

DIPLOMARBEIT

Development of a Monte Carlo dose calculation reference tool for proton beam therapy – beam modeling and validation

zur Erlangung des akademischen Grades
Diplom-Ingenieur
im Rahmen des Studiums
Technische Physik

ausgeführt am
Atominstitut der Fakultät für Physik der Technischen Universität Wien
in Zusammenarbeit mit der Abteilung für Radioonkologie
der Medizinischen Universität Wien

unter der Anleitung von
Univ.-Prof. Dipl.-Ing. Dr. Dietmar Georg
(Betreuer)
Dipl.-Ing. Hermann Fuchs, PhD
(Co-Betreuer)

eingereicht von
Christoph Lee, BSc
Matrikelnummer 01126869

Ort, Datum

Unterschrift Verfasser

Unterschrift Betreuer

Contents

1. Motivation and Objectives	1
2. Introduction	3
2.1. Radiotherapy	3
2.1.1. Historical overview of radiotherapy	3
2.1.2. Proton beam therapy	6
2.2. The MedAustron particle therapy accelerator	9
2.3. Treatment planning and quality assurance	14
2.4. Beam optics of charged particle beams	16
2.4.1. Concept and phase-space representation	16
2.4.2. Optical beam parameters	19
2.4.3. Reversible processes and Twiss parameters	22
2.4.4. Irreversible processes	26
2.4.5. Beam characteristics at MedAustron	28
2.5. Interactions of heavy charged particles with matter	30
2.5.1. Interaction processes	31
2.5.2. Stopping power and Bethe-Bloch equation	32
2.5.3. Multiple Coulomb scattering	33
2.5.4. Dose and IDD distributions	35
2.5.5. Range and range-energy relationship	37
2.5.6. Energy straggling and peak width	38
2.6. Monte Carlo methods	41
2.7. Beam model	42
3. Materials	45
3.1. Medical nozzle	45
3.2. Measurement devices	46
3.2.1. Water phantom MP3-PL	46
3.2.2. PinPoint chamber 31014	47
3.2.3. Lynx PT	48
3.3. Software	49
3.3.1. GATE	49
3.3.2. ROOT	50

3.3.3.	MoccaMed Cluster	50
3.3.4.	MATLAB	50
3.3.5.	RayStation TPS	51
4.	Beam modeling	53
4.1.	General concept	53
4.2.	GATE simulations	55
4.3.	Optimization	59
4.3.1.	General definitions	59
4.3.2.	Types of optimization problems	60
5.	Beam energy modeling	63
5.1.	Measurements	64
5.2.	MC simulations	65
5.3.	Automated optimization	67
5.3.1.	Objective functions	68
5.3.2.	Simulation uncertainties	69
5.3.3.	Solver selection and settings	71
5.3.4.	Implementation of optimization tool	74
5.3.5.	Energy optimization	74
5.4.	Beam model	76
5.4.1.	Uncertainties	76
5.4.2.	Model comparison	78
5.4.3.	Model validation	78
6.	Optical beam modeling	85
6.1.	Measurements	86
6.2.	MC simulations	88
6.3.	Automated optimization	90
6.3.1.	Objective function and constraints	90
6.3.2.	Simulation uncertainties	92
6.3.3.	Solver selection and settings	93
6.4.	Beam model	96
6.4.1.	Uncertainties	96
6.4.2.	Model comparison	96
6.4.3.	Model validation	99
7.	Treatment plan recalculation	109
7.1.	Treatment plans	109
7.2.	Simulation setup and settings	110
7.3.	Comparison of dose distributions	112

8. Summary and outlook	117
Bibliography	i
List of Figures	vii
List of Tables	ix
Appendix	xiii
A. Beam models	xiii
B. Remaining plots	xix
C. Calculation of beam parameters from Twiss parameters	xxiii
D. Source codes	xxv

Acknowledgements

I would like to express my sincere gratitude to my supervisor Univ.-Prof. Dr. DI Dietmar Georg for supporting and giving me the opportunity to pursue my Master's thesis as an external student at the Christian Doppler Laboratory for Medical Radiation Research for Radiation Oncology at the Medical University of Vienna and at MedAustron in Wiener Neustadt.

I am deeply grateful to my co-supervisor DI Hermann Fuchs, PhD. I am indebted to you for your continuous engagement and guidance throughout all phases of this work, from introducing me to the topic up to finalizing this thesis. Thank you for your helpful advice and for all the patience during the insightful discussions we had. Without you, this thesis would not have been possible.

My sincere thanks also go to Andreas Resch, MSc, for helping me out whenever I ran into technical problems with GATE or the MoccaMed cluster. Thank you for your prompt responses to my questions and your helpful advice.

I would like to thank Alessio Elia, MSc for the fruitful discussions on beam model parameters at MedAustron. Your advice was always valuable input for this work.

I would also like to thank Mag. Barbara Knäusl, PhD, Dr. Gerd Heilemann and Natalia Kostiukhina, MSc for the collaborative atmosphere during office Thursdays. Special thanks in this respect also to Mag. Peter Kuess, PhD and to my office mate Fatima Padilla, MSc for the welcoming atmosphere and the refreshing coffee breaks.

I would further like to thank all my friends and family members for their support over the past years.

In particular, I cannot find the words to express my gratitude to my wife Daye for her continued support and unfailing encouragement throughout the process of researching and writing this thesis. This would not have been possible without the countless sacrifices you have made. I will be forever grateful for your love.

Finally, I want to thank God for creating this magnificent universe and allowing it to be studied.

Abstract

Independent dose calculations (IDC) for quality assurance in proton beam therapy treatment planning become more and more interesting these days. As opposed to quality assurance (QA) measurements, IDCs free up valuable beam time for patient treatment and research and can help to increase the treatment efficiency of a therapy facility. A popular method for *precise* independent dose calculations are Monte Carlo (MC) simulations. Before the actual IDC can be performed a beam model is necessary to fully exploit the accuracy of MC simulations.

The aim of this thesis was to find a beam model for the research beamline at MedAustron for the MC toolkit GATE. Since beam modeling is usually performed manually and is thus a time-consuming and tedious task, a novel optimization tool was used that automated the process of finding a beam model. Beam modeling was split into two parts: beam energy modeling and optical beam modeling.

For beam energy modeling, proton beams (62.4, 97.4, 148.2, 198.0, 252.7 MeV) stopping in a 40x40x42 cm² water phantom were simulated using GATE v8.1. Resulting simulated Bragg curves were analyzed with regard to range and Bragg peak width and fitted to measured Bragg curves. Corresponding measurements were performed using an MP3-PL water phantom (PTW, Freiburg, Germany) incl. Bragg peak chambers.

For optical beam modeling, proton beams (with same energies as for beam energy modeling) traversing phase-space actors at selected isocenter distances (ISDs) were simulated using GATE v8.0. The FWHM (full width at half maximum) of the transverse intensity profiles was extracted and the simulated FWHM-vs-ISD curves were fitted to measured curves. Corresponding measurements were performed using a Lynx PT detector (IBA dosimetry, Schwarzenbruck, Germany).

Using the obtained beam model as basis, a special QA box treatment plan in water from the RayStation (RaySearch Laboratories, Sweden) treatment planning system (TPS) was recalculated in GATE. The resulting dose distribution was compared to the dose distributions predicted by the TPS.

The optimized beam parameters showed clinically acceptable agreement when validated with the respective measurements. Maximal deviations for Bragg curve ranges were less than 0.1%. Deviations in Bragg peak width reached up to 6%. Most absolute deviations for the FWHM lay within +0.5 mm and -0.5 mm. The

corresponding relative deviations for the FWHM lay within +4% and -6%. Comparison of the recalculated dose distribution in GATE with the TPS-predicted distribution showed relative deviations in dose of less than 2.5% for most cases.

Therefore, the obtained beam model can be used as the basis of an independent dose calculation tool for the research beam line at MedAustron.

The optimization tool allowed to automatically find an accurate beam model within 35 to 40 hours on a conventional notebook. Employment of the tool at MedAustron for carbon ions is planned where it will help to reduce the amount of try-and-error based beam modeling sessions. Current work in progress deals with testing the automated beam modeling approach also at other particle therapy centers.

1. Motivation and Objectives

Proton beam therapy exploits the physical properties of charged particles to create highly conformal dose distributions around the target volume, thereby sparing healthy tissue especially in the low- and mid-dose region. The high precision leads to an increased sensitivity towards changes within the patients' anatomy. In order to take into account anatomical changes throughout the therapy, proton beam therapy is often combined with adaptive radiotherapy (ART). In ART the treatment plan is re-optimized each time as reaction to patient-specific changes that are unaccounted by the initial treatment plan. As a result, quality assurance for treatment planning plays an essential role in ART. According to quality assurance standards, every initial as well as adapted treatment plan has to be checked against an independent reference method for possible errors in dose prediction prior to patient treatment. Several possibilities exist for this independent method, such as quality assurance measurements. However, nowadays independent dose calculations become more and more favorable. In clinical practice, a wide variety of dose calculation approaches exists, ranging from fast analytic algorithms to precise full Monte Carlo (MC) dose calculation. A popular framework for Monte Carlo simulations that is often used for medical applications is the simulation environment GATE (GEANT4 Application for Tomographic Emissions).

At the Austrian particle therapy facility MedAustron treatments and research with protons have recently started, while carbon ion therapy is planned to be available in the near future. At MedAustron researchers have the possibility to use a research version of the treatment planning system (TPS) RayStation (Raysearch Laboratories, Sweden) including novel tools for ART that are not released for clinical use yet. In order to evaluate these novel tools and to have an independent dose calculation tool for quality assurance purposes the Monte Carlo simulation toolkit GATE will be used as reference. In order to fully exploit the potential of precise Monte Carlo calculations a precise beam model is required.

The aim of this thesis was to develop a beam model for the research beam line at MedAustron in the Monte Carlo toolkit GATE. Since the process of beam modeling is usually carried out manually, it is a rather time-consuming and tedious task. For this reason, a novel automated optimization tool was developed that helps to find the optimal beam model given the respective sets of measured beam data.

The auto-optimized beam model was validated by comparison with measurement data. Furthermore, the MC dose calculation based on the established beam model was compared with the dose calculations performed with the RayStation TPS.

2. Introduction

2.1. Radiotherapy

Radiotherapy uses ionizing radiation to treat tumors [62]. Ionizing radiation, i.e. X-rays or particle radiation with a sufficient high energy, can break the chemical bonds of DNA in a patient's tumor cells. The breaking of a DNA bond can be a result of either direct ionization or indirect action that is caused by chemical reactions between the DNA and the products that are created when ionizing radiation acts on water molecules. If the damage to the DNA is beyond repair, the tumor cells stop dividing which further on can eliminate the tumor. Thus, radiotherapy can be used with a curative intent in order to eliminate a tumor. Furthermore, it can also be used with a palliative intent in order to improve a patient's quality of life [61]. A challenging part of radiotherapy is to deliver the maximal dose necessary to kill the tumor while sparing the healthy surrounding tissue. In order to reduce side effects, the dose deposited in normal tissue must not exceed a certain limit that differs for each type of tissue. This also sets constraints for treatment planning. In the past 20 years advances in technology have pushed the accuracy of dose deposition to a high level and made it possible to greatly reduce the dose to healthy tissue [62].

Radiation can be delivered internally or externally [30]. In internal radiotherapy (brachytherapy) a radiation source, often a radioactive material, is brought near or into the tumor. In contrast, external radiotherapy (teletherapy) uses external beams emerging from a source that is located at a certain distance from the patient. In the following a brief historical overview of external radiotherapy will be given in which two external therapy modalities, photon therapy (PT) and proton beam therapy (PBT), will be introduced and compared in more detail later on.

2.1.1. Historical overview of radiotherapy

In 1895 the German physicist Wilhelm C. Röntgen discovered a certain type of radiation that penetrates bones and human tissue [53]. Soon after the discovery of these so-called X-rays, it became apparent that this radiation could not only

be used for imaging and diagnostic purposes but also for treatment. In 1896, the Viennese physician Leopold Freund was the first one to treat a young patient's skin disease with X-rays [16]. Due to the limited knowledge about radiobiological and genetic effects and the absence of accurate dose measurement and calculation techniques, early patients had to suffer from severe side effects.

In 1928 the International Commission on Radiological Protection (ICRP) was founded to address these issues of radioprotection. From 1930 to 1950, the ongoing development of linear accelerators and the invention of powerful X-ray tubes allowed to produce X-rays up to 200keV which made treatment of deeper seated tumors possible. In the 1950s Cobalt therapy began to emerge in which γ -rays emitted from a cobalt-60 isotope are used to irradiate tumorous tissue. During the 1970s powerful medical LINACs (linear accelerators) were developed that made the production of MeV X-rays possible. Together with the emergence of these LINACs the development of advanced radiotherapy techniques like IMRT (intensity-modulated radiation therapy) and IGRT (image-guided radiotherapy) started. Photon therapy, as one possible modality of radiotherapy, became a recognized medical discipline [20].

In 1946, however, Robert Wilson brought up an entirely new concept. He proposed that accelerator-produced proton beams could be a favorable candidate for highly-localized tumor treatment. The reason is the high dose deposition occurring toward the end of a protons range in tissue [64]. Already in 1954 this modality, called proton beam therapy, was used to treat the first human at the Lawrence Berkeley Laboratory [35]. In the following 40 years, the improvement of key technologies for PBT e.g. synchrotrons, magnetically scanned proton beams and treatment planning systems led the way to the first commercial proton delivery systems that appeared in 2001 [42].

In the following section, the way of how photon therapy works and is currently applied will be described. Afterward the same will be done for proton beam therapy and the advantages and disadvantages between these two modalities will be contrasted.

Photon therapy

Nowadays in photon therapy, X-rays are produced in a linear accelerator (LINAC) which accelerates electrons up to the MeV range and lets them impinge on an anode [58]. As a result, characteristic X-rays and Bremsstrahlung with energies ranging from keV to MeV are produced. The X-ray beam then passes a flattening filter and collimators that flatten the dose profil of the beam and define the beam dimensions (see figure 2.1). Afterwards, the beam propagates toward the patient.

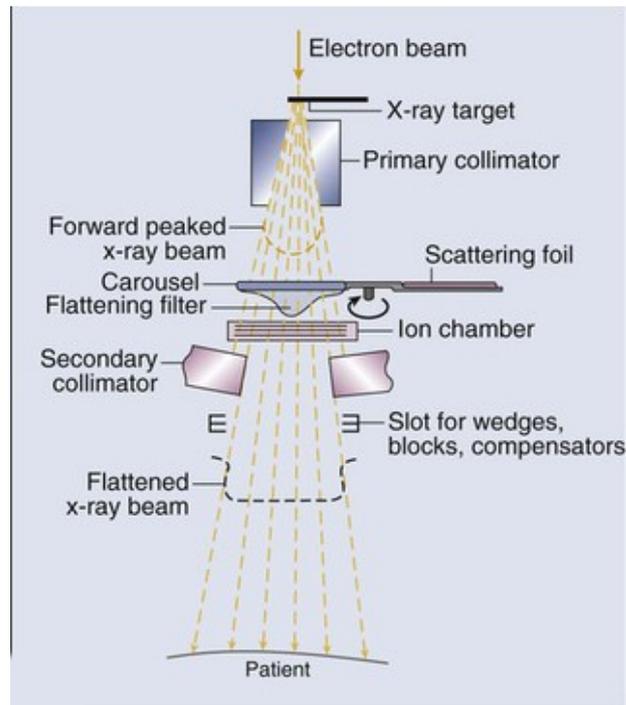


Figure 2.1.: Schematic of the treatment head of a LINAC for photon therapy. Image source: [65]

When the X-ray beam traverses the patient's tissue, it continuously deposits dose along its way. The amount of dose deposited as a function of the penetration depth can be displayed in a so-called integrated¹ depth-dose (IDD) distribution. The typical IDD profile of an X-ray beam in tissue is shown in figure 2.2. One can observe a build-up of dose up to the maximum dose at a depth of around 2.5 cm. After that point the dose decreases exponentially with penetration depth, which means that photons do not exhibit a distinct range in a patient's body. This is due to the photon-matter interaction: the major three processes of photon-matter interaction are the photoelectric effect, Compton scattering and pair production. The nature of these processes is stochastic which means a photon undergoes these processes with a certain probability. However, these interaction probabilities do not change significantly as the photons traverse through the tissue, as it is the case with charged particles (for Bragg peak, see subsection 2.5.4) [14]. This implies that the tissue beyond will also absorb a significant amount of dose from the X-ray beam that exits the patient, the so-called exit dose. In order to relatively

¹integrated, because the 3D dose distribution resulting from the penetrating radiation is integrated laterally such that the resulting integrated dose is only a function of penetration depth

decrease the exit dose and homogeneously deposit dose within the target volume (i.e. the volume to be irradiated within the patient), multiple X-ray beams from different directions and angles are used in clinical practice.

Prior to any treatment in radiotherapy, the respective clinical target volume has to be delineated for each patient individually. This is achieved using specific imaging methods, e.g. computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET) or ultrasound, or a combination of these. In particular, a 3D-CT scan yields the electron densities of the patient's tissue based on which dose calculations can be performed in order to determine the optimal treatment plan [47].

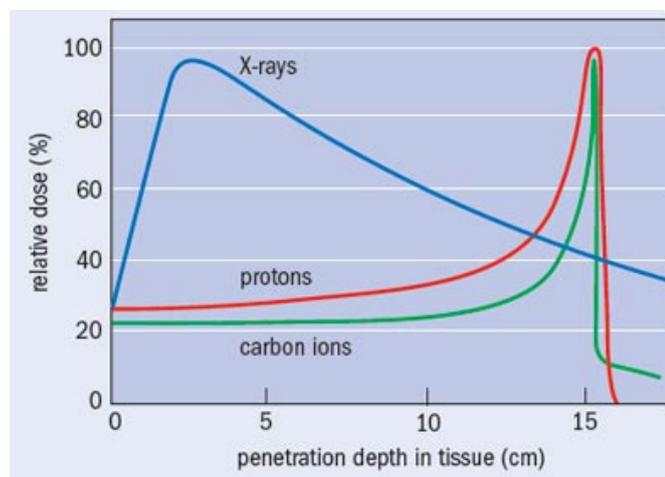


Figure 2.2.: Comparison of depth-dose profiles for a monochromatic photon, proton and carbon ion beam. Image source: [39]

2.1.2. Proton beam therapy

In proton beam therapy a synchrotron or cyclotron is used to accelerate protons to tens and hundreds of MeV such they can be used for tumor treatment. The produced proton beam is directed to the target volume in the patient and deposits dose as it traverses through the patient's tissue. The resulting effects that occur in comparison to photon therapy can be categorized whether they emerge from the physical characteristics or the radiobiological characteristics of protons. In the following the advantages and disadvantages of PBT will be discussed from these two viewpoints.

Physical characteristics

As mentioned before, a proton deposits most of its energy towards the end of its trajectory [31]. This phenomenon is a result of a charged particle's *physical characteristics* when traversing through matter. As opposed to photons, charged particles are increasingly more likely to interact with matter as they slow down. The accumulation of deposited dose right before the end causes a distinct dose peak, commonly referred to as Bragg peak (see figure 2.2). The dose increases only slowly in the region preceding the Bragg peak, whereas beyond the Bragg peak one can observe a steep dose fall-off to almost zero dose.

The position of the Bragg peak depends on the initial energy of the particles before penetrating the tissue. Therefore, the Bragg peak can be positioned at almost every point within the body if the initial energy and the lateral offset are accordingly varied. In clinical practice, the Bragg peak of a monochromatic proton beam is usually too narrow to cover the target volume homogeneously. For this reason several Bragg peaks from beams with different initial energies and intensities are superimposed to form a so-called spread-out Bragg peak (SOBP) [14]. The SOBP is shaped in such a way that the resulting dose plateau covers the target volume homogeneously (see figure 2.3). The Bragg peak enables a more confor-

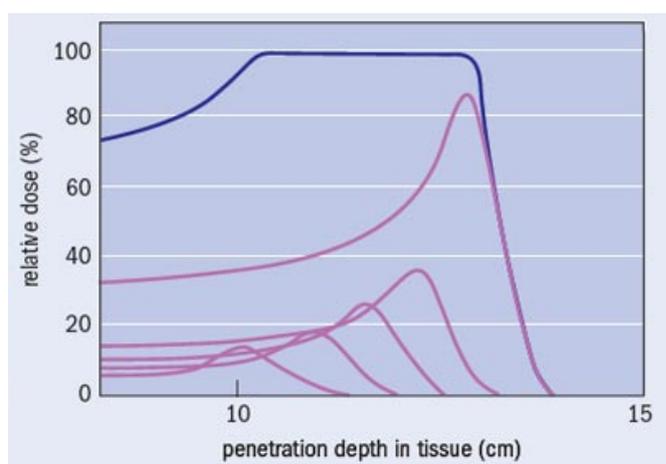


Figure 2.3.: Formation of a spread-out Bragg peak (SOBP) by superposition of Bragg peaks from monochromatic proton beams of different energy and intensity. Image source: [39]

mal dose distribution compared to photon therapy: it allows to escalate the dose within the target volume while minimizing the dose in healthy surrounding tissue,

completely eliminating exit dose². This constitutes the main clinical advantage of ion beam therapy over conventional photon therapy [63]. The higher accuracy in dose deposition allows to treat tumors close to organs at risk (OAR) and thus can facilitate the treatment of tumors that are difficult to treat surgically or with photon therapy. Furthermore, the steep dose gradient allows to reduce the overall dose deposited in the patient compared to photon therapy. This can be important for pediatric patients who are more susceptible for developing secondary tumors as a result of radiotherapy treatment.

Nevertheless, the physical characteristics of protons also imply certain disadvantages in PBT. The highly conformal dose distributions make treatment very sensitive to patient and organ motion. As a result, accuracy is more vital than in conventional photon therapy. It is crucial to take changes in anatomy during and in between treatment sessions into account. Breathing control is indispensable for a precise treatment of lung or abdominal tumors. Also other effects like tumor shrinkage throughout the therapy or weight loss of the patient have to be considered in order to avoid overshooting the Bragg peak into a critical structure [56]. Especially reproducible patient immobilization is a big issue. Similar to conventional PT, imaging techniques like CT and/or MRI are used prior to PBT treatment to obtain body scans which allow to delineate the target volume. After this stage of imaging, the patient gets irradiated. Once a patient positioning system cannot guarantee reproducible and sufficiently high accuracy between the stages of imaging and irradiation, the conformal dose distributions of PBT do not longer provide advantage³ [63].

Radiobiological characteristics

Up to now only the advantages of PBT resulting from its *physical characteristics* were discussed. However, there are also advantages emerging from a *radiobiological* perspective. Since protons show different ionization properties in comparison to photons, the same amount of dose deposited will result in different biological effects for ions and photons. A measure that quantifies the biological effect of a specific kind of radiation is the relative biological effectiveness (RBE). For a given type of radiation i , the RBE is defined as the ratio of the dose D_{ref} of a reference radiation producing a specific biological effect to the dose D_i of radiation i necessary to

²Exit dose is only eliminated for protons. In the case of carbon ions considerable dose deposition also occurs in the so-called fragmentation tail beyond the Bragg peak

³this is also true for the patient positioning systems used in PT

produce the same biological effect under the same conditions [31]

$$RBE = \left(\frac{D_{ref}}{D_i} \right)_{\text{same biol. effect}} \quad (2.1)$$

These days most proton therapy centers use a generic RBE value of 1.1 for protons [28], with photons used as reference radiation. This means that a dose delivered by protons causes the same biological effect as a 10% higher photon dose. This means that in PBT less physical dose needs to be given to the patient for the same biological effect.

It shall be mentioned here that carbon ions exhibit an even higher RBE than protons which is due to their increased mass. The RBE can take on values from 2.3 to 3.0 along the SOBP [60]. For that and other various reasons, also carbon ions were suggested for radiotherapy and in 1993, the first therapy-dedicated research facility for carbon ion radiotherapy (CIRT) was constructed in Japan [60].

Nowadays there are 79 particle beam therapy facilities under operation, 68 of which are proton beam facilities and 11 of which are carbon ion beam facilities [49]. Despite the advantages of PBT over conventional PT from a *physical* and *radiobiological* viewpoint, conventional photon therapy is still today the more cost-competitive and cost-effective modality [42]. In PBT the average costs per patient amount up to 20,000€ [12] which is about 2-3x more than for modern photon therapy. A more detailed break down of patient and investment costs can be found, for example, in [48].

2.2. The MedAustron particle therapy accelerator

MedAustron is an ion beam therapy and research facility situated in Wr. Neustadt, Austria. The idea of MedAustron goes back to the 1990s. Its construction eventually started in 2011, followed by a test-run phase. The first cancer patients were irradiated with protons in December 2016 and patient treatment using $^{12}\text{C}^{6+}$ carbon ions is planned to begin in the near future. The following section will give a brief overview about MedAustron and its facilities based on [4].

The main focus at MedAustron currently lies in the clinical treatment of cancer using proton beam therapy. Besides cancer therapy, MedAustron also engages in research, both clinical and non-clinical. The clinical side of research comprises effectiveness studies of PBT and CIRT and its side effects. Additionally, effort

is devoted to improve existing therapy techniques and develop new ones. The non-clinical side of research encompasses research in the fields of radiobiology and medical physics.

For proton beam therapy at MedAustron protons with a kinetic energy ranging from 60MeV to 250MeV are available. For carbon ion radiotherapy carbon ions will be accelerated to kinetic energies ranging from 120MeV/u to 400MeV/u. 250MeV protons have a velocity of around 0.6c and 400MeV/u carbon ions a velocity of about 0.7c. In order to accelerate protons and carbon ions to such relativistic energies, a synchrotron with pre-accelerator is used.

The layout of the MedAustron particle therapy accelerator (MAPTA) is shown in figure 2.4.

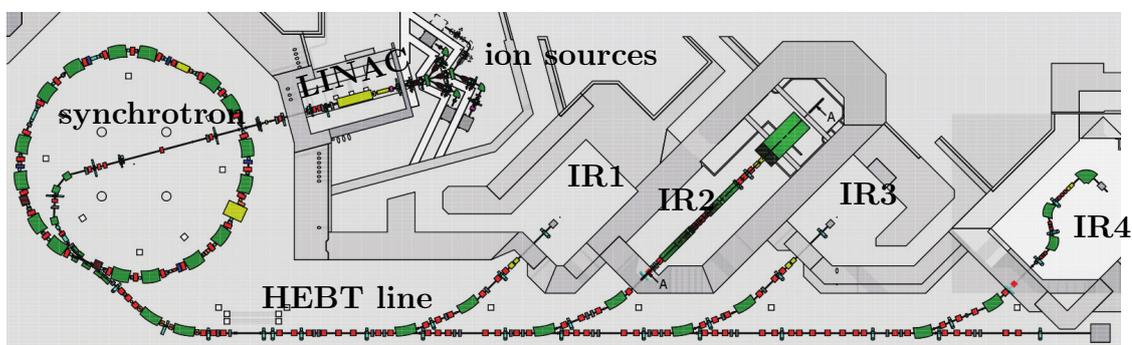


Figure 2.4.: Layout of the MedAustron particle therapy accelerator. Illustration is not to scale. The synchrotron has a circumference of about 77m. In comparison, the world's largest synchrotron, the LHC, has a circumference of about 27km. Image source: [19]

There are three ion sources that provide either H_2^+ or C^{4+} ions. These are pre-accelerated in the so-called low energy beam transfer (LEBT) line to 8keV and formed to a pulsed ion beam by passing a high-frequency chopper. Afterward, the ions are accelerated in a linear accelerator (LINAC) to 7MeV/u. In order to achieve the ion charge desired, the ions traverse a fixed target stripper foil. As a result, one electron is stripped from the H_2^+ or two from the C^{4+} ion. Hence, H^+ or C^{6+} ions are produced. Before the beam is injected into the synchrotron it passes a debuncher and a degrader. The debuncher allows to widen the beam pulse in longitudinal direction⁴ which results in a reduction of the spread Δp_z of the beam's longitudinal momentum distribution, thereby making a variation of the beam's energy spread ΔE possible. The degrader is comprised of 3 movable

⁴that is, in direction of propagation

copper plates with a regular pattern of holes having different diameters. Depending on which plate is being used, the beam current can be reduced by 50%, 80% or 90%, thereby allowing to vary the beam intensity without affecting other beam parameters. Once the beam is completely injected into the synchrotron, it is accelerated to clinical energies, that is, protons to 60 to 250MeV and carbon ions to 120MeV/u to 400MeV/u. Similar to most other particle accelerators, this is achieved using radiofrequency cavities. The frequency of the accelerating electric fields has to be increased proportionally to the circular frequency of the ions in the synchrotron. In order to keep the ions on circular track, bending magnets in the form of dipole magnets are used. Also the magnetic field of the bending magnets has to be increased synchronously with increasing kinetic energy of the ion beam. A magnetic structure consisting alternately of focusing and defocusing quadrupole magnets guarantee an overall focusing effect on the ion beam along the orbit. In order for the synchrotron to work properly, ultra high vacuum of 10^{-9} mbar is necessary.

The most important aspect for clinical operation regarding particle acceleration, however, is the extraction of the beam from the synchrotron into the high energy beam transfer (HEBT) line. In order to allow a precise dose delivery and to allow for online-dose measurements, the beam is extracted using the method of *resonant slow extraction*. As opposed to other extraction methods it produces relative stable beam intensities [66]. *Slow* refers to the rather long extraction time of about 5 seconds which extends over several revolutions in which the beam is gradually extracted from the synchrotron. For 250MeV protons the time period for one turn is about 0.4 μ s, consequently the beam turns 12.5 million times before it is completely extracted. Since a proton spill consists of around 10^{10} protons, the number of extracted protons per turn amounts only up to around 4000⁵. The word *resonant* refers to the way the beam is extracted. The single ions of an ion beam revolving in a synchrotron usually do not follow strictly a circular orbit but oscillate around the design orbit (see figure 2.5). These oscillations exist in both vertical and horizontal planes and are called betatron oscillations. This oscillating system of a single ion can be driven into resonance with the aid of a sextupole magnet and a betatron magnet. When the system approaches resonance, the amplitude of the Betatron oscillations increases gradually up to the point where the oscillations are so strong that the ion reaches the electrostatic septum⁶ which extracts the ion from the synchrotron (see figure 2.6). One can think of this extraction mechanism as a process of gradually "peeling off" the particles of the beam from outer to

⁵assuming a constant extraction rate

⁶a electrostatic septum is a device that separates the field regions of the circulating and extracted beam [5]

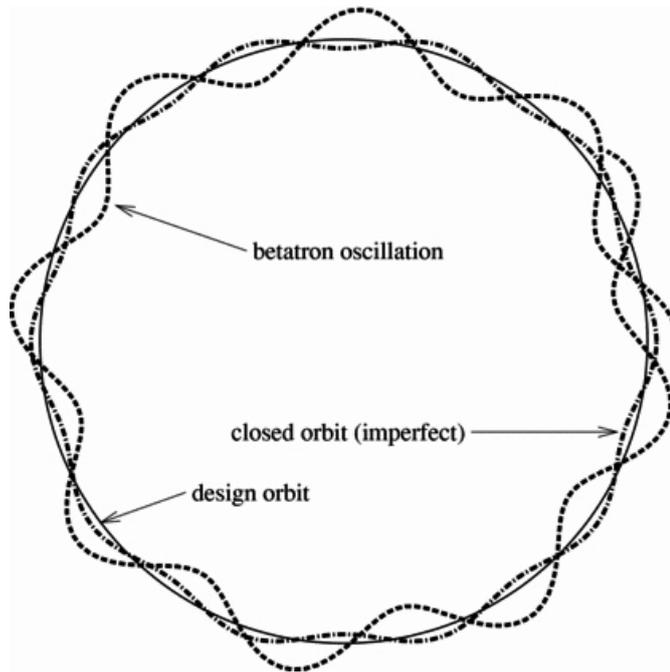


Figure 2.5.: Betatron oscillations of a single ion shown in the horizontal plane (y - z -plane). The oscillations typically form a non-closed trajectory. Image source: [38]

inner. Since this "peeling process" only happens in the horizontal plane, the beam intensity profiles along the horizontal and vertical axis differ (see section 2.4.5 for details). This difference is an important factor to take into account during beam modeling.

Once the beam is extracted, it is transported along the HEBT line toward the four irradiation rooms (IRs). Upon entering the individual IRs the beam has to pass two important devices prior to impinging on the patient's tissue.

First of all, it has to pass a pair of magnets that deflect the beam in both horizontal and vertical direction and thus allow the beam to be scanned laterally over the target volume. This technique is called *pencil beam scanning*. The term *pencil beam* describes the form of the narrow cylindrical beam pulses that are delivered to the target volume.

Secondly, after having passed the scanning magnets, the beam has to pass a medical nozzle. The nozzle acts as a control device that monitors the dose that is currently delivered and the current spot position of the beam. It also contains passive elements to modify the beam (for detailed information on the nozzle, see section 3.1).

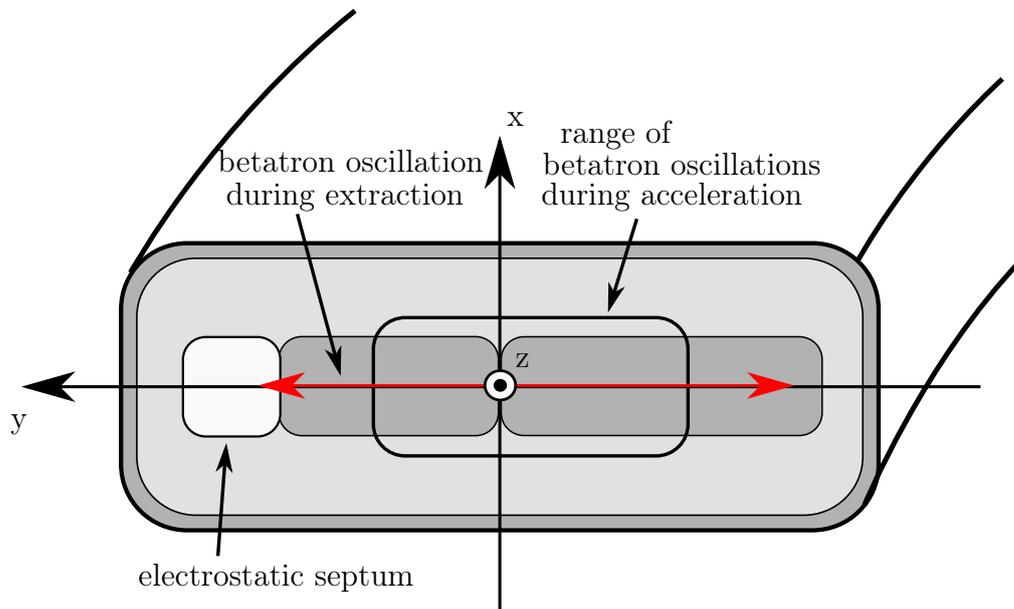


Figure 2.6.: Schematic showing the synchrotron cross section. The range of betatron oscillations is increased during the extraction process such that ions reach the electrostatic septum and thus are gradually extracted from the synchrotron ("peeling process").

After passing the nozzle the beam freely propagates toward the isocenter⁷ where the target volume is positioned. Figure 2.7 illustrates all this for the situation in IR2. The horizontal distance of a point on the undeflected beamline⁸ to the isocenter is referred to as isocenter distance (ISD). According to the convention used the ISD gets negative going downstream and positive going upstream.

The situation in other IRs is similar to the one described above. It shall be mentioned, however, that IR1 and IR3 only host a horizontal beamline, but no vertical one. In contrast, IR4 hosts a Gantry for proton treatment which is a rotating device that allows tumor irradiation from different directions. As a result, the patient does not have to be moved during treatment as it is the case in IR2 and IR3. However, carbon ions cannot be provided to this room because of the Gantry specifications.

⁷a defined point of reference

⁸beamline is synonym for beam trajectory

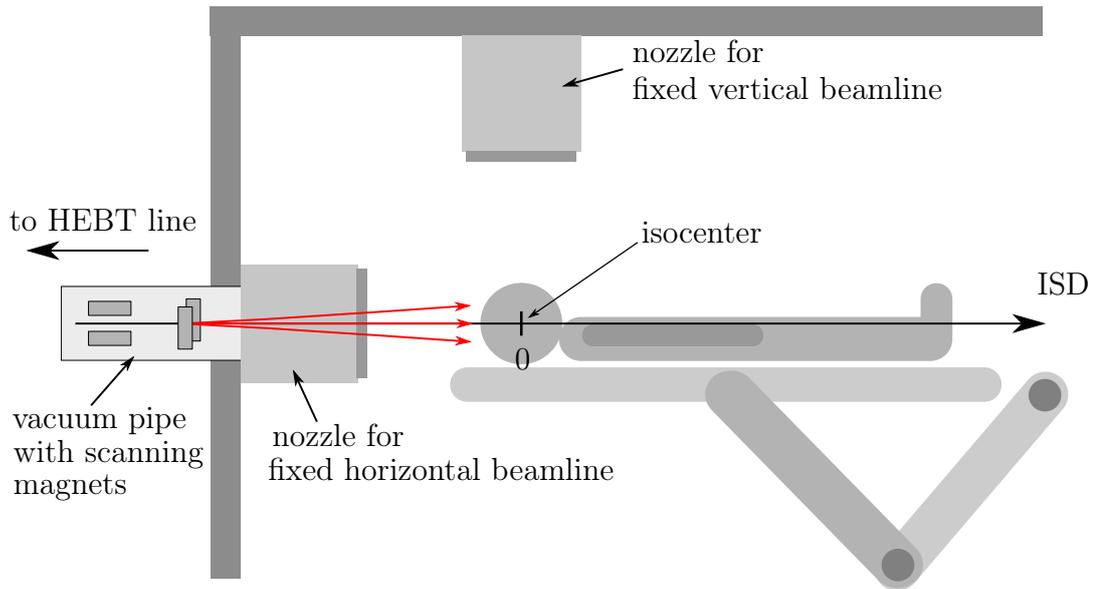


Figure 2.7.: Schematic depicting fixed horizontal and vertical beamlines as they can be found in IR2 (not to scale). The ion beam has to pass a medical nozzle before it is delivered to the target volume which is positioned at the isocenter. Analogous to the fixed horizontal beamline, the beam also passes a pair of scanning magnets when propagating along the vertical beamline, even though these are not shown in this figure.

2.3. Treatment planning and quality assurance

Prior to irradiation a treatment plan has to be created for each patient individually [30]. The responsible physician starts by delineating the target volume and possible organs at risk near to the target volume and defines the dose to be delivered to the tumor per treatment session. Based on this information a treatment planning system (TPS) determines the number of pencil beams, the energy of each of these pencil beams and the direction from which they have to be delivered to the target volume such that the resulting SOBP homogeneously covers the target volume and the surrounding healthy tissue is spared as much as possible.

Before the treatment plan can be used for patient irradiation, its resulting dose distribution has to be checked against the dose distribution obtained from an independent reference method under a selected set of treatment conditions. This allows to detect errors *before* irradiating the patient, e.g. errors that are accidentally introduced by the responsible physicians or medical physicists or errors in the propagation of information from the TPS to the delivery system [40]. As a result,

the comparison to an independent method helps to ensure accurate dose delivery to the patient.

The specific requirements for the independent reference method are usually regulated by the associated quality assurance (QA) framework. Quality assurance in general describes a set of policies and procedures to guarantee and maintain the quality of patient care [31]. It plays a crucial role especially in PBT where small changes in input parameters can have a large impact on treatment outcome due to the conformal dose distributions.

According to MedAustron's QA framework, the independent method of reference can either be

1. a QA measurement of the dose distribution resulting from the treatment plan, e.g. performed with a water phantom, or
2. a manual calculation of the dose distribution that is independent from the dose calculation of the TPS used to create the treatment plan

The second option is also referred to as *independent dose calculation* (IDC) and is getting more and more favorable these days compared to the first one [40]. This is because QA measurements of the treatment plan's dose distribution are very time-consuming and thus take away valuable beam time for patient treatment or research. As a result, they reduce the treatment efficiency of a therapy facility.

IDCs can be performed in several ways. The currently available methods can be roughly categorized into

- a) analytical dose calculation techniques
- b) methods that are based on Monte Carlo (MC) simulations
- c) methods that combine an analytical approach with MC simulations

Analytical dose calculation techniques are often used to speed up processes during treatment planning while maintaining a clinically acceptable accuracy. As an example for an analytical technique the so-called *pencil beam algorithm* (PBA) shall be mentioned. The PBA assumes that the actual beam is comprised of a superposition of small, cylindrical, narrow sub-beams, the so-called *pencil beams*. For every pencil beam a kernel is applied which describes the pencil beam's dose deposition in a given voxel. The total dose deposition can then be obtained by superposing the dose depositions of all sub-beams [18]. PBAs provide a good compromise between flexibility, accuracy and speed. However, not every effect can be modelled and therefore they exhibit limitations when applied to arbitrary geometries and tissue heterogeneities.

In contrast, MC based dose calculations are more universal, since every geometry and heterogeneity can be modelled. They also allow to model the beam delivery, i.e. the transportation of the proton beam through the nozzle and through the air before it impinges on the patient. It is generally accepted that MC based dose calculations provide the highest accuracy in comparison to other techniques [44] but are computationally also more time-consuming. Due to their high accuracy they are the method of choice for IDCs that are used to be checked against treatment plan results. MC based dose calculations will be treated later in section 2.6 in more detail.

Summing up, prior to patient irradiation an individual treatment plan has to be developed. Before the plan can be applied to the patient, its dose distribution has to be checked against the dose distribution obtained from an independent reference method. For this reference method MC based dose calculations are usually used. However, before the MC based dose calculation can be performed, a beam model is necessary that describes the optical and energy characteristics of the proton beam emerging from the HEBT line. Thus, the following section will introduce the physical foundations that are necessary to fully describe a beam and its behavior when traversing through matter.

2.4. Beam optics of charged particle beams

This section will introduce and discuss the properties of charged particle beams. First of all the concept of a particle beam is defined and associated terms and nomenclature will be explained. Afterward a particle beam's representation in phase-space and trace-space in the paraxial limit is motivated. Building on these representations, the three optical beam parameters spot size, beam divergence and beam emittance will be introduced and discussed. These parameters play a crucial role in beam modeling further on. The parameter emittance will be explained in more depth and its importance in beam transportation will be illustrated using the examples of a drift, a focusing process and a scatterer. This will eventually lead to the definition of the so-called *Twiss parameters* that are used in accelerator physics to describe beam transportation along an accelerator beamline.

If not stated otherwise, all the information in this section is based on [26].

2.4.1. Concept and phase-space representation

A beam of charged particles consists of a group of particles that have nearly the same kinetic energy and move nearly into the same direction. The kinetic energy

of every particle has to be much higher than its thermal energy at ordinary temperatures so that the beam can be regarded as an ordered flow of particles. Figure 2.8 shows a schematic of a charged particle beam.

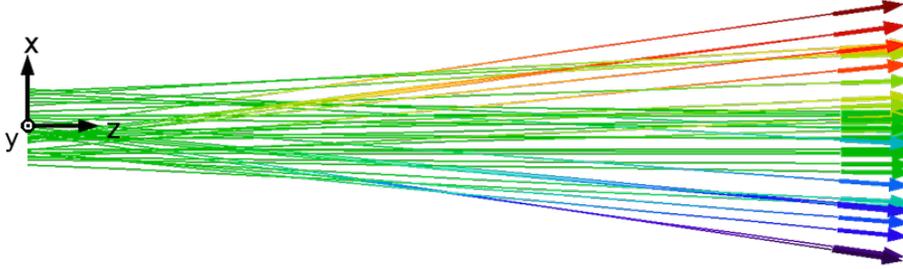


Figure 2.8.: Schematic of a diverging, charged particle beam. Lines indicate the particle trajectories and arrows the velocity of each particle. The color gradient maps to the temporal change of the x-coordinate, i.e. the transverse velocity v_x . The transverse scale is exaggerated. Image source: [7]

For the complete description of such a beam consisting of n particles one can use the total set of positions $\vec{r}_i(t)$ and velocities $\vec{v}_i(t)$ for each particle $i = 1, \dots, n$ at each point in time t

$$\vec{r}_i = (x_i, y_i, z_i)^T \quad \vec{v} = (v_{x,i}, v_{y,i}, v_{z,i})^T \quad (2.2)$$

The position and velocity along the beam's direction of motion, i.e. the *beam axis* (in figure 2.8 the z-axis), are called *axial* position and velocity. The positions and velocities perpendicular to the axial direction are referred to as *transverse* positions and velocities. In clinical practice, proton beams are paraxial, which means the axial velocity v_z is much greater than the transverse velocities $v_{x,i}$ and $v_{y,i}$. As a result, instead of using the transverse velocities $v_{x,i}$ and $v_{y,i}$ one can also use the corresponding *orbit inclination angles* x'_i and y'_i between transverse velocity direction and beam axis defined by

$$x'_i \approx \sin(x'_i) = \frac{v_{x,i}}{v_{z,i}} \quad y'_i \approx \sin(y'_i) = \frac{v_{y,i}}{v_{z,i}} \quad (2.3)$$

The approximation is only valid for the paraxial limit. Since the inclination of particle orbits can be measured directly, in accelerator physics usually angles instead of transverse velocities are used to describe the beam.

Now position $\vec{r}_i(t)$ and velocity coordinates $\vec{v}_i(t)$ can be visualized in 6-dimensional phase-space (\vec{r}, \vec{v}) for each particle i at each point in time t . Adding more particles to the beam in figure 2.8 and mapping their respective positions and velocities at the beam's very end to transverse phase-space (x, v_x) would result in the phase-space diagram shown in figure 2.9. The sum of all particles coordinates

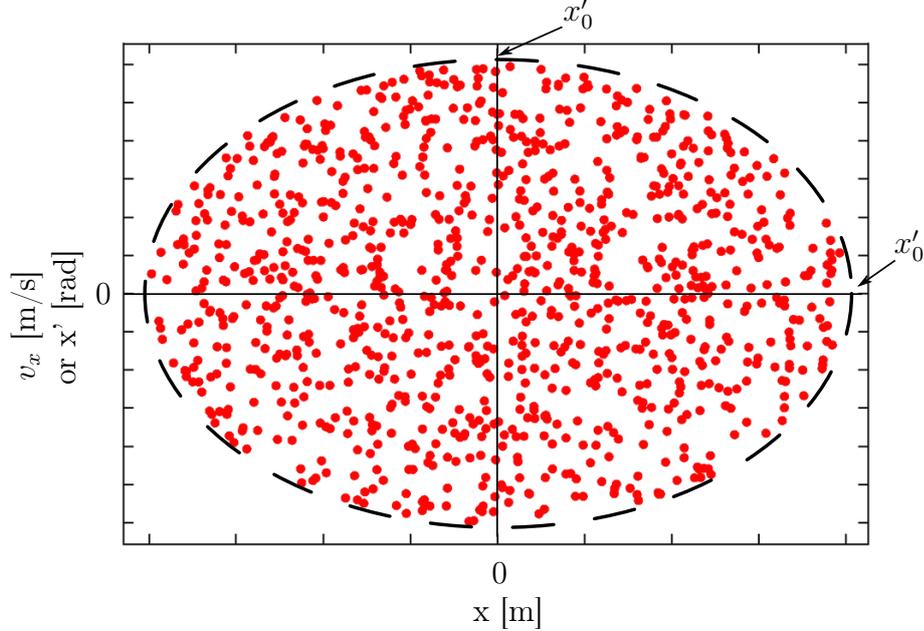


Figure 2.9.: Phase-space representation (when using v_x on y-axis) of a charged particle beam. The corresponding trace-space representation would look qualitatively the same, except that instead of v_x the inclination angle x' would be used on the y-axis. All points lie within an ellipse with semi-major axis x_0 and semi-minor axis x'_0 (in trace-space), here marked with a dashed line. Image source: [7]

in phase-space take an elliptical shape and can be circumscribed by a so-called *phase-space ellipse*. This means that particles closer to the beam axis exhibit a greater spread in their transverse velocities than particles in the periphery. The ellipse proportions and the ellipse's tilt, however, vary along the beamline due to beam propagation and focusing of the beam and thus are a function of time (see later).

Analogous to mapping $(x_i, v_{x,i})$ to the transverse phase-space, one can also use angles x'_i instead of velocity coordinates and map to the space (x_i, x'_i) . This space is referred to as *trace-space* and contains the same information as the phase-space since they only differ by a linear factor $v_{z,i}$ as shown by equation 2.3. The trace space is used later on for the definition of beam emittance. The same can be done

along the remaining transverse direction (y), leading to the trace-space (y_i, y'_i).

Since n is usually a large number⁹, it would be impossible to take into account the position and velocity coordinates of each particle. Therefore, one naturally chooses statistical root-mean-square (RMS) values to describe the beam.

2.4.2. Optical beam parameters

Spot size

The RMS value for the transverse position x_{RMS} is defined as

$$x_{RMS} = \sqrt{\langle x_i^2 \rangle} \quad (2.4)$$

with $\langle \dots \rangle$ denoting the mean value calculated over all particles $i = 1, \dots, n$. Since it is always possible to set the z -axis in such a way that it coincides with the beam axis, the mean over all transverse positions $\langle x_i \rangle$ in a charged particle beam can be assumed to be zero

$$\langle x_i \rangle = 0 \quad (2.5)$$

As a result, the RMS value for transverse position x actually equals the standard deviation of the particle distribution in x -direction.

$$x_{RMS} = \sqrt{\langle x_i^2 \rangle} = \sqrt{\langle x_i^2 \rangle - \langle x_i \rangle^2} = \sigma_x \quad (2.6)$$

Since σ_x is a measure for the lateral dimension of the beam in x -direction at a given axial position z , σ_x is also referred to as **x-spot size** at this axial position z . The same arguments can also be applied to the y -direction yielding that $y_{RMS} = \sigma_y$ which is referred to as **y-spot size**.

Beam divergence

In a similar way, the RMS value for the inclination angle x'_{RMS} is defined as

$$x'_{RMS} = \sqrt{\langle x_i'^2 \rangle} \quad (2.7)$$

Using the same argument as before, we can always set the z -axis such that the mean over all inclination angles $\langle x'_i \rangle$ equals zero

$$\langle x'_i \rangle = 0 \quad (2.8)$$

⁹ $n = 10^{10}$ for a pencil beam at emerging from MAPTA

which results in

$$x'_{RMS} = \sqrt{\langle x_i'^2 \rangle} = \sqrt{\langle x_i'^2 \rangle - \langle x_i' \rangle^2} = \sigma'_x \quad (2.9)$$

where σ'_x denotes the standard deviation of the distribution of the particles' inclination angles. Since σ'_x is an angle given in rad and acts as a measure of how fast the beam's x-spot size increases while propagating along the z axis it seems reasonable in the first instance to refer to σ'_x as beam divergence. However, given equation 2.9 it is not possible for σ'_x to take on negative values and it is therefore not possible to distinguish between divergence and convergence at a particular axial position z. Nevertheless, σ'_x in above equation can be extended with the sign of the particles' position-angle-covariance¹⁰, i.e. the sign of the correlation between position x and inclination angle x'.

$$\theta_x = \text{sgn}(\langle x_i x_i' \rangle) \sigma'_x \quad (2.10)$$

If particles with positive positions x_i tend to have a positive angle x_i' and particles with negative positions x_i a negative angle x_i' the beam is divergent and the sgn-function in above equation yields +1. In the other case where the correlation is negative, the beam is convergent and the sgn-function yields -1. Therefore, θ_x is also referred to as **x-divergence** of the beam at a particular axial position z. A negative divergence means a converging beam at this position z. The same arguments hold true for the y-direction with θ_y which is referred to as **y-divergence**.

Beam emittance

Additionally to spot size and divergence there is another important beam parameter: the beam emittance. It is an empirical quantity which has useful applications in accelerator physics. The beam emittance is generally defined as the area of the ellipse formed by the particles in trace-space. The transverse emittance ϵ_x in x-direction reads

$$\epsilon_x = \iint_{ell} dx dx' \quad (2.11)$$

in which the region of integration *ell* is the ellipse that encloses the points of all particles in trace-space as shown in figure 2.9. The emittance for the remaining transverse direction ϵ_y is defined analogous to the x-direction. In the special case of figure 2.9 in which the ellipse's major and minor axes coincide with the x- and x'-axes (i.e. the ellipse is not tilted) one can express the transverse emittance ϵ_x

¹⁰the covariance of the particles' x-position and inclination angle x' can be written as $\text{cov}(x_i, x_i') = \langle x_i x_i' \rangle - \langle x_i \rangle \langle x_i' \rangle$ in which the latter product cancels to zero because of equation 2.5 and 2.8

in terms of the maximum transverse position x_0 and inclination angle x_0'

$$\epsilon_x = \pi \cdot x_0 \cdot x_0' \quad (2.12)$$

It shall be emphasized that equation 2.11 is not the only convention to define emittance. In addition to equation 2.11 two other conventions in literature exist for the definition of emittance. They differ in whether the area of the trace-space ellipse is left unchanged or divided by π or 4π which is motivated by equation 2.12. Hence, the other two conventions read

$$\epsilon_x = \frac{\iint_{ell} dx dx'}{\pi} \quad (2.13)$$

$$\epsilon_x = \frac{\iint_{ell} dx dx'}{4\pi} \quad (2.14)$$

In order to not mix up these three conventions one usually flags the units according to the convention used. The emittance convention as defined in equation 2.11 has SI units of $m \cdot rad$ and is in the following referred to as 1-convention. Accordingly, the units of emittance defined in equation 2.13 has SI units of $\pi \cdot m \cdot rad$ and will be referred to as π -convention. Finally, the one defined in equation 2.14 has SI units of $4\pi \cdot m \cdot rad$ and will be referred to as 4π -convention. In the course of this thesis emittance will always be given in SI units of $m \cdot rad$. Therefore, all following formulas concerning emittance will be given in accordance with the 1-convention, even though [26] uses the π -convention. In order to transform equations containing emittance from 1-convention to π -convention one just has to consider a factor of π .

Since emittance is defined as the area of the ellipse in trace-space, it decreases when a beam is accelerated. This is due to adiabatic damping and can be understood by considering equation 2.3. During acceleration the longitudinal velocity $v_{z,i}$ of each particle increases while the transverse velocities $v_{x,i}$ and $v_{y,i}$ remain unchanged which results in shrinking inclination angles x'_i and y'_i for each particle. This in turn decreases the area of the circumscribed trace-space ellipse and thus the emittance. It is therefore useful to define a quantity that remains constant during acceleration: the so-called *normalized emittance* $\epsilon_{n,x}$ or $\epsilon_{n,y}$. It is defined according to

$$\epsilon_{n,x} = \beta\gamma \cdot \epsilon_x \quad (2.15)$$

in which $\beta = \frac{|\vec{v}|}{c}$ and $\gamma = (1 - \beta^2)^{-1/2}$. The normalized emittance for the remaining transverse direction $\epsilon_{n,y}$ is defined analogous to the x-direction. In order to distinguish between normalized emittance and emittance, the latter one is also often referred to as *geometric emittance*. In the following only processes without

acceleration will be examined and it is thus sufficient to only consider the geometric emittance.

2.4.3. Reversible processes and Twiss parameters

The reason why emittance is a helpful concept in accelerator physics is because it is a conserved quantity if the beam is subject to *reversible processes*. In the following the conservation of emittance will be illustrated by the means of two reversible processes: drift and linear focusing.

Drifts refer to the free propagation of a particle beam along its beam axis as a consequence of their velocity. Figure 2.10 shows the effect of a drift on the trace-space ellipse. The higher the inclination coordinate x' of a trace-space vector, the higher the change in its transverse position coordinate x along the drift. Points on the x -axis do not move since they exhibit zero inclination angle and thus do not change their transverse position. The maximal inclination angle x'_{max} does not change during a drift. The net result is a shear of the phase space ellipse that conserves its area and thus also the emittance. This is a special case of Liouville's Theorem [21].

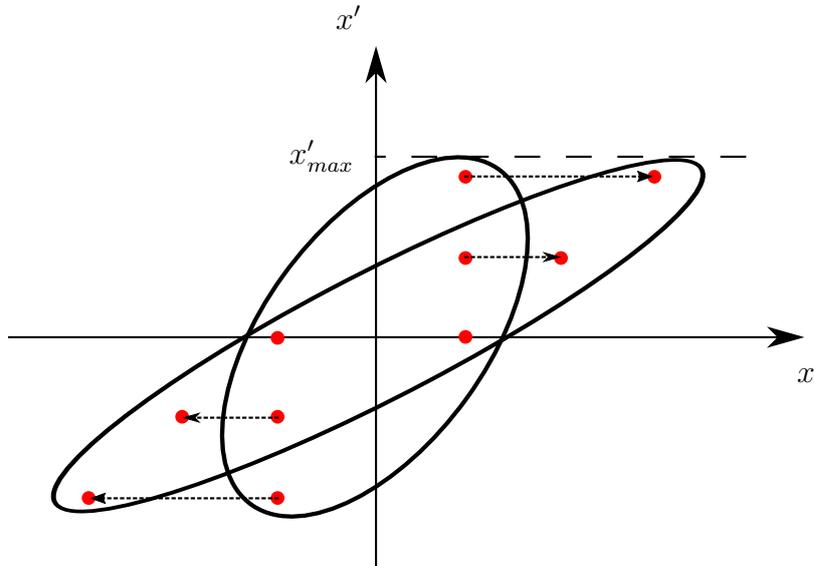


Figure 2.10.: Schematic depicting a drift of a diverging particle beam in trace-space. Trace-space vectors before and after the drift as well as their movement caused by the drift are visualized with red points and dashed arrows. The movement is directed in $+x$ or $-x$ direction, depending on the sign of the respective x' -coordinate of each particle. Image conceptually taken from [21]

The other reversible process stated above is linear beam focusing. It refers to the process of focusing a particle beam with a linear lens system. This is a system that applies a transverse force to each particle which is linearly proportional to the transverse distance of the particle to the beam axis. It can be shown that the effect of such a linear focusing system can be described by multiplying the trace-space vector $(x_0, x'_0)^T$ of each beam particle before entering the system with the so-called transfer matrix M

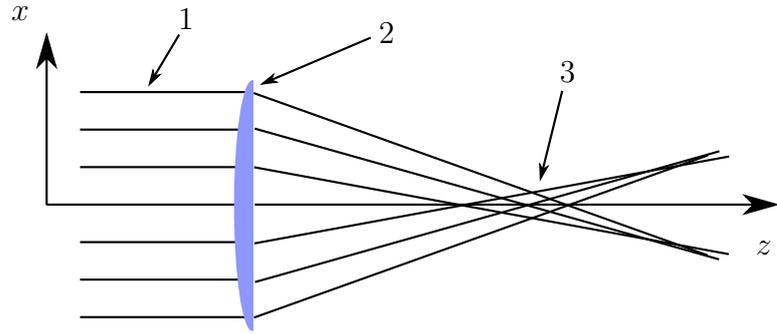
$$\begin{pmatrix} x_1 \\ x'_1 \end{pmatrix} = M \begin{pmatrix} x_0 \\ x'_0 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} x_0 \\ x'_0 \end{pmatrix} \quad (2.16)$$

Here $(x_1, x'_1)^T$ denotes the trace-space vector of a particle having passed the system. The m_{ij} depend on the distribution of forces. Figure 2.11 shows the effect of a linear focusing lens on an initially parallel particle beam in the x-z plane and in trace-space. The change of each particle's inclination angle x' due to the lens is directly proportional to the respective transverse position x of each particle. Points on the x' axis do not move since they lie directly on the beam axis. The maximal transverse position x_{max} is left unchanged directly after passing the focusing lens and first starts to shrink when drifting occurs. The net result is similar to the drift situation: the trace-space ellipse gets sheared, however, this time in x' -direction. During this process the area enclosed by the ellipse and thus the emittance are conserved.

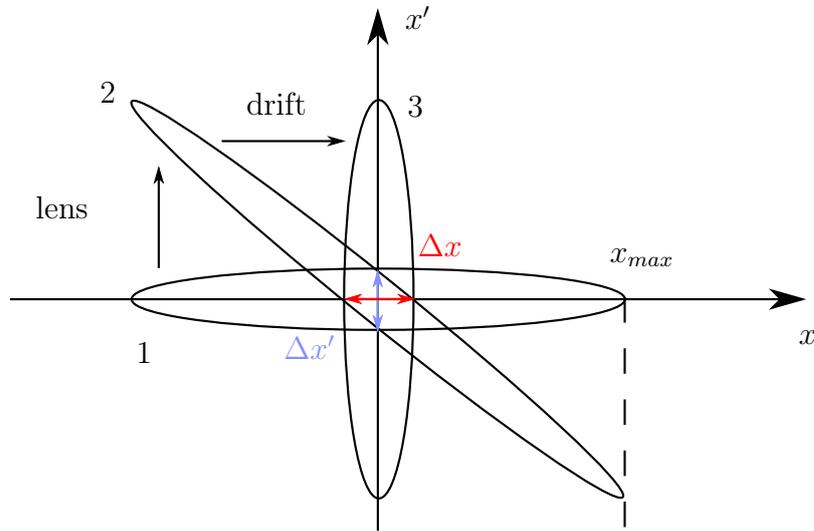
Analogous to focusing processes also drifts can be described by a transfer matrix M_d . This formalism proves extremely helpful for the mathematical description of a particle beam in a synchrotron where the beam is focused using focusing cells that usually consist of alternating focusing and defocusing quadrupole magnets with drift spaces in between. Each element in such a focusing cell can be described by a transfer matrix M or M_d . An example for a specific type of focusing cell is the so-called FODO cell. A FODO cell is a magnetic structure that consists of a subsequent arrangement of a focusing lens, followed by a drift space, a defocusing lens, another drift space and a second focusing lens. The beam passage through a FODO cell can thus be described by subsequent matrix multiplication

$$M_{\text{FODO}} \begin{pmatrix} x_0 \\ x'_0 \end{pmatrix} = M_{\text{foc}} M_d M_{\text{defoc}} M_d M_{\text{foc}} \begin{pmatrix} x_0 \\ x'_0 \end{pmatrix} \quad (2.17)$$

The transfer matrix formalism sketched here can be simplified as will be motivated in the following. As mentioned earlier in subsection 2.4.1, typically an *ellipse* is used to circumscribe the sum of all points in trace-space. In accelerator physics



(a)



(b)

Figure 2.11.: Schematic depicting the focusing process of an initially (almost) parallel particle beam in the x - z -plane (a) and in trace-space (b). Since the beam is not completely parallel and exhibits an initial spread in inclination angle $\Delta x'$ in stage 1 it cannot be focused to a single point and thus also exhibits a spread in transverse space Δx in stage 3 (minimum focal spot size). The spread in inclination angle is not visualized in subfigure (a).

this is a matter of pure convenience and is due to the fact that a linear transformation always transforms an elliptical distribution into another elliptical distribution. Hence, instead of describing the effect of reversible processes on all the points in trace-space, one can simply describe the change of the trace-space ellipse's geomet-

rical parameters. These parameters are the so-called Twiss parameters¹¹ α_x , β_x and γ_x . Together with the beam emittance ϵ_x they determine size and orientation of the ellipse. This means the four quantities $(\alpha_x, \beta_x, \gamma_x, \epsilon_x)$ describe a beam uniquely using the formalism of a circumscribed trace-space ellipse.

All quoted relations in the remaining part of this subsection also apply for the transverse direction y . The Twiss parameters are defined using the equation for a tilted ellipse in trace-space

$$\gamma_x x^2 + 2\alpha_x x x' + \beta_x x'^2 = \frac{\epsilon_x}{\pi} \quad (2.18)$$

in which the 1-convention of beam emittance ϵ_x is used. As carried out in [11] one can derive a relationship between the three Twiss parameters and the RMS values for position x and inclination angle x' . These read as follows

$$\sigma_x = \sqrt{\beta_x \frac{\epsilon_x}{\pi}} \quad (2.19)$$

$$\sigma_{x'} = \sqrt{\gamma_x \frac{\epsilon_x}{\pi}} \quad (2.20)$$

$$\langle x x' \rangle = -\alpha_x \frac{\epsilon_x}{\pi} \quad (2.21)$$

in which σ_x denotes the x -spot size and $\sigma_{x'}$ the absolute value of the x -divergence. As is apparent from the equations above, β_x is a measure of a beam's spatial spread in x -direction and γ_x a measure of its angular spread in x' -direction. The remaining parameter α_x gives information about the tilt of the ellipse and whether the beam is diverging or converging. Comparing equation 2.21 with 2.10 allows us to extract the sign of the x -divergence which is just the negative sign of α_x . This yields

$$\theta_x = -\text{sgn}(\alpha_x) \sqrt{\gamma_x \frac{\epsilon_x}{\pi}} \quad (2.22)$$

The considerations above prove that as an alternative to the set of transport parameters $(\alpha_x, \beta_x, \gamma_x, \epsilon_x)$ a beam can also be described by the three optical parameters $(\sigma_x, \theta_x, \epsilon_x)$. As shown in [26], the transfer matrix formalism can be extended to describe the evolution of the Twiss parameters along the beamline provided that only reversible processes are involved and the beam in trace-space can be circumscribed by an ellipse. However, the latter is not always the case and the transfer matrix formalism cannot be applied in its original form in these situations.

¹¹also often referred to as transport parameters or Courant-Snyder functions

Nevertheless, there are techniques to circumvent such problems as carried out in [4].

2.4.4. Irreversible processes

Up to now only reversible processes were considered which conserve emittance. As a result, a beam's emittance does not change when revolving in a synchrotron with an ideal linear focusing system. Even if the beam was accelerated in such a synchrotron, its emittance in terms of normalized emittance would be constant.

The normalized emittance, however, is subject to growth when passing an imperfect focusing device. By determining normalized emittance along the beamline, potential imperfections in the focusing system can be detected and corrected. In accelerator physics, usually a small normalized emittance is desirable because it implies good parallelism of the beam.

Another irreversible process that leads to emittance growth can be found if one leaves the regime of vacuum and examines a beam's behavior when traversing through matter: scattering. Charged particles can interact in different ways with the constituents of matter. At this point the effect of a thin scatterer on the trace-space ellipse and the resulting emittance growth shall be illustrated based on [21]. A thin scatterer refers to the process of charged particles undergoing multiple Coulomb scattering (see subsection 2.5.3 for details). If a beam particle undergoes a scatterer it gets deflected from its original direction and also loses some of its kinetic energy E during this process. This means that every point in trace-space receives random vertical kicks $\Delta x'_i$ in either x' or $-x'$ -direction. The magnitude of these kicks is Gaussian distributed around the mean $\mu_{\Delta x'} = 0$ rad with a standard deviation of θ_0 . The latter quantity θ_0 is also called *scatterer strength* and depends on the scatterer material, its thickness and the scattering particle's current kinetic energy E . The position coordinate x remains unchanged for every point because there is no drift involved in a thin scatterer. As a result, the area enclosed by the ellipse increases and so does the emittance. Figure 2.12 visualizes the emittance growth due to a single thin scatterer.

The emittance growth of a series of N thin scatterers that are separated by drift spaces can be calculated analytically within the framework of Fermi-Eyges theory and is given by (see [21])

$$\epsilon_x = \sqrt{\epsilon_{x,0}^2 + \sum_{i=1}^N (\pi \sigma_{x,i} \cdot \theta_{0,i})^2} \quad (2.23)$$

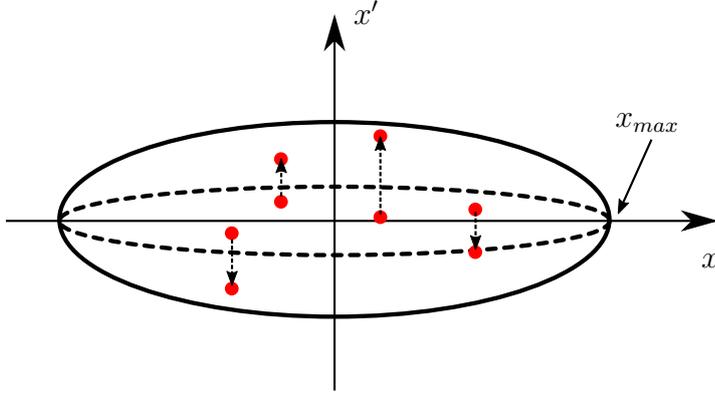


Figure 2.12.: Schematic depicting the process of a scatterer. The trace-space ellipse of the beam right before the scatterer is indicated with a dashed line, the one of the beam right after the scatterer with a solid line. Image conceptually taken from [21].

in which $\epsilon_{x,0}$ denotes the emittance (using 1-convention) of the incident beam, $\sigma_{x,i}$ the x-spot size of the beam when undergoing the i-th scatterer and $\theta_{0,i}$ the scatterer strength of the i-th scatterer.

Fermi-Eyges theory also allows to analytically describe the full transportation of the beam ellipse (i.e. determine its Twiss parameters as a function of penetration depth z) through a stack of homogenous slabs with infinite transverse dimensions. Unfortunately, the theory can only be applied to some very specific problems. It is not applicable to situations where non-elliptical¹² trace-space distributions, material heterogeneities or interaction effects other than multiple Coulomb scattering are involved. Regardless from fulfilling those requirements, the Fermi-Eyges theory breaks down near the stopping depth of the particles because the relationship between depth and energy of the particles is blurred (see section 2.5.6) [22]. This is, however, the area of most interest in proton beam therapy because it marks the region where the target volume is located. Consequently, MC simulations are normally used in practice in order to describe the transportation of the beam ellipse through matter.

In summary, it was shown that the optical properties of a charged particle beam can be fully and uniquely described with the help of 6 parameters ($\sigma_x, \theta_x, \epsilon_x, \sigma_y, \theta_y, \epsilon_y$), which are the three optical parameters spot size, beam divergence and beam emittance for each transverse direction x and y . The beam emittance is a conserved quantity for reversible processes like drifts in vacuum, but increases

¹²i.e. non-Gaussian trace-space distributions

during irreversible processes such as scatterers that occur when a beam traverses matter.

2.4.5. Beam characteristics at MedAustron

In this subsection the theoretical framework about beam optics introduced in the past section 2.4 will be applied to describe the treatment beam used at MedAustron. This will let the reader gain a better understanding of how the trace-space formalism is used in practice. First of all, the trace-space representation of the treatment beam at MedAustron will be discussed. Building on that, the transverse intensity profiles of the beam will be derived. Finally, the modification of these profiles due to scattering with the double vacuum windows will be presented.

When the treatment beam revolves in the synchrotron it initially takes the form of an ellipse in both horizontal and vertical trace-space. However, due to the *slow resonant extraction* mechanism¹³ the initially elliptical trace-space distribution in horizontal trace-space is converted to a rectangular trace-space distribution. In vertical trace-space the beam still has an elliptical distribution, thus only the extraction direction is affected [4]. Figure 2.13 visualizes all this. The rectangular

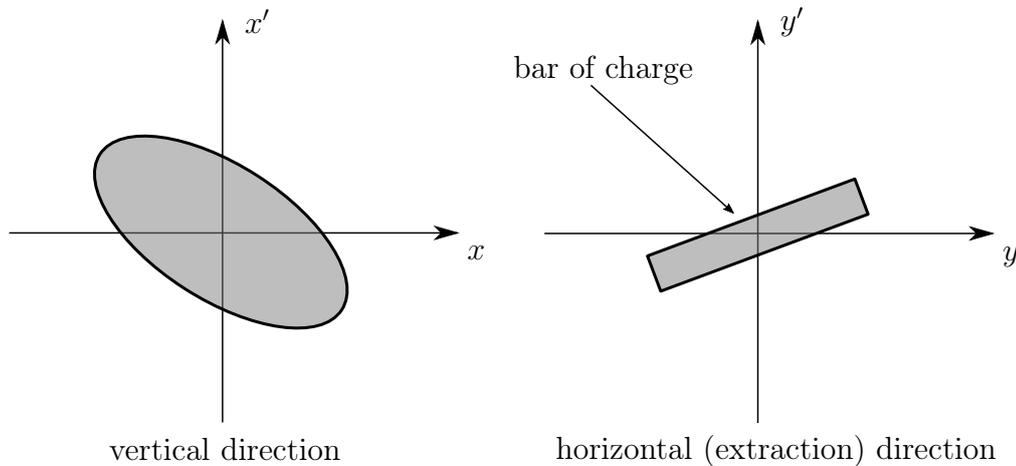


Figure 2.13.: Visualization of the transverse trace-spaces of a slow resonant extracted beam. Image conceptually taken from [4]

trace-space distribution is also referred to as "bar of charge". Its geometric emittance does not depend on kinetic energy [4], which is in stark contrast to the

¹³see section section 2.2

geometric emittance of the beam ellipse in horizontal trace-space. As shown in 2.4.2 the geometric emittance of a beam ellipse decreases with increasing kinetic energy due to adiabatic damping.

Now as a second step, the beam intensity profiles for vertical and horizontal direction will be obtained by orthogonal projection of all points within the ellipse or the bar of charge onto the x- or y-axis in the plots of figure 2.13, respectively. As a result, the beam intensity profile in vertical direction is Gaussian whereas for the horizontal direction it is trapezoidal. Measurements of the intensity profiles for a point in the HEBT, i.e. after extraction, support theory and are sketched in figure 2.14. Eventually, the modification of the obtained beam profiles due to scattering

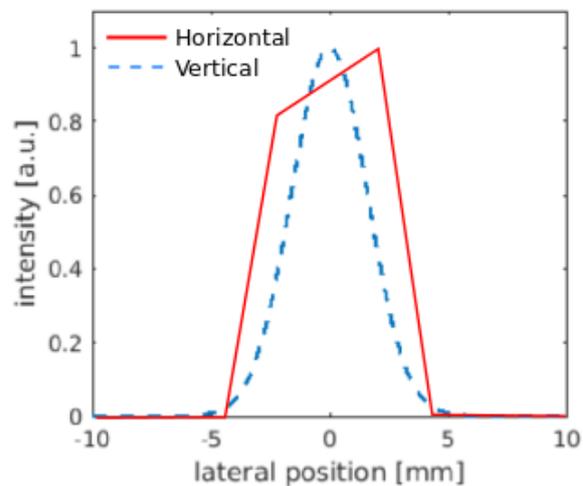


Figure 2.14.: Sketch of beam intensity profile measurements for horizontal and vertical direction after extraction. The plateau of the trapeze is not a horizontal line but slightly increases with increasing y-position. Image source: [9]

is considered. Measurements directly after the double vacuum window are shown in figure 2.15. The measurements show that scattering due to the vacuum windows shape the trapezoidal profile form to a more Gaussian form. This is also predicted by theory [21]. The horizontal profile is, however, positively skewed which might be due to the tilt of the trapeze’s plateaus in figure 2.14. In contrast, the vertical profile is Gaussian without skew. This asymmetry between horizontal and vertical profile is important to keep in mind for the beam modeling procedure.

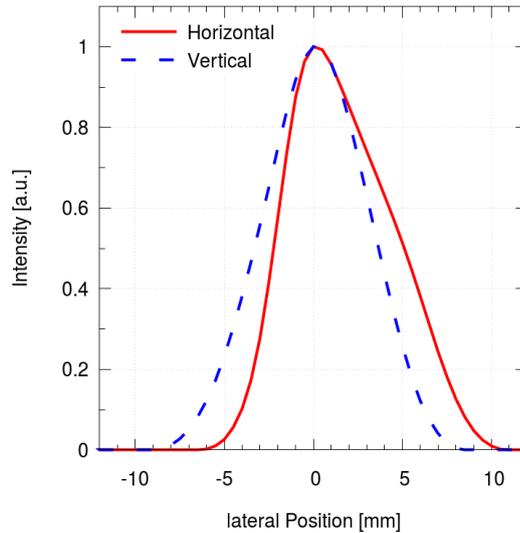


Figure 2.15.: Measurements of laterally integrated beam intensity profiles in both vertical and horizontal direction right after the double vacuum windows for 62 MeV protons. The nozzle was removed for these measurements. Image source: [17]

2.5. Interactions of heavy charged particles with matter

After having introduced the fundamentals of beam optics, this section will now treat the basic physics concepts that are necessary to describe the dose deposition of a charged particle beam in matter. First of all the relevant interaction processes of heavy charged particles with matter will be discussed. Building on these interactions, the stopping power S as a form of energy deposition rate will be introduced and its behavior for different particle velocities will be presented using the Bethe-Bloch equation. Afterward, multiple Coulomb scattering and its approximations will be discussed. In order to consider the radiobiological effects of energy deposited, the quantity *dose* will be defined and nomenclature for IDD distributions will be introduced. Eventually, range and energy straggling and its implications on the range of a particle beam and its Bragg peak width will be discussed.

If not stated differently all the information in this section is based on [36].

2.5.1. Interaction processes

When a charged particle passes through matter it can undergo different interactions with the matter's constituents, i.e. atoms and molecules. The passage of such an incident particle, in the following also called *primary*, through matter is generally characterized by two effects: a continuous energy loss of the primary and a deflection of the primary from its incident direction. In the following the interactions of only heavy charged particles with matter will be considered. This group contains protons, muons, pions, α -particles and other light ions. They show different behavior when traversing matter compared to the group of light charged particles, consisting of electrons and positrons. In the MeV range, i.e. the clinical energy range for proton beam therapy, the two most prominent¹⁴ interaction processes are

1. inelastic scattering with atomic electrons
2. elastic scattering with nuclei

Both are electromagnetic processes that occur with a certain quantum mechanical probability.

In the first one a primary transfers a certain amount of its energy to an atomic electron which is then either lifted to a higher energetic (but still bound) state or completely ionized if the transferred energy is sufficiently high. In some of the ionization processes the electron still has enough energy to cause secondary ionization. These electrons are also known as δ -rays.

In the second one a primary transfers energy to an atomic nucleus of the material without exciting it. These processes occur frequently but still not that often as the inelastic scattering processes with electrons. Since the nuclei of the material have in general a higher mass than the primaries, very little energy is transferred in the nucleus scattering processes. As a result, the electron scattering is responsible for the main energy loss of an incident primary. The nucleus scattering processes, however, repeatedly introduce a lot of small-angle deflections into the primaries' trajectory, leading to a net deflection from the incident direction. This is called multiple Coulomb scattering. On the contrary, a primary gets hardly deflected when scattering with an electron because the mass of the primary is much larger than the electron mass.

¹⁴There are of course other interaction channels, e.g. elastic scattering with atomic electrons, nuclear reactions or the deflection of a primary in the Coulomb field with subsequent Bremsstrahlung emission. However, these effects can be neglected for the purpose of beam modeling.

2.5.2. Stopping power and Bethe-Bloch equation

The amount of energy transferred per inelastic scattering process with an electron is usually a very small fraction of the primary's kinetic energy E . However, the number of these processes per unit path length is very high such that a substantial energy loss is observed even when only thin layers of matter are traversed. Thus, the fluctuations in energy loss with penetration depth are small and it is meaningful to define a quantity that gives the average kinetic energy loss per unit path length: the stopping power S . It is defined as

$$S(x) = -\frac{dE}{dx} \quad (2.24)$$

Stopping power is usually given in MeV/cm or keV/ μm . Often it is divided by the density ρ of the absorber material. This quantity is called mass stopping power S/ρ and has units of $\frac{\text{MeV}}{\text{g/cm}^2}$, i.e. energy per mass thickness. The advantage of mass stopping power is that it varies little over a wide range of materials.

A quantity that is closely related to stopping power is the linear energy transfer (LET). The LET is defined as the energy locally deposited in the medium per unit path length. It differs from the stopping power in that it generally does not include the emission of Bremsstrahlung because X-rays usually escape the region of the particle path. Stopping power on the contrary takes the emission of Bremsstrahlung into account.

The stopping power of heavy charged particles resulting from inelastic scattering with atomic electrons when traversing through a material can be calculated using the Bethe-Bloch formula

$$-\frac{dE}{dx} = 2\pi N_a r_e^2 m_e c^2 \cdot \rho \frac{Z}{A} \frac{z^2}{\beta^2} \left[\ln \left(\frac{2m_e \gamma^2 v^2 W_{max}}{I^2} \right) - 2\beta^2 - \delta - 2\frac{C}{Z} \right] \quad (2.25)$$

in which N_a is Avogadro's number, r_e the classical electron radius, m_e the electron rest mass, c the speed of light, ρ the density of the absorbing material, Z the atomic number and A the atomic weight of the material's atoms, z the charge of the projectile in terms of the elementary charge e , $\beta = v/c$ with v being the current velocity of the projectile and $\gamma = (1 - \beta^2)^{-1/2}$. I is the mean excitation potential of the material. In radiotherapy water is most often used for calculations and as a tissue substitute because human tissue consists mostly of water. Its mean excitation potential I is not exactly known but is somewhere between 69 and 78 eV [2]. W_{max} is the maximum energy transfer in a single collision which is reached in a head-on collision. For protons as projectiles and electrons as targets one can assume $m_p \gg m_e$ and approximate W_{max} with $W_{max} \approx 2m_e c^2 (\beta\gamma)^2$.

Finally, δ denotes the density correction and C the shell correction which are only important at very high or very low energies, respectively.

Figure 2.16 shows the mass stopping power as a function of kinetic energy for protons in water. The Bethe-Bloch equation is valid only for a specific kinetic energy range which is also indicated in the figure. When analyzing the Bethe-Bloch equation 2.25, one can observe that for high velocities the logarithmic term is dominant and stopping power decreases with decreasing velocity. For low projectile velocities the term $1/\beta^2$ in the prefactor is dominant and stopping power increases steeply with decreasing velocity. If the projectile velocity drops below a certain value similar to the velocity of the atomic electrons of the material, the stopping power reaches a sharp maximum and drops sharply again. This is where the Bethe-Bloch equation breaks down due to several effects. The most important one is that projectiles pick up atomic electrons which reduce their effective charge and thus their stopping power. The Bethe-Bloch equation provides the theoretical explanation of the Bragg peak: the energy loss of a heavy charged particle along its trajectory in a material reaches a maximum right before its velocity drops to zero, i.e. right before the particle stops.

2.5.3. Multiple Coulomb scattering

Other than the energy loss of the projectile by inelastic scattering with electrons, also the elastic scattering processes with nuclei play an important role. These processes repeatedly introduce many small-angle deflections which make the projectile follow a random zigzag path when traversing the material. This results in a projected¹⁵ net deflection angle θ_{net} from the incident direction, as figure 2.17 shows. The phenomenon is called multiple Coulomb scattering (MCS) and can currently be best described by the theory of Molière [6] which fits the experimental data up to a few percent [23]. The theory describes the distribution of the projected net deflection angle θ_{net} as a series expansion. Figure 2.18 shows Molière's distribution of θ_{net} for 158.6 MeV protons traversing 1 cm of water. For smaller net deflection angles the distribution can be well approximated by a Gaussian which results in Highland's theory, a simpler and parameterized version of Molière's theory¹⁶. In contrast to Highland's Gaussian distribution the distribution of Molière exhibits long broad tails for larger angles. The deflections at larger angles are generally a result of a single large-angle scattering event rather than the sum of multiple small-angle scatterings. For most problems in proton beam therapy it is sufficient to use the Gaussian approximation for multiple Coulomb scattering. This allows

¹⁵projected onto the beam axis

¹⁶there are also other theories of multiple Coulomb scattering in Gaussian approximation, see [37]

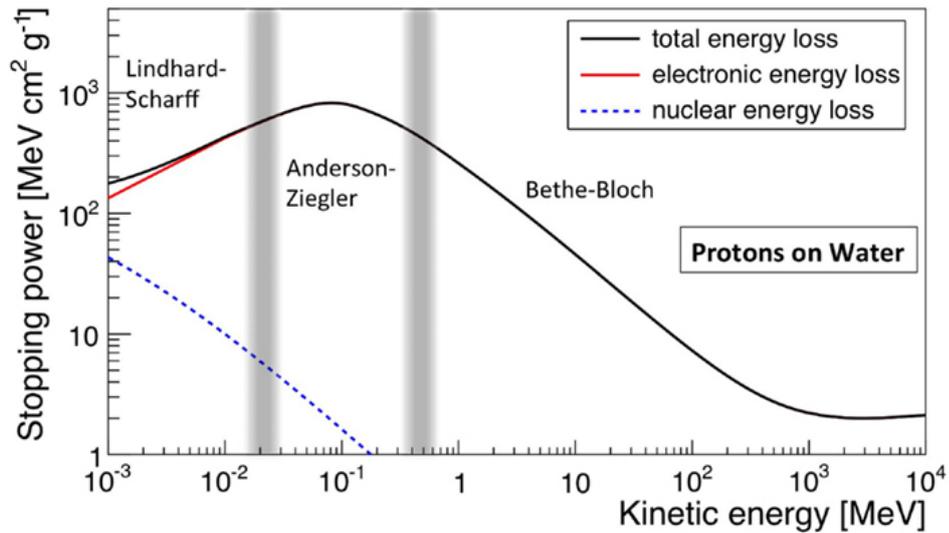


Figure 2.16.: Mass stopping power as a function of kinetic energy for protons in water. As can be seen in the figure, the energy loss due to nuclear scattering events is negligible for clinical energies. After the mass stopping power reaches a minimum in the GeV regime, it starts to increase again with increasing kinetic energy. The range of validity of the Bethe-Bloch formula is bounded below by the Anderson-Ziegler domain and above by an energy range which is dominated by radiative losses. The latter range, however, is not depicted in this figure. Image source: [33]

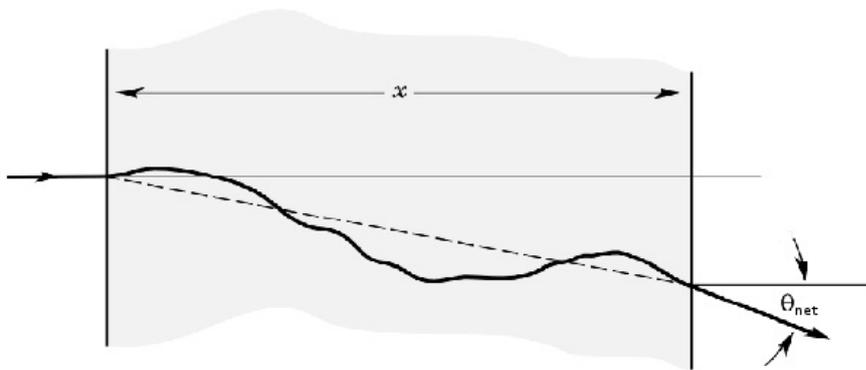


Figure 2.17.: A projectile traversing a material of thickness x experiencing repeated small-angle deflections caused by multiple Coulomb scattering. Adapted from [41].

the following argument: if the distribution of the projected net deflection angle θ_{net} is Gaussian and the spatial distribution of the transverse x direction of the incoming beam is Gaussian as well, then in the paraxial limit this spatial distribu-

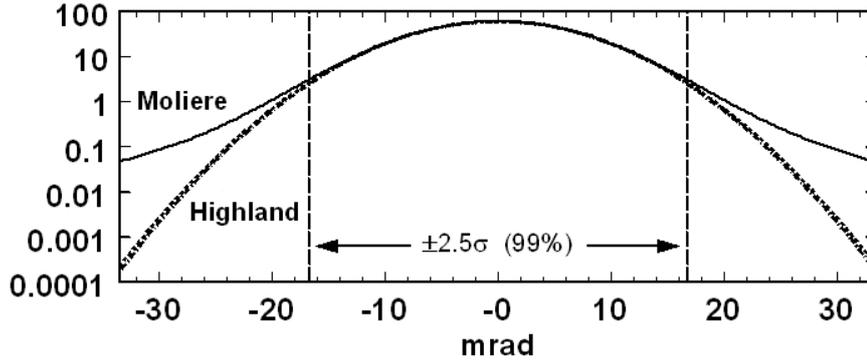


Figure 2.18.: Distribution of the projected net deflection angle θ_{net} for 158.6 MeV protons traversing 1 cm of water. The Gaussian approximation by Highland is correct for up to $\pm 2.5\sigma$ which includes 99% of scattered particles. Adapted from [21].

tion somewhere more downstream will also be Gaussian [21]. The same argument holds true for the distribution of the remaining transverse position y because of symmetry reasons. This means that lateral beam profiles that are Gaussian stay Gaussian when undergoing MCS in Gaussian approximation. A situation where the Gaussian approximation cannot be used is for example when one examines the nuclear halo of a pencil beam traversing through water [22]. The halo is a volume surrounding the core region of a pencil beam's lateral dose distribution. It consists of charged secondary particles stemming from primaries that underwent a single large-angle scattering with an H- or O-atom. The single-large angle scattering can be due to an electromagnetic elastic process and thus be described by the large-angle scattering tail of Molière's distribution or even to nuclear scattering processes. Even though the halo does not affect the shape of the core's dose distribution, it contains a significant fraction of the total dose deposited which is why it has to be taken into account during dose calculation for treatment planning.

2.5.4. Dose and IDD distributions

In order to discuss the radiobiological effects of charged particles losing energy when traversing through matter one has to consider the energy that is deposited in the material. The quantity used to describe the energy dE deposited in a infinitesimal mass element dm is the dose D

$$D = \frac{dE}{dm} \quad (2.26)$$

which is given in Gray [$\text{Gy} = \text{J/kg}$].

As already introduced in 2.1.1 the integrated depth-dose (IDD) distribution of a particle beam shows the laterally integrated dose deposited as a function of penetration depth z . Figure 2.19 illustrates the IDD distribution of a monoenergetic 62.4 MeV proton beam stopping in water. The whole curve is also called Bragg curve. It consists of a sub-peak region and the actual Bragg peak. Since stopping power increases with the inverse square of particle velocity, the dose increases gradually along the penetration depth, culminating in the Bragg peak. The Bragg peak is the maximum dose peak located just before the stopping range of the particles. After the maximum, i.e. at its distal end, the curve exhibits a sharp dose falloff.

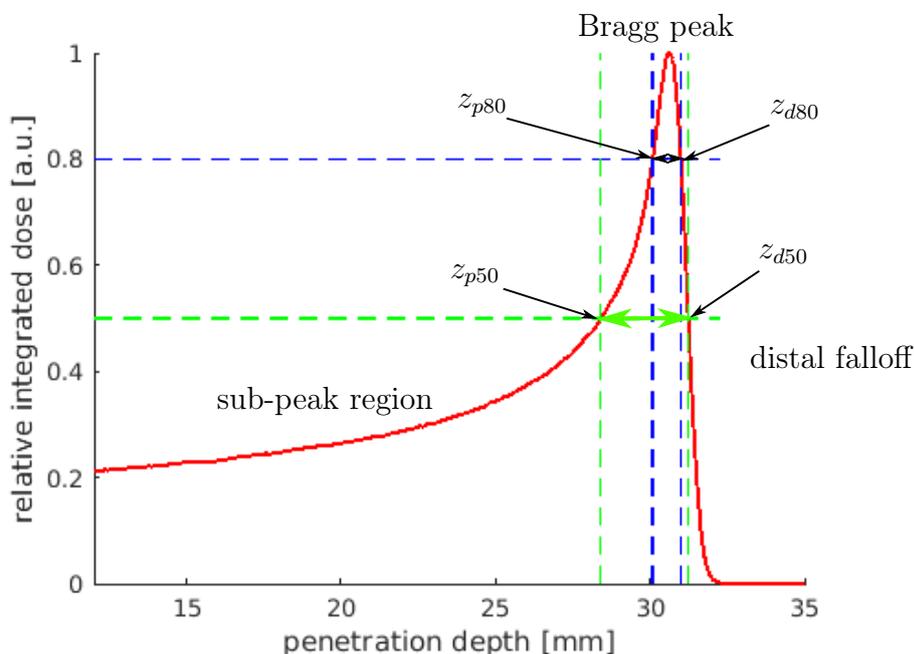


Figure 2.19.: Illustration of the integrated depth-dose distribution of a monoenergetic 62.4 MeV proton beam stopping in water. The dose is normalized to the dose maximum appearing at the Bragg peak.

Figure 2.19 also shows the definitions of four specific depths: z_{d50} , z_{p80} , z_{d80} and z_{p50} . The depth z_{d80} is called the distal-80% depth and refers to the distal-most depth at which the absorbed dose has dropped to 80% of the maximum dose at the Bragg peak. The depth z_{p80} is called the proximal-80% depth and refers to the proximal-most depth at which the absorbed dose equals 80% of the maximum dose.

The other two depths z_{d50} and z_{p50} are defined analogous to the former ones. The nomenclature was adapted from [42]. These depths allow to define two different measures of the Bragg peak width: the 50%-width w_{50} and the 80%-width w_{80} that can be calculated via

$$w_{50} = z_{d50} - z_{p50} \quad (2.27)$$

$$w_{80} = z_{d80} - z_{p80} \quad (2.28)$$

2.5.5. Range and range-energy relationship

When defining range in an absorbing medium, one has to distinguish between the range of a single particle and the range of a particle beam.

In the case of a single particle the range is defined as the path length it can travel in the medium until its kinetic energy drops to the thermal energy of the medium. It depends on the particle type, the initial energy of the particle and the type of the material. However, even when these 3 parameters are held constant, the particle range is not a well-defined number. This is due to the fact that the energy loss of a single particle travelling through a material is not a continuous but a statistical process. Two particles with the same energy and starting conditions will very likely not experience the same number of collisions, the same amount of energy loss per collision and very likely not follow the same zigzag path through the material caused by multiple Coloumb scattering. The Bethe-Bloch equation 2.25 thus only gives the mean stopping power of a particle. This phenomenon is called range straggling and results in a distribution of particle ranges, which is Gaussian in first approximation.

In contrast, the range R of a particle beam is usually defined as the depth at which 50% of the incident particles come to rest [42]. The higher the initial kinetic energy E_0 of the particle beam, the higher the beam range R . The range-energy relationship for clinical energies can be approximately described by a simple power law, called the Bragg-Kleeman rule [59]

$$R(E_0) = \alpha \cdot E_0^p \quad (2.29)$$

in which $p \approx 1.8$ for energies in the range of 10 to 250 MeV and α is a material-dependent constant.

However, since treatment planning systems in general deal with absorbed dose and not with number of particles, range is defined with respect to the distal depths of the Bragg peak curve within the scope of this thesis. This allows to define two measures of range: the 80%-range $R_{80} = z_{d80}$ and the 50%-range $R_{50} = z_{d50}$. Both ranges have a qualitatively similar range-energy relationship as expressed by equation 2.29. Figure 2.20 shows IDD profiles of three monoenergetic proton

beams with different initial kinetic energies E_0 stopping in water. One can observe

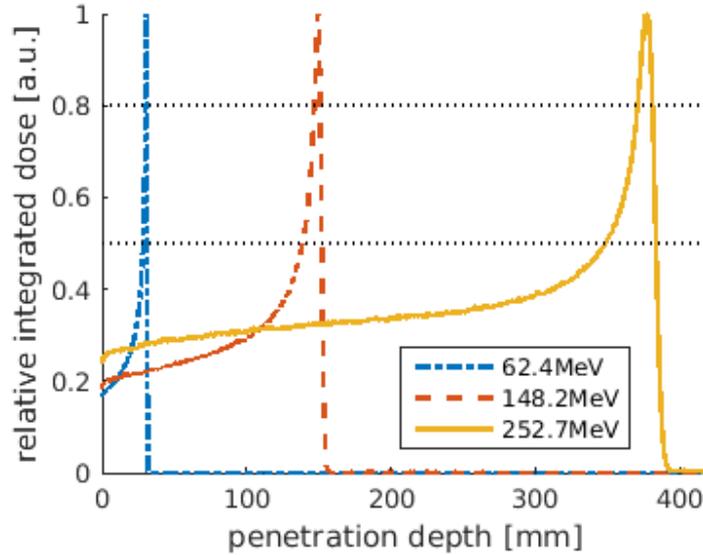


Figure 2.20.: IDD distributions of three monoenergetic proton beams with initial energies E_0 of 62.4MeV, 148.2MeV and 252.7MeV stopping in water. Horizontal lines at the 50% and 80% relative dose level are drawn for guiding the eye. Both the beam range and the beam width increase with increasing E_0 .

that not only range but also the Bragg peak width increases with increasing initial energy E_0 which is due to range straggling. The higher the initial energy E_0 the higher the range of the beam which in turn means that the particles travel a longer distance and thus have higher chances to develop energy deviations between each other. The quantification of this relationship will be covered in the following subsection.

2.5.6. Energy straggling and peak width

Energy straggling describes in principal the same phenomenon as range straggling, just from a different perspective. Instead of looking at the range deviations one considers directly the energy deviations between the particles resulting from the statistical nature of energy loss and multiple Coulomb scattering. The initially delta peak-shaped energy distribution of a monoenergetic beam does not remain monoenergetic when traversing through a material of a given thickness x . Once in the material, the delta-function starts to widen and at the same time is shifted

down along the energy axis by the mean energy loss $\bar{\Delta}$ given by

$$\bar{\Delta} = \int_0^x \left(\frac{dE}{dx'} \right) dx' \quad (2.30)$$

If the absorber material thickness x is high enough, i.e. if the particles experience many individual collisions while traversing the material, one can use Bohr's theory of energy straggling in order to analytically calculate the energy distribution function $f(x, \Delta)$ of the initially monoenergetic beam after having traversed the material with thickness x [42]

$$f(x, \Delta) = \frac{1}{\sqrt{2\pi}\sigma_E(x)} \exp \left[-\frac{1}{2} \left(\frac{\Delta - \bar{\Delta}}{\sigma_E(x)} \right)^2 \right] \quad (2.31)$$

This means Bohr's theory predicts a Gaussian probability density function with mean $\bar{\Delta}$ and standard deviation σ_E in which the energy loss Δ is the independent variable and the material thickness x is a parameter, that enters the formula through the energy spread $\sigma_E = \sigma_E(x)$. The theory yields [36]

$$\sigma_E(x) = 4\pi N_a r_e^2 (m_e c^2)^2 \rho \frac{Z}{A} \cdot x \quad (2.32)$$

where N_a denotes Avogadro's number, r_e the classical electron radius, m_e the electron rest mass, c the speed of light and ρ , Z , A the density, atomic number and atomic weight of the material, respectively. It theoretically explains why the effects of energy straggling increase with increasing distance x traveled in the material. If one deals with intermediate or thin absorbers, Bohr's theory cannot be applied. However, the Landau or Vavilov theory provide an appropriate framework to calculate the energy distribution function in these situations (see [36]). For most applications in proton therapy, however, it is sufficient to describe energy straggling using Bohr's theory [42].

Instead of working with the energy spread σ_E caused by energy straggling due to *traversing a material* one can also use the range spread σ_R of the range distribution of particles *stopping in the material*¹⁷. Instead of σ_E depending on material thickness x , σ_R now depends on the beam range R . For clinical proton beams stopping in water the range spread-range relationship can be well approximated using

¹⁷i.e. approaching the problem from the range straggling-point of view

a simple power law [59]

$$\sigma_{R, \text{straggling}}(R) = 0.012 \cdot R^{0.935} \quad (2.33)$$

The range spread σ_R is directly related to the Bragg peak widths w_{80} or w_{50} : the higher the beam range R , the higher the range spread σ_R , the higher the widths of the Bragg peak, as already figure 2.20 showed.

Up to now only monoenergetic beams were considered. However, in practice a beam extracted from a synchrotron does not have a delta-peak shaped energy distribution but any other energy distribution with a finite standard deviation. If one assumes the initial energy distribution of the beam to be Gaussian with mean energy E_0 and an initial energy spread σ_{E_0} , equation 2.33 can be extended such that the resulting squared range spread σ_R^2 now becomes the sum of variances [59]

$$\sigma_R^2(R, \sigma_{E_0}) = \sigma_{R, \text{straggling}}^2 + \left(\frac{dR}{dE_0} \right)^2 \sigma_{E_0}^2 = \dots = \sigma_R^2(E_0, \sigma_{E_0}) \quad (2.34)$$

in which $\sigma_{R, \text{straggling}}$ denotes the range spread for a monoenergetic beam due to range straggling (see equation 2.33). The dots indicate calculations in which the Bragg-Kleeman rule 2.29 was used to substitute the range R with the initial mean energy E_0 . The result shows that the range spread σ_R and thus the Bragg peak width in general depends on both the initial mean energy E_0 and the initial energy spread σ_{E_0} of the beam. It also points out that the finite width of a Bragg peak is in practice the combined result of two factors:

1. the phenomenon of range straggling
2. the existence of a finite spread of the initial energy distribution σ_{E_0}

In fact, if one was able to "turn off" energy straggling completely, the Bragg peak of a monoenergetic beam stopping in a material would degenerate to a singularity as shown in [59].

In summary, it was shown that the energetical parameters (E_0, σ_{E_0}) of a charged particle beam, that is the mean E_0 and the standard deviation σ_{E_0} of the beam's initial energy distribution, both influence the range of the Bragg curve and the width of the Bragg peak.

2.6. Monte Carlo methods

In general Monte Carlo (MC) methods are a broad group of algorithms that can provide a numerical solution to complex problems using a statistical approach. The algorithms are based on the repeated sampling process of suitable random numbers. Consequently, the solutions obtained by Monte Carlo methods obey the law of large numbers which states that as the number of samples grows larger the observed averaged results tend to approach the theoretical, expected results.

In radiotherapy, Monte Carlo methods model the temporal transportation of particles (protons, electron, photons, neutrons, etc.) considering the interactions with other particles (often the charged nuclei of the traversed material) using theoretical particle-particle interaction cross sections [55]. As a result, MC methods simulate from first principles, that is, every interaction is simulated on a microscopic level. The MC simulation generates a primary particle¹⁸ with a certain initial kinetic energy, position and velocity by sampling them from a user-specified energy and phase-space distribution. In the next step the drift path of the primary is followed until the primary encounters an interaction partner. Now the interaction channel, as well as the energy transfer and the deflection angles are sampled for the primary according to the quantum mechanical probability distributions. During such interactions also secondary particles like electrons or photons can be created. This drift-and-interaction process is repeated until the particle stops, leaves the region of interest or its energy drops below a user-specified cut-off value. Then the whole process is repeated by starting again with the generation of the next primary particle. Through this kind of simulation quantities of interest can be determined, e.g. the spatial distribution of dose in a certain geometry or the total set of phase-space coordinates of a charged particle beam at a certain axial position.

The accuracy of MC simulations that was already discussed in section 2.3 is a result of simulating from first principles. It is only limited by statistics and the underlying physics-related implementations, i.e. the theoretical interaction cross sections. In order to obtain a statistically acceptable result a sufficiently high number of primaries has to be simulated. More primaries, however, also imply a higher runtime of the simulation. As a result, computational resources and time pose a natural limitation of the number of primaries simulated. Although the runtime of simulation increases linearly with the number of particles, the obtained simulation statistics is only improved by the square root of the number of primaries [45]. This is why in practice always a trade-off between accuracy and simulation speed has to be found.

¹⁸in the following referred to as *primary*

For accurate independent dose calculations, MC simulations are currently the method of choice. Consequently, these days MC simulation *toolkits* that are geared toward applications in radiotherapy are already available. Most of them were originally developed for applications in high energy physics and popular existing codes include GEANT4 [1], GATE [24] and FLUKA [15].

2.7. Beam model

In section 2.3 it was stated that the predicted dose distribution of a treatment plan has to be checked against the dose distribution obtained from an independent reference method. As a reference method often IDCs in the form of MC simulations are used.

However, before the IDC can be performed a *beam model* is necessary that provides information about the used treatment beam to the MC simulation. The beam model consists of a certain number of beam parameters that allow to reconstruct the energy distribution and the phase-space of the beam at a defined point on the beam axis called the *source point*. In case of proton beam therapy the source point is often located near the medical nozzle. The MC simulation is finally provided with these beam parameters as initial parameters. Thus, the beam model constitutes the link between reality, i.e. the properties of the beam emerging from the treatment machine, and the MC simulation that is used to perform the IDC (the so-called dose calculation engine).

The beam parameters of the model have to fulfill certain requirements [32]. First of all, the extraction of the beam parameters from beam or dose measurements has to be possible. The parameters should also have a clear connection to reality. Additionally, the beam parameters should allow to reconstruct the phase-space of the beam at the source point in such detail so that it is possible to fulfill the accuracy demands for the comparison of dose distributions.

A possible set of beam parameters can be made up by the two energetical beam parameters (E_0, σ_{E_0}) introduced in section 2.5 and the six optical beam parameters ($\sigma_x, \theta_x, \epsilon_x, \sigma_y, \theta_y, \epsilon_y$) introduced in section 2.4. Putting these together, one obtains

$$(E_0, \sigma_{E_0}, \sigma_x, \theta_x, \epsilon_x, \sigma_y, \theta_y, \epsilon_y) \quad (2.35)$$

in which E_0 denotes the initial mean energy and σ_{E_0} the initial energy spread which are the mean and standard deviation of the initial energy distribution. σ_x and θ_x respectively denote the spread of the spatial and angular distribution in

x -direction while ϵ_x denotes its geometric emittance. The parameters for the y -direction are defined similarly.

This beam model consists of 8 beam parameters¹⁹. All 8 quantities in general depend on the axial position z and the system energy E_{sys} . The *system* energy E_{sys} refers to the energy that is given as input to the treatment *system*, i.e. the energy that is actually set for synchrotron acceleration. It is important to note that the system energy E_{sys} does not necessarily coincide with the mean energy E_0 of the energy distribution of accelerated particles and thus must not be confused with E_0 .

For the beam model one is only interested in the beam parameters evaluated at the axial position of the source point $z = z_{source}$. Finding a beam model thus means to determine the values of all 8 beam parameters as a function of system energy E_{sys} .

¹⁹this number, however, is not universal since the beam model can also consist of less or more parameters (see for example [32])

3. Materials

3.1. Medical nozzle

MedAustron's research room IR1¹ hosts a medical nozzle identical to the nozzles in the respective clinical IRs². However, in contrast to the other IRs the nozzle in IR1 is *removable* which is a unique feature for ion beam therapy facilities. Due to safety regulations a nozzle normally has to be used during patient treatment which is why in the clinical irradiation rooms the respective nozzles are not removable. The removable nozzle in IR1 allows to conduct unique experiments in which the effects of the nozzle on the treatment beam can be investigated.

An illustration of the nozzles' interior design is shown in figure 3.1. The figure also indicates the position of the source point that was chosen as the reference point for beam modeling. It is still located within the vacuum pipe on the beam axis with an isocenter distance of $ISD = -130$ cm. The nozzle is situated downstream of the vacuum pipe and the double vacuum windows. Both vacuum windows are separated by a 14.5 mm air gap and have a thickness of 0.19 mm each. The nozzle is about 44 cm long and consists of three fixed active elements and three passive elements which can be brought into and removed from the beamline using a remotely controlled motor. The space between these elements is filled with air.

The active elements include an independent termination system (ITS) and two dose delivery systems (DDS). The DDS monitor the applied dose and the spot position of the treatment beam and can correct small deviations of the spot position via a feedback loop to the scanning magnets. Since there is always the possibility of a sudden break down of the DDS during irradiation, safety regulations require a second dose measurement system to be used within the nozzle which is independent from the DDS. This system is the ITS that can also monitor the applied dose. The passive elements consist of two ripple filters (RiFi 1 + 2) and a range shifter (RS). The ripple filter widens the Bragg peak and is necessary for irradiation with carbon ions since these exhibit a relative sharp Bragg peak. Thus, a ripple filter helps to reduce the number of energy steps needed to create a SOBP³ [25]. The

¹IR ... irradiation room

²i.e. IR2, IR3 and IR4

³spread-out Bragg peak; see subsection 2.1.2

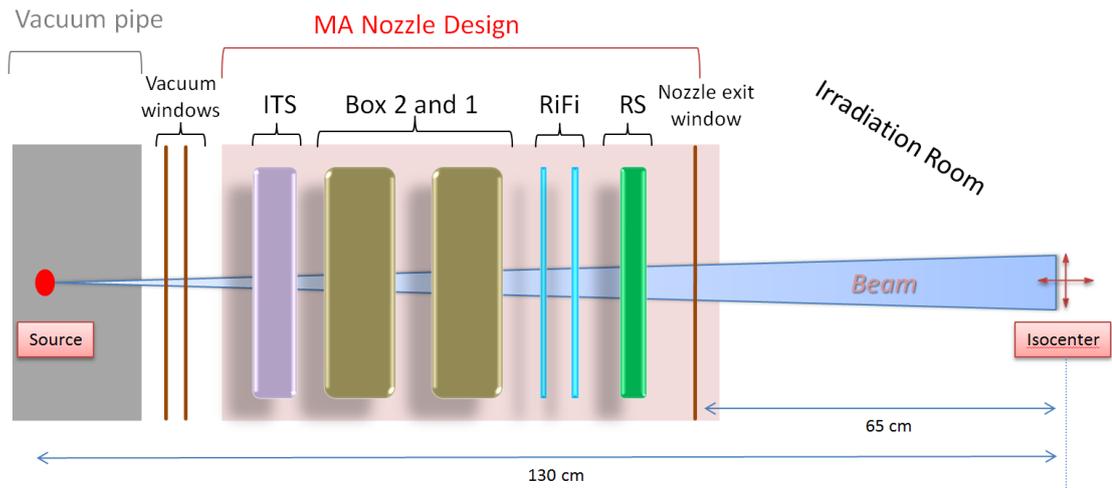


Figure 3.1.: Illustration of the MedAustron (MA) nozzle's interior design (not to scale). The dose delivery systems (DDS) are denoted as "Box 2 and 1". The source point that was chosen as the reference point for developing the beam model is also indicated. Image source: [17]

range shifter, consisting of a PMMA⁴ plate, allows to reduce the minimal beam range in tissue. This is necessary for treating superficial tumors since the lowest possible extraction energy at MedAustron of about 60 MeV corresponds to a beam range of already 3 cm in tissue [46]. After the beam has traversed the interior components of the nozzle it leaves the device via the nozzle exit window.

For the treatment beam the nozzle poses a scatterer that not only introduces divergence and emittance growth of the beam through scattering (see 2.4.4) but also energy loss and energy straggling (see 2.5.6). This means a beam that traversed the nozzle will have a higher transverse x- and y-emittance, a lower Bragg curve range and a larger Bragg peak width in comparison to a beam that only traversed the double vacuum window in the case of the nozzle removed.

3.2. Measurement devices

3.2.1. Water phantom MP3-PL

The water phantom MP3-PL (PTW, Freiburg, Germany) is a water tank that consists of 20 mm thick PMMA walls and is equipped with a 3D scanning system on which different types of detectors can be placed (see figure 3.2). It has

⁴Polymethyl methacrylate; better known as acrylic glass

an exchangeable 250x250x5 mm³ entrance window made of PMMA such that the phantom cannot only be used for vertical but also horizontal beams.

When used with a PinPoint chamber block, the phantom allows to automatically



Figure 3.2.: Water phantom MP3-PL from PTW. For measurements the tank is filled with water. Image source: [51]

acquire dose distribution measurements. When used with Bragg peak chambers the phantom allows to automatically measure the Bragg curve of charged particles stopping in water. In this case one chamber is mounted to the entrance window outside of the water tank as reference chamber and the second chamber is mounted to the moving bar for relative dose measurements at varying penetration depths. The moving bars are made of stainless steel and enable a scanning range of 405 mm vertically and 600x500 mm² horizontally. Stepper motors are used to position the detectors at any position within the phantom with a position accuracy of ± 0.1 mm [51]. The integrated MEPHYSTO software package allows to acquire and evaluate measured data.

3.2.2. PinPoint chamber 31014

The PinPoint chamber 31014 (PTW, Freiburg, Germany) is an ionization chamber that can be used for high-resolution dosimetry in radiotherapy (see figure 3.3). It is cylindrical shaped with a radius of 1 mm, a length of 5 mm and a nominal sensitive volume of 0.015 cm³. Like any other ionization chamber it consists of a gas-filled chamber with two electrodes known as cathode and anode. In the case of the PinPoint chamber 31014 the cathode surrounds the cylindrical chamber



Figure 3.3.: Picture of the PinPoint ionization chamber 31014. Image source: [50].

whereas the anode is implemented as a coaxially positioned wire. When ionizing radiation traverses the chamber, gas particles are ionized which leads to a production of negatively charged electrons and positively charged ions. Due to the applied electrode voltage these charge carriers are attracted toward the electrodes and can be measured as a current. The higher the deposited dose in the fill gas, the higher the measured current. In the case of the PinPoint chamber 31014 a nominal electrode voltage of 400 V is applied. The resulting nominal response is 400 pC/Gy [50].

3.2.3. Lynx PT

The Lynx PT (IBA Dosimetry, Schwarzenbruck, Germany) is a 2D high resolution detector that consists of a scintillator coupled with a CCD camera (see figure 3.4). It can be used for measuring the lateral profile and the spot position of a treatment beam, respectively. The active area of the scintillating screen is 300x300 mm². When a particle impinges on the scintillating screen it is converted into photons which are subsequently detected by the CCD camera. The resolution of the CCD camera results in a total resolution of 0.5x0.5 mm² of the whole measurement device. The original analog signal is digitized by an ADC (analog-to-digital converter) using a 12-bit format. The data is then transferred to a computer that is connected with a standard ethernet cable. Once image acquisition is completed, the IBA software allows several post-correction and beam profile analysis procedures.



Figure 3.4.: Picture of the Lynx PT scintillator-based detector. Image source: [13]

3.3. Software

3.3.1. GATE

GATE (**G**eant4 **A**pplication for **T**omographic **E**mission) is an open-source MC simulation toolkit which is developed by the international OpenGATE collaboration [10]. It was originally developed for MC simulations for emission and transmission tomography [29]. Still being used for its original purpose, GATE was later on extended to simulations of experiments in radiotherapy, especially in particle therapy.

As the acronym already suggests, GATE is based on the MC simulation framework GEANT4 (GEometry ANd Tracking) [1]. Developed by the international GEANT4 collaboration, GEANT4's current fields of application include high energy physics, nuclear physics, accelerator physics, medical science and space science. Even though GATE and GEANT4 are written in C++, GATE does not require any knowledge of C++. This is because GATE features a user-friendly macro language mechanism that is based on scripts. GATE therefore acts mostly as GUI (graphical user interface) and facilitates the development of MC simulations.

3.3.2. ROOT

ROOT is a data analysis framework that was developed by the high energy physics community at CERN in order to process the enormous amounts of data produced by LHC experiments.

Vast amounts of data are also created during GATE simulations since these take into account a lot of different particles and their interaction with each other. As a result, ROOT is today also used as a storing and data analysis tool for the outputs of GATE simulations.

Data is stored in so-called root files (.root) which use a compression algorithm to keep file size as low as possible. Furthermore, ROOT features a C/C++ interpreter called CINT, which allows the user to apply self-written C/C++ scripts to the data stored in root files [3].

3.3.3. MoccaMed Cluster

MoccaMed (**M**onte **C**arlo **c**alculations in **M**edicine) is a collaboration within the Medical University of Vienna that established an inhouse grid (MoccaMed cluster) used for Monte Carlo simulations and other computationally intensive research applications [43]. The MoccaMed cluster uses the high-throughput batch processing system Condor and allows to run a GATE simulation concurrently on multiple PCs. This is achieved with a tool called *GJS* (GATE job splitter) which splits an entire simulation into n subjobs that are sent to multiple CPU cores. After the simulation has ended, the n independent output files are merged (i.e. summed up) to a single output file used for further analysis.

3.3.4. MATLAB

MATLAB⁵ is a proprietary programming language developed by MathWorks, Inc. (Natick, USA). It is a very powerful application for numerical computing. Besides the development of algorithms, its environment also enables plotting of data and allows interfacing with other languages. Thus, MATLAB can call functions written in C, Fortran or can even invoke Linux shell scripts.

The current versions of MATLAB feature a so-called *Optimization Toolbox* that provide solvers for different kind of optimization problems ranging from linear programming, quadratic programming and non-linear programming to least square methods and data fitting. If the optimization problems contains multiple extrema

⁵abbreviation for **M**atrix **L**aboratory

MATLAB also offers the *Global Optimization Toolbox* that is licensed separately from the Optimization Toolbox solvers. It contains solvers that allow to search for global solutions of an optimization problem including particle swarm solvers, simulated annealing solvers and genetic algorithms.

Throughout this thesis, MATLAB 8.6 (R2015b) was used for the optimization of both energetic and optical beam model parameters.

3.3.5. RayStation TPS

The RayStation TPS (RaySearch Laboratories, Stockholm, Sweden) is a treatment planning system that can be used in treatment planning for several external beam therapy modalities. The RayStation TPS is currently used at MedAustron for treatment planning in proton beam therapy for which it offers various tools along the entire treatment planning workflow. Tools include multiple frameworks for treatment plan optimization, especially in combination with adaptive planning and the simulation of organ motion.

4. Beam modeling

Beam modeling describes the process of finding an optimal set of parameters, the so-called *beam parameters*, that describe the beam emerging from a treatment machine. The set of optimal beam parameters is also called *beam model*.

At the very beginning of beam modeling one has to decide on a set of beam parameters. The eight parameters ($E_0, \sigma_{E_0}, \sigma_x, \theta_x, \epsilon_x, \sigma_y, \theta_y, \epsilon_y$) that were introduced in equation 2.35 in section 2.7 were used for modeling the IR1 beamline. These consist of the two beam energy parameters (E_0, σ_{E_0}) which denote the mean energy and the energy spread of the beam's energy distribution and the six optical parameters ($\sigma_x, \theta_x, \epsilon_x, \sigma_y, \theta_y, \epsilon_y$) which denote the spot size, divergence and emittance for the transverse x- and y-direction, respectively.

The structure of the chosen parameters allowed to separate the actual process of beam modeling into two parts: *beam energy modeling* and *optical beam modeling*. The aim of beam energy modeling was to determine the initial energy parameters (E_0, σ_{E_0}) as a function of system energy E_{sys} ¹. Once the beam energy model was obtained² and validated, optical beam modeling could be performed by determining the initial optical beam parameters ($\sigma_x, \theta_x, \epsilon_x, \sigma_y, \theta_y, \epsilon_y$) as a function of E_{sys} .

Before the actual process of beam energy modeling and optical beam modeling will be presented, a general introduction into the topic of beam modeling is given in the following section.

4.1. General concept

The most-straightforward approach to determine beam parameters as a function of system energy E_{sys} would be to simply measure them for varying E_{sys} in the source

¹The *system* energy E_{sys} refers to the energy that is given as input to the treatment *system*, i.e. the energy that is actually set for synchrotron acceleration. It is important to note that the system energy E_{sys} does not necessarily coincide with the mean energy E_0 of the energy distribution of accelerated particles and thus must not be confused with E_0 .

²Beam energy modeling always has to precede optical beam modeling since the beam energy influences scattering and thus the optical beam parameters.

point. However, this approach turns out to be infeasible for two reasons. First of all, some of the chosen parameters are not straightforward to measure directly, e.g. divergence or emittance. Secondly, the source point may be not accessible with the necessary measurement devices.

As a consequence, one has to find a measurable quantity that is sensitive to changes of the beam parameter. The measurable quantity has to be functionally related to the beam parameter and respond to changes in the beam parameter. For instance, for the mean energy such a quantity could be the shape of the Bragg curve in water. When the mean energy E_0 is shifted, also the Bragg curve range changes. For the energy spread one could choose the Bragg peak width as a response quantity. For the six optical parameters one could consider spot size-vs-ISD curves³ since the shape of these curves changes when one or multiple of the optical parameters are varied.

As a next step these response quantities need to be measured on one hand and modeled on the other hand. For modeling, GATE simulations are especially suitable for this purpose because GATE allows to easily simulate particle beams using the selected eight beam parameters as input. Measurements and simulations are sequentially compared to each other while the beam parameters are iteratively varied. If the simulation results coincide with the measurements, consequently the beam parameters that were used as input also coincide with the beam parameters of the treatment machine. This of course requires that the simulations are reliable and accurate.

The process just described is illustrated in figure 4.1. One can think of the iterative process of finding the correct beam parameters as an optimization procedure: the beam parameters are iteratively varied until the difference between measured response quantity and simulated response quantity for a particular system energy E_{sys} is minimized (and reaches zero in the best case).

However, a complete beam model means to know the value of all eight beam parameters not for only one particular E_{sys} but for each E_{sys} within the clinical energy range (60-255 MeV). Therefore, one chooses a finite set of representative system energies for which the beam parameters are actually determined. Based on the obtained parameter values for each of these energies, interpolation can be used to determine the parameter values in between the representative energies. This way one can find a beam model with a finite number of energies.

³these are curves that describe the spot size of the beam as a function of isocenter distance (ISD)

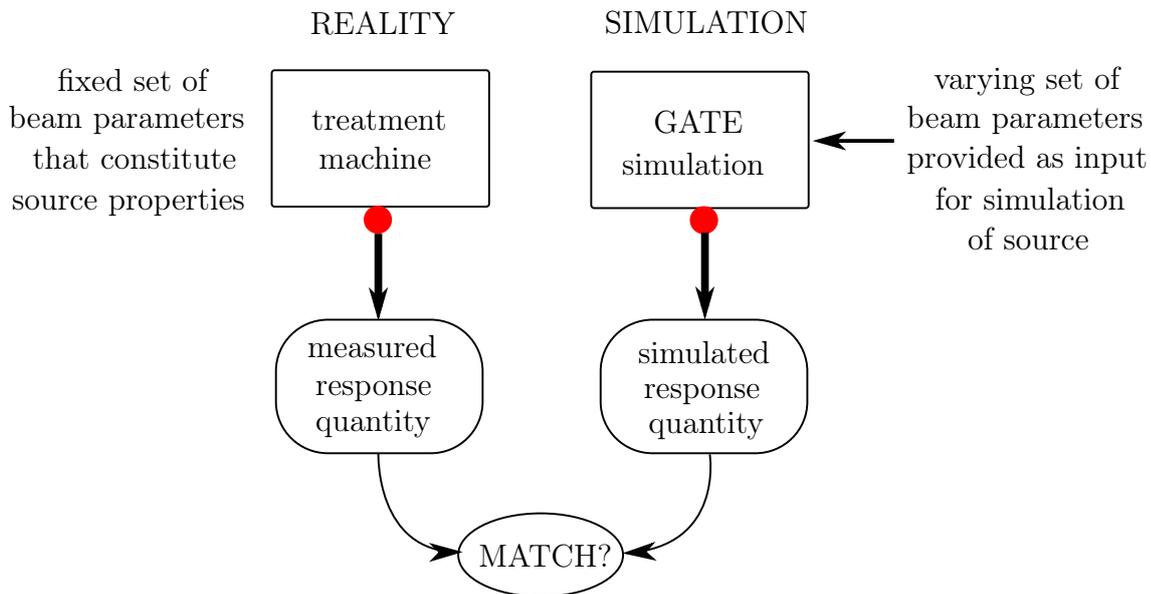


Figure 4.1.: Schematic illustrating the principle of beam modeling. The beam model constitutes the link between reality and simulation. The beam parameters are $(E_0, \sigma_{E_0}, \sigma_x, \theta_x, \epsilon_x, \sigma_y, \theta_y, \epsilon_y)$. The red dot symbolizes the source point for both the treatment machine and the simulation.

Since for beam modeling GATE simulations and optimization procedures are required, the next two sections will introduce their basics in more detail.

4.2. GATE simulations

GATE allows to simulate the passage of different particles through a variety of materials. Before the actual simulation, the materials and their dimensions have to be defined, the particles have to be generated and a lot of other user-specified parameters have to be set. This is done in so-called macro files (.mac).

All GATE simulations used for beam modeling had a similar structure which is illustrated in figure 4.2. The actual simulation file *main.mac* was split up in several sub-macro files which facilitated development and debugging of the source code. The file *main.mac* is the main file that connects the submodules with each other.

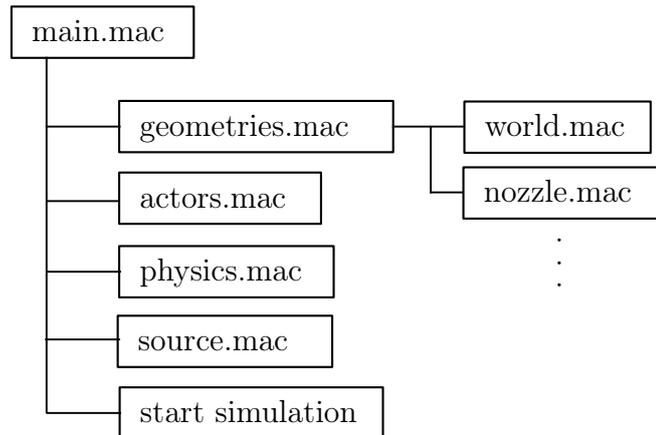


Figure 4.2.: Architecture of GATE simulations used for beam modeling.

This file was also used to set main simulation parameters such as the number of primaries. In the following, the role of each sub-macro file will be explained in further depth.

geometries.mac

The file `geometries.mac` allows to create geometries with user-specified shapes and dimensions and can be filled with a variety of materials. Every geometry was created within a separate macro file while the parent file `geometries.mac` connected the sub files with each other.

The first geometry that must be defined before other geometries is the *world* geometry. It basically sets the surrounding spatial limits for the following setup. An important geometry for beam modeling in our case is the MedAustron nozzle, for which a fully detailed model in GATE already existed at the beginning of this thesis. The nozzle could be simply added to or removed from the simulation setup by including or excluding the file `nozzle.mac` which describes the detailed nozzle setup. Since in IR1 the nozzle can be physically removed from the beam path, it was possible to find a beam model for both situations with and without nozzle.

Other important geometries for beam modeling include a box filled with water (modeling the water phantom) or planes of air with phase-space actors⁴ attached (modeling the Lynx PT), modeling the respective measurements.

⁴see next subsection

actors.mac

As stated before, also the measurement setup needs to be modeled in GATE. Geometries allow to set the spatial position and orientation of a measurement device in GATE. However, one additionally has to model the actual measurement functionality of such a measurement device. This can be implemented in GATE by so-called *actors*. Actors are attached to a previously defined geometry, collect information during the simulation and output this information to files. Two types of actors were used in this thesis for beam modeling: the dose actor and the phase-space actor.

Dose actor

In the case of modeling a water phantom, one needs to be able to also model the measurement of the dose distribution caused by the beam stopping in a water phantom. This can be done using the dose actor. First the actor has to be attached to the water phantom geometry. Then the water phantom needs to be divided into separate voxels with specific dimensions. The user can do so by setting the resolution of the dose actor in x-, y- and z-direction. The dose actor thus records the energy deposited in every voxel and outputs this information in form of a .txt file or .mhd/.raw file.

Phase-space actor

In order to measure lateral beam profiles (e.g. like the Lynx PT does) in GATE, the phase-space actor can be used. The phase-space actor has to be attached to a geometry. The actor then records the position $\vec{r} = (x, y, z)^T$ of each particle where it entered the geometry and stores this information in .root files. If the geometry is positioned perpendicular to the beam axis, the transverse intensity distribution of the beam is obtained. The distribution can be further analyzed and quantities of interest, e.g. the spot size, can be extracted.

The phase-space actor also allows to record the particle's direction of movement as well as its kinetic energy when entering the associated geometry. Hence, the entire phase-space of the simulated particle beam can be reconstructed.

physics.mac

GEANT4 as the basis of GATE allows to simulate a myriad of interactions between beam particles and the particles of the traversed material. These interactions cover electromagnetic processes, weak interaction processes and hadronic processes ranging from some eV to the TeV regime. The models that simulate these processes

contain several parameters and settings that can be changed by the user. The user does so by including a so-called *physics list* into the physics.mac file.

Other parameters that can be set in the physics.mac file and can severely influence the simulated physics are *production cuts* and the *maximal step size*.

Production cuts

Production cuts are necessary to avoid infrared-divergence. During a Monte Carlo simulation, primary particles can create secondary particles like electrons and photons through processes like ionization and emission of Bremsstrahlung. In order to avoid a diverging number of secondaries at low energies, the user has to set production cuts for electrons, positrons and photons. The cuts are set as a range value in a material, but are internally converted into a corresponding energy value. If the initial energy of a potential secondary particle drops below the cut value, it is not created at all and the remaining energy is deposited locally. Thus, one has to set cuts carefully, especially when modeling measurements where the energy deposition needs to be recorded with high resolution (e.g. for modeling beam profile measurements in a water phantom). For other measurements where energy deposition does not play a role (e.g. modeling scattering effects), cut values are not that important. In these situations they can be set rather high which increases simulation speed.

Maximal step size

When a particle traverses a material during a GATE simulation, it repeatedly experiences interactions with a particle of the material followed by a drift until the next interaction occurs. In reality the interaction probability and the stopping power of the particle changes along a drift, however, this is not considered within the numerical framework of a GATE simulation. Consequently, a maximal drift length between two subsequent interactions has to be defined. This length is the so-called maximal step size. The maximal step size has to be adjusted very carefully in situations where energy deposition (and thus the number of interactions per unit length) has to be recorded with high resolution (e.g. water phantom measurements). For other situations step size can be set rather high in order to gain simulation speed.

source.mac

Before the simulation can be started, a particle source has to be defined that generates the primary particles. GATE offers various types of sources, however, for

finding the beam model of a pencil beam scanning treatment machine (as it is used at MedAustron), the *PencilBeam source* (PBS) and the *TPS PencilBeam source* (TPS-PBS)⁵ are suited the best.

Both sources generate pencil beams with particular optical and energy properties. The user first has to set a source point, i.e. the point in space from which the primary particles emerge, and an emission direction. Then the source has to be provided with the desired energy parameters and optical parameters of the beam, that is with $(E_0, \sigma_{E_0}, \sigma_x, \theta_x, \epsilon_x, \sigma_y, \theta_y, \epsilon_y)$. The sources then internally construct a Gaussian energy distribution with mean E_0 and standard deviation σ_{E_0} and a phase-space ellipse for both transverse x- and y-phase-space. The primary particles are finally sampled from these two distributions and released in emission direction.

The difference between PBS and TPS-PBS lies in the way *in which form* beam parameters are passed to the respective source. For the PBS a single value for each of the eight parameters has to be passed, i.e. the value of the parameters for a certain system energy E_{sys} . In contrast, the TPS-PBS has to be provided with a set of polynomial coefficients for every parameter that describe the polynomial of the respective beam parameter as a function of E_{sys} . In other words, the TPS-PBS can only be used if the complete beam model has already been established. Thus for the actual task of finding a beam model the PBS is more convenient to use. In case one wants to validate the found beam model or perform a treatment plan calculation, the TPS-PBS is better suited.

4.3. Optimization

4.3.1. General definitions

When developing a beam model, one has to vary beam parameters $\mathbf{x} = (x_1, x_2, \dots, x_n)$ ⁶ in such a way that the simulation results converge to the measurement results. This can be achieved by defining the difference between the simulated response quantity $Q^{sim}(\mathbf{x})$ (which depends on the beam parameters) and the measured response quantity Q^{meas} as a function of \mathbf{x}

$$f(\mathbf{x}) = |Q^{sim}(\mathbf{x}) - Q^{meas}| \quad (4.1)$$

⁵TPS is the abbreviation for treatment planning system

⁶ \mathbf{x} denotes here a general set of beam parameters. In a specific case these can for example consist of the set $\mathbf{x} = (E_0, \sigma_{E_0}, \sigma_x, \theta_x, \epsilon_x, \sigma_y, \theta_y, \epsilon_y)$

and minimizing $f(\mathbf{x})$ by varying \mathbf{x} .

This is the mathematical concept of optimization. The optimal beam parameters \mathbf{x}^* can then be obtained as the minimum of the so-called objective function $f(\mathbf{x})$ [8]

$$\mathbf{x}^* = \min_{\mathbf{x}} f(\mathbf{x}) \quad (4.2)$$

Often the minimum has to be found under a set of specific constraints. These can be expressed in general as

$$\begin{aligned} G_i(\mathbf{x}) &= 0, & i &= 1, \dots, m_e \\ G_i(\mathbf{x}) &\leq 0, & i &= m_e + 1, \dots, m \\ l_i &\leq x_i \leq u_i, & i &= 1, \dots, n \end{aligned} \quad (4.3)$$

in which the first line refers to a set of equality constraints, the second line to a set of inequality constraints and the third line to lower bounds l_i and upper bounds u_i of the beam parameters. For instance, two inequality constraint that occur for the optimization of optical beam parameters are

$$\epsilon_x \leq \pi \cdot \theta_x \cdot \sigma_x \quad \text{and} \quad \epsilon_y \leq \pi \cdot \theta_y \cdot \sigma_y \quad (4.4)$$

which result from the definition of emittance (see 2.4.2).

4.3.2. Types of optimization problems

In beam modeling two different types of optimization problems can occur: stochastic and non-stochastic problems. These will be introduced and distinguished in the following. Furthermore, possible solvers that exist for these kind of problems will be presented.

Stochastic optimization problems

MC simulations are used to model real systems. Since a MC simulation relies on random sampling, the simulation outcome will usually differ for every run. This characteristic introduces fluctuations in the simulated response quantity: $Q^{sim}(\mathbf{x})$ will not be the same for every simulation run, even if the same set of beam parameters \mathbf{x} and simulation parameters is used. As shown in equation 4.1 these fluctuations are inherited by the objective function $f(\mathbf{x})$. Consequently, an objective function $f(\mathbf{x})$ that is based on simulation results will most of the times not be smooth, but rather stochastic. Such an objective function is called *stochastic objective function*. For example, one can think of a smooth deterministic objective function $f_d(\mathbf{x})$ to which random noise $f_r(\mathbf{x})$ is added, making the resulting

function $f(\mathbf{x}) = f_d(\mathbf{x}) + f_r(\mathbf{x})$ a stochastic objective function. The added random noise is a result of the limited accuracy of a MC simulation [52]. This so-called *simulation noise* can, however, be reduced by increasing the number of simulated primary particles.

In order to minimize such a stochastic optimization function not any solver can be used. This is because a stochastic objective function poses additional challenges to a solver in comparison to a non-stochastic one. For instance, due to the non-smooth behavior a stochastic objective function exhibits multiple local minima that might not be the global minimum and in which the solver might get trapped. Solvers that can handle these challenges are called *stochastic optimization solvers*. MATLAB provides several such methods, one of them being the **genetic algorithm**⁷ solver. It is part of the Global Optimization Toolbox and is based on the principle of natural selection. Like many other solvers, it has to be provided with an initial guess \mathbf{x}_0 of the optimal beam parameter set. The genetic algorithm-solver then uses random sampling during its optimization procedure, which is why it can also handle stochastic objective functions. However, this randomness can lead to different optimization outcomes, even if the ga-solver is passed the same initial beam parameter set \mathbf{x}_0 at the beginning of each optimization. The ga-solver additionally has to be passed user-specified termination criteria so that it knows when to stop optimizing.

Other methods outside of MATLAB to treat stochastic optimization problems include stochastic gradient methods, regeneration and selection methods (see [57] for a detailed description).

Non-stochastic optimization problems

Due to simulation noise most of the problems occurring in beam modeling are stochastic optimization problems. However, for some problems it might be that simulation noise is so low that the simulated response quantity $Q^{sim}(\mathbf{x})$ and thus the objective function $f(\mathbf{x})$ do not fluctuate significantly. In this case a stochastic optimization solver is not required, since the objective function is smooth and does not exhibit random noise. In such cases it is better to use a deterministic solver, since these are much faster in finding the optimal beam parameters.

MATLAB offers a variety of solvers for non-stochastic optimization problems. One example is the *fminsearch*-solver from the Optimization Toolbox. It is based on the Nelder-Mead simplex algorithm as described in [34] and uses a derivative-free

⁷in MATLAB abbreviated as *ga*

and heuristic search method which approximates a *local* minimum of the objective function. Consequently, `fminsearch` might not find the global minimum of the objective function if multiple minima exist. Similar to the `ga`-solver, `fminsearch` has to be provided with an initial guess \mathbf{x}_0 for the optimal parameter set \mathbf{x}^* by the user. Also, a termination criteria for the solver has to be specified.

5. Beam energy modeling

This section describes how the beam energy parameters E_0 and σ_{E_0} of the proton beam were determined as a function of system energy E_{sys} . As representative energies in the clinical energy range (60-255 MeV) five energies¹ were selected. These are referred to as *key energies* in the following. Based on the obtained parameter values for each of these key energies, polynomial interpolation was used to approximate the parameter values in between them.

As response quantities for the mean energy E_0 the Bragg curve ranges R_{80} and R_{50} was chosen. For the energy spread σ_{E_0} the Bragg peak widths w_{80} and w_{50} were selected as response quantities. These are well-suited, since the ranges increase monotonously with increasing E_0 and the widths increase monotonously with both increasing E_0 and increasing σ_{E_0} , as was shown in subsections 2.5.5 and 2.5.6.

Consequently, for each key energy, Bragg curve measurements were performed previously in a water phantom and measured ranges and widths were extracted. Afterward, the measurements were modeled in GATE to obtain the simulated ranges and widths. The following optimization process consisted in minimizing range and width difference between measurements and simulations by varying the beam parameters E_0 and σ_{E_0} given as simulation input. This optimization problem was solved using MATLAB. After the optimal beam parameters were obtained, they were validated on the measurements.

This procedure was carried out for both the situations with nozzle and without nozzle (i.e. nozzle removed from the beamline). As a result, two different sets of optimized beam energy parameters, $(E_0^*, \sigma_{E_0}^*)_{\text{with}}$ and $(E_0^*, \sigma_{E_0}^*)_{\text{without}}$ were obtained as a function of E_{sys} . These beam energy models will in the following be referred to as *nozzle model* and *vacuum model*, respectively.

In the following the methods outlined above will be explained in more detail.

¹these read as follows: 62.4 MeV, 97.4 MeV, 148.2 MeV, 198.0 MeV and 252.7 MeV

5.1. Measurements

All measurements had already been acquired before this thesis was started. For acquiring the measurement data, a proton beam of the specific key energy was directed to an MP3-PL water phantom. The center of the proximal plane of the phantom was positioned in the isocenter, such that the beam impinged perpendicular onto the phantom and stopped in the phantom for ISDs > 0 (see figure 5.1). The Bragg curve had then been acquired by measuring the relative IDD as

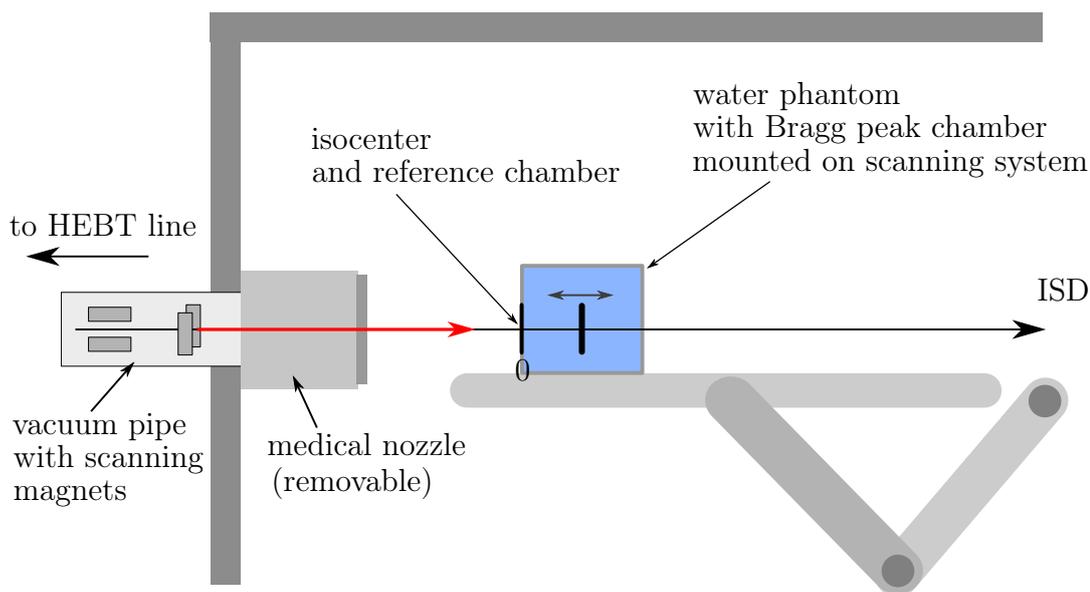


Figure 5.1.: Illustration of the measurement setup for Bragg curve acquisition.

a function of penetration depth using the Bragg Peak chambers from PTW. The step size of the scanning system was adapted for each system energy. Since the Bragg peaks get wider with increasing system energy, the axial distance between two dose measurements was increased such that no significant loss in relative resolution occurred. Table 5.1 summarizes the specific penetration depth resolution values used. For the situation with nozzle, Bragg curves were recorded for every of the 5 key energies. For the situation without nozzle, only curves for three key energies (62.4 MeV, 148.2 MeV and 198.0 MeV) were available.

The measured sets of Bragg curves were analyzed with regard to their ranges (R_{80} , R_{50}) and Bragg peak widths (w_{80} , w_{50}). This analysis was performed using a self-developed MATLAB function called *BP_analyzer.m*. The function has to be passed the filepath to the file that contains the Bragg peak curve (either .txt

system energy E_{sys} [MeV]	penetration depth resolution [mm]
62.4	0.1
97.4	0.2
148.2	0.2
198.0	0.3
252.7	0.3

Table 5.1.: Values of the penetration depth resolution that had been used during measurements for each key energy.

or .rsl) and returns the respective ranges and widths. It uses linear interpolation between the measurement points to allow a determination of ranges and widths with an accuracy that is higher than the depth resolutions listed in table 5.1. The source code of *BP_analyzer.m* can be found in the appendix (see D).

5.2. MC simulations

For all simulations related to beam energy modeling, the "GATE v8.1 developer version" with GEANT4 in version 10.03.p01 was used.

The simulation setups to model the measurement setups for the cases with and without nozzle are shown in figure 5.2.

The world geometry was created using the parameters in table 5.2. The isocenter was defined as the origin of the cartesian coordinate system at (0|0|0). A PBS with source point at (0|0|-130cm)² was created that simulated proton beams with beam parameters (E_0, σ_{E_0}) and $(\sigma_x, \theta_x, \epsilon_x, \sigma_y, \theta_y, \epsilon_y)$ ³.

The water box was created using the parameters in table 5.2. It was filled with G4_WATER⁴. The ionization potential of G4_WATER was set to 78.0 eV which is consistent with ICRU report 73 [27].

²this position is located within the vacuum pipe *before* nozzle and double vacuum windows

³The optical beam parameters of the IR3 beam model were used here for the optical portion of parameters. This approach is possible because for beam energy modeling the beam optics do not play a role as long as the water box has sufficiently large lateral dimensions.

⁴G4_WATER has the same physical properties as the material *Water*, however, when using G4_WATER GATE uses tabulated values for the stopping power in water instead of calculating them via the Bethe Bloch formula which increases simulation speed

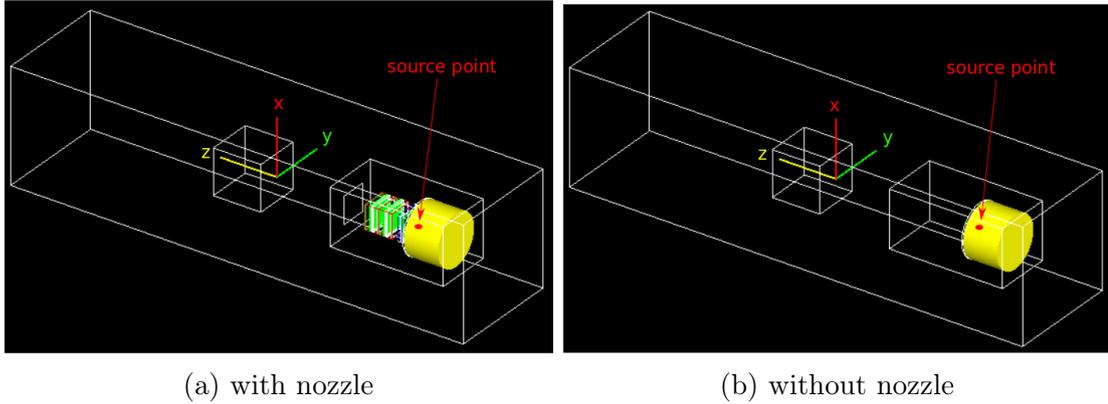


Figure 5.2.: Simulation setups for the situation with and without nozzle visualized in GATE. The yellow cylinder depicts the vacuum pipe in which the source point is located. The water box is depicted by the small white cube near the coordinate system. The world geometry is depicted by the white cuboid surrounding all elements.

A dose actor was attached to the water box geometry to record the IDD distribution of the protons stopping in it. The x-, y- and z- voxelsize of the actor was set to $400 \times 400 \times 0.1 \text{ mm}^3$. The IDD profile was eventually output in form of a .txt file.

This .txt file was subsequently analyzed by the script *BP_analyzer.m* that was used to analyze the measurements before. By doing so the simulated ranges (R_{80} , R_{50}) and Bragg peak widths (w_{80} , w_{50}) were obtained.

All simulation parameters are summarized in table 5.2 where also the used values for maximal step size and production cuts are listed.

The only parameter that has not been discussed so far, but is crucial for running a GATE simulation is the number of primaries. The choice of this parameter strongly impacts the simulation speed, but also the accuracy of the extracted ranges and widths and thus the resulting simulation uncertainty. Therefore, the choice of this parameter has to be considered in combination with the optimization procedure which will be addressed in the following subsection.

geometry	dimensions [cm ³]	material (ion. pot.)	step size	cuts	actor (voxelsize [mm ³])
world	100x100x400	air	5 cm	10 m	-
NozzleBox	-	various	5 cm	10 m	-
waterbox	40x40x42	G4_WATER (78.0 eV)	0.1 mm	0.1 mm	dose actor (400x400x0.1)

Table 5.2.: Summary of simulation parameters. The values displayed for the maximal step sizes are valid for protons and the production cut values are valid for electrons, positrons and gammas.

5.3. Automated optimization

The aim of the optimization algorithm is to find the optimal value pair for main energy and energy spread (E_0, σ_{E_0}) that minimizes the difference in Bragg curve ranges and Bragg peak widths. Thus, the objective function(s) of the optimization problem should contain the absolute range differences between measurements and simulations

$$|R_{80}^{meas} - R_{80}^{sim}| \quad \text{or} \quad |R_{50}^{meas} - R_{50}^{sim}| \quad (5.1)$$

and absolute width differences between measurements and simulations

$$|w_{80}^{meas} - w_{80}^{sim}| \quad \text{or} \quad |w_{50}^{meas} - w_{50}^{sim}| \quad (5.2)$$

or any linear combination of those absolute differences multiplied with different weighting factors. One could come up with several objective functions for this problem, however, a two-step-optimization process was used since it proved to be the fastest and simplest approach which provided a sufficiently high accuracy. The two-step-process split the problem into two separate parts: an optimization of mean energy E_0 followed by and independent optimization of energy spread σ_{E_0} .

It shall be emphasized that this two-step-process is only possible because of a special characteristic of the 80%-range R_{80} . The R_{80} is the only measure that almost solely depends on E_0 and is independent of the energy spread σ_{E_0} ⁵. Figure 5.3 visualizes this phenomenon. It shows the IDD profile of three proton beams stopping in water that have the same initial mean energy $E_0 = 62.4\text{MeV}$ but differ-

⁵this phenomenon stands in contrast to the Bragg peak widths w_{80} and w_{50} which depend on both E_0 and σ_{E_0} as shown in subsection 2.5.6

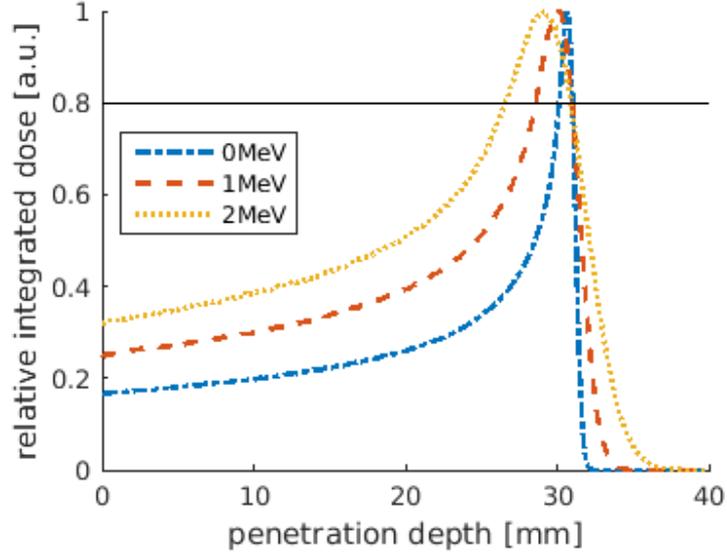


Figure 5.3.: IDD distributions of three proton beams stopping in water that all have the same initial energies $E_0 = 62.4\text{MeV}$, but different initial energy spreads σ_{E_0} of 0MeV, 1MeV or 2MeV. The horizontal line at the 80% relative dose level is drawn for showing that R_{80} is independent of changes in σ_{E_0} .

ent energy spreads σ_{E_0} . While the 50%-range R_{50} increases with increasing energy spread σ_{E_0} , the 80%-range R_{80} remains constant. Therefore, only R_{80} was used as the response quantity for the mean energy E_0 .

5.3.1. Objective functions

Due to the circumstances explained above, the objective function f_1 for the *mean energy optimization* (1st step) was defined as the absolute relative difference between measured and simulated R_{80}

$$f_1(E_0) = \left| \frac{R_{80}^{sim}(E_0) - R_{80}^{meas}}{R_{80}^{meas}} \right| \quad (5.3)$$

The optimal mean energy value E_0^* is then defined by

$$E_0^* = \min_{E_0} [f_1(E_0)] \quad (5.4)$$

which corresponds to finding the global minimum for $f_1(E_0)$ in a 1D parameter space. The only physical constraint for E_0 is $E_0 > 0$.

The objective function f_2 for the *energy spread optimization* (2nd step) should contain the absolute difference(s) of either w_{50} or w_{80} or any superposition of these width differences with a particular set of weighting factors. However, it was decided that only the w_{50} -difference should contribute to the objective function f_2 . The reason behind this choice is that w_{50} is more sensitive to variations in initial energy spread σ_{E_0} than w_{80} . This can be observed in figure 5.3: w_{50} increases more in absolute terms than w_{80} does for the same change in energy spread σ_{E_0} . A high energy spread-width-sensitivity will be especially important later on⁶ for modeling carbon ions since these exhibit even sharper Bragg peaks than protons. Objective function f_2 was thus defined as

$$f_2(E_0^*, \sigma_{E_0}) = \left| \frac{w_{50}^{sim}(E_0^*, \sigma_{E_0}) - w_{50}^{meas}}{w_{50}^{meas}} \right| \quad (5.5)$$

The optimal energy spread $\sigma_{E_0}^*$ is then defined by

$$\sigma_{E_0}^* = \min_{\sigma_{E_0}} [f_2(E_0^*, \sigma_{E_0})] \quad (5.6)$$

In principle this would correspond to a global minimum search in a 2D parameterspace. However, since in the first step E_0 has already been optimized, the problem reduces to a 1D minimization task. The only physical constraint for σ_{E_0} is $\sigma_{E_0} > 0$.

5.3.2. Simulation uncertainties

As already explained in subsection 4.3.2, for simulation-based optimization problems one has to consider the random *simulation noise*⁷ that is added to the objective function, making it a stochastic objective function. The simulation noise can be reduced when the number of primaries is increased, consequently increasing simulation runtime. In order to find a well-balanced trade-off between simulation noise and optimizer runtime, the simulation noise was estimated in terms of the standard deviation of the objective function fluctuations at a specific point in 2D parameterspace (E_0, σ_{E_0}) as a function of the number of primaries. Figure 5.4 and 5.5 show the standard deviation of f_1 and f_2 (each calculated for a sample size of $n = 100$) together with simulation runtime plotted as a function of number of primaries. The results were obtained from GATE simulations with nozzle.

⁶A beam model for carbon ions is not within the scope of this thesis. However, the developed optimization algorithms can be easily adapted to model carbon ions which will be helpful for establishing beam models for carbon ions at MedAustron in the near future.

⁷The terms simulation noise and simulation uncertainties are used interchangeably.

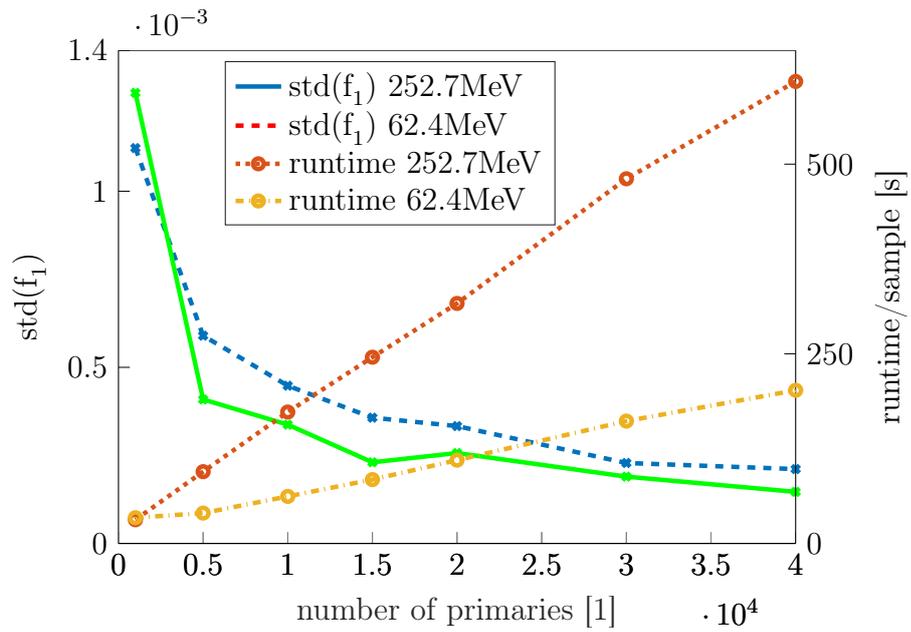


Figure 5.4.: Standard deviation-curves for f_1 and simulation runtime as a function of number of primaries for 62.4MeV and 252.7MeV.

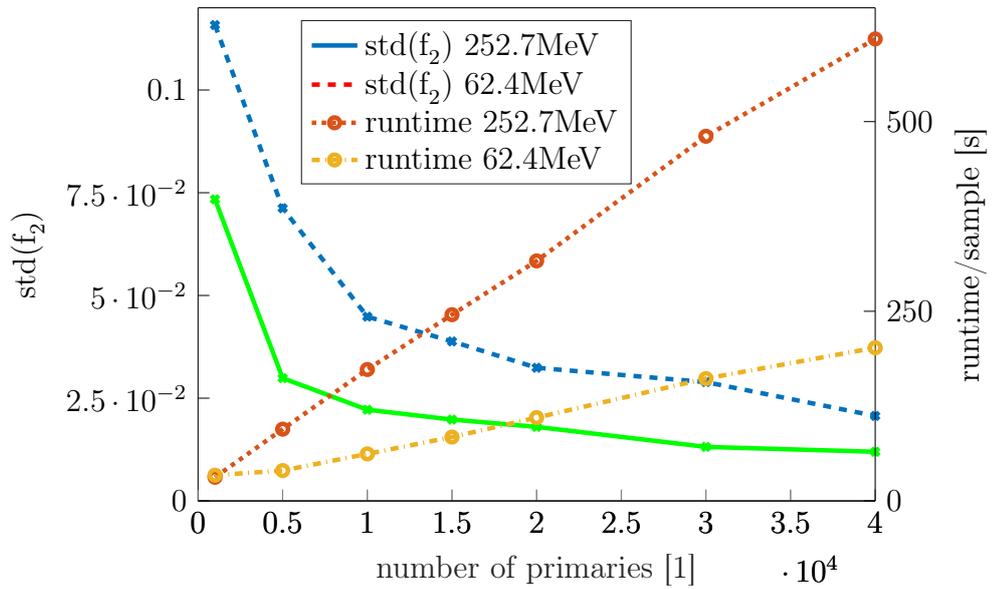


Figure 5.5.: Standard deviation-curves for f_2 and simulation runtime as a function of number of primaries for 62.4MeV and 252.7MeV.

The plots show a linearly increasing runtime and a standard deviation of objective function values that improves with the square root of number of primaries, just as theory in section 2.6 predicts.

Additionally, one can observe that the standard deviation-curves of 252.7 MeV lie above the curves for 62.4 MeV, which means that fluctuations increase with increasing system energy. This is due to the fact that a higher system energy implies more range straggling which in turn implies higher fluctuations in R_{80} and w_{50} . Since 62.4 MeV and 252.7 MeV are the lowest and highest value of all key energies, the standard deviation-curves for the remaining key energies are expected to lie between those two. Therefore, the number of primaries were adapted to each system energy for the actual optimization procedure. Using the number of primaries tabulated in table 5.3 guarantees standard deviations around $2 \cdot 10^{-4}$ for f_1 and $2 \cdot 10^{-2}$ for f_2 within the whole clinical energy range.

system energy E_{sys} [MeV]	number of primaries [1]
62.4	15,000
97.4	20,000
148.2	25,000
198.0	30,000
252.7	40,000

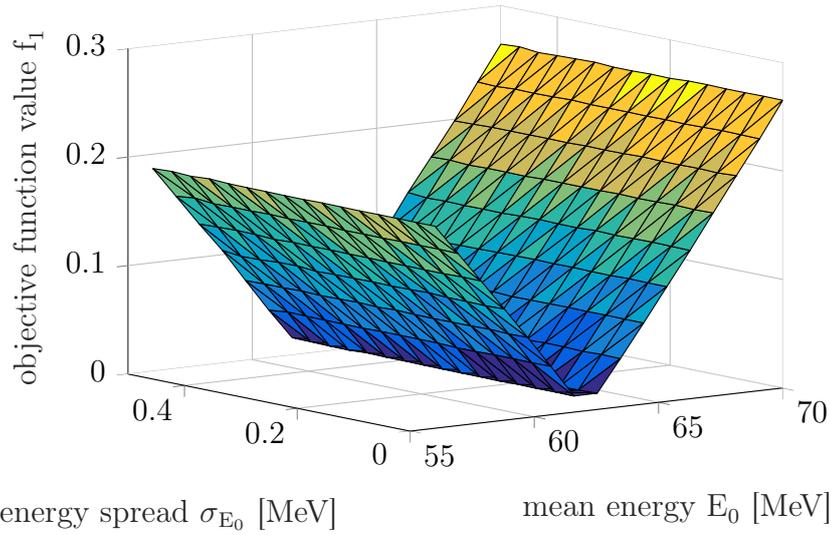
Table 5.3.: System energy-dependent values for the number of primaries that were used for the actual optimization procedure.

The most striking result of these plots, however, is that objective function f_2 fluctuates about 100x more than objective function f_1 . As a result, one expects f_2 to be less smooth than f_1 , which is visualized in figures 5.6 and 5.7.

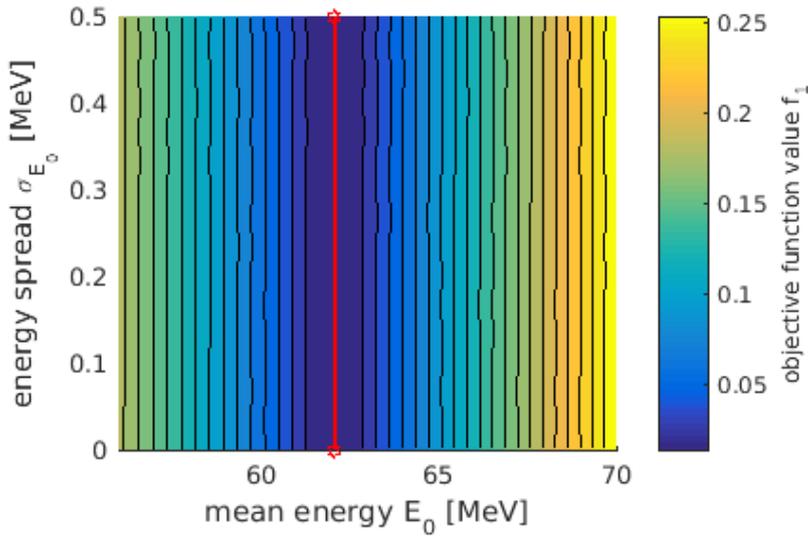
5.3.3. Solver selection and settings

In order to minimize the objective functions defined previously, appropriate MATLAB solvers had to be chosen that converge to a solution of the problems considering the properties of f_1 and f_2 .

Given the fact that f_1 varies smoothly and rather poses a non-stochastic objective function, it was decided to use the *fminsearch* solver for the first optimization

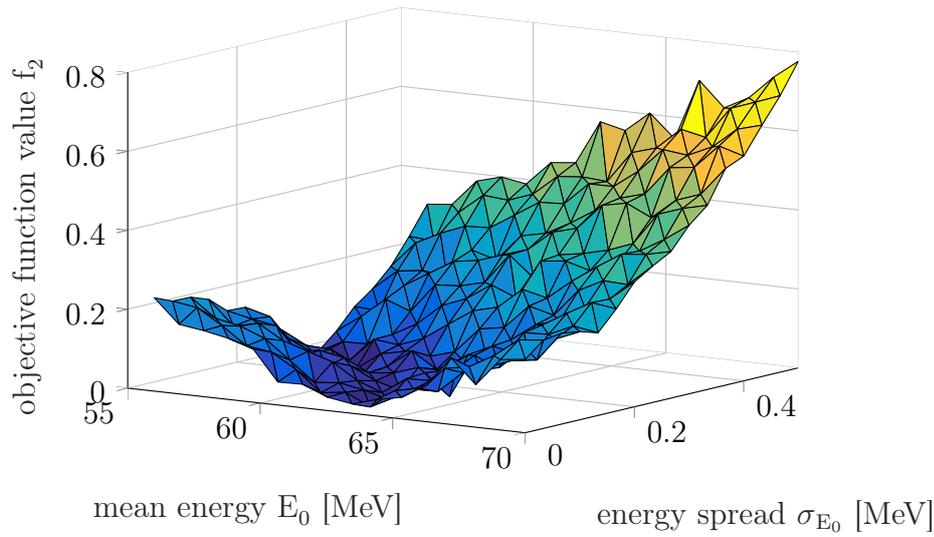


(a) 3D plot

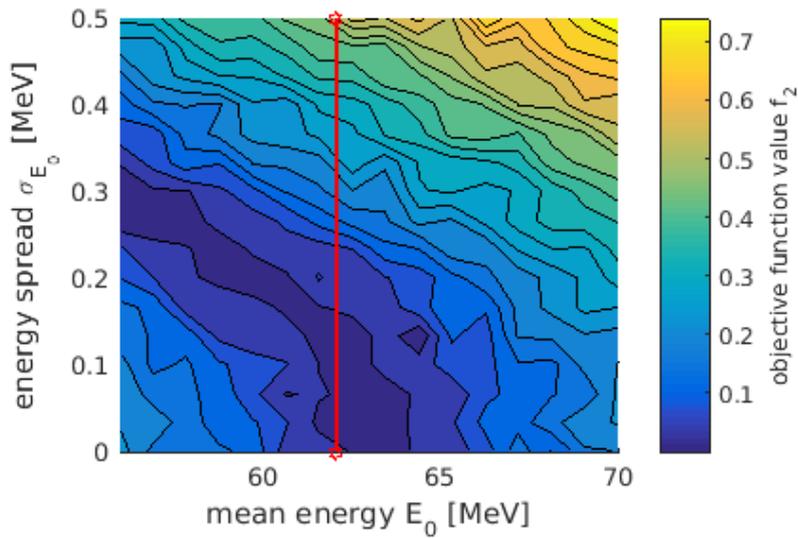


(b) Contour plot

Figure 5.6.: Representations of f_1 for the situation with nozzle and 62.4 MeV, evaluated on a regular grid in 2D parameterspace using 15000 primaries. The red line in the contour plot indicates the optimal mean energy E_0^* . The plots also support the fact that R_{80} does not depend on σ_{E_0} since varying σ_{E_0} does not lead to a change in f_1 . 15000 primaries are sufficient for 62.4 MeV as figure 5.5 shows: the fluctuations in f_2 cannot be reduced significantly even if the number of primaries is increased disproportionately.



(a) 3D plot



(b) Contour plot

Figure 5.7.: Representations of f_2 for the situation with nozzle and 62.4 MeV, evaluated on a regular grid in 2D parameterspace using 15000 primaries. The red line in the contour plot indicates the optimal mean energy E_0^* along which σ_{E_0} has to be minimized.

step of E_0 . Since f_1 has only one local minimum for varying E_0 and fixed σ_{E_0} ⁸, fminsearch cannot run into the problem of finding the wrong minimum.

⁸which is consequently at the same time the global minimum

For all E_0 -optimization runs an initial value of $\sigma_{E_0} = 0.2$ MeV was chosen. Regarding the termination criteria for `fminsearch`, the 'TolFun'-flag was set to 10^{-5} which is about the value for the standard deviation of f_1 . This means the solver stops when the change in objective function values f_1 between two subsequent iterations drops below 10^{-5} .

As opposed to f_1 , f_2 appears to be a rather non-smooth, stochastic objective function. Therefore, a stochastic optimization method is appropriate and it was decided to use MATLAB's *genetic algorithm* solver to minimize this objective function. Here the optimal mean energy value E_0^* obtained in the first step was used as an input parameter for the MC simulations of this second optimization step for σ_{E_0} . An initial value of $\sigma_{E_0} = 0.2$ MeV was passed to the solver as initial parameter to start the algorithm. Lower and upper bounds of $l = 0.0001$ MeV and $u = 0.3$ MeV were used for σ_{E_0} . The termination criteria 'TolFun' was set to 10^{-3} which is around the value for the standard deviation of f_2 . The 'Generation'-flag was set to 70 such that the solver stops after having created 70 generations. The 'PopulationSize'-flag was set to 10, which means that every generation consists of a population of 10 individuals.

Other than that the default settings of MATLAB 8.6 were used. For the situations both with and without nozzle the same solver settings were used.

5.3.4. Implementation of optimization tool

In figure 5.8 a chart flow is presented that summarizes and illustrates the workflow of the beam energy optimization tool and how the self-developed MATLAB scripts, Linux shell scripts and GATE codes are connected to each other. The computationally most time-consuming single part are the GATE simulations which account for the majority of total optimization runtime. The source code for the MATLAB scripts `optimizer.i.m` and `ObjFun.i.m` as well as the GATE source codes can be found in the appendix (see D).

5.3.5. Energy optimization

The two-step optimization procedure described before was repeated for the 5 key energies for the situation with nozzle and for the 3 key energies for the situation without nozzle. In both situations, `fminsearch` and genetic algorithm solvers converged toward a solution of minimal objective function value f_1 or f_2 within the underlying simulation uncertainties.

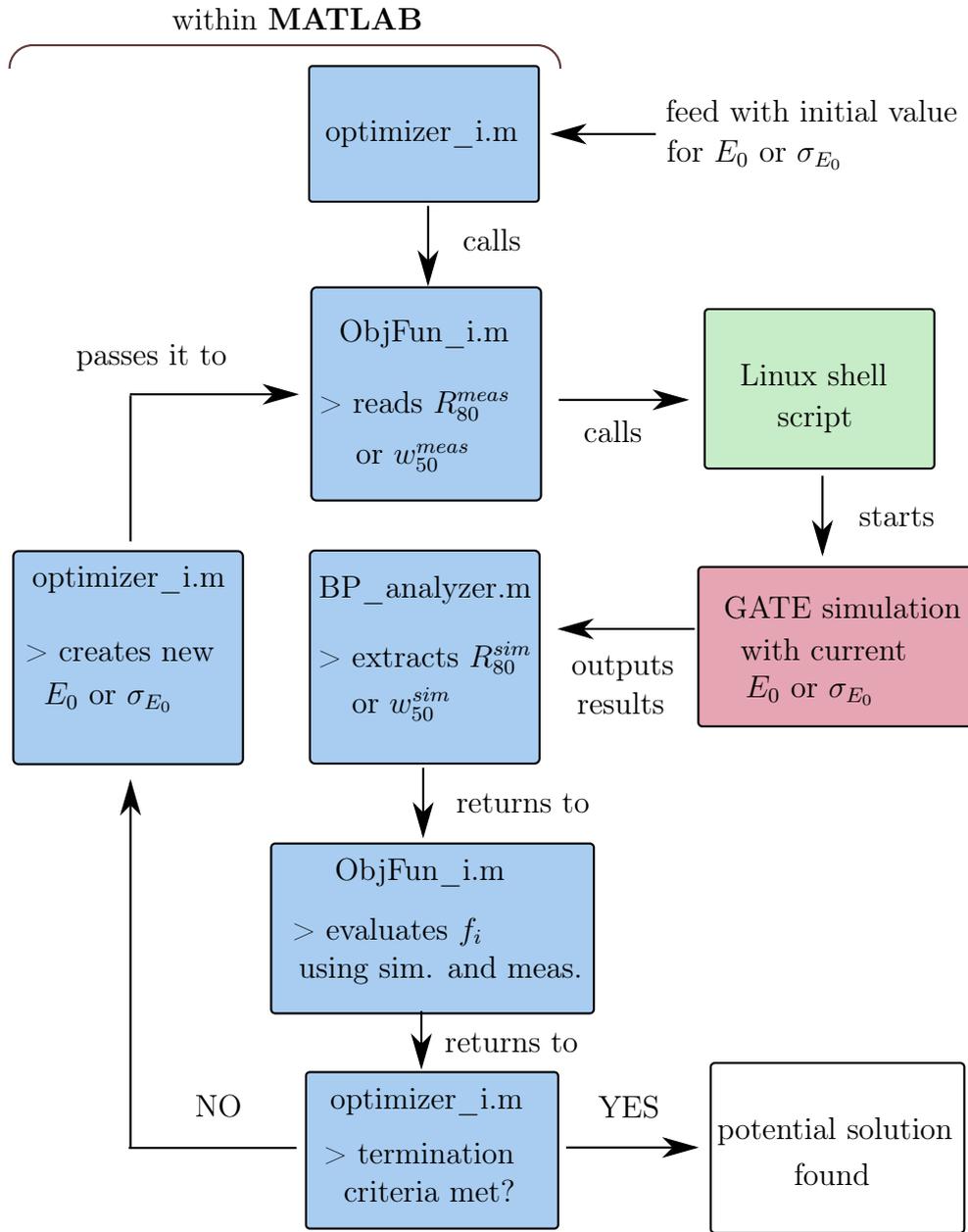


Figure 5.8.: Flowchart illustrating the working principle of the beam energy optimization tool. The subscript i can be either $i=1$ for E_0 -optimization or $i=2$ for σ_{E_0} -optimization.

The convergence process of `fminsearch` and genetic algorithm for the situation with nozzle and 62.4 MeV is illustrated in figure 5.9, respectively. When look-

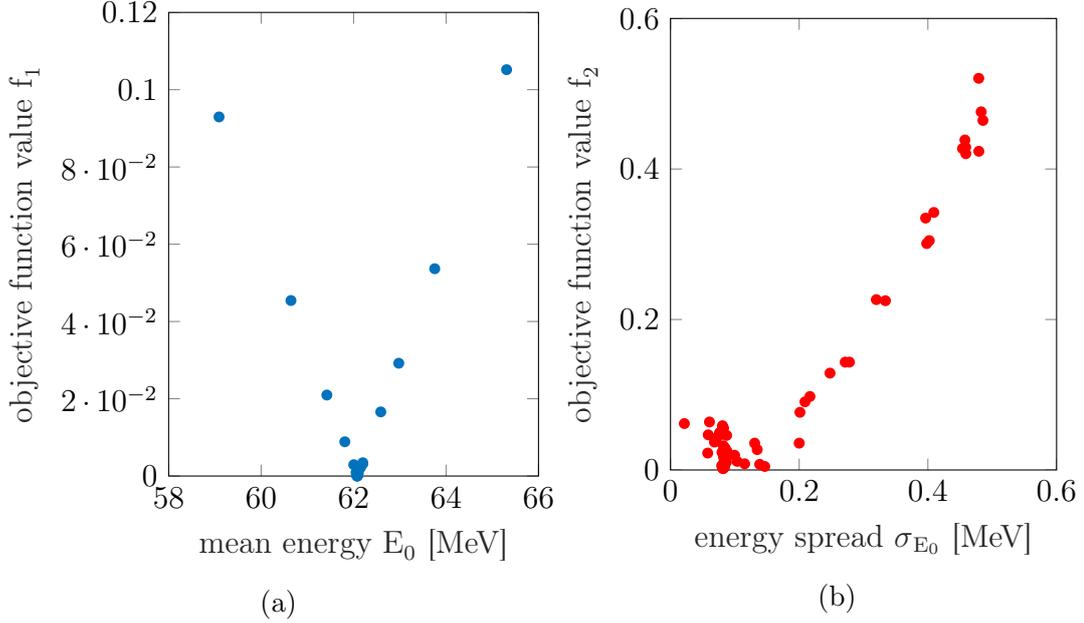


Figure 5.9.: Scatter plots showing the single iteration steps for the fminsearch-solver (a) and the ga-solver (b) for objective functions f_1 and f_2 .

ing at plot (b) one can observe that the objective function evaluations for energy spreads around 0.1 MeV are rather scattered and do not exhibit a precise, point-like minimum as it is the case for f_1 in plot (a). This phenomenon results from two factors: on the one hand from the simulation noise (which is about $100x$ higher for f_2 than for f_1) and on the other hand from the fact that a stochastic solver was used for optimization.

5.4. Beam model

5.4.1. Uncertainties

In order to quantify the fluctuations of the obtained optimized beam parameter $\sigma_{E_0}^*$ the σ_{E_0} -optimization procedure was repeated $m = 10$ times (only for the situation with nozzle). The best parameter $\sigma_{E_0, optimal}^*$ among this set of 10 optimized parameters $\sigma_{E_0}^*$ was chosen to be the one lowest objective function value f_2 . The fluctuations were quantified in terms of standard deviation of the set of $\sigma_{E_0}^*$.

The obtained optimal parameters E_0^* and $\sigma_{E_0, optimal}^*$ for the situation *with nozzle* are summarized in table 5.4 for each key energy. The table also includes the associated objective function values and the uncertainties of the optimized en-

ergy spreads. For 252.7 MeV no standard deviation was calculated, because the obtained optimal energy spread $\sigma_{E_0, optimal}^*$ for this key energy is not meaningful (see subsection 5.4.3). The optimal beam parameters and the associated objective function values for the situation *without nozzle* are summarized in table 5.5.

Energy parameters for nozzle model						
E_{sys} [MeV]	optimal parameters		objective functions		rel. uncertainties std($\sigma_{E_0}^*$) [%]	
	E_0^* [MeV]	$\sigma_{E_0, optimal}^*$ [MeV]	$f_1(E_0^*)$	$f_2(\sigma_{E_0, optimal}^*)$		
62.4	62.07	0.082	0	$1.78 \cdot 10^{-3}$	21.9	
97.4	96.78	0.061	$6.53 \cdot 10^{-4}$	$6.53 \cdot 10^{-4}$	45.9	
148.2	147.24	0.139	0	$3.17 \cdot 10^{-4}$	38.1	
198.0	196.77	0.058	$4.00 \cdot 10^{-6}$	$3.43 \cdot 10^{-4}$	129	
252.7	250.99	0.084	$2.63 \cdot 10^{-6}$	$2.63 \cdot 10^{-2}$	-	

Table 5.4.: Summary of the beam energy parameters for the nozzle model, including parameters, objective function values and uncertainties.

Energy parameters for vacuum model					
E_{sys} [MeV]	optimal parameters		objective functions		
	E_0^* [MeV]	$\sigma_{E_0, optimal}^*$ [MeV]	$f_1(E_0^*)$	$f_2(\sigma_{E_0, optimal}^*)$	
62.4	62.05	0.054	$1.10 \cdot 10^{-16}$	$7.30 \cdot 10^{-4}$	
148.2	147.21	0.109	0	$4.78 \cdot 10^{-3}$	
198	196.66	0.056	$3.98 \cdot 10^{-6}$	$1.18 \cdot 10^{-2}$	

Table 5.5.: Summary of the beam energy parameters for the vacuum model, including objective function values.

As both tables 5.4 and 5.5 display, the minimized objective function values altogether range from 0 to 10^{-4} for f_1 whereas the minimized values for f_2 range from 10^{-4} to 10^{-2} . As shown in 5.3.2, the simulation uncertainties for f_1 is about 10^{-4} and for f_2 about 10^{-2} . This suggests that the simulation uncertainties constitute a lower bound for the minimized objective function values.

Another remarkable result are the rather high relative uncertainties for the minimized energy spreads $\sigma_{E_0}^*$ for the nozzle model. They range from about 22% for 62.4 MeV up to 129% for 198 MeV and are a direct result from the relatively high simulation uncertainties of f_2 .

5.4.2. Model comparison

The optimal parameters from tables 5.4 and 5.5 are plotted as function of system energy in figure 5.10a and 5.10c for the nozzle model and the vacuum model. Uncertainties for the optimized energy spread were added as errorbars in figure 5.10c. In order to estimate the parameters E_0 between the key energies, a polynomial of order 2 was chosen for approximation. For the energy spread σ_{E_0} a polynomial of order 4 was interpolated. Additionally, the IR3 nozzle model has been added to all plots for comparison. Since in figure 5.10a no significant deviation of all three models with the system energy can be observed, the difference between system energy and obtained mean energy $E_{sys} - E_0$ has been plotted as a function of E_{sys} for better visualization. Plot (a) shows an overlap of E_0 for auto-optimized nozzle, auto-optimized vacuum model and IR3 nozzle model with no significant deviations. Plot (b) reveals maximal deviations of 0.1 MeV between IR3 nozzle model and the auto-optimized nozzle model. For the energy spread, plot (c) shows that the auto-optimized models are only a shifted-down version of the IR3 nozzle model with a shift of about 0.12 MeV. The points for the vacuum model lie mostly within the error bars for the nozzle model.

The relatively good overlap between the beam models obtained with and without nozzle suggests that the implementation of the MedAustron nozzle in GATE is accurately enough for beam energy modeling.

5.4.3. Model validation

The obtained nozzle model was validated on the measurements for the situation with nozzle (see figure 5.12). In 5.12a the measured Bragg curve ranges R_{80}^{meas} and R_{50}^{meas} are compared to the respective simulated ranges R_{80}^{sim} and R_{80}^{sim} that resulted from a GATE simulation with TPS-PBS using the obtained nozzle model as source input. Furthermore, the absolute differences and percentual differences between measured and simulated widths are illustrated in figure 5.12b and 5.12c, respectively. The same sort of plots for the comparisons between widths w_{80} and w_{50} are visualized in figure 5.13.

The plot (a) in figure 5.12 shows an overlap of measured and simulated ranges without observable deviations. The difference plots (b) and (c) in the same figure

exhibit range differences of up to 0.2 mm in absolute terms and up to 0.1% in percentual terms over the whole clinical energy range.

Higher deviations are revealed by the Bragg peak width validation plots in figure 5.13. In plot (a) a rather high w_{50} deviation for $E_{sys} = 252.7$ MeV can be observed. Plot (b) shows that this w_{50} deviation exceeds 2 mm, which is the highest absolute deviation. In comparison, the differences for other key energies are closely concentrated around 0. Plot (c) quantifies the w_{50} deviation at $E_{sys} = 252.7$ MeV to about 6%.

In order to further investigate the relatively high w_{50} deviation, a Bragg peak was simulated caused by a *monoenergetic*⁹ proton beam stopping in water for the situation at 252.7 MeV. The simulated Bragg peak was then compared to the measured one. The results are depicted in figure 5.11. It shows that even for a simulated monoenergetic proton beam the simulated 50%-width w_{50}^{sim} is already larger than the measured 50%-width w_{50}^{meas} . As derived in 2.5.6 the width w_{50} , however, depends not only on the energy spread σ_{E_0} , but also on the extent of energy straggling that the beam experiences. Consequently, two possible explanations for this result exist: either GATE overestimates energy straggling for this situation, or measurements for this key energy are not as accurate and mask a slightly wider Bragg peak. Either way, this result implies that the obtained nozzle model presented above is not applicable to the area around 252.7 MeV.

⁹i.e. with zero energy spread $\sigma_{E_0} = 0$ MeV

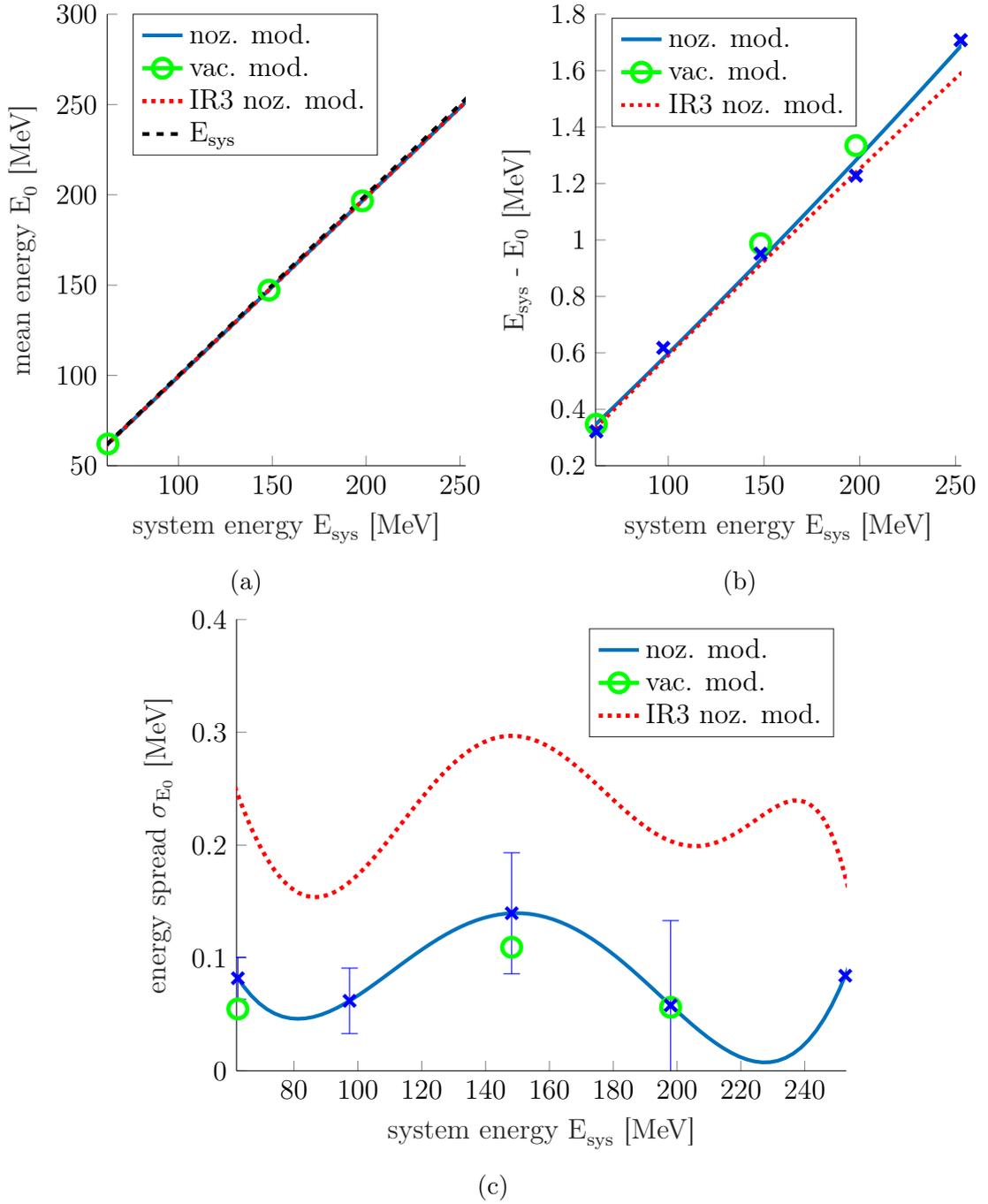


Figure 5.10.: Plots displaying the auto-optimized nozzle model (noz. mod.), the auto-optimized vacuum model (vac. mod.) and the IR3 nozzle model (IR3 noz. mod.). In subfigure (a) the curve for " E_{sys} " resembles the linear function $E_0(E_{sys}) = E_{sys}$ with slope 1.

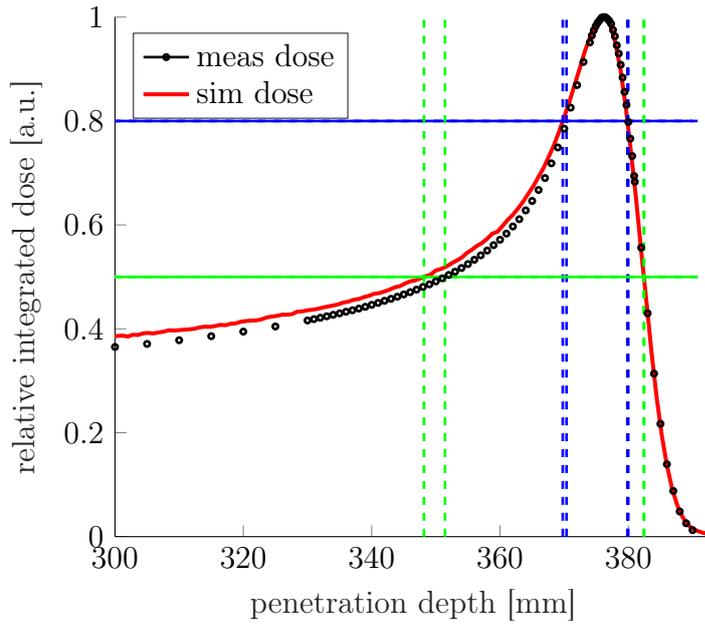


Figure 5.11.: Comparison of simulated and measured Bragg peak for the situation with nozzle and 252.7 MeV. The energy spread for the simulation was set to zero and the mean energy to the optimal value E_0^* obtained from the nozzle model. The number of primaries was set to 120,000 in order to guarantee minimal simulation uncertainties. The vertical, blue dashed lines show the proximal and distal 80%-depths, the green dashed lines the proximal and distal 50%-depths for simulation and measured Bragg peak each.

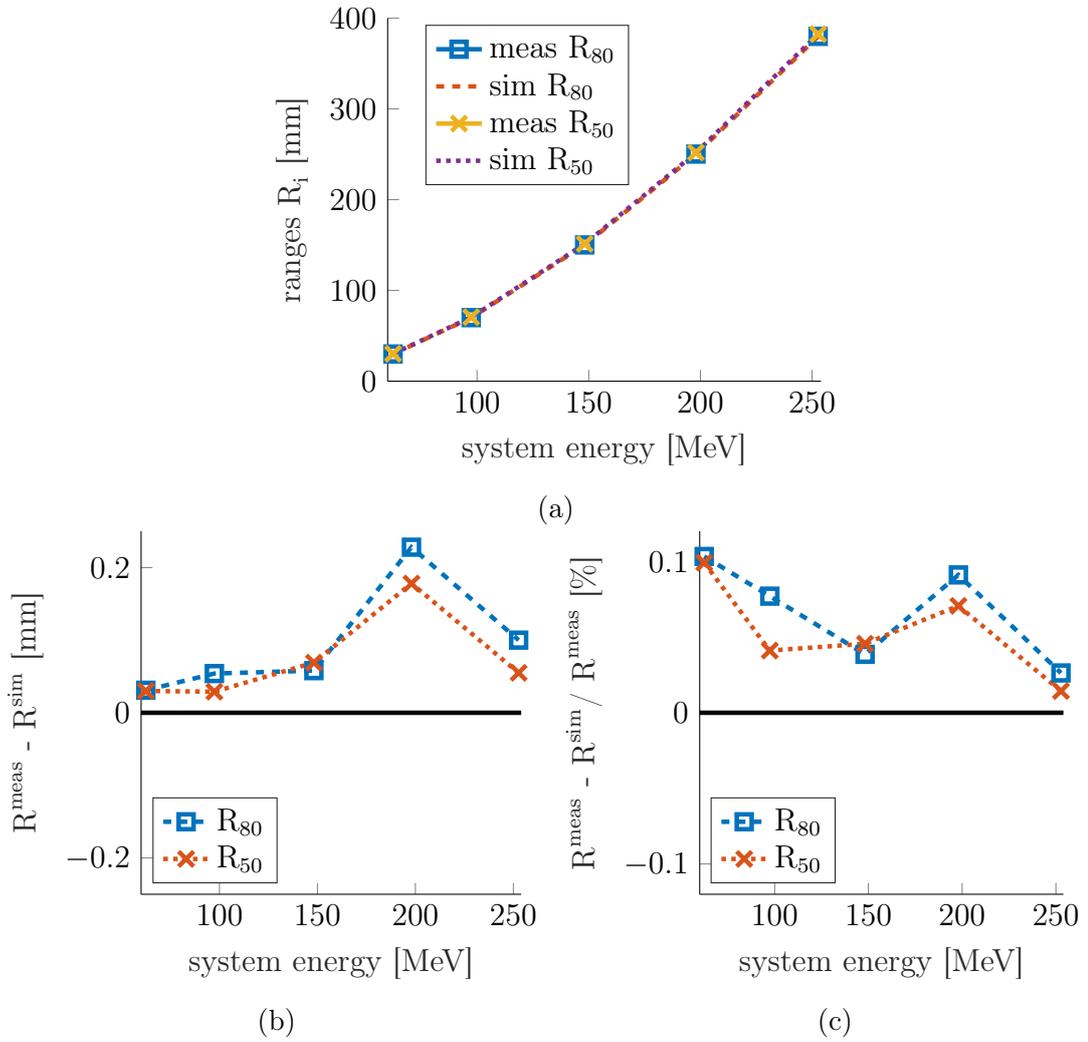


Figure 5.12.: Range validation plots for the nozzle model.

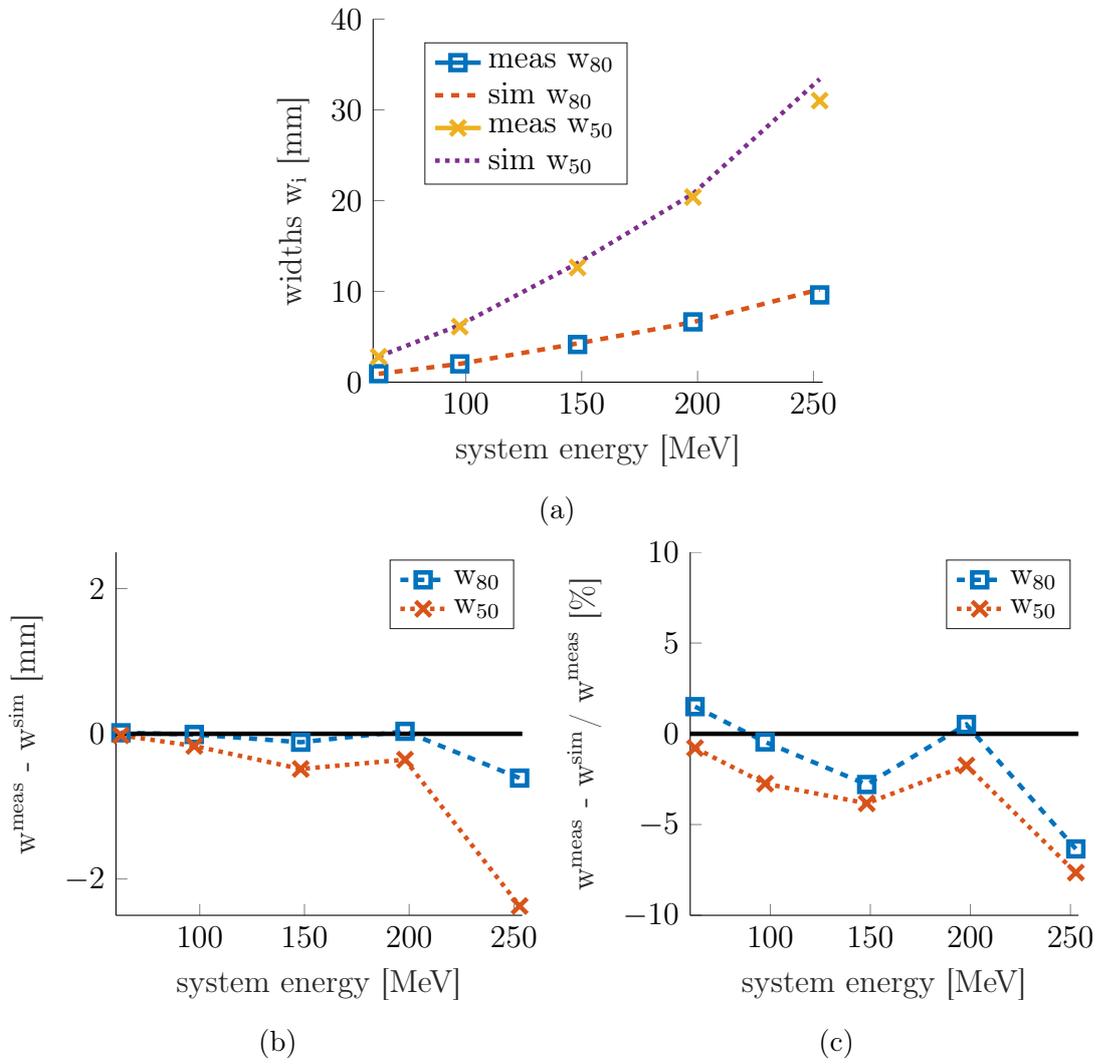


Figure 5.13.: Width validation plots for the nozzle model.

6. Optical beam modeling

This section describes how the optical beam parameters ($\sigma_x, \theta_x, \epsilon_x, \sigma_y, \theta_y, \epsilon_y$) of the proton beam were determined. The procedure is in many ways similar to the beam energy modeling procedure described in the previous chapter. The optical beam parameters were determined for the 5 key energies (62.4 MeV, 97.4 MeV, 148.2 MeV, 198.0 MeV and 252.7 MeV). In order to estimate the parameters for energies in between the key energies, polynomial interpolation was used.

As response quantity FWHM-vs-ISD curves for the x- and y-direction were chosen. These curves describe the spot size of the beam in terms of FWHM¹ as a function of isocenter distance (ISD).

The optical parameters were optimized for three different configurations, that differ in nozzle configuration and source point setting:

1. with nozzle in the source point before the vacuum windows at (0|0|-130 cm)²
2. without nozzle in the source point before the vacuum windows at (0|0|-130 cm)
3. with nozzle in a source point shifted downstream after the DDS boxes at (0|0|-82 cm)

In the following, these configurations are referred to as *nozzle model*, *vacuum model* and *in-nozzle model*, respectively. The source point for nozzle model and vacuum model is still located in the vacuum pipe upstream of the vacuum windows. In contrast, the source point of the in-nozzle model is located *in the nozzle* behind ITS and DDS boxes but before the ripple filters (see also figure 6.4 for the position of source points). Thus, the in-nozzle model allows to study the effects of scatterers on the optical beam parameters.

For each key energy the FWHMs in x- and y-direction were measured for a specific set of ISDs. Afterward, the measurements were modeled in GATE to obtain the simulated curves. The optimization procedure minimized the difference between

¹full width at half maximum of the lateral beam profile in the respective transverse direction

²this source point will in the following also be referred to as *nominal* source point

measured and simulated FWHM differences along the set of ISDs by varying the optical beam parameters. In order to simplify the optimization procedure, x- and y-directions were simulated and optimized separately. This is possible because the x- and y-profiles of the proton beam can be considered independently. Consequently, the six optical beam parameters can be divided into the two parameter sets $(\sigma_x, \theta_x, \epsilon_x)$ and $(\sigma_y, \theta_y, \epsilon_y)$ which are totally independent from each other. Therefore, optical beam modeling consists of two independent steps: optimizing the parameters for x *and* optimizing them for y, for which the order does not matter.

The optimization problems were solved using MATLAB. The working principle of the optical parameter optimization tool is identical to the tool for beam energy optimization (see figure 5.8), except that different solvers, objective functions and analysis tools were used. The resulting optimal beam parameters $(\sigma_x^*, \theta_x^*, \epsilon_x^*)$ and $(\sigma_y^*, \theta_y^*, \epsilon_y^*)$ were then validated on the measurements.

The methods outlined above will be now explained in more detail.

6.1. Measurements

All measurements were already acquired before this thesis was started. In order to measure the FWHM-vs-ISD curve for a specific key energy, a proton beam of this key energy was directed to the Lynx PT. The active area of the Lynx PT was positioned at a certain ISD such that the beam impinged perpendicular on it (see figure 6.1), acquiring 2D lateral beam profiles.

From these lateral beam profiles the x-FWHM and y-FWHM were extracted using the in-house developed software tool "Lynx QA". The extraction mechanism³ is the same as the one implemented in the *batchPhSEval.C*-script which was used to analyze the simulation results (see appendix D for the C-script and section 6.2 for simulations)

The resulting FWHM-vs-ISD curves for all 5 key energies are depicted in figures 6.2 and 6.3. For measurements without nozzle, FWHMs for ISDs from -118 cm to 40 cm could be acquired. For the situation with nozzle only measurements at ISDs from -57 cm to 40 cm were possible. The measurements without nozzle indicate

³For FWHM extraction, histograms displaying the number of detected particles as a function of x- and -y position were created with a binsize of 0.5 mm (coinciding with the Lynx PT resolution). The histograms were then fitted with Gaussian distributions for which the FWHM was extracted.

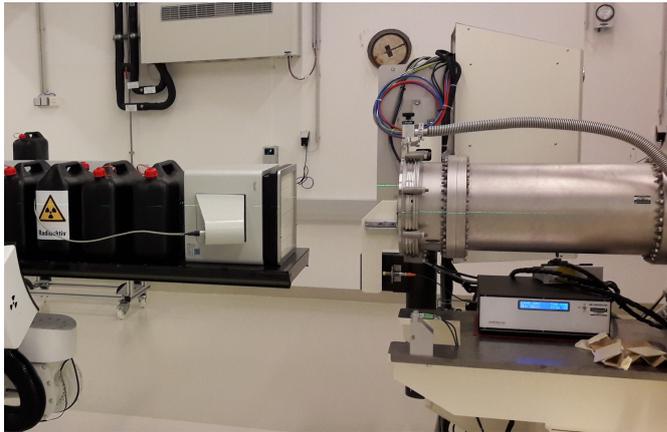


Figure 6.1.: Picture of a FWHM-vs-ISD measurement without nozzle. The vacuum pipe with removed nozzle can be seen on the right-hand side, the Lynx PT on the left-hand side. Water tanks were put behind the Lynx in order to stop the proton beam after the measurement.

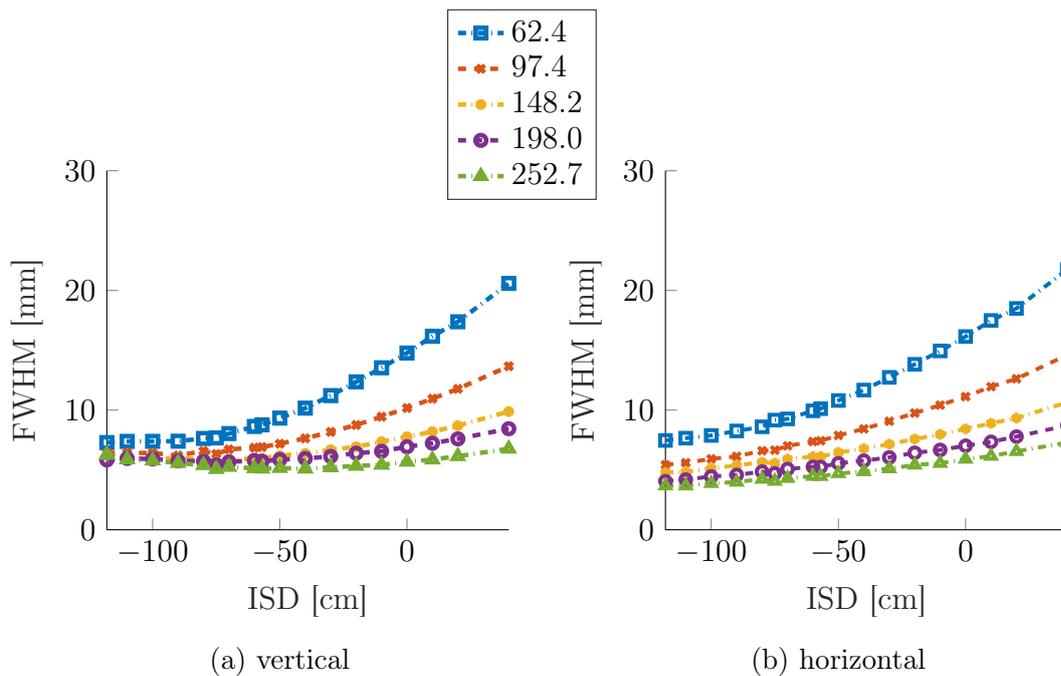


Figure 6.2.: Measured FWHM-vs-ISD curves without nozzle.

that the beam is convergent in the vertical plane and divergent in the horizontal plane for the nominal source point at $ISD = -130$ cm. For measurements including the nozzle, the vertical direction changes from a converging to a diverging behavior

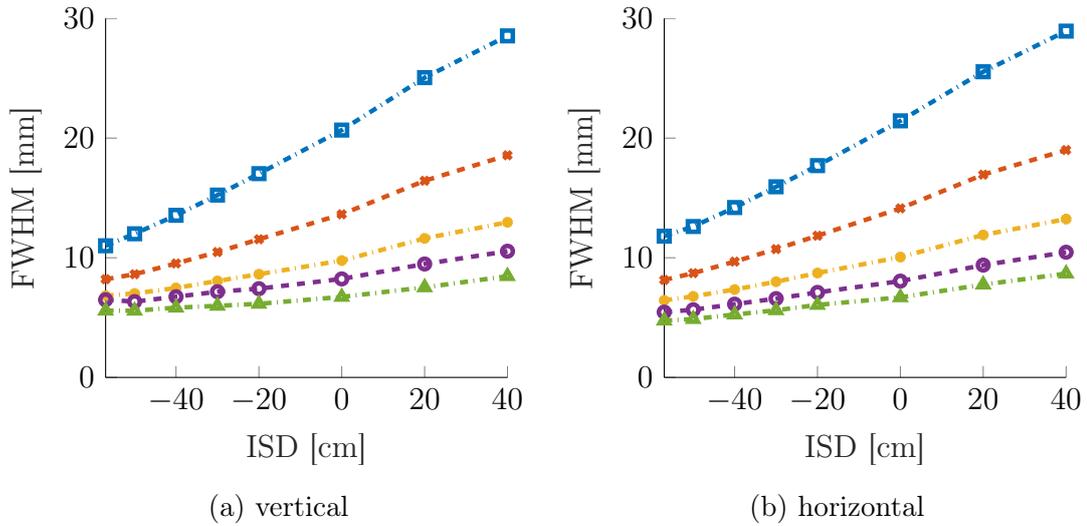


Figure 6.3.: Measured FWHM-vs-ISD curves with nozzle.

due to the additional scattering introduced by the nozzle elements.

6.2. MC simulations

For all simulations related to optical beam modeling, GATE v8.0 with GEANT4 in version 10.03.p01 was used.

The simulation setups to model the measurement setups for the three different configurations are shown in figure 6.4. The used coordinate system was set such that the x-direction coincides with the *vertical* direction and the y-direction with the *horizontal* direction.

The isocenter was defined as the origin of the cartesian coordinate system at (0|0|0), coinciding with the real-world isocenter. Depending on the configuration, the PBS was located at (0|0|-130 cm) or (0|0|-82 cm). The protons were moving along +z-direction.

Phase-space actors were used to measure the lateral beam profiles as a function of ISD. For each phase-space actor, a geometry called "PhS_x"⁴ was created (details can be found in table 6.1). The dimensions were chosen to coincide with those of the Lynx PT. The geometries were positioned such, that the proximal plane

⁴x denotes the associated ISD, e.g. PhS_-57 contains the phase-space actor that determines the FWHMs at ISD = -57 cm.

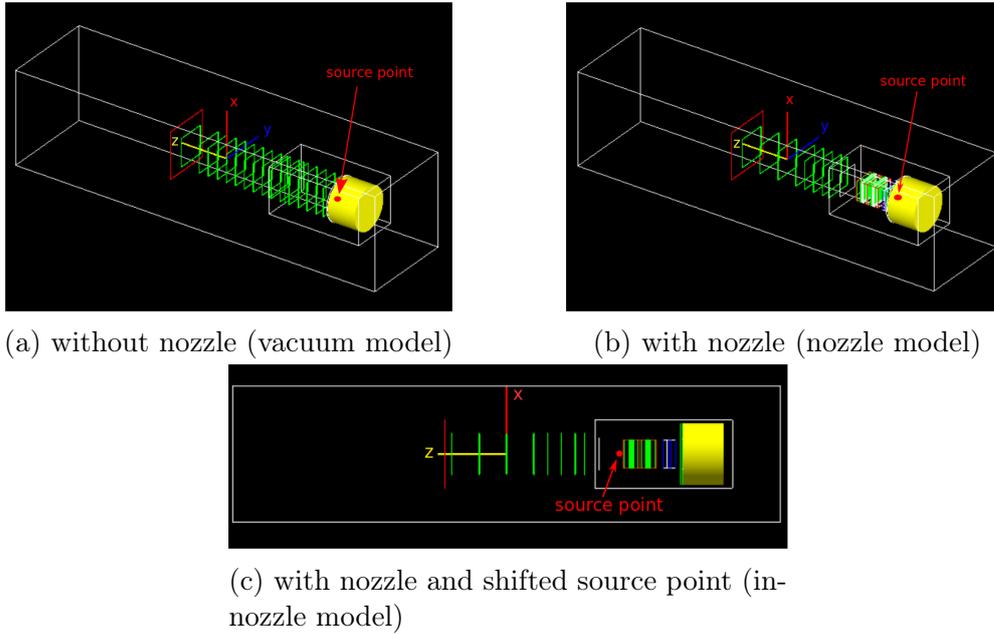


Figure 6.4.: GATE simulation setups for the three different configurations. The green quadratic planes depict the geometries for the phase-space actors. The red quadratic plate at the most downstream position depicts the geometry for the kill actor.

of each geometry was perpendicular to the z -axis and situated at the corresponding measurement positions. The phase-space actors, attached to these geometries, recorded the lateral (x, y) position of each proton entering the geometry. The resulting lateral beam profiles were stored as .root-files.

Using the script *batchPhSEval.C* the FWHM was extracted from the simulated beam profiles. The script can be found in the D. It copies the respective particle positions from the PhS-files into a 1D histogram with 600 bins and a binsize of 0.5 mm, coinciding with the dimension and resolution of the Lynx PT. A 1D Gaussian distribution is then fitted to the histogram from which the FWHM can be extracted.

In order to increase simulation speed, particles were removed from the simulation once they traversed all phase-space actors. This was achieved by a kill actor which was positioned behind the last phase-space actor geometry at $ISD = 45$ cm.

All simulation parameters are summarized in table 5.2 where also the used values for maximal step size and production cuts are listed.

geometry	dimensions [cm ³]	material	step size [cm]	cuts [m]	attached actor
world	100x100x400	air	5	10	-
NozzleBox	-	various	5	10	-
PhS_x	29.882x29.882 x0.5	air	5	10	PhaseSpaceActor
killVolume	50x50x5	air	5	10	killActor

Table 6.1.: Summary of simulation parameters. The values displayed for the maximal step sizes are valid for protons and the production cut values are valid for electrons, positrons and gammas.

6.3. Automated optimization

6.3.1. Objective function and constraints

The optimal beam parameters should minimize the difference between simulated and measured FWHMs along the whole beam path. Therefore, the objective function f_3 for optical beam modeling should contain the FWHM differences at *all* available ISDs. This can be achieved by summing up the squared relative differences over all ISDs in order to obtain f_3

$$f_3(\sigma, \theta, \epsilon) = \sum_{i \in \text{ISD}} w_i \cdot \left(\frac{\text{FWHM}_i^{\text{sim}}(\sigma, \theta, \epsilon) - \text{FWHM}_i^{\text{meas}}}{\text{FWHM}_i^{\text{meas}}} \right)^2 \quad (6.1)$$

Here i is the summation index which runs over all available ISDs. The objective function f_3 was used for both x- and y-optimization, respectively.

Additionally, weighting factors w_i were introduced to compensate for unequal spacing between the available ISDs. Neighboring ISDs that exhibit a distance of 10 cm were assigned weighting factors of $w_i = 1$. For ISDs that were closer together than 10 cm, the weighting factor was set to a value smaller than 1, and vice versa. The weighting factors were calculated using the formula

$$w_i = \frac{(\text{ISD}_i - \text{ISD}_{i+1})[\text{cm}]}{10 \text{ cm}} \quad (6.2)$$

Table 6.2 summarizes all available ISDs and their corresponding weighting factors for the situations without and with nozzle.

As already mentioned in section 4.3, for optical beam modeling there are con-

i	without nozzle		with nozzle	
	ISD_i	w_i	ISD_i	w_i
1	-118	0.8	-57	0.7
2	-110	1	-50	1
3	-100	1	-40	1
4	-90	1	-30	1
5	-80	1	-20	1
6	-75	0.5	0	2
7	-70	1	20	2
8	-60	1	40	2
9	-57	0.3		
10	-50	1		
11	-40	1		
12	-30	1		
13	-20	1		
14	-10	1		
15	0	1		
16	10	1		
17	20	1		
18	40	2		

Table 6.2.: Summary of available ISDs and corresponding weighting factors for the situations without and with nozzle.

straints that have to be taken into account while optimizing f_3 . The definition of emittance (see 2.4.2) implies the two inequality constraints

$$\epsilon_x \leq \pi \cdot |\theta_x| \cdot \sigma_x \quad \text{and} \quad \epsilon_y \leq \pi \cdot |\theta_y| \cdot \sigma_y \quad (6.3)$$

which were directly implemented into the objective function *ObjFun_3.m* within MATLAB. The source code of the objective function can be found in appendix D. Additional obvious constraints are $\sigma \geq 0$ and $\epsilon \geq 0$.

Finding the optimal optical beam parameters $(\sigma^*, \theta^*, \epsilon^*)$ corresponds to a global

minimum search in a 3D parameter space

$$(\sigma^*, \theta^*, \epsilon^*) = \min_{(\sigma, \theta, \epsilon)} f_3(\sigma, \theta, \epsilon) \quad (6.4)$$

6.3.2. Simulation uncertainties

Simulation uncertainties can be reduced by increasing the number of primary particles. However, this also increases the simulation runtime and thus the overall time required for optimization. In order to find an appropriate trade-off between accuracy and runtime, the uncertainty in FWHM for a specific parameter set $(\sigma_j, \theta_j, \epsilon_j)$ was examined as a function of the number of primaries. Figure 6.5 shows the standard deviation of the simulated FWHM for several ISDs using a sample size of $n=100$ and a system energy $E_{sys} = 62.4$ MeV. The figure shows

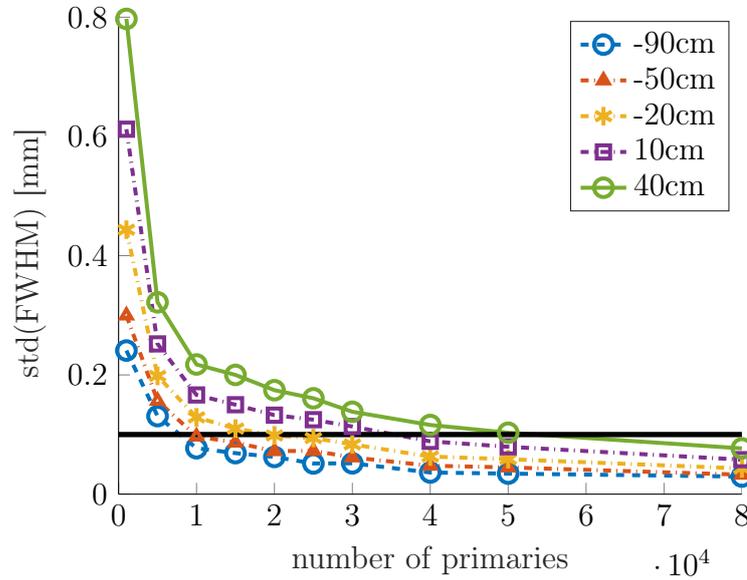


Figure 6.5.: Standard deviations (abbr. as $\text{std}(\dots)$) of FWHM for several ISDs using $E_{sys} = 62.4$ MeV and a sample size of $n = 100$. The black line shows the maximal expected uncertainty of 0.1 mm if the number of primaries are set to 50,000 for each simulation.

simulation uncertainties that decrease with the square root of the number of primaries as expected (see section 2.6). The absolute FWHM uncertainties increase with increasing ISDs since the absolute FWHM also increases with larger ISDs. For all remaining key energies the absolute uncertainties can be expected to be lower than the ones displayed for 62.4 MeV because the FWHM of the beam shrinks with increasing system energy due to adiabatic damping. Consequently, the curve

for $ISD = 40$ cm in figure 6.5 can be considered as an upper bound for the absolute FWHM uncertainty in the clinical energy range.

Considering runtime and FWHM uncertainties, 50,000 primaries were chosen for the simulations *regardless* of system energy. As a consequence of figure 6.5, this implies FWHM uncertainties of about 0.1 mm. In order to reduce simulation uncertainties and smoothen the objective function, the simulated FWHM values were rounded to a precision of 0.1 mm before they were used to calculate the objective function value. The rounding procedure is implemented in the source code of `ObjFun_3.m` (see appendix D). Using such a rounding procedure implies f_3 -uncertainties around $1 \cdot 10^{-3}$ for 50,000 primaries, which is depicted in figure 6.6.

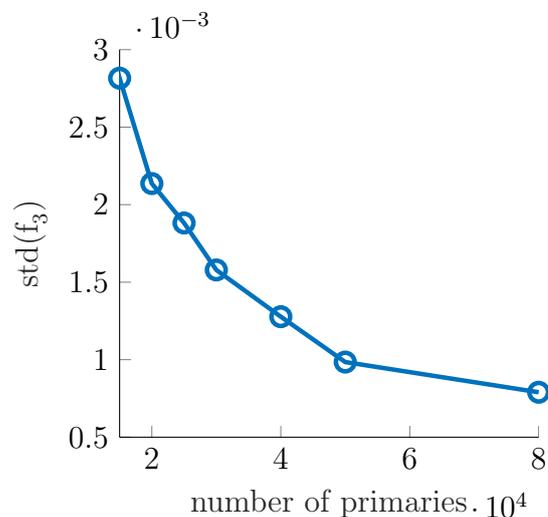


Figure 6.6.: Standard deviations of objective function evaluations f_3 for a certain point $(\sigma_j, \theta_j, \epsilon_j)$ in 3D parameter space, using $E_{sys} = 62.4$ MeV, a sample size of $n = 100$ and rounding of FWHMs to 0.1 mm; plotted over number of primaries.

6.3.3. Solver selection and settings

In order to choose an appropriate solver for the optimization procedure, the behavior of the objective function f_3 was examined regarding smoothness and the existence of multiple local minima. Figures 6.7 and 6.8 show objective function evaluations of f_3 for three different situations: without nozzle, with nozzle (but no passive elements) and with nozzle including the range shifter.

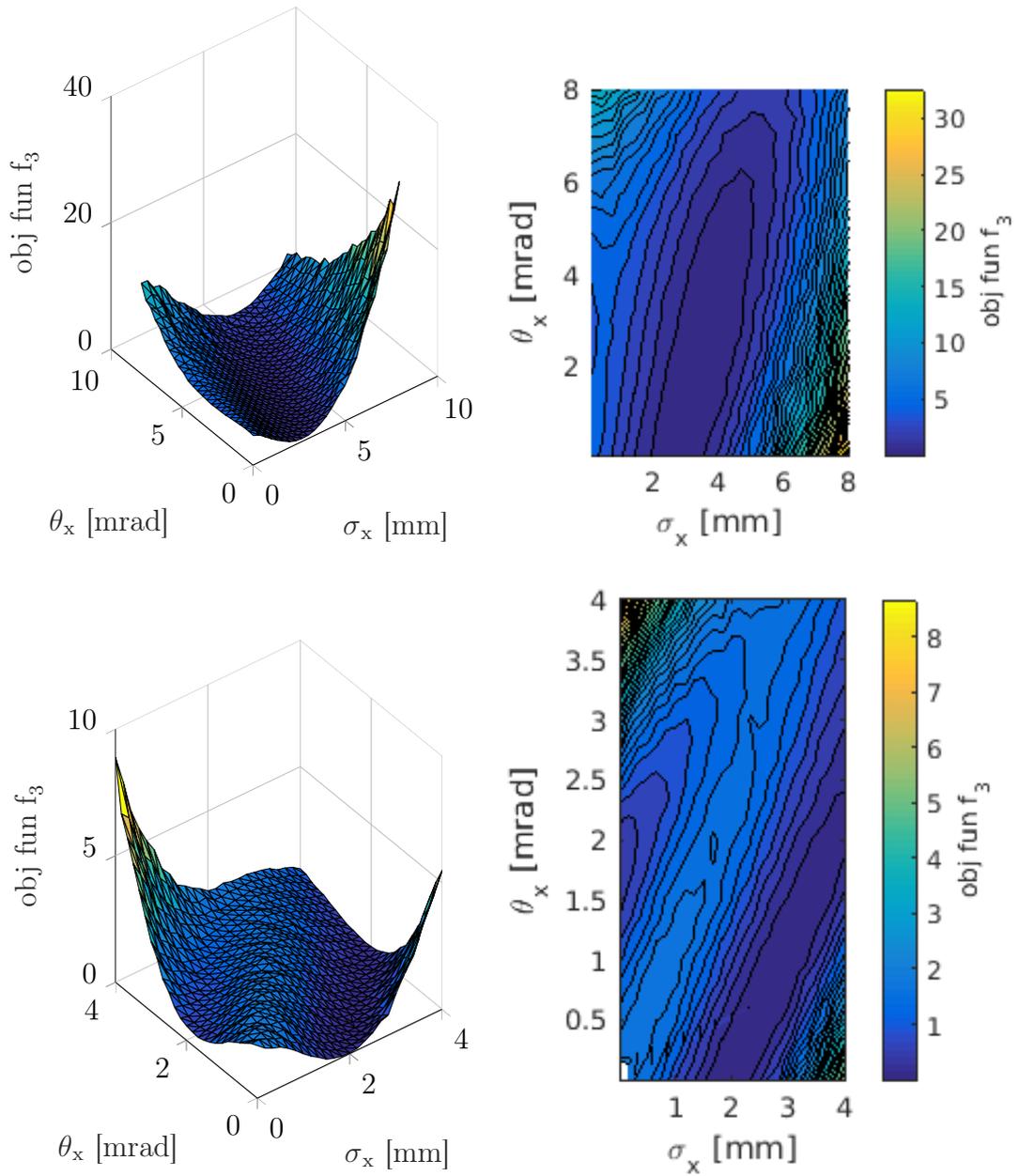


Figure 6.7.: 3D plots (left) and contour plots (right) showing the evaluated obj. fun. f_3 for 2 different configurations using $E_{sys} = 62.4$ MeV, $\epsilon_x = 3$ mm*mrad and 50,000 primaries: the 1st row was obtained for the situation *without nozzle*, the 2nd row for the situation *with nozzle*.

The figures show two interesting characteristics. First, the simulation uncertainty

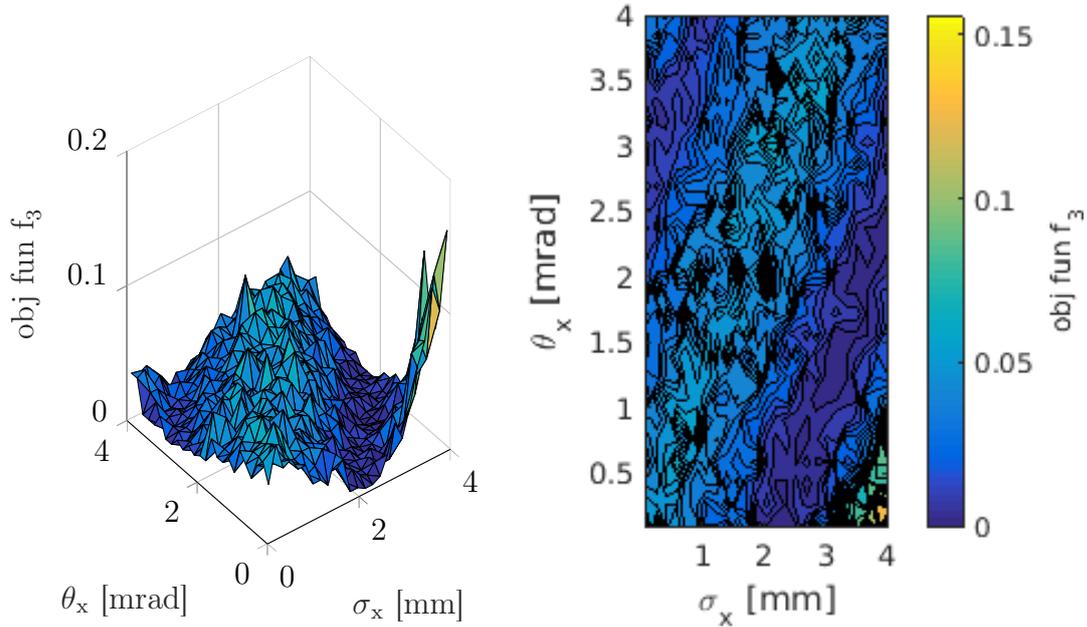


Figure 6.8.: 3D plot (left) and contour plot (right) showing the evaluated obj. fun. f_3 for the situation with nozzle with *range shifter* included; using $E_{sys} = 62.4$ MeV, $\epsilon_x = 3$ mm*mrad and 50,000 primaries.

increases the more scattering material is introduced in the beam path. For the situation without nozzle f_3 appears to be relatively smooth, whereas for the situation with nozzle and range shifter f_3 looks rather like a stochastic objective function. The major reason for this behavior are the absolute FWHM values for a particular ISD which increase with an increasing number of scatterers in the beam path. If the absolute FWHM values increase, also the FWHM uncertainty increases and consequently the absolute uncertainties in f_3 do as well. Second, the objective function forms a new local minimum if scattering material is added. For the situation without nozzle there exists only one local minimum which coincides with the global minimum. When the nozzle is added, a second local minimum is formed for spot sizes $\sigma_x < 1$ and a divergence of around $\theta_x = 2$ mrad. This local minimum gets shifted towards higher θ_x when the range shifter is additionally included.

Since the objective function to minimize may experience stochastic fluctuations and as shown above exhibits multiple local minima, MATLAB's *genetic algorithm*-solver was used for optimizing the optical beam parameters. The optimization source code *optimizer_3.m* can be found in appendix D. The used solver setting regarding initial values, lower and upper bounds can be found in table 6.3. The ter-

mination criteria 'TolFun' was set to 10^{-3} which is around the maximal standard deviation for f_3 . The 'Generation'-flag was set to 70 and the 'PopulationSize'-flag to 10. Other than that the default settings of MATLAB 8.6 were used. For all three modeled configurations⁵ the same solver settings were used.

	σ [mm]	θ [mrad]	ϵ [mm · mrad]
initial value	3	2	5
lower bounds	0	0	0
upper bounds	10	10	20

Table 6.3.: Summary of solver settings for both x- and y-direction.

6.4. Beam model

6.4.1. Uncertainties

Using the solver settings specified before, the optical beam parameters were optimized for 5 key energies for three different configurations, each in both planes. Since a stochastic solver was used for optimizing a possibly stochastic objective function, the optimized beam parameters vary for each optimization run. In order to get a better understanding of these fluctuations, the optimization procedure was repeated $m = 10$ times which yielded 10 different sets of optimized parameters $(\sigma^*, \theta^*, \epsilon^*)$ for each case. The fluctuations of these were quantified in terms of standard deviation⁶. The best parameter set $(\sigma^*, \theta^*, \epsilon^*)_{optimal}$ was chosen to be the one which had the lowest objective function value f_3 .

The results for the nozzle model are displayed in table 6.4. The results for the vacuum model and the in-nozzle model can be found in appendix A.

6.4.2. Model comparison

The parameters of the three optimized models are plotted as a function of system energy in figures 6.9 and 6.10. The uncertainties for the vacuum model parameters were added as errorbars. In order to estimate the parameters in between key

⁵i.e. for vacuum model, nozzle model and in-nozzle model, cf. figure 6.4

⁶This is of course only a preliminary assessment since the sample size $m = 10$ might be too low to obtain good statistics.

Optical nozzle model				
E_{sys} [MeV]	optimal parameters			objective function
	σ_x [mm]	θ_x [mrad]	ϵ_x [mm · mrad]	$f_3(\sigma_x, \theta_x, \epsilon_x)$ [10^{-3}]
x-axis				
62.4	-3.75	1.99	4.45	0.6
97.4	-3.74	2.01	4.78	1.1
148.2	-3.25	1.50	7.13	2.0
198.0	-3.03	1.30	7.62	2.1
252.7	-2.65	1.01	4.97	0.4
y-axis				
62.4	2.00	0.84	1.28	3.0
97.4	1.89	0.48	0.76	1.4
148.2	1.61	0.59	1.61	2.5
198.0	1.57	0.51	2.13	1.2
252.7	1.51	0.23	0.15	0.8

Table 6.4.: Summary of the optical nozzle model

energies, a 4th order polynomial was chosen for each parameter and each model. These interpolation polynomials are required by the GATE PBS-TPS source in order to calculate treatment plans. The polynomials are also displayed in figures 6.9 and 6.10.

Additionally, a manually optimized model⁷ for the situation with nozzle that already existed for irradiation room 1 *before* this thesis was started was added to the plots for comparison.

Furthermore, the predictions of beam parameters in the nominal source point at (0|0|-130 cm) from the design-phase of MedAustron were added to the plots (IR1 predictions). These predictions were originally only available in the form of Twiss parameters including emittance, and thus were transformed into the optical beam parameters spot size, divergence and emittance. The corresponding calculations are explained in detail in appendix C.

With regard to spot size σ , figure 6.9 shows that the auto-optimized models

⁷this is referred to as *preliminary manual nozzle model* (prel. man. nozzle model) in the following

exhibit the expected physical behavior as function of E_{sys} : the spot sizes of both planes decrease with increasing E_{sys} as it is also displayed by the IR1 predictions for the x-axis. The IR1 predictions for the y-axis (horizontal direction) are constant lines which is due to the "bar of charge"-behavior of the slow resonant extracted beam (see C for details). Since the Twiss formalism cannot be directly applied for such a situation, the IR1 predictions for the horizontal direction in figures 6.9 and 6.10 should be considered with caution.

Regarding divergence θ , the optimized vacuum and nozzle model get relatively close to the IR1 predictions for divergence. The high deviations of the in-nozzle model can be explained by taking into account that this model was optimized for a source point located *in* the nozzle after the DDS boxes, while the IR1 predictions show the parameters for the nominal source point in the vacuum pipe *before* the nozzle. Hence, the beam has already undergone scattering in the nozzle and therefore displays a magnified divergence. Also, the divergence of the in-nozzle model is positive throughout the clinical energy range, which means the beam diverges at the source point *in* the nozzle. This is in contrast to the nozzle and vacuum model which display negative x-divergence, i.e. the beam converges in x-direction at the nominal source point before the nozzle. These results are confirmed by the FWHM-vs-ISD measurements for both the situations with and without nozzle (cf. 6.1).

A novel feature of the nozzle model is that it uses a converging beam in x-direction at the nominal source point while the preliminary nozzle model uses a diverging beam. FWHM-vs-ISD measurements without nozzle showed (cf. section 6.1), that the beam is indeed converging in x-direction at the nominal source point. This implies that considering only "with nozzle"-information when developing a beam model might lead to a different assumption about the divergence of the beam before the nozzle.

The highest uncertainties for beam parameters were obtained for the emittance ϵ , as the relative large errorbars in 6.10 show. Due to the large errorbars, one could also model the emittances of vacuum and nozzle model using constant functions around $\epsilon_x = 6 \text{ mm} \cdot \text{mrad}$ and $\epsilon_y = 2 \text{ mm} \cdot \text{mrad}$. The majority of the polynomials of the auto-optimized models do not display a physically meaningful behavior which would be a decreasing emittance for an increasing system energy (see IR1 prediction for x-emittance in figure 6.9). The only exception here is the y-emittance of the in-nozzle model which shows such a behavior.

Another remarkable feature of the in-nozzle model is the magnified y-emittance

in comparison to the y-emittance of vacuum and nozzle model. As theory in subsection 2.4.4 showed, a beam experiences emittance growth when undergoing a series of scatterers. Since the in-nozzle model describes a beam that has already traversed the nozzle and thus experienced scattering, it displays a larger emittance in comparison to nozzle and vacuum model.

In order to examine the reason for the rather large errorbars in figure 6.10, the objective function f_3 was evaluated for varying emittance. Figure 6.11 shows that f_3 does not respond strongly to variations in emittance as opposed to variations in spot size or divergence. As a consequence, the stochastic genetic algorithm-solver experiences larger fluctuations in the optimized emittance values ϵ^* which in turn explains the large errorbars.

6.4.3. Model validation

For model validation the measured FWHMs were compared to the corresponding simulated FWHMs. The latter were obtained by using GATE's pencil beam-TPS source and the respective optical model as input.

Figures 6.12 and 6.13 show the validation of the *vacuum model*. Simulations were performed without nozzle and compared to the corresponding measurements. The graphs show absolute deviations ranging from +0.2 mm to -0.8 mm for both axis, which is well within the clinical acceptable range⁸. The percentual deviations go from +4% to -6% which is also clinically acceptable.

Figure 6.14 shows the validation of the *nozzle model*. The graphs show that most absolute deviations lie within +0.5 mm and -0.75 mm, while all percentual deviations range from 4% to -5% for both axis. This is for both cases within the clinical acceptable range. A rather large deviation can be observed for the y-FWHM difference for 62.4 MeV which reaches up to -1.4 mm at ISD = 40 cm. The reason for this rather large deviation are either inaccuracies for multiple Coulomb scattering in GATE or uncertainties in the FWHM measurements at ISD = 40 cm. A similar behavior at this ISD is also exhibited by the absolute differences in the case for the vacuum model for the 62.4 MeV-curve (see figures 6.12 and 6.13).

Furthermore, as a cross check the vacuum model was compared to the measurements with nozzle. This was only possible because a fully detailed implementation of the MedAustron nozzle existed in GATE. Figure 6.15 illustrates this kind of

⁸The clinical acceptable range used at MedAustron is either ± 1 mm in absolute deviations or $\pm 10\%$ in percentual deviations, depending on which constraint is more stringent.

"cross-validation". Simulations were performed using the vacuum model as source input and the nozzle for the simulation setup. The graphs show absolute deviations from +0.4 mm to -0.7 mm for most ISDs, again with larger deviations at $ISD = 40$ cm exceeding 1 mm. The percentual deviations range from +6% to about -7%. Figure 6.15 is thus a validation of the nozzle implementation in GATE.

The remaining validation for the in-nozzle model can be found in appendix B. Its absolute and percentual deviations also lie within the clinically acceptable range.

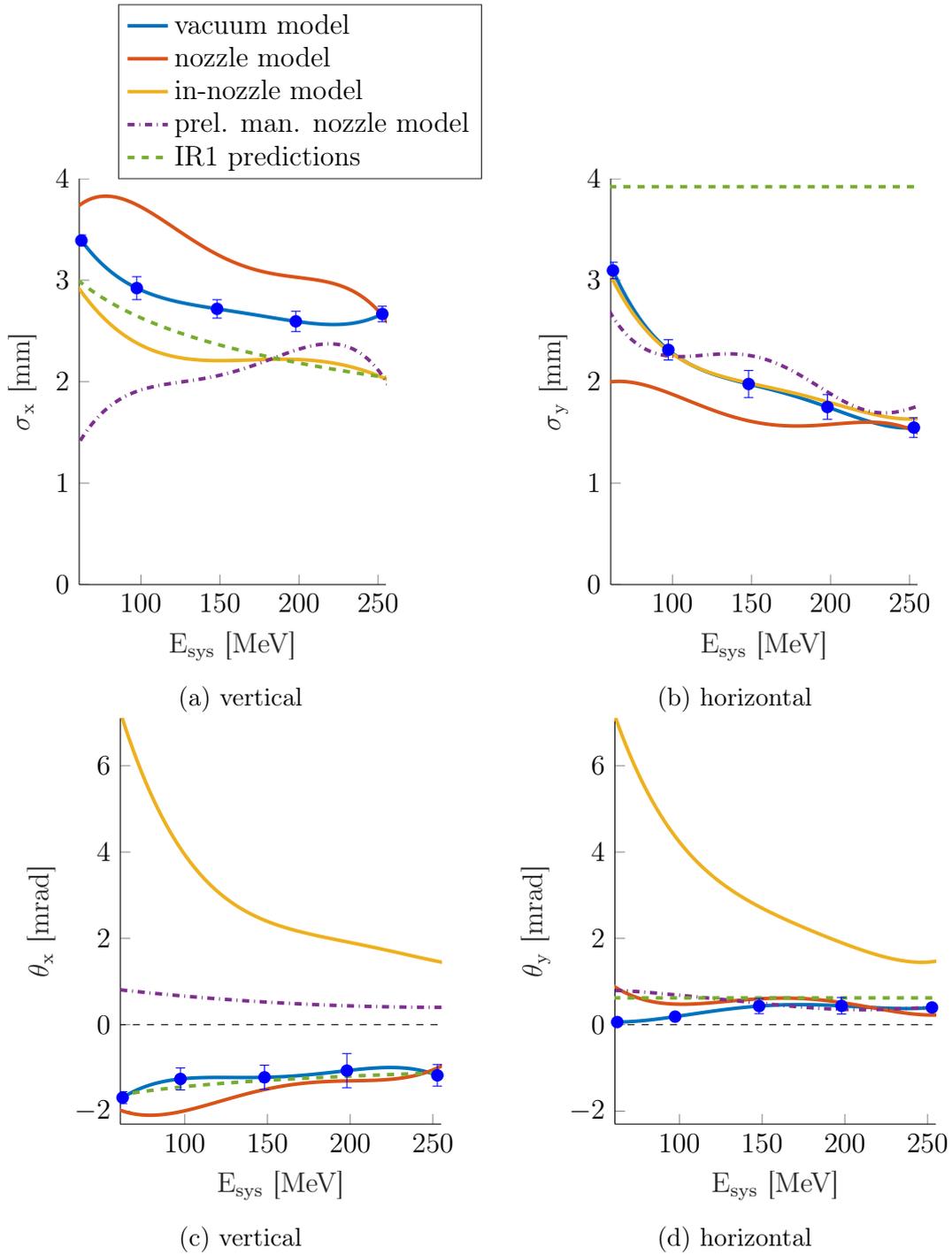


Figure 6.9.: Comparison of optimized optical beam models with the preliminary manual nozzle model and IR1 predictions, showing spot size and divergence. The in-nozzle model is not directly comparable to the other models since it was optimized for a different source point.

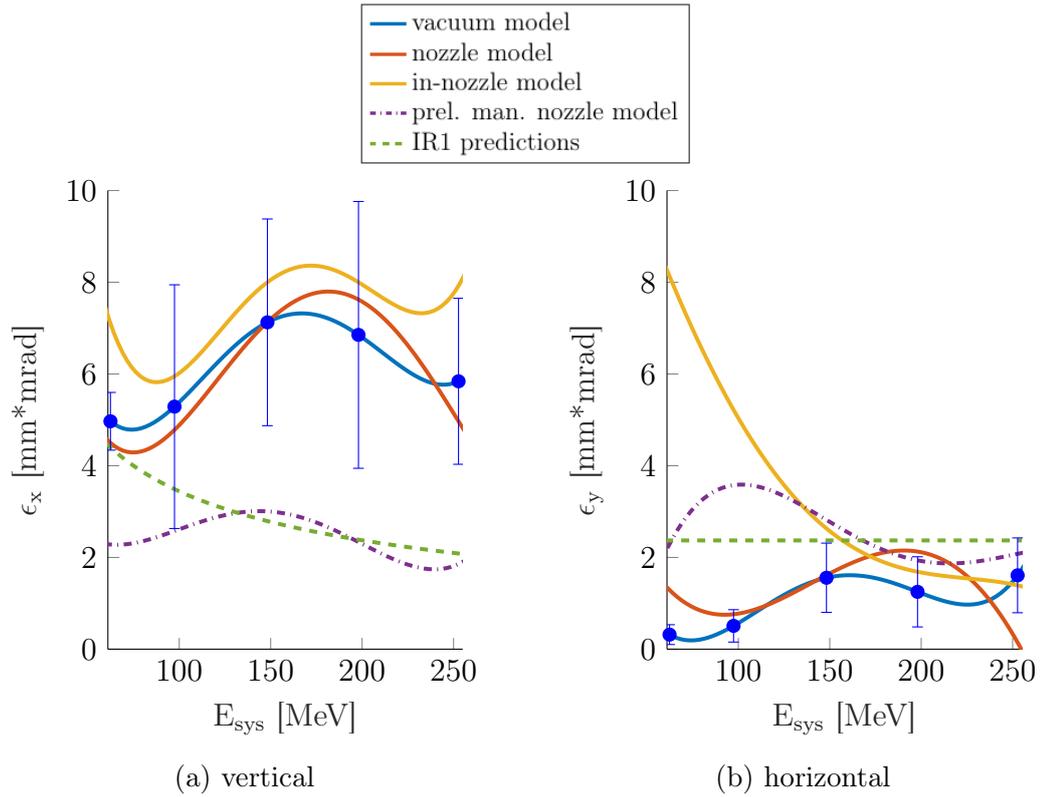


Figure 6.10.: Comparison of optimized optical beam models with the preliminary manual nozzle model and IR1 predictions, showing emittance. The in-nozzle model is not directly comparable to the other models since it was optimized for a different source point.

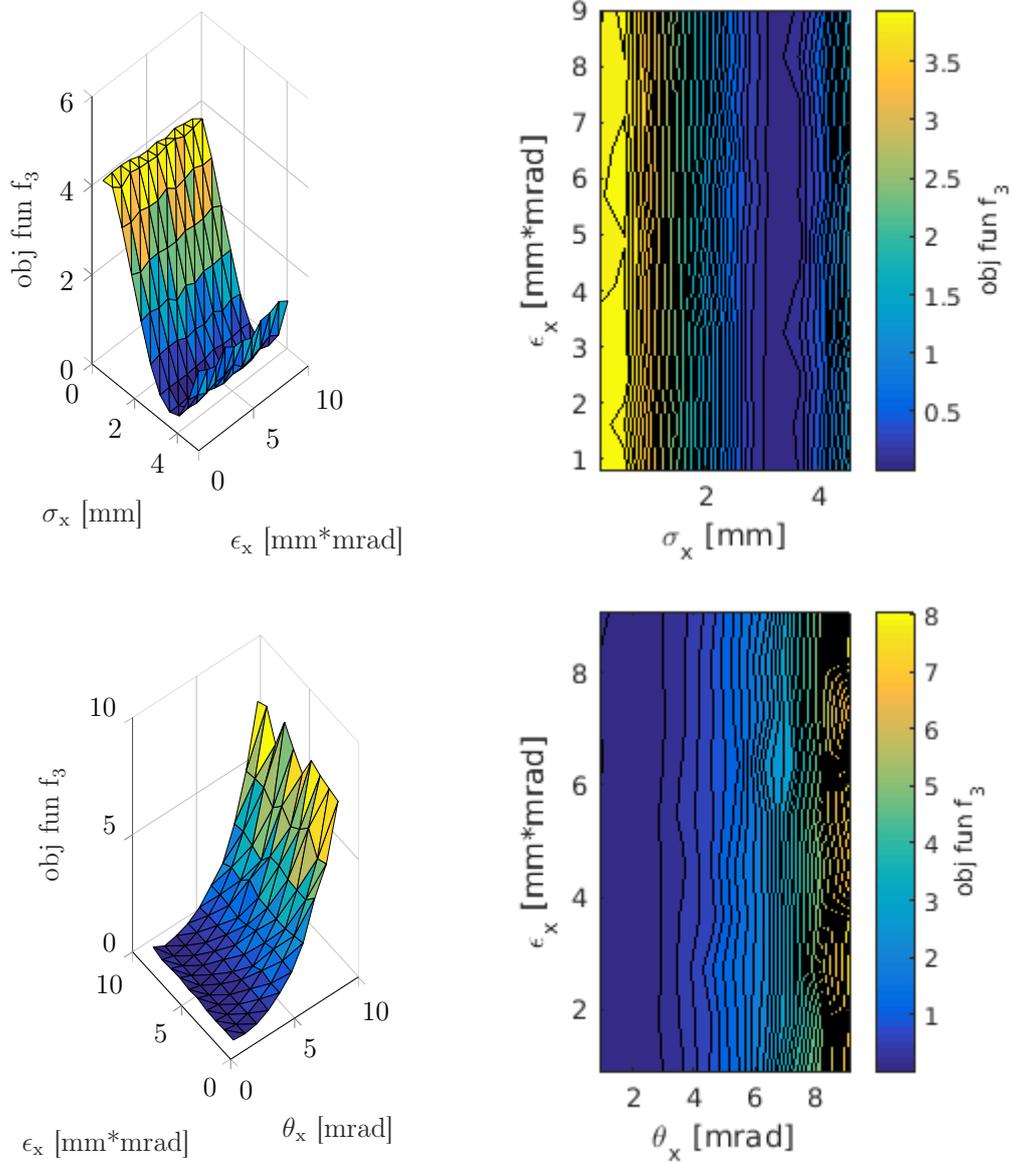


Figure 6.11.: 3D plots (left) and contour plots (right) showing the evaluated obj. fun. f_3 as a function of (σ_x, ϵ_x) and (θ_x, ϵ_x) for the situation without nozzle using $E_{sys} = 62.4$ MeV and 50,000 primaries: for the plots in the 1st row $\theta_x = 1.6$ mrad and for the plots in the 2nd row $\sigma_x = 3.3$ mm was used.

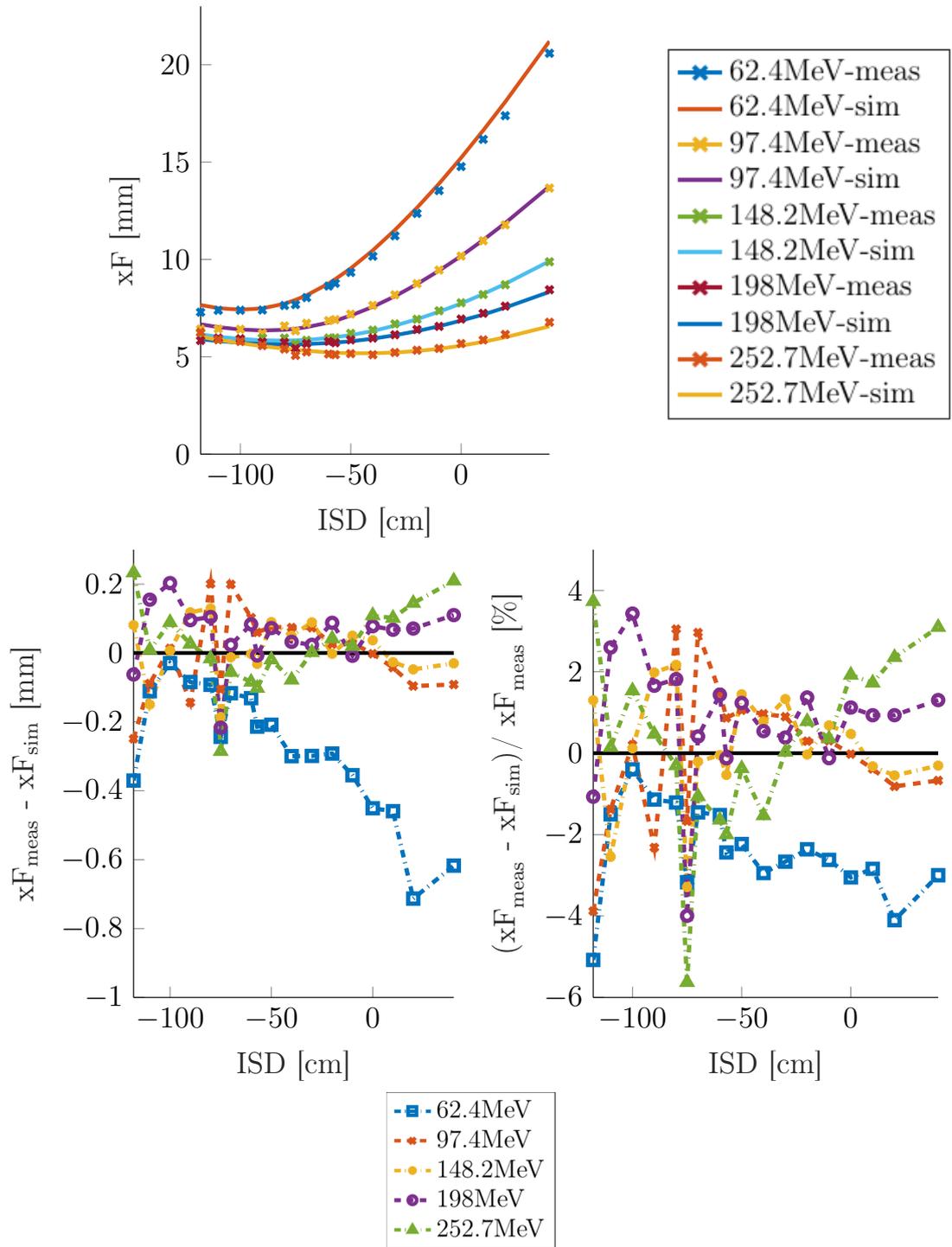


Figure 6.12.: Plots displaying the validation of the vacuum model (using no nozzle for simulation) on the measurements without nozzle for the vertical plane. The 1st row shows the comparison of measured and simulated FWHMs, the 2nd row the absolute difference, the 3rd row the percentual differences. F is the abbreviation for FWHM.

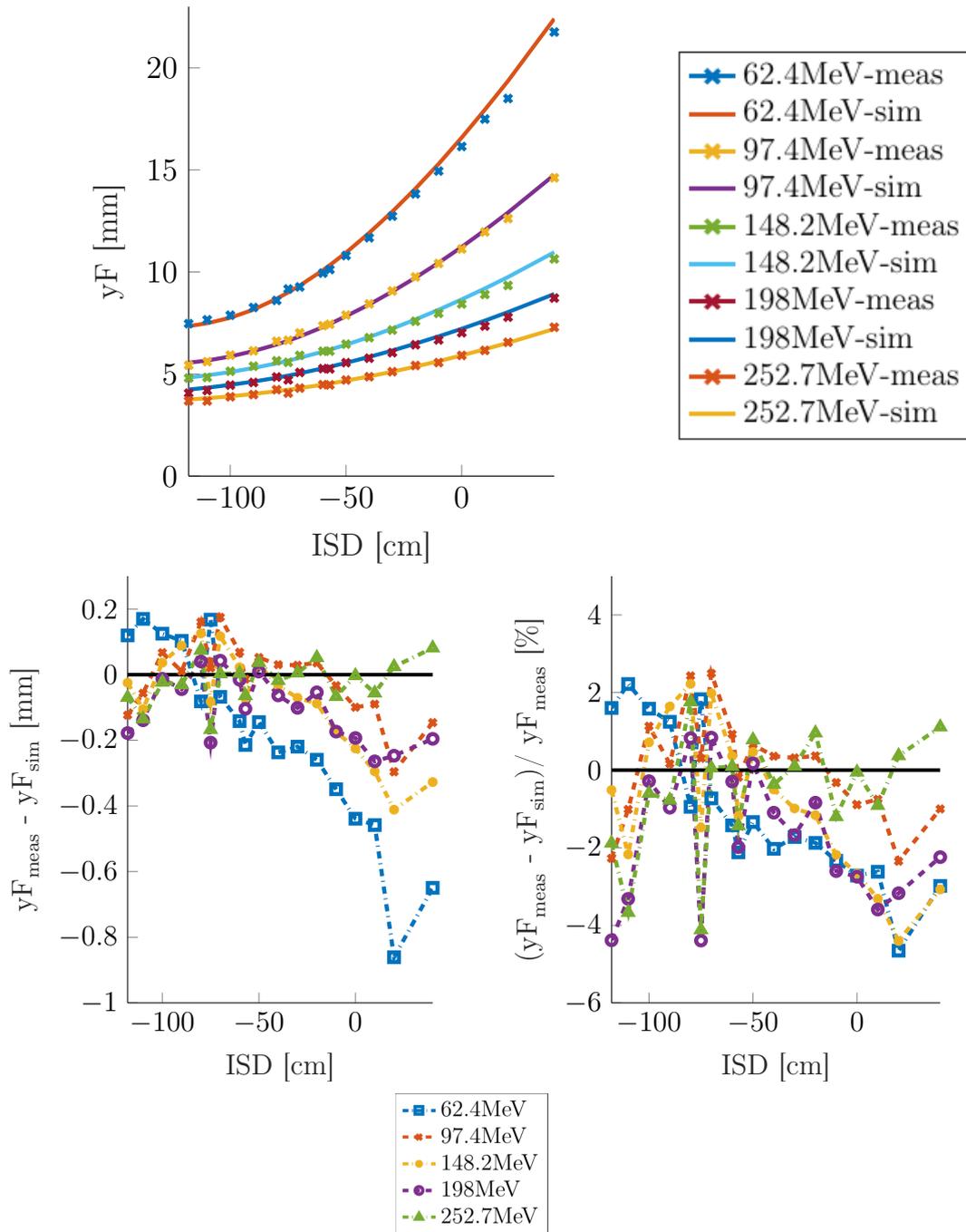


Figure 6.13.: Plots displaying the validation of the vacuum model (using no nozzle for simulation) on the measurements without nozzle for the horizontal plane. The 1st row shows the comparison of measured and simulated FWHMs, the 2nd row the absolute difference, the 3rd row the percentual differences. F is the abbreviation for FWHM.

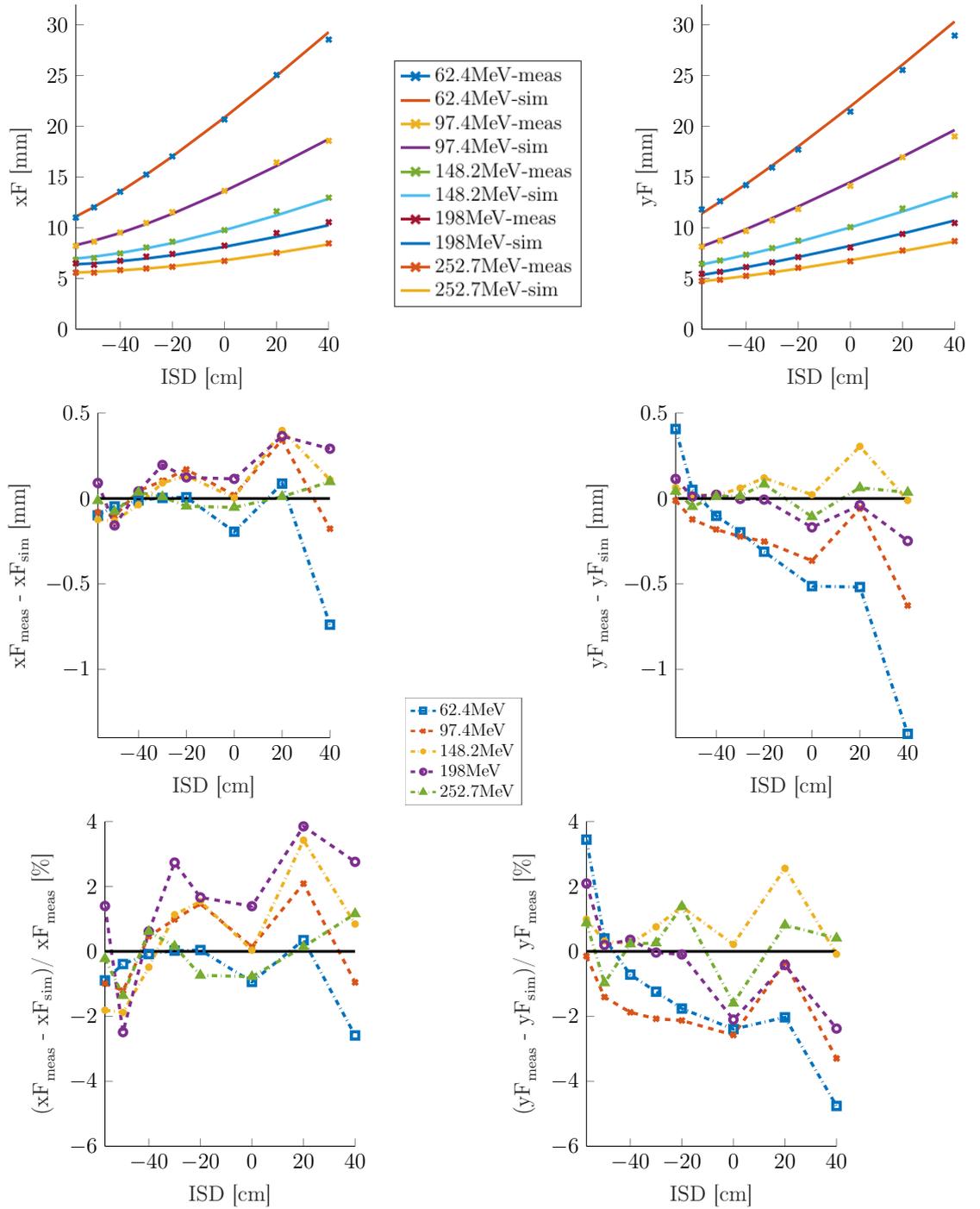


Figure 6.14.: Plots displaying the validation of the nozzle model (using the nozzle for simulation) on the measurements with nozzle. The 1st row shows the comparison of measured and simulated FWHMs, the 2nd row the absolute difference, the 3rd row the percentual differences. F is the abbreviation for FWHM.

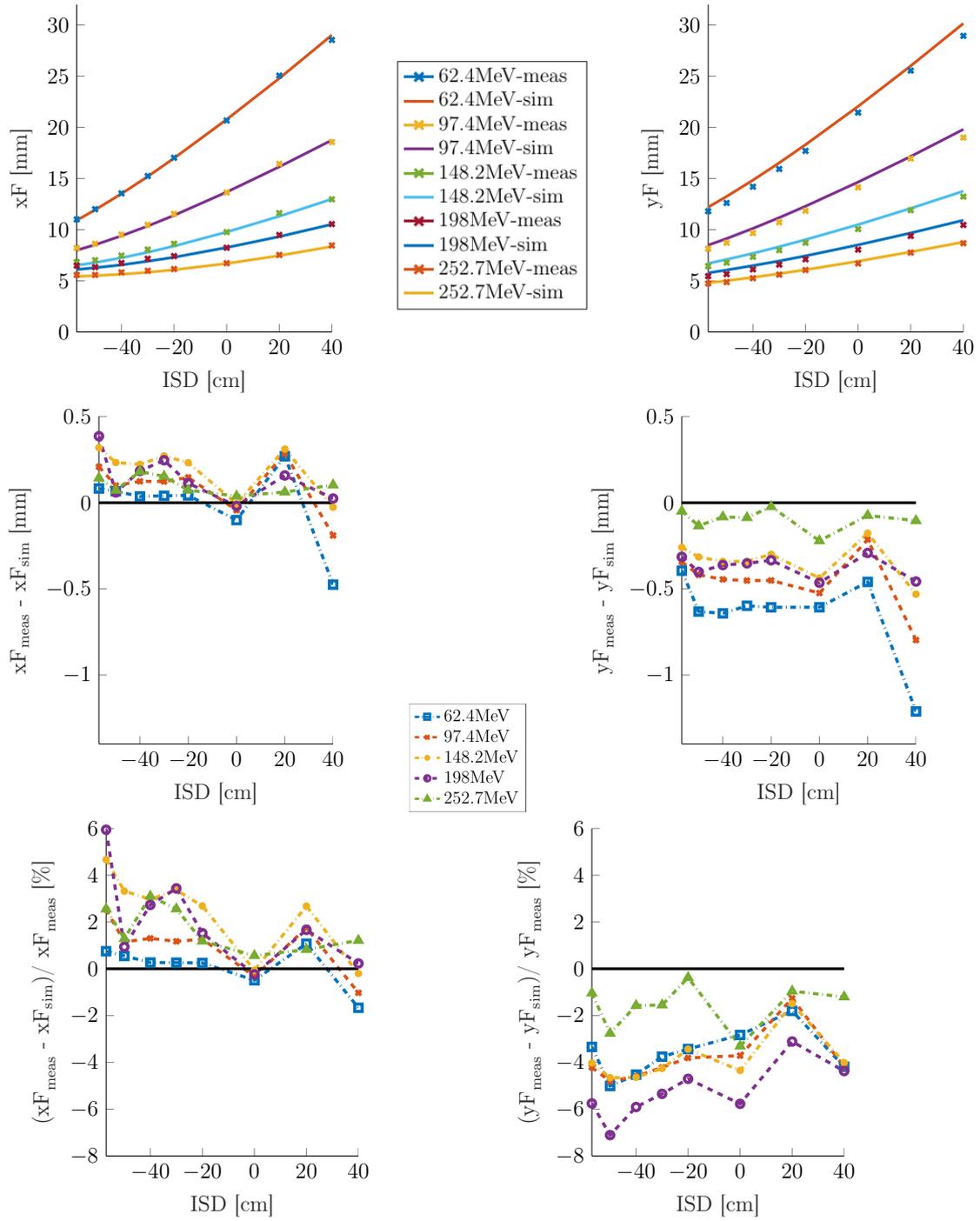


Figure 6.15.: Plots displaying the cross-validation of the vacuum model (using the nozzle for simulation) on the measurements with nozzle. The 1st row shows the comparison of measured and simulated FWHMs, the 2nd row the absolute difference, the 3rd row the percentual differences. F is the abbreviation for FWHM.

7. Treatment plan recalculation

In the previous two chapters the development of the beam energy model and the optical beam model was described and discussed. The combined set of optimized parameters $(E_0, \sigma_{E_0}, \sigma_x, \theta_x, \epsilon_x, \sigma_y, \theta_y, \epsilon_y)$ together with the respective interpolation polynomials constitute the beam model of IR1¹. With this beam model the proton beam in IR1 can now be completely simulated in GATE.

As was discussed in section 2.3, the beam model is a necessary ingredient for independent dose calculations (IDCs) which play a crucial role in quality assurance (QA). The result of such an IDC is a dose distribution which can be compared against the dose distribution predicted by the treatment planning system (TPS).

In this chapter the performance of the obtained *nozzle model*² for independent MC based dose calculations is assessed. Three different RayStation treatment plans were independently recalculated in GATE using the pencil beam-TPS source and the obtained nozzle model as input. The resulting dose distributions were checked against the dose distributions calculated by the treatment planning system (TPS).

In the following the methods outlined above will be explained in more detail.

7.1. Treatment plans

The treatment plans (TPs) that were used are special *QA box treatment plans* in water (no patient treatment plans). The resulting dose distribution of these TPs exhibit a box-shaped high-dose region, hence the name *box* treatment plans.

The three TPs differ in that the center of this box-shaped high-dose region is located at different penetration depths for each TP. The centers are situated at 6 cm, 15 cm, or 25 cm. The TPs are accordingly referred to as *TP6*, *TP15* and *TP25* in the following. Figure 7.1 visualizes the dose distribution of TP15 as predicted by the RayStation treatment planning system.

¹irradiation room 1

²i.e. the combination of beam energy-nozzle model and optical nozzle model

The RayStation TPs were originally available only in Dicom format (.dcm-files) and thus converted to *GATE-plan description files* (.txt-files) using an open source tool from CLITK [54]. Each TP file contains information about the number of pencil beams, their respective intensity and energy as well as their position.

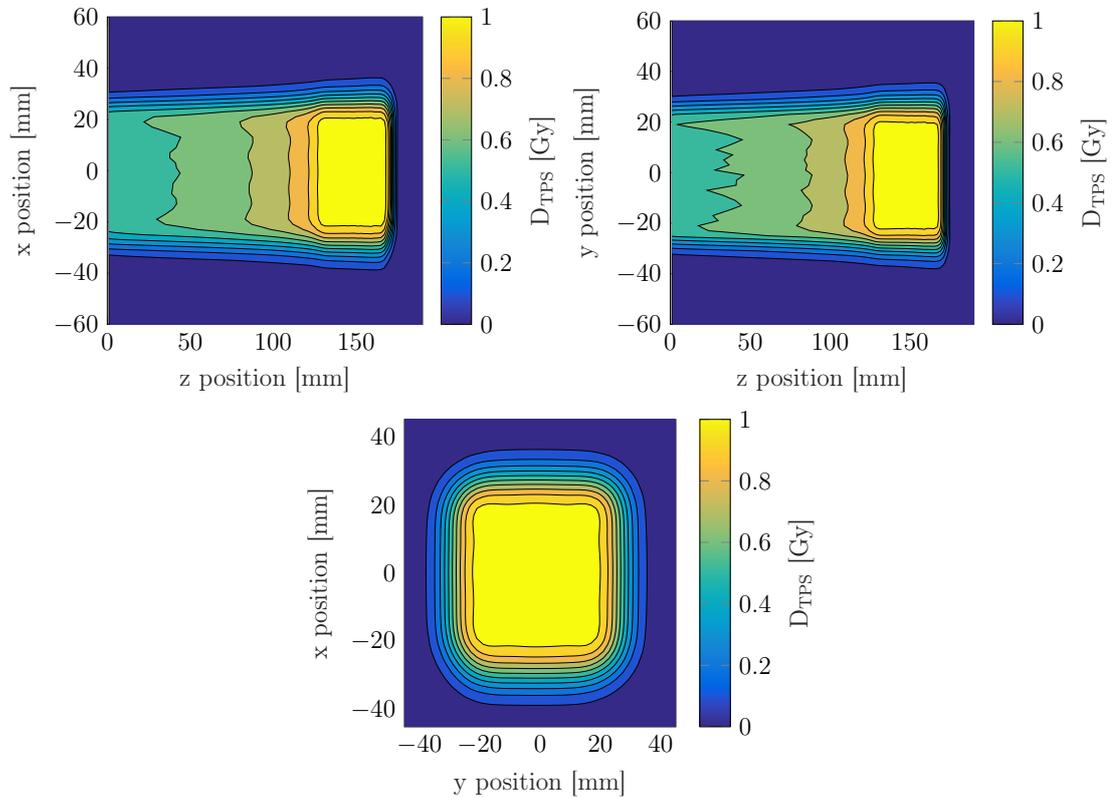


Figure 7.1.: Contour plots of the dose distributions of TP15 in a water phantom, as predicted by the RayStation treatment planning system. The upper two plots show the lateral dose distributions in the x-z-plane and y-z-plane. The lower plot shows the transverse dose distribution in the x-y-plane located in the middle of the SOBP.

In the following section these dose distributions will be recalculated in GATE by making use of the obtained nozzle model.

7.2. Simulation setup and settings

For all simulations regarding treatment plan recalculation the "GATE v8.1 developer version" with GEANT4 in version 10.03.p01 was used.

The simulation setup in GATE is shown in figure 7.2. It is very similar to the one used for beam energy modeling.

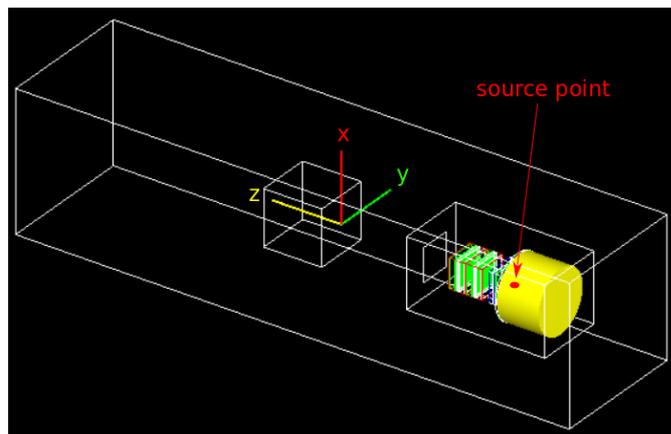


Figure 7.2.: Simulation setup for TP recalculation. The yellow cylinder depicts the vacuum pipe in which the source point is located. The water box is depicted by the small white cube near the coordinate system. The isocenter is located in the origin $(0|0|0)$ of the coordinate system. The world geometry is depicted by the white cuboid surrounding all elements.

The world geometry was created using the parameters given in table 7.1. The isocenter was defined as the origin of the cartesian coordinate system at $(0|0|0)$. The pencil beam-TPS source was used to generate the proton beams. Its source point was set to $(0|0|-130\text{ cm})$ which is located within the vacuum tube before the vacuum windows. The *nozzle model* source description file (see appendix A) was used as input to simulate the pencil beams and the respective plan description file was used to run the treatment plan.

The water phantom was created using the parameters in table 7.1. A dose actor was attached to the waterbox geometry recording the deposited dose using a voxel size of $2 \times 2 \times 2\text{ mm}^3$. The dose map was output as a `.mhd/.raw`-file.

The number of primaries was set to 10^8 for each treatment plan. Table 7.1 summarizes all simulation parameters.

geometry	dimensions [cm ³]	material (ion. pot.)	step size	cuts	actor (voxelsize [mm ³])
world	100x100x400	air	5 cm	10 m	-
NozzleBox	-	various	5 cm	10 m	-
waterbox	40x40x40	G4_WATER (78.0 eV)	0.2 mm	0.2 mm	dose actor (2x2x2)

Table 7.1.: Summary of simulation parameters. The values displayed for the maximal step sizes are valid for protons and the production cut values are valid for electrons, positrons and gammas.

7.3. Comparison of dose distributions

The simulated dose distribution D_{sim} of TP15 is visualized with the TPS-predicted dose distribution D_{TPS} for comparison in figure 7.3. The plots show qualitatively that there are no significant deviations between the independent GATE calculation and the TP prediction.

In order to have a more quantitative comparison, the dose distributions were compared in 1D along a variety of selected profiles (see figure 7.4). Furthermore, *dose measurements* at certain positions were added to the plots. The measurements were performed using 24 PinPoint chambers 31014 mounted on a movable 3D block in an MP3-PL water phantom (PTW, Freiburg, Germany; see section 3.2 for more information on the measurement devices). Figure 7.4 shows that most dose differences between GATE simulation and TPS prediction do not exceed 2.5%. An exception is the plot for varying y position in figure 7.4: dose differences reach up to 7.5%. However, this is due to the fact that the simulated distribution and the TPS-predicted distribution are slightly displaced to each other. The small lateral offset might be due to the fact that the TPS uses a different number of voxels for the y-direction than for the x-direction, resulting in slight voxel disalignment between GATE and the TPS.

The smaller dose differences of up to 2.5% can be explained when considering the simulation uncertainties arising from the finite number of primaries used. Figure 7.5 shows the relative dose uncertainty for each voxel for a transverse cut in the x-y-plane at $z = 150$ mm. The dose uncertainties reach about 2 to 3% at the steep fall-off region located at the "edges" of the box-shaped dose distribution.

This is also the region where the relative dose differences reach their maximum in figure 7.4. Simulation uncertainties can of course be reduced when a larger number of primaries is used for simulation. Another way to increase the accuracy of the treatment plan recalculation can be to use a larger number of representative key energies for beam modeling which increases the accuracy of the polynomial interpolation procedure.

The comparison plots for TP6 and TP25 show similar results and can be found in appendix B. Thus, the nozzle model seems to be suitable³ for independent MC based dose calculations.

³This is only a short test and not a full validation of the nozzle model for IDCs.

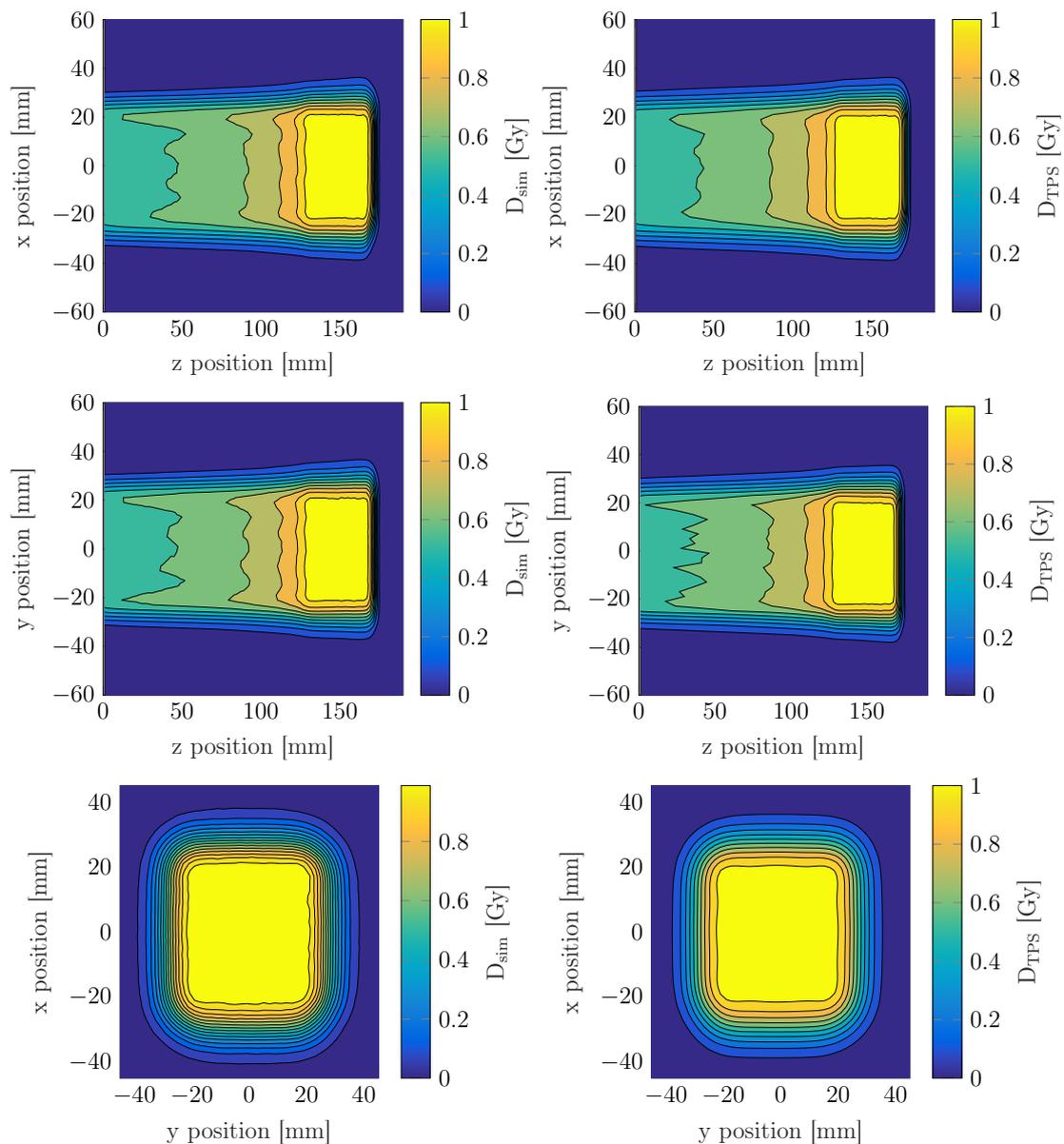


Figure 7.3.: Comparison of dose distributions between the GATE simulation (left column) and the TPS prediction (right column). The 1st row shows the lateral dose distribution for a cut in the x-z-plane, the 2nd row for a cut in the y-z-plane. The 3rd row shows the transverse dose distribution for a cut in the x-y-plane at $z = 150$ mm (located in the middle of the SOBP maximum).

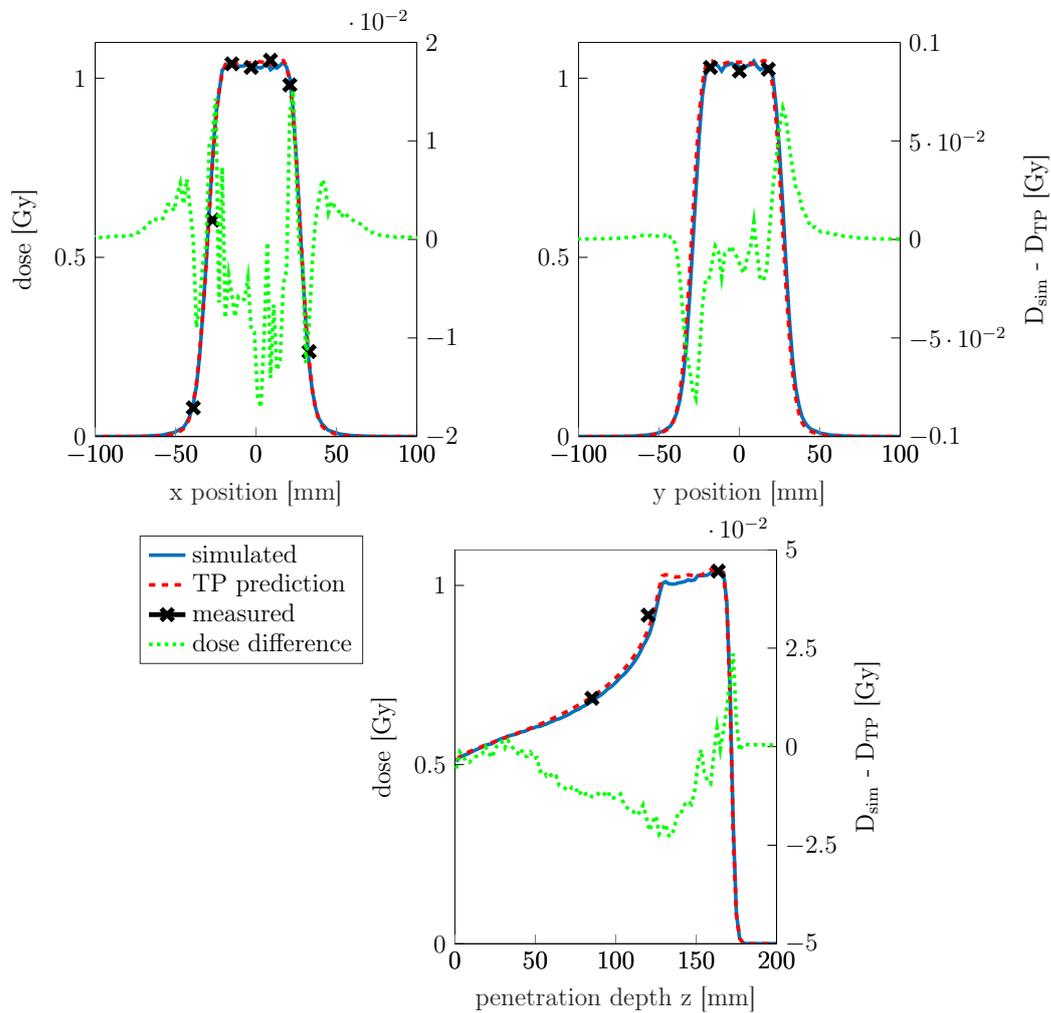


Figure 7.4.: 1D dose distributions along selected profiles. The upper plots show the dose distribution along $y=0$ and $x=0$ in the transverse dose distribution plots of figure 7.3. The lower plot shows the dose distributions along the z -axis. Measurements are depicted as crosses. The green, dotted line displays the absolute dose difference between GATE simulation and TPS prediction $D_{sim} - D_{TPS}$ (on the right y -axes). Since the treatment plan was calculated for 1 Gy, the absolute dose differences are at the same time also relative differences with regard to the dose maximum.

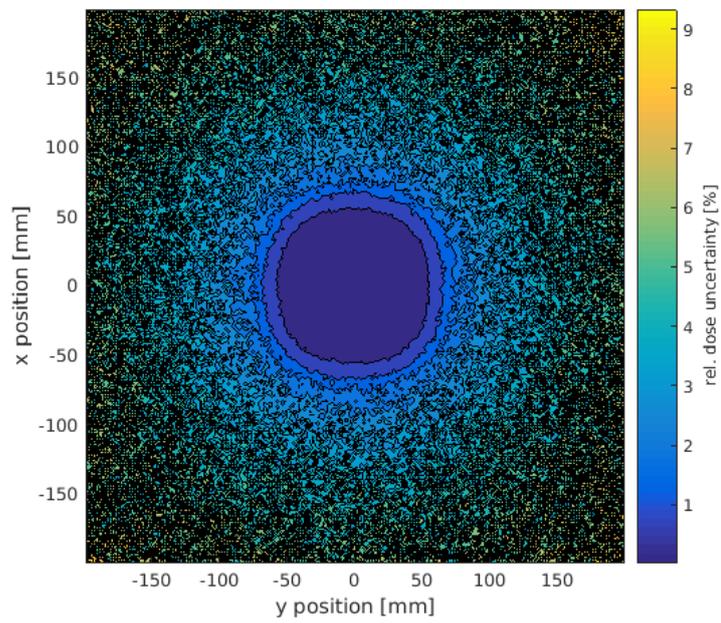


Figure 7.5.: Contour plot showing the relative dose uncertainties in the transverse plane at a depth of $z = 150$ mm (located within the maximum of the SOBP)

8. Summary and outlook

The aim of proton beam therapy is to treat cancer patients using proton beams. Due to their physical characteristics, proton beams allow highly conformal dose distributions for treatment. The high precision leads to an increased sensitivity towards changes within the patients' anatomy, which makes quality assurance (QA) essential. According to quality assurance standards, every treatment plan has to be checked against an independent reference method for possible errors in dose prediction prior to patient treatment. Several possibilities exist for this independent method, e.g. QA measurements. However, nowadays independent dose calculations using Monte Carlo simulations become more and more interesting. In order to fully exploit the potential of accurate Monte Carlo dose calculations a precise beam model is required.

The aim of this thesis was to develop a beam model for the research beam line at MedAustron for the Monte Carlo toolkit GATE. Since the process of beam modeling is usually carried out manually, it is a time-consuming and tedious task. For this reason, a novel automated optimization tool was used that helped to find the optimal beam model given the respective sets of measured beam data. The tool was provided with initial values for the beam model and subsequently performed a sequence of GATE simulations until the difference between simulation results and measurements was minimized. The optimization procedure was performed in two parts: beam energy modeling and optical beam modeling. For the first part, simulated Bragg curves were fitted to measured ones. For the second part, FWHM-vs-ISD curves were used for fitting.

Overall, the optimization algorithm converged to the respective global minimum of the objective function within the given simulation uncertainties. These uncertainties, which arose, for example, from using a finite number of primaries, were at the same time the biggest challenge for the optimization procedure. Additionally, the stochastic genetic algorithm-solver introduced random fluctuations. Energy spread uncertainties ranged from 22% to 46% (corresponding to 0.01 to 0.05 MeV) for the low energy range. For the high energy range uncertainties were even higher. Energy spread optimization did not work for 252.7 MeV since simulations with a monoenergetic beam already resulted in a width larger than the measured one. Further investigations are necessary to determine whether this problem is caused

by overestimated energy straggling in GATE or problematic measurement results. Despite these uncertainties, the optimized energy parameters showed clinically acceptable agreement when validated with the respective measurements. Maximal deviations for Bragg curve ranges were less than 0.1%. Deviations in Bragg peak width reached up to 6%.

Regarding optical parameters, uncertainties for the optimized spot size went up to about 6.9%. The respective uncertainties for the divergence ranged from 20% to 40%. The highest uncertainties were exhibited by the emittance parameter (30% to 70%). Further research should seek to reduce these rather high emittance uncertainties, e.g. by using FWHM measurements at ISDs going beyond +40 cm.

The beam modeling tool allowed to automatically find a beam model within 35 to 40 hours on a conventional notebook. Future research efforts may include to improve the total optimization runtime, e.g. by parallel computing or using faster stochastic optimization methods.

Due to the flexibility of the automated beam modeling tool, not only one but various beam models that differed in nozzle configuration and source point could be created during this thesis. Despite the uncertainties in optimized beam parameters, all found beam models showed clinically acceptable agreement when validated with the respective measurements. Most absolute deviations for the FWHM lay within +0.5 mm and -0.5 mm. Only in a few cases for 62.4 MeV and for high ISDs the deviations exceeded this range which is most likely due to increased measurement uncertainties for ISDs around +40 cm. The corresponding relative deviations for the FWHM lay within +4% and -6%.

The unique feature of a removable nozzle in MedAustron's research room made it possible to independently develop a beam model for the situation with nozzle and a beam model for the situation without nozzle. Validation of the model without nozzle on the FWHM measurements with nozzle was performed by using a fully detailed GATE model of the MedAustron nozzle. The result showed an agreement within clinically acceptable deviations for the comparison of both Bragg curves and FWHM-vs-ISD curves. Consequently, this validation is at the same time a validation for the GATE model of the MedAustron nozzle.

Using the beam model that was optimized for the situation with nozzle, a QA treatment plan of a 3D box in water was recalculated to assess the performance of the model for independent dose calculations. The dose distributions calculated in GATE showed relative deviations in dose of less than 2.5% for most cases. There-

fore, this beam model can be used as the basis of an independent dose calculation tool for the research beam line at MedAustron. Points of improvement include the optimization of beam parameters for more than five key energies, which will result in an even more accurate beam model.

The automated optimization algorithms that were developed throughout this thesis will provide a helpful tool for future applications in beam modeling. It is foreseen to test this automated approach at other particle therapy centers. Furthermore, the tool will be employed at MedAustron for carbon ions where it could be of assistance as it allows to *automatically* find an optimal beam model, reducing the amount of try-and-error based beam modeling sessions.

Bibliography

- [1] S. Agostinelli et al. GEANT4 - a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A*, 506(3):250–303, 2003.
- [2] P. Andreo. On the clinical spatial resolution achievable with protons and heavier charged particle radiotherapy beams. *Phys Med Biol.*, 54(11):N205–15, June 2009.
- [3] I. Antcheva et al. Root — a c++ framework for petabyte data storage, statistical analysis and visualization. *Computer Physics Communications*, 182(6):1384–1385, 2011.
- [4] T. Auberger et al. *Das Projekt MedAustron*. T. Auberger and E. Griesmayer, 2nd edition, 2007.
- [5] M. J. Barnes, J. Borburgh, B. Goddard, and M. Hourican. Injection and extraction magnets: septa. *ArXiv e-prints*, March 2011.
- [6] H. A. Bethe. Molière’s theory of multiple scattering. *Phys. Rev.*, 89:1256–1266, Mar 1953.
- [7] C. Boucher. Phase Space Distributions and Emittance in 2D Charged Particle Beams. <https://www.comsol.com/blogs/phase-space-distributions-and-emittance-in-2d-charged-particle-beams/>. [Online; accessed April 17, 2018].
- [8] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [9] P. Bryant and U. Dorda. MedAustron HEBT design report. ES-081208-d-PBR, ebg MedAustron, 2010-11-29.
- [10] OpenGATE collaboration. Forewords. <http://www.opengatecollaboration.org/>. [Online; accessed May 17, 2018].
- [11] R.C. Davidson and H. Qin. *Physics of Intense Charged Particle Beams in High Energy Accelerators*. Imperial College Press, 2001.

- [12] Hauptverband der österreichischen Sozialversicherungsträger. Erstattungskodex MedAustron. <http://www.erstattungskodex.at/cdscontent/load?contentid=10008.564558>. [Online; accessed March 31, 2018].
- [13] IBA Dosimetry. Lynx - Images and Videos. <http://www.iba-dosimetry.com/solutions/radiation-therapy/particle-therapy-qa/lynx-pt/>. [Online; accessed May 14, 2018].
- [14] M. Durante and J. Loeffler. Charged particles in radiation oncology. *Nature reviews. Clinical oncology*, 7:37–43, Dec 2009.
- [15] A. Ferrari, P.R. Sala, A. Fassio, and J. Ranft. FLUKA: a multi-particle transport code. *CERN 2005-10 (2005), INFN/TC 05/11, SLAC-R-773*.
- [16] L. Freund. Ein mit Röntgen-Strahlen behandelter Fall von Naevus pigmentosus piliferus. *Wiener Medizinische Wochenschrift*, 47(10):428–434, 1897.
- [17] H. Fuchs, A. Elia, A. Resch, C. Lee, L. Grevillot, and D. Georg. Impact of a medical treatment nozzle on beam optics. ESTRO 37 Conference. April, 2018.
- [18] H. Fuchs et al. A pencil beam algorithm for helium ion beams. *Medical Physics*, 39(11):6726–37, 2012.
- [19] A. Garonna et al. Status of Proton Beam Commissioning of the MedAustron Particle Therapy Accelerator. Conference Proceedings of IPAC2016, Busan, Korea.
- [20] S. Gianfaldoni, Gianfaldoni, R., U. Wollina, J. Lotti, G. Tchernev, and T. Lotti. An Overview on Radiotherapy: From Its History to Its Current Applications in Dermatology. *Open Access Maced J Med Sci.*, 5(4):521–525, 2017.
- [21] B. Gottschalk. Techniques of Proton Radiotherapy: Transport Theory. *ArXiv e-prints*, April 2012.
- [22] B. Gottschalk, E. W. Cascio, J. Daartz, and M. S. Wagner. On the nuclear halo of a proton pencil beam stopping in water. *Physics in Medicine and Biology*, 60:5627, July 2015.
- [23] B. Gottschalk, A. M. Koehler, R. J. Schneider, J. M. Sisterson, and M. S. Wagner. Multiple Coulomb scattering of 160 MeV protons. *Nuclear Instruments and Methods in Physics Research B*, 74:467–490, June 1993.

- [24] L. Grevillot et al. GATE as a GEANT4-based Monte Carlo platform for the evaluation of proton pencil beam scanning treatment plans. *Physics in Medicine and Biology*, 57(13):4223–4244, 2012.
- [25] L. Grevillot, M. Stock, and S. Vatnitsky. Evaluation of beam delivery and ripple filter design for non-isocentric proton and carbon ion therapy. *Physics in Medicine and Biology*, 60(20):7985, 2015.
- [26] S. Humphries. *Charged Particle Beams*. Dover Books on Physics Series. Dover Publications, Incorporated, 2013.
- [27] ICRU. *Report 73, Stopping of ions heavier than helium*. International Commission on Radiation Units and Measurements, 2005.
- [28] ICRU. *Report 78, Prescribing, Recording and Reporting Proton-Beam Therapy*. International Commission on Radiation Units and Measurements, 2007.
- [29] S. Jan et al. GATE - Geant4 Application for Tomographic Emission: a simulation toolkit for PET and SPECT. *Physics in Medicine and Biology*, 49(19):4543–4561, 2004.
- [30] F.M. Khan. *Treatment Planning in Radiation Oncology*. LWW Medical Book Collection. Lippincott Williams & Wilkins, 2007.
- [31] F.M. Khan and J.P. Gibbons. *Khan's The Physics of Radiation Therapy*. Wolters Kluwer Health, 5th edition, 2014.
- [32] P. Kimstrand. Beam modelling for treatment planning of scanned proton beams.
- [33] A.C. Kraan. Range verification methods in particle therapy: underlying physics and Monte Carlo modeling. <https://www.frontiersin.org/articles/10.3389/fonc.2015.00150/full>. [Online; accessed May 19, 2018].
- [34] J.C. Lagarias, J.A. Reeds, M.H. Wright, and P.E. Wright. Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions. *SIAM Journal of Optimization*, 9:112–147, 1998.
- [35] J.H. Lawrence et al. Pituitary irradiation with high-energy proton beams: a preliminary report. *Cancer Res.*, 18(2):121–34, 1958.
- [36] W.R. Leo. *Techniques for Nuclear and Particle Physics Experiments: A How-to Approach*. Springer, 1994.

- [37] A. Makarova, B. Gottschalk, and W. Sauerwein. Comparison of Geant4 multiple Coulomb scattering models with theory for radiotherapy protons. *Physics in Medicine and Biology*, 62:5959, August 2017.
- [38] S.R. Mane, Y.M. Shatunov, and K. Yokoya. Spin-polarized charged particle beams in high-energy accelerators. *Reports on Progress in Physics*, 68(9):1997, 2005.
- [39] MedicalPhysicsWeb. Damage limitation: the case for particle therapy. <http://images.iop.org/objects/med/talkingpoint/5/9/1/dam2.jpg>. [Online; accessed March 27, 2018].
- [40] G. Meier et al. Independent dose calculations for commissioning, quality assurance and dose reconstruction of PBS proton therapy. *Phys Med Biol.*, 60(7):2819–36, 2015.
- [41] S. Meroli. Multiple scattering for particles in the matter. http://meroli.web.cern.ch/lecture_multiple_scattering.html. [Online; accessed April 28, 2018].
- [42] W.D. Newhauser and R. Zhang. The physics of proton therapy. *Physics in Medicine and Biology*, 60(8):R155, 2015.
- [43] Medical University of Vienna. Monte carlo calculations in Medicine. <https://www.meduniwien.ac.at/hp/moccamed/>. [Online; accessed June 13, 2018].
- [44] H. Paganetti, H. Jiang, S.Y. Lee, and H.M. Kooy. Accurate Monte Carlo simulations for nozzle design, commissioning and quality assurance for a proton radiation therapy facility. *Med Phys.*, 31(7):2107–18, 2004.
- [45] H. Paganetti, H. Jiang, K. Parodi, R. Slopsma, and M. Engelsman. Clinical implementation of full Monte Carlo dose calculation in proton beam therapy. *Phys Med Biol.*, 53(17):4825–53, 2008.
- [46] M. Palm, M. Benedikt, and A. Fabich. Design choices of the MedAustron nozzles and proton gantry based on modeling of particle scattering. Proceedings of IPAC2011, San Sebastián, Spain.
- [47] A.C. Paulino. *PET-CT in Radiotherapy Treatment Planning*. Elsevier Health Sciences, 2008.
- [48] A. Peeters et al. How costly is particle therapy? Cost analysis of external beam radiotherapy with carbon-ions, protons and photons. *Radiotherapy and Oncology*, 95:45–53, 2010.

- [49] PTCOG. Particle therapy facilities in operation (last update: March 2018). <https://www.ptcog.ch/index.php/facilities-in-operation>. [Online; accessed March 31, 2018].
- [50] PTW. Detectors. http://www.emf-japan.com/ptw/img/PDF/PTW_all_detectors.pdf. [Online; accessed May 25, 2018].
- [51] PTW. Particle therapy QA tools. http://www.ptw.de/typo3conf/ext/naw_secured1/secure.php?u=0&file=ZmlsZWZkbnwluL2ludGVybmFscy9yYWRfdGhlcmFweS9QYXJ0aWNsZV9UaGVyYXB5X1FBX1Rvb2xzX0Jyb&t=1526566995&hash=9e54b8ee030b1e4fbc0e6709271a23f4. [Online; accessed May 22, 2018].
- [52] Huashuai Qu. Simulation optimization: new approaches and an application, 2014.
- [53] W.C. Röntgen. Über eine neue Art von Strahlen. *Annalen der Physik*, 30(1):1–11, 1898.
- [54] D. Sarrut et al. VV, the 4D Slicer. <https://www.creatis.insa-lyon.fr/rio/vv>. [Online; accessed June 13, 2018].
- [55] J. Seco and F. Verhaegen. *Monte Carlo Techniques in Radiation Therapy*. CRC Press, 2013.
- [56] P.G. Seiler, H. Blattmann, S. Kirsch, R.K. Muench, and C. Schilling. A novel tracking technique for the continuous precise measurement of tumour positions in conformal radiotherapy. *Phys Med Biol.*, 45(9):N103–10, 2000.
- [57] J.C. Spall. *Introduction to Stochastic Search and Optimization*. Wiley-Interscience, 1 edition, 2003.
- [58] A. Sulieman. *In Vivo Dosimetry in Radiotherapy and Radiology: Basics, Concepts and Techniques of Patients and Staff Radiation Dosimetry*. Lambert Academic Publishing, 2011.
- [59] Bortfeld T. An analytical approximation of the bragg curve for therapeutic proton beams. *Med Phys.*, 24(12):2024–33, 1997.
- [60] Hirohiko Tsujii. Overview of Carbon-ion Radiotherapy. *Journal of Physics: Conference Series*, 777(1):012032, 2017.
- [61] M. Wannemacher, J. Debus, and F. Wenz. *Strahlentherapie*. Springer Berlin Heidelberg, 2006.

- [62] C.M. Washington and D.T. Leaver. *Principles and Practice of Radiation Therapy*. Mosby, 2015.
- [63] C. Wild, K. Hintringer, and M. Narath. *Protonen und Kohlenstoff-Ionen. Eine Übersicht: Refundierungsstatus Evidenz und Forschungsstand*. Ludwig Boltzmann Institut für Health Technology Assessment, 2013.
- [64] R.R. Wilson. Radiological Use of Fast Protons. *Radiology*, 47(5):487–491, 1946.
- [65] E.M. Zeman, E.C. Schreiber, and J.E. Tepper. Basics of Radiation Therapy. <https://clinicalgate.com/basics-of-radiation-therapy-2/>. [Online; accessed February 17, 2018].
- [66] Y. Zou, J. Tang, and J. Yang. Resonant slow extraction in synchrotrons using anti-symmetric sextupole fields. *Nuclear Instruments and Methods in Physics Research A*, 830:150–162, September 2016.

List of Figures

2.1.	Schematic of a LINAC treatment head for photon therapy	5
2.2.	Comparison of IDD profiles: photons, protons, carbon ions	6
2.3.	Formation of an SOBP	7
2.4.	Layout of the MedAustron accelerator	10
2.5.	Betatron oscillations	12
2.6.	MedAustron synchrotron cross section	13
2.7.	Beamlines in the irradiation rooms	14
2.8.	Schematic of a charged particle beam	17
2.9.	Phase-space representation of a particle beam	18
2.10.	Drift in trace-space representation	22
2.11.	Focusing process in geometrical space and in trace-space	24
2.12.	Process of a scatterer in trace-space	27
2.13.	Trace-spaces of a slow resonant extracted beam	28
2.14.	Beam intensity profile after extraction	29
2.15.	Beam intensity profile after vacuum windows	30
2.16.	Mass stopping power for protons in water	34
2.17.	Small-angle deflections caused by MCS	34
2.18.	Distribution of Molière and Highland	35
2.19.	IDD of a proton beam stopping in water	36
2.20.	IDD profile of proton beams with different initial energy	38
3.1.	Illustration of the MedAustron nozzle	46
3.2.	Water phantom MP3-PL from PTW	47
3.3.	PinPoint chamber from PTW	48
3.4.	Lynx PT from IBA dosimetry	49
4.1.	The principle of beam modeling	55
4.2.	Architecture of GATE simulations used for beam modeling.	56
5.1.	Measurement setup for Bragg curve acquisition	64
5.2.	GATE simulation setup for beam energy modeling	66
5.3.	IDD profile of proton beams with different initial energy spreads	68
5.4.	Standard deviation and simulation runtime for f_1	70
5.5.	Standard deviation and simulation runtime for f_2	70

5.6.	3D and contour plot of f_1	72
5.7.	3D and contour plot of f_2	73
5.8.	Working principle of the beam energy optimization tool	75
5.9.	Scatter plots for single iteration steps for f_1 and f_2	76
5.10.	Comparison of beam energy models	80
5.11.	Simulated Bragg peak width for 252.7 MeV	81
5.12.	Range validation plots for nozzle model	82
5.13.	Width validation plots for nozzle model	83
6.1.	Picture of FWHM-vs-ISD measurements	87
6.2.	Measured FWHM-vs-ISD curves without nozzle	87
6.3.	Measured FWHM-vs-ISD curves with nozzle	88
6.4.	GATE simulation setups for optical beam modeling	89
6.5.	Standard deviations for FWHMs	92
6.6.	Standard deviations of f_3	93
6.7.	3D and contour plots of f_3 for the situations without and with nozzle	94
6.8.	3D and contour plots of f_3 for the situation with nozzle incl. RS	95
6.9.	Comparison of optical beam models: spot size and divergence	101
6.10.	Comparison of optical beam models: emittance	102
6.11.	3D and contour plots of f_3 as a function of emittance	103
6.12.	Validation plots of optical vacuum model for vertical plane	104
6.13.	Validation plots for optical vacuum model for horizontal plane	105
6.14.	Validation plots for optical nozzle model	106
6.15.	Cross-validation plots for optical vacuum model	107
7.1.	Contour plots of the TPS-predicted TP15 dose distribution	110
7.2.	GATE simulation setup for TP recalculation	111
7.3.	Dose distribution comparison between GATE simulation and TPS prediction (TP15)	114
7.4.	1D dose distributions along selected profiles for TP15	115
7.5.	Relative dose uncertainties for TP15	116
B.1.	Validation plots for optical in-nozzle model	xx
B.2.	1D dose distributions along selected profiles for TP6	xxi
B.3.	1D dose distributions along selected profiles for TP25	xxii

List of Tables

5.1. Penetration depth resolution used for Bragg curve acquisition . . .	65
5.2. Summary of simulation parameters for beam energy modeling . . .	67
5.3. Number of primaries used for beam energy optimization	71
5.4. Energy parameters for nozzle model	77
5.5. Energy parameters for vacuum model	77
6.1. Summary of simulation parameters for optical beam modeling . . .	90
6.2. ISD weighting factors	91
6.3. Summary of solver settings for optical beam modeling	96
6.4. Summary of the optical nozzle model	97
7.1. Summary of simulation parameters for TP recalculation	112
A.1. Summary of optical vacuum model	xiv
A.2. Summary of optical in-nozzle model	xv
C.1. Summary of IR1 design parameters	xxiv

Appendices

A. Beam models

Tables for optical beam models

Optical vacuum model								
E_{sys} [MeV]	optimal parameters			objective function	rel. uncertainties [%]			
	σ_x [mm]	θ_x [mrad]	ϵ_x [mm · mrad]	$f_3(\sigma_x, \theta_x, \epsilon_x)$ [10^{-3}]	std(σ_x)	std(θ_x)	std(ϵ_x)	
x-axis								
62.4	3.39	-1.68	4.97	2.1	1.6	8.2	12.6	
97.4	2.92	-1.25	5.29	2.6	3.8	20.2	50.2	
148.2	2.71	-1.22	7.13	2.4	3.3	23.0	31.6	
198.0	2.59	-1.06	6.85	2.4	3.9	37.3	42.4	
252.7	2.66	-1.17	5.84	3.3	2.8	21.2	30.9	
y-axis								
62.4	3.10	0.06	0.31	2.2	2.6	20.7	68.1	
97.4	2.31	0.18	0.51	2.7	4.3	40.3	69.6	
148.2	1.97	0.43	1.56	3.5	6.7	40.6	48.5	
198.0	1.75	0.44	1.25	4.1	6.9	27.7	61.2	
252.7	1.54	0.39	1.61	2.1	6.2	9.6	50.6	

Table A.1.: Summary of the optical vacuum model, including optimal parameters, associated objective function values and uncertainties in percent.

Optical in-nozzle model				
E_{sys} [MeV]	optimal parameters			objective function
	σ_x [mm]	θ_x [mrad]	ϵ_x [mm · mrad]	$f_3(\sigma_x, \theta_x, \epsilon_x)$ [10^{-3}]
x-axis				
62.4	2.88	7.08	7.1	1.1
97.4	2.38	4.08	5.95	2.1
148.2	2.21	2.42	8.1	3.1
198.0	2.19	1.92	7.95	3.7
252.7	2.04	1.47	7.92	3.1
y-axis				
62.4	3.0	7.0	8.13	2.6
97.4	2.32	4.36	5.22	1.4
148.2	1.99	2.74	2.63	1.2
198.0	1.80	1.90	1.69	1.5
252.7	1.63	1.46	1.39	2.4

Table A.2.: Summary of the optical in-nozzle model

Source description files

In this section the GATE-source description files for the auto-optimized nozzle model and vacuum model are listed. These contain the respective polynomial coefficients.

Auto-optimized nozzle model

```
# MA BEAM PARAMETERS
# Nozzle entrance to Isocenter distance
1300
# SMX to Isocenter distance
6700
# SMY to Isocenter distance
7420
# ENERGY
# mean energy
# polynomial order
2
# polynomial parameters
1.34533191e-06
0.9929853798
0.09696886526
# energy spread
```

```

# polynomial order
5
# polynomial parameters
-6.418086744e-11
5.407225281e-08
-1.753032839e-05
0.00271578693
-0.2004698617
5.828744027
# SPOT x y theta phi
# x
# polynomial order
4
# polynomial parameters
-8.68935129030683e-09
5.57650593617634e-06
-0.00126093265841783
0.111301814901103
0.493647781484948
# Theta
# polynomial order
4
# polynomial parameters
-9.61120308864926e-12
6.21490453775891e-09
-1.41136968253855e-06
0.000125607114779005
-0.00170874459934202
# y
# polynomial order
4
# polynomial parameters
-4.55154809482997e-09
2.84515927433158e-06
-0.000606817700412440
0.0481977817503223
0.733132433271404
# Phi
# polynomial order
4
# polynomial parameters
7.63405441380489e-12
-5.26884148171193e-09
1.27531263404439e-06
-0.000128019722037364
0.00502973395262726
# SPOT emittances
# x-theta emittance
# polynomial order
4
# polynomial parameters
2.45829579448800e-11
-1.83874704290875e-08
4.50790291544557e-06
-0.000407058368830991
0.0164405597040246
# y-phi emittance
# polynomial order
4
# polynomial parameters
4.01046066845342e-12
-5.24600387094583e-09
1.72641692808219e-06
-0.000197475244204149
0.00810345907223398

```

Auto-optimized vacuum model

```

# MA BEAM PARAMETERS
# Nozzle entrance to Isocenter distance
1300
# SMX to Isocenter distance
6700
# SMY to Isocenter distance
7420

# ENERGY
# mean energy
# polynomial order
2
# polynomial parameters
1.34533191e-06
0.9929853798
0.09696886526

# energy spread
# polynomial order
5
# polynomial parameters
-6.418086744e-11
5.407225281e-08
-1.753032839e-05
0.00271578693
-0.2004698617
5.828744027

# SPOT x y theta phi
# x
# polynomial order
4
# polynomial parameters
4.34036890346491e-09
-2.88759758625903e-06
0.000716006439585628
-0.0809070991397489
6.28706621392367
# Theta
# polynomial order
4
# polynomial parameters
8.60876580918452e-12
-5.52776558675402e-09
1.27742355247061e-06
-0.000127332312889381
0.00587161958925583
# y
# polynomial order
4
# polynomial parameters
5.90263506127951e-09
-4.18985511249232e-06
0.00108982462532650
-0.127616659976697
7.74383556909331
# Phi
# polynomial order
4
# polynomial parameters
3.67322004242937e-12
-2.30809488845621e-09
4.86100077257995e-07
-3.70441681604046e-05
0.000986206228907976

# SPOT emittances
# x-theta emittance
# polynomial order
4
# polynomial parameters
3.81077996105761e-11
-2.46662272313169e-08
5.43175970180001e-06
-0.000460390752763939
0.0179625666050274

```

```
# y-phi emittance
# polynomial order
4
# polynomial parameters
3.03373655812145e-11
-1.86230153577827e-08
3.93802142722504e-06
-0.000326345356505637
0.00941233394889312
```

B. Remaining plots

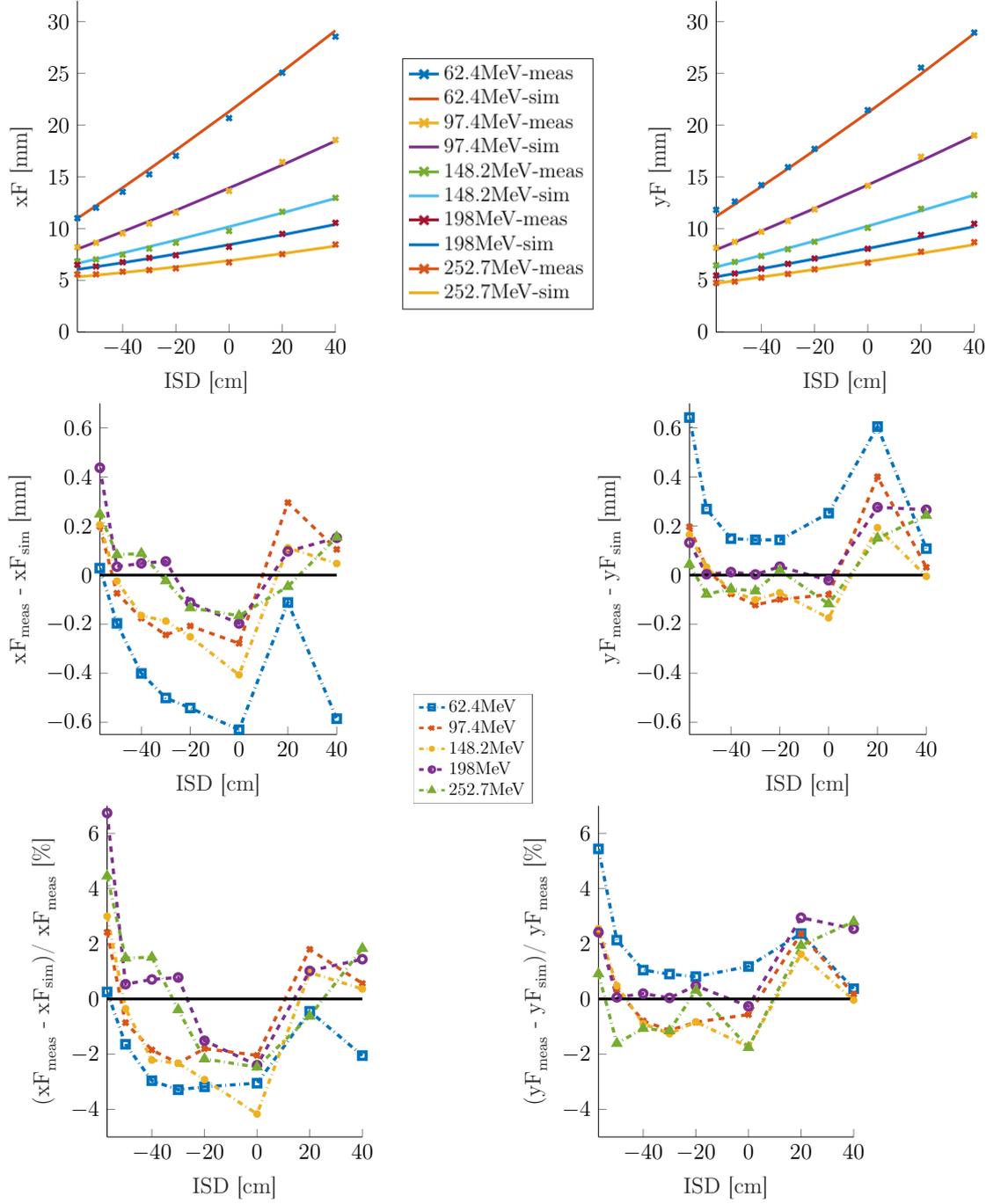


Figure B.1.: Plots displaying the **validation of the in-nozzle model** (using the nozzle for simulation) on the measurements with nozzle. The 1st row shows the comparison of measured and simulated FWHMs, the 2nd row the absolute difference, the 3rd row the percentual differences. F is the abbreviation for FWHM.

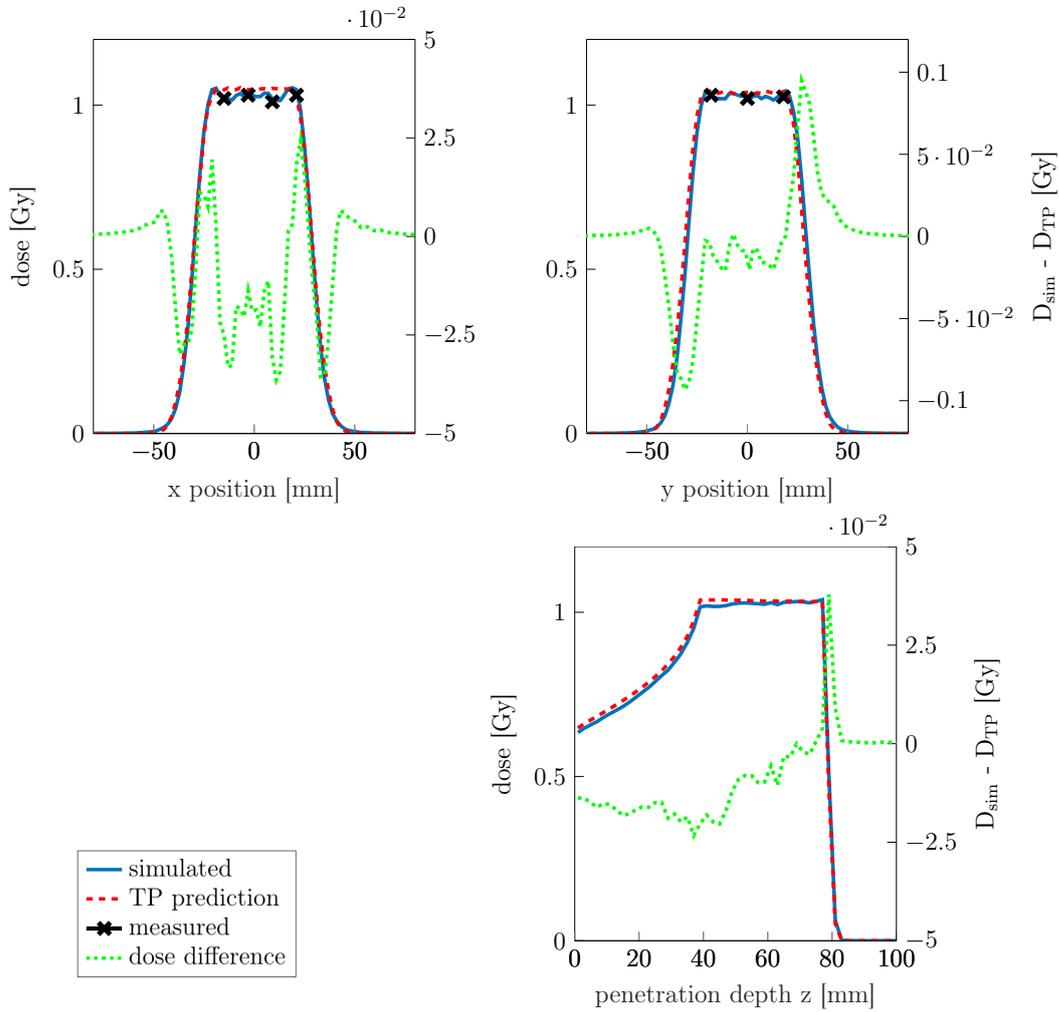


Figure B.2.: **Dose comparison for TP6:** 1D dose distributions along selected profiles. The upper plots show the dose distribution along $y=0$ and $x=0$ for a transverse cut at $z = 60$ mm. The lower plot shows the dose distributions along the z -axis. Measurements are depicted as crosses. The green, dotted line displays the absolute dose difference between GATE simulation and TPS prediction $D_{sim} - D_{TPS}$ (on the right y -axes). Since the treatment plan was calculated for 1 Gy, the absolute dose differences are at the same time also relative differences with regard to the dose maximum.

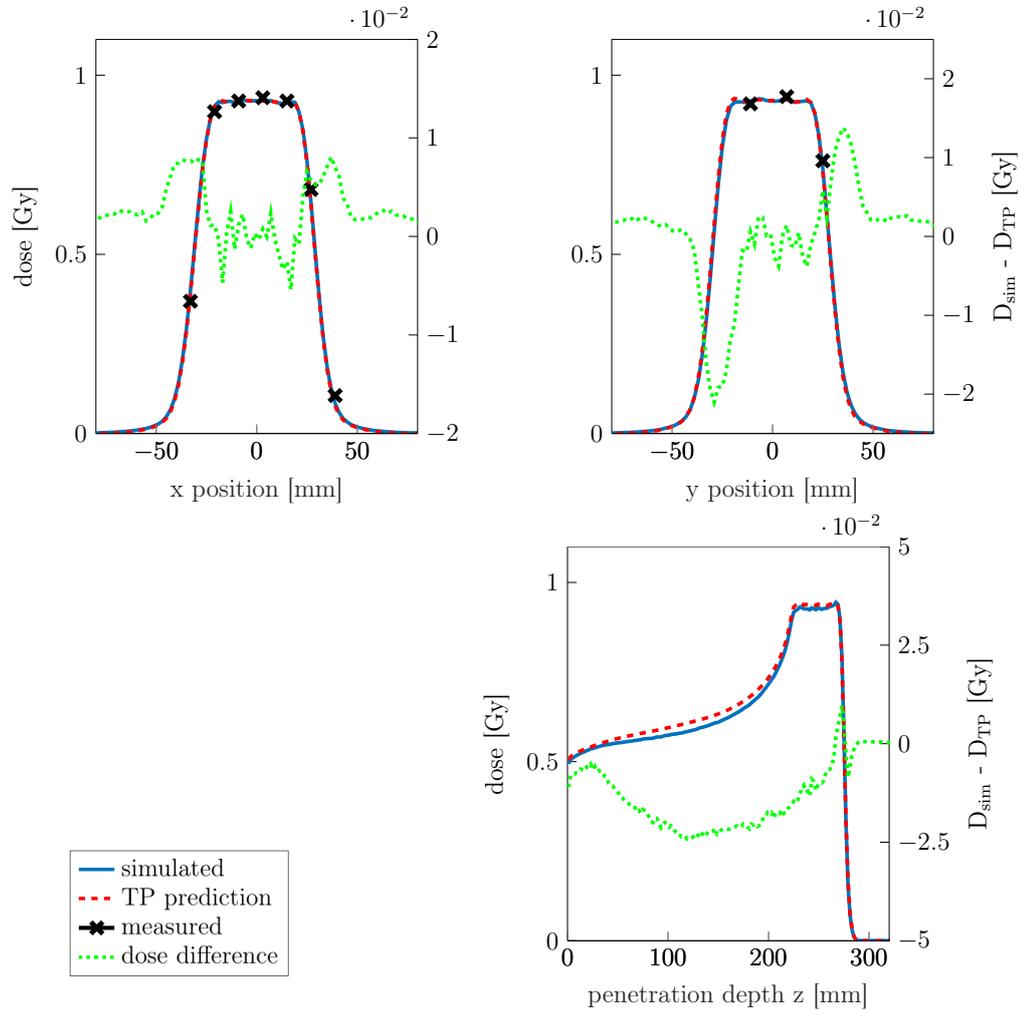


Figure B.3.: **Dose comparison for TP25:** 1D dose distributions along selected profiles. The upper plots show the dose distribution along $y=0$ and $x=0$ for a transverse cut at $z = 250$ mm. The lower plot shows the dose distributions along the z -axis. Measurements are depicted as crosses. The green, dotted line displays the absolute dose difference between GATE simulation and TPS prediction $D_{sim} - D_{TPS}$ (on the right y -axes). Since the treatment plan was calculated for 1 Gy, the absolute dose differences are at the same time also relative differences with regard to the dose maximum.

C. Calculation of beam parameters from Twiss parameters

In this chapter the transformation from Twiss parameters $(\alpha_j, \beta_j, \gamma_j, \epsilon_j)$ to optical beam parameters $(\sigma_j, \theta_j, \epsilon_j)$ ¹ will be explained in detail. These transformations were necessary in order to calculate the beam parameter predictions from MedAustron's design phase which were originally only available in Twiss form.

As was shown in 2.4.3, the two sets $(\alpha_j, \beta_j, \gamma_j, \epsilon_j)$ and $(\sigma_j, \theta_j, \epsilon_j)$ can be used interchangeably for describing a beam. We recall three equations from subsection 2.4.3. The first one is the relationship between geometric emittance ϵ_j and normalized emittance $\epsilon_{n,j}$

$$\epsilon_j(E_{sys}) = \frac{\epsilon_{n,j}}{\beta\gamma} \quad (\text{C.1})$$

in which $\gamma = (1 - \beta^2)^{-1/2}$ and $\beta = \frac{v}{c}$ are relativistic coefficients and must not be mixed up with the Twiss parameters γ_j and β_j . The relativistic coefficients are dependent on the system energy E_{sys} , whereas the normalized emittance is independent from E_{sys} per definition.

The other two equations link Twiss parameters to optical parameters

$$\sigma_j = \sqrt{\beta_j \frac{\epsilon_j(E_{sys})}{\pi}} \quad \text{and} \quad \theta_j = -\text{sgn}(\alpha_j) \sqrt{\gamma_j \frac{\epsilon_j(E_{sys})}{\pi}} \quad (\text{C.2})$$

The normalized x-emittance (i.e., in vertical direction) from the design-phase is

¹the subscript j denotes the transverse directions and can thus take on the values $j = x, y$

given as $\epsilon_{n,x}$ in table C.1. By using²

$$(\beta\gamma)(E_{sys}) = \sqrt{\frac{2E_{sys}}{m_0c^2} + \left(\frac{E_{sys}}{m_0c^2}\right)^2} \quad (\text{C.3})$$

and equation C.1, the x-emittance $\epsilon_x(E_{sys})$ can be calculated as a function of E_{sys} . Here c denotes the speed of light, m_0 denotes the rest mass of a proton and E_{sys} is used synonymously for the kinetic energy E_{kin} of a proton.

Using both formulas in equation C.2, the remaining two optical parameters $\sigma_x(E_{sys})$ and $\theta_x(E_{sys})$ can be calculated as a function of E_{sys} .

In order to calculate the optical parameters for the y-axis (i.e., in horizontal direction), one has to consider the "bar of charge"-behavior of the beam which was theoretically treated in section 2.4.5. As a consequence, the geometric y-emittance ϵ_y as given in table C.1 is constant for varying E_{sys} . This also implies that σ_y and θ_y are constants which can be calculated using the transformation formulas in equation C.2.

j	α_j [rad]	β_j [m]	γ_j [1/m]	emittance [mm · mrad]
x	0.445	2.008	0.597	$\epsilon_{n,x} = 1.63$
y	-0.235	6.483	0.163	$\epsilon_y = 2.371$

Table C.1.: Summary of IR1 design parameters in Twiss form, predicted for the nominal source point at (0|0|-130 cm).

²this equation can be obtained by using the definitions of β and γ , as well as $E_{kin} = (\gamma - 1)m_0 c^2$

D. Source codes

Codes for beam energy modeling

Bragg curve analysis script *BP_analyzer.m*

```
1 function [normDose R_80 width_80 R_50 width_50] = ...
2     BP_analyzer(filepath, energy, ACCURACY)
3 %% read simulation data from file
4 file = filepath;
5 rawdata = dlmread(file, ',', 6, 0); % skipping the first 6 rows
6 % read voxel size in mm
7 filetext = fileread(file);
8 voxelsize_line = regexp(filetext, 'VoxelSize .+)', 'match');
9 voxelsize = regexp(voxelsize_line, '\d[.]\d+', 'match');
10 voxelsize = str2num(cell2mat(voxelsize{1}));
11 % store voxel size and dose data in a 2-column array called "dose"
12 %% 1st column: depth in water in mm; the position in the middle
13 % of a voxel is chosen as the associated depth of that voxel.
14 %% 2nd column: relative integrated dose in arbitrary units
15 depth = linspace(voxelsize/2, numel(data)*voxelsize - ...
16     voxelsize/2, numel(data))';
17 dose = [depth data];
18 clear data depth file filetext rawdata voxelsize_line;
19 %% calculate R_80, R_50, width_80, width_50
20 % linear interpolation used. Thus, the accuracy of the
21 % determination of ranges and widths is equal to the set
22 % parameter ACCURACY.
23 % search dose maximum
24 [maxDose, maxIndex] = max(dose(:, 2));
25 % normalize IDD curve
26 normDose = dose;
27 normDose(:, 2) = normDose(:, 2)/maxDose;
28 % -----
29 % determine R_80
30 % linear interpolation used in order to more accurately
31 % estimate depth where normDose = 0.8!
32 afterpeak = normDose(maxIndex:RmaxIndex, :); % truncate array
33 cmp = 0.8; %set dose value for comparison
34 fineAfter = afterpeak(1,1):ACCURACY:afterpeak(end,1);
35 interpolAfter = interp1(afterpeak(:, 1), afterpeak(:, 2), ...
36     fineAfter);
37 % search for index where the interpol normDose after the peak
38 % best fits cmp=0.8
39 [dummy, auxIndex] = min(abs(interpolAfter - cmp));
40 R_80 = fineAfter(auxIndex);
41 % -----
42 % determine width_80
43 beforepeak = normDose(1:maxIndex, :);
44 cmp = 0.8; %set dose value for comparison
45 fineBefore = beforepeak(1,1):ACCURACY:beforepeak(end,1);
```

```

43 interpolBefore = interp1( beforepeak(:, 1), beforepeak(:, 2), ...
    fineBefore);
44 [ dummy , auxIndex] = min(abs(interpolBefore - cmp));
45 R_80_before = fineBefore(auxIndex);
46 width_80 = R_80 - R_80_before;
47 % -----
48 % determine R_50
49 cmp = 0.5; % set dose value for comparison
50 [ dummy , auxIndex] = min(abs(interpolAfter - cmp));
51 R_50 = fineAfter(auxIndex);
52 % -----
53 % determine width_50
54 cmp = 0.5; %set dose value for comparison
55 [ dummy , auxIndex] = min(abs(interpolBefore - cmp));
56 R_50_before = fineBefore(auxIndex);
57 width_50 = R_50 - R_50_before;
58 % -----
59 clear interpolAfter interpolBefore fineBefore fineAfter ...
    auxIndex afterpeak beforepeak cmp dummy k;

```

optimizer_1.m

```

1 % set energy and nozzle specification globally
2 global ENERGY
3 global NOZZLE
4 global PRIMS
5 ENERGY = 62.4;
6 NOZZLE = 'w';
7 PRIMS = 15000;
8 % initial parameters
9 X0 = 62.4; % in MeV
10 ObjectiveFunction = @ObjFun_1.m;
11 % set optimizer options and start
12 options = optimset('TolFun', 1E-5, 'TolX', 1E-5);
13 [x,fval,exitflag,output] = fminsearch(ObjectiveFunction, X0, ...
    options);

```

ObjFun_1.m

```

1 function F = ObjFun_1(meanE)
2 % meanE and ESig in MeV
3 % declare the global variables ENERGY, NOZZLE and AXIS
4 global ENERGY
5 global NOZZLE % w or wo
6 global PRIMS
7 % if OptimData.mat already exists load data
8 if exist('OptimData.mat', 'file') == 2
9     load('OptimData.mat');
10 end
11 if ~exist('ParAndChiSquare')
12     ParAndChiSquare = [];
13 end
14 if ~isempty(ParAndChiSquare)
15     [lia, locb]=ismember(meanE, ParAndChiSquare(:,1), 'rows');

```

```

16     if lia
17         disp('point already exists!')
18         result = ParAndChiSquare(locb,:);
19         F = ParAndChiSquare(locb,2); return;
20     end
21 end
22 % differentiate the cases with different nozzle configuration
23 switch NOZZLE
24 case 'w'
25     % load measured range according to energy and nozzle setting
26     tbl = readtable('meas/meas_ranges.txt', 'Delimiter', '\t');
27     measRanges = table2array(tbl(:, 3:8));
28     R80Meas = measRanges(2);
29     % load the latest 6 optical parameters for nozzle model
30     tbl = readtable('optical_Parameters_final.xls');
31     X = tbl( strcmp(tbl.Axis, 'x'), :);
32     Y = tbl( strcmp(tbl.Axis, 'y'), :);
33     RotX = X.RotationNorm; RotY = Y.RotationNorm;
34     ESig = 0.2; % set ESig constantly to 0.2MeV
35     % run GATE8.1 simulation
36     MatlabPath = getenv('LD_LIBRARY_PATH'); % Save library paths
37     setenv('LD_LIBRARY_PATH',getenv('PATH')); % Make Matlab ...
        use system libraries
38     % run bash script and pass all the parameters
39     command = ['bash startGateWithParam_wnozzle.sh '[meanE,' ...
        num2str(meanE,15) ']' [ESig,' num2str(ESig,15) ']' ...
        [prims,' num2str(PRIMS,15) ']' ....'''];
40     [status,cmdout] = system(command); cmdout;
41     setenv('LD_LIBRARY_PATH',MatlabPath); % Reassign old ...
        library paths
42     %% read sim. Bragg peak and analyze
43     [simDose R80Sim w80Sim R50Sim w50Sim] = ...
        BP_analyzer(outputfile, ENERGY, 1E-3);
44 case 'wo'
45     ...
46 end
47 %% calculate obj fct value F
48 F = abs( (R80Sim - R80Meas)/ R80Meas );
49 resultN = [meanE F]
50 ParAndChiSquare = [ParAndChiSquare; resultN];
51 save('OptimData.mat', 'ParAndChiSquare');

```

GATE simulation for beam energy optimization

The following GATE source code is written in "GATE v8.1 developer version" and was used to simulate a proton beam stopping in a water phantom (with nozzle).

```

# SETTING MATERIALS DATABASE
/gate/geometry/setMaterialDatabase data/GateMaterials.db

# LOAD GEOMETRY
/control/execute mac/MA-NOZZLE-DESIGN-V6.0/alias-nozzle.mac

# geometry world
/gate/world/geometry/setXLength 400 cm
/gate/world/geometry/setYLength 100 cm
/gate/world/geometry/setZLength 100 cm

```

```

/gate/world/setMaterial Air
/gate/world/vis/forceWireframe
/gate/world/vis/setVisible 1
/gate/world/vis/setColor blue

/control/execute mac/MA-NOZZLE-DESIGN-V6.0/Nozzle.mac

# geometry for waterbox
/gate/world/daughters/name waterbox
/gate/world/daughters/insert box
/gate/waterbox/geometry/setXLength 40 cm
/gate/waterbox/geometry/setYLength 40 cm
/gate/waterbox/geometry/setZLength 42 cm
/gate/waterbox/setMaterial G4_WATER
/gate/waterbox/vis/forceWireframe
/gate/waterbox/vis/setVisible 1
/gate/waterbox/vis/setColor blue
/gate/waterbox/placement/setRotationAxis 0 1 0
/gate/waterbox/placement/setRotationAngle 90 deg
/gate/waterbox/placement/setTranslation -21 0.0 0.0 cm

# ACTORS
/gate/actor/addActor SimulationStatisticActor Statistics
/gate/actor/Statistics/saveEveryNSeconds 30
/gate/actor/Statistics/save output/Statistics.txt

/gate/actor/addActor                               DoseActor   dose1DLong
/gate/actor/dose1DLong/attachTo                     waterbox
/gate/actor/dose1DLong/stepHitType                  random
/gate/actor/dose1DLong/setResolution                1 1 4200
/gate/actor/dose1DLong/enableEdep                   true
/gate/actor/dose1DLong/enableDose                    false
/gate/actor/dose1DLong/enableUncertaintyDose        false
/gate/actor/dose1DLong/enableNumberOfHits           false
/gate/actor/dose1DLong/save                          output/DoseLongitudinal.txt
/gate/actor/dose1DLong/saveEveryNSeconds            3600

# PHYSICS
/gate/physics/addPhysicsList QBBC_EMZ

# set step size and production cuts
/gate/physics/SetMaxStepSizeInRegion world 5 cm
/gate/physics/Gamma/SetCutInRegion world 10 m
/gate/physics/Electron/SetCutInRegion world 10 m
/gate/physics/Positron/SetCutInRegion world 10 m
#-----
/gate/physics/SetMaxStepSizeInRegion NozzleBox 5 cm
/gate/physics/Gamma/SetCutInRegion NozzleBox 10 m
/gate/physics/Electron/SetCutInRegion NozzleBox 10 m
/gate/physics/Positron/SetCutInRegion NozzleBox 10 m
#-----
/gate/physics/SetMaxStepSizeInRegion waterbox 0.1 mm
/gate/physics/Gamma/SetCutInRegion waterbox 0.1 mm
/gate/physics/Electron/SetCutInRegion waterbox 0.1 mm
/gate/physics/Positron/SetCutInRegion waterbox 0.1 mm
#-----
/gate/physics/ActivateStepLimiter proton
/gate/physics/ActivateStepLimiter deuteron
/gate/physics/ActivateStepLimiter triton
/gate/physics/ActivateStepLimiter alpha
/gate/physics/ActivateStepLimiter GenericIon

```

```

# set ionisation potential of water manually
/gate/geometry/setIonisationPotential G4_WATER 78.0 eV

# INITIALISATION
/gate/run/initialize

# CHOOSE BEAM INPUT (SOURCE) FILE
# parameters in {...} are passed from MATLAB
/gate/source/addSource PBS PencilBeam
/gate/source/PBS/setParticleType proton
/gate/source/PBS/setEnergy {meanE} MeV
/gate/source/PBS/setSigmaEnergy {ESig} MeV
/gate/source/PBS/setPosition 1300 0 0 mm
/gate/source/PBS/setSigmaX {SigX} mm
/gate/source/PBS/setSigmaY {SigY} mm
/gate/source/PBS/setSigmaTheta {Theta} mrad
/gate/source/PBS/setSigmaPhi {Phi} mrad
/gate/source/PBS/setEllipseXThetaEmittance {EmmX} mm*mrad
/gate/source/PBS/setEllipseYPhiEmittance {EmmY} mm*mrad
/gate/source/PBS/setEllipseXThetaRotationNorm {RotX}
/gate/source/PBS/setEllipseYPhiRotationNorm {RotY}
/gate/source/PBS/setRotationAxis 0 1 0
/gate/source/PBS/setRotationAngle -90 deg

# START SIMULATION OR GEOMETRY TEST
/gate/application/noGlobalOutput
/gate/random/setEngineName MersenneTwister
/gate/random/setEngineSeed auto

# CHANGE PARTICLE NUMBER FOR SIMULATION
/gate/application/setTotalNumberOfPrimaries {prims}
/gate/application/start

```

Codes for optical beam modeling

batchPhSEval.C

```

1 #include <iostream>, #include <iomanip>, #include <fstream>
2 #include <iostream>, #include <string>, #include <sstream>
3 #include <iomanip>, #include <fstream>, #include <string>
4 #include <cstdlib>, #include <TH1D.h>, #include <TH2D.h>
5 #include <TF2.h>, #include <TStyle.h>, #include <TString.h>
6 #include <TCanvas.h>, #include <TMath.h>, #include <TChain.h>
7
8 //compile to standalone program using
9 //g++ -o atry.out batchPhSEvalV4_improved_binsize.C `root-...
10 config --cflags --glibs`
11 using namespace std;
12 void PhSEval(string FileToAnalyse="", string FileOutAnalyse="...
13     Analyse.txt", string correspondingDepth="0", string part="...
14     proton"){
15 //-----
16 float xPHS;
17 char particleNamePHS[64];
18 double XSigma;
19 double EPHS;
20 string temp="";

```

```

19 // -----
20 // PHS DECLARATION
21 cout<<"Loading " <<FileToAnalyse<<endl;
22 TChain * T_PHS = new TChain("PhaseSpace");
23 T_PHS->Add(FileToAnalyse.c_str());
24 T_PHS->SetBranchAddress("ParticleName",&particleNamePHS);
25 T_PHS->SetBranchAddress("Ekine",&EPHS);
26
27 T_PHS->SetBranchAddress("Y",&xPHS);
28 double weightPHS=1;
29
30 int nPHS = T_PHS->GetEntries();
31 int jPHS = 0;
32 double Xmoy=0.0;
33 double Emoy=0.0;
34
35 for(int i = 0 ; i< nPHS; i++){
36   T_PHS->GetEntry(i);
37   TString pName = particleNamePHS; // ensures that only ...
   // protons are counted
38   if (pName.Contains(part))
39   {
40     jPHS++;
41     Xmoy=Xmoy+xPHS;
42     Emoy=Emoy+EPHS;
43   }
44 }
45 Emoy=Emoy/jPHS; // calculate average energy
46 Xmoy=Xmoy/jPHS; // calculate average x value (should be around ...
   // zero)
47 // -----
48 // HISTOGRAMS
49 // Defining the bin sizes
50 int nBins=600;
51 int n=3;
52 double xBin=150;
53 // xBin is 15 cm because Lynx dimensions are about 30x30cm
54 // i.e.: since nBins=600, the binsize is 0.5mm
55 // -----
56 // HISTOGRAMS
57 // Declaring the histograms
58 TCanvas * tmpCanvas = new TCanvas("tmpCanvas","",1);
59 TH1D *histoBeamProfile;
60 histoBeamProfile = new TH1D("histoBeamProfile", "", nBins, ...
   Xmoy- xBin, Xmoy+xBin);
61 // -----
62 // PROCESSING
63 jPHS = 0;
64 for(int i = 0 ; i< nPHS; i++){
65   T_PHS->GetEntry(i);
66   TString pName = particleNamePHS;
67   if (pName.Contains(part)) // just consider protons as ...
   // particles
68   {
69     jPHS++;
70     histoBeamProfile-> Fill(xPHS, weightPHS);
71   }
72 }
73 // Gaussian Fit
74 histoBeamProfile->Fit("gaus","WWQ");
75 TF1 *fitProfile = histoBeamProfile->GetFunction("gaus");
76 if(histoBeamProfile->GetEntries()>0)

```

```

| 77 {
| 78 XSigma=fitProfile->GetParameter(2); // 0..amplitude, 1...μ, ...
| 79     2...sigma
| 80 }
| 81 else XSigma=0; // if there are no entries in in ...
| 82     histoBeamProfile
| 83 //SAVE to file
| 84 //check whether file already exists and has nonzero number ...
| 85     of rows
| 86     Int_t numberOfLines=0;
| 87 //check if file exists and count number of rows FILE
| 88 cout<<"storing: "<<FileOutAnalyse<<endl;
| 89 ifstream inputfile;
| 90 inputfile.open(FileOutAnalyse.c_str(),std::ios::in);
| 91 if (inputfile.is_open())
| 92 {
| 93     while(!inputfile.eof())
| 94     {
| 95         string tmp;
| 96         std::getline(inputfile, tmp);//read one line of file
| 97         numberOfLines++;
| 98     }
| 99     inputfile.close();
| 100 //if file has already content, append
| 101 //else add header information
| 102     ofstream ofs;
| 103     if( numberOfLines>0)
| 104     {
| 105         ofs.open (FileOutAnalyse.c_str(), std::ofstream::app);
| 106     }
| 107     else
| 108     {
| 109         ofs.open (FileOutAnalyse.c_str(), std::ofstream::out);
| 110         ofs<<"#filename\tcorrespondingDepth (passed)
| 111         \t Nb_particles\tX(FWHM) (mm)"<<endl;
| 112         ofs<<FileToAnalyse<<"\t"<<correspondingDepth<<"
| 113         \t "<<jPHS<<"\t"<<XSigma*2.3548<<endl;
| 114         ofs.close();
| 115     }
| 116 delete fitProfile;
| 117 delete histoBeamProfile;
| 118 delete tmpCanvas;
| 119 return;}

```

optimizer_3.m

```

| 1 clear all;
| 2 close all;
| 3 clc;
| 4 % set energy, nozzle and axis specification globally
| 5 global ENERGY
| 6 global NOZZLE
| 7 global AXIS
| 8 ENERGY = 62.4;
| 9 NOZZLE = 'WITHOUT';
| 10 AXIS = 'x';

```

```

11 % set initial parameters
12 initSig = 3.0;
13 initTheta = 2.0;
14 initEmm = 5.0;
15 ObjectiveFunction = @ObjFun_3; % set obj fun
16 nvars = 3; % number of variables
17 LB = [0 0 0]; % lower bound
18 UB = [10 10 20]; % upper bound
19 X0=[initSig, initTheta, initEmm]; % start value to begin the ...
    algorithm
20 % set options
21 options = gaoptimset(...
22     'MutationFcn',@mutationadaptfeasible,...
23     'CreationFcn',@gacreationlinearfeasible,...
24     'Display','diagnose',...
25     'PopulationSize',20,...
26     'Generations',70,...
27     'InitialPopulation',X0)
28 %start optimization
29 [x,fval,exitflag,output] = ga(ObjectiveFunction,nvars,...
30 [],[],[],[],LB,UB,ConstraintFunction,options)

```

ObjFun_3.m

```

1 function f = ObjFun_3(x)
2 % declare the global variables ENERGY, NOZZLE and AXIS,
3 % in order to get access to it in this function
4 global ENERGY
5 global NOZZLE
6 global AXIS
7 if exist('OptimData.mat', 'file') == 2
8     load('OptimData.mat');
9     end
10    if ~exist('ParAndChiSquare')
11        ParAndChiSquare = [];
12    end
13    %length(ParAndChiSquare)
14    if ~isempty(ParAndChiSquare)
15        [lia, locb]=ismember(x,ParAndChiSquare(:,1:3),'rows');
16        if lia
17            disp('point already exists!')
18            result=ParAndChiSquare(locb,:);
19            f=ParAndChiSquare(locb,4);
20            return
21        end
22    end
23 % check, if parameters obey "Emm <= pi*Theta*Sig"
24 % if not, set f = inf and return
25 if x(3) > pi*x(2)*x(1)
26     disp('CAVE: Emm > pi*Theta*Sig');
27     f = inf;
28     return
29 end
30 % differentiate the cases with different nozzle configuration:
31 % here only shown for without nozzle
32 switch NOZZLE
33 case 'WITHOUT'
34     tbl = ...
        readtable(['meas_data/WONozzle_BeamDataMeasured_', ...
            num2str(ENERGY), 'MeV.txt'], 'Delimiter', '\t');

```

```

35     ISD = table2array(tbl(:, 2));
36     if( strcmp(Axis, 'x') )
37         FWHMMeas = table2array(tbl(:, 3));
38     elseif( strcmp(Axis, 'y') )
39         FWHMMeas = table2array(tbl(:, 4));
40     else
41         disp('ERROR: Axis was not set correctly!')
42         return
43     end
44     % obtain optimized mean energy E and energy spread ...
45     % ESig from outside
46     [~, E, ESig] = getEnergy(ENERGY);
47     % Save library paths
48     MatlabPath = getenv('LD_LIBRARY_PATH');
49     % Make Matlab use system libraries
50     setenv('LD_LIBRARY_PATH',getenv('PATH'));
51     % run GATE simulation by calling shell script
52     command = ['bash startGateWithParam.sh '[Sig,' ...
53               num2str(x(1),15) ' ] [Theta,' num2str(x(2),15) ' ] ...
54               [Emm,' num2str(x(3),15) ' ] [E,' num2str(E,15) ' ] ...
55               [ESig,' num2str(ESig,15) ' ]'''];
56     [status,cmdout] = system(command);
57     cmdout;
58     % Reassign old library paths
59     setenv('LD_LIBRARY_PATH',MatlabPath);
60     tbl = readtable('Analyse-PhS.txt', 'Delimiter', '\t');
61     ISDTemp = table2array(tbl(:, 2));
62     FWHMTemp = table2array(tbl(:, 4));
63     % calculate objective function value
64     FWHMTemp=round(FWHMTemp*10)/10;
65     if sum(FWHMTemp) <= 0
66         f=inf;
67     else
68         tmp=( (FWHMTemp-FWHMMeas) ./FWHMMeas) .^2;
69         % assign weighting factors for ISDs
70         tmp=tmp.*[0.8 1 1 1 1 0.5 1 1 0.3 1 1 1 1 1 1 ...
71                 1 2].';
72         f = sum(tmp);
73     end
74     resultN = [x,f]
75     ParAndChiSquare = [ParAndChiSquare; resultN];
76     save('OptimData.mat', 'ParAndChiSquare');

```