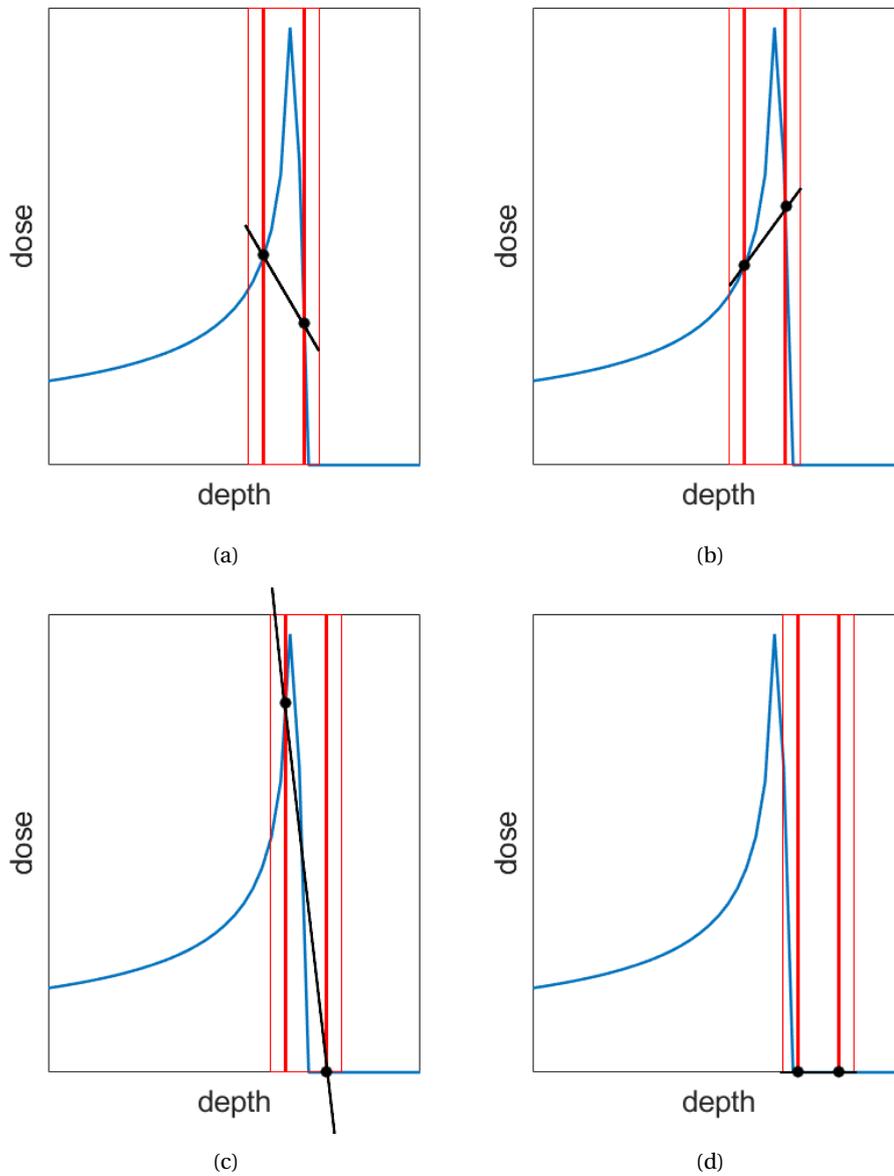


Figure 4.6: The calculated dose distribution is sensitive for the location of the QP relative to the Bragg peak: an 1D representation. In blue the traced dose distributions; in red the borders of the final integration domain, with in thick red the locations of the QP. The black dots represent the dose scored at the QP, and the black line the dose distribution produced by the Gaussian quadrature method. **a.** and **b.** yield acceptable, but very different $D(\mathbf{r})$. In **c.** must be used because $D(\mathbf{r}) < 0$ in the distal edge. In **d.** the quadrature method yields $D(\mathbf{r}) = 0$, while energy was deposited in the domain.

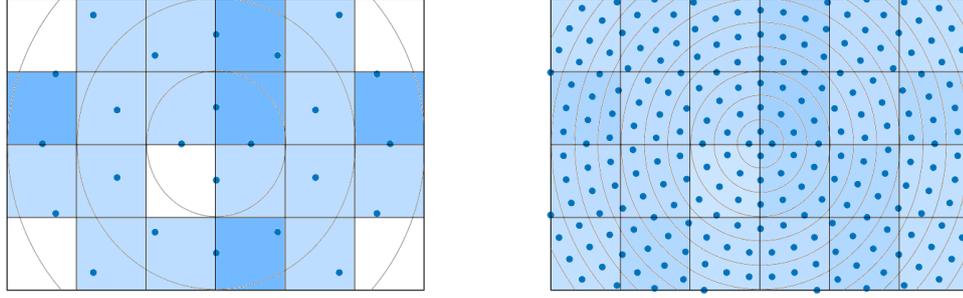


Figure 4.7: 2D representation of needle beams crossing voxels, perpendicular to the beam direction. The intensity of the color markings represents the number of beams per voxel, not taking the needle beam weight into account. **a.** ~ 1 needle beam per voxel yields severe deviations in the lateral dose distribution. **b.** ~ 9 needle beams per voxel yield a more homogeneous lateral distribution of dose.

- (b) The path is divided into m intervals, with step size $h = l/m$.
- (c) For each interval, $S(\mathbf{r}, E)$ is determined based on E , the voxel's ρ and \mathbf{f} , and the cross section library for stopping power, using equation 4.1.
- (d) The beam propagates $h\hat{\Omega}$. E is re-evaluated using $S(\mathbf{r}, E)$, based on an integration method from section 4.3. When absorption cross sections are included, ϕ_u is updated too.
- (e) Steps c-d are repeated until the path is crossed, or until $E < 1$ keV, in which case the remaining energy is immediately added to the current voxel.
- (f) The dose D is scored in the voxel:

$$D = W_{nb} (\phi_u(\mathbf{r}_{in})E(\mathbf{r}_{in}) - \phi_u(\mathbf{r}_{out})E(\mathbf{r}_{out})) \quad (4.16)$$

- (g) Steps a-f are repeated for each needle beam in the pencil beam. The energy depositions are accumulated.

3. Steps 1-3 are repeated for each pencil beam. The energy depositions are accumulated.

4.4.3. Comparison between quadrature and needle beam methods

The advantage of the needle beam method as opposed by the Gaussian quadrature method, is that the needle beams are traced in a forward manner. The quadrature method on the other hand, requires each node to be traced to from the beam entry, even though the same path is crossed multiple times.

To illustrate the effect on calculation time, the number of voxels crossed for both methods are determined for one pencil beam in a test grid. The grid has voxel size $0.1 \times 0.1 \times 0.3$ cm³, where the 0.1×0.3 cm² face is oriented perpendicularly to the beam direction. The pencil beam has a width of $\sigma = 0.5$ cm, and the cut-off radius is $r_c = 3.8 \sigma$. This means that the beam area perpendicular to the beam direction covers $3.8 \cdot 10^2$ voxels. Lastly, a proton range of 20 cm, or 200 voxels, is assumed.

Needle beam method: For 10 needle beams per perpendicular voxel area, the needle beam method requires $3.8 \cdot 10^3$ beams to describe this pencil beam. For the proton range of 20 cm, this requires tracing through $7.6 \cdot 10^5$ voxels.

Gaussian quadrature method: The beam covers $3.8 \cdot 10^2 \cdot 200 = 7.6 \cdot 10^4$ voxels. At 8 QP per voxel, this means that $6.0 \cdot 10^5$ QP must be traced towards. The QP most proximal to the beam source only require a single voxel crossing, but the most distal QP require 200 crossings. The average QP requires 100.5 voxels traced through before it is evaluated. Therefore, this method requires $6.1 \cdot 10^7$ voxels crossed. For this test case, this method is over 80 times more expensive than the needle beam method.

4.5. Reduction of computation time

The computation time of the BS is an important factor for its clinical viability. Therefore three methods to further reduce the number of computational steps are discussed.

4.5.1. Euler versus Heun

As the tracing algorithm is the most frequently called subroutine in the code, it is the most dominant factor in computation time. Because the Heun integration method is a two-step procedure, an iteration takes roughly twice the calculation time of an Euler iteration. Therefore it might be feasible to apply the Euler explicit integration method with one interval per crossed voxel for high beam energies. The beam is then traced until it reaches a certain lower energy limit, after which the Euler method starts showing significant deviations and the slower Heun method, also with one interval per voxel, will be applied instead.

The value of the lower energy limit depends on the deviation in dE/dx allowed. The 'Euler 1' data of figure 4.3b show a 0.15% deviation at $E = 100$ MeV. Because this might yield significant deviations already, the idea of using the single Euler method was abandoned.

4.5.2. Cut-off radius

Another way to reduce calculation time is to reduce the number of voxels handled. As the lateral beam profile in the BS is a 2D Gaussian distribution, the code includes a cut-off distance r_c perpendicularly from the beam axis. In polar coordinates, when $\sigma_x = \sigma_y = \sigma$ and $\mu_x = \mu_y = 0$, the beam intensity at distance r is described as

$$\frac{1}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}} \quad (4.17)$$

and the fraction covered in case of cut-off radius r_c

$$\int_0^{r_c} \int_0^{2\pi} \frac{1}{2\pi\sigma^2} e^{-\frac{r^2}{2\sigma^2}} r d\theta dr = \left[-e^{-\frac{r^2}{2\sigma^2}} \right]_0^{r_c} = 1 - e^{-\frac{r_c^2}{2\sigma^2}}. \quad (4.18)$$

To cover 99.9% of the beam's effect, a cut-off distance of $r_c = 3.8 \sigma$ suffices. For energy conservation purposes the last 0.1% is included through a normalization factor is applied to the whole beam.

The distance $r_c = 3.8 \sigma$ is somewhat arbitrarily chosen. As long as the beam intensity is normalized, smaller r_c may be used, e.g. 3.2σ , which covers 99.4% of the beam intensity, or 2.6σ , which covers 96.6%. In these case, the calculation time would be reduced because fewer needle are used. For $r_c = 3.2 \sigma$ the computation time would be reduced by a factor $(3.2^2/3.8^2) = 0.7$ and for $r_c = 2.6 \sigma$, by a factor 0.47.

For single pencil beams, the dose distribution produced with $r_c = 2.6 \sigma$ may differ significantly from the case with $r_c = 3.8 \sigma$. However, when many pencil beams are packed close together, as occurs during proton therapy, the penumbra of the beams overlap, so the effect of lower r_c becomes insignificant.

4.5.3. Binary search

Lastly, the looking-up of cross sections requires a significant number of computation time. Because the beam energy is continuous, the cross sections are linearly interpolated between the predecessor (next smallest) and successor (next highest) energy levels. A top-down row search algorithm identifies the corresponding row numbers.

The tabulated cross sections are sorted from low to high energy. The stopping power tables are comprised of 157 energy levels each, of which the 100 MeV level is the 130th. As an alternative to the time-intensive top-down search method, a binary search algorithm was implemented. For n rows, this method requires $O(\log(n))$ steps, which is significantly faster than the top-down approach, which requires $O(n)$ steps.

The algorithm uses support variables L , H and M to represent the lower, higher and middle row numbers of a table with n rows. To find the predecessor, the algorithm is described by the following pseudocode:

```

L = 0;
H = n;
while L < H do
  M = floor((L + H)/2);
  if  $E_M < E$  then
    | L = M + 1
  else
    | H = M
  end
end
return L

```


5

Simulation set-up

To test the reliability of the output of the Boltzmann solver, its dose distributions were compared to the simulations produced by other toolboxes. The results of three set-ups were analyzed: (1) a single pencil beam in a homogeneous muscle tissue phantom, (2) a single pencil beam in a muscle tissue phantom with an insert of bone and (3) a full treatment plan in a volume based on the planning CT. The former two were compared to simulations produced by Topas MC [23] and the latter to the planned dose, calculated by the treatment planning system.

5.1. Single beam on homogeneous phantom

The homogeneous target was a $6 \times 6 \times 10 \text{ cm}^3$ volume comprised of $120 \times 120 \times 200$ voxels of 0.5 mm in all dimensions. The characteristics of the tissue were based on the Material Composition Data from NIST PSTAR [47]. A mono-energetic, uni-directional pencil beam with initial energy $E_0 = 100 \text{ MeV}$, intensity $I_0 = 10^7$ protons, width $\sigma = 0.3 \text{ cm}$ and cut-off radius $r_c = 3.8 \sigma$ was traced through the volume.

For the simulation in the Boltzmann solver a 'CT scan' with $120 \times 120 \times 200$ voxels, each assigned with a mock CT value of 0, was constructed. By default, the BS converts HU into tissue characteristics using Schneider's stoichiometric calibration [45]. However, to properly compare the BS and Topas MC, the BS code was manipulated so ρ and \mathbf{f} for CT value 0 were overwritten with the material characteristics for skeletal muscle tissue from NIST, with $\rho = 1.05 \text{ g/cm}^3$ and the weight vector as described in table 5.1.

The pencil beam was divided into needle beams with a resolution of 8 per dimension, yielding ~64 needle beams per voxel and $1.0 \cdot 10^5$ needle beams in total.

	G4_MUSCLE_SKELETAL_ICRP [% <i>mass</i>]	G4_BONE_COMPACT_ICRU [% <i>mass</i>]
f_H	10.2	6.4
f_C	14.3	27.8
f_N	3.4	2.7
f_O	71.0	41
f_{Na}	0.1	-
f_{Mg}	-	0.2
f_P	0.2	7.0
f_S	0.3	0.2
f_{Cl}	0.1	-
f_K	0.4	-
f_{Ca}	-	14.7

Table 5.1: Mass fractions of the components of skeletal muscle and compact bone from NIST PSTAR [47]. These values are used to overwrite the algorithm by Schneider et al. for sections 5.1 and 5.2.

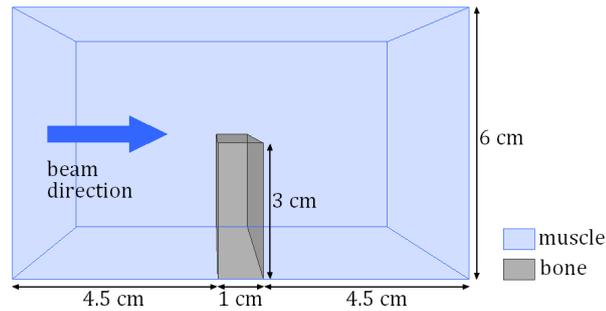


Figure 5.1: The design of the muscle tissue phantom with bone insert.

Topas MC version 3.1.p2 was applied with default settings. The target was a $6 \times 6 \times 10 \text{ cm}^3$ box volume. Topas employs the Geant4 Material Database, which includes NIST compounds, so the proper material characteristics were called by assigning `G4_MUSCLE_SKELETAL_ICRP` to the volume. The dose was scored on a $120 \times 120 \times 200$ grid, which exactly overlapped the phantom.

The initial proton energy was set to 100 MeV, with an energy spread of 0.0 MeV. The beam `BeamPositionDistribution` was set "Gaussian", with a 0.3 cm standard deviation. The `BeamPositionCutoffShape` was an "Ellipse", with cut-off radius $3.8 \sigma = 1.14 \text{ cm}$.

Unlike the BS, Topas MC does not allow strictly uni-directional beams. Therefore, to approach the singular beam direction of the Boltzmann solver, the `BeamAngularSpread` was set at $1^\circ \cdot 10^{-14}$, with a `BeamAngularCutoff` of $1^\circ \cdot 10^{-4}$. The simulation consisted of 10^7 protons.

5.2. Single beam on inhomogeneous phantom

The second geometry is included to illustrate the Boltzmann solver's handling of inhomogeneities. The phantom was comprised of `G4_MUSCLE_SKELETAL_ICRP` with an insert made of compact bone (`G4_BONE_COMPACT_ICRU`) with $\rho = 1.85 \text{ g/cm}^3$. The bone slab had a volume of $3 \times 6 \times 1 \text{ cm}^3$ and was located in the phantom as shown in figure 5.1.

In the 'CT' for the BS, the insert was given another mock CT value. The code was further manipulated to assign compact bone material features, according to the NIST database, to voxels with this CT value. Furthermore, the beam set-up was exactly the same as in the homogeneous run.

In Topas MC a box of the `G4_BONE_COMPACT_ICRU` was placed in the muscle tissue volume. Apart from that, the set-up was equal to the previous simulation in Topas MC.

5.3. Full treatment plan on a planning CT

To analyze the performance of the BS in a more realistic set-up, the planning CT of a prostate cancer patient was used as a target. The used CT consists of 108 slices of 0.3 cm thickness, covering $40.95 \times 26.25 \text{ cm}^2$ in 468×323 voxels of $0.875 \times 0.875 \text{ mm}^2$ each.

The CT was placed in a cartesian coordinate system so that the volume had 108 voxels in the x-direction, 323 in the y-direction and 468 in the z-direction. Furthermore, the plan defines an isocenter of the CT at the center of mass of the prostate, which was a few cm from the geometric CT center. The origin of the coordinate system was placed at the isocenter.

The corresponding treatment plan consists of 1521 pencil beams distributed over two lateral source (or gantry) locations. The gantry locations both lie on the z-axis. The source-to-axis distance (SAD) of the gantry was set at 195 cm.

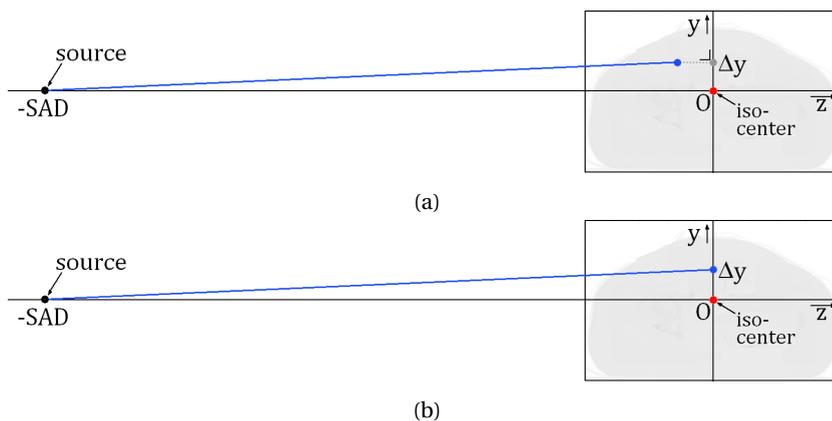


Figure 5.2: **a.** In the treatment plan, the beam direction is defined using an x- and y-coordinate, here illustrated in 2D with y-coordinate Δy only. The beam (in blue) reaches Δy at a certain depth z . However, the interpretation of z from the treatment plan is unclear. **b.** In the Boltzmann solver Δy is projected on the iso-center instead of depth z . Because $SAD \gg \Delta y$, this results in a very small deviation in the beam angle relative to image a.

A treatment plan assigns each beam with an initial energy, direction, intensity and width. The energy and width can be applied directly in the BS. However, the direction and beam intensity required conversion before application.

In the treatment plan, the beam direction is described by x- and y-coordinates, which the beam reaches after it has crossed its water-equivalent path length (WEPL) in the CT volume (figure 5.2). Due to limited communication with the party that provided the treatment plan, some uncertainty about the exact interpretation of these directions exists. However, because the $SAD \gg \Delta x, \Delta y$, the deviations due to this uncertainty are very small. Therefore the x- and y-coordinates were projected on the isocenter of the CT.

The treatment plan describes the intensity in monitor units (MU). The number of protons per MU is energy dependent. A separate table of proton energies and the corresponding number is employed for conversion of MU to number of protons per beam. Linear interpolation was used for energy levels in between tabulated values.

The Boltzmann solver was assigned a needle beam resolution of 2, which yields $(0.0875 \cdot 0.3) / (0.0875/2)^2 \approx 14$ beams per perpendicular voxel area. In order to study the effect of the cut-off radius r_c , three runs with $r_c = 3.8 \sigma; 3.2 \sigma; 2.6 \sigma$ were simulated, yielding $6.9 \cdot 10^6, 4.9 \cdot 10^6$ and $3.3 \cdot 10^6$ needle beams respectively.

The simulation of the Boltzmann solver was compared to the planned dose which is enclosed in the treatment plan. The treatment plan was designed by the Erasmus-iCycle treatment planning system [51].

6

Results

6.1. Single beam on homogeneous phantom

The set-up described in section 5.1 in the BS is was run in 16.0 s. The total energy deposited was exactly 10^9 MeV, which corresponds to 10^7 protons of 100 MeV. This proves that the BS is a conservative tool. Figure 6.1a depicts the dose distribution of the central slice of the volume.

The simulation in Topas MC following section 5.1 took 4.43 h to run, using 4 threads. The total energy deposited was $9.78 \cdot 10^8$ MeV, which indicates that secondary particles have exited the scoring volume. The dose distribution in the central slice is shown in figure 6.1b.

From the figures it stands out that the Bragg peak in the MC simulation is distally more spread out than the result from the BS: in the BS the energy is largely delivered in the final voxels. In the BS each needle beam has the exact same path length, and because the beams are uni-directional, they all reach the same depth. In Topas MC range straggling is included: each proton is individually traced and undergoes different interactions, causing the path lengths of protons to differ. Furthermore, because lateral scatter is allowed, the path length and depth of the proton differ. This phenomenon increases to the broadening of the Bragg peak. The difference in Bragg peak handling was expected, because the BS uses the continuous slowing down approximation and excludes scatter interaction.

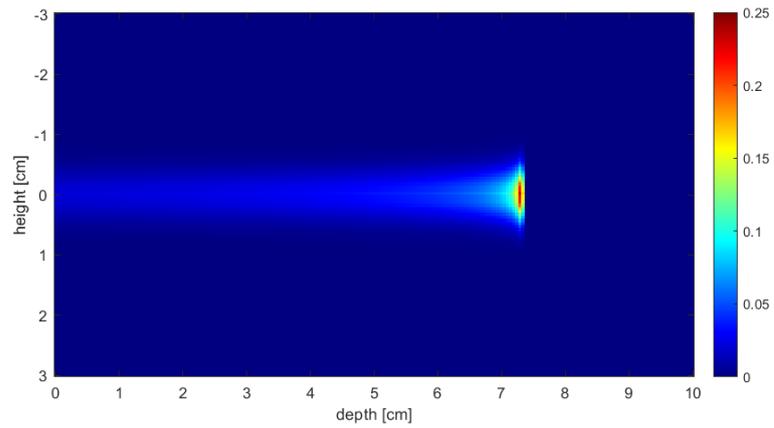
The absolute difference between the central slices is depicted in figure 6.1c. This image shows that the Bragg peak from the BS has a steeper rise than the MC, and shows a very abrupt transition between overestimation and underestimation when the fluence in the BS stops. Furthermore, the BS is expected to underestimate lateral dose because of the lack of scatter in the model. However, the lateral deviations are small relative to the distal deviations. In figure 6.2 the underestimation relative to the maximum value of the MC code is indicated, on logarithmic scale. Here the lateral deviations that are of the order 0.01-0.1%, become visible.

Figure 6.3 shows the integral depth dose (IDD) in [J] for Topas MC and the BS. The IDD from the BS shows a higher Bragg peak and a steeper rise and fall-off, which are all results from the lack of scatter interactions.

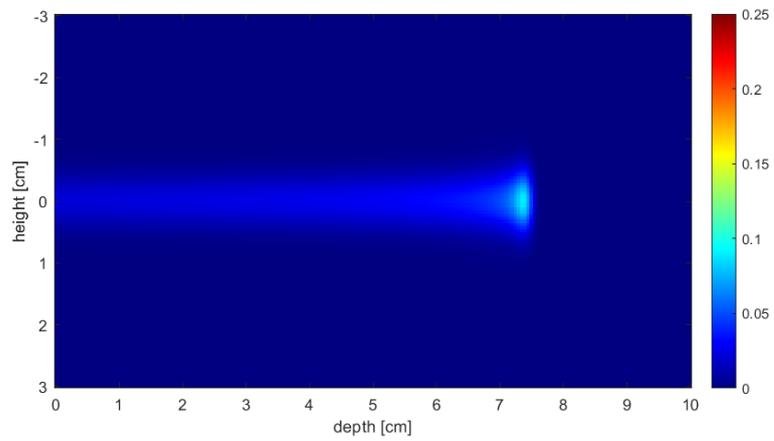
The range simulated by the BS is 7.3 cm, which is shorter than the range from Topas MC: 7.5 cm. This seems counter-intuitive, because protons in the BS do not deviate from their initial direction and they are thus expected to travel deeper than protons simulated in MC at the same path length. The phenomenon of range decrease due to lateral scatter is described by the detour factor F . If the range including angular deviations, is R_i and the range excluding scatter R_e , $F = R_i/R_e < 1$. NIST PSTAR states $F = 0.9987$ for a 100MeV proton beam in skeletal muscle, indicating that the range from the BS is expected to be larger than that of Topas MC [47].

The underlying factor of range underestimation is the overestimation of stopping power. This can have multiple reasons in the BS. First of all, the linear interpolation to determine $S(E)$ based on tabulated cross section overestimates $S(E)$, because the $(E, S(E))$ graph is largely concave (figure 6.4). The magnitude of this error was estimated by comparison with 3rd order polynomial interpolation. For energy levels halfway in-between tabulated values, the $S(E)$ determined by linear interpolation was $\sim 0.2\%$ higher than the data interpolated with the polynomial method. The effect hereof is little: upon re-running the simulation with the polynomial interpolation, the protons reached the exact same depth.

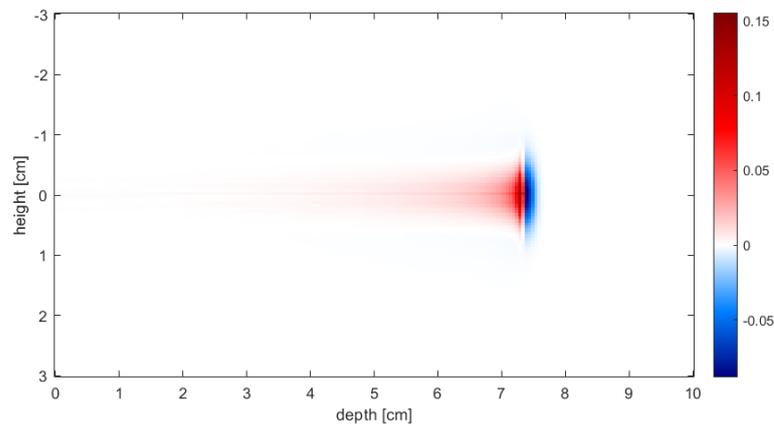
Secondly, the energy integration method used to re-evaluate E on each tracing step may cause deviations. However, the integration methods have been thoroughly reviewed in section 4.3.2. The range calculated with



(a)



(b)



(c)

Figure 6.1: Dose [Gy] deposited in the central slice of the muscle phantom. A 100 MeV pencil beam of 10^7 protons, $\sigma = 0.3$ cm was traced through the volume. **a.** Simulated by the Boltzmann solver (BS). The energy deposited in this slice is $7.11 \cdot 10^7$ MeV. **b.** Simulated by Topas MC. The energy deposited in this slice is $5.94 \cdot 10^7$ MeV. **c.** Deviation [Gy] of the BS simulation relative to the Topas MC simulation (a-b). The red indicates that the BS overestimated the dose relative to Topas MC; blue indicates underestimation.

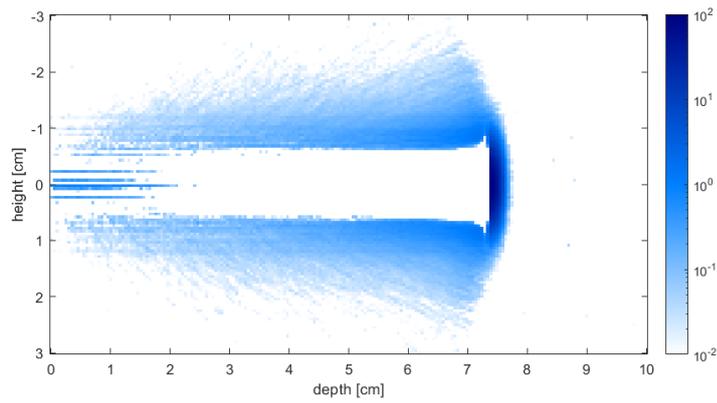


Figure 6.2: The underestimation of dose in figure 6.1a as a percentage of the maximum dose from figure 6.1b. Overestimations are not included in this figure.

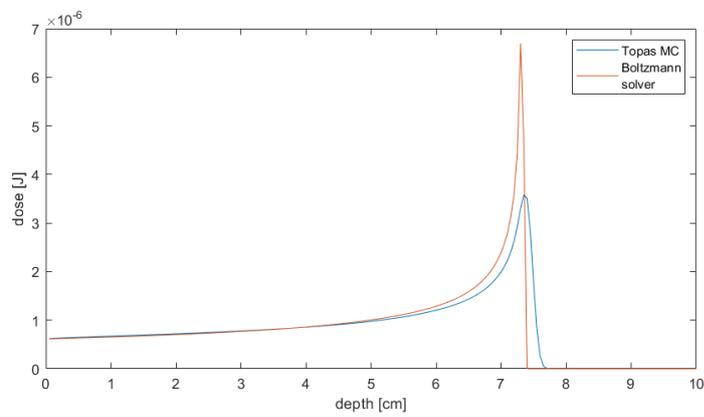


Figure 6.3: Integrated depth dose [J] of a 100MeV pencil beam in a homogeneous skeletal muscle phantom, simulated by Topas MC and the Boltzmann solver.

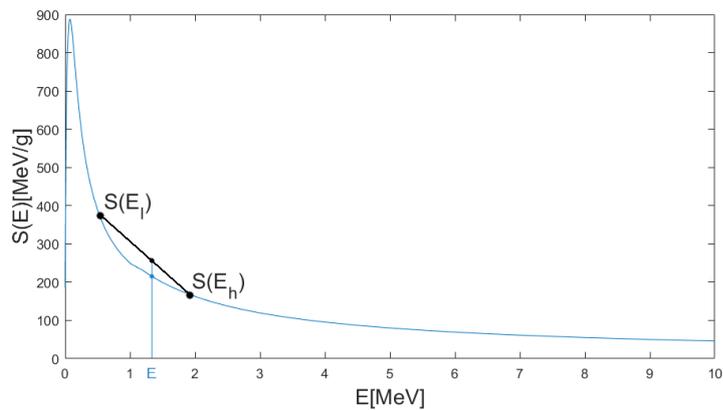


Figure 6.4: Relation between E and $S(E)$ for skeletal muscle [47]. Using linear interpolation between E_{low} and E_{high} to find the stopping power for the E in blue causes an overestimation in $S(E)$.

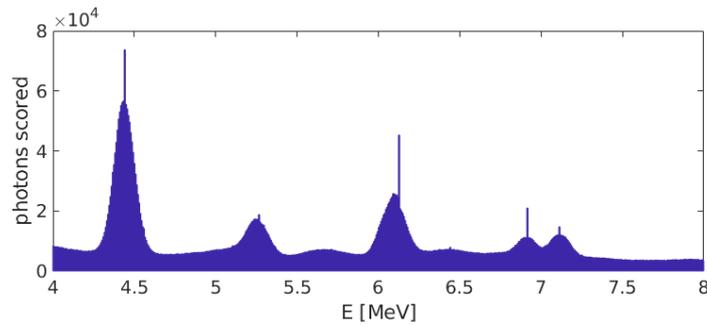


Figure 6.5: The energy spectrum of photons scored by Topas MC. A 100 MeV proton beam was traced through a homogeneous muscle tissue phantom.

single-interval Heun method did not show significant deviation simulations with smaller intervals.

Lastly, the range difference between Topas MC and the BS can be caused by different cross sections. Figure 4.1b shows that stopping power tables from various sources can differ significantly. The physical models employed by Topas MC may therefore produce stopping powers that substantially deviate from the SRIM tables included in the BS.

6.1.1. Prompt gamma emission

Apart from the dose distribution, the prompt gamma emission of the 6.13 MeV gamma line, produced by O-16, was simulated. In the BS this data set was computed simultaneously with the dose distribution.

A separate Topas MC simulation was run to produce reference data, because higher statistics were required than for the dose deposition. $2 \cdot 10^8$ protons were simulated and scored on spherical surface that lied around the dose scoring volume. The scored data included the origin of the PG, from which a PG source distribution was produced. The computation time was 13.0 h. The energy spectrum of the scored photons is shown in figure 6.5. All photons with $6.1 < E < 6.2$ MeV were considered PG.

The PG productions in the central slice for the BS and Topas MC are shown in figure 6.6 and the integral depth production in figure 6.7. Apart from the deviations caused by the lack of scatter in the BS, these images show that there is a factor ~ 2 in the emission yields. This is an unexpected result, because the cross sections in the BS are based on simulations in Topas MC.

The difference can be explained by the criteria based on which photons are counted as PG. Tolboom does not elaborate on the method to select PG from the photon energy spectrum. Therefore, further work is necessary to determine whether the BS contains an error in the PG cross section handling, or if the methods of categorization of photons differ.

6.2. Single beam on inhomogeneous phantom

The set-up for the single pencil beam in the inhomogeneous phantom, described in section 5.2, took 15.6 s to run in the BS. The difference in the calculation time relative to the homogeneous phantom is a random deviation in computation time.

In Topas MC the computation time was 4.53 h. The 0.10 h additional run time relative to the homogeneous slab can be due to statistical deviations in computation speed, but can also be caused by the different interactions because of the different f in the bone slab. The energy scored was $9.88 \cdot 10^8$ MeV. This indicates that less secondary particles exited the scoring volume than in the case of the homogeneous phantom. This is expected, because the total mass of the inhomogeneous phantom is higher than that of the muscle tissue phantom.

Figures 6.1a, 6.8b and 6.8c show the dose profile in the central slice for the BS, Topas MC and the difference, respectively. In figure 6.1a the dose becomes a step function after the the beam has entered the bone slab. Because there are no lateral scatter effects, the insert only causes a shorter range for the portion of the beam that crosses it.

Furthermore, the results largely indicate the same phenomena as figure 6.1. The BS again shows sharper peaks due to the lack of scatter, no range straggling occurs and the range is underestimated, as shown in figure 6.9.

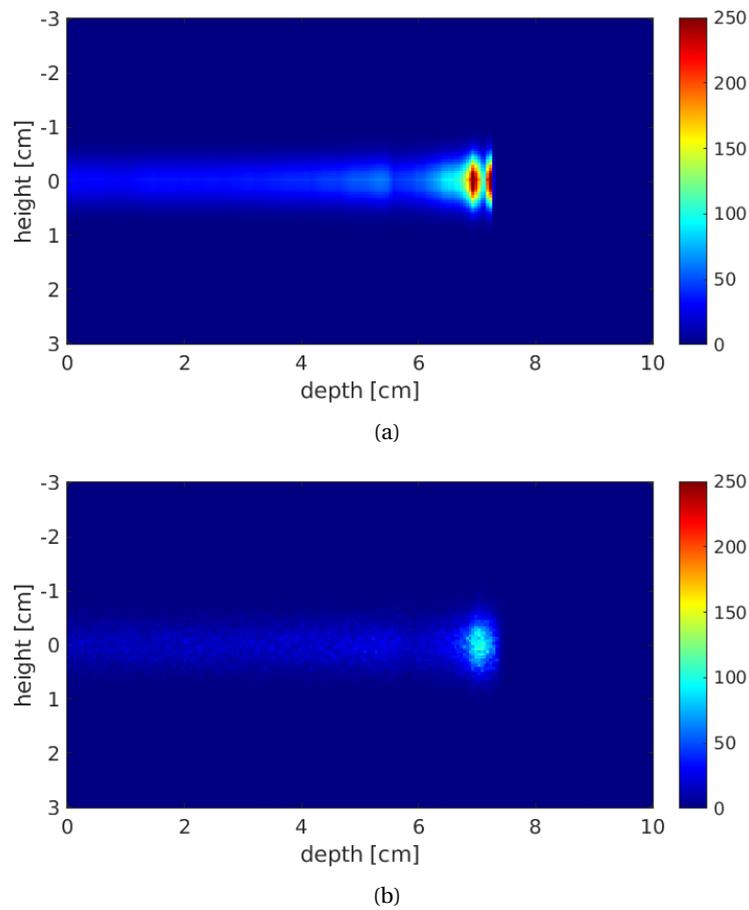


Figure 6.6: The prompt gamma activity of the 6.13 MeV gamma line. A beam of $2 \cdot 10^8$ 100 MeV protons was traced through a homogeneous muscle tissue phantom. **a.** Simulation by the Boltzmann solver. **b.** Simulation by Topas MC.

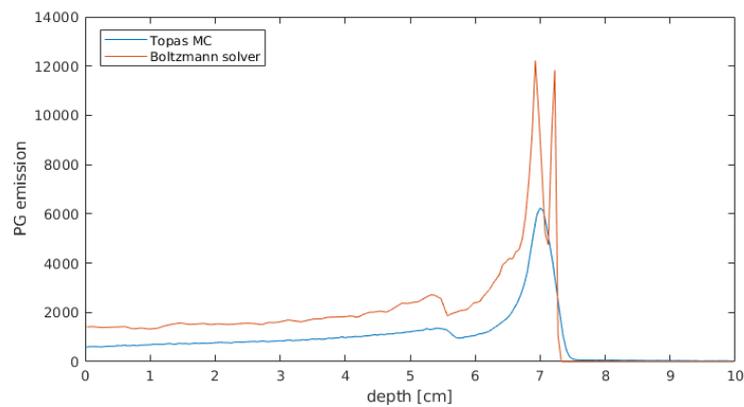
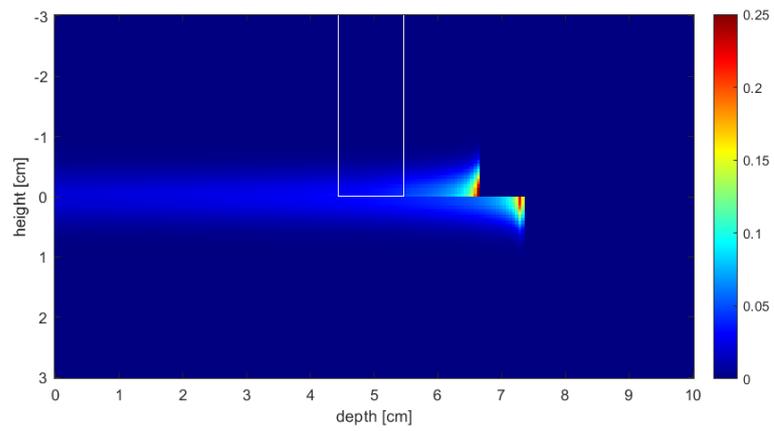
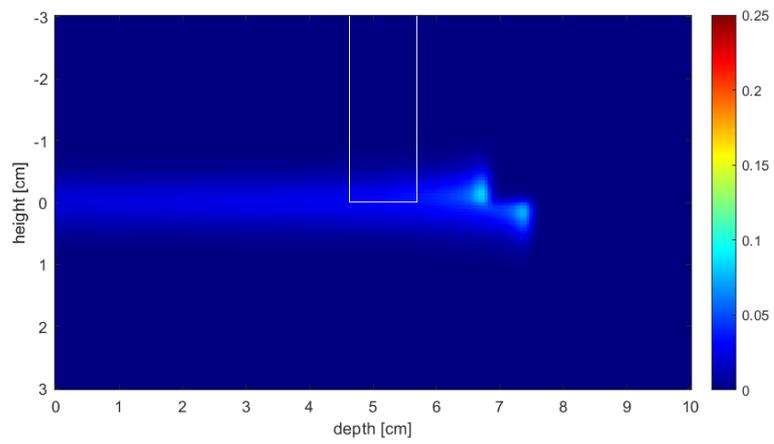


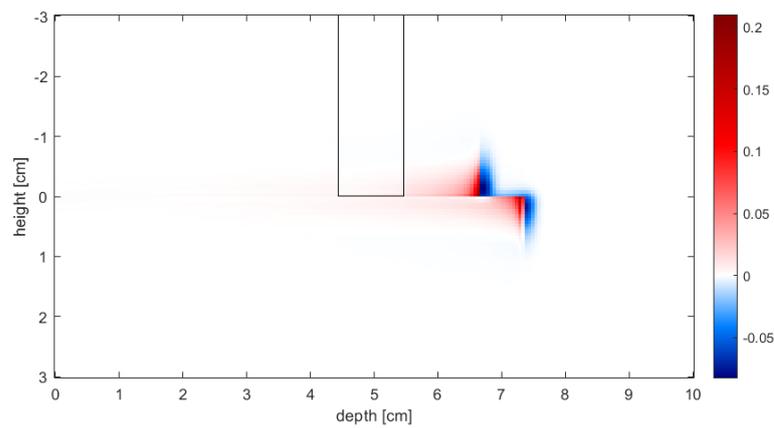
Figure 6.7: Integrated depth activity from the simulation of 100 MeV protons in a homogeneous muscle tissue phantom.



(a)



(b)



(c)

Figure 6.8: Dose [Gy] deposited in the central slice of the muscle phantom from a 100 MeV pencil beam of 10^7 protons, $\sigma = 0.3$ cm. The insert is indicated with an outline. **a.** Simulated by the BS. **b.** Simulated by Topas MC. **c.** Deviation [Gy] of the BS simulation relative to the Topas MC simulation (**a-b**). The red indicates that the BS overestimated the dose relative to Topas MC; blue indicates underestimation.

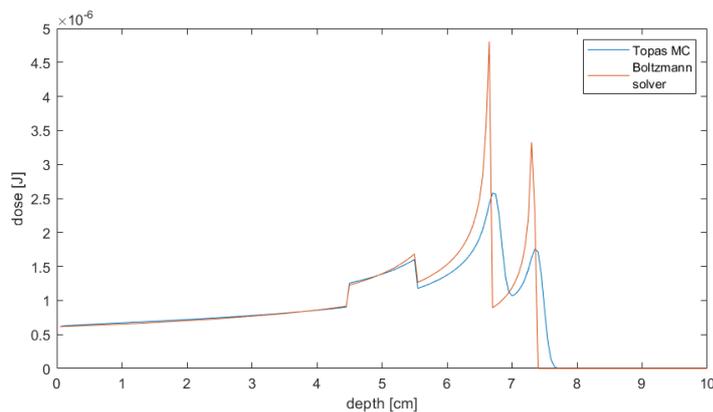


Figure 6.9: Integrated depth dose [J] of a 100MeV pencil beam in an inhomogeneous phantom comprised of skeletal muscle with a compact bone insert, simulated by Topas MC and the Boltzmann solver.

6.3. Full treatment plan on a planning CT

A full treatment plan was run in the BS according to the set-up discussed in section 5.3. The computation time was 0.43 h. The dose distribution is compared to the planned dose from the treatment plan. If the full plan consisting of $3.4 \cdot 10^{12}$ were run in Topas MC, at the same computation rate as the single pencil beam in the homogeneous muscle tissue phantom, the calculation time would be $1.7 \cdot 10^2$ years.

This section discusses the dose deposition in two slices. Slice 46 contains the tumor volume for which a dose of 71.8 Gy was planned. Slice 65 contains lymph nodes with a planned dose of 54.5 Gy.

First of all, the total energy deposited by the BS was 90.3 J, which equals the number of protons multiplied by their energy. However, the energy in the planned dose was much larger: 197.4 J. This number was calculated by converting the planned dose file, which was in [Gy], into [J/voxel] using a mass density map calculated from the planning CT based on Schneider's method and the voxel size. To properly compare the distributions in this section, the dose calculated from BS was multiplied by a factor $197.4/90.3$.

The dose to slice 46 as calculated by the BS is depicted in figure 6.10a and the equivalent from the treatment plan in figure 6.10b. Figure 6.10c shows the absolute difference between these simulations. From these images it stands out that the dose calculated by the treatment plan was homogeneous in the PTV, while BS, in that same area, produced a dose profile with very high dose peaks and underestimation of dose near the edges of the PTV due to range underestimation.

The deviations between the simulations are further illustrated by the γ -indices. Figure 6.11 shows the γ (3 Gy, 3 mm), calculated using the code of Geurts [52]. The figure largely shows $\gamma < 1$. Two areas with high γ stand out: the PTV and two lateral areas near $z = 0$ cm and $z = 40$ cm with $\gamma \approx 2$. The PTV has high γ because the difference between the BS method and MC-based simulations is most severe near the end of range, as shown in the previous sections. The lateral high- γ areas exist because the treatment planning system uses a body contour around the body area in the CT scan, and all voxels outside the contour are excluded in the dose calculation. Such an algorithm does not exist in the BS, so the dose in the air-area is non-zero.

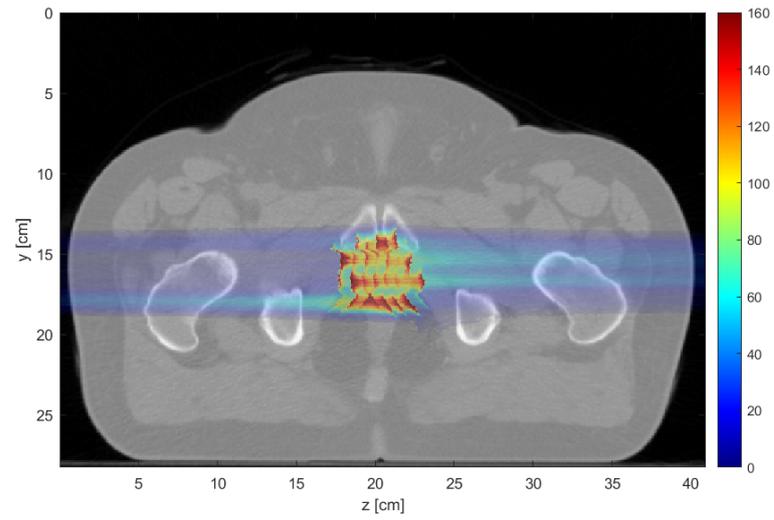
The dose distribution in slice 65 is shown in figure 6.12. Based on the amount of Bragg peaks in the BS distribution, it can be estimated that more beams have crossed this slice than slice 46. The dose is therefore more uniformly distributed, and the deviations from the treatment plan are smaller. However, figure 6.12c shows a significant underestimation of the proton range that was less visible in slice 46.

The dose distribution in slice 46 appears more sensitive for the differences between Topas MC and the BS. Therefore this slice is used to analyze smoothing and lateral beam cut-off in the following sections.

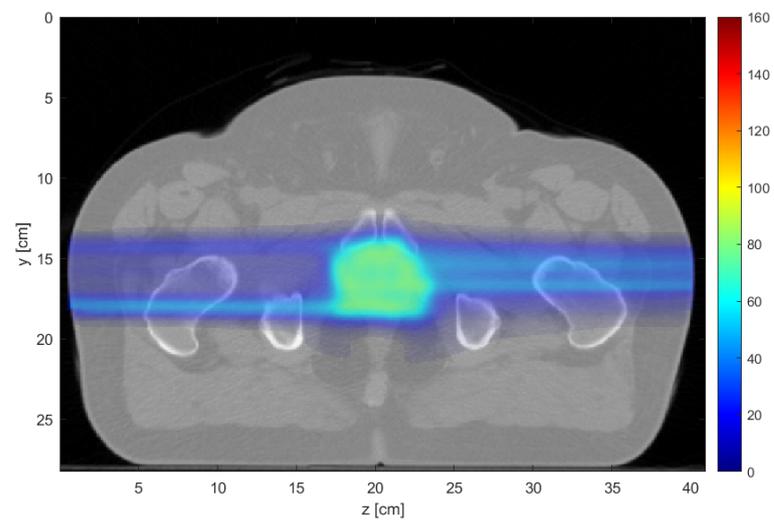
6.3.1. Gaussian blur

To mimic the effect of range straggling and get a more homogeneous covering of the target volume, the whole dose profile was convolved with a 3D Gaussian kernel with $\sigma = 0.3$ cm. The resulting dose profile in slice 46 and the difference the planned dose is shown in figures 6.13a and 6.13b-c.

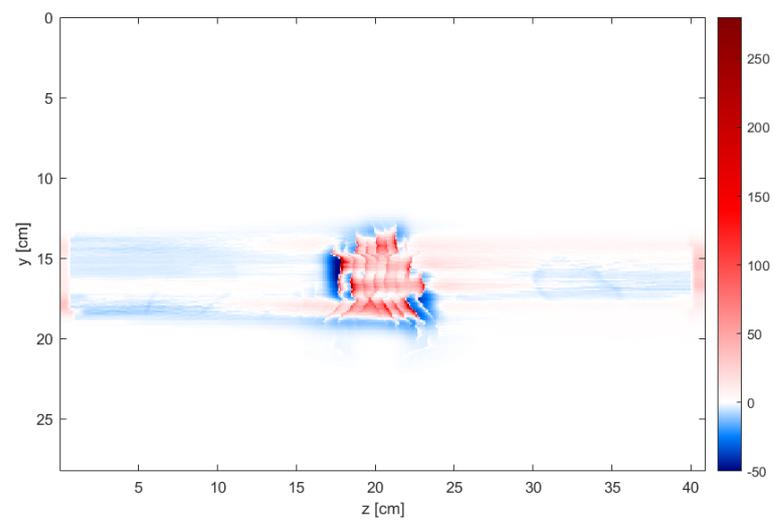
When blurred, the dose is more homogeneously spread over the target volume. However, the dose in the



(a)



(b)



(c)

Figure 6.10: Dose deposited from a full treatment plan. **a.** Dose [Gy] deposited in CT slice 46, simulated by the BS, pasted over an image of CT slice. **b.** Dose [Gy] deposited in CT slice 46 as calculated by the treatment planning system, pasted over an image of CT slice. **c.** Deviation [Gy] of the BS simulation relative to the planned dose (a-b). Red indicates that the BS overestimated the dose relative the planned dose; blue indicates underestimation.

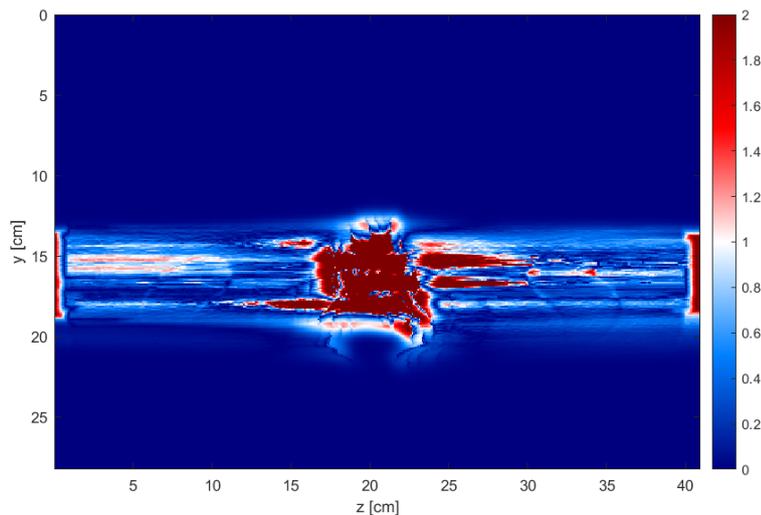


Figure 6.11: γ indices (3 Gy,3 mm) using slice 46 of the planned dose as reference data, and the dose in the same slice simulated by the BS as target data.

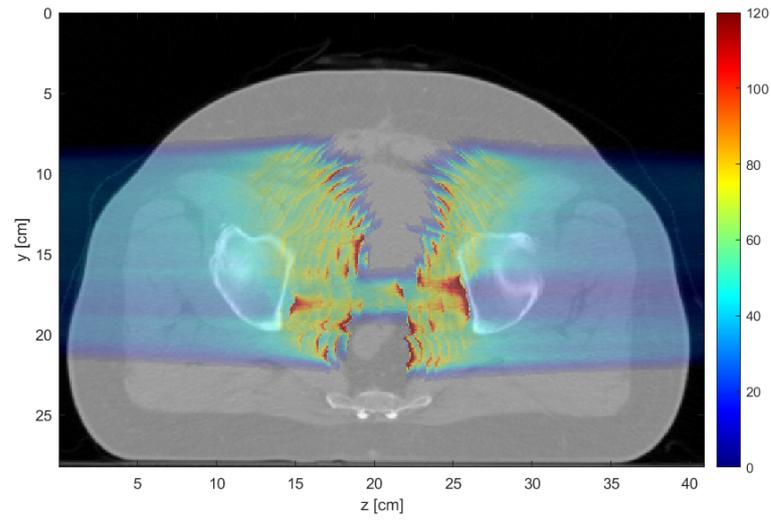
center of the PTV is still overestimated. The blurring is thus not a reliable method to mimic scatter. Meanwhile, the dose on the edges is underestimated. This was to be expected because the blurring does not compromise the short range.

Convolution with a Gaussian kernel does not accurately model the effect of scatter interactions. The spread of the dose depends on the proton energy and tissue characteristics, which are not saved in the dose distribution files produced by the BS.

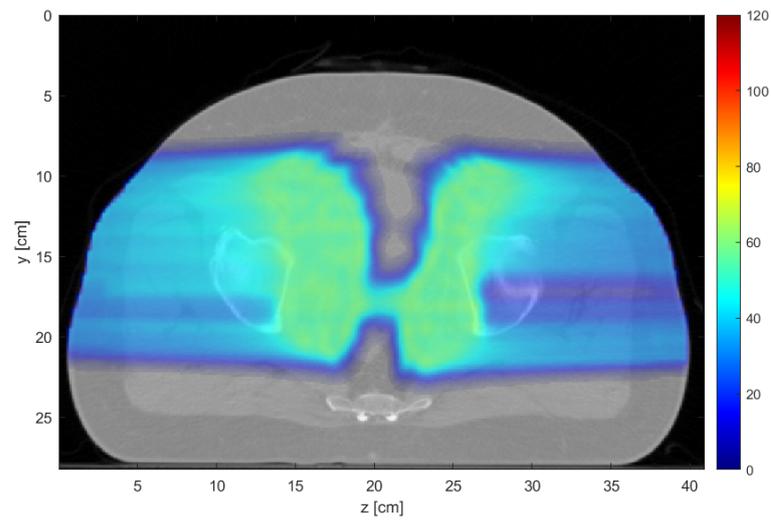
6.3.2. Cut-off radius

The treatment plan was re-run using different cut-off distances, as suggested in section 5.3. The calculation time reduced from 0.43 h for $r_c = 3.8 \sigma$ to 0.31 h for $r_c = 3.2 \sigma$ and to 0.21 h for $r_c = 2.6 \sigma$.

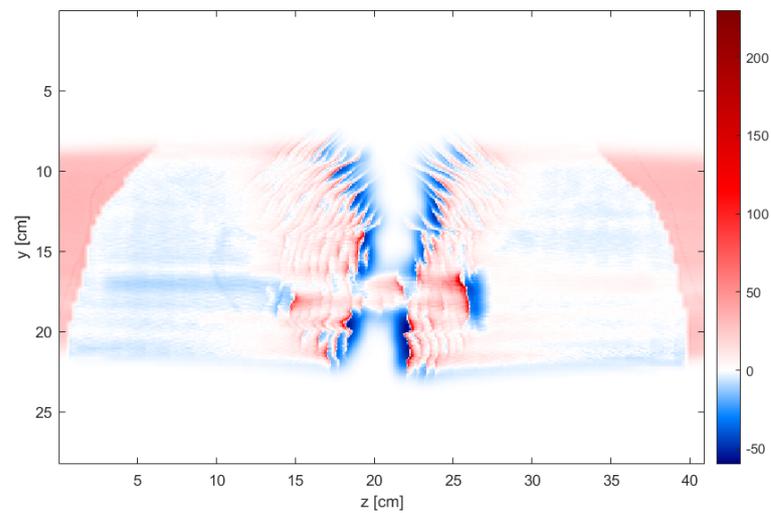
An analysis using the γ -index (1Gy,3mm) is shown in figure 6.14. For the former case, only γ well below 1 are found, which indicates that this cut-off radius yields relatively small deviations. For $r_c = 2.6 \sigma$, the deviations become more profound. The γ values are mostly below 1, except for a few voxels in the PTV, that have $\gamma = 2$. Based on these findings, a smaller cut-off radius of 3.2σ is a good candidate to reduce calculation time at low cost of outcome reliability.



(a)

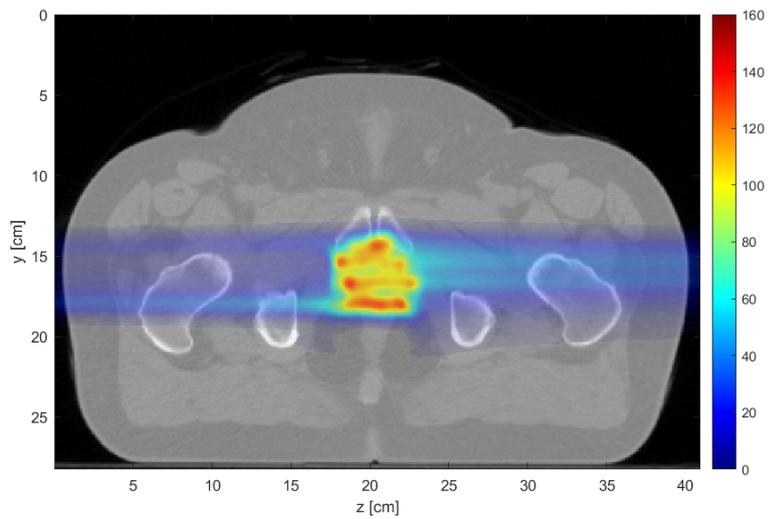


(b)

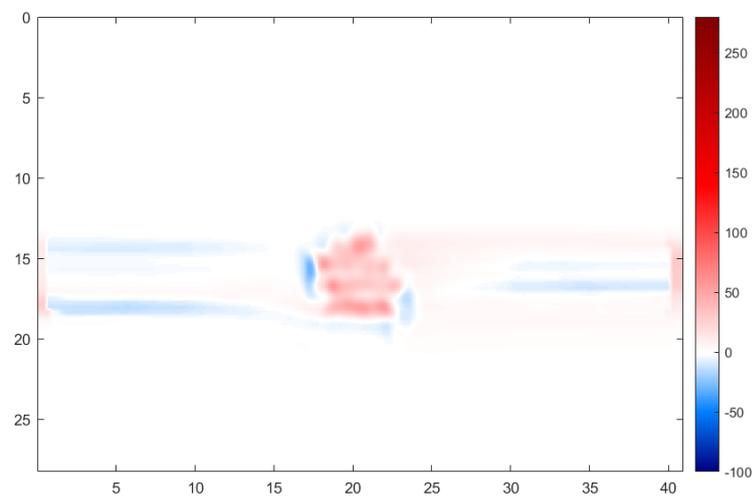


(c)

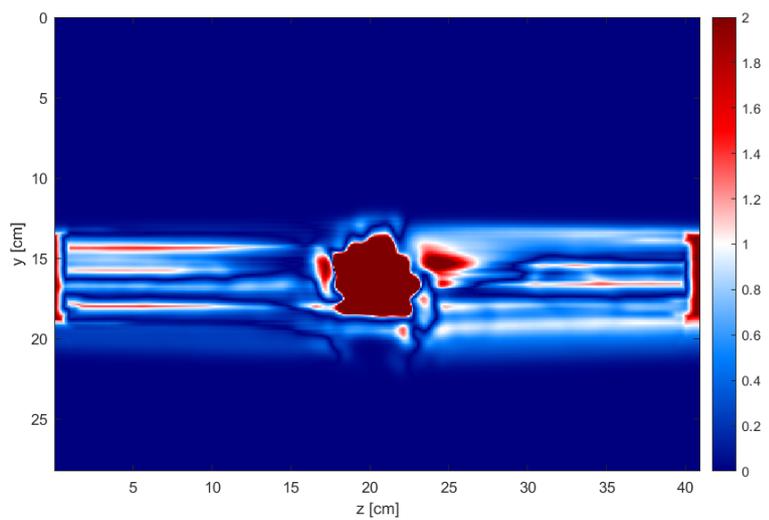
Figure 6.12: Dose deposited from a full treatment plan. **a.** Dose [Gy] deposited in CT slice 65, simulated by the BS, pasted over an image of CT slice. **b.** Dose [Gy] deposited in CT slice 65 as calculated by the treatment planning system, pasted over an image of CT slice. **c.** Deviation [Gy] of the BS simulation relative to the planned dose (a-b). Red indicates that the BS overestimated the dose relative to the planned dose; blue indicates underestimation.



(a)

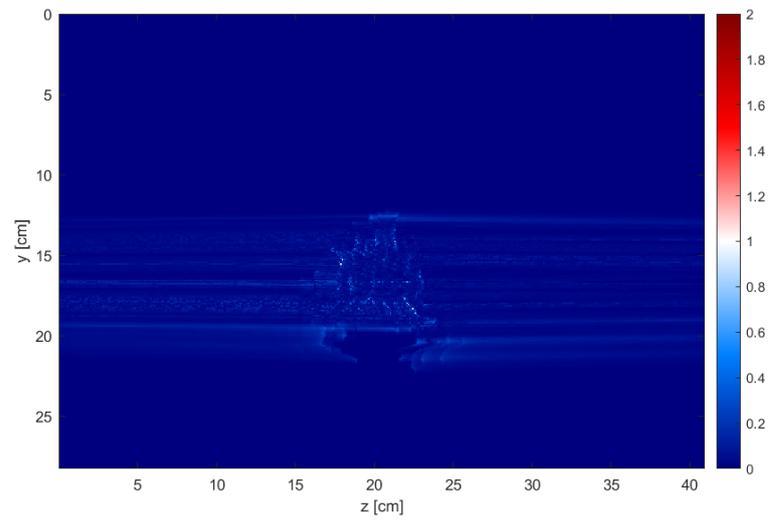


(b)

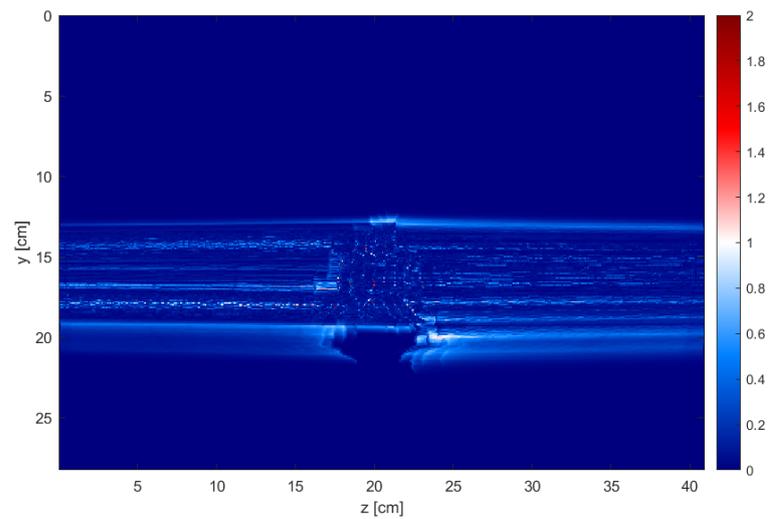


(c)

Figure 6.13: **a.** Dose [Gy] deposited in slice 46, simulated by the BS, pasted over an image of the CT slice. The dose was smoothed with a 3D Gaussian kernel. **b.** Difference [Gy] between the dose in figure **a** relative to figure 6.10b. The red indicates that the BS overestimated the dose relative to the planned dose; blue indicates underestimation. **c.** γ -index (3 Gy, 3 mm) of filtered dose deposition, using the planned dose as reference data set.



(a)



(b)

Figure 6.14: γ -indices (1Gy,3mm) to compare the dose deposited in slice 46 of the CT, simulated by the BS, with different cut-off radii. **a.** $r_c = 3.2\sigma$ relative to $r_c = 3.8\sigma$. **b.** $r_c = 2.8\sigma$ relative to $r_c = 3.8\sigma$.

7

Discussion

Accurate simulation of proton interactions is an important step towards dose control for proton therapy. The MC-based simulation tools that are used frequently in the literature are not clinically applicable because of the long computation time. This study aims to provide proton therapy simulations of high reliability at short calculation times.

The Boltzmann solver simulates the dose profile for a set of proton pencil beams in a 3D grid of which the characteristics are described by a CT image. The pencil beams are converted into mathematical, one-dimensional beams that are forward-traced through the medium until they run out of energy. This study involves the unscattered solutions only. A scattered fluence step may be added, which uses the unscattered fluence as input.

The BS has a powerful mathematical foundation and can be applied to produce 3D distributions. The original goal was to produce PG emission source terms. However, the code is more versatile. Because the proton energy, proton fluence and local material characteristics are known for each location along the beam path, the code can use other cross section libraries to produce emission profiles. It can therefore also be applied to simulate positron emission or the production of other secondary particles.

The reliability of the outcomes of the BS depends fully on the quality of the cross section libraries provided. Cross sections are extremely difficult to measure and therefore no golden standard library exists. This is a burden to the whole field of proton therapy simulation. Therefore the comparison with Topas MC and the planned dose is not indisputable either: deviations between the results from the BS and Topas MC do not necessarily indicate errors in the BS. The extensive validation of the BS code is a complex task and requires further research.

When the BS is provided with a dependable cross section library, it can be applied in a range verification system. These systems in general compare the measured PG profile to a simulated PG profile. Therefore an additional code will be necessary to convert the source terms produced by the BS into a simulated detector readout. However, apart from that, the multiple concepts of range verification methods are studied. This may affect the prerequisites of the Boltzmann solver.

One clinical approach to improve dose control is to commence the administering of a treatment fraction with a few high-intensity proton beams that are part of the treatment plan. The ranges of these beams are verified, and when they are found satisfactory, the remaining portion of the dose is administered. When the BS is to support this method, it only requires the calculation of the first few pencil beams. This means that the calculation time can be relatively long, but the results must be very accurate.

The trade-off between computation time must be different for other applications of the BS. Another approach could be the building of a library of PG source terms for each treatment plan, including various cases of anatomical changes of the patient and set-up errors. In this case the measured data from the full treatment plan will be matched with a case in the library. Because many runs of the BS will be required to build the library, the computation time must be drastically reduced.

Though the BS is a promising code, it requires work to develop it into a clinically applicable tool. The first difficulty encountered with the BS is the interpretation of the treatment plan it simulates. Two aspects have remained unclear over the course of this research: the interpretation of the x- and y-coordinate that define

the beam direction, and the weighting of the beam. The former is expected to cause insignificant deviations and can therefore be ignored. However, the error in the beam weight calculation yields an error in the order of 50% of the planned energy deposition. The multiplication by a single upscaling factor is naive for two reasons. Firstly, it cannot be expected that the dose scored in a MC-based dose distribution exactly equals the energy that was put into the system, because secondary particles may have left the scoring volume. Secondly, the assumption that the weighting method had a linear offset, has no foundation. Because of the severity of the error, the interpretation of the treatment plan should be clarified.

The dose distributions simulated show that the BS underestimates the proton range relative to the calculations of Topas MC. The cross section library employed by the BS contains relatively high stopping powers. Though the use of the SRIM database was carefully considered, it contains $S(E)$ that are significantly larger than e.g. the NIST PSTAR database. Therefore the stopping powers might have to be reconsidered.

Furthermore, the lack of scatter causes both lateral and distal deviations. Because no elastic electronic interactions occur in the Boltzmann solver, all protons deposit their dose on their straight path only, causing the simulation to underestimate laterally delivered dose. When many beams near each other are considered, these lateral deviations become less significant because of the penumbra overlap. Furthermore, the Bragg peaks produced by the BS are higher than those simulated by MC-based codes, causing less homogeneous coverage of dose in high-dose areas of the treatment plan. The results produced by the BS differ from reality but are no source of error because the scattered part of the fluence is to be added to the BS in the future.

To mimic the effect of scatter, the dose distributions were smoothed with a 3D Gaussian kernel. This improved the homogeneity of the high-dose region, but the used kernel did not include enough information about scatter. Firstly, to mimic the distal aspect of range straggling, the kernel should operate in the beam direction only. However, this is complex because the beams all have a slightly different direction.

Secondly, the behaviour of lateral scatter depends on local medium characteristics (mass and component fractions) and the beam energy. Therefore, the shape of the convolution kernel should depend on those characteristics. The development of an algorithm to design sophisticated convolution kernels is complex but it might be rewarding: if these kernels produce good estimations, the addition of a scattered fluence step to the BS may become obsolete.

Furthermore, the analysis of a single-gamma line PG source distribution shows that there is a significant difference between the proton yields simulated by Topas MC and the BS. This is an unexpected result, because the cross sections were produced with the same version of Topas MC. Whether this is due to an error in the BS or a different method to identify PG from the scored photon energy spectrum, is to be studied.

The BS requires much lower computation times than MC-based codes. Whether or not the calculation time of the BS suffices for clinical application, depends on the range verification tool it serves. It holds the potential to become faster. The following methods may yield time reductions in the order of 10¹%. Firstly, the trade-off between the cut-off radius of individual pencil beams and the output quality should be further researched. Especially when the penumbra overlap this holds the potential to be a relatively cheap time reducer.

A second opportunity lies in the reduction of the number of needle beams. To ensure a homogeneous covering over voxels for each individual pencil beam, the amount of needle beams per voxel area was set at 14, but when penumbra overlap, fewer beams may not compromise the outcome.

Furthermore, the looking-up of cross sections in tables can be further optimized. The current tables include low energy levels that are hardly used and reduce loop-up speed. Apart from that, for high proton energies looking up the stopping power for each tracing step might be unnecessary, as the energy levels remains between the same energy levels from the stopping power table.

Lastly, the body contour used in the treatment plan can be applied in the BS to exclude CT voxels that contain air from the tracing volume. Due to the low mass density, the energy loss in air is very low, but the BS handles these voxels at the same computation speed as relevant voxels.

In conclusion, this study added to the development of a simulation tool for proton therapy. The tracing of mathematical pencil beams was found to be the most time-efficient method to model proton interaction in an inhomogeneous medium. For numerical integration and interpolation methods a trade-off between calculation time and reliability of the simulation was made. The code simulates full treatment plans in ~0.4 h, and holds the potential to further reduce calculation time.

Though the developed tool is mathematically well-endorsed, it heavily depends on the reliability of its cross section library, which has not been found satisfactory. Furthermore, significant time reduction can still be achieved.

Bibliography

- [1] J. Efstathiou, P. Gray, and A. Zietman, "Proton beam therapy and localised prostate cancer: current status and controversies," *British journal of cancer*, vol. 108, no. 6, p. 1225, 2013.
- [2] M. Wedenberg and I. Toma-Dasu, "Disregarding rbe variation in treatment plan comparison may lead to bias in favor of proton plans," *Medical physics*, vol. 41, no. 9, p. 091706, 2014.
- [3] M. Goitein, "Calculation of the uncertainty in the dose delivered during radiation therapy," *Medical physics*, vol. 12, no. 5, pp. 608–612, 1985.
- [4] H. Paganetti, "Range uncertainties in proton therapy and the role of monte carlo simulations," *Physics in Medicine & Biology*, vol. 57, no. 11, p. R99, 2012.
- [5] J. Schuemann, S. Dowdell, C. Grassberger, C. Min, and H. Paganetti, "Site-specific range uncertainties caused by dose calculation algorithms for proton therapy," *Physics in Medicine & Biology*, vol. 59, no. 15, p. 4007, 2014.
- [6] A. C. Kraan, "Range verification methods in particle therapy: underlying physics and monte carlo modeling," *Frontiers in oncology*, vol. 5, p. 150, 2015.
- [7] Z. Tan, Y. Xia, M. Zhao, and X. Liu, "Proton inelastic mean free path in a group of bioorganic compounds and water in 0.05–10 mev range—including higher-order corrections," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 268, no. 14, pp. 2337–2342, 2010.
- [8] H.-M. Lu, "A potential method for in vivo range verification in proton therapy treatment," *Physics in Medicine & Biology*, vol. 53, no. 5, p. 1413, 2008.
- [9] H.-M. Lu, "A point dose method for in vivo range verification in proton therapy," *Physics in Medicine & Biology*, vol. 53, no. 23, p. N415, 2008.
- [10] C.-H. Min, C. H. Kim, M.-Y. Youn, and J.-W. Kim, "Prompt gamma measurements for locating the dose falloff region in the proton therapy," *Applied physics letters*, vol. 89, no. 18, p. 183517, 2006.
- [11] K. Parodi, F. Ponisch, and W. Enghardt, "Experimental study on the feasibility of in-beam pet for accurate monitoring of proton therapy," *IEEE Transactions on Nuclear Science*, vol. 52, no. 3, pp. 778–786, 2005.
- [12] M. Moteabbed, S. España, and H. Paganetti, "Monte carlo patient study on the comparison of prompt gamma and pet imaging for range verification in proton therapy," *Physics in Medicine & Biology*, vol. 56, no. 4, p. 1063, 2011.
- [13] K. Parodi, H. Paganetti, H. A. Shih, S. Michaud, J. S. Loeffler, T. F. DeLaney, N. J. Liebsch, J. E. Munzenrider, A. J. Fischman, A. Knopf, *et al.*, "Patient study of in vivo verification of beam delivery and range, using positron emission tomography and computed tomography imaging after proton therapy," *International Journal of Radiation Oncology• Biology• Physics*, vol. 68, no. 3, pp. 920–934, 2007.
- [14] A. Sanchez-Crespo, P. Andreo, and S. A. Larsson, "Positron flight in human tissues and its influence on pet image spatial resolution," *European journal of nuclear medicine and molecular imaging*, vol. 31, no. 1, pp. 44–51, 2004.
- [15] A.-C. Knopf, K. Parodi, H. Paganetti, T. Bortfeld, J. Daartz, M. Engelsman, N. Liebsch, and H. Shih, "Accuracy of proton beam range verification using post-treatment positron emission tomography/computed tomography as function of treatment site," *International Journal of Radiation Oncology• Biology• Physics*, vol. 79, no. 1, pp. 297–304, 2011.

- [16] J. Smeets, F. Roellinghoff, D. Prieels, F. Stichelbaut, A. Benilov, C. Fiorini, R. Peloso, M. Basilavecchia, T. Frizzi, J. Dehaes, *et al.*, “Prompt gamma imaging with a slit camera for real-time range control in proton therapy,” *Physics in Medicine & Biology*, vol. 57, no. 11, p. 3371, 2012.
- [17] P. Gueth, D. Dauvergne, N. Freud, J. Létang, C. Ray, E. Testa, and D. Sarrut, “Machine learning-based patient specific prompt-gamma dose monitoring in proton therapy,” *Physics in Medicine & Biology*, vol. 58, no. 13, p. 4563, 2013.
- [18] E. Draeger, D. Mackin, S. Peterson, H. Chen, S. Avery, S. Beddar, and J. Polf, “3d prompt gamma imaging for proton beam range verification,” *Physics in Medicine & Biology*, vol. 63, no. 3, p. 035019, 2018.
- [19] J. M. Verburg, H. A. Shih, and J. Seco, “Simulation of prompt gamma-ray emission during proton radiotherapy,” *Physics in Medicine & Biology*, vol. 57, no. 17, p. 5459, 2012.
- [20] J. C. Polf, R. Panthi, D. S. Mackin, M. McCleskey, A. Saastamoinen, B. T. Roeder, and S. Beddar, “Measurement of characteristic prompt gamma rays emitted from oxygen and carbon in tissue-equivalent samples during proton beam irradiation,” *Physics in Medicine & Biology*, vol. 58, no. 17, p. 5821, 2013.
- [21] A. Koning and J. Akkermans, “Pre-equilibrium nuclear reactions: An introduction to classical and quantum-mechanical models,” in *Nuclear Reaction Data And Nuclear Reactors-Physics, Design And Safety: Proceedings Of The Workshop*, p. 143, World Scientific, 1999.
- [22] S. Jan, G. Santin, D. Strul, S. Staelens, K. Assie, D. Autret, S. Avner, R. Barbier, M. Bardies, P. Bloomfield, *et al.*, “Gate: a simulation toolkit for pet and spect,” *Physics in Medicine & Biology*, vol. 49, no. 19, p. 4543, 2004.
- [23] J. Perl, J. Shin, J. Schümann, B. Faddegon, and H. Paganetti, “Topas: An innovative proton monte carlo platform for research and clinical applications,” *Medical physics*, vol. 39, no. 11, pp. 6818–6837, 2012.
- [24] T. Aso, A. Kimura, S. Tanaka, H. Yoshida, N. Kanematsu, T. Sasaki, and T. Akagi, “Verification of the dose distributions with geant4 simulation for proton therapy,” *IEEE transactions on nuclear science*, vol. 52, no. 4, pp. 896–901, 2005.
- [25] C. Z. Jarlskog and H. Paganetti, “Physics settings for using the geant4 toolkit in proton therapy,” *IEEE Transactions on nuclear science*, vol. 55, no. 3, pp. 1018–1025, 2008.
- [26] F. Le Foulher, M. Bajard, M. Chevallier, D. Dauvergne, N. Freud, P. Henriquet, S. Karkar, J. Létang, L. Lestand, R. Plescak, *et al.*, “Monte carlo simulations of prompt-gamma emission during carbon ion irradiation,” *IEEE Transactions on Nuclear Science*, vol. 57, no. 5, pp. 2768–2772, 2010.
- [27] M. Pinto, D. Dauvergne, N. Freud, J. Krimmer, J. M. Létang, and E. Testa, “Assessment of geant4 prompt-gamma emission yields in the context of proton therapy monitoring,” *Frontiers in oncology*, vol. 6, p. 10, 2016.
- [28] L. Lestand, G. Montarou, P. Force, and N. Pauna, “In-beam quality assurance using induced β^+ activity in hadrontherapy: a preliminary physical requirements study using geant4,” *Physics in Medicine & Biology*, vol. 57, no. 20, p. 6497, 2012.
- [29] A. Schumann, J. Petzoldt, P. Dendooven, W. Enghardt, C. Golnik, F. Hueso-González, T. Kormoll, G. Pausch, K. Roemer, and F. Fiedler, “Simulation and experimental verification of prompt gamma-ray emissions during proton irradiation,” *Physics in Medicine & Biology*, vol. 60, no. 10, p. 4197, 2015.
- [30] L. Grevillot, T. Frisson, N. Zahra, D. Bertrand, F. Stichelbaut, N. Freud, and D. 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*, vol. 268, no. 20, pp. 3295–3305, 2010.
- [31] B. F. Huisman, J. M. Létang, É. Testa, and D. Sarrut, “Accelerated prompt gamma estimation for clinical proton therapy simulations,” *Physics in Medicine & Biology*, vol. 61, no. 21, p. 7725, 2016.
- [32] M. Testa, C. H. Min, J. M. Verburg, J. Schümann, H.-M. Lu, and H. Paganetti, “Range verification of passively scattered proton beams based on prompt gamma time patterns,” *Physics in Medicine & Biology*, vol. 59, no. 15, p. 4181, 2014.

- [33] L. S. Waters, G. W. McKinney, J. W. Durkee, M. L. Fensin, J. S. Hendricks, M. R. James, R. C. Johns, and D. B. Pelowitz, "The mcnp monte carlo radiation transport code," in *AIP conference Proceedings*, vol. 896, pp. 81–90, AIP, 2007.
- [34] A. K. Biegun, E. Seravalli, P. C. Lopes, I. Rinaldi, M. Pinto, D. C. Oxley, P. Dendooven, F. Verhaegen, K. Parodi, P. Crespo, *et al.*, "Time-of-flight neutron rejection to improve prompt gamma imaging for proton range verification: a simulation study," *Physics in Medicine & Biology*, vol. 57, no. 20, p. 6429, 2012.
- [35] K. Parodi, A. Ferrari, F. Sommerer, and H. Paganetti, "Clinical ct-based calculations of dose and positron emitter distributions in proton therapy using the fluka monte carlo code," *Physics in Medicine & Biology*, vol. 52, no. 12, p. 3369, 2007.
- [36] G. Battistoni, J. Bauer, T. T. Boehlen, F. Cerutti, M. P. Chin, R. Dos Santos Augusto, A. Ferrari, P. G. Ortega, W. Kozłowska, G. Magro, *et al.*, "The fluka code: an accurate simulation tool for particle therapy," *Frontiers in oncology*, vol. 6, p. 116, 2016.
- [37] C. Robert, G. Dedes, G. Battistoni, T. Böhlen, I. Buvat, F. Cerutti, M. Chin, A. Ferrari, P. Gueth, C. Kurz, *et al.*, "Distributions of secondary particles in proton and carbon-ion therapy: a comparison between gate/geant4 and fluka monte carlo codes," *Physics in Medicine & Biology*, vol. 58, no. 9, p. 2879, 2013.
- [38] A. Koning, S. Hilaire, and M. Duijvestijn, "Talys 1.8 - user manual," *A nuclear reaction program, User manual, NRG-1755 ZG PETTEN, The Netherlands*, 2015.
- [39] M. Herman, R. Capote, B. Carlson, P. Obložinský, M. Sin, A. Trkov, H. Wienke, and V. Zerkin, "Empire: nuclear reaction model code system for data evaluation," *Nuclear Data Sheets*, vol. 108, no. 12, pp. 2655–2715, 2007.
- [40] K. Parodi and T. Bortfeld, "A filtering approach based on gaussian–powerlaw convolutions for local pet verification of proton radiotherapy," *Physics in Medicine & Biology*, vol. 51, no. 8, p. 1991, 2006.
- [41] A. Miyatake, T. Nishio, and T. Ogino, "Development of activity pencil beam algorithm using measured distribution data of positron emitter nuclei generated by proton irradiation of targets containing ^{12}C , ^{16}O , and ^{40}Ca nuclei in preparation of clinical application," *Medical physics*, vol. 38, no. 10, pp. 5818–5829, 2011.
- [42] J. J. Duderstadt, L. J. Hamilton, *et al.*, *Nuclear reactor analysis*, vol. 84. Wiley New York, 1976.
- [43] S. B. Uilkema, "Proton therapy planning using the s_n method with the fokker-planck approximation," Master's thesis, TU Delft, 2012.
- [44] C. L. Leakeas and E. W. Larsen, "Generalized fokker-planck approximations of particle transport with highly forward-peaked scattering," *Nuclear science and engineering*, vol. 137, no. 3, pp. 236–250, 2001.
- [45] W. Schneider, T. Bortfeld, and W. Schlegel, "Correlation between ct numbers and tissue parameters needed for monte carlo simulations of clinical dose distributions," *Physics in Medicine & Biology*, vol. 45, no. 2, p. 459, 2000.
- [46] W. H. Bragg and R. Kleeman, "Xxxix. on the α particles of radium, and their loss of range in passing through various atoms and molecules," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 10, no. 57, pp. 318–340, 1905.
- [47] M. Berger, J. Coursey, M. Zucker, and J. Chang, "Stopping-power and range tables for electrons, protons, and helium ions (2005)," Available on <http://physics.nist.gov>, 2015.
- [48] J. Ziegler, J. Biersack, and M. Ziegler, "The stopping and range of ions in matter, srim, 2013," *Actual website: <http://www.srim.org>*, 2013.
- [49] M. H. H. Tolboom, "Using cramér-rao theory to optimize the detection of prompt gamma profiles," Master's thesis, TU Delft, 2017.
- [50] J. Van Kan, A. Segal, and F. Vermolen, "Numerical methods in scientific computing, department of applied mathematics," *Delft University of Technology*, 2008.

-
- [51] S. van de Water, A. Kraan, S. Breedveld, W. Schillemans, D. Teguh, H. Kooy, T. Madden, B. Heijmen, and M. Hoogeman, "Improved efficiency of multi-criteria impt treatment planning using iterative resampling of randomly placed pencil beams," *Physics in Medicine & Biology*, vol. 58, no. 19, p. 6969, 2013.
- [52] M. Geurts, "Calcgamma," Software, 2014.