TECHNISCHE
UNIVERSITÄT
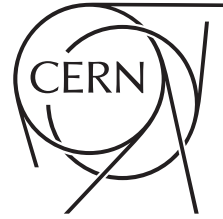WIEN
WIEN
Vienna University of Technology

CERN

DISSERTATION

Performance requirements of the MedAustron beam delivery system

Ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Doktors
der technischen Wissenschaften unter der Leitung von

PRIVATDOZ. DIPL.-ING. DR.TECHN. MICHAEL BENEDIKT
Institutsnummer: E141
Atominstitut der Österreichischen Universitäten

eingereicht an der Technischen Universität Wien
FAKULTÄT FÜR PHYSIK

von

Marcus Palm

Matrikelnummer 0827852
732 Rue de Paris, F-01170 Gex

Wien, im Februar 2013

_____

Marcus Palm

Performance requirements of the MedAustron beam delivery system
Marcus Palm
Atominstitut
Technische Universität Wien

ABSTRACT

The Austrian hadron therapy center MedAustron is currently under construction with patient treatment planned to commence in 2015. Tumors will be irradiated using proton and carbon ions, for which the steeply rising Bragg curve and finite range offer a better conformity of the dose to the geometrical shape of the tumor compared to conventional photon irradiation.

The current trend is to move from passive scattering toward active scanning using a narrow pencil beam in order to reach an even better dose conformation and limit the need of patient specific hardware. The quality of the deposited dose will ultimately depend on the performance of the beam delivery chain: beam profile and extraction stability of the extracted beam, accuracy and ramp rate of the scanning magnet power supplies, and precision of the beam monitors used for verifying the delivered dose. With a sharp lateral penumbra, the transverse dose fall-off can be minimized. This is of particular importance in situations where the lesion is adjacent to critical and radiosensitive structures. Multiple Coulomb scattering upstream of the patient in e.g. air gap, monitors and vacuum windows will limit the sharpness of the beam, in particular for protons in the lower therapeutic energy range.

In this work, clinical requirements on target dose distribution are translated into specifications along the beam delivery chain. By using an analytical model for third integer resonant extraction in combination with particle tracking, the effect of power supply ripple in the main ring magnets on extraction stability has been analyzed. A model for calculating beam growth caused by scattering in a heterogeneous slab geometry along a transfer line (taking focusing and defocussing of the beam into account) has been implemented and used for optimizing the MedAustron nozzles and proton gantry with respect to beam growth at the isocenter. By combining the scattering model with parameterized proton and carbon ion Bragg curves, a three dimensional dose calculator for spot scanning has been implemented to perform a dose distribution error analysis with respect to the *combined* effect of various beam delivery imperfections and limitations. This approach enables for identification of "bottlenecks" at an early stage in the design procedure while avoiding unnecessarily strict specifications. Different scanning modes have been compared and preventive means of suppressing the impact of beam delivery imperfections already on the treatment planning stage are discussed.

Acknowledgments

First of all, I would like to thank my supervisor Privatdozent. Dr. Michael Benedikt, MedAustron project leader at CERN, for his kind support and valuable advice the last years and the opportunity to return to CERN for doctoral studies. A special thanks also to Dr. Adrian Fabich for encouraging me to apply for a PhD position in the MedAustron team in the first place and his continuous support and feedback throughout the PhD time in both work related and private matters.

I would also like to express my gratitude to Dr. Stanislav Vatnitsky for sharing his many years of experience and putting me on the right track.

I am very grateful to Dr. Ulrich Dorda for numerous rewarding discussions and invaluable feedback on beam optics, requirements issues and proofreading.

A warm thanks also to Fabian Moser for his help with the Bragg curve simulations and an enjoyable time together in the office.

I would also like to thank the PSI proton therapy team for the opportunity to do measurements in Gantry 2 and kindly sharing their expertise and experience. Lots of support were also given by the MedAustron team at PSI as well: Albin, Alex, Stefania, Maarten and Olaf.

The Prevessin lunch crew; Alex, Florian, Claire, Angela, Morten and Jeff: I will miss you and our unforgettable (for better or worse) lunch conversations.

All friends and family: thank you for being there and making me the one I am.

Lastly, to Christina: thank you for caring. I can't wait to spend the rest of my life with you and our William.

Contents

Chapter 1

Introduction and motivation

1.1 Proton and carbon ion therapy

Photons have been used for treatment of cancer since 1896, when Emil Grubbe applied X-rays to a female patient suffering from breast cancer, only weeks after Röntgen's publication of the newly discovered rays [1]. Half a century later, Robert Wilson published a seminal paper on the advantages of using protons instead of X-rays for irradiation of cancer [2]. The fundamental difference is illustrated by the depth-dose curves shown in Fig. 1.1. While X-rays (and electrons) deposit the highest dose close to the surface, the proton dose is lowest at the surface, ending in a sharp peak - the *Bragg peak*. For non-superficial tumors, the Bragg curve of ions allow for a superior dose localisation, i.e. the dose can be concentrated to where the tumor is: the dose given to healthy tissue in front of the tumor can be reduced, and, unlike photons which show an exponential dose reduction with depth, protons deposit virtually no dose at all beyond the Bragg peak due to their finite range.

Since the sharp Bragg peak typically is much smaller than the actual size of the tumor, the energy of the incoming beam must be modulated to create a Spread Out Bragg peak (SOBP). Up until the 1990's, proton therapy was mainly executed by means of *passive scattering*: a mono-energetic beam is passed through a range modulator (e.g a fast rotating wheel with varying thickness), introducing an energy spread, and then scattered and collimated such that the lateral extent of the field matches the transverse shape of the tumor. Although this provides dose sparing of normal tissue, compared to photons, there are disadvantages: patient-specific hardware is required and there are limitations in conforming the dose in three dimensions to the tumor [4].

Since 1994, proton therapy by means of *active scanning* has been performed at PSI (Paul Scherrer Institute), Switzerland [5]. A cyclotron delivers a mono-energetic proton beam to the irradiation room, where it passes through a stack of range shifters. By inserting or removing range shifters, the penetration depth of the beam in the patient can be controlled in steps of about 5 mm. The entire tumor is thus divided into *layers*, where each layer corresponds to the penetration depth of a particular beam energy. Each layer, in turn, is divided into a two-dimensional map
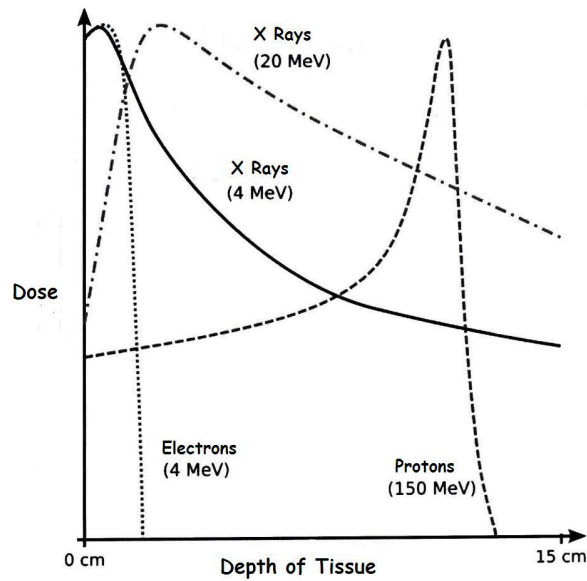
Figure 1.1: Depth dose of electrons, X-rays and protons entering human tissue for the purpose of cancer treatment [3].

of *spots*, where each spot is to receive a certain number of particles at a certain energy. Dipole magnets upstream of the patient direct the beam towards a spot until it has received its prescribed dose, and then moves the beam to the next spot. Once all spots in a layer are fully irradiated, the energy of the beam is changed by inserting or removing range shifters, and irradiation continues with the next layer. A conceptual overview of an active scanning system is shown in Fig. 1.2.
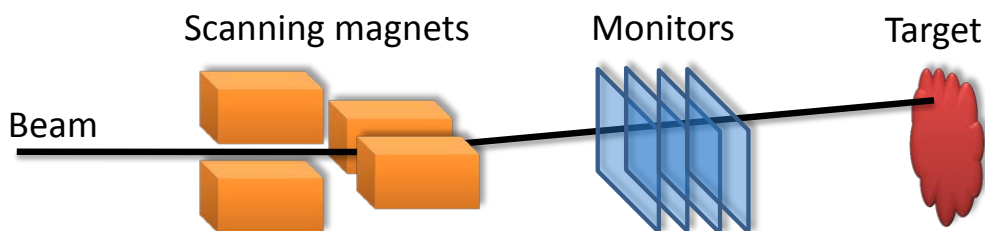
Figure 1.2: Active scanning principle. A narrow beam is deflected horizontally and vertically by two scanning magnets towards the current spot. Beam monitors in front of the target verify that the beam intensity, position and size is correct. When one spot is fully irradiated, the scanning magnets move the beam to the next spot.

Active scanning offers a better dose conformation than passive scattering, without the need of patient-specific hardware in the beam path. Additionally, the patient

will not receive any neutron dose from scattering elements [4][1].

One way of quantifying the biological effect of an ion beam is its *Relative Biological Effectiveness* (RBE). The RBE is defined as the ratio between the photon dose, D_X, given under some reference conditions, and the ion beam dose, D_{ion}, that is required to achieve *the same* biological effect as the photon dose:

$$RBE = \frac{D_X}{D_{ion}} \tag{1.1}$$

The biological effect to be reached could be e.g. 50% survival of irradiated cells, or a certain probability to develop tumors.

The exact RBE value depends on a multitude of variables such as cell type, tissue type, dose rate, fractionation scheme (how often and in which portions the total dose is delivered), biological effect under investigation, and the beam energy. The RBE of protons is typically taken to be close to unity (around 1.1 [4]) along the Bragg curve, i.e. the biological effect of protons and photons is similar. For carbon ions, on the other hand, the RBE values in use show a larger variation: in [6], RBE values between 0.2 and 9.6 are mentioned. HIMAC (Heavy Ion Medical Accelerator in Chiba) - an active scanning carbon ion facility in Japan - reports using RBE values between 2 and 4: the higher value in the distal part of the SOBP and around 3 in the center of the SOBP [7]. With an RBE that increases with depth, carbon ions offer a potential advantage over protons (similarly to protons over photons) in that the same biological effect can be achieved in the tumor while reducing the effective dose to surrounding healthy tissue.

Regardless of the type of radiation, the goal is the same: to kill the cancerous cells by inflicting irreparable damage to their DNA, while keeping the dose given to healthy tissue at an acceptable level. Damage to only one strand of the DNA double helix can be repaired while if both strands are broken, the chance of repair is much smaller. Carbon ions are advantageous in that their ionization tracks are about the same size as the DNA double helix, as illustrated in Fig. 1.3. The effect of this is that carbon ions have a higher probability of breaking *both* sides of the DNA helix (a "double strand break") [8].
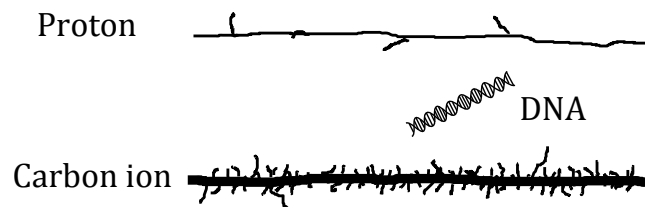

Figure 1.3: Proton and carbon ionization track and DNA double helix.

---

[1]If the range shifters are placed far away from the patient

1.2 Cyclotrons vs. synchrotrons

Existing active scanning hadron therapy facilities can primarily be sorted in two categories: cyclotron facilities and synchrotron facilities. To date, all cyclotron facilities produce proton beams exclusively. This is due to technical limitations: the energy required for a carbon ion beam to reach the same penetration depth as a proton beam is many times higher. Most synchrotron facilities provide carbon ion beams (and possibly other ions up to neon) for therapeutic use[2]. While synchrotrons have the advantage of energy selection directly from the synchrotron, cyclotrons allows for more flexibility in beam intensity control and stability. This makes it possible to turn the beam off between spots ("discrete scanning" [4]) while verifying that the last spot was delivered correctly. Pausing the extraction on a spot-to-spot basis with a synchrotron is more challenging, and the beam is therefore kept on while moving from spot to spot ("quasi-discrete scanning" [4]).

	Cyclotron	Synchrotron
Facilities	PSI, Switzerland Skandion, Sweden (under construction)	MedAustron, Austria CNAO, Italy HIMAC, Japan HIT, Germany Loma Linda, USA M.D. Anderson, USA
Particles	p	p, C (He,Ne,O)
Scanning technique	Discrete scanning	Quasi-discrete scanning
Energy selection	Range shifter, movable wedges	Via synchrotron

Table 1.1: Comparison between active scanning cyclotron and synchrotron facilities.

1.3 The MedAustron project

MedAustron is a synchrotron based hadron therapy facility currently under construction in Wiener Neustadt, Austria, 40 km south-west of Vienna. The decision for construction was taken in 2006 by the County of Lower Austria in partnership with the Austrian federal government.

With a capacity of 24,000 fractions per year, between 1200 and 1400 patients per year (depending on fractionation scheme) can be treated with protons and carbon ions. Irradiation of the patients will be performed in three different treatment rooms (see Fig. 1.4):

IR2: Fixed horizontal + fixed vertical beam line (protons and carbon ions)

---

[2]One exception is Loma Linda, Houston, which uses a synchrotron for proton use only [9]

IR3: Fixed horizontal beam line (protons and carbon ions)

IR4: Gantry (protons only)

In addition, one irradiation room (IR1) is dedicated to non-clinical research.
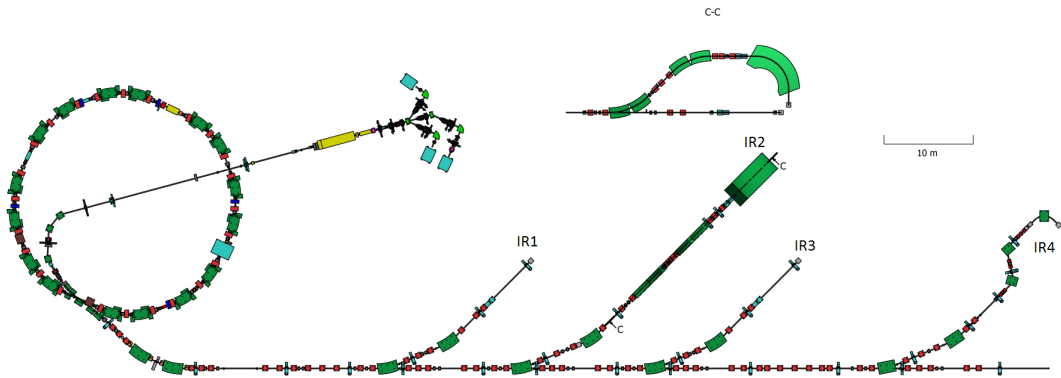

Figure 1.4: Layout of the MedAustron beam lines.

Collaboration with several other actors in the field of hadron therapy and synchrotrons has been established [10]: the design, construction and follow-up of the manufacturing of the MedAustron accelerator facility is carried out in partnership with the CERN, the world's largest particle physics research laboratory, located in Geneva, Switzerland[3]. From the Italian hadron therapy facility CNAO (Centro Nazionale di Adroterapia Oncologica[4], documentation concerning injector, synchrotron, high-energy transfer line and technical infrastructure has been made available: both CNAO and MedAustron are based on the PIMMS (Proton Ion Medical Machine Study) design [11, 12]. A collaboration agreement has also been signed with PSI[5] (PSI) for construction of the MedAustron proton gantry, which will be a hardware copy of the state-of-the-art Gantry 2, designed and developed at PSI [13, 14].

1.4 The goal of this thesis

The quality of the dose delivered to the patient is ultimately determined by the performance of all components preceding the patient: the quality of the beam extracted from the accelerator, the transport line from the accelerator to the treatment room and the Beam Delivery System (BDS) (scanning magnets, beam monitors and passive elements in the nozzle). Although clinical specifications on the patient dose quality are readily available, a direct translation into performance requirement on the beam delivery chain is elusive: a requirement that the delivered dose should be accurate to so and so many percent tells little about how fast the scanning magnet

---

[3]http://www.cern.ch/
[4]http://www.cnao.it/
[5]http://p-therapie.web.psi.ch

must move the beam, or the required stability of the extracted particle flux. Even less apparent is the *combined* effect of different kinds of beam delivery imperfections, such as beam positioning errors at the patient or beam flux fluctuations, on target dose quality.

The goal of this thesis is to link the clinically motivated dose requirements to performance specifications on the different parts of the beam delivery chain. This obviously includes active elements, such as the scanning magnets and the accelerator itself, but also evaluation of e.g. beam growth due to scattering in the nozzle and other passive elements upstream of the patient, since the shape of the beam reaching the patient affects the dose distribution.

Having linked the performance of different parts of the the beam delivery chain to resulting target dose quality, one can produce a *consistent* set of requirements along the beam delivery chain. Consistent here means that different parts of the beam delivery system contribute about equally to dose inaccuracies (within clinical specifications). This approach avoids specifications that are unnecessarily strict and costly. Additionally, one can early identify and avoid "bottlenecks". One such example is the velocity of the scanned beam: if the beam movement is too slow, the highest beam fluxes the accelerator is designed for cannot be utilized without causing systematic overdosage.

Alternatively, if requirements on one part of the beam delivery chain are already fixed (for example, in an existing facility that considers upgrading e.g. the scanning magnets), a direct link between performance of the different beam delivery components and dose quality is a useful tool for determining whether an upgrade would have any substantial improvement on dose quality or scanning performance.

The details of unraveling the relation between beam delivery components and dose quality obviously depends on the scanning facility. However, the methods used in this thesis are general and can be applied to any hadron therapy facility.

1.5 Thesis outline

Chapter 2 describes the relevant clinical requirements for active scanning.

Chapter 3 describes the theoretical models used for evaluating the target dose distribution: extraction from the synchrotron, multiple coulomb scattering (causing lateral beam growth) and a parameterized Bragg curve model. This chapter also describes the MedAustron scanning technique, the proton gantry and the nozzle elements.

A schematic overview of the beam delivery chain is shown in Fig. 1.5. The synchrotron delivers a beam with a certain shape and time profile, described and evaluated in Chapter 4. From the synchrotron, the beam is guided either to one of the fixed beam line rooms or the proton gantry where two scanning magnets deflect the beam in the proper direction, before entering the nozzle.

Chapter 5 contains an evaluation of different nozzle and proton gantry options, with respect to transverse and longitudinal beam profile at the patient.

Finally, an error analysis of the resulting target dose distribution is made in Chap-

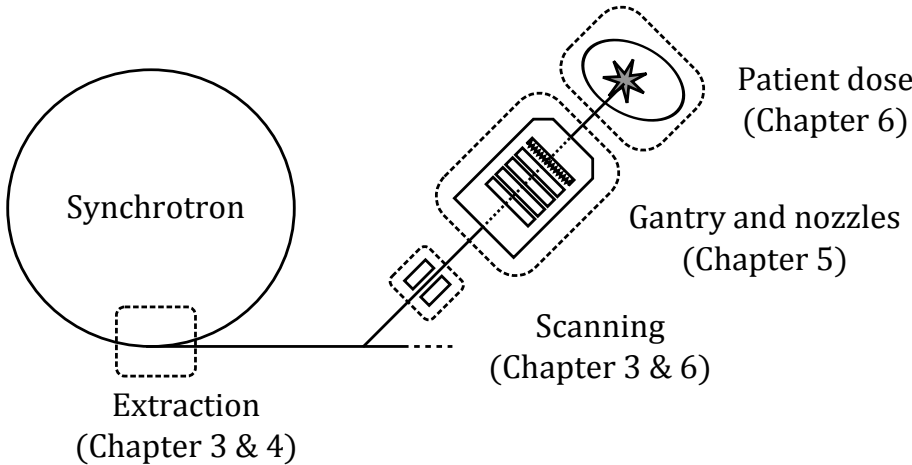ter 6, resulting in requirements on the beam delivery chain.

Figure 1.5: Schematic overview of the beam delivery chain (not to scale) and overview of thesis layout.

Chapter 2

Clinical requirements

2.1 Beam range

A beam range of up to 25 g/cm^2 (25 cm in water) is sufficient to satisfy most therapeutic requests - even for deep seated tumors in obese patients [15]. The available beam ranges in water for the MedAustron synchrotron are up to 27 cm for carbon ions and up to 37 cm for protons. Since protons are lighter than carbon ions, they can be accelerated to higher energies in the synchrotron, which enables a longer proton range.

2.2 Irradiation time

A fractional dose of 2 Gy must typically be given to a 2 liter volume in less than 3 minutes.

2.3 Dose homogeneity

Dose homogeneity requirements set a limit on how much the delivered dose in the target volume is allowed to deviate from the approved treatment plan. Many different formulations of dose homogeneity specifications exists; both in terms of formulation and in terms of limits.

For example, the Tera[1] report *Clinical requirements and physical specifications of therapeutical proton beams* [15] from 1994 requires that the ratio between maximum and minimum absorbed dose, P_{max}/P_{min}, in a plane transverse to the beam is between 1 and 1.05. The term *absorbed dose* here refers to the averaged absorbed dose in an area not larger than 0.25 cm^2.

In the longitudinal direction, the Tera report requires a homogeneity better than 111%.

In discussion with the medical expertise of MedAustron, it has been decided *not* to adopt these requirements for two reasons:

---

[1] http://www.tera.it/

- From a medical point of view, there is no apparent reason to employ different homogeneity requirements in the transverse and longitudinal directions. One target is typically irradiated from different angles in the same fraction, so a target plane that is transverse with one beam entrance angle can very well be a longitudinal plane when using a different beam entrance angle.

- Minor random beam delivery system errors in e.g. beam positioning or spot weight accuracy will result in a target dose distribution that is normal-distributed around some value (see Chapter 6). With the Tera requirements, even the tails of this dose distribution must be within specified limits, regardless of how small fraction of the target the tails contain. More natural would be to specify a maximum width of the resulting dose error distribution.

The *Performance Specifications for a Proton Medical Facility* [16] from 1993 by Chu et. al does not discern between transverse and longitudinal direction and recommends a dose compliance (maximum variation from specified dose) better than 2.5% within the treatment field (excluding the lateral, distal and proximal target edges). Like the Tera requirements, this is an absolute requirement that does not take into account the dose error distribution tails caused by random beam delivery system or uncertainties in the determination of the absorbed dose.

To address these issues, the principles outlined in the document *What accuracy is required and can be achieved in radiation therapy* by Wambersie [17] has been used as a starting point in determining feasible dose homogeneity requirements for MedAustron. These requirements on dose homogeneity will be the basis for technical specifications on various aspects of, on the one hand, the beam delivery system and the synchrotron, and, on the other hand, the Beam Verification System (BVS) that ensures that the beam specifications are within tolerances.

Wambersie considers two types of dose uncertainties: Type A (random) and Type B (systematic).

Type A errors (random) can be estimated from repeated independent measurements and quantified by e.g. the standard deviation, σ_A. A typical Type A error is beam position errors caused by dipole current fluctuations, or spot weight errors caused by fluctuations in the particle flux[2] from the accelerator.

The influence of systematic uncertainties (Type B) remains the same in repeated observation and an interval which contains the "true" value in about 70% of the cases (the 70%-interval corresponds roughly to the $\pm 1\sigma$ interval of a normal distribution) can therefore not be estimated statistically. An appropriate interval should therefore be based on a careful analysis of the procedures used. A typical Tybe B error source would be systematic extraction energy errors. Another source is the RBE values used in the treatment planning which influence the biological dose delivered to the tumour.

Another source of dose uncertainty is **patient positioning**: if the patient is misplaced with respect to the beam (or the beam with respect to the patient), the entire

---

[2]Flux is defined as particles per second, $\mathrm{d}N/\mathrm{d}t$ [s^{-1}] [18].

dose distribution will be shifted. Depending on the amplitude of misalignment, this can increase the dose given to healthy tissue and reduce the dose given to the tumor.

For proton or carbon ion therapy, **range uncertainty** is of high importance. One element of uncertainty stems from converting Hounsfield units from CT-images to water equivalent density. This uncertainty is typically about 1-2% (1 s.d.) [4, Sec. 6.4.6.1.4], which corresponds to a range uncertainty of 1-2 mm for proton beam with 10 cm water equivalent range. Uncertainties in the CT-measurements themselves can add another 1-2% of uncertainty. Furthermore, **density heterogeneities** along the beam path can degrade the sharp distal falloff of the Bragg peak dramatically, as demonstrated by Urie et.al [19]. The main effects of this is that (a) a relatively high dose can be deposited beyond the assumed location of the most distal Bragg peak - something to carefully consider if the orientation of the incident beam is chosen such that the distal part of the target is adjacent to a critical organ - and (b) the flatness of the Spread Out Bragg Peak (SOBP) is deteriorated.

The capability of ions to spare healthy tissue distal to the target from dose, compared to photons, and the range uncertainty can be summarized in: "The advantage of ions is that they stop. The disadvantage is that we don't always know where" (Anthony Lomax, PSI [20]).

Range uncertainty can also be caused by **beam energy uncertainties**: an SOBP is generated by a superposition of many individual Bragg curves, weighted such that the total (biological) dose is flat in the target region (see Sec. 3.3.6). Systematic or random errors of the entrance energy of each individual Bragg curve can cause dose inhomogeneities along the SOBP.

In active scanning, each iso-energy layer is divided into several spots, the distance between two neighboring spots being smaller than the Full Width at Half Maximum (FWHM) of the beam. The treatment planning system assigns each spot a transverse width, weight (number of particles) and a lateral coordinate (x,y) to produce a homogeneous dose in the target. If any of the delivered spots deviates from one or several of these prescribed quantities due to imperfections of the **BDS**, the target dose distribution will be inhomogeneous and/or systematically over- or under-dosed. Inhomogeneities could thus be caused by e.g. inaccurate scanning magnets, while a systematic over- or under-dosage could be the effect of wrongly calibrated beam intensity monitors.

Based on the steepness of dose-response curves (i.e. the change in either tumor control or normal tissue damage due to a change in dose), Wambersie recommends a *total* dose uncertainty better than 3.5% (1 s.d.) at the "ICRU Reference Point", and 5% in other points. The ICRU (International Commission on Radiation Units and Measurements) Reference Point is a term that has been used in conventional radiotherapy for a long time and is defined as [21]:

- The dose at that point should be clinically relevant and representative of the dose distribution throughout the Planning Target Volume (PTV).

- The point should be easy to define in a clear and unambiguous way.

- The point should be selected where the dose can be accurately determined.

- The point should be in a region where there is no large dose gradient.

The purpose of the ICRU Reference Point is to simplify prescription and reporting of dose. The "ICRU Reference Dose" is simply the dose to be delivered to the ICRU Reference Point. For conventional radiotherapy (photons), recommendations for choosing the Reference Point are [22]:

- For a single beam: the point on the central axis at the center of the target volume.

- For parallel opposed equally weighted beams: the point on the central axis midway between the beam entrance points.

- For parallel opposed unequally weighted beams: the point on the central axis at the center of the target volume.

- For other combinations of intersecting beams: the point at the intersection of the central axes (insofar as there is no dose gradient at this point).

These recommendations are adapted to conventional radiotherapy, where the incoming radiation field is uniform in intensity and energy, and compensation for e.g. patient contour irregularities and "missing tissue" is made by use of e.g. wedges and boluses [22]. However, more modern irradiation techniques, such as Photon Intensity Modulated Radiation Therapy (IMRT), offer an "improved dose homogeneity inside the target volume and the potential for limited irradiation of surrounding sensitive structures" [22, Sec. 15.7.6]. In IMRT, the treatment is performed using a large number of beams from different angles. Each individual beam delivers a non-uniform dose to the PTV, but together they produce a (near-)uniform dose within the PTV [4, Sec. 6.6.2.2]. With IMRT, it is therefore difficult to choose a point in the target volume which is "representative of the dose distribution throughout the PTV" [23]. Dose prescription and reporting is therefore moving towards a volumetric approach, using e.g. dose-volume histograms and requiring that e.g. the entire PTV should receive at least 95% of the prescribed dose and/or that the homogeneity of the planned dose distribution should be such that 95% of the PTV receives within -5% to +7% of the prescribed dose [4, Sec. 6.6.1.3]. This is the case for photon IMRT, but even more so for actively scanned protons beams, which have an extra degree of freedom - the penetration depth - compared to photon beams.

To conclude, the relaxed conditions in other points than the ICRU Reference Point, as proposed by Wambersie, will not be adopted by the active scanning facility MedAustron. Instead, an overall dose homogeneity requirement of 3.5% (1 s.d.) in all points of the target volume is aimed for, taking both Type A and Type B uncertainties into account. Treating all uncertainties as independent random variables, the combined, or total, uncertainty, σ_{tot}, can be estimated by quadratic addition of all sources of dose uncertainty [17]:

$$\begin{aligned}\sigma_{tot} &= \sqrt{\sigma_{\text{CT-conversion}}^2 + \sigma_{\text{Patient positioning}}^2 + \sigma_{\text{Energy errors}}^2 + \sigma_{\text{BDS}}^2 + ...} \\ &\leq 3.5\%\end{aligned} \tag{2.1}$$

A thorough evaluation of all the individual components contributing to σ_{tot} is beyond the scope of this thesis. Focus in put on uncertainties originating from imperfections of vital parts of the beam delivery system:

- Beam energy errors

- Spot weight errors

- Beam size errors

- Beam positioning errors

The impact of these types of errors on dose homogeneity will be investigated, based on the specific properties of the MedAustron facility: synchrotron extraction mechanism technique, beam properties at the isocenter and beam delivery nozzles. Once these relations are known, acceptable error tolerances can be defined. These error tolerances form the basis for performance requirements of e.g. synchrotron extraction stability and beam delivery system. But in order for these requirements to be meaningful, they must also be verifiable: this means that specifications must also be made on the BVS, which is responsible for verifying the properties of the scanned beam during patient irradiation. The resulting dose distribution from the measured beam properties should still satisfy the clinical requirements on dose homogeneity. This procedure is outlined in Fig. 2.1.
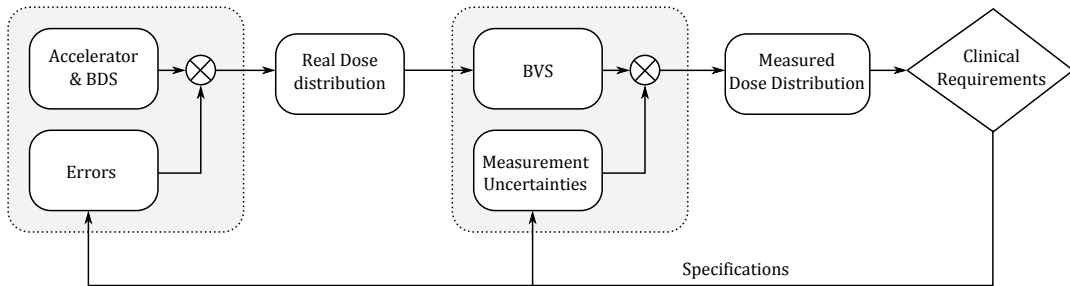
Figure 2.1: Schematic requirement specification process.

2.3.1 Uncertainty weighting

As seen in Eq. 2.1, the overall uncertainty of the target dose is given by the quadratic sum of all contributions, σ_i. Two contributors are considered in this thesis: σ_{BDS} and $\sigma_{\Delta E}$, which concern the following beam properties, discussed in the previous section:

$$
\begin{aligned}
\sigma_{\mathrm{BDS}} : &\left\{ \begin{array}{l} \text{Beam positioning} \\ \text{Beam size} \\ \text{Spot weight} \end{array} \right. \\
\sigma_{\Delta E} : &\left\{ \text{Energy errors} \right.
\end{aligned}
\tag{2.2}
$$

13

Setting an upper limit on one source of uncertainty without knowing the magnitude of all the other is impossible. However, as a starting point, the following limits have been set:

$$\sigma_{\text{BDS}} \leq 2.0\%$$
$$\sigma_{\Delta E} \leq 1.0\% \tag{2.3}$$

These numbers are chosen somewhat arbitrarily, but:

- They leave a margin of $\sqrt{3.5^2 - 2.0^2 - 1.0^2} = 2.7\%$ to other sources of dose uncertainty, which is not a severe restriction compared to the maximum 3.5%.

- Choosing too small values would be pointless, since σ_{tot} then would either be unnecessarily small, or completely dominated by other sources of uncertainty.

Nevertheless, the exact numbers chosen in Eq. 2.3 can always be discussed, and should be taken as a *starting point of an iterative procedure* where deduced requirements on e.g. synchrotron extraction stability or scanning magnet accuracy are checked against technical feasibility. Specifications that can be tightened without any (significant) additional cost leave room for a relaxation on other parts.

For this reason, dose error analysis will be done in a manner such that all resulting specifications easily can be rescaled to another choice of initial constraints.

2.4 Range accuracy

An energy error ΔE of the extracted beam will cause an error in range, ΔR, of the beam as it is stopped in the patient, causing the Bragg peak to be shifted toward the surface or deeper into the patient. The effect of an energy error would be largest for the most distal, high weighted, Bragg curve of an SOBP: if the energy is too high, the sharp Bragg peak could be moved into healthy tissue behind the tumor; if the energy is too low, the distal part of the target could be severely under-dosed. The sharper the Bragg peak (carbon ions, low energy protons), the stronger the effect of an energy error.

Apart from extraction energy errors, uncertainties in conversions from CT-images to beam range also contribute to uncertainties of the beam range in the patient. In order not to risk e.g. irradiation of critical organs in the proximity of the tumor, the beam entrance angle(s) is preferably chosen such that critical organs are not at risk in case of beam range errors.

Beam energy errors can not only damage healthy organs distal to the tumor, or cause significant underdosage in the distal part of the tumor; they will also deteriorate the homogeneity of the SOBP, since the peak-to-peak distances of the Bragg peaks no longer correspond to the planned ones.

The beam range accuracy required varies between different centers. One example is CNAO, which in their functional specifications [24] require an accuracy of ± 0.025 g/cm^2 (± 0.25 mm in water), one quarter of the minimum range adjustment step.

Another example is the proton therapy facility Loma Linda [25], which describe two differently deduced energy precision requirements in [26]. On the one hand, the range accuracy is required to be better than 0.5 mm for *patch fields*. Patch fields are combinations of fields where the distal edge of one field overlaps the lateral edge of another field. On the other hand, the *relative* energy accuracy - i.e. the maximum energy error when switching from one energy to the next when building up the SOBP - specified by the same center varies between ± 0.1 MeV for the sharp, low-energy Bragg peaks, to ± 0.25 MeV for the broadened, high-energy Bragg peak.

While the conversion from range accuracy (given in g/cm^2 or mm in water) to energy accuracy is straightforward (see Sec. 3.3.3), energy accuracy requirements for energy stacking requires a more careful analysis, since it will depend on the shape of the Bragg peak and the ridge filter used (carbon ions). An analysis of this is made in Sec. 6.3.2.2

Chapter 3

Theory and main concepts

3.1 Accelerator physics

3.1.1 Coordinate system and nomenclature

The horizontal, vertical and longitudinal coordinates of a single particle, with respect to the trajectory of an ideal zero-amplitude, on-momentum particle, in the synchrotron or along a transfer line are denoted x, y and s, respectively. The transverse direction of a single particle is defined as the derivative of x and y with respect to s:

$$x' = \frac{dx}{ds} \qquad y' = \frac{dy}{ds} \tag{3.1}$$

The relative momentum offset of a single particle is denoted $\frac{\delta p}{p}$ (zero for on-momentum particles). Each particle i is represented by the vector $\vec{r}_i$:

$$\vec{r}_i = (x_i, x_i', y_i, y_i', \delta p_i/p)^{\mathrm{T}} \tag{3.2}$$

In this section, z and z' are general coordinates used for either x and x' or y and y'.

3.1.2 Betatron motion

The transverse motion of a single particle around the equilibrium orbit in the synchrotron is called betatron motion, and has been parameterized by Courant and Snyder [27]:

$$z(s) = A_z \sqrt{\beta_z(s)} \cos \left[\int_0^s \frac{1}{\beta_z(s)} ds + \varphi_z \right] \tag{3.3}$$

A_z and ϕ_z are two constants determined by the starting coordinate of the particle. $\beta_z(s)$ is defined by the optics of the synchrotron. The three derived auxiliary vari-

ables, $\alpha_z(s)$, $\gamma_z(s)$ and $\mu_z(s)$, to be used later, are defined as[1]:

$$\alpha_z(s) = -\frac{1}{2}\frac{d\beta_z}{ds} \tag{3.4}$$

$$\mu_z(s) = \int_0^s \frac{1}{\beta_z(s)}ds \tag{3.5}$$

$$\gamma_z(s) = \frac{1+\alpha_z^2}{\beta_z} \tag{3.6}$$

$$\tag{3.7}$$

with which the oscillatory particle motion in the transverse plane (z,z') can be written as:

$$z(s) = A_z\sqrt{\beta_z(s)}\cos(\mu_z(s)+\varphi_z) \tag{3.8}$$

$$z'(s) = -\frac{A_z}{\sqrt{\beta_z(s)}}\left[\alpha_z(s)\cos(\mu_z(s)+\varphi_z)+\sin(\mu_z(s)+\varphi_z)\right] \tag{3.9}$$

The normalized coordinates (Z,Z') are defined as:

$$Z = A_z\cos(\mu_z(s)+\varphi_z) \tag{3.10}$$

$$Z' = -A_z\sin(\mu_z(s)+\varphi_z) \tag{3.11}$$

and in normalized phase space, (Z,Z'), it is clear that all particles with a given amplitude A_z will lie on a circle of radius A_z. A_z is denoted the (normalized) particle amplitude. The transformation from real phase space to normalized phase space is linear and can be written as (omitting the argument s):

$$\begin{pmatrix} Z \\ Z' \end{pmatrix} = \begin{pmatrix} 1/\sqrt{\beta_z} & 0 \\ \alpha_z/\sqrt{\beta_z} & \sqrt{\beta_z} \end{pmatrix} \begin{pmatrix} z \\ z' \end{pmatrix} \tag{3.12}$$

or, $\vec{Z} = B\vec{z}$. The inverse transformation, $\vec{z} = B^{-1}\vec{Z}$, is given by:

$$\begin{pmatrix} z \\ z' \end{pmatrix} = \begin{pmatrix} \sqrt{\beta_z} & 0 \\ -\alpha/\sqrt{\beta} & 1/\sqrt{\beta_z} \end{pmatrix} \begin{pmatrix} Z \\ Z' \end{pmatrix} \tag{3.13}$$

Fig. 3.1 shows the possible locations of particles of a given amplitude A_z in real and normalized phase space. Since $\det(B) = 1$, the area enclosed by the circle in normalized phase space is preserved by the linear transformation. This area is called the single-particle emittance, ε_z, and, since it depends only on the particle amplitude A_z, it is constant in both phase spaces:

$$\varepsilon_z = \pi A_z^2 \tag{3.14}$$

Via Eq. 3.12, this gives a motion invariant in real phase space:

$$\frac{\varepsilon_z}{\pi} = Z^2 + Z'^2 = \ldots = \beta_z z'^2 + 2\alpha_z zz' + \gamma z^2 \tag{3.15}$$

where the latter is the equation for an ellipse.

Note that particles with the same amplitude A_z but different phase φ_z will follow the same trajectory.

---

[1] $\alpha(s)$, $\beta(s)$ and $\gamma(s)$ are called Twiss functions.

(a) Real phase space, (z, z').
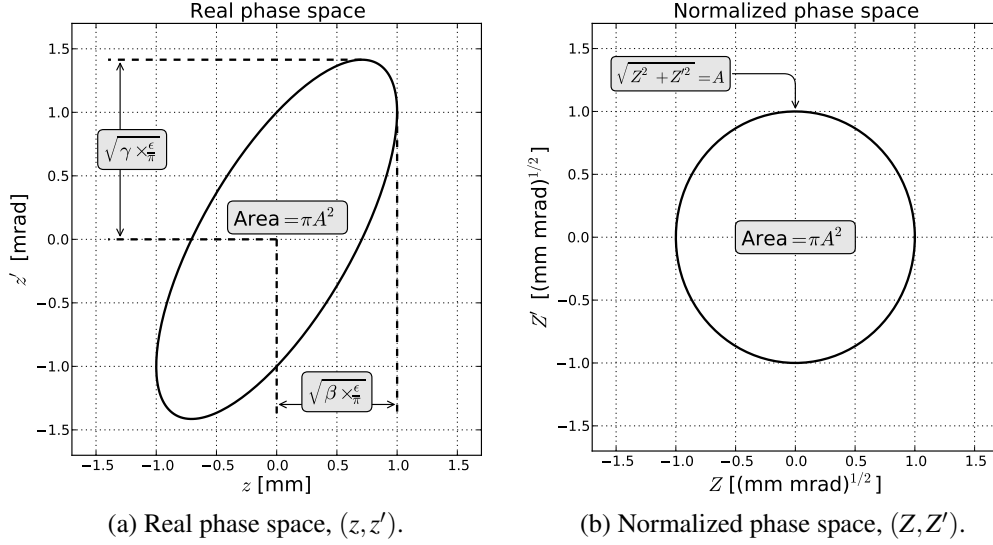
(b) Normalized phase space, (Z, Z').

Figure 3.1: Permissible coordinates for particles of a given amplitude A_z in real and normalized phase space. The single particle emittance is $1.0\,\pi$ mm mrad, $\beta_z = 1.0$ m and $\alpha_z = 1$ (subscript z omitted in figure).

3.1.3 Gaussian beam distribution

Prior to extraction, the distribution of particles in the synchrotron is well described by a two-dimensional Gaussian in the horizontal and in the vertical plane. The density of particles in either plane, denoted $\rho_z(z, z')$, is in normalized phase space given by:

$$\rho_z(Z, Z') = \frac{1}{2\pi\sigma_{norm}^2} \exp\left(-\frac{Z^2 + Z'^2}{2\sigma_{norm}^2}\right) \tag{3.16}$$

which in real phase space transforms to:

$$\rho_z(z, z') = \frac{1}{2\pi\sigma_{norm}^2} \exp\left(-\frac{\gamma_z z^2 + 2\alpha_z z z' + \beta z'^2}{2\sigma_{norm}^2}\right) \tag{3.17}$$

where α_z, β_z and γ_z are the Twiss parameters, which describe the tilt, stretch and height of the elliptical phase space distribution, and σ_{norm} defines the width of the distribution in phase space. An example of a particle distribution is shown in Fig. 3.2, where color indicates particle density. The isoline $\sqrt{Z^2 + Z'^2} = \sigma_{norm} \Leftrightarrow \rho_z = e^{-1/2}/2\pi\sigma_{norm}^2$ has been indicated as a black curve.

The area enclosed by this isocurve is called the 1σ-emittance of the Gaussian beam, denoted ε_z. Note that the emittance is here defined as the total area enclosed by the curve:

$$\sqrt{Z^2 + Z'^2} = \sigma_{norm} \tag{3.18}$$

In other literature, the emittance can be defined as the enclosed area divided by π. This thesis follows the recommendation by e.g. Phil Bryant [28] and explicitly
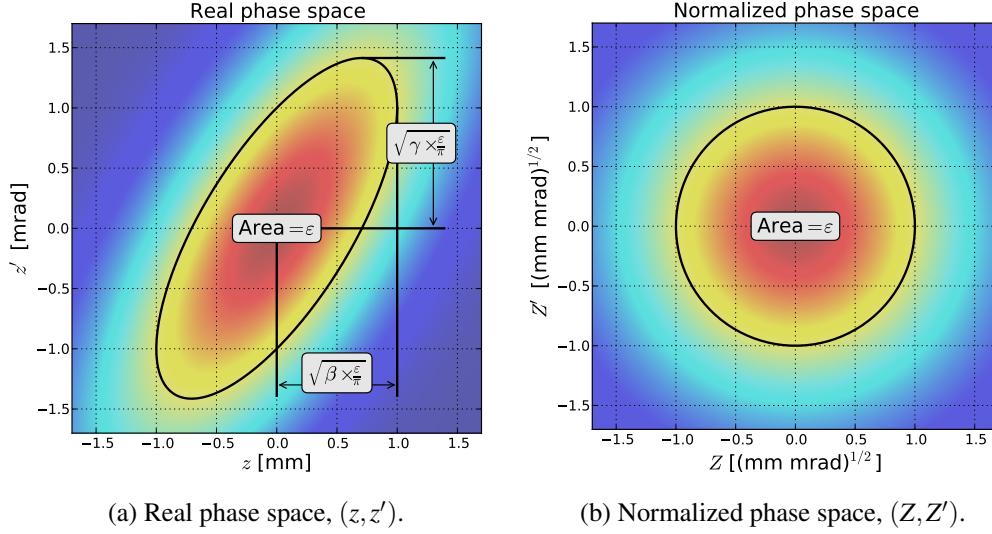
(a) Real phase space, (z, z').   (b) Normalized phase space, (Z, Z').

Figure 3.2: Example of beam distribution in a transverse plane. The 1σ-emittance, ε_z, is $1.0\ \pi$ mm mrad, $\beta_z = 1.0$ m and $\alpha_z = 1$ (subscript z omitted in figure). The solid line indicates the isolevel where $\rho = e^{-0.5}/2\pi\sigma_{norm}^2$.

writes out the emittance in units of π mm mrad, or π m rad. The factor π shows that emittance is defined as the *area* of the circle, and not the radius squared.

A projection of the beam distribution on the z-axis gives the 1σ-beam width:

$$\sigma_z = \sqrt{\beta_z \times \frac{\varepsilon_z}{\pi}} \tag{3.19}$$

while a projection on the z'-axis gives the 1σ-beam divergence:

$$\sigma_{z'} = \sqrt{\gamma_z \times \frac{\varepsilon_z}{\pi}} \tag{3.20}$$

Note that the division by π simply cancels out the factor π present in the unit of the emittance (π m rad or π mm mrad).

If not explicitly mentioned otherwise, the term emittance always means the 1σ-emittance. One could equally well choose the area of e.g. the circle $\sqrt{Z^2 + Z'^2} = \sqrt{5}\sigma_{norm}$ as an emittance definition, which would include a larger fraction of the beam.

3.1.4 Amplitude distribution

Particles with the same amplitude $A_z = \sqrt{Z^2 + Z'^2}$ are distributed along a circle with radius A_z in normalized phase space. The number of particles within the amplitude range $[A_z, A_z + dA_z]$ (dA_z small) is given by the number of particles in the circular band with width $\int dA_z$ and radius A_z. The amplitude distribution $\rho(A_z)$ can thus be

20

calculated from the 2-dimensional Gaussian distribution $\rho(Z,Z')$:

$$\rho(A_z)dA_z = 2\pi A_z dA_z \times \rho(\sqrt{Z^2 + Z'^2} = A_z) \Rightarrow$$

$$\rho(A_z) = \frac{A_z}{\varepsilon_z/\pi} \exp\left(-\frac{A_z^2}{2 \times \varepsilon_z/\pi}\right) \tag{3.21}$$

where ε_z/π is the transverse 1σ-emittance of the beam (horizontal or vertical). The amplitude distribution is shown schematically in Fig. 3.3a for a truncation level at $A_z = \sqrt{5\varepsilon_z/\pi}$. Particles with a single-particle emittance larger than $5\varepsilon_z$ are assumed to be collimated.

3.1.5 Momentum distribution

The momentum spread of the beam, just before extraction, is assumed to be uniform, with a total spread of $\Delta p/p$, as shown in Fig. 3.3b.
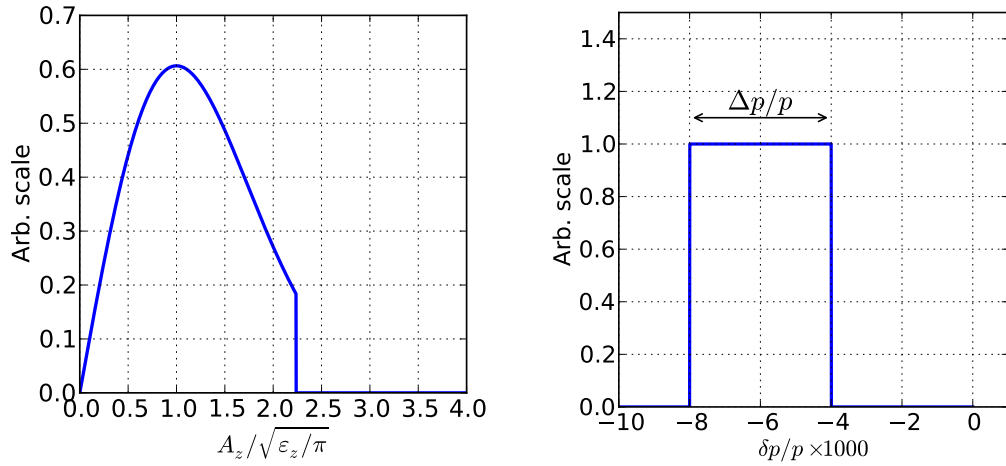
(a) Amplitude distribution, $\rho(A_z)$, truncated at $A_z = \sqrt{5\varepsilon_z/\pi}$.

(b) Uniform momentum distribution. $\Delta p/p$ denotes the momentum spread of the entire beam.

Figure 3.3: Amplitude and momentum distribution examples.

3.1.6 Transfer matrix formalism

Since Eq. 3.3 governs the motion of all particles along the synchrotron, the Twiss functions along the synchrotron can be used to calculate the particle motion from $s = s_1$ to $s = s_2$. The transfer matrix of (z,z') from s_1 to s_2 for uncoupled motion is given by [28]:

$$\begin{pmatrix} z \\ z' \end{pmatrix}_{s_2} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \begin{pmatrix} z \\ z' \end{pmatrix}_{s_1} = M_{s_1 \to s_2} \vec{z}_1 \tag{3.22}$$

where the elements of the transfer matrix are (subscript z omitted):

$$m_{11} = \sqrt{\frac{\beta_2}{\beta_1}} \left(\cos(\Delta\mu) + \alpha_1 \sin(\Delta\mu) \right)$$

$$m_{12} = \sqrt{\beta_1 \beta_2} \sin(\Delta\mu)$$

$$m_{21} = - \frac{1}{\sqrt{\beta_1 \beta_2}} \left((1 + \alpha_1\alpha_2) \sin(\Delta\mu) + (\alpha_2 - \alpha_1) \cos(\Delta\mu) \right)$$

$$m_{22} = \sqrt{\frac{\beta_1}{\beta_2}} \left(\cos(\Delta\mu) - \alpha_2 \sin(\Delta\mu) \right) \tag{3.23}$$

Subscript 1 or 2 on β and α indicates their value at s_1 or s_2 and $\Delta\mu$ is the phase advance from s_1 to s_2:

$$\Delta\mu = \mu(s_2) - \mu(s_1) \tag{3.24}$$

When applied to a single turn around the synchrotron (i.e. $\alpha_1 = \alpha_2$, $\beta_1 = \beta_2$), the transfer matrix becomes, in real phase space (subscript z):

$$M_z^1(Q_z) = \begin{pmatrix} \cos(2\pi Q_z) + \alpha \sin(2\pi Q_z) & \beta \sin(2\pi Q_z) \\ -\gamma \sin(2\pi Q_z) & \cos(2\pi Q_z) - \alpha \sin(2\pi Q_z) \end{pmatrix} \tag{3.25}$$

where $2\pi Q_z = \Delta\mu$, i.e. the phase advance over one turn. Q is the (horizontal or vertical) betatron tune of the synchrotron: the number of betatron oscillations a single particle makes per turn. In normalized phase space (subscript Z), the corresponding transfer matrix for a single turn is:

$$M_Z^1(Q_z) = B M_z^1 B^{-1} = \ldots = \begin{pmatrix} \cos(2\pi Q_z) & \sin(2\pi Q_z) \\ -\sin(2\pi Q_z) & \cos(2\pi Q_z) \end{pmatrix} \tag{3.26}$$

which is the standard matrix for a *clock-wise* rotation $2\pi Q$ radians. Thus, the transfer matrix for n turns around the synchrotron in normalized phase space coordinates is:

$$M_Z^n(Q_z) = \begin{pmatrix} \cos(2\pi n Q_z) & \sin(2\pi n Q_z) \\ -\sin(2\pi n Q_z) & \cos(2\pi n Q_z) \end{pmatrix} \tag{3.27}$$

3.1.7 Dispersion

The magnetic rigidities of off-momentum particles are different to on-momentum particles. The deflection angle of an off-momentum particle through e.g. a dipole will be different from that of an on-momentum particle. The orbit of a zero-amplitude, off-momentum particle will therefore be different from the reference orbit. The ratio between the phase space displacement and momentum deviation is the *dispersion vector*, (D_z, D_z'):

$$\Delta z = D_z \frac{\delta p}{p} \tag{3.28}$$

$$\Delta z' = D_z' \frac{\delta p}{p} \tag{3.29}$$

3.1.8   Chromatic tune shift

The deflection and focusing strength in the synchrotron magnets depends on the energy of the particle. Therefore, off-momentum particles will behave differently to on-momentum particles. The *chromaticity* (horizontal or vertical) of the synchrotron is the ratio between the tune shift of an off-momentum particle and its relative momentum deviation, $\delta p/p$:

$$Q'_z = \frac{\Delta Q_z}{\delta p/p} \tag{3.30}$$

where ΔQ is the difference between the particle tune and the synchrotron tune (on-momentum particle tune). The momentum dependent tune of a single particle can thus be written as:

$$Q_z = Q_{res} + Q'_z \frac{\delta p}{p} \tag{3.31}$$

assuming an on-momentum, zero-amplitude, particle ($\delta p/p = 0$) has the resonance tune, Q_{res}.

The MedAustron synchrotron has a negative horizontal chromaticity, i.e. during acceleration (increasing $\delta p/p$) the horizontal tune of the beam *decreases*. Before extraction, the horizontal tune of the beam is thus *above* the resonance tune.

3.2   Third integer resonance extraction

3.2.1   Resonance sextupole

The magnetic elements of the MedAustron synchrotron are tuned such that an on-momentum particle has a horizontal tune close to 5/3 ($Q_x = 1.6666$). The beam is extracted by using a resonance sextupole, which introduces a perturbation to the machine. Every time a particle passes through the resonance sextupole, it will receive a "kick" that shifts its horizontal and vertical direction (x' and y') by $\Delta x'$ and $\Delta y'$.

The horizontal and vertical components of a magnetic field , B_x, B_y, can always be written as the differentiation of scalar magnetic potiential Φ:

$$B_x(x,y) = -\frac{\partial \Phi}{\partial x}$$
$$B_y(x,y) = -\frac{\partial \Phi}{\partial y} \tag{3.32}$$

The scalar potential of a field with $2m$ poles can be written as the sum of two components:

$$\Phi(x,y) = \underbrace{A_m \Re\{(x+iy)^m\}}_{\text{Skew component}} + \underbrace{B_m \Im\{(x+iy)^m\}}_{\text{Normal component}} \tag{3.33}$$

where the relative weights of the terms A_m and B_m determines the orientation of the field in the transverse plane. For a normal magnet, $A_m = 0$, while for a skew magnet

$B_m = 0$. The magnetic field components of the resonance sextupole ($m = 3$, normal magnet) thus are:

$$B_x(x,y) = -B_3 \frac{\partial}{\partial x}(3x^2 y - y^3) = -6B_3 \times xy$$

$$B_y(x,y) = -B_3 \frac{\partial}{\partial y}(3x^2 y - y^3) = -3B_3 \times (x^2 - y^2) \tag{3.34}$$

A Taylor expansion of e.g. $B_y(x, y = 0)$ around $x = 0$ gives the term B_3:

$$-3B_3(x^2 - 0^2) \equiv \sum_{k=0}^{\infty} \frac{x^k}{k!}\left(\frac{\partial^k B_y}{\partial x^k}\right)_0 \Rightarrow B_3 = -\frac{1}{6}\left(\frac{\partial^2 B_y}{\partial x^2}\right)_0 \tag{3.35}$$

and the horizontal and vertical magnetic fields can be rewritten as:

$$B_x(x,y) = \left(\frac{\partial^2 B_y}{\partial x^2}\right)_0 xy$$

$$B_y(x,y) = \frac{1}{2}\left(\frac{\partial^2 B_y}{\partial x^2}\right)_0 (x^2 - y^2) \tag{3.36}$$

The bending angle α of a positively charged particle passing through a dipole field of strength B and length l is, using a thin lens approximation:

$$\alpha = \frac{Bl}{|B\rho|} \tag{3.37}$$

$|B\rho|$ is the magnetic rigidity of the particle, related to the particle momentum p and charge number q as:

$$|B\rho| = \frac{p}{q} \quad \text{[Tm]} \tag{3.38}$$

Since the variation of the transverse displacement in a sextupole is negligible, a thin lens approximation can be used. The "kick" (change in direction, $\Delta x'$ and $\Delta y'$) is given by [11, 29]:

$$\Delta x' = \frac{1}{2}\frac{l_s}{|B\rho|}\left(\frac{\mathrm{d}^2 B_y}{\mathrm{d}x^2}\right)_0 (x^2 - y^2) = \frac{1}{2}l_s k' \times (x^2 - y^2)$$

$$\Delta y' = -\frac{l_s}{|B\rho|}\left(\frac{\mathrm{d}^2 B_y}{\mathrm{d}x^2}\right)_0 xy = -l_s k' \times xy \tag{3.39}$$

where l_s is the length of the sextupole magnet and k' is the normalized sextupole gradient:

$$k' = \frac{1}{B\rho}\left(\frac{\mathrm{d}^2 B_y}{\mathrm{d}x^2}\right)_0 \tag{3.40}$$

Combining Eq. 3.12 and Eq. 3.40 gives the kick in normalized phase space:

$$\Delta X' = \beta_x^{3/2}\frac{l_s k'}{2}\left(X^2 - \frac{\beta_y}{\beta_x}Y^2\right) = S\left(X^2 - \frac{\beta_y}{\beta_x}Y^2\right)$$

$$\Delta Y' = -\beta_y \beta_x^{1/2} l_s k' XY = -2S\frac{\beta_y}{\beta_x}XY$$

$$\tag{3.41}$$

where the normalized sextupole strength, S, has been introduced:

$$S = \frac{1}{2}\beta_x^{3/2} l_s k' \tag{3.42}$$

The resonance sextupole will introduce a coupling between the horizontal and vertical plane if $Y > 0$. However, in the MedAustron synchrotron, Y is small compared to X at the resonance sextupole and, to a first order, the vertical motion can be neglected[2]. The simplified normalized sextupole kick thus becomes:

$$\begin{aligned} \Delta X' &= SX^2 \\ \Delta Y' &= 0 \end{aligned} \tag{3.43}$$

3.2.2 Kobayashi Hamiltonian

The horizontal tune of the MedAustron synchrotron is close to 5/3, i.e. after three turns, a particle is close to previous phase space coordinates. Let the horizontal tune of a particle be:

$$Q = Q_{res} + \delta Q \tag{3.44}$$

where $Q_{res} = 5/3$ is the resonance tune and δQ a small tune deviation. The operator $\hat{S}$ is the effect of the resonance sextupole on a single particle, i.e:

$$\hat{S}\begin{pmatrix} X \\ X' \end{pmatrix} = \begin{pmatrix} X \\ X' \end{pmatrix} + \begin{pmatrix} 0 \\ SX^2 \end{pmatrix} \tag{3.45}$$

The shift of a particle's position in horizontal phase space over three turns can, to a first order, be estimated by adding the effect of:

1. A non-zero δQ over three turns (if zero, the particle returns to its original position)

2. the sextupole kick after the first turn (neglecting δQ)

3. the sextupole kick after the second turn (neglecting δQ)

4. the sextupole kick after the third turn (neglecting δQ)

The total coordinate shift over three turns, $\Delta \vec{X}$, can be calculated from:

$$\begin{aligned} \Delta\vec{X} = & M_Z^3(Q_{res} + \delta Q)\vec{X} - \vec{X} + \\ & M_Z^2(Q_{res})\hat{S}\left\{M_Z^1(Q_{res})\vec{X}\right\} - \vec{X} + \\ & M_Z^1(Q_{res})\hat{S}\left\{M_Z^2(Q_{res})\vec{X}\right\} - \vec{X} + \\ & \hat{S}\left\{M_Z^3(Q_{res})\vec{X}\right\} - \vec{X} = \ldots \\ = & \begin{pmatrix} 6\pi\delta Q X' + \frac{3}{2}SXX' \\ -6\pi\delta Q X + \frac{3}{4}S(X^2 - X'^2) \end{pmatrix} \end{aligned} \tag{3.46}$$

---

[2]The optics has been designed to achieve exactly this.

The term $6\pi\delta Q$ is called *modified tune distance* and is in this document denoted ξ[3]:

$$\xi = 6\pi\delta Q \qquad (3.47)$$

Thus, the change of a particle's coordinates in horizontal phase space over three turns is, to first order:

$$\Delta X_{3\text{ Turns}} = \xi X' + \frac{3}{2}SXX' \qquad (3.48)$$

$$\Delta X'_{3\text{ Turns}} = -\xi X + \frac{3}{4}S(X^2 - X'^2) \qquad (3.49)$$

where $\Delta X_{3\text{ Turns}}$ is known as the *spiral step* and $\Delta X'_{3\text{ Turns}}$ the *spiral kick*. During slow extraction, the spiral step and kick are small and the time needed for a particle to travel around the synchrotron is short compared to the overall spill time. One can therefore treat X and X' as continuous variables with a time derivative of (omitting the subscript "3 Turns"):

$$\dot{X} \approx \frac{\Delta X}{\Delta t'} = \xi X' + \frac{3}{2}SXX'$$
$$\dot{X}' \approx \frac{\Delta X'}{\Delta t'} = -\xi X + \frac{3}{4}S(X^2 - X'^2) \qquad (3.50)$$

where the time t' is time in units of three times the revolution period:

$$t' = \frac{t}{3 \times t_{rev}} \qquad (3.51)$$

Now, assume there is a Hamiltonian $\mathcal{H}$ of which X and X' are the canonical variables. Then [30]:

$$\dot{X} = \frac{\partial \mathcal{H}}{\partial X'}$$
$$\dot{X}' = -\frac{\partial \mathcal{H}}{\partial X} \qquad (3.52)$$

$\mathcal{H}(X, X')$ can be resolved by partial integration of Eq. 3.50 (setting the integration constant to zero):

$$\mathcal{H}(X, X') = \frac{\xi}{2}(X^2 + X'^2) + \frac{S}{4}(3XX'^2 - X^3) \qquad (3.53)$$

This is the Kobayashi Hamiltonian.

---

[3]In other literature the letter ε is used [11] for modified tune distance, but to avoid confusing it with the emittance, the symbol ξ is used in this document.

3.2.3 Stable region

Of particular interest are the isocurves of the Kobayashi Hamiltonian where $\mathcal{H} = (2\xi/3)^3/S^2$, for which the equation is [29]:

$$\left(\frac{S}{4}X + \frac{\xi}{6}\right)\left(\sqrt{3}X' + X - \frac{4\xi}{3S}\right)\left(\sqrt{3}X' - X + \frac{4\xi}{3S}\right) = 0 \qquad (3.54)$$

This is the equation for three straight lines, forming a triangular region. These lines are called the *separatrices* and the size of the triangular region is determined by the ratio $|\xi/S|$.

The particle trajectories in normalized phase space will follow the isocurves of $\mathcal{H}$, shown in Fig. 3.4. Inside the triangular region, the curves are closed - the particles are stable. Outside the triangle, the trajectories continue to infinity - the particles are unstable, as they will spiral outwards from the origin. There are four fixed points $P_0 - P_3$ where $\dot{X} = \dot{X}' = 0$:

$$h = \frac{2\xi}{3S}; \quad \begin{cases} P_0 = (0,0) \\ P_1 = (2h,0) \\ P_2 = (-h, -\sqrt{3}h) \\ P_3 = (-h, \sqrt{3}h) \end{cases} \qquad (3.55)$$

and the area of the stable triangle is given by:

$$\varepsilon_{triangle} = 3\sqrt{3}h^2 = \frac{48\sqrt{3}\pi}{S^2}(\delta Q)^2\pi \qquad (3.56)$$

Particles that are close to the fixed points will move slower than particles that are in the mid-region of the sides of the triangle. Due to this effect, the density of particles will be higher at the corners of the triangle than along the sides of the triangle.

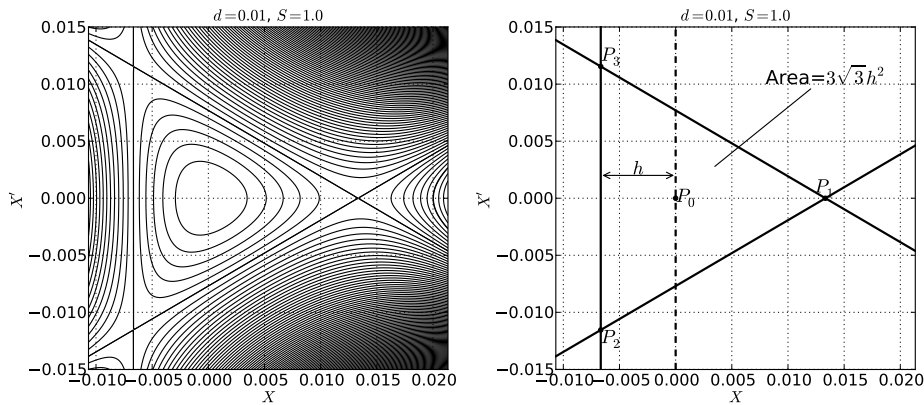Note that a change of sign of ξ will flip the triangle $180°$ in phase space.

Figure 3.4: Left: Particle trajectories in normalized horizontal phase space ($S = 1.0$, $\xi = 0.01$). Right: The stable triangle defined by the separatrices and fixed points $P_0 - P_3$.

The resonance sextupole is located in a dispersion-free region ($D_z = D'_z = 0$ in the horizontal and vertical plane) of the synchrotron.

3.2.4 Adiabatic ramping of resonance sextupole

The resonance sextupole will distort the elliptical phase space trajectories into triangular ones. In order to preserve the emittance, the sextupole must be ramped up to its nominal value S slowly enough that the elliptical beam distribution in horizontal phase space is "squeezed" into a triangular distribution. If the single particle emittances are preserved, the sextupole is ramped *adiabatically*. In order not to extract any particles during ramping, the full emittance of the beam $\varepsilon_{x,\text{beam}}$ must be smaller than the area of the stable triangle after ramping:

$$\varepsilon_{x,\text{beam}} \leq \varepsilon_{triangle} \tag{3.57}$$

An example of a single-particle trajectory during adiabatic and non-adiabatic ramping is shown in Fig. 3.5 (see Sec. 4.3 on particle tracking).

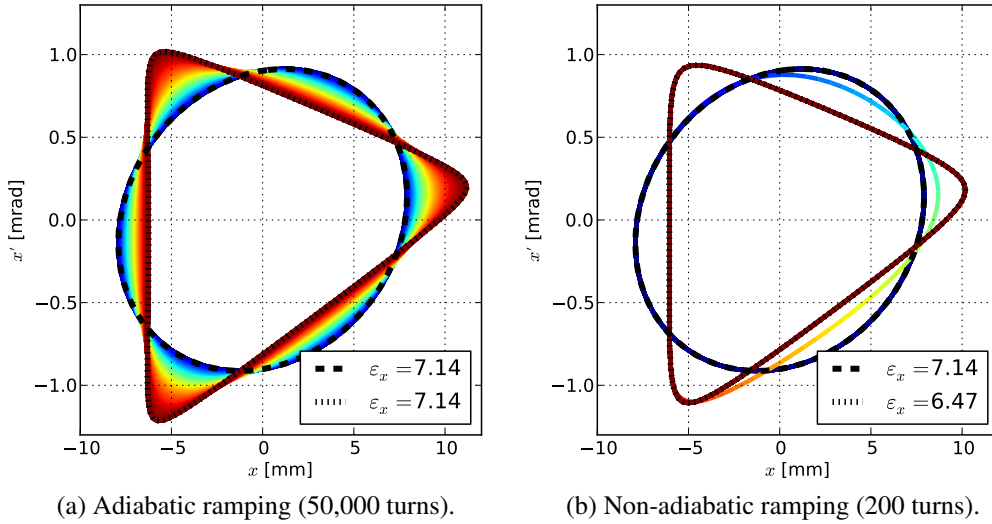(a) Adiabatic ramping (50,000 turns).　(b) Non-adiabatic ramping (200 turns).

Figure 3.5: Phase space trajectory of a single particle at the resonance sextupole during ramping of resonance sextupole. Closed trajectories before and after ramping are indicated with dashed and dotted lines. ε_x in π mm mrad. Color indicates sextupole strength.

3.2.5 Steinbach diagram

The area of the stable triangle in Fig. 3.4 depends on the modified tune distance $\xi = 6\pi\delta Q$. The factor δQ, in turn, depends on the particle momentum (Eq. 3.30):

$$\delta Q = Q'_x \frac{\delta p}{p} \tag{3.58}$$

Combining this with Eq. 3.56 gives the following stability criterion for a particle with emittance ε_x and momentum deviation $\delta p/p$:

$$\sqrt{48\sqrt{3}\pi} \left| \frac{Q'_x}{S} \frac{\delta p}{p} \right| \geq \sqrt{\varepsilon_x/\pi} \tag{3.59}$$

The linear relation between $|\delta p/p|$ and $\sqrt{\varepsilon_x/\pi} = A_x$ can be illustrated in a *Steinbach diagram* (Fig. 3.6), where the abscissa is the single-particle tune (or, via the chromaticity, momentum offset). The ordinate is the normalized particle amplitude. Particles inside the "V"-shaped region are unstable.
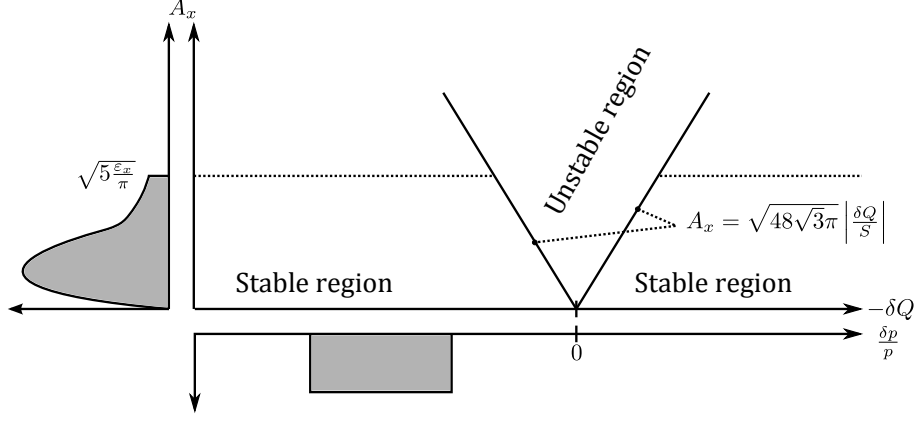

Figure 3.6: Steinbach diagram: graphical representation of Eq. 3.59. Particles with a given tune and amplitude A_x are stable only if they are outside the "V"-shaped region. Projected beam distribution in momentum and amplitude space shown below and to the left of the Steinbach diagram.

3.2.6 Acceleration-driven extraction

One way to extract the beam is to change its momentum (energy) such that the particles, via the chromaticity, are pushed into the resonance (the "V"-shaped region in the Steinbach diagram). In MedAustron and the Italian hadron therapy facility CNAO, this is done with an inductive element called betatron core. The principle is shown in Fig. 3.7: current coils parallel to the beam path induce a magnetic field $B(t)$, coaxial with the beam path. By gradually increasing the current, the magnetic field will increase and an electric field that accelerates the beam is induced via Maxwell's 3rd equation:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{3.60}$$

When a stable particle is accelerated, the stability criterion (Eq. 3.59) will at some point be violated. The particle is then inside the unstable region, spiraling outward. In principle, the particle can be anywhere along the triangular border when this happens, but as seen in Fig. 3.4, the trajectory of the particle will asymptotically approach one of the separatrices. The smaller the single-particle emittance, the closer to the resonance momentum the particle must be before it becomes unstable.

3.2.7 Electrostatic and magnetic septa

Downstream of the resonance sextupole is an electrostatic septum (ES) (see Fig. 3.8a), where a horizontal electric field $\mathbf{E} = E\hat{x}$, is present at $x > x_{ES}$. The stable particles
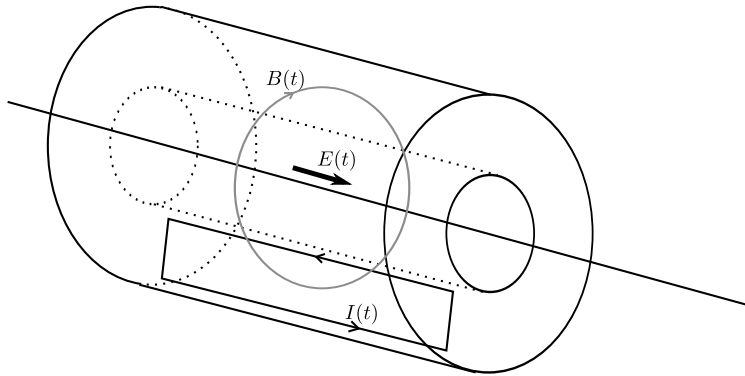
Figure 3.7: Betatron core principle: current coils parallel to the beam trajectory induces a coaxial magnetic field. Increasing the current induces an electric field parallel to the beam, which is accelerated.

are not affected by the septum, but unstable particles which have reached a horizontal coordinate larger than x_{ES} will be accelerated outwards, i.e. receive a "kick" φ in horizontal phase space, as illustrated in Fig. 3.9. The amplitude of the kick is given by:

$$\varphi = \frac{qE_{ES}L_{ES}}{pv} \tag{3.61}$$

where q is the charge of the beam, E_{ES} the electric field, L_{ES} the length of the field (along s). p and v are the beam momentum and velocity.

Further downstream, the phase advance has transformed this kick into a spatial jump, which creates a horizontal gap void of particles. At this location, a magnetic septum (MS) is inserted, which deflects the previously kicked particles into the High Energy Beam Transfer line (HEBT). The schematics of the magnetic septa is shown in Fig. 3.8b.

(a) Electrostatic septa, principle.
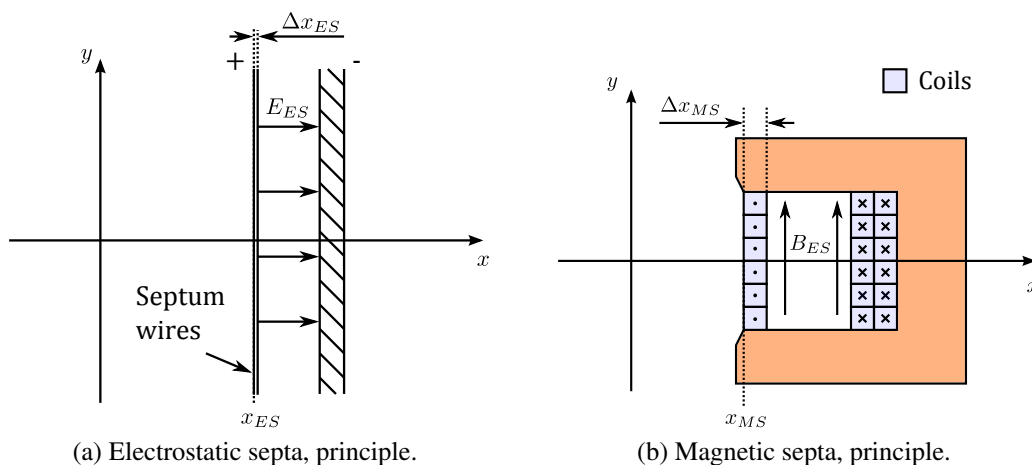

(b) Magnetic septa, principle.

Figure 3.8: Principle of electrostatic and magnetic septa.

Some particles will inevitably hit the wires of the electrostatic septum, and be lost. To maintain a high extraction efficiency, the thickness, Δx_{ES} must be small.

This limits the electric field that can be sustained, and a second kick by the magnetic septum is necessary to deflect the beam to the HEBT. However, if only a MS were used, the extraction efficiency would be reduced, since the inner coils of the MS are significantly thicker than the wires of the ES.
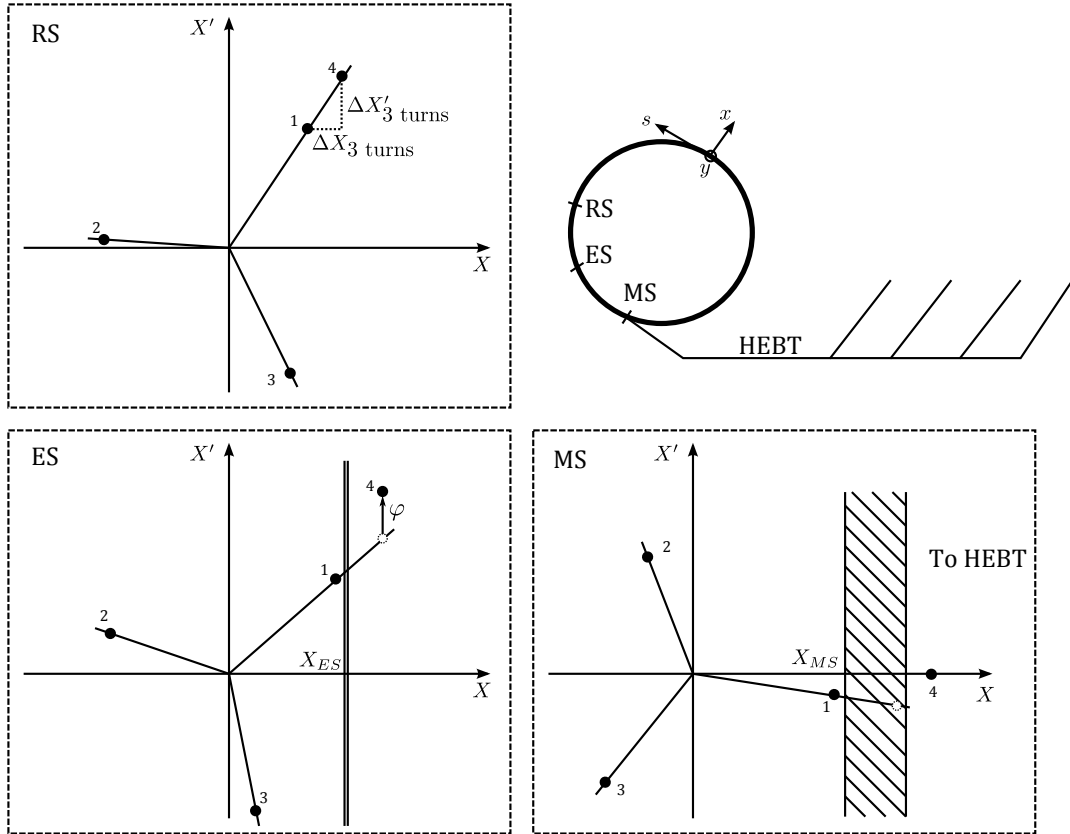

Figure 3.9: Schematic location of the resonance sextupole (RS), electrostatic septum (ES) and magnetic septum (MS) in the synchrotron (top right) and the phase space coordinates over the last three turns of a particle that is to be extracted to the High Energy Beam Transfer line (HEBT). For simplicity, the separatrices for an on-momentum particle ($\xi = 0$) are shown (i.e. the area of the stable triangle is zero).

3.3 The Bragg curve

3.3.1 Stopping power

Charged particles moving through matter lose their energy primarily by ionization and atomic excitation. The *stopping power* is defined as the average energy loss per unit length traversed (MeV/cm). For therapeutic energies, electronic interactions is the dominant contributor to energy loss. Toward the end of the particle range - where the energy is low - nuclear interactions become more important. An example of electronic and nuclear stopping power is shown in Fig. 3.10.
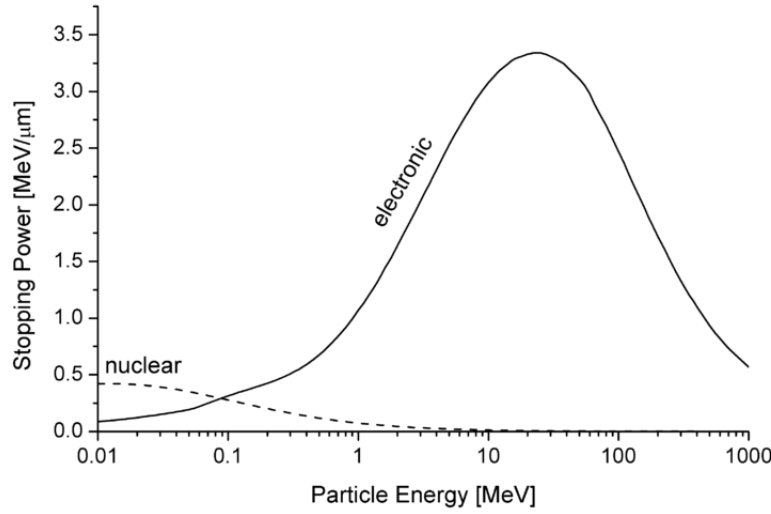
Figure 3.10: Electronic and nuclear stopping power of aluminum ions in aluminum [31].

The following expression (the Bethe formula) for energy loss was derived by Hans Bethe, using relativistic quantum mechanics [32]:

$$-\frac{\mathrm{d}E}{\mathrm{d}z} = \frac{4\pi k_0^2 Q^2 e^4 n}{m_e c^2 \beta^2} \left[\ln \frac{2 m_e c^2 \beta^2}{I(1-\beta^2)} - \beta^2 \right] \tag{3.62}$$

where:

- $k_0 = 8.99 \times 10^9$ N m^2 C^{-2}

- $Q =$ the atomic number of the charged particle

- $e =$ the electron charge

- $n =$ number of electrons per unit volume in the medium

- $m_e =$ the electron mass

- $c =$ speed of light in vacuum

- $\beta = v/c =$ speed of particle relative to c

- $I =$ the mean excitation energy of the medium

The *mass stopping power* is the stopping power divided by the density of the material, $-\mathrm{d}E/\rho\mathrm{d}z$. The mass stopping power is often useful since it is independent of e.g. the pressure of a gas. Additionally, materials with a similar atomic composition have similar mass stopping power.

32

The mean excitation energy, I, for a material with atomic number Z can be estimated to:

$$I \approx \begin{cases} 19.0 \text{ eV} & Z = 1 \\ 11.2 + 11.7 \times Z \text{ eV} & 2 \leq Z \leq 13 \\ 52.8 + 8.71 \times Z \text{ eV} & Z > 13 \end{cases} \qquad (3.63)$$

For compound materials, I can be estimated by adding the individual contributions as:

$$\ln I = \frac{1}{n} \sum_i N_i Z_i \ln I_i \qquad (3.64)$$

where n is the total number of electrons cm^{-3} and N_i the number of atoms cm^{-3}. This gives for water (H_2O, $Z_1 = 1$, $Z_2 = 8$, $N_1/n = 2/10$, $N_2/n = 8/10$):

$$\ln I_{water} \approx \frac{2 \times 1}{10} \ln 19 + \frac{1 \times 8}{10} \ln 105 = 4.312 \Rightarrow I_{water} \approx 74.6 \text{ eV} \qquad (3.65)$$

The stopping power for protons, helium and carbon ions in water, as calculated with the Bethe formula, are shown in Fig. 3.11.
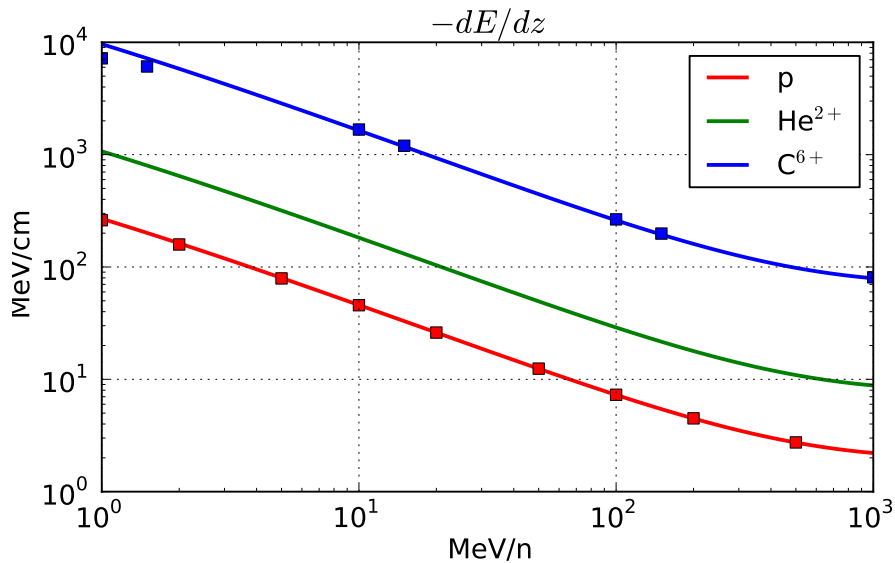

Figure 3.11: Stopping power for protons, helium and carbon ions in the energy range 1-1000 MeV/n. Proton and carbon ion data points from [4] and [33].

3.3.2 Analytical Bragg curve

In the therapeutic energy range (up to a few hundred MeV per nucleon), the stopping power increases with decreasing particle energy. The energy deposition will therefore be lowest at the entrance region. Just before the particle is stopped, it will deposit its remaining energy in a "spike". This spike is referred to as the Bragg peak.

When the kinetic energy of a particle has been reduced from the initial energy T_0 to $T_0 - T$, its current depth can be calculated by integrating the inverse stopping power:

$$z(T_0 - T) = \int_{T_0 - T}^{T_0} \left(-\frac{\mathrm{d}E}{\mathrm{d}z} \right)^{-1} \mathrm{d}E \qquad (3.66)$$

Eq. 3.66 gives the particle energy at any depth, from which the *linear energy transfer*, LET can be calculated. The LET is the average energy deposited along the particle track and is equal to the stopping power[4].

An example of the Bragg curves (LET vs. depth) received by integration of the Bethe formula is shown in Fig. 3.12. The flat region close to the entrance is called the *plateau* region. As the particle loses energy, the LET increases, and ends in a sharp Bragg peak.
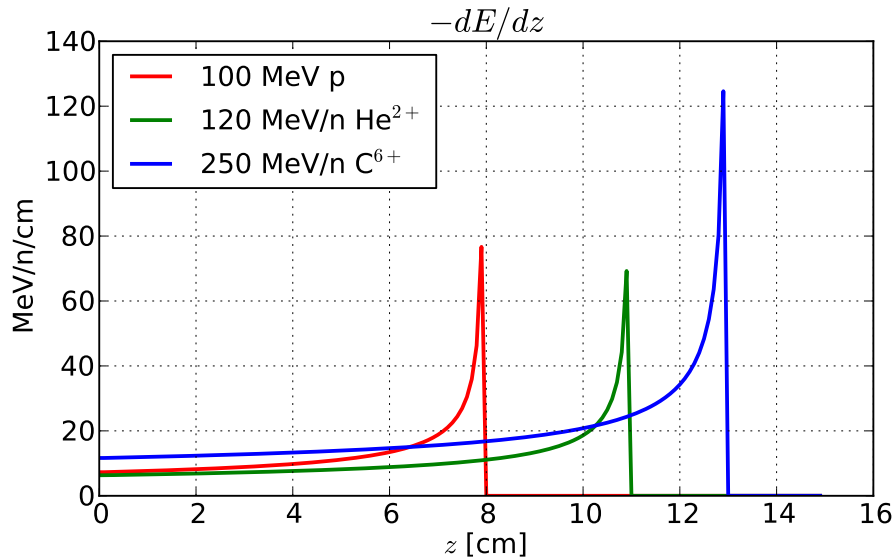

Figure 3.12: Example of proton, helium and carbon ion Bragg curves, calculated by integration of the Bethe formula.

3.3.3 Range

Analytically, the range R of a particle with initial kinetic energy T_0 is given by integrating Eq. 3.66 from zero kinetic energy:

$$R_{csda} = \int_0^{T_0} \left(-\frac{\mathrm{d}E}{\mathrm{d}z} \right)^{-1} \mathrm{d}E \qquad (3.67)$$

This definition of range is called the *continuous slowing down approximation* range, or csda range [33].

---

[4]If not explicitly mentioned otherwise, the LET always refers to the *unrestricted* LET [32], L_∞, which is equal to the stopping power.

However, the ionization and excitation processes which primarily cause energy loss are random. When two initially identical particles reach a certain depth, they will not necessarily have undergone the same number of interactions, and will therefore not have the same residual energy. This process is called *energy straggling*. Additionally, multiple scattering will randomly deflect the ions at each collision, with the result that the path taken by different ions will differ, as shown in Fig. 3.13. As a result, there will be a spread in the total path length when a group of initially mono-energetic particles reaches a depth d. This process is referred to as *range straggling*[5]. Both energy straggling and range straggling cause a spread in the final range of a group of particles with identical starting conditions.
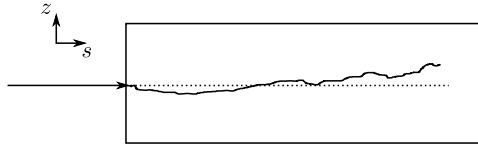

Figure 3.13: Beam path of a scattered particle.

A phenomenological expression for the standard deviation of the energy straggling is [34, 35]:

$$\sigma_{strag} = 0.012 \times R_0^{0.95} A^{-0.5} \tag{3.68}$$

where R_0 is the *mean* range and A the particle mass number. Thus, for proton ($A = 1$) and carbon ion ($A = 12$) beams of equal range, the energy straggling will be about a factor $\sqrt{12} \approx 3.5$ higher for the proton beam, compared to the carbon ion beam.

In order to estimate the mean ranges for beams of different energies, the SRIM[6] program [36] has been used, rather than analytical integration of the stopping power. SRIM produces e.g. range and stopping power tables for ions of different energies in different materials. The difference in range compared to e.g. ICRU energy-range tables [33] is below 2-3% [35].

Shown in Fig. 3.14 is the mean range, as calculated by SRIM [36], with exponential fits (R_0 in cm and E in MeV/n):

$$R_0 = \alpha \times E^k \quad \begin{cases} \text{Protons:} & \alpha = 0.00262, \quad k = 1.734 \\ \text{Carbon ions:} & \alpha = 0.00123, \quad k = 1.675 \end{cases} \tag{3.69}$$

One effect of straggling is that the sharp Bragg peak calculated by integration of the Bethe formula will be "smoothed", since not all particles will stop at the same depth.

3.3.4 Water equivalent thickness

The *Water Equivalent Thickness* (WET) of a slab of some material with thickness d_m is defined as the required thickness of a water column to cause the same energy

---

[5]Note that the term range straggling sometimes is used for what is here called energy straggling. This document follows the terminology of [32].
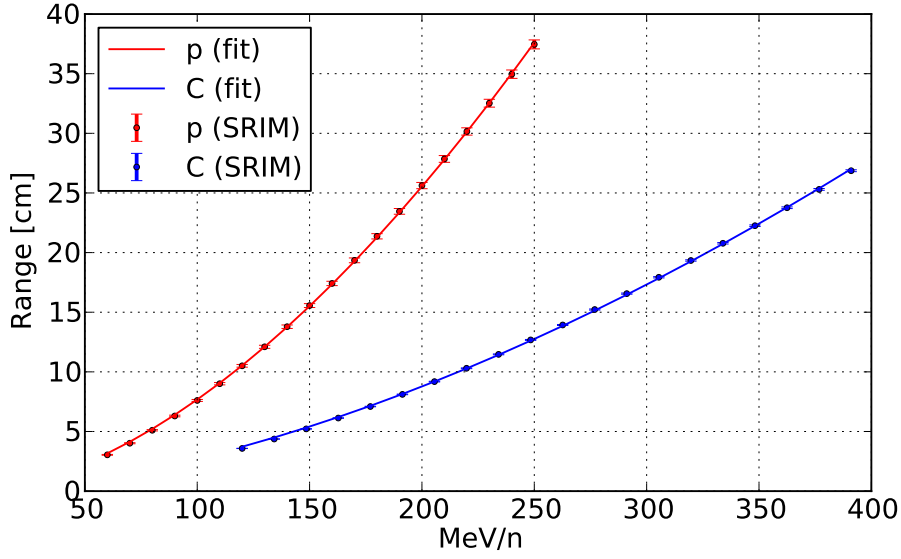
[6]Available at http://www.srim.org

Figure 3.14: Range vs. energy for protons and carbon ions in water (SRIM). Error bars indicate $\pm\sigma_{strag}$. Solid lines are exponential fits.

loss as when a beam passes through the slab:

$$WET = d_m \frac{\rho_m}{\rho_w} \frac{\bar{S}_m}{\bar{S}_w} \tag{3.70}$$

ρ_m and ρ_w are the density of the slab and water, and $\bar{S}_m$ and $\bar{S}_w$ are the average stopping power for the incoming beam in the slab and in water. If the beam enters and exits the slab with kinetic energies T_{in} and T_{out}, respectively, the average stopping power is defined as [37]:

$$\bar{S} = \frac{\int_{T_{in}}^{T_{out}} \left(\frac{dE}{dz}\right) dT}{\int_{T_{in}}^{T_{out}} dT} \tag{3.71}$$

3.3.5 Bragg curve parameterization

In order to model the depth-dose distribution from a proton or carbon ion beam in water, Monte Carlo simulations in GEANT4 have been performed over the therapeutic energy ranges by Dr. F. Moser [38]. The simulated Bragg peaks have been parameterized according to Bortfeld's model [39]. An example of the parameterized Bragg curves is shown in Fig. 3.15. As seen, the widening of the Bragg peak due to straggling is more pronounced for protons than carbon ions.

3.3.6 Spread Out Bragg Peak

In active scanning, a homogeneous target dose is produced by superposition of several Bragg curves of different weights and energies. Assume a target located between depths z_{min} and z_{max}, with infinite transverse dimensions. In depth, the target
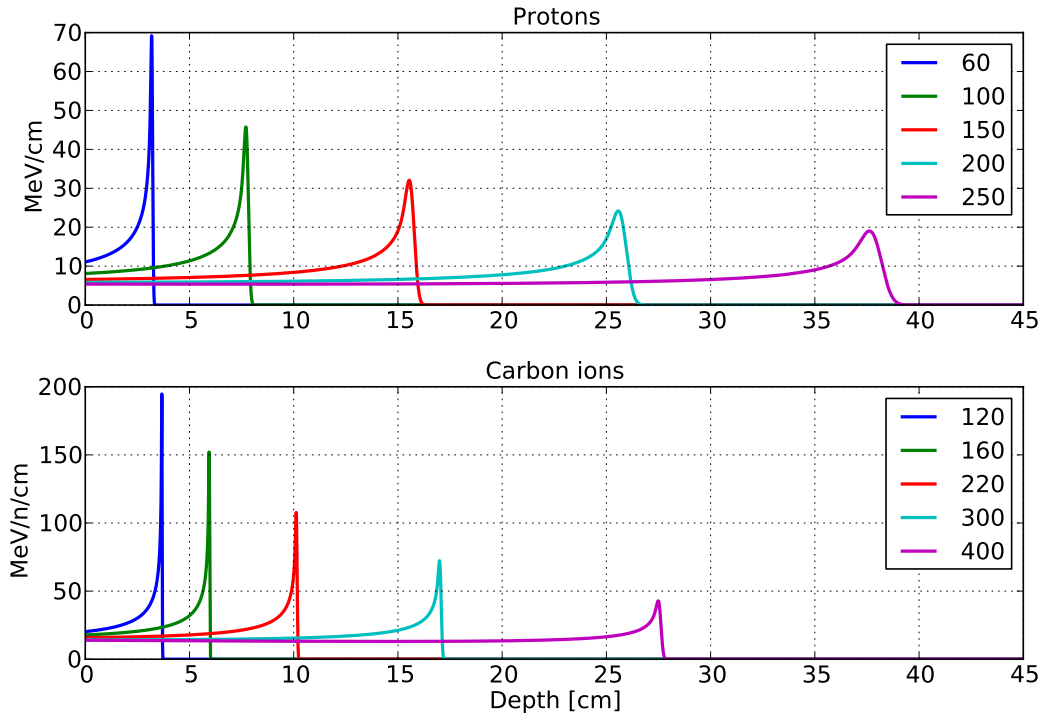
36

Figure 3.15: Proton and carbon ion Bragg curves parameterized from GEANT4 simulations [38]. Legends indicate initial beam energy in MeV/n.

is divided into N layers, each layer corresponding to the penetration depth of a beam with initial kinetic energy T_i, with a weight λ_i (particles per unit area). The LET of each Bragg curve is denoted $L_i(z)$ (energy deposition per unit length). The target dose, $D(z)$, is given by:

$$D(z) = \sum_i \lambda_i L_i(z)/P \quad [\frac{\text{particles}}{\text{cm}^2} \times \frac{\text{J}}{\text{cm}} \times \frac{\text{cm}^3}{\text{kg}} = \text{Gy}] \qquad (3.72)$$

where P is the target density. There are different ways of optimizing the layer weights, but the simplest method is to aim at a dose that is as flat as possible between z_{min} and z_{max}[7]. For carbon ions beams, or low energy proton beams, the Bragg peaks are very sharp: the distance between two consecutive layers must then be short in order to avoid "spikes" in the longitudinal dose distribution. An example of a "Spread Out Bragg Peak" (SOBP) for a carbon ion target between 10 and 14 cm depth is shown in Fig. 3.16, using layer thicknesses of 1 and 2.5 mm. In the former case, the target dose is homogeneous to $\pm 1\%$, while in the latter case, the layers are so sparsely spaced that the sharp Bragg peaks cause local "hot" and "cold" spikes at up to $\pm 20\%$.

---

[7]In fact, it is the biological dose that should be optimized, for which the RBE must be taken into account. This is neglected here, since, from a beam delivery system point of view, it is the quality of the measurable (physical) dose that is primarily of interest.
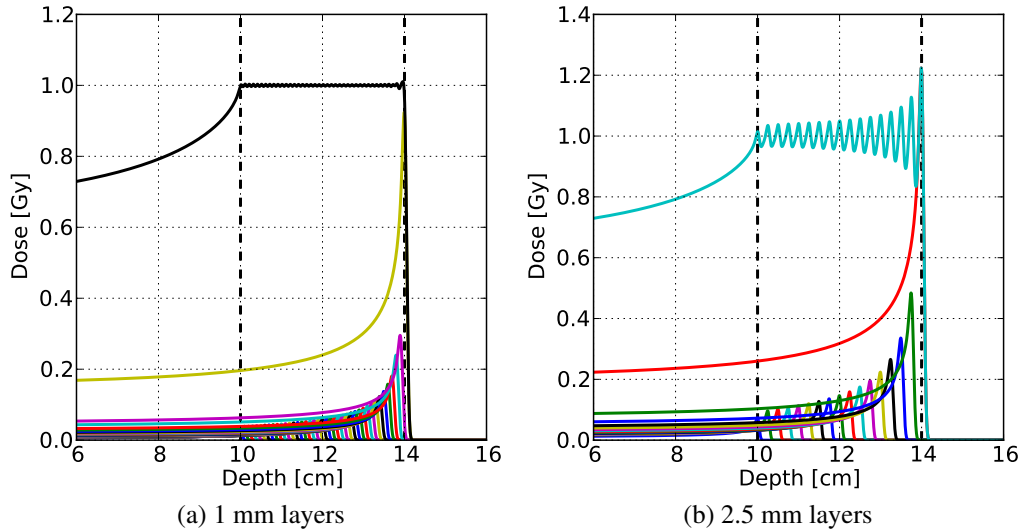
(a) 1 mm layers

(b) 2.5 mm layers

Figure 3.16: Carbon ion SOBP ($z_{min} = 10$ cm, $z_{max} = 14$ cm).

3.4 Scattering

In this section, a model for scattering along a beam line (taking beam optics into account) and in the patient will be elaborated. This model will be used for evaluating different nozzle designs, with respect to achievable beam sizes (and lateral penumbra) at the patient, and to model the dose distribution in the patient during active scanning.

3.4.1 Theory

An ion traversing a slab of some material will repeatedly change its direction mainly due to collisions with the nuclei of the slab. Most collisions result in small-angle deflections, but the accumulated effect will at some penetration depth d have led to a transverse displacement z and shift in angular direction z', as illustrated in Fig. 3.13.

The primary cause of deflection is Coulomb scattering from the nuclei, but there is also a contribution from strong interaction between the incoming ion and the nuclei [6, 40].

3.4.1.1 Molière theory

Consider a mono-energetic point-beam that passes through a slab that is thick enough that the average number of collisions is sufficiently high to allow for a statistical treatment of the distribution of z' at the exit, but thin enough that the energy lost is negligible. According to the scattering theory of Molière [41] (a condensed summary can be found in [42]), the distribution of the *space angle*, $\theta = \sqrt{x'^2 + y'^2}$ can be expressed as a power series in $1/B$, a beam and material dependent constant (taking

38

only the first three terms):

$$f(\theta') = \frac{1}{2\pi\theta_M^2}\frac{1}{2}\left[f^{(0)}(\theta') + \frac{f^{(1)}(\theta')}{B} + \frac{f^{(2)}(\theta')}{B^2}\right] \tag{3.73}$$

The constant B is given by solving the equation:

$$B - \ln B = b \tag{3.74}$$

where b is the logarithm of the effective number of collisions in the target (see [42] for details) and the *reduced angle* θ' is a scaling of the space angle θ:

$$\theta' = \frac{\theta}{\sqrt{2}\theta_M} \tag{3.75}$$

θ_M is the characteristic (projected) multiple scattering angle, using Gottschalk's notation [42], and related to B as:

$$\theta_M^2 = \frac{1}{2}\chi_c^2 B \tag{3.76}$$

The term χ_c is the *characteristic single scattering angle*, proportional to:

$$\chi_c^2 \propto \left(\frac{Q}{pc\beta}\right)^2 \times \frac{Z^2}{A}D\rho \tag{3.77}$$

Here, Q is the charge number of the particle, p its momentum and $c\beta$ its velocity. Z and A are the atomic number and weight of the material and $D\rho$ the target thickness multiplied by its density.

The functions $f^{(n)}$ are defined as:

$$f^{(n)} = \frac{1}{n!}\int_0^\infty yJ_0(\theta'y)e^{-y^2/4}\left(\frac{y^2}{4}\ln\frac{y^2}{4}\right)^n dy \tag{3.78}$$

where J_0 is a Bessel function of the first kind. The first term in Eq. 3.73 is a central Gaussian:

$$f^{(0)}(\theta') = 2 \times e^{-\theta'^2} \tag{3.79}$$

At small angles, $f(\theta')$ is nearly Gaussian. However, the best fit to $f(\theta')$ at small angles is *not* given by simply omitting the two terms $f^{(1)}$ and $f^{(2)}$, since the central part of f is somewhat narrower than $f^{(0)}$ ($f^{(1)}$ and $f^{(2)}$ can be negative). A better fit to the $1/e$ space angle, $\theta_{1/e}$ is given by:

$$\theta_{1/e} = \theta_M \times \sqrt{\frac{B-1.2}{B}} \tag{3.80}$$

with which the fitted Gaussian of the projected angular distribution has a standard deviation θ_0 of:

$$\theta_0 = \frac{\theta_{1/e}}{\sqrt{2}} \tag{3.81}$$

An example of $f(\theta')$, the dominant Gaussian term $f^{(0)}$ and the Gaussian fit to $f(\theta')$ is shown in Fig. 3.17 (160 MeV protons scattered by 30 μm copper). The total distribution is slightly more narrow than the Gaussian term, and exhibits longer tails. The Gaussian fit, using θ_0, agrees well with the central part of the Moliere distribution, and it is only at large angles that the difference is visible. However, at large angles, the angular density is only a few percent of the central peak, and for any practical purposes the Gaussian fit using Eq. 3.81 is adequate.

The angle θ_0 is in this document simply called the *scattering angle*, referring to the standard deviation of the fitted Gaussian to the projected angular distribution of a point beam after being scattered.
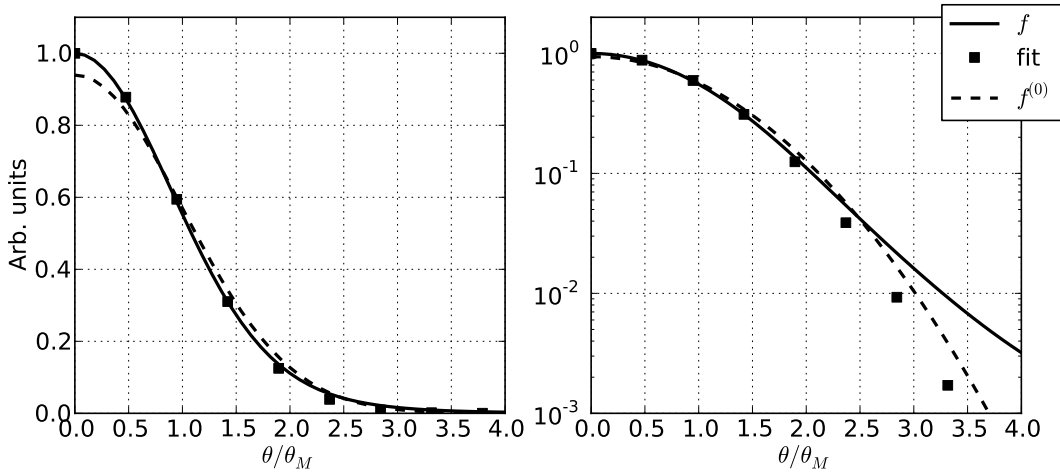

Figure 3.17: $f(\theta')$, fitted Gaussian and $f^{(0)}(\theta')$ for a 160 MeV proton beam scattered in 29 μm copper ($B = 8.761$, $\theta_M = 1.651$ mrad [42]).

3.4.1.2 Highland formula

The scattering theory of Molière is fairly complicated and in 1975, Higland suggested a considerably simpler empirical expression for "the average experimentalist" to quickly calculate the scattering angle (of a proton beam) [43]:

$$\theta_0 = \frac{13.9\,\text{MeV}}{p\beta c}\sqrt{\frac{D\rho}{X_0}}\left(1 + \frac{1}{9}\log_{10}\frac{D\rho}{X_0}\right) \tag{3.82}$$

X_0 is the *radiation length* of the material (see Sec. 3.4.1.3) (g/cm^2, independent of beam). The constant value 13.9 MeV would in fact depend on the material but Highland suggests using a single value, fitted to an element in the middle of the periodic table (silver, Z=47). The accuracy of this empirical fit is claimed to be about 5%, "except for very light elements, or low velocity, where it is 10-20%".

The Highland formula should not be applied to thick scatterers where the energy loss is significant, but is otherwise accurate to 11% in the range $10^{-3} < D\rho/X_0 < 100$ [40][8].

---

[8]Singly charged particles, $\beta = 1$.

3.4.1.3 Radiation length

The radiation length X_0 is a material dependent constant with units of g/cm^2. Materials with a short radiation length scatter a beam more than materials with a long radiation length. In general, materials with a high atomic number have a short radiation length (see Eq. 3.77). Tab. 3.1 displays the radiation lengths and nominal densities of the materials used in this thesis.

Material	X_0 [g/cm^2]	ρ [g/cm^3]	X_0/ρ [cm]
Water	36.08	1.00	36.08
Air	36.66	1.205 g/l	304 m
Copper	12.86	8.96	1.43
Mylar	39.95	1.39	28.7
Kapton (polyimide)	40.56	1.42	28.6
Helium	94.32	0.1786 g/l	5280 m
Aluminum	24.01	2.7	8.89

Table 3.1: Radiation length, density (ρ) and their ratio in cm. Numbers are taken from [40, Chapter 6].

3.4.2 Scattering power

Although Highland's formula is accurate and easy to use, two things are missing:

- For thick scatterers (e.g. the patient, where the beam stops completely), the energy loss must be taken into account.

- It is not applicable to heterogeneous geometries, such as a nozzle with different materials (monitors, air gap, ridge filters etc.), where the radiation length X_0 varies longitudinally.

It may be tempting to divide the geometry into smaller parts, each of a homogeneous material and reducing the beam energy only negligibly, and calculate the scattering angle $d\theta_i$ for each part individually. Since the Highland formula returns the standard deviation for a Gaussian beam, the contributions could be added in quadrature to yield the total scattering angle:

$$\theta_0^2 = \sum_i d\theta_i^2 \tag{3.83}$$

However, this approach will give erroneous results, since (a) the logarithmic term diverges if the scatterer is divided into too thin parts, and (b) the Molière theory does not describe a Gaussian beam: straightforward (quadratic) addition of small angles will always lead to a Gaussian beam [42].

To take energy loss into account, Gottschalk proposed in 1992 [42] an extension of Highland's formula, by moving the beam energy dependence inside an integral:

$$\theta_0 = \left(1 + \frac{1}{9}\log_{10}\frac{D\rho}{X_0}\right)\sqrt{\int_0^D \left(\frac{14.1\ \text{MeV} \times Q}{p(s)c\beta(s)}\right)^2 \frac{\rho\, ds}{X_0}} \tag{3.84}$$

Note that the beam charge number Q is apparent, and that Gottschalk uses 14.1 MeV, instead of Highland's 13.9 MeV.

With negligible energy loss, Gottshalk's expression reproduces Highland's formula[9]. Gottschalk's choice of keeping the logarithmic term outside the integral is somewhat arbitrary, but necessary in order to make the integral converge. Gottschalk's formula agrees with measurement within a few percent.

Still, an integral formulation of the resulting scattering angle would be convenient for modeling of scattering along an arbitrary geometry. This is the purpose of the *scattering power*, $S(s)$, which is a measure of how quickly the 1-σ divergence of a beam increases due to scattering. With a Gaussian beam model, the standard deviation of the projected angle is equal to:

$$\sigma_{z'} = \sqrt{\langle z'^2 \rangle} \tag{3.85}$$

i.e. the rms width of the beam divergence is (for a Gaussian beam) the expectation value of z'^2. The scattering power S is defined as the derivative of $\langle z'^2 \rangle$ with longitudinal coordinate s:

$$S(s) = \frac{d \langle z'^2 \rangle}{ds} \tag{3.86}$$

The scattering power should be formulated such that an integration of $S(s)$ over e.g. a single, thin, slab yields the same result as e.g. the Highland formula.

At the heart of the Highland formula and Gottschalk's integral formulation is the beam specific term $1/pc\beta$ and the material dependent term $1/X_0$. In [44], Kanematsu formulates the scattering power accordingly:

$$S(s) = f_{dH}(l) \frac{E_s^2}{X_0} \left(\frac{Q}{pc\beta} \right)^2 \tag{3.87}$$

The term $f_{dH}(l)$ is a *correction term*, defined below, and the constant E_s used by Kanematsu is:

$$E_s = 15.0 \text{ MeV} \tag{3.88}$$

which is also the value used throughout this thesis.

The correction term is a function of l, which is defined as the integral path length in radiation lengths:

$$l(s) = \int_0^s \frac{\rho(s') ds'}{X_0(s')} \tag{3.89}$$

The correction term is given by the extra condition that the *average* correction term over a homogeneous slab should coincide with the square of Highland's original correction (apart from a factor 14.1 MeV, as used by Gottschalk, instead of 13.9 MeV,

---

[9]Apart from the different values used for the constant energy.

as used by Highland). This gives:

$$\frac{1}{l}\int_0^l f_{dH}(l')\mathrm{d}l' = \left(1 + \frac{1}{9}\log_{10} l\right)^2 \times \left(\frac{14.1}{E_s}\right)^2 \Rightarrow$$

$$f_{dH}(l) = \left(\frac{14.1\,\text{MeV}}{E_s}\right)^2 \left(1 + \frac{\log_{10} l}{9}\right)\left(1 + \frac{2}{9\ln 10} + \frac{\log_{10} l}{9}\right) \quad (3.90)$$

with which the scattering power $S(s)$ is completely defined, and the scattering angle θ_0 of a scattered point beam is:

$$\theta_0 = \sqrt{\int_0^d S(s)\mathrm{d}s} = \sqrt{\langle z'^2\rangle(d)} \quad (3.91)$$

Note that the scattering power as defined here is *non-local*: the scattering power at one point depends not only on the beam energy and material at that point, but also on how much material the beam has already traversed via Eq. 3.89. Physically, this does not make sense, but one should keep in mind that the scattering power formulation is an extension of the Highland formula, which in itself is an empirical fit.

3.4.3 Scattering in a drift space

The transverse phase space of a Gaussian beam is uniquely defined by the three statistical quantities $\langle z'^2\rangle$, $\langle zz'\rangle$ and $\langle z^2\rangle$:

$$\langle z'^2\rangle = \frac{1}{N}\sum_i z_i'^2 = \iint z'^2 \rho(z,z')\mathrm{d}z\mathrm{d}z'$$

$$\langle zz'\rangle = \frac{1}{N}\sum_i z_i z_i' = \iint zz' \rho(z,z')\mathrm{d}z\mathrm{d}z'$$

$$\langle z^2\rangle = \frac{1}{N}\sum_i z_i^2 = \iint z^2 \rho(z,z')\mathrm{d}z\mathrm{d}z' \quad (3.92)$$

where the sum is a sum over all N particles and $\rho(z,z')$ is the particle density in the horizontal or vertical phase space (see Eq. 3.17). Carrying out the integrals in Eq. 3.92 gives the simple relations between the expectation values and the conventional Twiss parameters:

$$\langle z'^2\rangle = \ldots = \frac{\varepsilon_z}{\pi} \times \gamma_z$$

$$\langle zz'\rangle = \ldots = -\frac{\varepsilon_z}{\pi} \times \alpha_z$$

$$\langle z^2\rangle = \ldots = \frac{\varepsilon_z}{\pi} \times \beta_z \quad (3.93)$$

For beam modeling whenever scattering is present, it is in general more convenient to work with the expectation values than the Twiss functions, since the emittance is not conserved.

The three quantities $\langle z'^2 \rangle$, $\langle zz' \rangle$ and $\langle z^2 \rangle$ are in the following referred to as the (squared) *divergence*, *covariance* and (squared) *width* of the beam.

Scattering in a thin slab increases the divergence of the beam, but not covariance nor the beam width: it is only further downstream that the scattering is measurable as an increase in beam size. A beam with initial phase space given by $\langle z'^2 \rangle_a$, $\langle zz' \rangle_a$ and $\langle z^2 \rangle_a$ at $s = a$ going through some multi-slab geometry has at $s = b$ reached a phase space [45]:

$$\langle z'^2 \rangle_b = \langle z'^2 \rangle_a + \int_a^b \mathcal{S}(s')\mathrm{d}s'$$

$$\langle zz' \rangle_b = \langle zz' \rangle_a + (b-a)\langle z'^2 \rangle_a + \int_a^b (b-s')\mathcal{S}(s')\mathrm{d}s'$$

$$\langle z^2 \rangle_b = \langle z^2 \rangle_a + 2(b-a)\langle zz' \rangle_a + (b-a)^2 \langle z'^2 \rangle_a + \int_a^b (b-s')^2 \mathcal{S}(s')\mathrm{d}s' \quad (3.94)$$

3.4.4 Energy loss

In order to calculate the scattering power at $s = s'$, the kinetic energy of the particle at that point must be known. The momentum, p, and velocity, v, can be calculated from the kinetic energy T:

$$p = \sqrt{T^2 + 2E_0 T} \quad (3.95)$$

$$v = c \times \sqrt{1 - \frac{1}{(1 + E_k/E_0)^2}} \quad (3.96)$$

c is the speed of light and $E_0 = m_0 c^2$, with m_0 being the particle rest mass. In order to calculate the energy at any depth, energy-range tables from SRIM have been used, which (by interpolation) provides the beam range as a function of initial energy, $R(T_0)$, and the inverse, $R^{-1}(R_0)$, in any homogeneous material. The method for calculating the beam energy at any depth is illustrated in Fig. 3.18: at a depth $s = a$, the residual range is $R_0 - a$, and the remaining kinetic energy of the beam at depth a is thus given by:

$$T(a) = R^{-1}(R_0 - a) = R^{-1}(R(T_0) - a) \quad (3.97)$$

The same principle is applied to heterogeneous multi-slab geometries, but requires calculating the beam energy at the exit of each slab, which is used as initial energy for the next slab. One example of the energy loss for a proton beam in a multi-slab geometry is shown in Fig. 3.19.

3.5 Beam profile terminology

The term *beam size*, or beam width is, if not explicitly mentioned otherwise, used for the FWHM of the beam. The beam FWHM is denoted W. A Gaussian beam
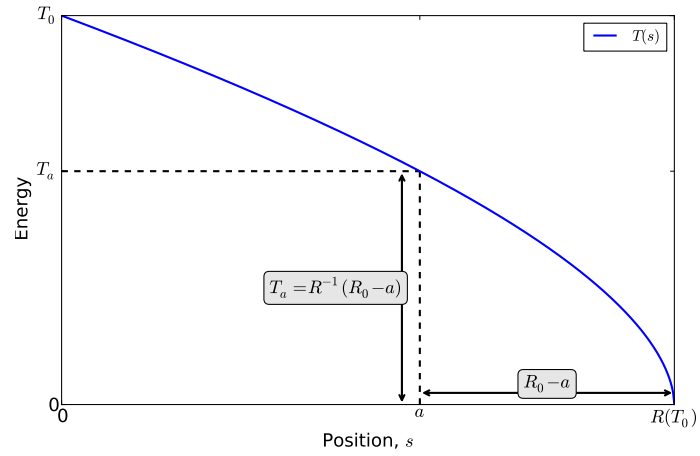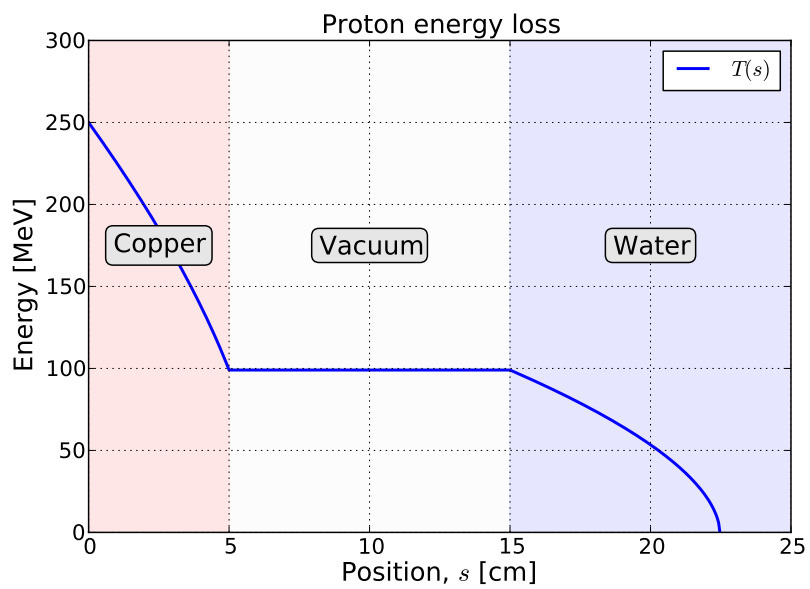
Figure 3.18: Energy as a function of depth.

Figure 3.19: The kinetic energy T at different depths s for a proton beam traversing 5 cm copper, 10 cm vacuum and stops in water.

profile can be described as:

$$\rho(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{3.98}$$

where σ is the standard deviation of the beam width. The FWHM is related to σ as:

$$W = 2\sqrt{2\ln 2} \approx 2.35\sigma \tag{3.99}$$

The lateral *penumbra* of the beam is defined as lateral distance from 20% to 80% of the peak dose:

$$\text{Penumbra} = \sqrt{2}\left(\sqrt{\ln\frac{1}{0.2}} - \sqrt{\ln\frac{1}{0.8}}\right)\sigma \approx 1.13\sigma \approx 0.48W \tag{3.100}$$

i.e. the 80-20 penumbra of a Gaussian beam is roughly half the FWHM.

3.6 Active scanning

3.6.1 General

In depth, the target is divided into layers of a few mm, each layer corresponding to the penetration depth of a specific extraction energy. Every layer is divided into spots and two orthogonal dipoles (scanning magnets) guide the beam from spot to spot during irradiation. Apart from the particle type (proton or carbon ions), every spot is defined by:

1. The extraction energy E_i

2. The transverse beam size (FWHM), W. Although the horizontal and vertical beam size can be set individually, it is in this work assumed that the beam is symmetric, i.e $W_x = W_y = W$, and that the same initial beam size is used for all layers of the target.

3. The transverse coordinates (X_i, Y_i) of the spot center at the Bragg peak in a plane orthogonal to the direction of the non-scanned beam.

4. The prescribed spot dose (number of particles), n_i

If the accelerator is a cyclotron, the beam is typically switched off when the prescribed dose to a single spot has been delivered. While the beam is turned off, the current through the scanning magnets is updated to match the position of the next spot. Also, it is verified that the delivered dose was within tolerance. If this is the case, the dose counter is reset to zero[10], and irradiation of the next spot is permitted as soon as the scanning magnet current matches the position of the next spot. This

---

[10]Private communication with the Proton Therapy team at PSI.

typically requires turning the beam on and off at a high rate (in the kHz range) with a switching time of 100 μs or less.

With a synchrotron using a betatron core for extraction, it is challenging to pause the extraction at such a high frequency. The alternative would be to insert a kicker magnet (a fast ramped dipole) somewhere along the extraction line to deflect the beam onto a beam dump to create a beam-pause between spots. The disadvantage with this is that a large fraction of the accelerated and extracted beam would be dumped. Since there is a dead time of about 1.5 s between every layer[11], it is important to utilize the extracted beam as efficiently as possible to reduce the overall irradiation time. Therefore, therapeutic synchrotron beams are typically *not* turned off when moved from spot to spot.

Scanning step The distance between two neighboring spots, the *scanning step*, is denoted Δ. To produce a homogeneous dose, the scanning step must be smaller than the extent of the beam. For a Gaussian beam profile, the following condition is sufficient to produce a homogeneous dose within 0.1% [46]:

$$\Delta \leq 1.6\sigma \approx 0.7W \quad \text{[Gaussian profile]} \tag{3.101}$$

This condition applies to the vertical profile of the beam extracted from the MedAustron synchrotron (see Sec. 4.3.3). In the horizontal plane, however, the beam profile is more trapezoidal, with sharp edges. In order for the horizontal beam profiles produce a homogeneous dose, it is necessary that the edges of one spot coincides with another spot. This can be expressed as:

$$\Delta = \frac{W}{m} \quad \text{[Rectangular or trapezoidal profile]} \tag{3.102}$$

where m is an integer number. For an equal scanning step in *x* and *y*, it follows that $W/m \leq 0.7W$, i.e. $m = 2, 3, 4, \ldots$ are valid steps. The scanning step should thus be half, one-third, etc. of the beam FWHM.

Spot weight Assuming that the target is rectangular, and that layer *i* is to be irradiated homogeneously with λ_i particles/cm^2, the number of particles per spot, n_i, is related to the scanning step as:

$$n_i = \lambda_i \left(\frac{W}{m}\right)^2 \quad \text{[Spot weight]} \tag{3.103}$$

The tighter the spot grid (i.e. the larger the value of *m*), the lower the number of particles per spot.

Single spot irradiation time With a nominal particle rate $\dot{N}_0$, the nominal irradiation time of a single spot (including the time it takes to move between spots) is:

$$T_{spot} = \frac{n_i}{\dot{N}_0} = \frac{\lambda_i}{\dot{N}_0} \left(\frac{W}{m}\right)^2 \tag{3.104}$$

---

[11]Between two spills, a hysteresis cycle of the synchrotron magnets is performed. This cycle takes about 1 s.

Transition time When the trigger to move the beam to the next spot is given, there will be a delay τ before the beam starts to move (see Sec. 3.6.2.2). The transition time is the time between trigger and the beam being centered upon the next spot:

$$T_{trans} = \tau + T_{move} = \tau + \frac{\Delta}{v_{beam}} \qquad (3.105)$$

v_{beam} is the beam velocity during ramping. In order not to systematically overdose any spots, it must be ensured that $T_{spot} \geq T_{trans}$. With $T_{spot} = T_{trans}$, the beam must be swept continuously in order not to overdose.

Dose rate at Bragg peak The *dose rate*, in a small mass element Δm at the Bragg peak is given by:

$$\frac{\Delta D}{\Delta t} = \frac{\Delta E / \Delta t}{\Delta m} = \frac{\Delta E / \Delta t}{P \Delta V} \qquad (3.106)$$

with P being the target density. $\Delta E / \Delta t$ is the energy deposition per unit time:

$$\frac{\Delta E}{\Delta t} = \frac{\dot{N} \Delta t \delta E}{\Delta t} = \dot{N} \delta E \qquad (3.107)$$

where $\dot{N}$ is the particle rate and δE the energy deposition of a single particle at the Bragg peak:

$$\delta E = \left| \frac{dE}{ds} \right|_{BP} \Delta s \qquad (3.108)$$

With $\Delta s \approx 5$ mm, δE is approximately 21 MeV for a single proton, and about 40 MeV/n for a carbon ion. Now, let ΔV be a small volume element, 5 mm in depth (Δs), with transverse dimensions similar to the beam size, W^2. The dose rate at the Bragg peak is then given by:

$$\frac{\Delta D}{\Delta t} = \frac{\dot{N} \delta E}{P W^2 \Delta s} \qquad (3.109)$$

Numerical values of the dose rate at highest extraction intensity are shown in Tab. 3.2.

Sweeping dose In a real treatment plan, the spots in one layer can be clustered into spatially separated "islands", depending on the geometry of the target. If no dose is prescribed to the region between islands, the beam should ideally be turned off. If it is not, a "sweeping dose" will be deposited along a line between two consecutive, but separated, spots. A volume element at the Bragg peak, with dimensions $W^2 \Delta s$ is exposed to the beam during a time:

$$\Delta t = \frac{W}{v_{beam}} \qquad (3.110)$$

where v_{beam} is the velocity of the scanned beam. The dose deposited along a line is:

$$D_{sweep} = \frac{\Delta D}{\Delta t} \times \frac{W}{v_{beam}} = \frac{\dot{N} \delta E}{P W \Delta s v_{beam}} \qquad (3.111)$$

Numerical values for proton and carbon ions at highest extraction intensity are shown in Tab. 3.2.

		Protons	Carbon ions
Energy deposition around Bragg peak ($\Delta s = 5$ mm)	δE	21 MeV	40 MeV/n
Max beam intensity (A.2)	$\dot{N}$	2×10^{10} s^{-1}	1×10^{9} s^{-1}
Dose rate in $10 \times 10 \times 5$ mm volume at Bragg peak	$\Delta D / \Delta t$	133 Gy/s	154 Gy/s
Sweeping dose	D_{sweep}	7 cGy	8 cGy
Time to deposit 2 Gy		15 ms	13 ms

Table 3.2: Dose rate and sweeping dose at highest extraction intensity, 10×10 mm beam.

3.6.2 Scanning magnets

3.6.2.1 Accuracy

A hypothetical scanning magnet current profile during irradiation of a spot is shown in Fig. 3.20. An average current error of δI is present during the flat part, as well as some high-frequency ripple[12]. δI results in a beam position error, while the high-frequency ripple will effectively broaden the spot size.
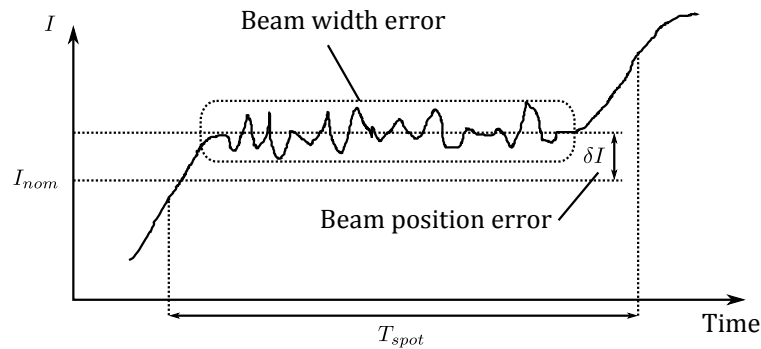

Figure 3.20: Example of scanning magnet current profile during one spot.

Since the deflection angle of a beam passing through the scanning magnets is inversely proportional to its magnetic rigidity $B\rho$, the relation between transverse beam position at the isocenter, x, and scanning magnet current I is:

$$I(x, B\rho) = \frac{B\rho}{(B\rho)_{max}} \frac{I_{max}}{x_{max}} x \tag{3.112}$$

Here, x_{max} is the coordinate at the edges of the scanning field (10 cm) and I_{max} the required current to move a 400 MeV/n carbon ion (with magnetic rigidity $(B\rho)_{max}$) to x_{max}.

---

[12]High frequency compared to the inverse spot time.

3.6.2.2 Agility

The communication rate with the scanning magnet power supplies, f_{PS}, sets the limit to how often a "Move Beam"-trigger can be acknowledged. The foreseen communication rate is 50 kHz, which means that triggers will only be registered with a time granularity of:

$$T_{PS} = \frac{1}{f_{PS}} = 20\,\mu s \tag{3.113}$$

The trigger to move the beam is generated by the beam delivery system, based on the delivered dose registered by the beam intensity monitors in the nozzle. Since the sampling rate of these monitors is higher than the power supply communication rate[13], the trigger can be assumed to occur anywhere within one period T_{PS}, assuming the control system for the monitors is fast enough. This means that it will take between 0 and 1 periods before the trigger is acknowledged. An additional period is then required before the current is ramped to the next set point, which corresponds to the center position of the next spot. Thus, the delay between trigger and ramp start is uniformly distributed as:

$$\tau \in \mathcal{U}[T_{PS}, 2T_{PS}] \tag{3.114}$$

with mean and max values of:

$$\langle\tau\rangle = 1.5T_{PS}$$
$$\tau_{max} = 2T_{PS} \tag{3.115}$$

The required current step ΔI to move the beam a distance Δx at the isocenter is:

$$\Delta I = \frac{B\rho}{(B\rho)_{max}} \frac{I_{max}}{x_{max}} \Delta x \tag{3.116}$$

Fig. 3.21a shows the time profile of a typical current step ($\Delta I = 15$ A, which in this case corresponds to about 3 mm for a 400 MeV/n carbon ion beam) and the applied magnet voltage. After an initial delay of about 20 μs (T_{PS}), the magnet voltage ramps up to maximum value (which takes about another 20 μs). Then, the maximum voltage is applied for some time, in order to ramp the current as fast as possible. Before the beam reaches the next spot, the voltage is successively reduced, which slows down the scanning velocity as the beam approaches the next spot. Overall, it takes about 160 μs (subtracting the initial delay) to reach the next spot, including the initial delay and round-off at the end of the step. This is the transition time T_{move}.

Smaller current steps can be completed faster, but the relation between ΔI and T_{move} is not linear. The voltage rise time and the round-off at the end of the step will take approximately the same time for all current steps. An example of a smaller current step ($\Delta I = 1$ A) is shown in Fig. 3.21b. Although this step is 15 times smaller, the transition time is only reduced by a factor 2, approximately.

---

[13]E.g. the CNAO beam intensity monitors work at about 1 MHz [47].
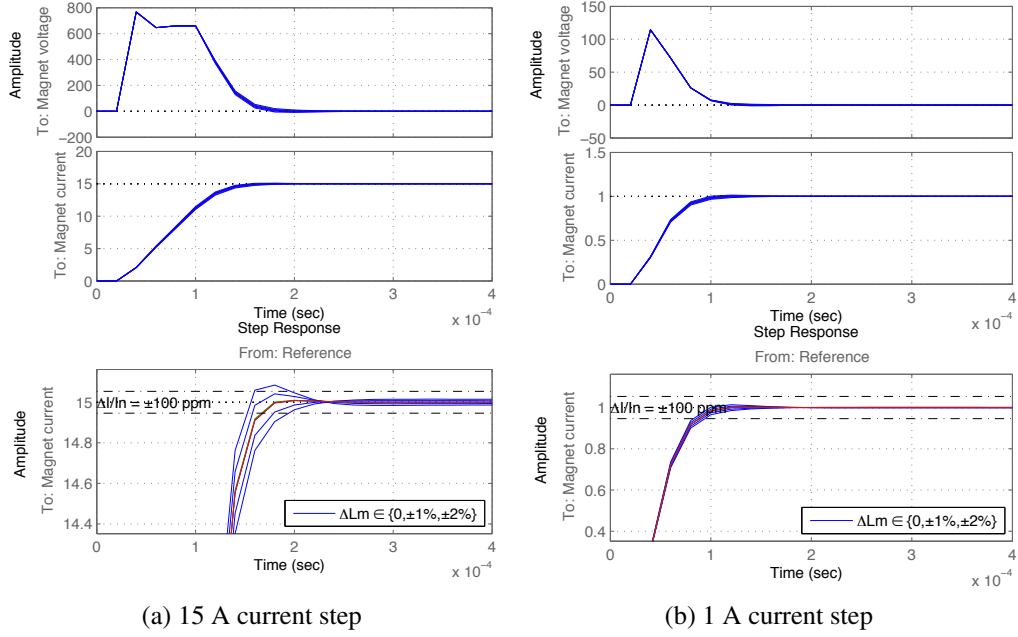
(a) 15 A current step (b) 1 A current step

Figure 3.21: Simulation of current step time profile. Top: magnet voltage. Middle: current step. Bottom: current step (zoom). Courtesy of Philippe Fraboulet, MedAustron.

3.6.3 Effect of beam intensity fluctuations

3.6.3.1 Quasi-discrete scanning

The tolerance to beam intensity ripple is closely linked to the dynamic performance of the scanning magnets. Beam intensity fluctuations that occur while waiting for the current spot dose to reach the threshold value for "Move Beam"-trigger will not affect the final dose of the spot: if the beam intensity is too high, the trigger will simply occur earlier, and vice versa. If the spot is to receive n_i particles, the trigger should be issued when a dose of:

$$n_{trigger} = n_i - n_{lag} \tag{3.117}$$

has been delivered. While the beam moves between spots, the particles are distributed between the two spots: 100% to the current spot at the start of the ramping, and 0% when the beam is centered upon the next spot. During beam movement, the ratio decreases linearly. The dose n_{lag} corresponds to the average number of particles delivered during the delay and the time it takes to reach halfway to the next spot:

$$n_{lag} = \dot{N}_0 \times \left(\langle \tau \rangle + \frac{T_{move}}{2} \right) = \dot{N}_0 \times T_{lag} \tag{3.118}$$

However, if the beam intensity $\dot{N}(t)$ is higher or lower than anticipated during the transition, the final dose of the spot will be too high or too low, as illustrated in Fig. 3.22. Likewise, the uncertainty of the delay τ will influence the final spot dose.

51
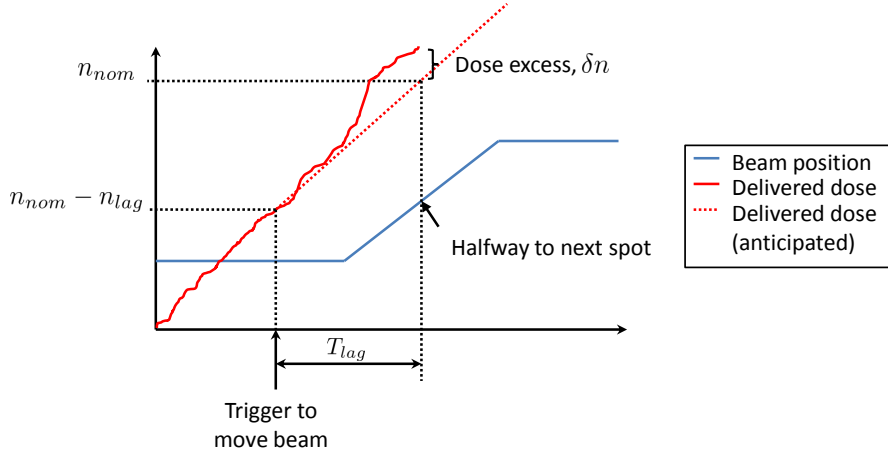
Figure 3.22: Beam intensity fluctuation causing a spot weight error δn.

An upper limit of the final excess of particles given to one spot can be estimated by the following:

$$
\delta n = \overbrace{\int_{t_0}^{t_0+\tau} \dot{N}(t)\mathrm{d}t - \dot{N}_0\langle\tau\rangle}^{\text{Particle excess during delay}} + \overbrace{\int_{t_0+\tau}^{t_0+\tau+T_{move}} \dot{N}(t) \times \underbrace{\left[1 - \frac{t-(t_0+\tau)}{T_{move}}\right]}_{\text{Decreases from 1 to 0 during transition}} \mathrm{d}t - \dot{N}_0 \times \frac{T_{move}}{2}}^{\text{Particle excess during beam movement}}
$$

$$
\leq (1+\xi_{max})\dot{N}_0\tau_{max} - \dot{N}_0\langle\tau\rangle + (1+\xi_{max})\dot{N}_0 \times \frac{T_{move}}{2} - \dot{N}_0 \times \frac{T_{move}}{2} \qquad (3.119)
$$

ξ_{max} is to be interpreted as the maximum relative beam intensity deviation during an integration window in the order of $100\,\mu$s (the time scale for moving between spots).

The *relative* spot weight error, η_i, is given by the ratio between δn and the prescribed spot weight $n_i = \dot{N}_0 T_{spot}$:

$$
\eta_i = \frac{\delta n}{\dot{N}_0 T_{spot}} \leq \frac{(1+\xi_{max}) \times \tau_{max} - \langle\tau\rangle + \xi_{max} \times \frac{T_{move}}{2}}{T_{spot}} = \eta_{max} = E_n \qquad (3.120)
$$

The term E_n is the maximum relative spot weight error, estimated as a function of the scanning magnet agility (τ, T_{move}) and the synchrotron extraction ripple.

The lower limit can similarly be estimated to $-E_n$ if $E_n \leq 1.0$ (a spot can naturally not be underdosed by more than 100%).

3.6.3.2 Discrete scanning

With discrete scanning, the impact of beam intensity fluctuations can be analysed in a similar manner: the decision to turn off the beam is taken slightly before the spot has been fully irradiated. The term n_{lag} then corresponds to the number of particles typically delivered during the time it takes to turn off the beam, rather than the time it takes to move halfway to the next spot. $T_{move}/2$ should therefore be replaced by the switch-off time.

3.6.4 Scanning mode and weight error

As explained in Sec. 3.6.1, once a spot is finished, one typically resets the dose counter before starting irradiation of the next spot. Spot weight errors are caused by variations of the switch-off time and beam intensity fluctuations during the switch-off time. If, for example, the beam intensity is systematically higher than anticipated, every single spot will be overdosed. Resetting the dose counters between spots consequently has the effect that all spot dose errors are independent:

$$n'_i = n_{nom}(1 + \eta_i) \tag{3.121}$$

With quasi-discrete scanning, however, one does not reset the dose counter *per se*. If too few particles were delivered during the transition from spot $i-1$ to spot i, spot $i-1$ will be underdosed. However, the beam will still remain at spot i until the proper integral dose to trigger movement to spot $i+1$ has been reached. A lack of particles at one spot will therefore cause an excess of particles at the next spot. This is illustrated schemtically in Fig. 3.23.
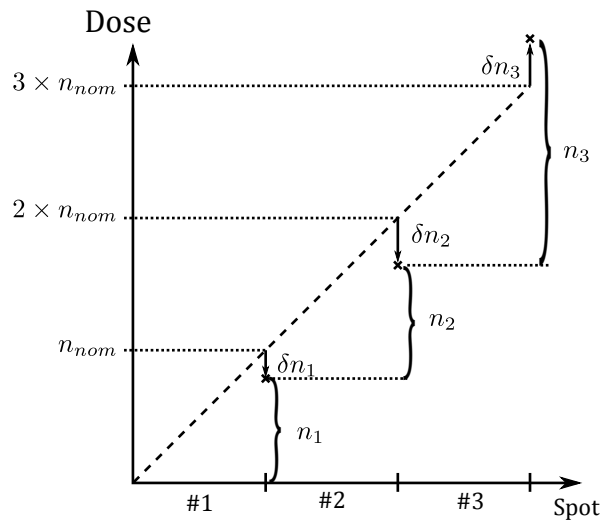

Figure 3.23: Integral dose-driven scanning: the trigger to move from one spot is only generated when the proper integral dose has been delivered.

The effect of integral-dose driven scanning is that a weight error made at spot $i-1$ is automatically compensated for at spot i:

$$n'_1 = n_{nom}(1 + \eta_1)$$
$$n'_2 = n_{nom}(1 + \eta_2 - \eta_1)$$
$$\cdots$$
$$n'_i = n_{nom}(1 + \eta_i - \eta_{i-1}) \quad \text{(Quasi-discrete scanning)} \tag{3.122}$$

From this, it follows immediately that with integral-dose driven scanning, *static, or slow beam intensity variations, where $\eta_i \approx \eta_{i-1}$ will have no or little effect on the dose homogeneity*, since these error are canceled out from one spot to the next.

This is obviously an important advantage. However, for more random spot-to-spot variations of the beam intensity, the effect is not obvious. Subtracting the error made at one spot from the next spot intuitively seems like a reasonable strategy if the spots are very close to each other, with a high degree of overlap. However, if the spots are sparsely placed, this sort of feedback with "spatial delay" could even make things worse.

In principle, it would be possible to implement an advanced quasi-discrete scanning mode algorithm that mimics the behavior of discrete scanning (i.e. no automatic spot-to-spot correction). In order to judge whether such an effort would be useful, the resulting dose homogeneity from an easy-to-implement integral-dose driven scanning ($n'_i = n_{nom}(1 + \eta_i - \eta_{i-1})$) and discrete scanning mode ($n'_i = n_{nom}(1 + \eta_i)$) have been compared for different spot-to-spot distances in Chapter 6.

3.6.5 Spot grid styles

The spot coordinates (X_i, Y_i) are generally arranged in a Cartesian grid by the treatment planning system [4]. For homogeneous irradiation, the horizontal spot-to-spot distance must exactly match the horizontal FWHM (Eq. 3.102). This does not only pose strict requirements on the scanning magnet beam positioning accuracy, but also on the extracted beam profile and optics along the transfer line, since the actual beam width must match the spot grid. Minor errors of the horizontal beam FWHM will cause hot and cold "spikes", where the edges no longer overlap, as illustrated in Fig. 3.24. A width error of the Gaussian beam profile has no effect on the homogeneity of the dose; only the penumbra of the field is affected.

In a Cartesian spot grid, the lateral edges of the trapezoidal beam profile are aligned vertically. Minor width errors would then result in hot and cold "stripes", where the edge matching is disrupted. A minor modification of the spot grid would mitigate the effect of a static (i.e. persistent over all spots in one spill) horizontal beam width error. By moving every second row half a scanning step to the right, as shown in Fig. 3.25, all the edges would no longer be aligned, which would reduce the amplitude of the hot and cold "stripes".

The Cartesian and shifted spot grid will be compared in Sec. 6.3, as well as different scanning steps.

3.7 Nozzle

3.7.1 General

The nozzle contains beam intensity and position monitors for verifying the delivered dose and position of the scanned beam. Additional passive elements are ridge filters, for widening of the Bragg peak, and (optionally) a range shifter for irradiation of superficial tumors.
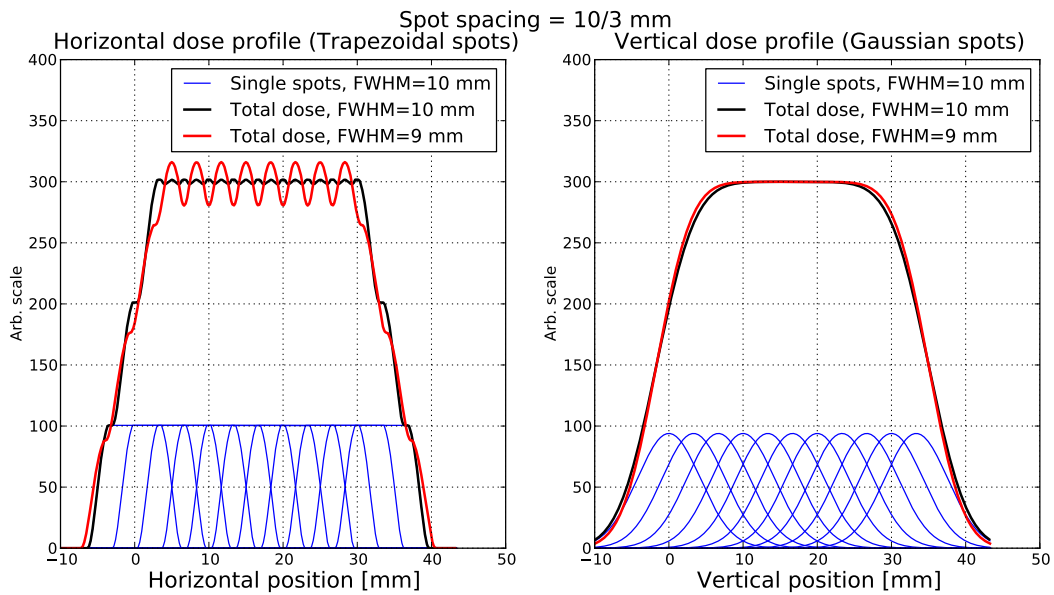
Figure 3.24: The impact of a fixed beam width error (Nominal width is 10 mm, black lines, actual width is 9 mm, red lines). Scanning step is 10/3 mm (blue curves indicated dose profile of each spot) Left: trapezoidal beam profile. Right: Gaussian beam profile (pristine, non-scattered, beam profile).
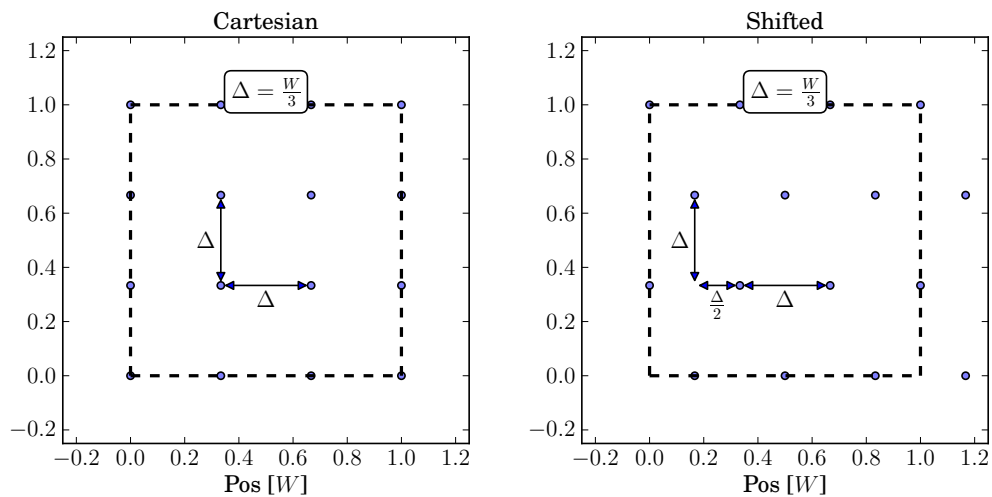

Figure 3.25: Left: Cartesian spot grid. Right: Shifted spot grid. "o" indicate spot center positions and dashed square the beam size ($W \times W$).

3.7.2 Monitors

Beam intensity monitors are used to monitor the dose delivered to the patient. Parallel plate ionization chambers with a short drift time compared to the typical irradiation time per spot are normally used [47–49].

3.7.3 Range shifter

For superficial tumors close to the skin, the penetration depth of even the lowest energy beams is too high. To reduce the range of the beam, one can insert a *range shifter* into the beam path, which reduces the energy of the beam entering the patient. The range shifter is a rectangular slab of constant thickness. In order to minimize lateral beam growth due to scattering, one typically uses a low-Z material, such as Plexiglas. However, the divergence caused by a few centimeter thick range shifter is still severe, and the range shifter must therefore be as close to the patient as possible, in order to keep the beam width (and lateral penumbra) small. The impact on isocenter beam size from a range shifter is evaluated in Chapter 5.

3.7.4 Ridge filter

Every time the extraction energy of the synchrotron is changed, the remaining particles must be dumped, the synchrotron dipoles must complete a hysteresis cycle and the new beam must be injected and accelerated. This causes a beam-off time of about 1.5 seconds between two layers. Since the irradiation time should be kept short to reduce the inconvenience for the patient (and to increase patient throughput), it is of interest to reduce the total number of layers that the tumor is divided into. In order to do this without compromising the dose homogeneity (compare figures 3.16a and 3.16b) one uses a *ridge filter*

A ridge filter is a static device that the pencil beam passes through before reaching the patient. It is typically constructed as a rectangular plate, with parallel grooves, or ridges, which have a width a few times smaller than the beam (see. Fig. 3.26). In the case of MedAustron, with beam sizes as small as 4 mm FWHM (W), the width of each ridge should not be larger than about 1 mm:

$$d_{RF} \approx 1 \text{ mm} \ll W \tag{3.123}$$

The particles that crosses the thickest part of the ridge filter will lose some energy, and their range is reduced by the water equivalent peak thickness of the ridge filter. Particles passing through the thinner part of the ridge filter will lose less energy. The ridge filter thus introduces a range spread. As mentioned, this occurs naturally in the target by range and energy straggling, but the ridge filter provides a way of widening the Bragg peak even for carbon ions and low energy proton beams, thereby reducing the number of layers required to produce a homogeneous dose.

Mathematically, the effect of the ridge filter on the Bragg curve can be modeled as a convolution between the pristine Bragg curve, $L(s)$ and a function $w(t)$ that
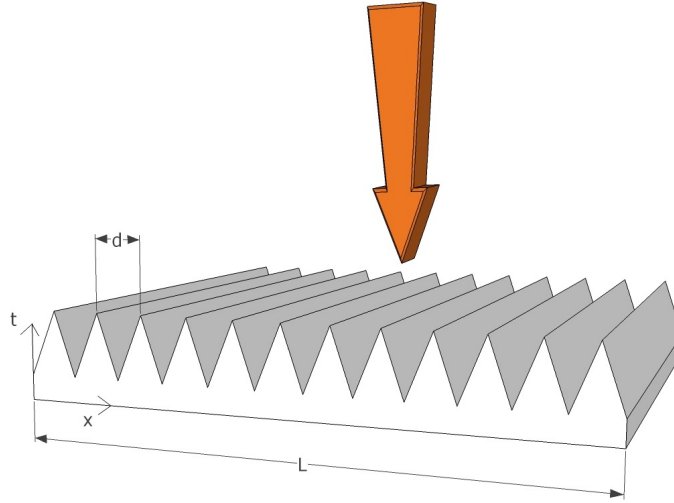
Figure 3.26: Sketch of ridge filter and the impending pencil beam.

describes the distribution of range reduction caused by the ridge filter [46]:

$$L_{mod}(s) = \int_{t_{min}}^{t_{max}} w(t) \times L(s+t)\mathrm{d}t \qquad (3.124)$$

$L_{mod}(s)$ is the Bragg curve filtered by the ridge filter (see Sec. 5.3 for examples). The function $w(t)$ is referred to as the *range reduction distribution* and depends on the shape of the ridges. The fraction of the beam with a range reduction in the interval $[t, t + \mathrm{d}t]$ is given by $w(t)\mathrm{d}t$. t_{min} and t_{max} are the minimum and maximum water equivalent thickness of the ridge filter, and $L_{mod}(s)$ the filtered Bragg curve (in water).

The range reduction distribution of a range shifter is a dirac delta function, $w(t) = \delta(T)$, which simply shifts the Bragg curve a distance T towards the surface.

An example of a triangular ridge profile and resulting range reduction distribution is shown in Fig. 3.27. If the ridge filter height t is expressed as a function of the transverse coordinate x, the range reduction distribution is given by the derivative of the inverse function, $x(t)$:

$$w(t) = \kappa \times \frac{\mathrm{d}}{\mathrm{d}t}x(t) \qquad (3.125)$$

where κ is a normalization constant (for the integral in Eq. 3.124 to be physically meaningful, the integral of $w(t)$ must be equal to one). Note that the transverse dimension, x, is plotted in arbitrary units: as long as the ridges are more narrow than the beam, the range reduction distribution will be the same (apart from the normalization constant).

The thicker the ridge filter, the fewer layers are necessary to produce a homogeneous SOBP, which reduces the overall irradiation time. The disadvantage with a thick ridge filter is that it inevitably scatters the beam, causing large beam sizes at the patient - especially if used with low energy protons (see Chapter 5). A compro-
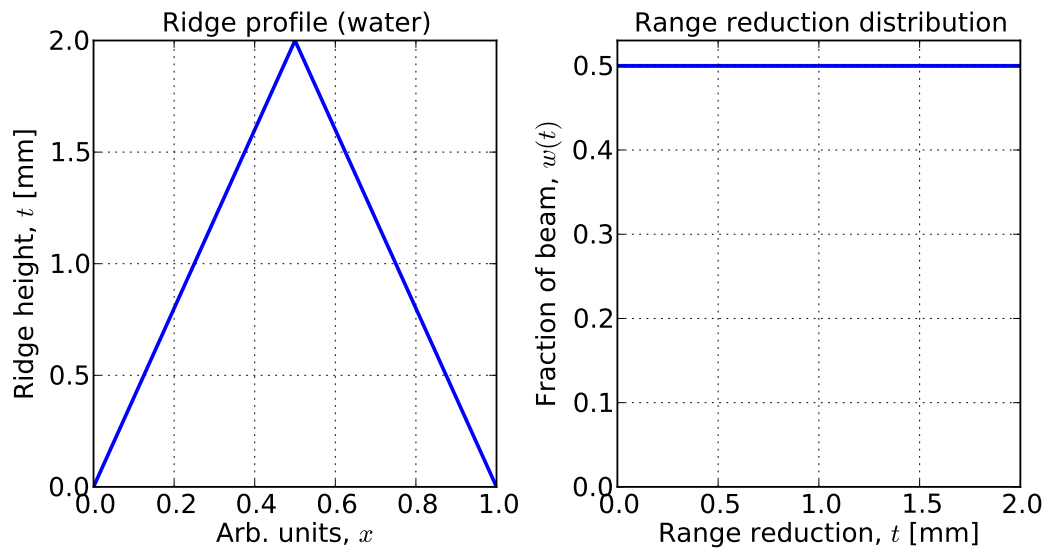
Figure 3.27: Left: triangular ridge filter profile, 2 mm peak height. Right: resulting range reduction distribution, $w(t)$.

mise must therefore be found between the number of layers and the resulting beam sizes at different energies.

3.8 Proton gantry

The proton gantry is a 15.6 m long (beam path), rotating construction that allows for irradiation of the patient from any direction. Fig. 3.28 shows the MedAustron proton gantry, which is a hardware copy of the state-of-the art Gantry 2 developed at PSI [13, 14]. The rotating vacuum pipe of the gantry needs to be separated from the fixed vacuum pipe upstream. Two solutions have been considered:

1. Separation of the vacuum pipes by breaking the vacuum. This is the simplest solution, used at PSI, but requires two vacuum windows.

2. Using a windowless joint. This is a more expensive solution (but commercially available), but avoids scattering of the proton beam.

The 90° bending dipole at the end of the gantry, just before the nozzle, has no dedicated vacuum chamber. Instead, the magnet yoke serves as vacuum chamber in order to increase the aperture and maximize the scanning area. In the event of a vacuum leakage, reparations of the 70 ton heavy dipole could be time-consuming. By filling the entire dipole with helium under atmospheric pressure (which has a lower scattering power than air), it would be robust against minor leakage. The helium could be contained by a large "balloon", or by a helium chamber.

Even though helium has a lower scattering power than air, the total beam path across the dipole is 2.36 m (402 cm from start of dipole to isocenter), which is a

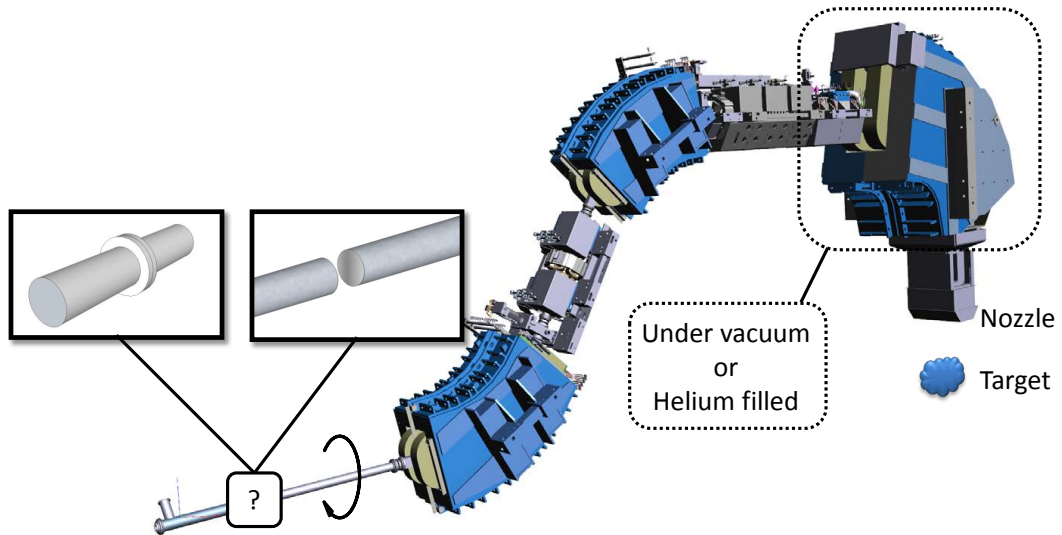considerable distance, which calls for an evaluation of the impact on the isocenter beam size.

Figure 3.28: The PSI/MedAustron proton gantry and scattering considerations. At the coupling point: windowless joint or broken vacuum with two vacuum windows. The 90° degree bending dipole: under vacuum or filled with helium. Total beam path from coupling point to isocenter is 15.58 m.

Chapter 4

Characteristics of extracted beam

4.1 Spill characteristics

4.1.1 Spill profile

The transit time is the time it takes for an unstable particle to reach the electrostatic septum after it has become unstable. Fig. 4.1 shows the horizontal phase space at the electrostatic septum during extraction. The acceleration pushes the beam closer to the resonance tune, shrinking the area of the stable triangle. Shown in the figure is one band of particles (between A and B) with identical amplitude and momentum that become unstable at the same moment. These particles will follow more or less the same trajectory and finally reach the electrostatic septum where they are extracted. However, they will not reach the electrostatic septum at the same moment. Particles becoming unstable close to B will reach ES first, followed by particles becoming unstable in the center region of the side AB.

Particles becoming unstable close to point A are still close to the fix-point, moving very slowly in the horizontal phase space (if tracked every 3 turns). But as the beam is accelerated, the separatrices are constantly shrinking the stable triangle. The moving separatrices will "catch up" on particles that became unstable close to point A, moving them into region 1 (indicated in the figure). Two turns later, these particles will be in region 2, close to the fixed point B. From there, they will soon reach the electrostatic septum and be extracted.

The density of particles along the line AB is as highest close to the corners of the triangle. An example of the horizontal particle distribution at the electrostatic septum before and after ramping of the sextupole is shown in Fig. 4.2. Particles are piling up close to the corner, due to their slower motion, whereas they are sparsely spaced along the sides of the triangle.

Quantitatively, the time profile of a group of particles with the same horizontal amplitude and momentum that are extracted can be described as an initial "spike", followed by a longer "tail". The spike corresponds to the corner particles, which reaches the septum first, followed by the particles along the side of the triangle, which taker a longer time to reach the septum. This profile is referred to as the *spill*
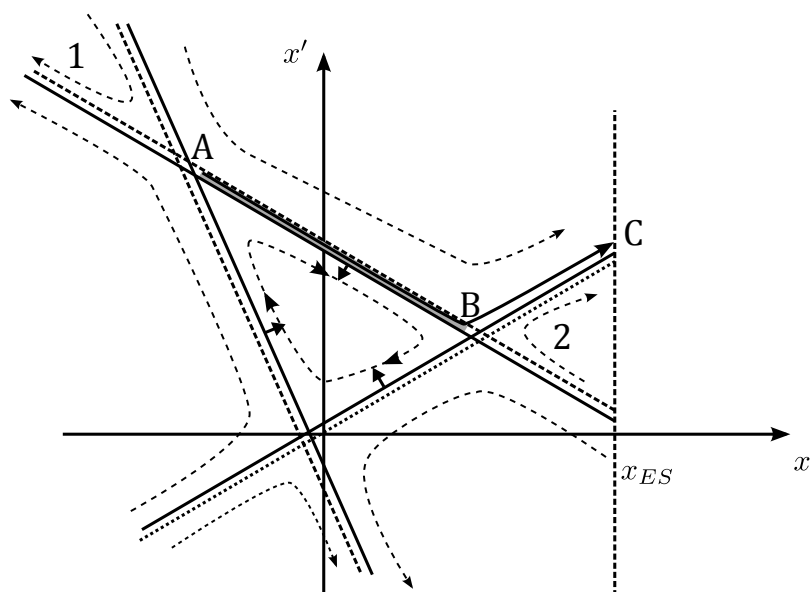
Figure 4.1: Horizontal phase space at electrostatic septum during extraction. Acceleration of the beam shrinks the stable triangle and the strip of particles between A and B becomes unstable. Note that the dispersion at the ES shifts the center of the stable triangle away from the origin.
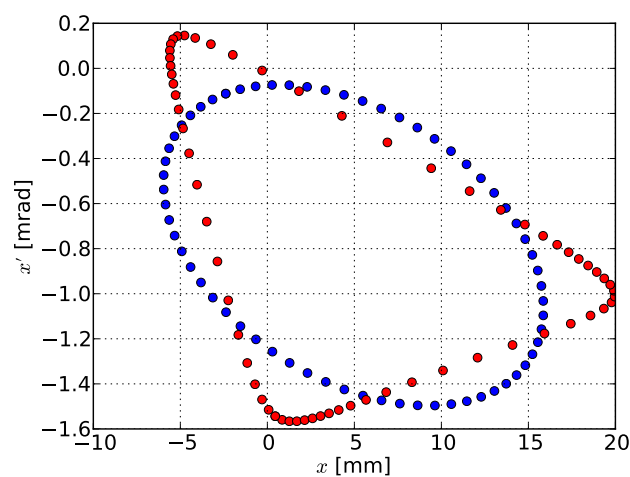

Figure 4.2: Horizontal particle distribution at the electrostatic septum before and after ramping of the resonance sextupole. All particles have initially the same emittance ($\varepsilon_y = 0$) and momentum. Note that dispersion at the septum shifts the stable triangle away from origin.

profile.

A rigorous mathematical analysis of the spill profile can be found in e.g. [11,29], leading to the conclusion that about half the particles will reach the septum in a spike after a time $T(\xi)$, whereas the other half will be distributed over a tail between $T(\xi)$ and $2 \times T(\xi)$, as shown in Fig. 4.3.

(a) Two groups of particles of different emittance and momentum becoming unstable at the same moment.
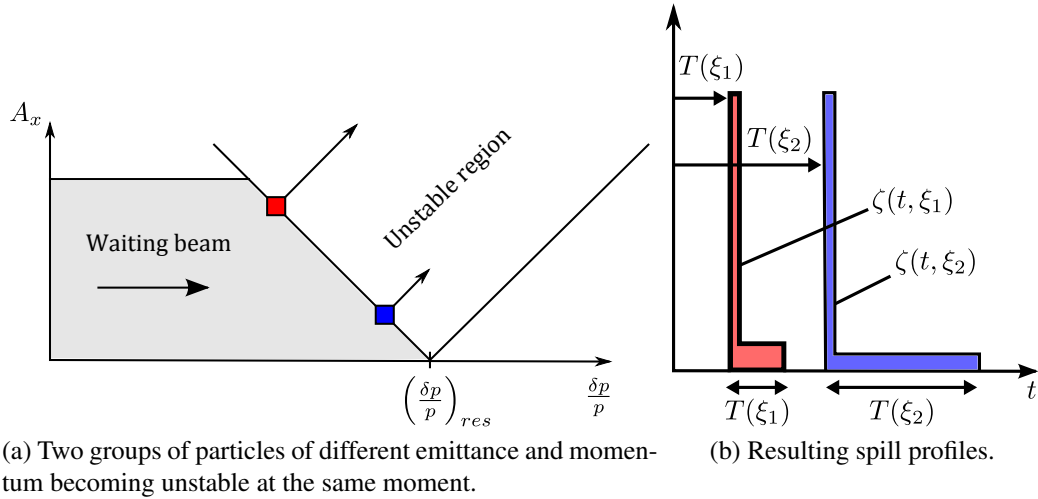
(b) Resulting spill profiles.

Figure 4.3: Resulting spill profiles from two groups of particles along the limit of stability that become unstable at the same moment.

$T(\xi)$ is called the *transit time*, and is the time it takes a corner particle to reach the electrostatic septum:

$$T(\xi) = \frac{3T_{rev}}{\sqrt{3}\xi} \ln \left| \frac{3\sqrt{3}}{|\dot{\xi}_0'|} \xi^2 \right| \tag{4.1}$$

Using normalized time units, t':

$$t' = \frac{t}{3T_{rev}} \tag{3.51}$$

the transit time formula is:

$$T'(\xi) = \frac{1}{\sqrt{3}\xi} \ln \left| \frac{3\sqrt{3}}{|\dot{\xi}_0'|} \xi^2 \right| \tag{4.2}$$

$\dot{\xi}_0'$ is the nominal *modified tune velocity* - a measure of how fast the betatron core

shifts the tune of the beam:

$$\dot{\xi}_0' = \frac{d\xi}{dt'}$$
$$= 3T_{rev} \times \frac{d\xi}{dt} =$$
$$= 3T_{rev} \times \frac{d}{dt}(6\pi\delta Q)$$
$$= 3T_{rev} \times 6\pi Q_x' \frac{d}{dt}\left(\frac{\delta p}{p}\right)$$
$$\approx 3T_{rev} \times 6\pi Q_x' \frac{\Delta p/p}{T_{extr}} \tag{4.3}$$

where the final approximation is made by setting the ramp rate of the betatron core such that the momentum increment during the extraction time, T_{extr}, is equal to the momentum spread of the beam, $\Delta p/p$. Note that, since the chromaticity of the MedAustron synchrotron is negative (Tab. A.2), $\dot{\xi}_0'$ is also negative. This means that accelerating beam will approach the resonance tune from a higher tune. This technicality is neglected in the Steinbach diagrams, where the waiting beam is drawn *below* the resonance tune.

The transit time formula is divergent at $\xi = 0$, and in order for it to be valid, the ratio between the change of the area of the stable triangle during transition and the tune distance to the resonance should not be too large. This is formulated by the condition [11]:

$$\frac{|\dot{\xi}_0'|}{\xi^2} < 0.05 \tag{4.4}$$

Fig. 4.4 shows an example of the transit time as a function of the modified tune distance, using parameters from App. A and 1 s extraction time.
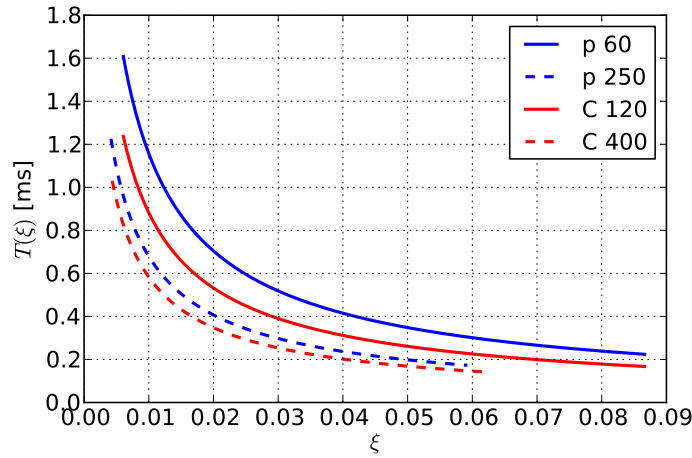

Figure 4.4: Transit time as a function of modified tune distance.

4.1.2 Instantaneous tune shift

The effect of an instantaneous, infinitesimal tune shift ΔQ (either by shifting the tune of the machine, or the tune of the beam by accelerating it towards the resonance) is illustrated in the Steinbach diagram in Fig. 4.5. A band of particles along the limit of stability will enter the resonance and be extracted.
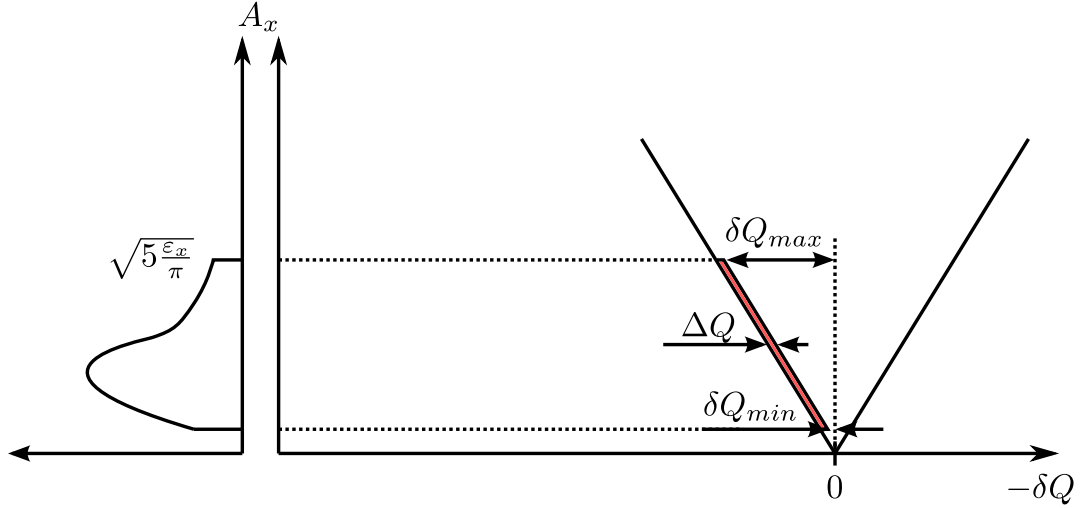

Figure 4.5: At an instantaneous tune shift, a "band" of particles (marked red) along the limit of stability becomes unstable.

The time profile of the extracted band can be calculated by integrating the spike and tail profiles along the entire band in the Steinbach diagram [11, 29]. Assuming the number of particles in the band is N_{band}, the extracted time profile is:

$$N_{band} \times \zeta(t) = N_{band} \times \int_{\xi_{min}}^{\xi_{max}} \rho(\xi)[\zeta_{spike}(t,\xi) + \zeta_{tail}(t,\xi)]\mathrm{d}\xi$$
$$= N_{band} \times (\zeta_{spike}(t) + \zeta_{band}(t)) \tag{4.5}$$

where $\rho(\xi)$ is the particle density at a modified tune distance ξ. Note that since the particle distribution is uniform in momentum, the particle density along the limit of stability is completely defined from the particle distribution in A_x. Since:

$$\rho(A_x) = \frac{A_x}{\varepsilon_x/\pi} \exp\left(-\frac{A_x^2}{2 \times \varepsilon_x/\pi}\right) \tag{3.21}$$

and, along the limit of stability (Eq. 3.59):

$$A_x = \sqrt{48\sqrt{3}\pi} \left|\frac{\delta Q}{S}\right| = \sqrt{\frac{4}{\sqrt{3}\pi}} \left|\frac{\xi}{S}\right| \tag{4.6}$$

the particle density along the limit of stability as a function of ξ is:

$$\rho(\xi) = \frac{\xi}{\sigma_\xi} \exp\left(-\frac{\xi^2}{2\sigma_\xi^2}\right) \tag{4.7}$$

65

where:

$$\sigma_\xi = S\sqrt{\frac{\sqrt{3}\pi}{4} \times \frac{\varepsilon_x}{\pi}} \tag{4.8}$$

The upper integration limit, ξ_{max}, is given by the truncation level $A_x = \sqrt{5\varepsilon_x/\pi}$:

$$A_{x,max} = \sqrt{5\varepsilon_x/\pi} \Rightarrow \xi_{max} = \sqrt{5}\sigma_\xi \tag{4.9}$$

In order for the transition formula to be valid, Eq. 4.4 must not be violated for the lower integration limit. A lower limit of:

$$\xi_{min} = \kappa \times \xi_{max} \tag{4.10}$$

$$\kappa = 0.07 \tag{4.11}$$

has been used in this document. With this value of κ, Eq. 4.4 is fulfilled for all foreseen beam configurations (highest and lowest proton and carbon energies, shortest and longest extraction time). This creates a "hollow" beam, where particles with an amplitude lower than $\kappa\sqrt{5\varepsilon_x/\pi}$ are neglected. However, this approximation has little effect since the neglected fraction is only about 1.3% of the total beam:

$$\text{Neglected fraction:} \quad \frac{\int_0^{\xi_{min}} \rho(\xi)d\xi}{\int_0^{\xi_{max}} \rho(\xi)d\xi} \approx 1.3\% \tag{4.12}$$

The spike profile, $\zeta_{spike}(t,\xi)$ is modeled as a Dirac-delta function, and the tail $\zeta_{tail}(t,\xi)$ as a flat plateau, each with weight $1/2$:

$$\zeta_{spike}(t,\xi) = \frac{1}{2}\delta(t - T(\xi))$$
$$\zeta_{tail}(t,\xi) = \frac{1}{2} \times \frac{H(t - T(\xi)) - H(t - 2T(\xi))}{T(\xi)} \tag{4.13}$$

$H(t)$ is the Heaviside step function:

$$H(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases} \tag{4.14}$$

The sum of the spike and tail is the spill profile, $\zeta(t,\xi)$, shown in Fig. 4.3b:

$$\zeta(t,\xi) = \zeta_{spike}(t,\xi) + \zeta_{tail}(t,\xi) \tag{4.15}$$

The integration of all spikes along the limit of stability can now be calculated to:

$$\zeta_{spike}(t_i) = \int_{\xi_{min}}^{\xi_{max}} \rho(\xi)\delta(t_i - T(\xi))d\xi =$$
$$= \begin{cases} \frac{1}{2} \times \frac{\rho(\xi_i)}{\frac{dT}{d\xi}\big|_{\xi_i}} & \text{if } T(\xi_{max}) \leq t_i \leq T(\xi_{min}) \\ 0 & \text{otherwise} \end{cases} \tag{4.16}$$

where ξ_i is given by the inverse transit time:

$$\xi_i = T^{-1}(t_i) \tag{4.17}$$

and the derivative of the transit time is:

$$\frac{\mathrm{d}T}{\mathrm{d}\xi} = 3T_{rev} \times \frac{2}{\sqrt{3}\xi^2} \times \left[1 - \ln\left|\sqrt{\frac{3\sqrt{3}}{|\dot{\xi}_0'|}}\xi\right|\right] \tag{4.18}$$

The tail profile can be rewritten as:

$$\begin{aligned}
\zeta_{tail}(t_i) &= \int_{\xi_{min}}^{\xi_{max}} \frac{1}{2} \times \frac{H(t_i - T(\xi)) - H(t_i - 2T(\xi))}{T(\xi)}\rho(\xi)\mathrm{d}\xi \\
&= \begin{cases} \int_{\xi_{low}}^{\xi_{high}} \frac{\rho(\xi)}{T(\xi)}\mathrm{d}\xi & \text{if } T(\xi_{max}) \leq t_i \leq 2T(\xi_{min}) \\ 0 & \text{otherwise} \end{cases}
\end{aligned} \tag{4.19}$$

where ξ_{low} and ξ_{high} are given by the inverse of the transit time formula:

$$\begin{aligned}
\xi_{low} &= \max\{\xi_{min}, T^{-1}(t_i)\} \\
\xi_{high} &= \min\{\xi_{max}, T^{-1}(t_i/2)\}
\end{aligned} \tag{4.20}$$

Fig. 4.6 shows $\zeta(t)$, $\zeta_{spike}(t)$ and ζ_{tail} for highest and lowest proton and carbon ion extraction energies. $\zeta_{spike}(t)$ and ζ_{tail} have been calculated numerically from Eq. 4.16 and Eq. 4.19.

High energy carbon ions exhibits the narrowest spill profile, while it is as widest for low energy proton beams.

4.2 Tune ripple spill generator

Current fluctuations in the synchrotron magnets will transform into ripple of the machine tune. The analytical spill model in [11] has been extended in order to evaluate the effect of current ripple on extraction stability.

4.2.1 Tune ripple

As mentioned previously, the beam is extracted by acceleration, pushing the particles into the unstable region in the Steinbach diagram. However, this is in principle equivalent to shifting the machine tune towards the waiting beam. As mentioned, the chromaticity of the MedAustron synchrotron is negative: during extraction, the beam approaches the resonance tune from above. In the analytical spill model, the beam is instead extracted by changing the tune of the machine[1].

---

[1]The momentum (energy) of the extracted beam would then be time dependent, since the changing machine tune "scrapes" the but for a study of the extraction stability, this difference has no importance.
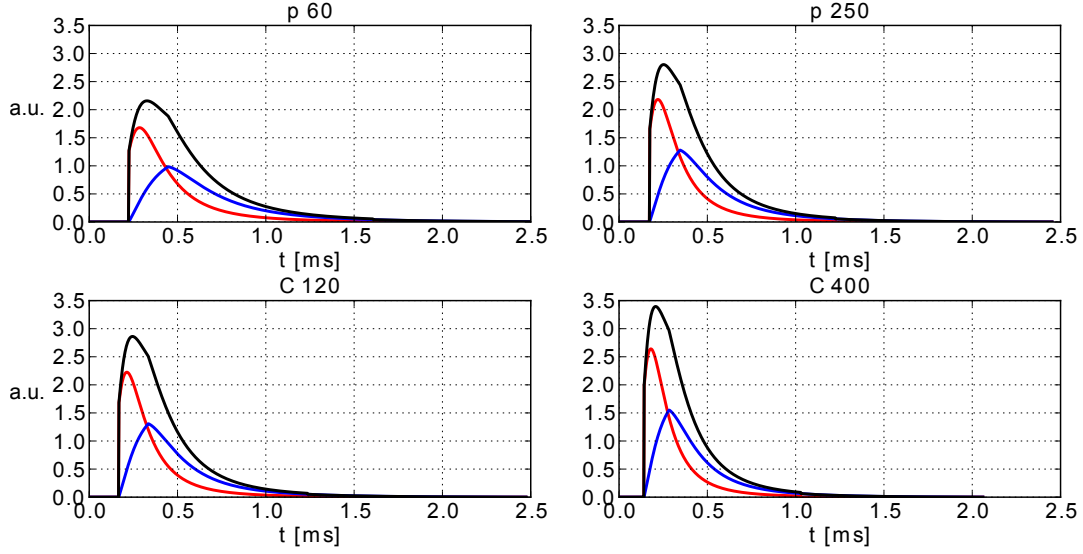
Figure 4.6: $\zeta(t)$ (black) is the extraction time profile from a narrow band of particles along the limit of stability. Red is contribution from spikes, $\zeta_{spike}(t)$ and blue contribution from tails, $\zeta_{tail}(t)$. Subfigures show cases for highest and lowest proton and carbon extraction energies.

With no tune ripple, the horizontal tune of the machine is:

$$Q_0(t) = Q_0 + \dot{Q}_0 t \qquad (4.21)$$

where Q_0 is the resonance tune and $\dot{Q}_0$ the nominal tune velocity, estimated to (see Eq. 4.3):

$$\dot{Q}_0 \approx -Q'_x \frac{\Delta p/p}{T_{extr}} \qquad (4.22)$$

The minus-sign is inserted since it is now the machine tune moving towards the beam (increasing tune), instead of the beam being pushed towards the resonance tune via the chromaticity (decreasing tune of the beam). The *relative* movement between beam and resonance tune is unaffected.

During a short time Δt, the tune shifts by an amount:

$$\Delta Q = \dot{Q}_0 \Delta t \qquad (4.23)$$

This is the tune width of the red band shown in Fig. 4.5. The number of particles contained in this band is:

$$\Delta N_{in} = \frac{dN}{dQ} \times \Delta Q \qquad (4.24)$$

Since the particle density is taken to be uniform in momentum space, dN/dQ is also taken to be constant due to the interchangeability of tune and momentum via the chromaticity.

Now, assume that a sinusoidal tune ripple with frequency $f_r = \omega_r/2\pi$ and amplitude and phase δQ_r and φ_r is added to the nominal tune. The tune and tune velocity

are:

$$Q(t) = Q_0 + \dot{Q}_0 t + \delta Q_r \sin(\omega_r t + \varphi_r)$$
$$\dot{Q}(t) = \dot{Q}_0 + \omega_r \delta Q_r \cos(\omega_r t + \varphi_r) \qquad (4.25)$$

If $|\omega_r \delta Q_r| \geq |\dot{Q}_0|$, the tune velocity will periodically be zero, or even negative. When this happens, the beam does not come closer to the resonance tune and no new particles will enter the resonance. As soon as all particles still inside the resonance have been extracted, the extraction intensity will be zero, and the beam is said to be chopped.

A schematic example of the tune curve is shown in Fig. 4.7. At $t = t_1$, the tune velocity is zero, and a band containing N_{in} particles have just entered the resonance. At $t = t_2$, the tune is moving away from the beam, and no particles are entering the resonance ($\dot{N}_{in} = 0$). At $t = t_3$, the tune moves towards the beam, but no particles will be extracted until the tune has returned to its previous maximum, $Q(t_1)$.
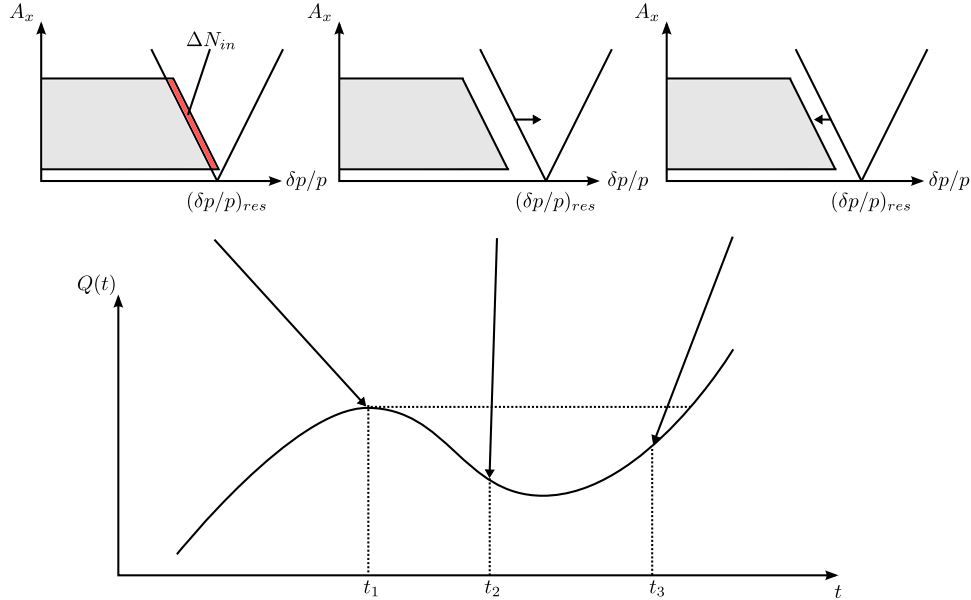
Figure 4.7: Schematic illustration of tune ripple, when $|\omega_r \delta Q_r| \geq |\dot{Q}_0|$: while the tune is lower than any previous value, no particles will enter the resonance.

4.2.2 Extracted particle rate

In order for any particles to enter the resonance, the tune must change to a new maximum level. Mathematically, the rate of particles entering the resonance, $\dot{N}_{in}(t)$ (particles per second) can thus be described by:

$$\dot{N}_{in}(t) = \dot{Q}(t) \times H \left[Q(t) - \max_{t'} \left\{ Q(t' \in [0,t[) \right\} \right] \times \frac{dN}{dQ} \qquad (4.26)$$

The Heaviside function is equal to 1 only when the tune is increasing to a new maximum value, otherwise it is zero.

The extracted spill profile $\zeta(t)$ from an instantaneous tune shift has been described in Sec. 4.1.2. By treating the synchrotron as a linear system with impulse response $\zeta(t)$ and input signal $\dot{N}_{in}(t)$, the extracted particle rate, $\dot{N}_{extr}(t)$ is given by a convolution between $\dot{N}_{in}(t)$ and $\zeta(t)$:

$$\dot{N}_{extr}(t) = (\dot{N}_{in} * \zeta)(t) \tag{4.27}$$

The width of the spill response $\zeta(t)$ is, as seen in Fig. 4.6 around 0.5 ms, which has the effect that ripple frequencies above a few kHz will be filtered by the finite spill response, while the extracted particle rate at lower frequencies is close to the rate of particles entering the resonance.

4.2.3 Current ripple to tune ripple conversion

4.2.3.1 Overview

Ripple in the current of the synchrotron magnets causes ripple in the magnetic field, which is transformed into tune ripple of the synchrotron. A simplified block diagram of how the reference current, I_0, given to the synchrotron magnet power supplies, is transformed into beam tune is shown in Fig. 4.8. A current reference value, I_0, is given to the magnet power supplies, which calculate the voltage (U_{ref}) that needs to be applied over the magnet in order to reach I_0. The voltage is generated via pulse width modulation (PWM).
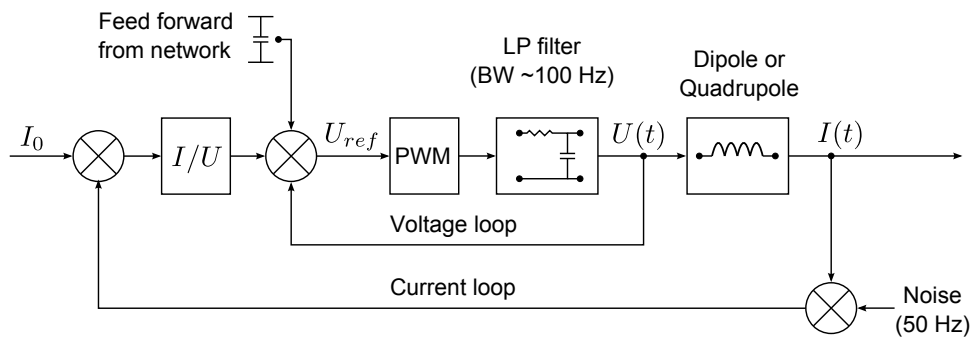
(a) Overview of the magnet power supply functionality
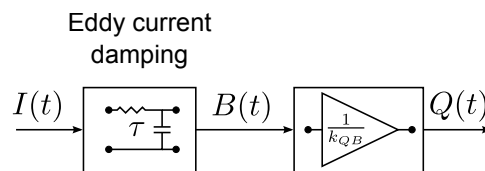
(b) Current to tune conversion

Figure 4.8: Sources of current ripple and transformation into tune ripple.

In order to suppress voltage ripple due to the finite frequency of the PWM (kHz range), it is followed by a low-pass filter with a bandwidth in the order of 100 Hz. Further noise reduction is made by using the output voltage in a feedback

loop (*voltage loop*, Fig. 4.8a). Additionally, the network voltage is continuously monitored and fluctuations are taken into account when calculating U_{ref}.

The resulting voltage $U(t)$ is applied over the magnets to generate a current $I(t)$. Measurement of $I(t)$ is used in another feedback loop (*current loop*, Fig. 4.8a) into which additional noise may seep.

The current flowing through the magnet generates a magnetic field, $B(t)$. Fluctuations in the current will be damped by eddy currents, with a time constant τ.

Finally, the field in the magnetic elements of the synchrotron will determine the tune of the beam. Ripple in the magnetic field will be transformed to tune ripple.

The bandwidth of the voltage loop is about 1 kHz, while the bandwidth of the current loop is in the order of 1 kHz[2].

4.2.3.2 Main current ripple source

The magnets of the synchrotron are divided into families, according to their type (one dipole family, three quadrupole families etc.). All magnets belonging to the same family are powered in series, such that the current of e.g. all dipole magnets is the same.

The dominant current ripple source is expected to be that from the pulse width modulation in the magnet power supplies, which have a known frequency. For the synchrotron dipoles, the frequency f_{PWM} is assumed to be 2 kHz, while the power supplies for the main ring quadrupoles are assumed to have a PWM frequency of about 12 kHz.

4.2.3.3 Power converter accuracy

Typically, the accuracy of the current generated by the power supplies is specified to be better than some fraction $\delta I_{r,max}/I_{max}$ of the *maximum* operating current that the power supplies can deliver. The required current is proportional to the magnetic rigidity of the beam. For a beam with magnetic rigidity $B\rho$, nominal current I_0 and a current accuracy of δI_r, the *relative* current accuracy can be written as:

$$\frac{\delta I_r}{I_0} = \frac{\delta I_r}{I_{max}} \times \frac{(B\rho)_{max}}{(B\rho)_0} \tag{4.28}$$

where $(B\rho)_{max}$ is the highest foreseen magnetic beam rigidity (that of 400 MeV/n carbon ions). If the power converter follows specifications, the fraction $\delta I_r/I_{max}$ will never be larger than the specified current accuracy $\delta I_{r,max}/I_{max}$.

The magnetic rigidity ratio between the foreseen lower energy proton beams and highest energy carbon ion beams is about 6 (see Tab. A.3). Thus, the relative power supply accuracy for low energy proton beams will be 6 times worse compared to a carbon ion beam at highest extraction energy.

---

[2]Private communication, Power Converter group.

4.2.3.4 Eddy current damping

Sinusoidal ripple of the current $I(t)$ will be transformed into sinusoidal variations of the magnetic field, $B(t)$. However, the relative amplitude of the ripple will be damped by eddy currents in the magnets' yoke-lamination and the vacuum chamber. The magnitude of the damping is frequency dependent. Assume that the steady-state current I_0 gives a static magnetic field of strength B_0, and that a sinusoidal ripple of amplitude δI_r and frequency $f_r = \omega_r/2\pi$ is added to the current:

$$I(t) = I_0 + \delta I_r \sin(\omega_r t) \tag{4.29}$$

The magnetic field will follow the current (with some phase shift ϕ_r):

$$B(t) = B_0 + \delta B_r \sin(\omega_r t + \phi_r) \tag{4.30}$$

The relation between $\delta I_r/I_0$ and $\delta B_r/B_0$ is given by the eddy current time constant τ (see e.g. [11, Sec. 5.11]):

$$\frac{\delta B_r}{B_0} = \frac{\delta I_r}{I_0} \frac{1}{\sqrt{1 + (\omega_r \tau)^2}} \tag{4.31}$$

where the factor $\sqrt{1 + (\omega_r \tau)^2}$ is referred to as the damping factor. The time constant τ is estimated to about $100\,\mu$s, which gives a damping factor of about 1.6 at a ripple frequency of $f_r = 2$ kHz, and 7.6 at $f_r = 12$ kHz.

4.2.3.5 Magnetic field to tune conversion

Fluctuations in the magnetic field, δB_r, are assumed to be instantaneously transferred to variation in the beam tune, δQ_r. The ratio is assumed to be linear, with a proportionality constant k_{QB}:

$$\frac{\delta B_r}{B_0} = k_{QB} \frac{\delta Q_r}{Q_0} \Leftrightarrow \delta Q_r = \frac{Q_0}{k_{QB}} \frac{\delta B_r}{B_0} \tag{4.32}$$

The constant k_{QB} depends on magnet type and the lattice parameters of the synchrotron. The lower the value of k_{QB}, the higher the impact on particle rate stability (given by $\delta Q_r/Q_0$) for a given value of $\delta I_r/I_{max}$. It is estimated that k_{QB} will be two orders of magnitude larger for the synchrotron sextupoles, compared to the dipoles and quadrupoles [12]. For this reason, current ripple in the sextupoles has not been evaluated.

Values of k_{QB} for the dipoles and quadrupole families (from [11]) are cited in Tab. 4.1.

4.2.3.6 Summary

To summarize, the tune ripple caused by a current ripple of relative amplitude $\delta I_r/I_0$ and frequency $f_r = \omega_r/2\pi$ in one of the magnet families can be written as:

$$Q_r(t) = \frac{Q_0}{k_{QB}} \frac{(B\rho)_{max}}{B\rho} \frac{\sin(\omega_r t + \phi_r)}{\sqrt{1 + (\omega_r \tau)^2}} \frac{\delta I_r}{I_{max}} \tag{4.33}$$

Combining this with Eq. 4.25 gives:

$$Q(t) = Q_0 + \dot{Q}_0 t + \frac{Q_0}{k_{QB}} \frac{(B\rho)_{max}}{B\rho} \frac{\sin(\omega_r t + \varphi_r)}{\sqrt{1 + (\omega_r \tau)^2}} \frac{\delta I_r}{I_{max}} \tag{4.34}$$

and a tune derivative of:

$$\dot{Q}(t) = \dot{Q}_0 + \frac{\omega_r}{\sqrt{1 + (\omega_r \tau)^2}} \frac{Q_0}{k_{QB}} \frac{(B\rho)_{max}}{B\rho} \frac{\delta I_r}{I_{max}} \cos(\omega_r t + \varphi_r) \tag{4.35}$$

which sets the rate of particles entering the resonance. The tune ripple amplitude δQ_r is equal to:

$$\delta Q_r = \frac{1}{\sqrt{1 + (\omega_r \tau)^2}} \frac{Q_0}{k_{QB}} \frac{(B\rho)_{max}}{B\rho} \frac{\delta I_r}{I_{max}} \tag{4.36}$$

The width of the spill profile, $\zeta(t)$ is in the order of half a millisecond (Fig. 4.6). At low frequency tune ripple ($\ll 2$ kHz), the convolution between the narrow spill profile and the rate of particle entering the resonance (Eq. 4.27) will leave $\dot{N}_{in}(t)$ more or less unaffected, i.e. $\dot{N}_{extr}(t) \approx N_{in}(t)$.

At lower frequencies, the width of the spill profile $\zeta(t)$ is negligible, and the extracted particle rate is approximately equal to the rate of particles entering the resonance. Fluctuations of the extracted particle rate are then approximately given by the ratio between $\omega_r \delta Q_r$ and $\dot{Q}(t)$:

$$\text{Low frequencies:} \quad \frac{\dot{N}_{extr}(t)}{\dot{N}_0} \approx \left(1 + \frac{\omega_r \delta Q_r}{\dot{Q}_0}\right) \cos(\omega_r t) \tag{4.37}$$

where $\dot{N}_0$ is the nominal particle extraction rate. Beam chopping ($\dot{N}_{extr} = 0$) can occur if:

$$\left| \frac{\omega_r \delta Q_r}{\dot{Q}_0} \right| > 1 \tag{4.38}$$

If relative fluctuations of the particle rate must be smaller than $\varepsilon_{\dot{N}}$ (e.g. max 50% variations of the particle rate) at a particular frequency $f_r = \omega_r / 2\pi$ in order to guarantee a homogeneous target dose distribution, the accuracy of the synchrotron magnet power supplies must meet the following condition:

$$\frac{\delta I_r}{I_{max}} < \frac{\dot{Q}_0}{Q_0} \frac{\sqrt{1 + (\omega_r \tau)^2}}{\omega_r} \frac{B\rho}{(B\rho)_{max}} k_{QB} \varepsilon_{\dot{N}} \tag{4.39}$$

As an example, consider a 400 MeV carbon ion beam restricted to 50% intensity fluctuations at $f_r = 100$ Hz. Using tune ripple parameters from Tab. 4.1 and considering only contribution from the quadrupoles (QF1+QF2+QD) gives ($k_{QB} = 1$, $\tau = 60~\mu$s, $B\rho/(B\rho)_{max} = 1$, $\omega_r = 200\pi$ s^{-1}, $Q_0 = 5/3$, $\dot{Q}_0 = 4.4 \times 10^{-3}$):

$$\frac{\delta I_r}{I_{max}} < 2.1~\text{ppm}$$

At higher frequencies, the finite width of the spill profile will filter out the tune ripple, and the maximum fluctuations of $\dot{N}_{extr}$ should be evaluated via Eq. 4.27.

Fig. 4.9 shows an example of the particle rate entering the resonance and the extracted particle rate when a 60 MeV proton beam is subject to a tune ripple caused by 0.5 ppm sinusoidal current ripple in the synchrotron dipoles. The rate of particles entering the resonance varies between 0% and more than 2.5 times the nominal rate. The extracted particle rate, however, is filtered by the spill response, and varies "only" between arond $\pm 30\%$ of nominal particle rate. It takes, in this case, about 1 ms for $\dot{N}_{extr}$ to stabilize around the nominal value.
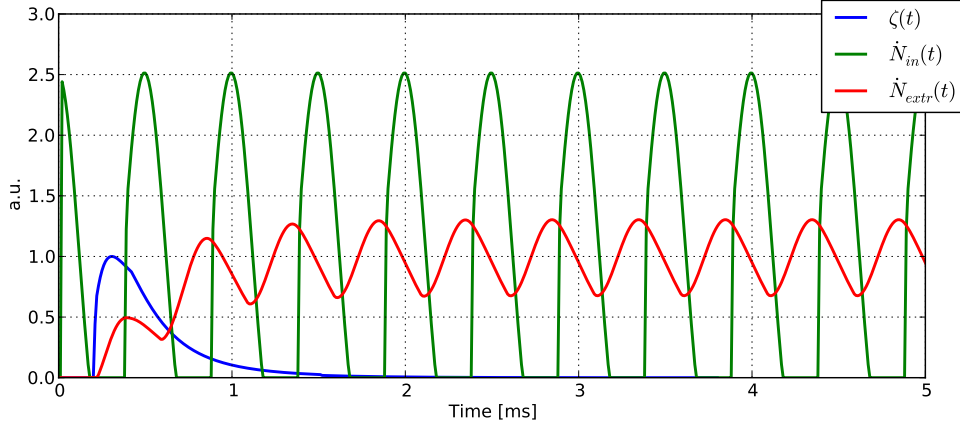

Figure 4.9: Comparison between $\dot{N}_{in}(t)$ and $\dot{N}_{extr}(t)$ for a 60 MeV proton beam subject to 0.5 ppm tune ripple at 2 kHz tune ripple and resulting particle rates entering and leaving the resonance (60 MeV protons). For comparison, the spill response $\zeta(t)$ is also shown.

4.2.4 RF-Channelling

A detailed description of RF-channelling can be found in [11, 50]. The general idea is to increase the local value of the tune velocity $\dot{Q}_0$ for particles that are close to the resonance, as illustrated in Fig 4.10. The impact of tune ripple caused by power supply ripple would then be mitigated, since the ratio $\delta Q_r / \dot{Q}_0$ is reduced for particles close to the resonance.

A simplified model of RF-channelling is to multiply the nominal tune velocity $\dot{Q}_0$ with a factor K_{RF}, where K_{RF} is the relative increase in tune velocity introduced by RF-channelling. The expression for the tune then becomes:

$$Q(t) = Q_0 + K_{RF}\dot{Q}_0(t) + \frac{Q_0}{k_{QB}} \frac{(B\rho)_{max}}{B\rho} \frac{\sin(\omega_r t + \varphi_r)}{\sqrt{1 + (\omega_r \tau)^2}} \frac{\delta I_r}{I_{max}} \qquad (4.40)$$

From this, one can calculate the rate of particles entering the resonance exactly as before. However, the transition formula will no longer be valid, since the time the particles spend in the resonance is expected to be significantly shorter than before due to the increased tune velocity. To take this into account, the spill response $\zeta(t)$ is
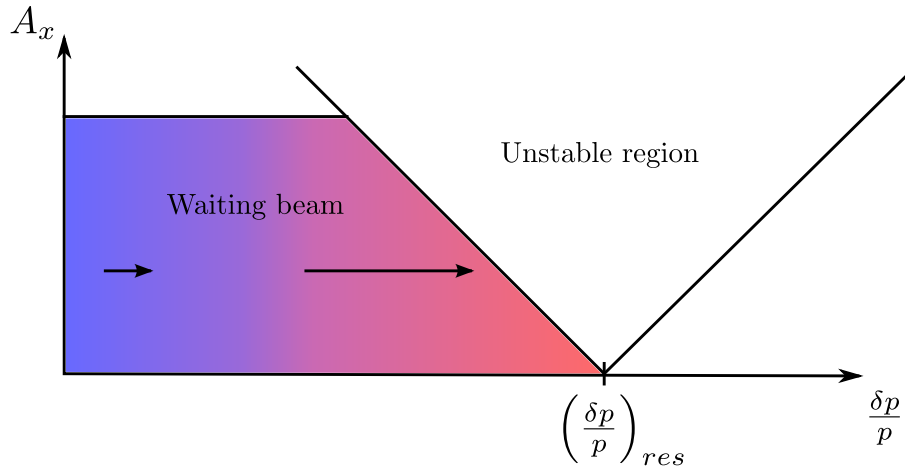
Figure 4.10: Principle of RF-channelling: particle close to the resonance tune have a higher tune velocity

replaced by a dirac-delta function, which has the effect of making the rate of particles leaving the resonance (extracted) exactly equal to the rate of particles entering the resonance:

$$\zeta_{RF}(t) = \delta(t) \tag{4.41}$$

The maximum relative fluctuations in the extracted particle rate is then given by the ratio between the amplitudes $K_{RF}\dot{Q}_0$ and $\dot{Q}_r$:

$$\max\left(\frac{\dot{N}_{extr}}{\dot{N}_0}\right) = \frac{\max(\dot{Q}_r)}{K_{RF}\dot{Q}_0} = \frac{1}{K_{RF}\dot{Q}_0}\frac{Q_0}{k_{QB}}\frac{\delta I_r}{I_{max}}\frac{(B\rho)_{max}}{B\rho}\frac{1}{\sqrt{1+(\omega_r\tau)^2}} \tag{4.42}$$

The factor K_{RF} is given by [11]:

$$K_{RF} = \sqrt{\frac{\pi}{\sin\varphi_s}} \tag{4.43}$$

where:

$$\sin\varphi_s = \frac{C \times B\rho\frac{\Delta p}{p}}{V_{RF}T_{extr}} \tag{4.44}$$

C is the synchrotron circumference and V_{RF} the voltage used for the RF-channelling. K_{RF} is typically of the order of 100 (see Tab. 4.1 for numerical values for different beam configurations). The value of $V_{RF} = 3$ kV is equivalent to the voltage anyway required for the foreseen beam acceleration at a dipole ramp rate of 3 T/s. Up to 5 kV may be available.

4.2.5 Tune ripple evaluation

Using the described spill generator, the impact of magnet current ripple on extraction rate has been evaluated for different current ripple frequencies and amplitudes.

	Comment	Value
$\lvert \dot{Q}_0 \rvert$		0.016 s^{-1}
f_{PWM}	Dipoles	2 kHz
	Quadrupoles (QF1, QF2, QD)	12 kHz
$\lvert k_{QB} \rvert$	Dipoles	1.5
	QF1	0.5
	QF2	0.7
	QD	0.4
	QF1+QF2+QD	≈ 1
τ	Dipoles	$\approx 100 \, \mu\text{s}$
	Quadrupoles	$\approx 60 \, \mu\text{s}$
ξ_{max}	60 MeV protons	0.0866
	250 MeV protons	0.0592
	120 MeV carbon ions	0.0866
	400 MeV carbon ions	0.0620
V_{RF}		3 kV
K_{RF}	60 MeV protons	166
	250 MeV protons	114
	120 MeV carbon ions	98
	400 MeV carbon ions	70

Table 4.1: Tune ripple parameters ($T_{extr} = 1.0$ s).

The maximum extraction ripple amplitude for a sinusoidal current ripple amplitude of 2 ppm is shown in Fig. 4.11 for the protons and carbon ions at lowest and highest extraction energy. Dipole parameters from Tab. 4.1 have been used. The 2 ppm current ripple causes the largest extraction ripple in the region around 1 kHz. Most sensitive (highest extraction ripple amplitude) is the 60 MeV proton beam. This is to be expected, due to its low magnetic rigidity. At lower frequencies, the extraction ripple increases linearly with ripple frequency, as predicted in Eq. 4.38. At higher frequencies, eddy current damping and the filtering effect of the spill response suppresses the current ripple.

A contour plot of the extraction ripple amplitude as a function of both current ripple frequency and amplitude (in the dipoles) is shown in Fig. 4.12. In general, low energy protons are the most sensitive to power supply ripple. An exception is at high frequencies (above $\sim$10 kHz), where the wider proton spill response (see Fig. 4.6) has a stronger filtering effect on the current ripple, compared to carbon ions.

At high frequencies (few kHz), the extraction ripple saturates: an increased current ripple amplitude for a given frequency does not affect the beam intensity ripple noticeably. This is due to the filtering effect of the spill response. In the extreme case, $\dot{N}_{in}(t)$ is a "dirac comb" (periodic spikes of infinite amplitude). A convolution with $\zeta(t)$ still gives a finite response $\dot{N}_{extr}(t)$. As seen in the figure, an increment of the current ripple from 10^{-5} to 10^{-4} barely affects the particle rate ripple amplitude for frequencies above 1 kHz.

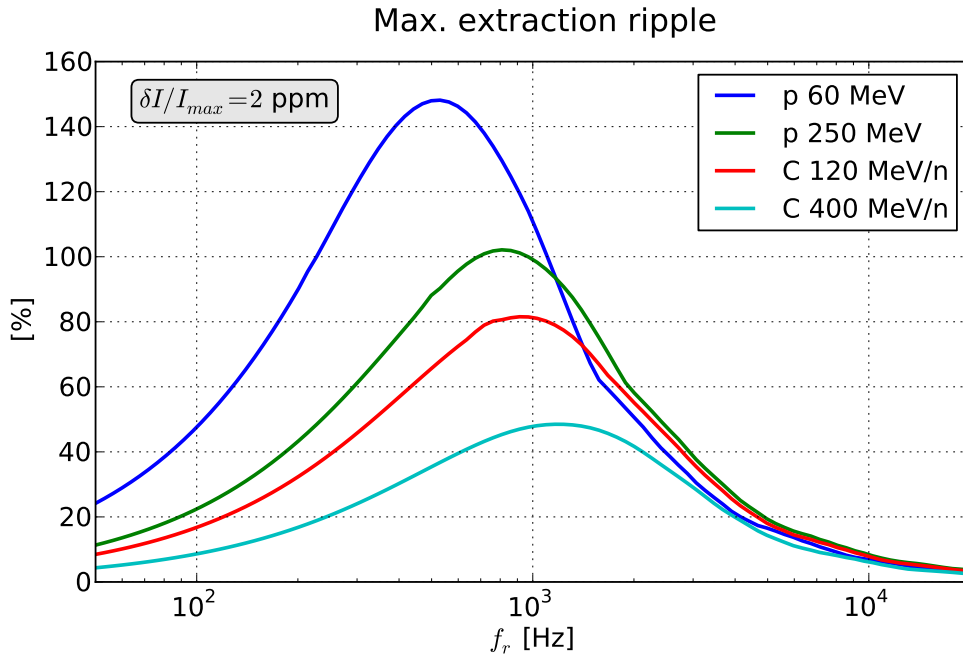Figure 4.11: Maximum relative extraction ripple amplitude in percent for dipole current ripple frequencies between 50 Hz and 20 kHz. The current ripple amplitude is 2 ppm.
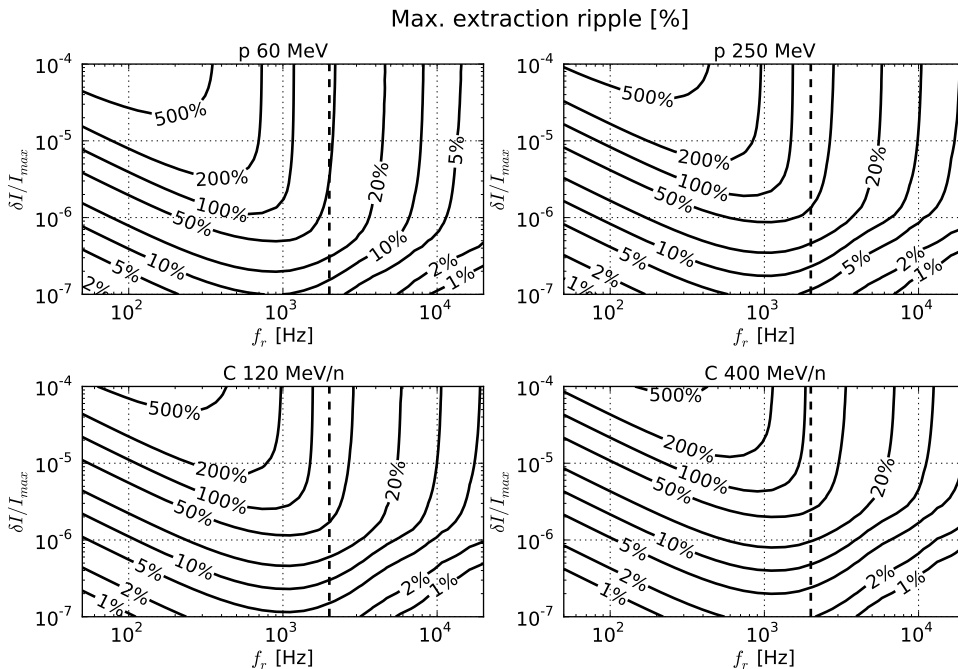

Figure 4.12: Maximum relative extraction ripple amplitude in percent for dipole current ripple amplitudes between 0.1 and 100 ppm of I_{max} and frequencies between 50 Hz and 20 kHz. Dipole f_{PWM} indicated as a dashed line.

Fig. 4.13 shows a similar evaluation for defocussing quadrupole ripple (QD). The general characteristics are similar to the dipole evaluation. However, the quadrupole PWM frequency (12 kHz) is sufficiently high that quadrupole current ripple will have a limited effect ($\lesssim$10%) on the extraction stability, compared to dipoles at 2 kHz.
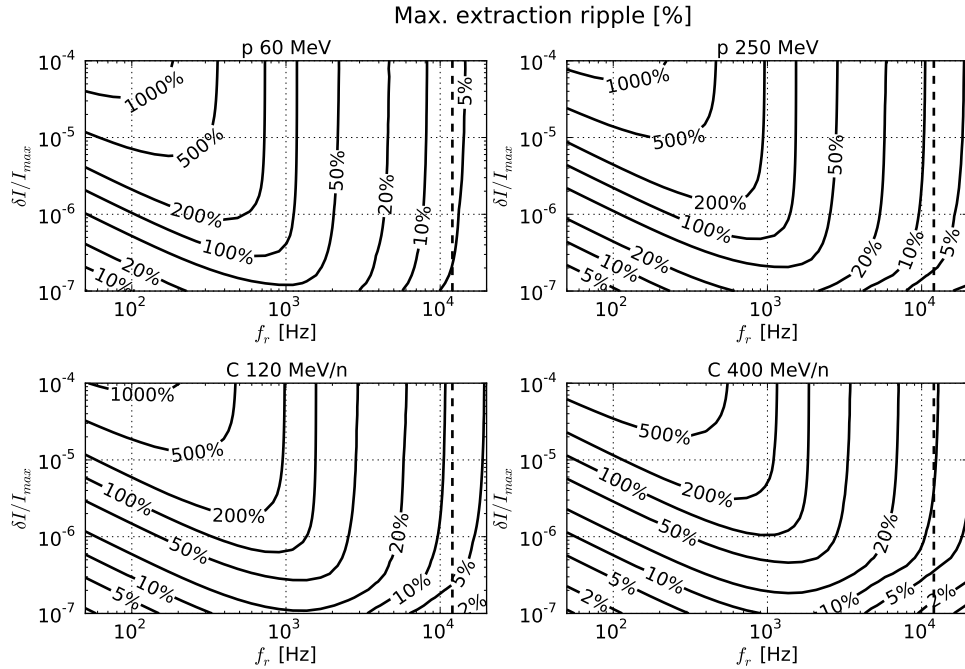

Figure 4.13: Maximum relative extraction ripple amplitude in percent for defocusing quadrupole current ripple amplitudes between 0.1 and 100 ppm of I_{max} and frequencies between 50 Hz and 20 kHz. Quadrupole f_{PWM} indicated as a dashed line.

Quadrupole current ripple at 12 kHz will have a negligible effect on the particle rate stability (see Fig. 4.13).

At a given current ripple amplitude, frequencies in the 1 kHz region will have the largest impact on extraction stability, and it is therefore concluded that the synchrotron dipole power supplies with PWM frequency of 2 kHz will be the most critical element for a stable extraction.

Using RF-channelling, the situation is significantly improved at low and medium frequencies, where the extraction ripple is small compared to the case without RF channelling (shown in Fig. 4.14). At higher frequencies, RF-channelling lacks the filtering effect from the finite width of the spill response and only Eddy current damping limits the sensitivity to very high frequency ripple.

4.3 Particle tracking

Particle tracking has been performed in order to complement and verify the analytical model of the extraction mechanism, with respect to e.g. the effect of dipole
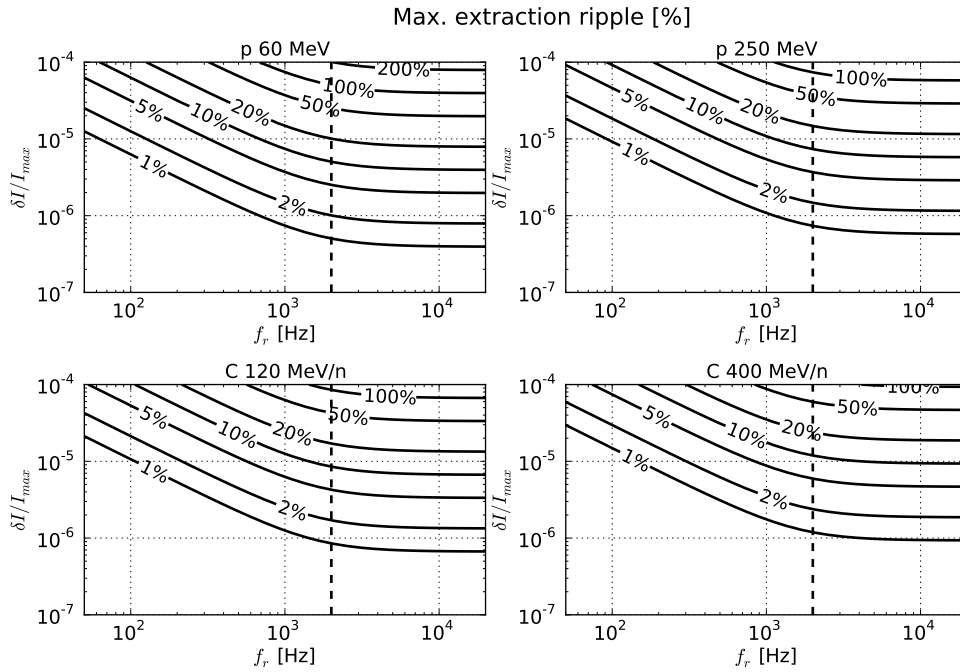
Max. extraction ripple [%]

Figure 4.14: Same as Fig. 4.12, but with RF channelling (dipole parameters).

current ripple on extraction stability. The change in phase space coordinates and momentum of a single particle over one turn in the synchrotron can be modeled by applying the effect of the resonance sextupole (Eq. 3.39), a chromatic rotation (Eq. 3.25 and Eq. (3.30)) and the acceleration by the betatron core. By doing this repeatedly for a large number of particles from a suitable initial distribution, the quality of the extracted beam can be studied.

Singe particle representation   The phase space coordinates and momentum of each particle are bundled into one vector $\mathbf{r}$, which represents the particle at the resonance sextupole:

$$\mathbf{r} = \begin{pmatrix} x \\ x' \\ y \\ y' \\ \delta p/p \end{pmatrix} \tag{4.45}$$

Sextupole kick   Using a thin lens approximation, the change in particle direction when passing through the resonance sextupole is (operator $\hat{S}$):

$$\hat{S}(t)\mathbf{r} = \begin{pmatrix} x \\ x' \\ y \\ y' \\ \delta p/p \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{1}{2}l_s k'(t) \times (x^2 - y^2) \\ 0 \\ -l_s k'(t) \times xy \\ 0 \end{pmatrix} \tag{4.46}$$

79

Note that the resonance sextupole ramps up from zero strength before extraction, which is why the normalized sextupole gradient k' is time dependent.

Betatron kick The effect of the betatron kick ($\hat{B}$) is modeled as an increment in momentum, $\delta_p(t)$:

$$\hat{B}(t)\mathbf{r} = \begin{pmatrix} x \\ x' \\ y \\ y' \\ \delta p/p + \delta_p(t) \end{pmatrix} \tag{4.47}$$

Single turn chromatic rotation Finally, after the particle has passed through the infinitesimally thin resonance sextupole and being accelerated by the betatron core, it is transported one turn round the synchrotron, $\hat{M}_{RS \to RS}(t)$, back to the resonance sextupole:

$$\hat{M}_{RS \to RS}(t)\mathbf{r} = \begin{pmatrix} m_{11,x}(t) & m_{12,x}(t) & 0 & 0 & 0 \\ m_{21,x}(t) & m_{22,x}(t) & 0 & 0 & 0 \\ 0 & 0 & m_{11,y}(t) & m_{12,y}(t) & 0 \\ 0 & 0 & m_{21,y}(t) & m_{22,y}(t) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ x' \\ y \\ y' \\ \delta p/p \end{pmatrix} \tag{4.48}$$

where (Eq. 3.25):

$$\begin{aligned} m_{11,z}(t) &= \cos(2\pi Q_z(t)) + \alpha_{RS} \sin(2\pi Q_z(t)) \\ m_{12,z}(t) &= \beta_{RS} \sin(2\pi Q_z(t)) \\ m_{21,z}(t) &= -\gamma_{RS} \sin(2\pi Q_z(t)) \\ m_{22,z}(t) &= \cos(2\pi Q_z(t)) - \alpha_{RS} \sin(2\pi Q_z(t)) \end{aligned} \tag{4.49}$$

and the time dependent tune $Q_z(t)$ given by:

$$Q_z(t) = Q_{0,z} + Q_{r,z}(t) + Q'_z \times \frac{\delta p}{p} \tag{4.50}$$

$Q_{0,z}$ is the nominal synchrotron tune (i.e. 5/3 in the horizontal plane). The term $Q_{r,z}(t)$ is introduced to allow for *tune ripple* of the synchrotron, which can be caused by e.g. current ripple in the synchrotron magnets.

At the electrostatic septum To determine whether a particle will exit the synchrotron, or continue circulating, its coordinates must be transferred from the resonance sextupole to the electrostatic septum, using the general transfer matrix from Eq. 3.22. If, at the electrostatic septum, $x > x_{ES}$, the particle will be extracted.

As mentioned, the resonance sextupole is placed in a dispersion free region, but the electrostatic septum is not. The transfer matrix will therefore contain dispersion

terms:

$$\mathbf{r}_{ES} = \begin{pmatrix} k_{11,x} & k_{12,x} & 0 & 0 & D_x \\ k_{21,x} & k_{22,x} & 0 & 0 & D'_x \\ 0 & 0 & k_{11,y} & k_{12,y} & D_y \\ 0 & 0 & k_{21,y} & k_{22,y} & D'_y \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{r}_{RS} \tag{4.51}$$

where $k_{ij,z}$ are given by inserting the horizontal/vertical Twiss functions at the resonance sextupole and electrostatic septum into Eq. 3.23. A particle will be extracted if:

$$k_{11,x}x + k_{12,x}x' + D_x\frac{\delta p}{p} > x_{ES} \tag{4.52}$$

Summary The phase space trajectory of a single particle k at the resonance sextupole is given by:

$$\mathbf{r}_k(t + T_{rev}) = \hat{M}_{RS \to RS}(t)\hat{B}(t)\hat{S}(t)\mathbf{r}_k(t) = \hat{R}(t)\mathbf{r}_k(t) \tag{4.53}$$

where T_{rev} is the time it takes for the particle to circulate the synchrotron once. Discretization gives:

$$\mathbf{r}_k^{i+1} = R^i\mathbf{r}_k^i \tag{4.54}$$

4.3.1   Initial and final distribution

The initial horizontal and vertical distribution of particles are given by inserting the resonance sextupole twiss functions into Eq. 3.17 (the momentum distribution is assumed to be uniform). An example of the initial and final distribution (after extraction) in amplitude-momentum space of 3000 particles is shown in Fig. 4.15. The slope of the initial distribution is exactly that of the theoretical limit of stability. As seen, the extracted particles are distributed along the limit of stability, except at higher amplitudes where they deviate slightly.

The (intentional) momentum gap between the initial distribution and the limit of stability ensures that no particles are extracted during ramping of the resonance sextupole, while the initial width of the momentum distribution is wide enough to produce a "flat top" of extracted particle rate[3].

4.3.2   Tune ripple

The time profile of the extracted beam during tune ripple of amplitude $\delta Q_r = 10$ ppm, $f_r = 200$ Hz from tracking is shown in Fig. 4.16[4]. After a "build up" phase of about 10 ms, where the initial particles reach the resonance and are extracted, the extraction rate oscillates regularly around the nominal level.

---

[3]The slight deviation between the limit of stability and the particle momenta after extraction causes the high-amplitude particles of the initial distribution to be extracted later than the low-amplitude particles.

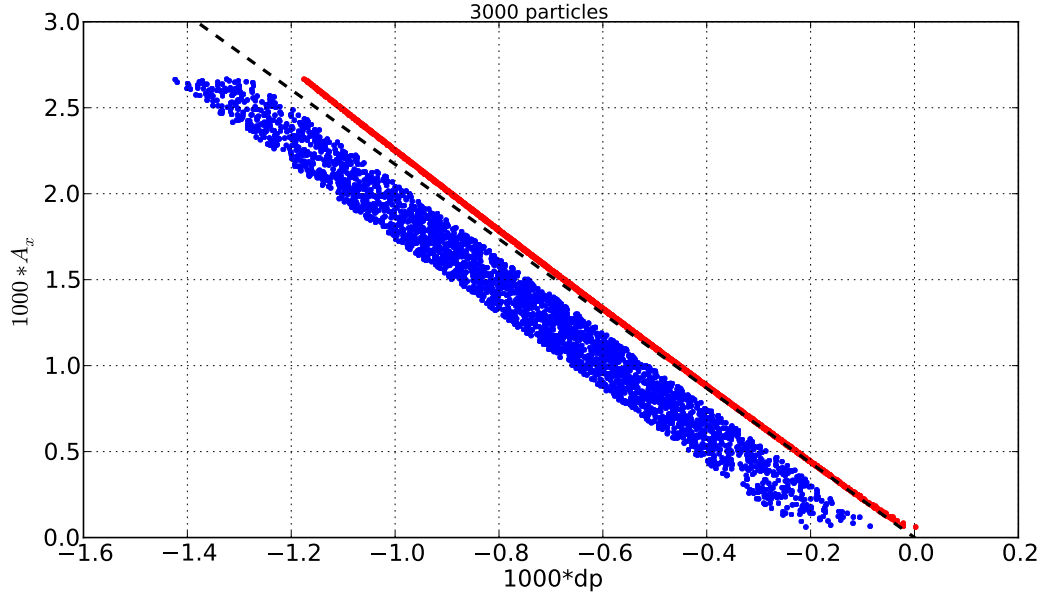[4]Extracted particle rate is a sliding average over 100 μs.

Figure 4.15: Initial (blue) and final (red) particle distribution in horizontal amplitude/momentum space. Dashed line shows the theoretical limit of stability. Particles are plotted vs. their *initial* horizontal amplitude, i.e. the increased amplitude during extraction is not shown.

From Eq. 4.36, this corresponds to a current ripple of about 1.6 ppm (dipole parameters). The peak ripple amplitude is about 80%, which is well in agreement with the low frequency approximation (79%, Eq. 4.38) and the evaluation shown in Fig. 4.12 (about 70%).

At higher frequencies, the spill model fails to account for particles that are thrown in and out of the resonance by the high amplitude tune velocity $\dot{Q}_r$. A comparison of the extraction ripple amplitude as calcualted by the spill model and tracking is shown in Tab. 4.2. At low- and medium frequencies, tracking confirms the accuracy of the spill model, but at higher frequencies, the spill model underestimates the ripple amplitude.

f_r [Hz]	δQ_r [ppm]	$\frac{\delta I_r}{I_{max}}$ (D) [ppm]	$\frac{\delta I_r}{I_{max}}$ (Q) [ppm]	Spillgen	Tracking
200	1	0.16	0.11	7%	10%
200	10	1.6	1.1	70%	80%
200	100	16	11	420%	550%
2000	1	0.26	0.14	16%	40%
2000	10	2.6	1.4	51%	130%
2000	100	26	14	57%	200%

Table 4.2: Column 1 and 2: tune ripple frequency and amplitude. Column 3 and 4 show corresponding dipole and quadrupole current ripple for a 60 MeV proton beam. Column 5 and 6 show the resulting extraction ripple ($\varepsilon_{\dot{N}}$) in %, using the analytical spill model and particle tracking.
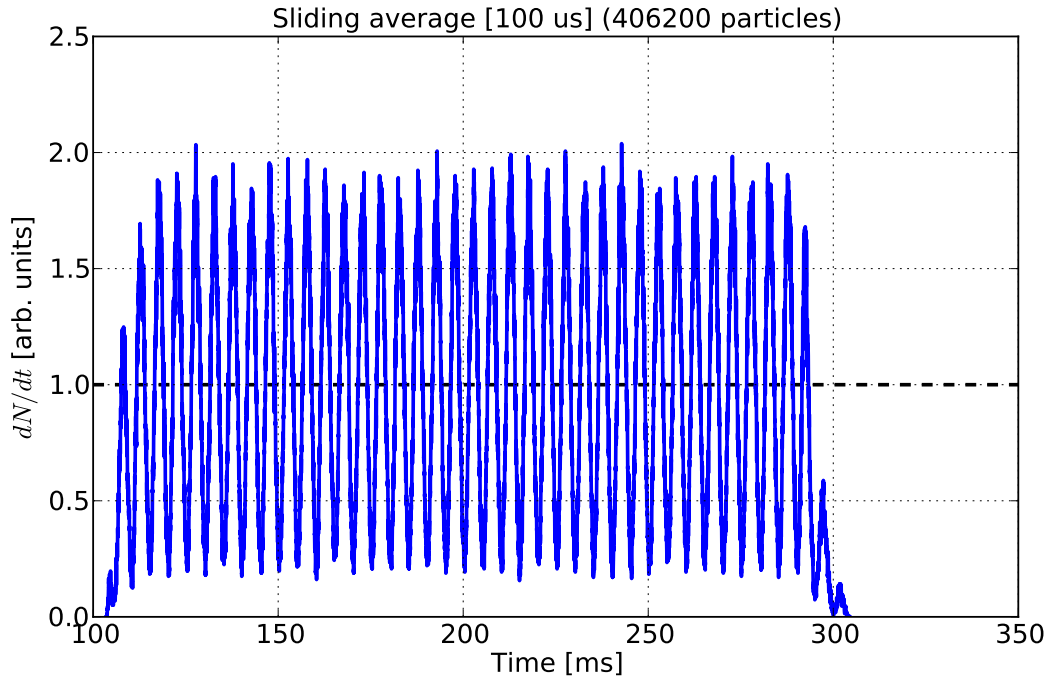
Figure 4.16: $\dot{N}(t)$, 200 Hz tune ripple of amplitude 10 ppm (60 MeV proton parameters).

At 12 kHz tune ripple, the extracted beam is still stable to within $\pm 10\%$[5] even at a tune ripple amplitude of 100 ppm.

4.3.3 Extracted beam profile

The phase space[6] of the beam extracted *after* the "build up" is shown in Fig. 4.17. The vertical beam profile is, as expected, a truncated Gaussian, while the horizontal phase space is a narrow bar of charge of about 10 mm FWHM. Fig. 4.18 also shows the extracted beam profile, but during 2 kHz tune ripple of 10 ppm. Although this causes over 100% beam rate modulation, the beam profile is unaffected. The "skew trapezoid" horizontal profile is also in agreement with the extracted beam profile at CNAO [51].

The skewness is largely caused by off-momentum particles taking shorter spiral steps over the extraction turn. In order to produce a symmetric horizontal beam profile at the patient, the correlation between skewness and momentum is planned to be corrected with a dispersion difference from extraction point to isocenter.

---

[5]With a 100 μs sliding average.
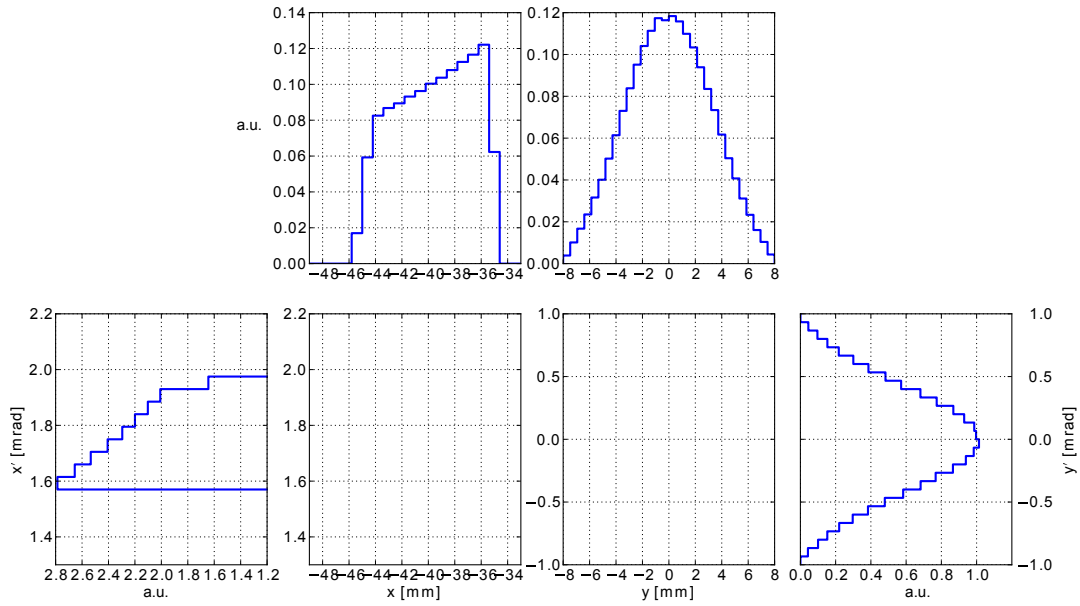[6]At the resonance sextupole.

83

Figure 4.17: Horizontal (left) and vertical (right) phase space and projected beam profiles at the resonance sextupole. No tune ripple.
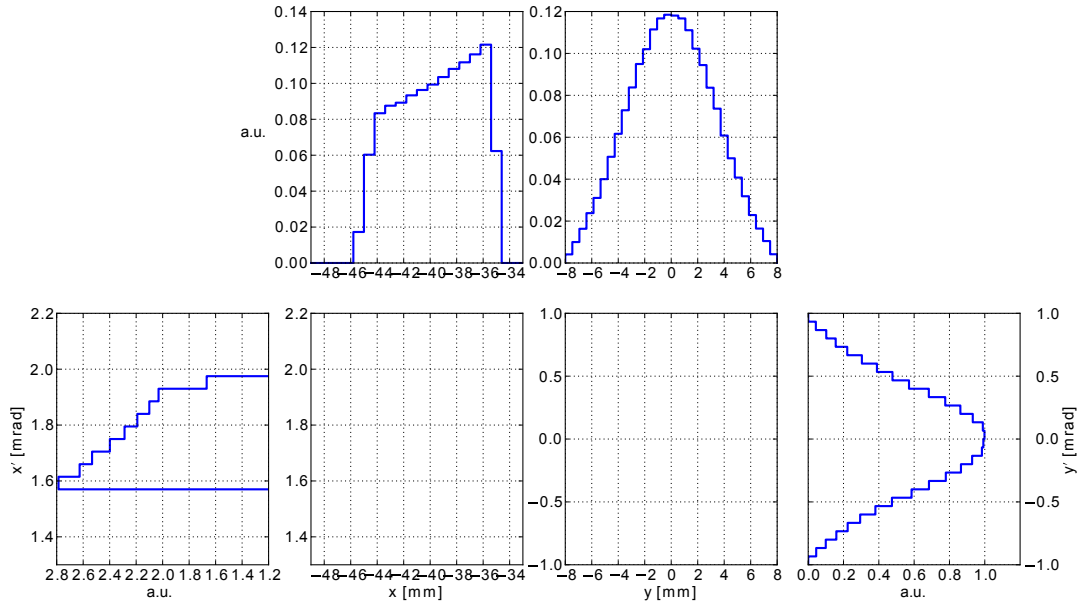
Figure 4.18: Horizontal (left) and vertical (right) phase space and projected beam profiles at the resonance sextupole. 2 kHz tune ripple, $\delta Q_r = 10$ ppm.

4.4 Results and discussion

A tune ripple spill generator has been implemented to model the extracted spill profile when the synchrotron is subject to tune ripple caused by current ripple in the main ring magnets. This allows for translating extraction stability requirements into magnet current stability requirements.

At low and medium frequencies below the kHz range, the spill generator is in agreement with particle tracking, while at higher frequencies, it underestimates the extraction ripple.

The impact of synchrotron tune ripple on extracted beam profile is confirmed by tracking to be negligible.

Chapter 5

Proton gantry and nozzles

In this chapter, the general layout of the proton gantry and the nozzles will be evaluated from a scattering point of view, in order to determine the smallest beam sizes that can be achieved at the patient. For scattering along the proton gantry, in elements that are far upstream of the patient, it is important to also take the beam optics into account, since the beam can be focused and defocussed between the point of scattering and the patient. For this reason, a general scattering model applicable to transfer lines will be presented.

A comparison of different kinds of ridge filters and their efficiency is also made.

This chapter ends with an analysis of measurement limitations of the beam position and beam size, that ought to be considered not only when choosing nozzle monitors, but also in the treatment plan.

5.1 Scattering model for a transfer line

5.1.1 Matrix notation

Let us start with rewriting the conventional drift space scattering integrals from Eq. 3.94 in matrix notation:

$$\mathbf{z}_b = M_{a \to b}^{Drift} \mathbf{z}_a + \int_a^b M_{s' \to b}^{Drift} \begin{pmatrix} \mathcal{S}(s') \\ 0 \\ 0 \end{pmatrix} \mathrm{d}s' \tag{5.1}$$

where:

$$\mathbf{z} = \begin{pmatrix} \langle z'^2 \rangle \\ \langle zz' \rangle \\ \langle z^2 \rangle \end{pmatrix} \tag{5.2}$$

and the drift space *transport matrix*, $M_{s \to b}^{Drift}$ is defined as:

$$M_{s \to b}^{Drift} = \begin{pmatrix} 1 & 0 & 0 \\ b-s & 1 & 0 \\ (b-s)^2 & 2(b-s) & 1 \end{pmatrix} \tag{5.3}$$

In this notation, the interpretation of the scattering integrals is straightforward: the initial phase space $\mathbf{z}_a$ is simply transported from a to b, while the integral with the transport matrix transports small divergence increments $\mathcal{S}(s')\mathrm{d}s'$ from s to b. At $s = b$, the contribution from the initial phase space and scattering along the transfer line are added.

The coefficients of the transport matrix are easily resolved by considering the motion of a single particle in a vacuum drift space, as illustrated in Fig. 5.1:

$$\begin{pmatrix} z_b \\ z_b' \end{pmatrix} = \begin{pmatrix} 1 & b - a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} z_a \\ z_a' \end{pmatrix} \tag{5.4}$$
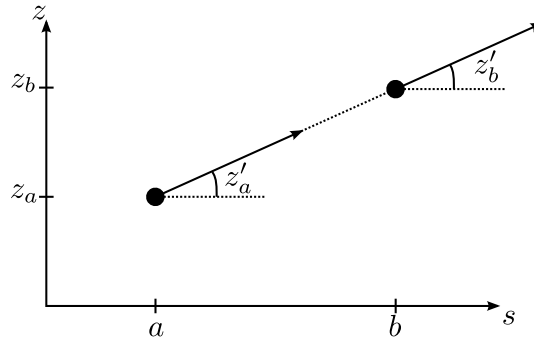

Figure 5.1: Single particle motion in vacuum (drift space).

From the single particle motion matrix, the evolution of $\mathbf{z}$ from a to b can be calculated from:

$$\begin{aligned} \langle z_b'^2 \rangle &= 1 \times \langle z_a'^2 \rangle \\ \langle z_b z_b' \rangle &= \langle (z_a + (b-a)z_a') \times z_a' \rangle = (b-a) \times \langle z_a'^2 \rangle + 1 \times \langle z_a z_a' \rangle \\ \langle z_b^2 \rangle &= \langle (z_a + (b-a)z_a')^2 \rangle = \\ &= (b-a)^2 \times \langle z_a'^2 \rangle + 2(b-a) \times \langle z_a z_a' \rangle + 1 \times \langle z_a^2 \rangle \end{aligned} \tag{5.5}$$

which gives exactly the coefficients in the elements of $M_{a \to b}^{Drift}$.

5.1.2 Generalization to transfer line

In order to generalize the scattering model to a transfer line with known optics, we modify the transport matrix in accordance with the single particle motion given by the Twiss functions. Repeating the procedure in Eq. 5.5, but with the general single particle transfer matrix from Eq. 3.22 gives the elements of the general transport matrix $M_{a \to b}$:

$$M_{a \to b} = \begin{pmatrix} m_{22}^2 & 2m_{21}m_{22} & m_{21}^2 \\ m_{12}m_{22} & m_{11}m_{22} + m_{12}m_{21} & m_{11}m_{21} \\ m_{12}^2 & 2m_{11}m_{12} & m_{11}^2 \end{pmatrix} \tag{5.6}$$

where m_{ij} is given by inserting the Twiss functions at $s = a$ and $s = b$ into Eq. 3.23. The phase space of the beam at any location b along the transfer line can thus be written as:

$$\mathbf{z}_b = \underbrace{M_{a \to b} \mathbf{z}}_{\text{Unscattered beam}} + \underbrace{\int_a^b M_{s' \to b} \begin{pmatrix} S(s') \\ 0 \\ 0 \end{pmatrix} \mathrm{d}s'}_{\text{Contribution from scattering}} \quad (5.7)$$

or even shorter:

$$\mathbf{z}_b = \mathbf{z}_0 + \mathbf{z}_s \quad (5.8)$$

where $\mathbf{z}_0$ is the *unscattered* phase space of the beam at b and $\mathbf{z}_s$ the scattering contribution. The implication of this is that one does not actually need to know the phase space of the initial beam to evaluate the phase space further downstream. If b is chosen as the isocenter in one of the irradiation rooms, it suffices to calculate the scattering integral and add the unscattered beam width quadratically to evaluate the scattered beam size. Writing out the scattering equation for the beam size only gives:

$$\langle z^2 \rangle_b = \langle z^2 \rangle_0 + \underbrace{\int_a^b \beta_z(s') \beta_z(b) \sin^2(\mu_z(b) - \mu_z(s')) \times S(s') \mathrm{d}s'}_{= \sigma_s^2} \quad (5.9)$$

The (square root of the) integral is denoted the *scattering term*, σ_s and the 1-σ beam size at b is thus given by a quadratic addition of the unscattered 1-σ beam size and the scattering term.

5.1.3 Remarks

5.1.3.1 Twiss functions during scattering

The Twiss functions along the transfer line depend on the initial Twiss functions which are matched to the phase space of the incoming beam. The ideal Twiss functions calculated in an optics program for a vacuum transfer line will therefore be different from the "real" Twiss functions of a scattered beam. However, the ideal Twiss functions used as input for Eq. 5.7 are only used to calculate the single particle transfer matrices[1]. If the energy loss due to scattering is negligible (e.g. scattering in a vacuum window or a short air gap), the single transfer matrices calculated from the ideal Twiss functions from one point to another are not affected.

If the energy loss due to scattering is considerable, the downstream optics must in any case be adjusted (e.g. by reducing the focusing strength) such that the ideal single particle transfer matrices are recovered.

---

[1]The Twiss functions along the transfer line depend on the phase space of the initial beam. However, the single particle transfer matrices calculated from the Twiss functions are independent of the phace space of the initial beam.

5.1.3.2   Scattering of non-Gaussian beam profile

The scattered non-Gaussian horizontal beam profile at the isocenter is resolved by convoluting the pristine trapezoidal beam profile, $\rho_x(x)$, with a Gaussian profile with a standard deviation corresponding to the scattering term[2]:

$$\tilde{\rho}(x,s) = \rho_x(x) * G(x, \sigma_s(s)) \tag{5.10}$$

5.2   Proton gantry evaluation: scattering

5.2.1   Gantry optics

The Twiss functions along the gantry are shown in Fig. 5.2. The optics are optimized such that there is a 1:1 image of the vertical phase space of the incoming beam to the isocenter. This is not the case in the horizontal plane.
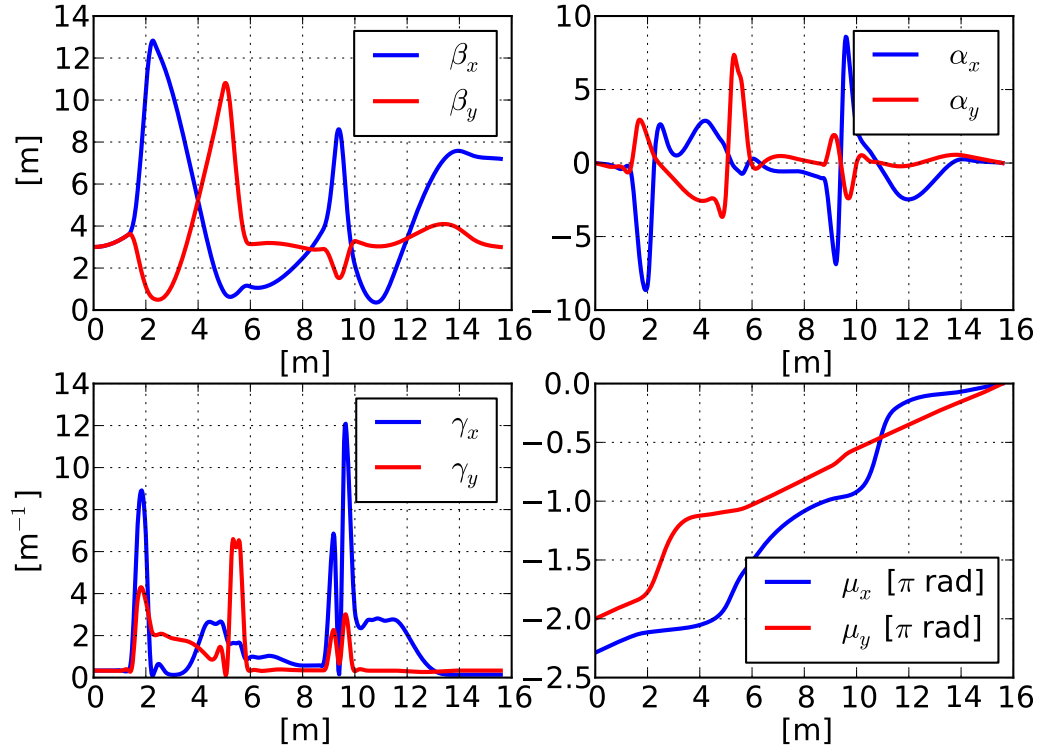
Figure 5.2: Horizontal and vertical Twiss functions along proton gantry (coupling point at 0 m, isocenter at 15.58 m). The phase advance is shown relative to the isocenter.

The phase space of the incoming beam is taken to be that which gives the smallest beam size at the isocenter (4 mm FWHM), for which the incoming phase space

---

[2]This is implied in the scattering model, since a convolution between two Gaussians results in a new Gaussian with quadratic addition of the initial widths.

ellipses are:

$$\mathbf{x}_{CP} = \frac{\varepsilon_x}{\pi} \times \begin{pmatrix} 1/3.0 \\ 0 \\ 3.0 \end{pmatrix} ; \quad \mathbf{y}_{CP} = \frac{\varepsilon_x}{\pi} \times \begin{pmatrix} 1/2.0 \\ 0 \\ 2.0 \end{pmatrix} \qquad (5.11)$$

with dimensions of (rad^2, m-rad, m^2). The horizontal and vertical emittance, ε_x and ε_y are geometric emittances at 60 MeV, given in Appendix A.

Note that the bar of charge in the horizontal phase space (Fig. 4.17) is not well represented by an ellipse. The quoted horizontal phase space ellipse refers to the "unfilled ellipse", which circumscribes the bar of charge.

At the isocenter (IC), the unscattered phase space ellipses are:

$$\mathbf{x}_{IC} = \frac{\varepsilon_x}{\pi} \times \begin{pmatrix} 0.139 \\ 0 \\ 7.12 \end{pmatrix} ; \quad \mathbf{y}_{IC} = \frac{\varepsilon_x}{\pi} \times \begin{pmatrix} 1/2.0 \\ 0 \\ 2.0 \end{pmatrix} \qquad (5.12)$$

5.2.2 Scattering geometry

From a scattering point of view, there are four different combinations to compare: with/without vacuum windows at the coupling point (CP) and with/without helium in the dipole.

In the case of no helium in the dipole, the last vacuum window should be placed at least 92 cm upstream of the isocenter, to give sufficient space for the treatment nozzle (Sec. 5.4). Likewise, when the dipole is filled with helium, the helium container must end at least 92 cm before the isocenter (the last vacuum window being placed just before the dipole).

Tab. 5.1 summarizes the gantry dimensions and the materials used in the scattering evaluation.

The last vacuum window must be thicker than the two vacuum windows at the coupling point. This is required since the vacuum pipe aperture close to the patient need to be almost as large as the scanning region. The pressure difference and the larger vacuum pipe cross section requires about 180 μm of mylar, in order not to risk breaking.

At the proton therapy facility M. D. Anderson, a helium chamber similar to what is considered here is installed upstream of the nozzle [52]. The reason is that a beam profile monitor is installed just before the scanning magnets. The monitor enables continuous monitoring of the beam profile at a stable location, but breaks the vacuum. The beam profile monitor is directly followed by a helium chamber at atmospheric pressure, which reduces scattering of the beam compared to air. The helium chamber windows at M.D. Anderson are made of kapton, coated on both sides with copper and aluminum[3]. Identical windows have been assumed in the scattering evaluation of the MedAustron proton gantry. These windows are substantially thinner than typical vacuum chamber windows, which favours using a helium chamber rather than a second vacuum chamber. An additional disadvantage with a vacuum

---

[3]Private conversation with M.D. Anderson employee.

chamber is that in case the window breaks, it can cause a loud noise which may frighten the patient.

Dimensions	
Total gantry length (beam path)	15.58 m
Start of dipole to isocenter	4.02 m
Dipole length (beam path)	2.36 m
Minimum space required for nozzle	0.92 m
Material	
Vacuum windows at CP	$2\times 50\,\mu$m kapton
Last vacuum window	$180\,\mu$m mylar
Helium chamber windows	$\begin{cases} 12.5\,\mu\text{m kapton} \\ 2\times 0.2\,\mu\text{m copper} \\ 2\times 0.1\,\mu\text{m aluminum} \end{cases}$

Table 5.1: Proton gantry dimensions and material used in the scattering evaluation.

Radiation lengths and nominal densities of all material can be found in Tab B.1.

5.2.3   Beam FWHM along proton gantry

Shown in Fig. 5.3 is the horizontal and vertical FWHM along the gantry, from the coupling point to the isocenter for the following geometries:

1. Only vacuum

2. Windowless gantry joint, single vacuum window after dipole

3. Vacuum windows at coupling point and after dipole

4. Windowless gantry joint, helium in dipole

5. Vacuum windows at coupling point, helium in dipole

The incoming beam is a 60 MeV proton beam, which is the worst case scenario: higher energies and carbon ions are less sensitive to scattering.

The beam size along the gantry has been evaluated via Eq. 5.8, approximating the horizontal bar of charge with a Gaussian beam profile. This is not strictly correct: one should rather do a convolution between a projection of the bar of charge on the x-axis and a Gaussian profile with 1-σ width equal to the scattering term, as mentioned in Sec. 5.1.3.2. However, if the FWHM of the scattering contribution is approximately equal to, or larger than, the unscattered beam profile, the difference is negligible.

(a) Five different geometry scenarios.
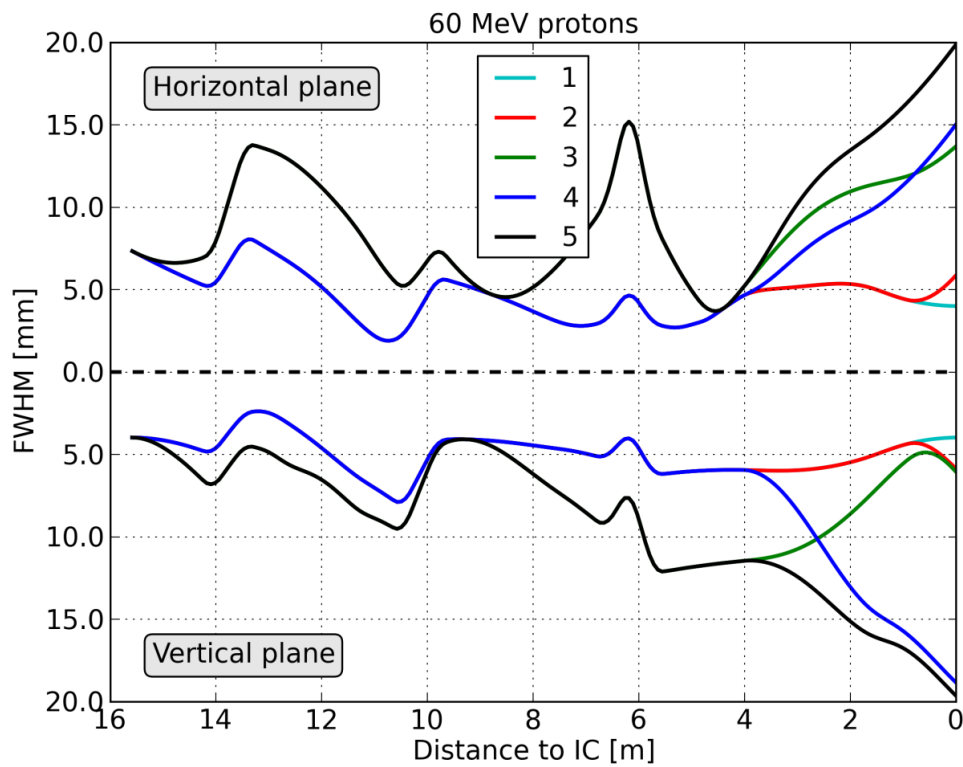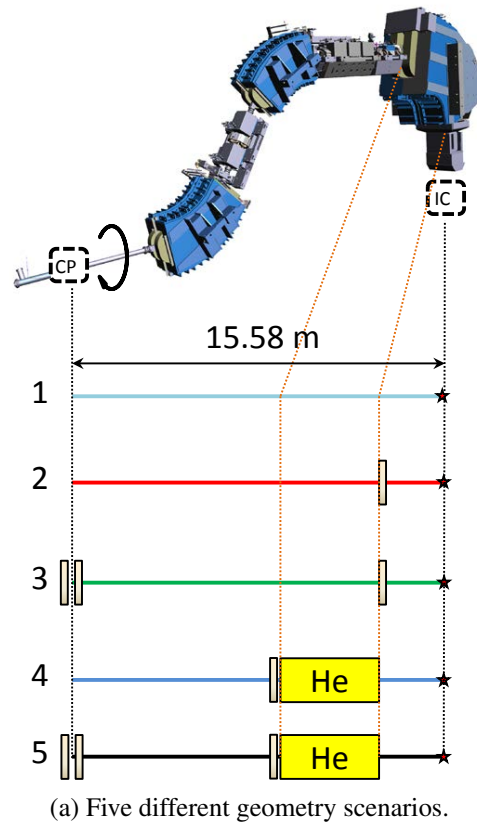
(b) Horizontal (upper half) and vertical (lower half) FWHM of a 60 MeV proton beam along the gantry

Figure 5.3: Scattering along the gantry.

5.2.4 Gantry design choices

With vacuum all the way to the isocenter (Geometry 1), the beam would be focused at the isocenter in both the horizontal and vertical plane simultaneously, with a FWHM of 4 mm. Insertion of a single vacuum window before the nozzle (Geometry 2) increases the isocenter beam size to some 6 mm in both planes, and shifts the focal point (local beam size minimum) from the isocenter towards the vacuum window. Smaller beam sizes than this can never be achieved for low energy proton beams in practice, since the vacuum window is absolutely necessary.

Inserting vacuum windows at the coupling point (Geometry 3) dramatically increases the horizontal FWHM, while the vertical FHWM is barely affected. This is due to the different optics used in the horizontal and vertical plane. Since there is a 1:1 image in the vertical plane from the coupling point to the isocenter, the vertical isocenter beam size is unaffected by scattering at the coupling point (within reasonable limits). This can be seen directly in Eq. 5.9: the phase advance is from coupling point to isocenter is 2π, so the vertical sin-term will be zero, while the horizontal beam size increases to about 13 mm.

It should be stressed here that the gantry optics of PSI and MedAustron are different: at PSI, there is a 1:1 image in *both* planes. This allows for breaking the vacuum at the coupling point, without affecting the beam size at the isocenter.

Omitting the coupling point vacuum windows, but filling the last dipole with helium (Geometry 4) has a similar effect in both planes: from 4 mm FWHM to 15 mm (x) and 18 mm (y).

Using both coupling point vacuum windows and dipole helium (Geometry 5) results in low energy beam sizes of about 20 mm in both planes.

To conclude: in order to produce beam profiles at the proton gantry isocenter that are both small and symmetric, vacuum should be kept throughout the gantry, all the way to the nozzle.

5.3 Ridge filter comparison

Three types of ridge filters have been investigated. In principle, they only differ in characteristic groove profile (and resulting range reduction distribution):

- Triangular ridge filters: used at CNAO, Italy [53]

- Optimized ridge filter: used at GSI, Germany [46]

- Gaussian ridge filter: used in Japan [54]

5.3.1 Single triangular ridge filter

A triangular ridge filter has straight ridge sides, as shown in Fig. 3.27. The straight ridges produces a uniform range reduction distribution between t_{min} and t_{max}.

The main advantage with a triangular ridge filter is that it is relatively easy to manufacture, compared to ridge filters with curved grooves described in the coming sections (Gaussian and optimized).

5.3.2 Double triangular ridge filter

The modulated Bragg curves produced by a single triangular ridge filter are not ideal for energy-stacking [46]. At CNAO, this problem is resolved by using double parallel triangular ridge filters, which has the equivalent effect of filtering the pristine Bragg curve with the characteristic range reduction distribution twice. However, in order to achieve a "double-filtering" effect from the two ridge filters, they must be separated by some distance. If the ridge filters placed close together, there will be a strong correlation between a particle's range reduction in the first ridge filter and in the second ridge filter. Placed peak-to-peak, the two ridge filters are equivalent to a single triangular filter with twice the peak height, as illustrated in Fig. 5.4. Placed peak-to-valley, they are equivalent to a range shifter.
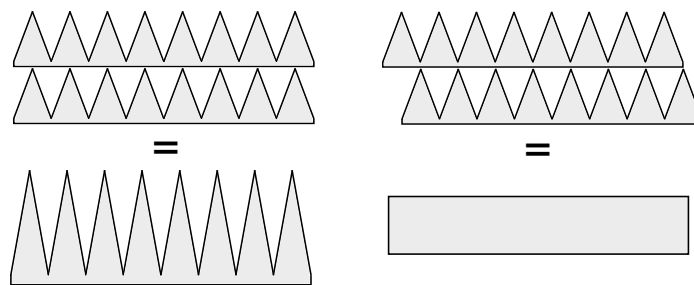
Figure 5.4: Effect of double ridge filters with parallel ridges, stacked close together.

By separating the ridge filters, one takes advantage of the divergence of the beam. Some distance after the first ridge filter, the divergence of beam will wash out the correlation between range reduction in the first ridge filter and position at the second ridge filter (schematically illustrated in Fig. 5.5).

Figure 5.5: Double parallel ridge filters, separated by some distance.

The double-filtering produces more "rounded" Bragg peaks. The range reduction distribution and equivalent single-filter groove profile of double triangular filters

are shown in Fig. 5.6. Fig. 5.7 shows an example of a carbon ion Bragg curve being filtered by one and two triangular ridge filters.

Note that there is a minimum distance required to achieve the desired "double filtering" effect.

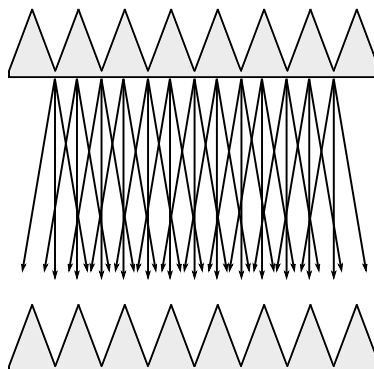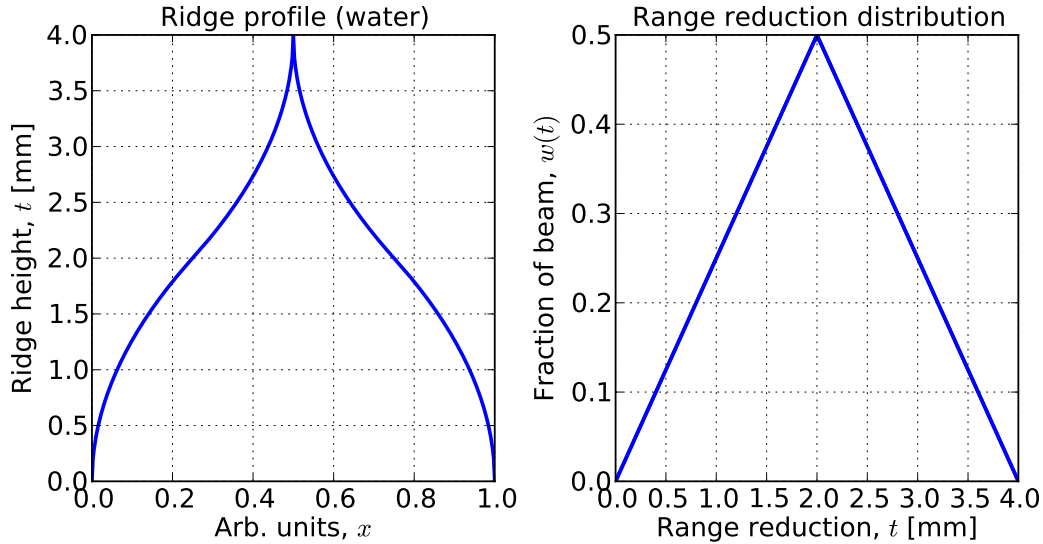Figure 5.6: Right: range reduction distribution from double triangular ridge filters, 2 mm peak height. Left: single-ridge filter profile required to generate the equivalent range reduction distribution.

Rather than aligning the grooves of the two ridge filter, they can instead be oriented orthogonally, as shown schematically in Fig. 5.8. This option has the advantage that the "double filtering" effect can be achieved without any drift space between the ridge filters. This configuration is favorable from a scattering point of view, as will be elaborated in Sec. 5.4.5.

5.3.3 Gaussian ridge filter

Another ridge profile is described in [54]. The shape is designed to create a Gaussian range reduction distribution, truncated at ± 2-σ. Ridge profile and range reduction distribution for $\sigma = 1$ mm are shown in Fig. 5.9, while a filtered Bragg curve is shown in Fig. 5.10. The filtered Bragg curve is similar to the one produced by double ridge filters.

The Gaussian ridge filter as described in [54] is intended for broad proton beams with a Gaussian transverse profile (60 mm FWHM). The peak height of the ridges are up to 7.2 mm, with a 10 mm peak-to-peak distance. No motivation of the choice of ridge profile is given, but the Gaussian range reduction distribution mimics the natural Bragg peak widening that occurs in the target, which is also normal-distributed [46].
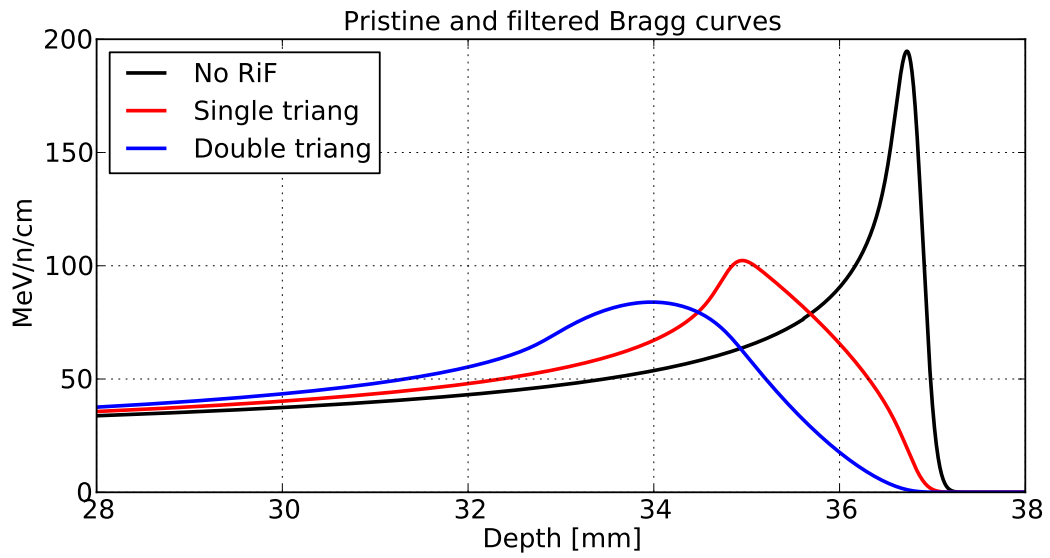
Figure 5.7: Pristine 120 MeV/n carbon ion Bragg curve ("No RiF") filtered once and twice with a triangular ridge filter with 2 mm peak height.

(a) Parallel ridge filters.   (b) Orthogonal ridge filters.
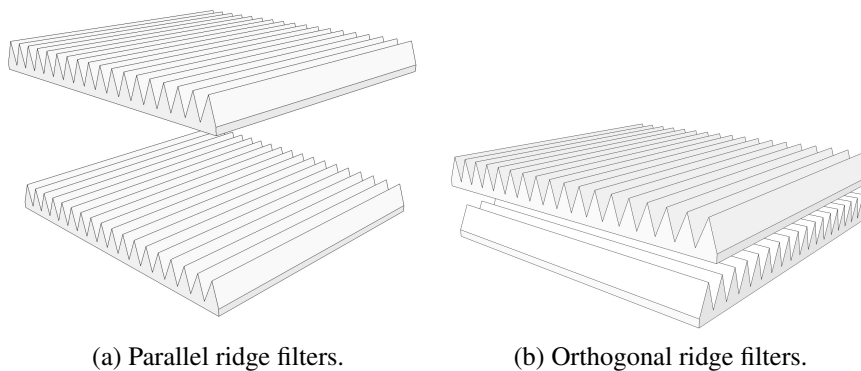
Figure 5.8: Two ridge filter options.
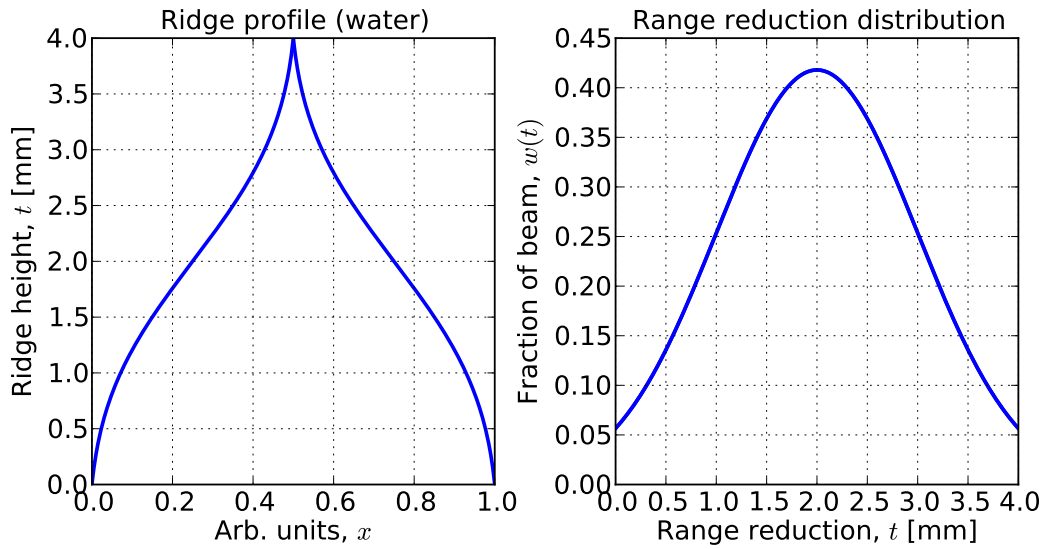
Figure 5.9: Ridge profile and range reduction distribution of a Gaussian ridge filter, truncated at ±2-σ.
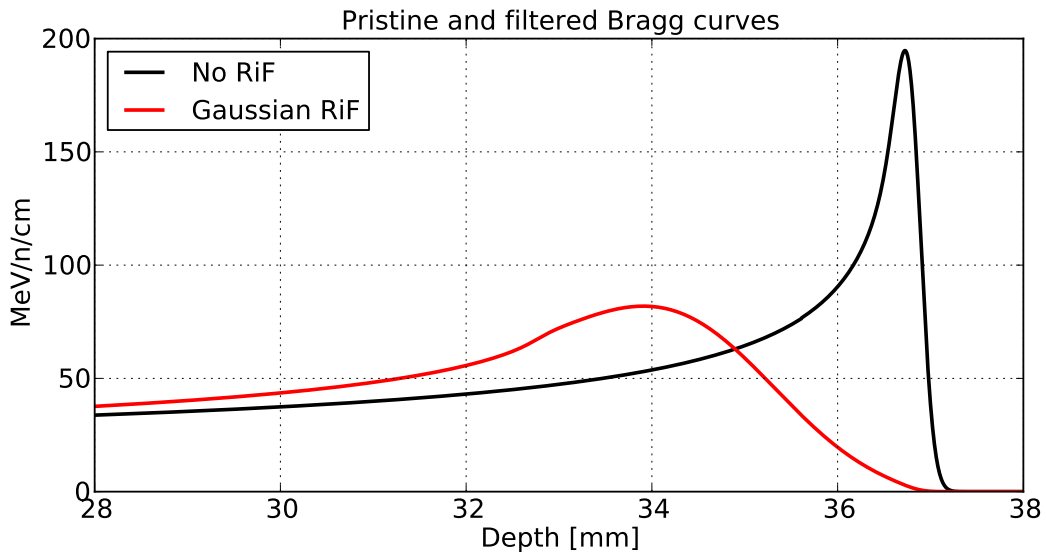
Figure 5.10: Pristine and filtered 120 MeV/n carbon ion Bragg curve, Gaussian ridge filter (σ = 1 mm, truncated at ±2-σ).

5.3.4 Optimized ridge filter

Weber et.al [46] propose a ridge filter shape that is optimized to transform the pristine Bragg peak of a 90 MeV/n carbon ion beam into a Gaussian peak with a 1-σ width of 1 mm. Apart from the shape of the non-filtered Bragg curve, the exact ridge profile will depend on a number of input parameters to the optimization algorithm.

The desired shape of the Bragg peak is a Gaussian centered at depth $s = s_0$, with a standard deviation σ_0: $G(s - s_0, \sigma)$. With the modulated Bragg curve as defined in Eq. 3.124, the aim is to optimize $w(t)$ such that the modulated Bragg peak mimics a Gaussian in some fitting region $[s_a, s_b]$:

$$\min_{w(t)} \left\{ \frac{1}{s_b - s_a} \int_{s_a}^{s_b} \left(C \times G(s - s_0, \sigma_0) - \int_{t_{min}}^{t_{max}} L_0(s + t) w(t) \mathrm{d}t \right)^2 \mathrm{d}s \right\} \quad (5.13)$$

The constant C scales the normalized Gaussian function to the height of the filtered Bragg curve. Weber et.al choose an asymmetric fitting interval around the peak of the filtered Bragg curve:

$$s_a = s_0 - 0.4 \times \sigma_0 \quad (5.14)$$
$$s_b = s_0 + 1.6 \times \sigma_0 \quad (5.15)$$
$$(5.16)$$

For the optimization procedure, the range reduction distribution $w(t)$ has in this work been modeled as a polynomial of degree $(N - 1)$ with N fit parameters a_i:

$$w(t) = \sum_{i=0}^{N-1} a_i t^i \quad (5.17)$$

and normalized such that:

$$\int_{t_{min}}^{t_{max}} w(t) \mathrm{d}t = 1 \quad (5.18)$$

Note that this method is different from the one used in [46], where $w(t)$ is discretized in steps of 10 μm and then least-square optimized.

Fig. 5.11 shows the pristine and optimized Bragg curve, using $\sigma_0 = 1$ and 3 mm and $N = 5$. The ridge profiles (Fig. 5.12) are essentially identical, which means that a ridge filter optimized to widen the Bragg peak to a Gaussian peak, 1 mm σ, can simply be scaled on the height by a factor 3 to produce 3 mm σ peaks.

The motivation for aiming at a Gaussian-shaped Bragg peak is that the combined dose of Gaussian profiles regularly juxtaposed by a distance λ is homogeneous (within 0.1%) for all values of λ in the interval $]0, 1.6\sigma]$. However, this is only the case for pure Gaussian profiles. Weber et.al do not address the fact that only the *peak* of the filtered Bragg curve is Gaussian: at the distal side of the fit region, the filtered Bragg curve is steeper than a Gaussian, while at the proximal edge it is flatter.

With the ridge filter profile described in [46], up to 2 mm layers can be used, guaranteeing a dose ripple below 4% along the SOBP.
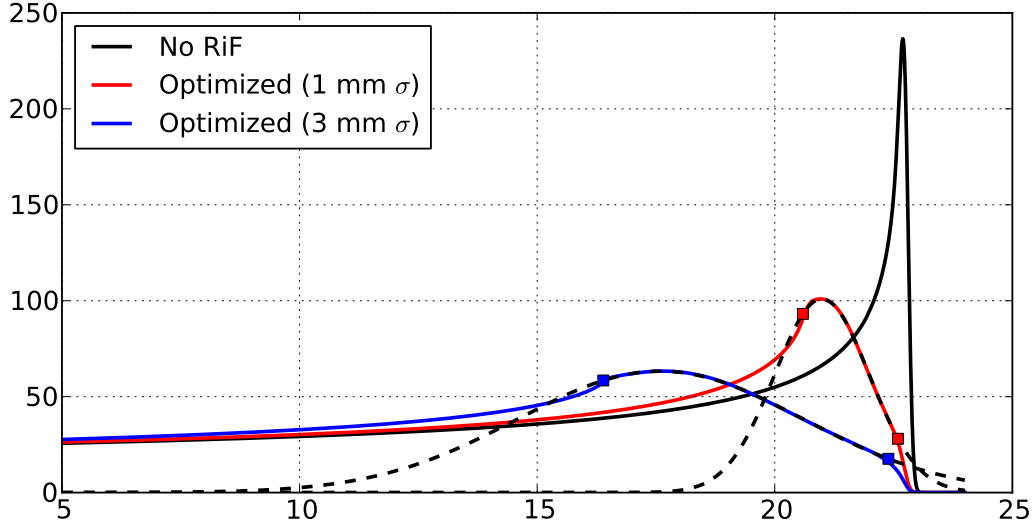
Figure 5.11: Pristine and filtered Bragg curve, 90 MeV/n carbon ions, optimized ridge filter ($\sigma_0 =$1 and 3 mm). Fit region, s_a and s_b indicated with squares. Dashed line is Gaussian target function.

5.3.5 Evaluation

5.3.5.1 Average thickness

The characteristics of each type of ridge filter depends primarily on its average thickness: the thicker it is, the more filtered the Bragg curve will be, which enables for fewer target layers, which in turn means a shorter irradiation time. On the other hand, a thick ridge filter comes at the cost of scattering, resulting in larger beam sizes. Thus, when comparing different ridge filters, it is required that they have the same average thickness $\langle t \rangle$:

$$\langle t \rangle = \int_{t_{min}}^{t_{max}} t \times w(t) \mathrm{d}t \tag{5.19}$$

For evaluation of different ridge filters, an average water equivalent ridge filter thickness of $\langle t \rangle$=1 mm has been used, which is the same as that used in [46], allowing for a layer thickness of about 2 mm.

Note that a double triangular ridge filter with $\langle t \rangle$=1 mm is equivalent to two triangular ridge filters with $\langle t \rangle$=0.5 mm.

5.3.5.2 Distal falloff

The *distal falloff* is defined as the longitudinal distance between the distal 80% and 20% levels of the Bragg peak. The distal falloff for the different types of ridge filters is shown in Fig. 5.13 for protons and carbon ions over the available energy range. With no ridge filter, the distal falloff exhibits a linear increase with beam range.

It is clear that a ridge filter for protons would only be useful at the lowest energies: at higher energies, the Bragg peak widening in the patient will be dominant,

(a) 1 mm σ

(b) 3 mm σ
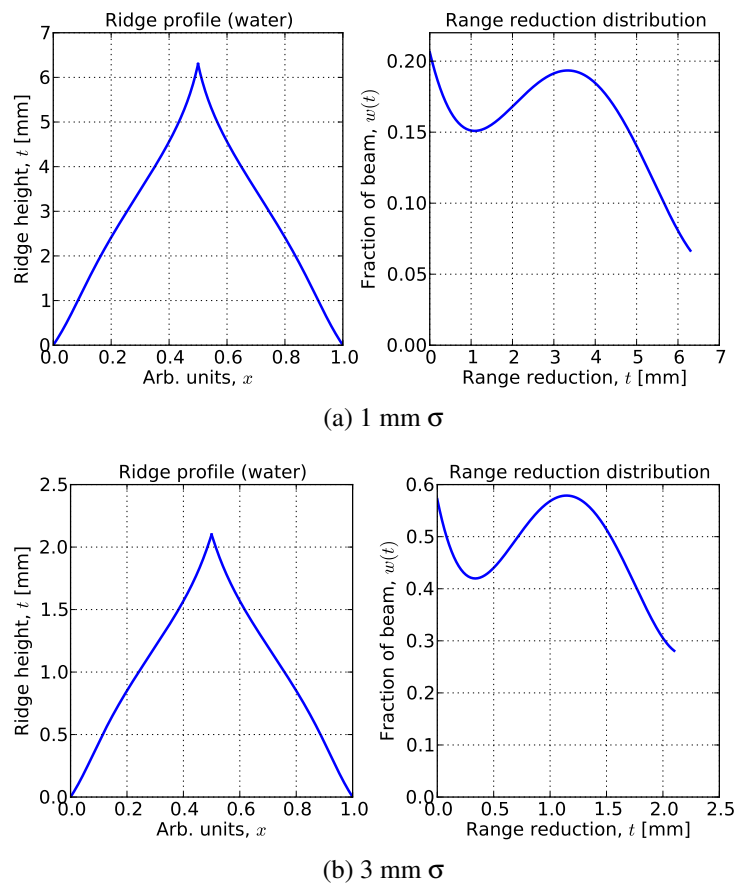
Figure 5.12: Ridge profile and range reduction distribution of an optimized ridge filter, 1 and 3 mm σ. The ridge profiles are practically identical, only scaled on the height.

and the distal falloff is almost the same whether or not a ridge filter is used.

For carbon ions, there is a significant widening of the Bragg peak over the entire therapeutic energy range, resulting in a longer distal falloff.
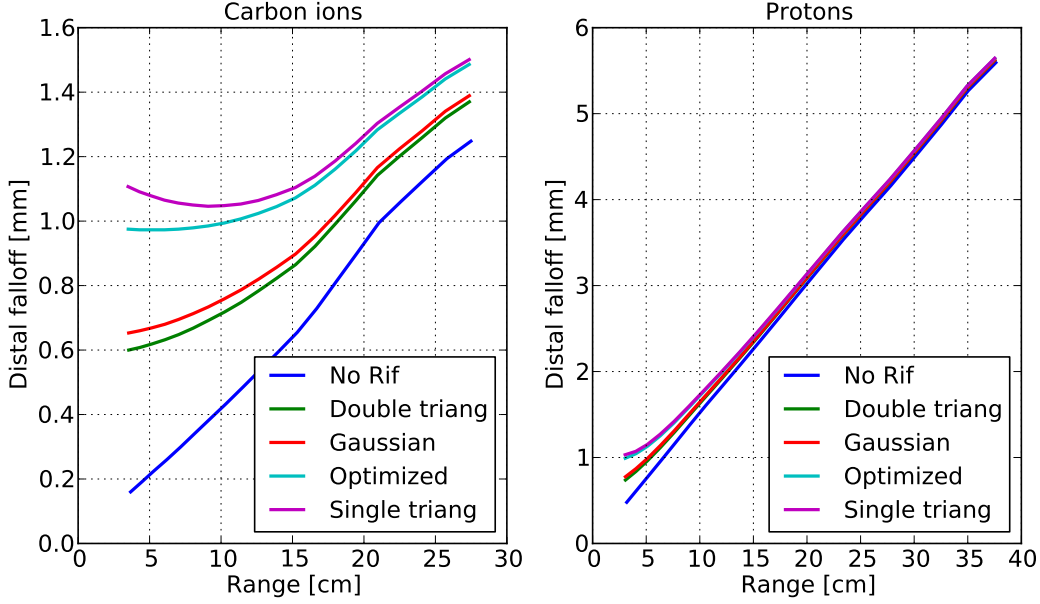
Figure 5.13: Distal falloff (80-20%) vs. beam range in water. $\langle t \rangle$=1 mm.

5.3.5.3 SOBP homogeneity

All ridge filters   The SOBP for a given ridge filter is constructed by weighting the individual layers such that the dose along the SOBP is as homogeneous "as possible". "As possible" is here defined as minimizing the (normalized) integral S^2 of the relative dose error between the SOBP, $D(s)$, and the nominal dose D_0:

$$S^2 = \min_{\lambda_i} \left\{ \frac{1}{s_{max} - s_{min}} \int_{s_{min}}^{s_{max}} \left(\frac{D(s)}{D_0} - 1 \right)^2 \mathrm{d}s \right\} \tag{5.20}$$

$D(s)$ is defined as in Eq. 3.72. s_{min} and s_{max} are set to the depth of the Bragg peak of the lowest and highest energy.

Apart from the dose error variance S^2, the maximum relative dose error, ε_{max}, along the SOBP is also of interest:

$$\varepsilon_{max} = \max \left\{ \left| \frac{D(s)}{D_0} - 1 \right| \right\} ; \quad s \in [s_{min}, s_{max}] \tag{5.21}$$

$\sqrt{S^2}$ and ε_{max} for carbon ions are plotted vs. layer thickness in Fig. 5.14, assuming an average ridge filter thickness of 1 mm[4], and a target from s_{min}=4 cm to s_{max}=9 cm. Some conclusions that can be drawn from the figure are:

---

[4]In the case of double triangular ridge filter, this means that $\langle t \rangle = 0.5$ mm for the individual ridge filters.

- All ridge filters improve the quality, in terms of $\sqrt{S^2}$, of the SOBP by about an order of magnitude, compared to the case without any ridge filter.

- The maximum local dose deviation (ε_{max}) is fairly insensitive to the layer thickness for the optimized ridge filter up to about 2 mm layers.

- With 2 mm layers, a single triangular ridge filter is actually the best choice. This can be explained by the fact that its peak height (2 mm) then coincides with the Bragg peak spacing: the triangular ridge filter effectively filters the SOBP with a rectangular window. When the window width is exactly equal to the Bragg peak spacing, the dose peaks are highly suppressed.

- The Gaussian ridge filter is the better choice for suppressing S^2 at layer thicknesses up to about $1.5 \times \langle t \rangle$.

- The quality of the SOBP from a double triangular ridge filter is comparable to (or only slightly "worse" than) that from a Gaussian ridge filter.
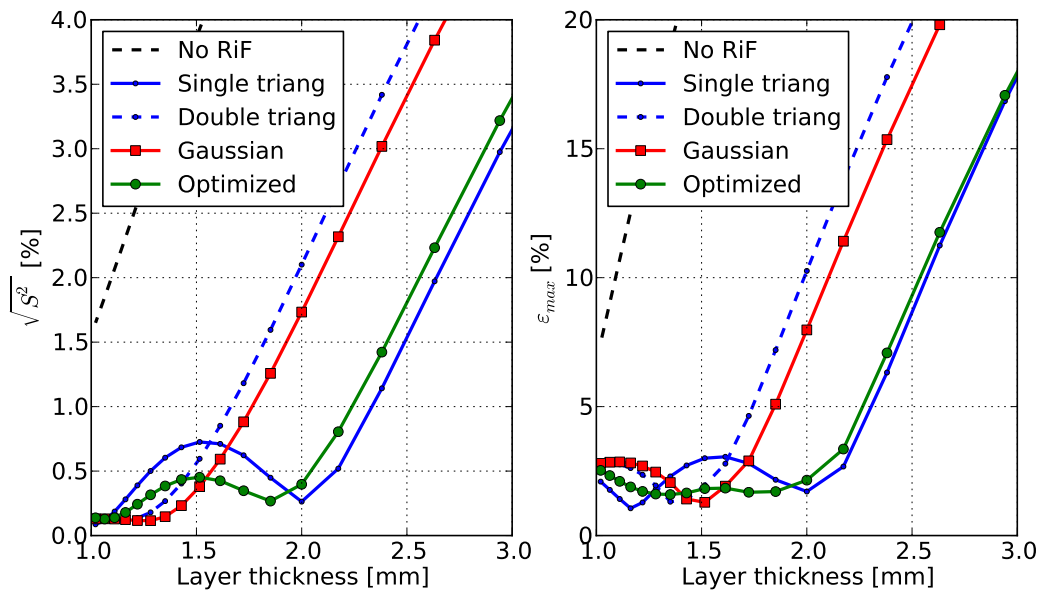
Figure 5.14: $\sqrt{S}$ and ε_{max} for different ridge filters (carbon ions) and layer thicknesses between 1 and 3 mm (s_{min}=4 cm, s_{max}=9 cm). $\langle t \rangle$=1 mm.

Triangular ridge filters   Due to the difficulties in manufacturing ridge filters with curved profiles (see Sec. 5.3.6), and the fact that a single triangular ridge filter is comparable to an optimized ridge filter if the layer thickness is similar to the ridge filter peak height, it is of interest to have an extra look at triangular ridge filters.

A ridge filter used for low-energy protons should be as thin as possible, in order not to scatter the beam. Carbon ions are less sensitive to scattering. One solution

could therefore be to equip the dual particle nozzles with two individually remov-able triangular ridge filters. During proton operation, one or both ridge filters are removed, while both ridge filters are inserted for carbon operation.

A study of the SOBP quality for 0, 1 and 2 triangular ridge filters for a target at depth 4 to 9 cm is shown in Fig. 5.15.

Up to about 2 mm layers, the maximum dose deviation ε_{max} is below 3% for sin-gle and double triangular ridge filters. At 3 mm layers, the gain with a double ridge filter for carbon ions is significant (ε_{max} =3% vs. ε_{max} =15%). Likewise, the dose variance is smaller than 1% for layers up to 3 mm (0.5% at 2.5 mm) for the double filter, while a single filter causes $\sqrt{S^2}$ =3% at 3 mm layers. Thus, double triangular ridge filters for carbon ions is strongly motivated to allow for layers of 3 mm. At 2 mm layers (the ridge filter peak height), a single triangular ridge filter is sufficient. However, to allow for some flexibility in the treatment planning system (other layer thicknesses than 2 mm), two triangular ridge filters is preferred for carbon ions.

For protons, the difference between single and double triangular ridge filter is smaller than for carbon ions. The SOBP variance is similar with and without ridge filter up to about 2.5 mm layer thickness, and the maximum dose deviation up to about 3.5 mm. To keep the dose variance below 1%, layers up to 3 mm are ac-ceptable with a single triangular ridge filter (ε_{max} below 3%). Thus, when small beam sizes and short irradiation times are important during proton beam irradiation of superficial tumors, a single triangular ridge filter would be a good compromise.
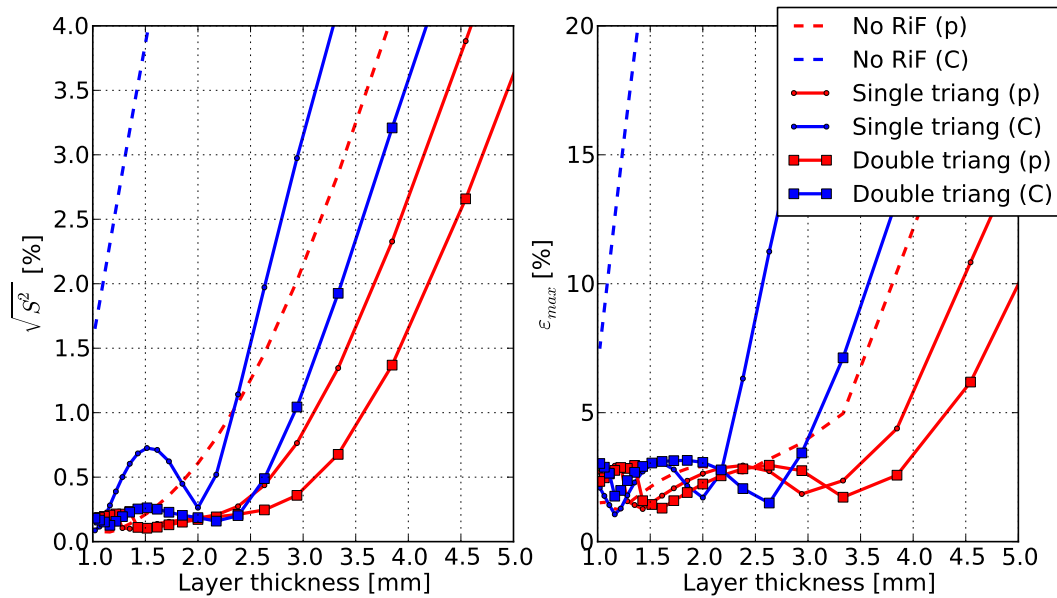

Figure 5.15: $\sqrt{S^2}$ and ε_{max} for no ridge filter, one triangular ($\langle t \rangle$=1 mm) ridge filter and one double triangular ridge filter ($\langle t \rangle$=2 mm), protons (p) and carbon ions (C).

5.3.6 Manufacturing

Accurate manufacturing of ridge filters with curved ridges (i.e. all except the triangular ones) can prove difficult, considering the relatively short peak-to-peak distances. Weber et.al mentions that the required profile accuracy is in the order of 5-10 μm [46]. A comparison between the optimized ridge profile and the triangular one is shown in Fig. 5.16. The tiny difference illustrates that the homogeneity of the SOBP is very sensitive to the exact shape of the ridge profile.
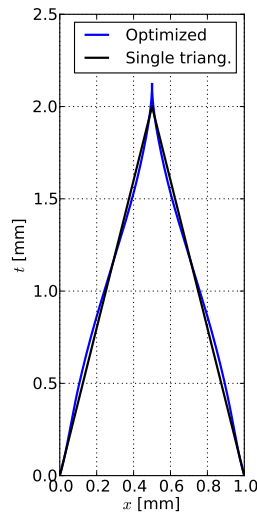
Figure 5.16: Comparison between optimized and triangular ridge profile (1 mm average thickness).

The ridge filter profile manufactured by Weber et. al deviates somewhat from the ideal, optimized shape, which can be seen in measurements of the filtered Bragg curve [46, Fig. 2]. In [53], the main motivation for choosing triangular ridge filters is motivated by the difficulties in producing optimized ridge filters accurately.

The manufactured Gaussian ridge filter described in [54] is intended for uniform scanning with proton beams, and tested with beam sizes much larger than the smallest beams foreseen in MedAustron (60 mm vs. 4 mm).

Although an optimized ridge filters theoretically is the best choice, it may be a safer option to use triangular ridge filters which are easier to manufacture[5].

5.3.7 Ridge filter choices

Due to the strict requirements of a few micrometer accuracy in manufacturing of the ridge filters with curved profiles, ridge filters with triangular profiles are preferred. During irradiation with carbon ion beams, double triangular ridge filters with a water equivalent peak height of at least 2 mm are recommended, allowing for layers up to about 3 mm. Thicker ridge filters with a higher peak would allow for even thicker layers.

---

[5]Confirmed in private discussion with workshop engineers at CERN.

For protons, ridge filters are only necessary for superficial tumours, since the natural widening of the Bragg peak is dominating at higher energies. Since low energy protons are sensitive to scattering, a single triangular ridge filter is a reasonable compromise between the desired longitudinal broadening of the Bragg peak and the undesired lateral widening of the beam profile.

5.4 Nozzle optimization

5.4.1 Nozzle options

Fig. 5.17 shows schematic models of the considered nozzle options, containing beam intensity, position and profile monitors (1.1 mm water equivalent thickness (WET) [47] in total.). Ridge filters to widen the Bragg peak can be inserted when irradiating with carbon beams. The overall length from the last vacuum window to the IC is 92 cm. It ends with a 0.5 mm thick Plexiglas window, preventing the patient from accidentally reaching into the nozzle. Two methods for minimizing the beam growth in the nozzle are considered:

1. Minimizing the nozzle-to-isocenter air gap by moving the monitors as close as possible to the patient (Nozzle 1)[6].

2. Inserting a helium-filled bellow between the vacuum window and the monitors to reduce scattering in air (Nozzle 2).

Additionally, the use of a movable Plexiglass Range Shifter (RS) to reduce the beam range below the limit set by the minimum extraction energy is evaluated (Nozzle 3). The RS thickness (3 cm) roughly corresponds to the minimum proton and carbon ion beam penetration depths, which allows for irradiation of superficial tumors. For safety reasons, the gap between nozzle (or RS) and patient should be at least 10 cm.

5.4.2 Beam size at isocenter, in air

The horizontal and vertical beam FWHM at the isocenter, in air, have been calculated for the three nozzle options for all beam energies. Results are presented in Fig. 5.18. With Nozzle 1, beam sizes of 9-11 mm can be achieved for low-energy protons, depending on air gap. Adding the helium chamber reduces the beam sizes by at most 2 mm. At higher energies, the scattering effect decreases and the FWHM at the IC approaches 4 mm.

Fig. 5.18 also shows beam sizes for Nozzle 3 for RS-IC air gaps of 0, 10 and 20 cm (No air gap would be equivalent to attaching the range shifter/pre-absorber directly on the patient's skin). Although the primary purpose of the range shifter is to reduce the penetration depth, it could also be used to beam sizes of target sizes up to 6-7 cm depth, if the RS-IC air gap can be made sufficiently small (compare curves for Nozzle 1 and Nozzle 3 at 0 cm air gap).

---

[6]This functionality is available in Gantry 2 at PSI [13].
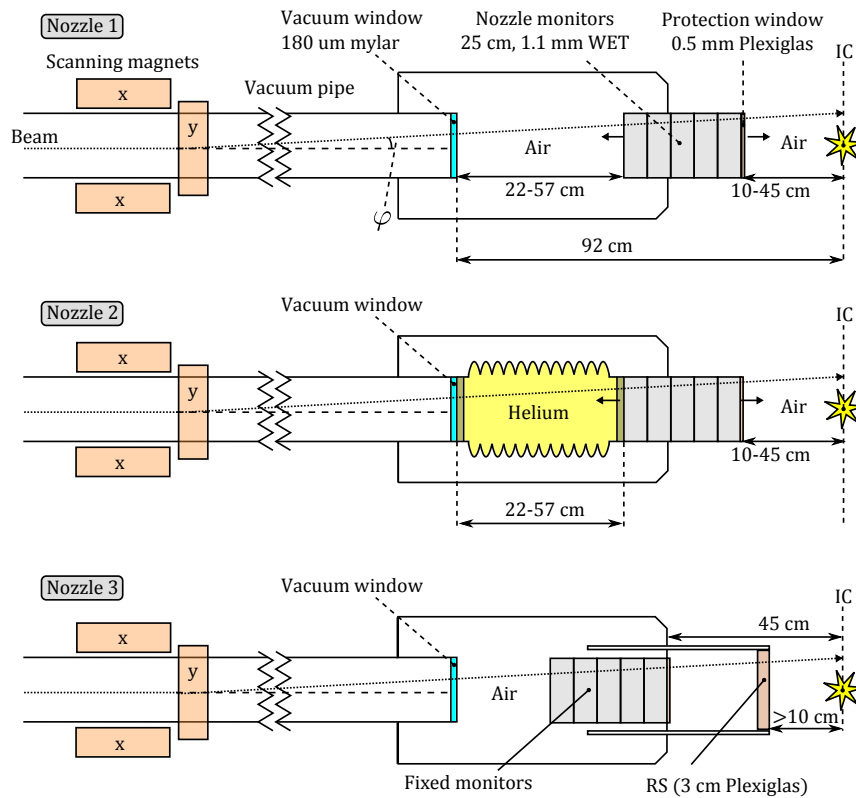
Figure 5.17: Different nozzle options. 1: Movable monitors. 2: Movable monitors and helium chamber after vacuum window. 3: Fixed monitors, movable (and removable) range shifter.
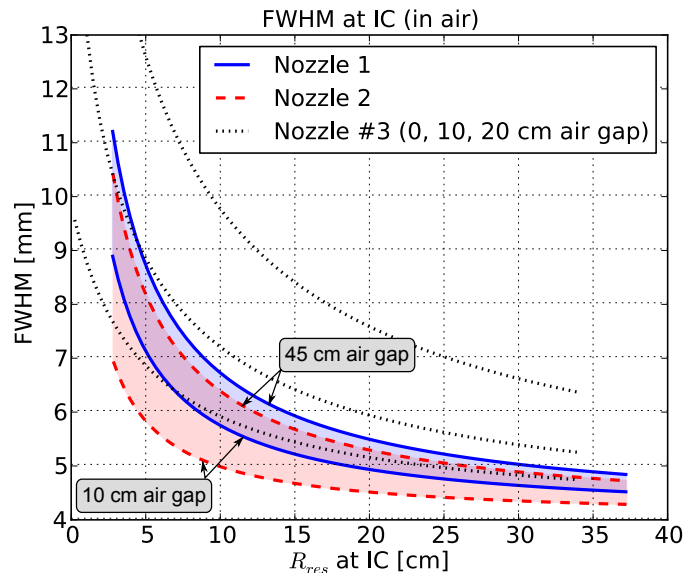
Figure 5.18: Proton beam FWHM (horizontal and vertical) in air vs. residual beam range in water, R_{res}, at the IC. The three dotted lines indicate IC beam sizes for Nozzle 3, assuming 0, 10 and 20 cm air gap between RS and IC. Colored bands indicate the FWHM span achievable by moving the nozzle.

The primary use of the RS is to reduce the penetration depth below 3 cm, but, compared to Nozzle 1, it could also be used to reduce beam sizes of targets at up to 6-7 cm depth, if the RS-IC air gap is short enough.

5.4.3 Beam size at Bragg peak

By placing a water phantom at the IC, as shown in Fig. 5.19, the beam sizes at the Bragg peak (in water) can be evaluated. The amount of water in front of the IC is matched to the energy of the incoming beam such that the Bragg peak is located exactly at the IC, where the unscattered beam is focused. Resulting beam sizes at the IC are presented in Fig. 5.20.
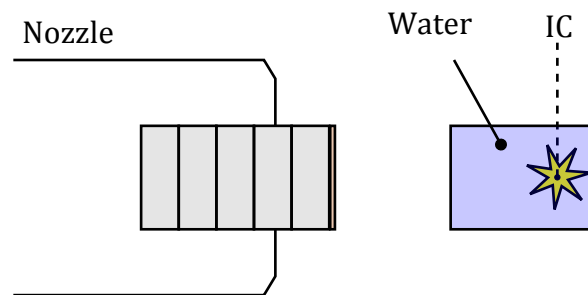
Figure 5.19: A water slab is placed over the isocenter. The amount of water upstream of the IC is matched to the energy of the beam such that the beam stops at the IC.

At low energies, the beam size is approximately equal to the beam size in air: scattering in the nozzle and air gap is dominant. With increasing energy, scattering in the nozzle decreases and the Bragg peak beam sizes feature a minimum at a depth of 5-8 cm. At high energies, scattering in the water phantom is dominant. The gain in optimizing the nozzle and reducing the air gap would be less than 1 mm for proton Bragg peaks deeper than 10 cm.

As shown in the figure, carbon ions are largely insensitive to the layout of the nozzle. Even at the lowest energy (120 MeV/n), the difference between the "best" and "worst" case is less than 1 mm.

5.4.4 Scattering contribution from individual elements

In order to better understand the different nozzle models, and clearly see which elements have the largest impact on beam size, it is of interest to evaluate the scattering term from individual elements. Fig. 5.21 shows the scattering term σ_s (scaled by a factor 2.35 for conversion to Gaussian FWHM) for single elements as a function of x, where x is defined in Tab. 5.2. For most elements, x is simply the distance between the element and the isocenter, but for the helium/air column in front of the monitors, x denotes the thickness of this column.

The vacuum window, at $x = 92$ cm has an individual scattering contribution just above 4 mm. The air gap between vacuum window and monitors (air gap 1) also shows a comparatively large scattering contribution: 4 to 5.5 mm depending on
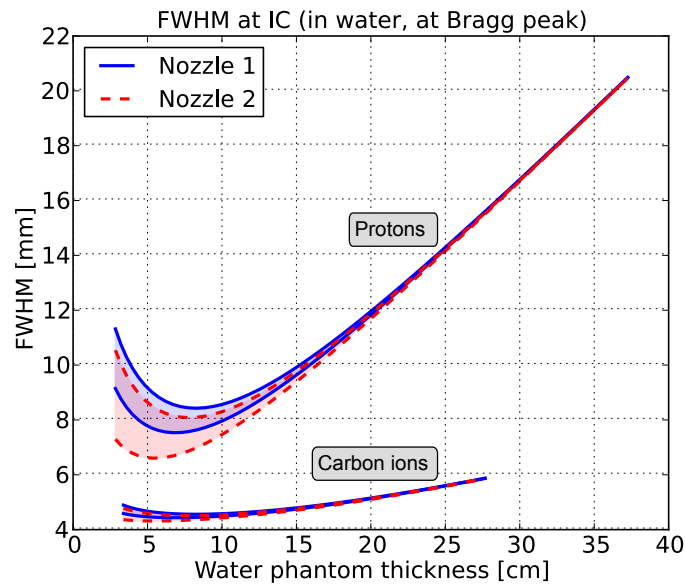
Figure 5.20: As Fig. 5.18, but at IC, in water (at Bragg peak). Upper bands: 60-250 MeV protons, lower bands: 120-400 MeV/n carbon ions.

Element	Sketch	Range x (movable nozzle)
Vacuum window		92 cm
Monitors		10-45 cm
Protection window		10-45 cm
Helium window		35-92 cm
Helium		22-57 cm
Ridge filter		10-45 cm
Air gap 1 (before monitors)		22-57 cm
Air gap 2 (after monitors)		10-45 cm

Table 5.2: Individual nozzle elements, sketch of the parameter x in Fig. 5.21 and nominal values of x.
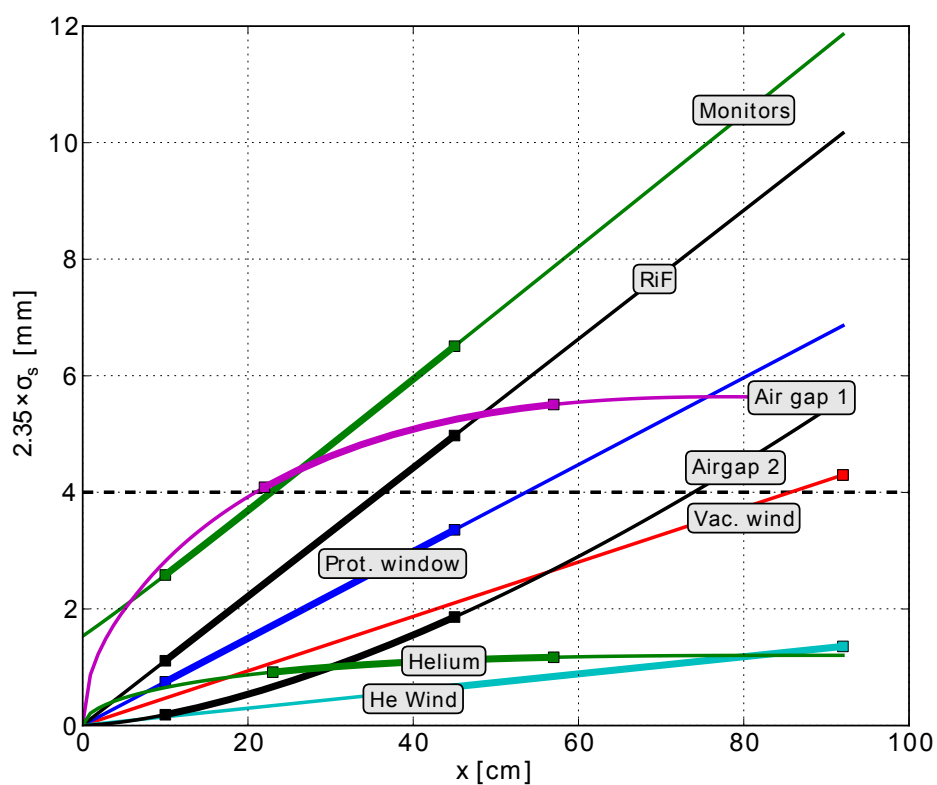
Figure 5.21: Scattering term (multiplied by 2.35) for single elements as a function of *x* (60 MeV proton beam). Nominal values are indicated with squares and the unscattered beam width (4 mm) by a dashed line.

nozzle position. The air gap between monitors and patient (air gap 2) has a minor effect on the beam size, just like the helium and helium windows.

5.4.5 Ridge filter placement in nozzle

In an evaluation of the CNAO ridge filters [53], it is mentioned that a distance of several decimeters between the two triangular parallel ridge filters is necessary in order to achieve the desired modulation of the pristine Bragg peak. These ridge filters are parallel, as shown in Fig. 5.8a. As explained in Sec. 5.3.2, a minimum distance between the two parallel ridge filters is necessary to "blur" the correlation between position at the second ridge filter and range reduction introduced by the first ridge filter. A similar reasoning can be applied to the minimum distance required between double orthogonal ridge filters and the patient. If the patient is too close to the ridge filter, the structure of the grooves will be reflected in the dose distribution in the patient - something that should be avoided.

The minimum distance condition for a homogeneous target dose distribution is as follows: a point beam leaving a ridge filter "valley" must have grown to a 1-σ width of at least $\lambda/1.6$ when reaching the target. A conservative estimate of the minimum distance can be made by considering the 1-σ beam divergence, $\sigma_{x'}$, of a point beam leaving the last ridge filter. Neglecting further divergence increase due to scattering in air, this beam will at the target entrance have grown to a 1-σ beam size of (linear beam growth in a drift space) [46]:

$$\sigma_x = \sigma_{x'} \times L \tag{5.22}$$

where L is the distance from last ridge filter to target. The minimum distance is thus:

$$\sigma_x \geq \frac{\lambda}{1.6} \Leftrightarrow L \geq \frac{\lambda}{1.6\sigma_{x'}} \tag{5.23}$$

For a completely homogeneous dose distribution, one should consider $\sigma_{x'}$ for the particles that have passed through the valley part of both ridge filters (minimum amount of scattering). For a zero-emittance incoming beam, $\sigma_{x'}$ after the last ridge filter is given by scattering in the nozzle and vacuum window. At this stage, one should also take into account the non-zero divergence of the incoming beam. However, since the horizontal phase space distribution is a bar of charge, the incoming horizontal beam divergence will be very small for large beam sizes, when the bar of charge is "lying down".

Tab. 5.3 summarizes the minimum distances required between last ridge filter and target, in order to produce homogeneous entrance doses. The beam divergence has been calculated by taking scattering in the main nozzle elements into account (see Sec. 5.4):

- 180 μm mylar vacuum window

- monitors with a water equivalent thickness of 1.1 mm

- Ridge filter base plate

    - 2×0.3 mm Plexiglas (minimum base plate thickness of ridge filters [46]) for carbon ions

    - 1×0.3 mm Plexiglas for protons

Beam type	Energy	$\sigma_{x'}$ [mrad]	L_{min} [cm]
Carbon ions	120 MeV/n	1.7	36
Carbon ions	400 MeV/n	0.57	109
Protons	60 MeV	6.1	10
Protons	100 MeV	3.7	17

Table 5.3: Required distance from last ridge filter to isocenter.

From Tab. 5.3, it is clear that for protons, the minimum distance is easily achievable, and the proton ridge filter is preferrably placed as the last nozzle element. For low energy carbon ions, the minimum distance is achievable as well. High energy carbon ions, on the other hand, require a longer distance (about 1 m to generate a homogeneous surface dose for high energies). To resolve this, the following is proposed:

> **Proposal:** Place the two ridge filters at the end of the nozzle. One ridge filter (to be used for protons and carbon ions) should have a thin base plate. The other one, used exclusively for carbon ions, has a thicker base plate to intentionally increase the beam divergence in order to reduce L_{min} to approximately 45 cm.

The base plate thickness of the carbon ion dedicated ridge filter would have to be 7 mm to allow a L_{min} of 45 cm for 400 MeV/n. Unfortunately, such a thick base plate would inevitably blow up a 120 MeV/n point beam to 4.6 mm FWHM at the isocenter ($\sqrt{4.0^2 + 4.6^2} = 6.1$ mm minimum beam size, adding the 4 mm unscattered beam size in quadrature). But this drawback is difficult to circumvent: if one of the ridge filters are moved back to 109 cm upstream of the isocenter, the low energy carbon ion beam size will still be large, due to the longer drift space.

Therefore, the author's recommendation is to foresee one dual particle ridge filter with a thin base plate, and one carbon ion ridge filter with a possibly thicker base plate to increase the beam divergence. Since the exact thickness easily can be modified any time (e.g. by simply adding millimeter-thick Plexiglas plates until satisfactory results are achieved), it is best determined at a later stage, when real depth-dose curves and lateral dose distributions can be measured.

5.5 Beam position and measurement accuracy: limitations

Even if the beam position and profile monitors were perfectly accurate, there are two fundamental limitations in measuring the beam position and width. The first is the *pitch*: the strip width (if strip chambers are used), or the wire spacing (if wire chambers are used). The second is the number of particles upon which the measurement is based. With a low number of particles, the granularity of the beam is apparent. This causes a larger uncertainty of the number of particles per strip which will affect the accuracy of the evaluated beam position and beam size. For completeness, a brief evaluation of the achievable beam position and beam width accuracy and precision will be made, with respect to these two factors.

Method For a given value of pitch and number of particles, a total of N_{meas} independent measurements are simulated. At each measurement, particles are randomly distributed over the strips (following a normal distribution with a given FWHM), giving a number of particles per strip. The beam position and width is calculated from the number of particles per strip. The accuracy and precision of the measurement is deduced from the distribution of measured values.

Monitors The monitors are assumed to be strip ionization chambers with a pitch Δx and N_s strips (a similar reasoning can be applied to wire scanners). The coordinates of the strip edges are denoted $e_0, e_1, \ldots, e_{N_s}$, with strip center coordinates, c_i:

$$c_i = \frac{e_i + e_{i-1}}{2} \tag{5.24}$$

Beam profile The beam profile at measurement k is assumed to be Gaussian with a standard deviation σ_{beam}, centered at μ_k and containing M particles:

$$M\rho_k(x) = \frac{M}{\sqrt{2\pi}\sigma_{beam}} \exp\left(-\frac{(x-\mu_k)^2}{2\sigma_{beam}^2}\right) \tag{5.25}$$

During irradiation, the beam moves and the central position of the beam (μ_k) is not correlated to the strip edges. The center of the beam profile, μ_k, is therefore taken from a uniform distribution over one strip width:

$$\mu_k \in \mathcal{U}[-\frac{\Delta x}{2}, \frac{\Delta x}{2}] \tag{5.26}$$

The probability of a particle to hit strip i at measurement k, p_{ki}, is:

$$p_{ki} = \int_{e_{i-1}}^{e_i} \rho_k(x)\mathrm{d}x = \Phi\left(\frac{e_i - \mu_k}{\sigma}\right) - \Phi\left(\frac{e_{i-1} - \mu_k}{\sigma}\right) \tag{5.27}$$

113

where Φ is the cumulative distribution function of a normal distribution. The number of particles detected on one strip, n_{ki} follows a binomial distribution:

$$n_{ki} \in \text{Bin}[M, p_{ki}] \tag{5.28}$$

Since n_{ki} are drawn from a random distribution, the exact number of particles in one measurement (the sum of all strips) will be:

$$M_k = \sum_{i=1}^{N_s} n_{ki} \tag{5.29}$$

Beam position evaluation   At each measurement, the beam position is evaluated via a Center-Of-Gravity (COG) calculation:

$$\bar{x}_k = \frac{1}{M_k} \sum_{i=1}^{N_s} c_i n_{ki} \tag{5.30}$$

Beam width evaluation   The beam width (1-σ) is evaluated as:

$$\sigma_{meas,k} = \sqrt{\sum_{i=1}^{N_s} \frac{n_{ki}}{M_k}(c_i - \bar{x}_k)^2} \tag{5.31}$$

A multiplication by 2.35 gives the FWHM of the beam (Sec. 3.5). Note that other ways of evaluating the beam width are possible (e.g. by fitting, or by using only a few strips close to the beam center). The advantage with a simple rms-calculation, compared to fitting, is that $\bar{x}_k$ is in any case evaluated for the beam position.

Measurement accuracy   The average error in the COG measurement over N_{meas} measurements is:

$$\Delta(COG) = \frac{1}{N_{meas}} \sum_{k=1}^{N_{meas}} \bar{x}_k - \mu_k \tag{5.32}$$

Likewise, the average error in the beam width (1-σ) measurement is:

$$\Delta(RMS) = \frac{1}{N_{meas}} \sum_{k=1}^{N_{meas}} \sigma_{meas,k} - \sigma_{beam} \tag{5.33}$$

Measurement precision   The standard deviation (rms) of the beam COG measurement error is:

$$\sigma(COG) = \sqrt{\frac{1}{N_{meas}} \sum_{k=1}^{N_{meas}} (\bar{x}_k - \mu_k - \Delta(COG))^2} \tag{5.34}$$

and the rms of the beam width measurement error is:

$$\sigma(RMS) = \sqrt{\frac{1}{N_{meas}} \sum_{k=1}^{N_{meas}} (\sigma_{meas,k} - \sigma_{beam} - \Delta(RMS))^2} \tag{5.35}$$

Random strip gain error   The impact of random gain errors of the strips can been evaluated by multiplying each strip weight n_{ki} with a factor $1 + g_{ki}$, where g_{ki} is drawn from a normal distribution.

Evaluation   Shown in Fig. 5.22 and 5.23 are $\Delta(COG)$, $\Delta(RMS)$, $\sigma(COG)$ and $\sigma(RMS)$ for a 4 and 10 mm beam (FWHM) without gain errors ($g_{ki} = 0$). Within the given pitch range (0.5-4 mm) and spot weight range (10^3-10^6 particles), the *accuracy* of the COG calculation is in the μm range: there is no reason that the COG algorithm should systematically over- or under-estimate the beam position.

The *precision* (standard deviation) of the calculated COG is more or less independent of the pitch, up to about 3.5 mm pitch for a 4 mm beam. Below 3.5 mm pitch, the wider beam shows the "worst" precision ($\sigma(COG)$ is proportional to beam size).

The beam width - as calculated via the standard deviation of the strip weight distribution - shows a systematic over-estimation ($\Delta(RMS) > 0$). This systematic offset is fairly constant if the number of particles is more than 1000, though, and can therefore simply be subtracted from the estimated beam width (Eq. 5.31). As with the beam position, the uncertainty in the beam width measurement is proportional to the beam width.
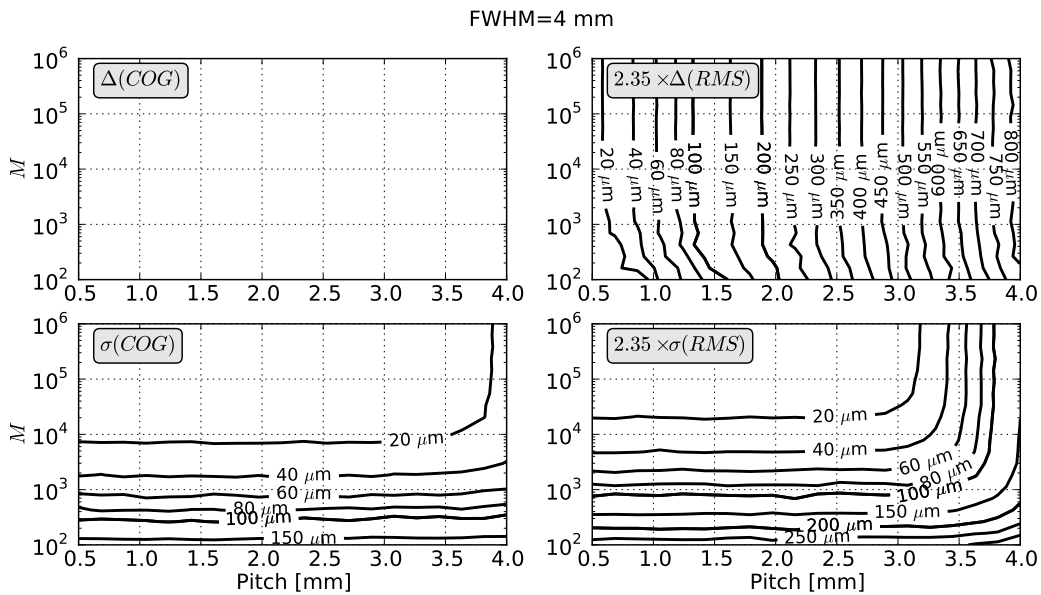

Figure 5.22: Lower limit of beam position and width measurement accuracy and precision, 4 mm beam. $\Delta(COG) \approx 0$ for all cases.

The accuracy and precision of the beam position and beam size measurement have also been evaluated with non-zero random gain errors. The conclusion is that gain errors up to ~2% has a negligible impact on the results shown in Figures 5.22 and 5.23
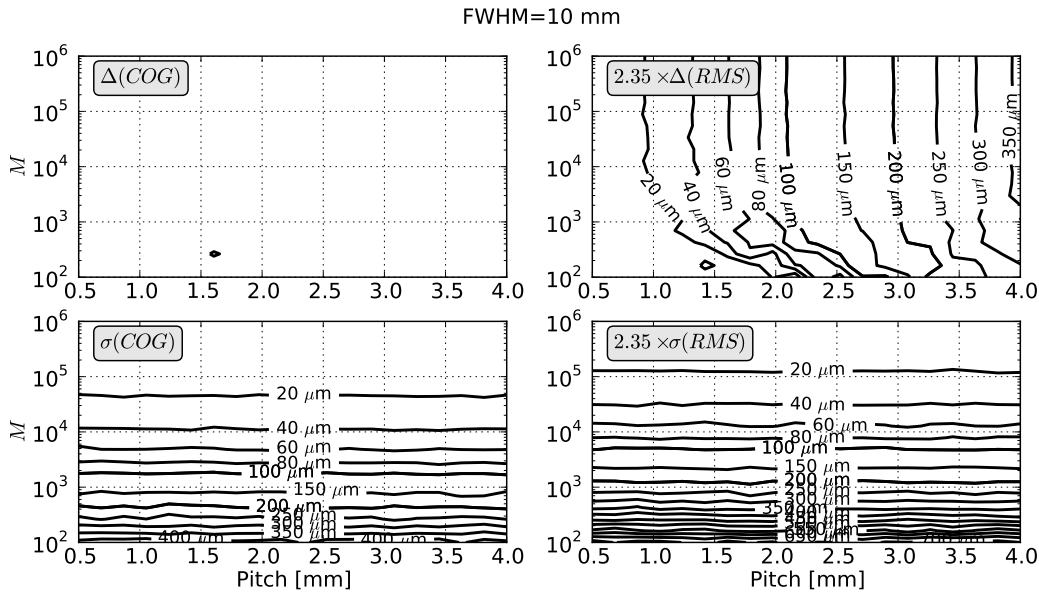
Figure 5.23: Lower limit of beam position and width measurement accuracy and precision, 10 mm beam. $\Delta(COG) \approx 0$ for all cases.

5.6 Results and discussion

With the foreseen beam optics along the proton gantry, vacuum windows at the coupling point would increase the horizontal beam FWHM by about 1 cm at the isocenter, while leaving the vertical beam size unaffected. An even larger increase would be caused by filling the 90° dipole at the end of the gantry with helium. Thus, in order to produce isocenter beam profiles that are small and symmetric, vacuum should be kept all the way to the nozzle.

In order to reduce the number of scanning layer for carbon ions, double triangular ridge filters with a ridge peak height of 2 mm WET (about 1.7 mm Plexiglas), or $\langle t \rangle = 1$ mm, are recommended, as these are easiest to manufacture. Maximum dose ripple along the SOBP would be less than 2% for 2.5 mm layers. Nothing prevents using double triangular ridge filters also for low energy protons, but since they are more sensitive to scattering, a single triangular ridge filter would suffice, in situations where both small beams and few layers are important.

No distance between the ridge filters is required if they are positioned orthogonally. The proposed solution is to put both ridge filters at the end of the nozzle, after the monitors. The main advantage with this solution is that double calibration of the beam monitors is avoided: beam profile and energy deposition at the monitors will differ if a ridge filter is inserted upstream. This would require up to three sets of calibrations for 0, 1 or 2 ridge filters inserted.

A minimum distance from ridge filter to patient is necessary to avoid reflecting the ridge filter structure in the dose distribution. For protons and low-energy carbon ions, an air gap of 45 cm between ridge filter and isocenter (maximum available) is fully sufficient to "blur" the correlation between range reduction and beam position

caused by the ridge filters. This would permit placing the ridge filters after the monitors. However, in the high-energy carbon ion case, additional scattering would be necessary, since 45 cm is not enough. Additional scattering can easily be achieved by making the base plate of one of the ridge filters thicker. This ridge filter is moved out for proton use.

The question whether is is clinically necessary to produce homogeneous doses at the target *entrance* for high energies (i.e. deep-seated targets) must be raised, though. If not, the thicker base plate could be omitted.

The reduction of beam size growth due to scattering that can be achieved by movable monitors and a helium chamber for a 60 MeV proton beam can also be achieved (within 1 mm) by placing a range shifter directly on the patient skin, and extract at a higher energy. Although this would require some patient-specific hardware, it could in the end be a simpler solution than using moving monitors. Moving monitors without a helium chamber would "only" gain about 2 mm FWHM, or about 1 mm lateral penumbra. With movable monitors, the deflection angle of the scanned beam (φ, Fig. 5.17) necessitates correlating the measured beam position to the longitudinal position of the monitors. This increases the complexity of beam verification during scanning and certification of the beam delivery system. In the proton gantry, though, the optics is such that $\varphi \approx 0$, even when irradiating at the edges of the field. A compromise could therefore be to use moving monitors in the proton gantry only. Patients for which small, low-energy protons beams are important could be treated in the gantry, while deep-seated tumors could be irradiated in any of the other rooms.

By gathering the removable ridge filters and (re-)movable range shifter at the end of the nozzle, construction could possibly be simplified and movable parts inside the nozzle avoided. If the ridge filters were simply sliding in and out of the beam path, the transverse dimensions of the nozzle would have to be at least twice the scanning field size. With the ridge filters (and range shifter) at the end of the nozzle, they could simply be installed as "flaps", resting flat along the external sides of the nozzle when not used, and inserted into the beam path by a 270° rotation, to allow for a compact nozzle that can be brought close to the patient.

Chapter 6

Patient dose evaluation

In this chapter, the scattering model and Bragg curve parameterization will be combined to a single *spot dose model*, which allows for evaluation of the dose distribution at the patient, taking scattering in the optimized nozzle and patient and ridge filter Bragg peak modulation into account. With this dose model, the relation between different kinds of beam delivery imperfections and dose inhomogeneities can be studied systematically. These relations are the basis for the requirements on the entire beam delivery chain.

6.1 Target dose calculator

A module for 3-dimensional dose calculations has been implemented in Python, using the framework previously described. This module ("dosecalc") allows for performing dose calculations on a spot-by-spot basis with the following input:

- Longitudinal slab geometry (material, density, thickness of each slab)

- Beam optics (Twiss function or simple drift space)

- Ridge filters

- Unscattered horizontal and vertical beam profile

6.1.1 Single spot

The longitudinal dose profile of a single spot is given by the Bragg peak parameterization described in Sec. 3.3.5. The 2-dimensional transverse profile at any depth s (taking multiple Coulomb scattering in nozzle, air gap and patient into account) can be calculated via Eq. 5.10. Combining the longitudinal and transverse beam models gives the 3-dimensional dose distribution, $D_i(x, y, s)$ of a single spot with initial energy E_i containing n_i particles:

$$D_i(x, y, s) = n_i \times \rho_x(x, s) \times \rho_y(y, s) \times \frac{L_i(s)}{P} \qquad (6.1)$$

Figure 6.1: Deposited dose per proton. Incoming beam is 6 mm FWHM in both planes. Range from 5 to 25 cm.
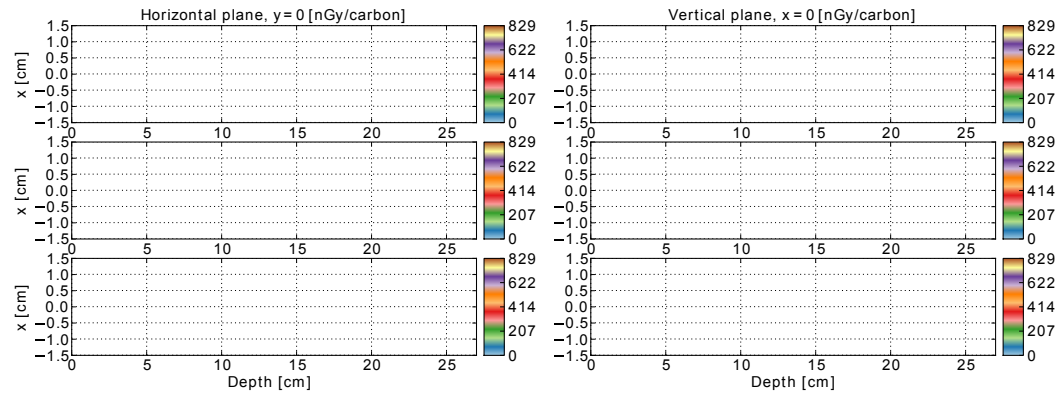
Figure 6.2: Deposited dose per carbon ion. Incoming beam is 6 mm FWHM in both planes. Range from 5 to 25 cm.

with units of: $\mathrm{cm}^{-1} \times \mathrm{cm}^{-1} \times \mathrm{J/cm} \times \mathrm{cm}^3/\mathrm{kg} = \mathrm{Gy}$.

Examples of the spot profiles in a transverse plane are shown in Fig. 6.1 (protons) and 6.2 (carbon ions). For low energy proton beams, the longitudinal Bragg peak is very pronounced, while at higher energies, multiple coulomb scattering and range straggling diffuses the Bragg peak so much that the physical dose is very low. Range straggling is in fact the main contributor to this effect [55].

The Bragg peak of carbon ions, on the other hand, is clearly visible even for high energies.

In order to reduce calculation time, the transverse beam profiles $\rho_x(x,s)$ and $\rho_y(y,s)$ (trapezoidal/Gaussian) are interpolated from pre-calculated beam profile tables[1].

---

[1]3D-interpolation, where scattering term, initial beam FWHM and transverse coordinate are input parameters. The implemented module supports parallel computing. Time-consuming computations are made in C++, embedded in Python.

6.1.2 Target volume

For dose calculation of an entire target, the nominal weights (particles/cm^2) of each layer is first calculated by making an SOBP optimization according to the desired depths of the Bragg peaks of each layer and the specific ridge filter to be used. From this the nominal number of particles per spot, n_i, are calculated, and the total target dose is given by a sum over the dose contribution from each spot:

$$D_{tot}(x,y,s) = \sum_i n_i \times D_i(x - X_i, y - Y_i, s) \tag{6.2}$$

(X_i, Y_i) are the transverse center coordinates of spot i.

6.2 Dose calculation assumptions

6.2.1 Target volume definitions

The quality of the target dose depends not only on the beam delivery system, but also on the target type. A superficial tumor treated with carbon ions will be more sensitive to e.g. beam size errors than a deep-seated tumors irradiated with protons: in the latter case, initial beam width errors are "washed away" by multiple coulomb scattering in the patient, and the degree of overlap between neighboring spots is high.

A target with a large extent in depth will be divided into many layers. While the dose to the most distal layer can be delivered during a single spill, the dose to the proximal layers will be delivered over several spills, as the SOBP is built up. Statistically, random errors will therefore be more suppressed in the proximal part of the target, roughly as $1/\sqrt{N}$, where N is the number layers behind the considered layer. A "worst case" would be a target that is narrow in depth, where the total dose can be delivered in a single spill, and the "statistical smoothing" of random errors is low.

In order to evaluate the sensitivity to different types of beam delivery errors, two "standard targets", with an extent of 5 cm in depth have been studied:

1. *Superficial target* at 3.5 - 8.5 cm depth. 3.5 cm is the minimum beam range in water for carbon ions (120 MeV/n).

2. *Deep-seated target* at 22 - 27 cm depth. 27 cm is the maximum beam range in water for carbon ions (400 MeV/n).

Additionally, the dose homogeneity in a transverse plane of a single-layer target has also been studied:

3. At the Bragg peak at 7 cm depth. Bragg peak beam sizes around this depth are as smallest for proton beams.

For Target 1 and Target 2, the layer thickness has been set to 4 mm for protons and 2.5 mm for carbon ions.

6.2.2 Nozzle

The design of the nozzle will influence the transverse beam size at the patient: the more the nozzle scatters the beam, the larger the spot sizes at the patient will be, resulting in more overlap between neighboring spots (for any given spot-to-spot distance). As mentioned previously, higher overlap will mitigate the effect of random errors. Therefore, a "minimal" nozzle design (Nozzle 2, 10 cm air gap, Fig. 5.17) has been used in the following dose error analysis, to reach conservative requirements on the beam delivery chain. Any beam delivery requirements deduced from a minimal nozzle design will hold also for other nozzle designs, which scatter the beam more. The 0.5 mm Plexiglas protection window at the end of the nozzle has been omitted in these simulations, as it is not decided whether it is absolutely necessary during irradiation. However, as seen in Fig. 5.21, this has a negligible effect on the lateral beam profile, even for low energy proton beams, for which the vacuum window and the monitors are the main contributors to beam growth before the patient.

6.3 Dose error analysis

6.3.1 Method

6.3.1.1 Relative dose error

Not even the dose distribution of a perfectly executed irradiation, without any kinds or errors, will be homogeneous. The primary cause of unavoidable dose inhomogeneities will be longitudinal dose variations along the SOBP, due to the limited number of layers. Minor transverse dose variations will also occur due to the non-straight slopes of the trapezoidal beam profile. However, these dose variations are known in advance and can (at least in principle) be taken into account already in the treatment plan. For a dose error analysis, we are only interested in dose inhomogeneities that are due to imperfections of the beam delivery system. In order to isolate the effect of the beam delivery imperfections, the relative dose error, ε, in any point (x, y, s) in the target volume is defined as the relative deviation from the *nominal* dose distribution:

$$\varepsilon(x, y, s) = \frac{D(x, y, s) - D_{nom}(x, y, s)}{D_{nom}(x, y, s)} \tag{6.3}$$

6.3.1.2 Target dose inhomogeneity: definition

The inhomogeneity of the target dose, H, is in this work defined as the standard deviation of the relative dose error in the entire target volume V, i.e:

$$H = \sqrt{\int_V (\varepsilon - \bar{\varepsilon})^2 dV} \tag{6.4}$$

The inhomogeneity is used to compare target dose distributions resulting from different types of beam delivery errors. The transverse size of the target is large enough to give statistically reliable values.

6.3.1.3 Lateral dose frame

In each layer, spots are placed in a cartesian or shifted grid from (X_{min}, Y_{min}) to (X_{max}, Y_{max}). At the edges of the field, the dose will drop to zero, where the *relative* dose error potentially can be very large due to the low nominal dose D_0. To avoid extremely large values of ε, which would affect the homogeneity evaluation, a lateral frame is introduced, such that the dose is only calculated in the center region of the target, where the nominal dose is reasonably flat. The width of the frame F_z is defined as:

$$F_z = 3.7 \frac{W_z}{2.35} \tag{6.5}$$

where W_z is the horizontal or vertical FWHM of the unscattered beam and the term $W_z/2.35$ corresponds to the initial 1-σ beam width. In the case of a staggered grid, an extra margin of $\Delta/2$ is added to the horizontal frame.

The dose frame is illustrated schematically in Fig. 6.3
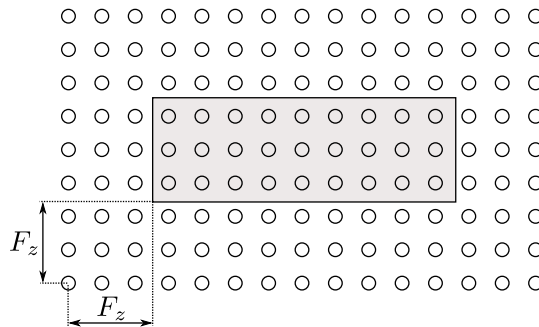
Figure 6.3: Lateral dose frame example (not to scale). Spots (spot centers indicated as rings) are placed in a transverse grid, Cartesian or shifted. The dose is only calculated in the central region (shaded) to avoid edge-effects.

6.3.2 Beam energy errors

6.3.2.1 Range accuracy

The displacement in range ΔR to an energy error ΔE can be calculated via the exponential energy-range fit from Eq. 3.69:

$$\Delta R = \alpha k E^{k-1} \Delta E \tag{6.6}$$

For a given tolerance in range, the energy error must be smaller than:

$$\Delta E \leq \frac{\Delta R}{\alpha k E^{k-1}} \tag{6.7}$$

Tab. 6.1 summarizes the resulting energy requirements for proton and carbon ions for a range accuracy of $\Delta R = 0.25$ mm in water, i.e. identical to the clinical requirements of the CNAO hadron therayp facility (see Sec. 3.3.3). It is notable that

the resulting energy accuracy is stricter at higher energies. This can be understood from the $1/\beta^2$-term in the Bethe-Block formula: the lower the initial energy of the particle, the higher the stopping power. An energy error at 60 MeV will thus have a lower impact on the particle range than the same (absolute) energy error at a high energy.

Beam	ΔR	ΔE	$\Delta E/E$
Proton, 60 MeV	0.25 mm	0.27 MeV	4.5×10^{-3}
Proton, 250 MeV	0.25 mm	0.10 MeV	0.4×10^{-3}
Carbon ion, 120 MeV/n	0.25 mm	0.48 MeV/n	4.0×10^{-3}
Carbon ion, 400 MeV/n	0.25 mm	0.21 MeV/n	0.5×10^{-3}

Table 6.1: Energy accuracy requirements for a range accuracy of 0.25 mm.

The required energy accuracy of down to 0.1 MeV for the specified range accuracy is much stricter than e.g. the 0.4 MeV specified by Chu et.al in [16], but still more relaxed than the 80 keV seen in the TERA report [15], or the 40 keV by Arduini et.al in [56][2].

6.3.2.2   SOBP homogeneity

The impact of random energy errors on the SOBP homogeneity has been studied for the superficial and deep-seated target (Target 1 and 2), using the parameterized Bragg curves from Sec. 3.3.5. For carbon ions, the Bragg curve has been filtered with double triangular ridge filters (each with a 2 mm water equivalent peak height).

For a given target (deep-seated or superficial, protons or carbon ions), the dose along ideal SOBP (no energy errors) is first calculated according to:

$$D_{nom}(s) = \sum_i \lambda_i \left. \frac{-\mathrm{d}E}{\mathrm{d}s} \right|_{E_0 = E_j} \times \frac{1}{P} \tag{6.8}$$

where the layers weights λ_i are optimized to give as flat dose as possible. Then, random energy errors ΔE_{ik} are applied to each layer (not changing the layer weights) and the resulting dose along the SOBP, $D_k(s)$ is then given by:

$$D_k(s) = \sum_i \lambda_i \left. \frac{-\mathrm{d}E}{\mathrm{d}s} \right|_{E_0 = E_i + \Delta E_{ik}} \times \frac{1}{P} \tag{6.9}$$

where the energy errors are randomly distributed:

$$\Delta E_{ik} \in [-\Delta E, +\Delta E] \tag{6.10}$$

The relative dose error along the SOBP, $\varepsilon_k(s)$ is calculated from Eq. 6.3 (transverse variations are not considered in the energy error analysis).

---

[2]40 keV corresponds to 10% of the smallest energy step.

To quantify the impact of energy errors on the SOBP quality, we consider the maximum over- and under-dosage anywhere along the SOBP, $\max\{\varepsilon_k(s)\} = \varepsilon_{max,k}$ and $-\min\{\varepsilon_k(s)\} = \varepsilon_{min,k}$, as well as the SOBP inhomogeneity, H_k. As the energy errors ΔE_{ik} are chosen randomly, these three quantities will vary significantly from time to time, due to the limited number of layers.

For a given value of maximum energy error amplitude, ΔE, a large number of SOBPs have been calculated, which allows for evaluation of the mean value and standard deviation of the maximum and minimum dose error and dose inhomogeneity, respectively.

Figures 6.4 to 6.7 show the mean maximum over- and under-dosage, and the SOBP homogeneity as a function of maximum energy error amplitude for a superficial and deep-seated target, protons and carbon ions. The mean of the maximum dose error and the inhomogeneity are linear to ΔE, but the spread (standard deviation) is considerable, especially for ε_{max} and ε_{min}.

A superficial proton target (Fig. 6.6) is more sensitive to energy errors than a deep-seated targets (Fig. 6.4). This may seem to contradict the findings in Tab. 6.1, where the energy accuracy requirements are higher at high energies. However, one should keep in mind that the unfiltered low-energy proton Bragg peaks are significantly sharper than the high-energy Bragg peaks. A longitudinally misplaced Bragg peak will therefore have a larger impact on dose homogeneity at lower energies, compared to higher energies.
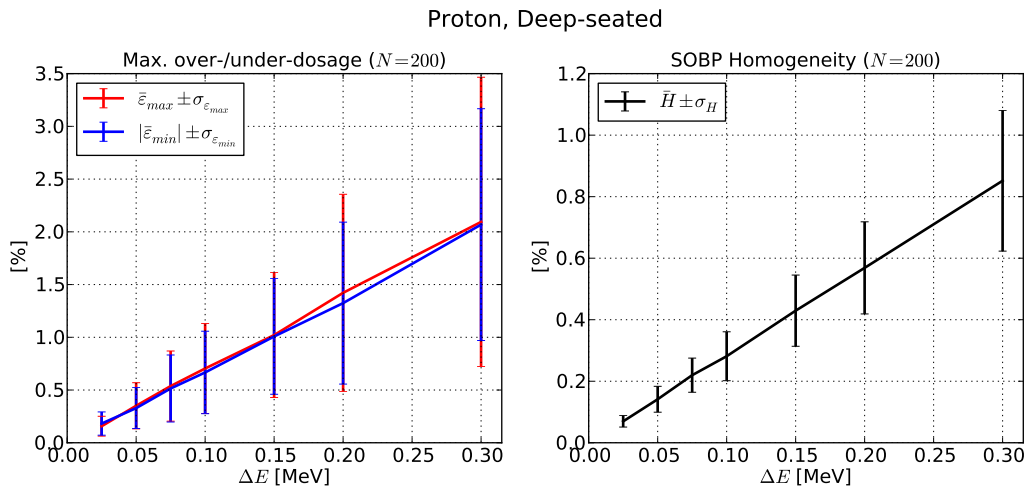
Figure 6.4: Protons, deep-seated target. Left: highest relative over- and under-dosage along SOBP vs. ΔE. Right: SOBP inhomogeneity vs. ΔE. Error bars indicate $\pm$ one standard deviation.

One way to specify energy accuracy requirements for the homogeneity of the SOBP (which has been set to better than 1% in Sec. 2) would be to specify an upper limit of ΔE such that $\bar{H} \leq 1\%$ for all four target configurations. However, about half of the SOBP's would then have a inhomogeneity worse than 1%. For this reason,
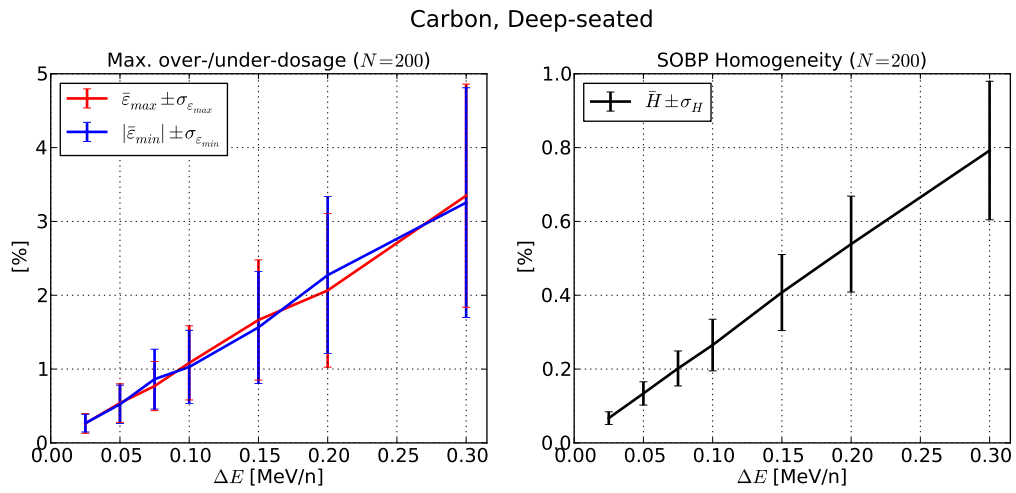
Figure 6.5: As Fig. 6.4, carbon ions, deep-seated target.
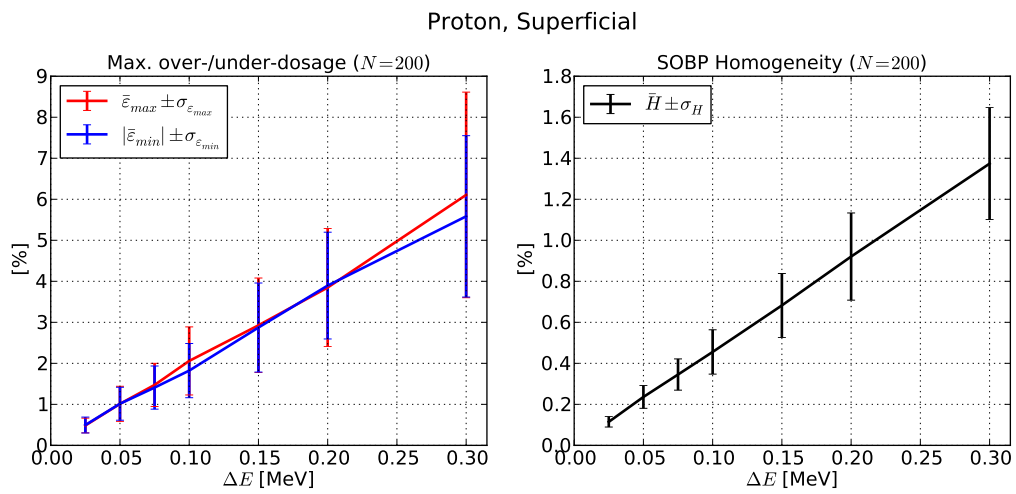
Figure 6.6: As Fig. 6.4, protons, superficial target.
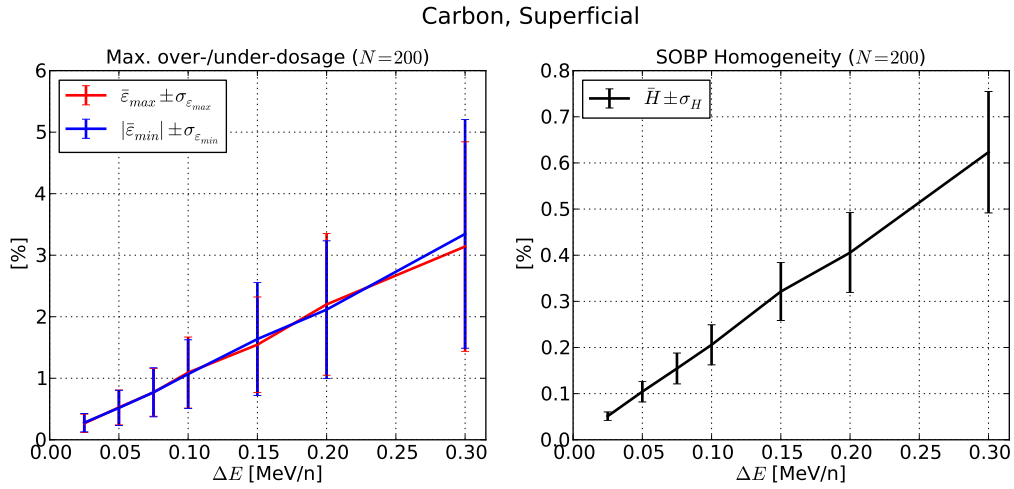
Figure 6.7: As Fig. 6.4, carbon ions, superficial target.

the energy accuracy with respect to SOBP homogeneity is chosen such that:

$$\bar{H} + \sigma_H \leq 1\% \tag{6.11}$$

In this case, the SOBP homogeneity will be better than 1% in about 86% of all cases. Tab. 6.2 summarizes the mean inhomogeneity, $\bar{H}$ and the max/min dose error at the 1-σ level, for a 1% dose homogeneity at the 1-σ level ($\bar{H} + \sigma_H = 1\%$).

Target	ΔE	$\bar{H}$	$\bar{H} + \sigma_H$	$\bar{\varepsilon}_{max} + \sigma_{\varepsilon_{max}}$	$|\bar{\varepsilon}_{min}| + \sigma_{\varepsilon_{min}}$
p, Superficial	0.18 MeV	0.8%	1.0%	4.9%	4.7%
p, Deep-seated	0.28 MeV	0.8%	1.0%	3.1%	3.0%
C^{6+}, Superficial	0.40 MeV/n	0.8%	1.0%	6.5%	6.6%
C^{6+}, Deep-seated	0.30 MeV/n	0.8%	1.0%	4.7%	4.8%

Table 6.2: Energy accuracy requirements to achieve a 1% dose homogeneity at the 1-σ level. The two last columns show the resulting maximum over- and under-dosage error at the 1-σ level.

6.3.3 Transverse errors: distribution

For a given target volume, an optimized treatment plan for a rectangular target has been generated. To evaluate the impact of errors in spot weight, beam positioning

and beam size, intentional errors were added to every spot:

$$n_i' = \begin{cases} n_i \times (1+\eta_i) & \text{Discrete mode} \\ n_i \times (1+\eta_i-\eta_{i-1}) & \text{Integral-dose driven} \end{cases}$$
$$X_i' = X_i + \delta x_i$$
$$Y_i' = Y_i + \delta y_i$$
$$W_{x,i}' = W_x + \delta W_{x,i} + \Delta W_x$$
$$W_{y,i}' = W_y + \delta W_{y,i} + \Delta W_y \tag{6.12}$$

η_i, δx_i, δy_i, $\delta W_{x,i}$, $\delta W_{y,i}$ are random errors that vary from spot to spot while ΔW_x, ΔW_y are static beam width errors that are the same at all spots. The random errors are assumed to be uniformly distributed:

$$\eta_i \in \mathcal{U}[-E_n, +E_n]$$
$$\delta x_i \in \mathcal{U}[-E_{pos,x}, +E_{pos,x}]$$
$$\delta y_i \in \mathcal{U}[-E_{pos,y}, +E_{pos,y}]$$
$$\delta W_{x,i} \in \mathcal{U}[-E_{W,x}, +E_{W,x}]$$
$$\delta W_{y,i} \in \mathcal{U}[-E_{W,y}, +E_{W,y}] \tag{6.13}$$

The motivation for choosing a uniform distribution deserves an explanation. The simplest possible beam verification strategy would be to accept all spots with an error (spot weight, position or size) within some tolerance band $\pm E_k$ and generate an interlock for all spots with an error larger than E_k. The goal of this evaluation is to find an upper limit of acceptable spot errors, $E_{k,tol}$. At the moment, we do not care how often an interlock would occur, only that the "approved" spots guarantee an acceptable target dose homogeneity. If the distribution of spot weight errors were normal distributed, the approved spots would be distributed as a Gaussian curve truncated at $\pm E_{k,tol}$. Now, the acceptance limit $E_{k,tol}$ would obviously depend on the width of the Gaussian: the narrower the error distribution, the larger the acceptable value of $E_{k,tol}$. A very narrow distribution would even allow for infinite values of $E_{k,tol}$, which makes sense: if large amplitude errors are very rare, they will not affect the *global* homogeneity of the target. However, relating the error tolerance limit to the width of the error distribution has this problem: in order to guarantee that the delivered dose is homogeneous, it is not enough to verify that the error of all spots are within $\pm E_{k,tol}$. One must also verify - while scanning - that the *distribution* of the spot errors is sufficiently narrow. This would be complicated - especially at the beginning of a spill, when only a few spots have been delivered and there is no reliable statistics available on the error distribution.

Thus, the conservative assumption is made that the distributions of spot errors (weight, position and width) are wide enough that they are practically uniform within $\pm E_{k,tol}$. Of course, in practice, one must ensure that it is more narrow in order to have a sufficiently low number of spots causing an interlock.

6.3.4 Effect of individual errors

We first look at the effect of a single type of error, setting all other errors to zero.

6.3.4.1 Weight errors

Fig. 6.8 shows an example of $\varepsilon(x, y, s = 7 \text{ cm})$ in the center region of Target 3 (single layer) for carbon ions (discrete scanning mode). Each spot is subject to a random weight error of up to $\pm 20\%$, causing a few local hot and cold spots with about 8% dose error.
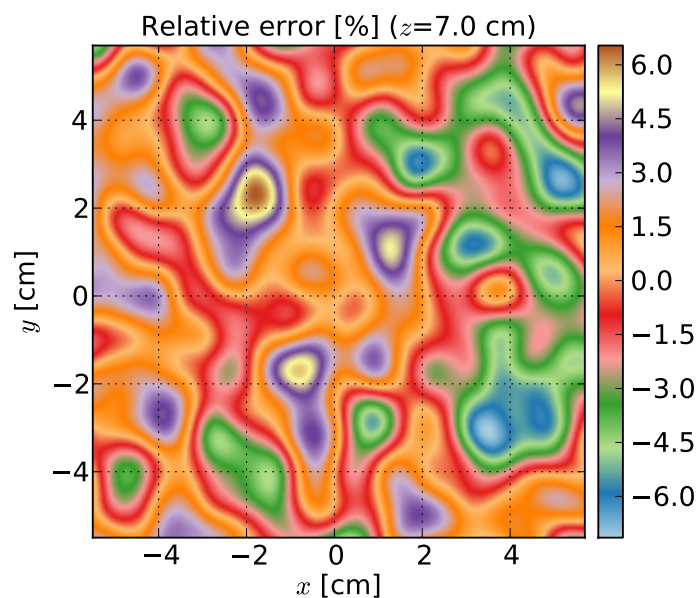
Figure 6.8: $\varepsilon(x, y, s = 7 \text{ cm})$, Target 3, carbon ions, discrete scanning mode ($W_x = W_y = 1.0 \text{ cm}$, Cartesian grid, $\delta x = \delta y = 3.33 \text{ mm}$). "+" indicate spot center positions.

A histogram over the relative dose error, ε, is shown in Fig. 6.9. Although the individual spot weight errors are uniformly distributed, the dose error distribution is more Gaussian, due to the overlapping between spots.

With a multi-layer target, the total dose to the proximal layers is given in several smaller portions. Statistically, random errors are therefore "smoothed out" in the proximal part. This is clearly seen in Fig. 6.10a, where $\varepsilon(x, y, s)$ for Target 1 (superficial) in a large number of randomly selected target points are plotted as a function of depth (spot weight errors up to $\pm 10\%$). The spread is smaller at the proximal part of the target, compared to the distal part, which behaves more like a single-layer target. The overall distribution is still well described by a Gaussian, as shown in Fig. 6.10b.

The target inhomogeneity H has been evaluated for different spot weight error amplitudes E_n, showing that the inhomogeneity is linear to the error amplitude.

Figure 6.9: Histogram of ε from Fig. 6.8.

(a) ε vs. depth.

(b) Histogram of ε
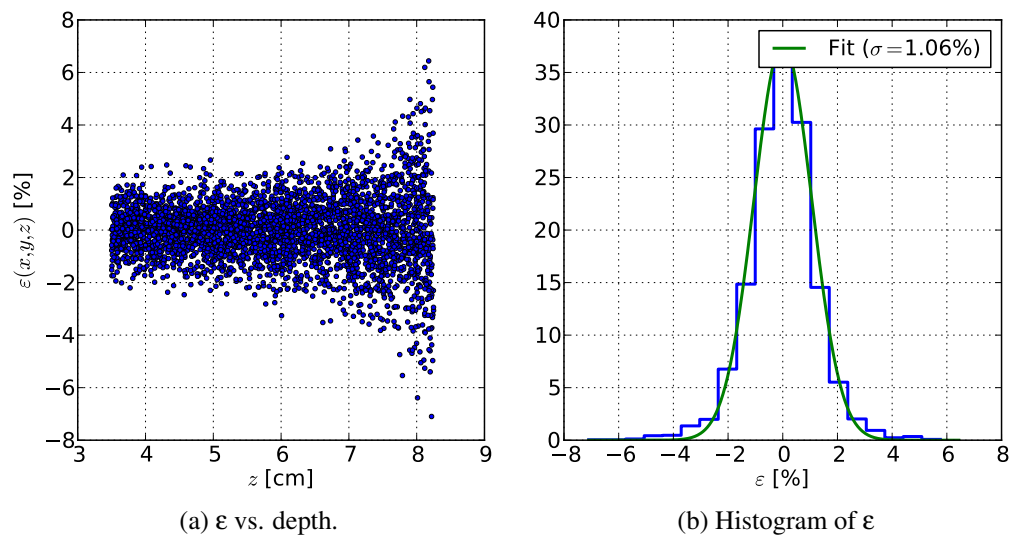
Figure 6.10: Relative dose error, Target 1, carbon ions, $W = 10$ mm, $E_n = 10\%$, discrete scanning mode. Points are chosen randomly in target volume.

Thus, the inhomogeneity in a target that is subject to only spot weight errors can be predicted by the simple relation:

$$H_n = k_n \times E_n \qquad (6.14)$$

The coefficient k_n will depend on target type (1-3), beam type, beam size, spot-to-spot distance and scanning mode (quasi-discrete or discrete).

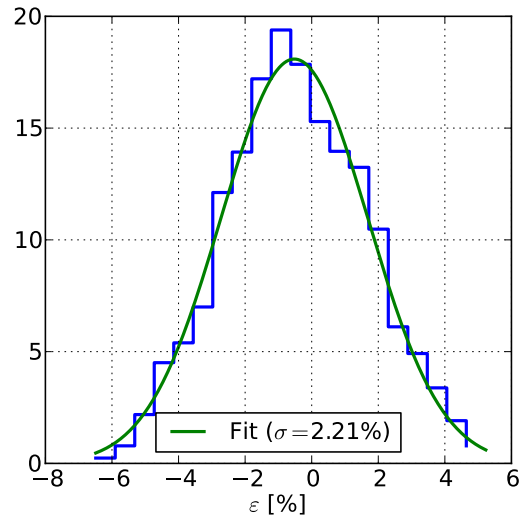For all considered spot-to-spot distances and beam sizes, integral-dose driven scanning is superior to a discrete scanning mode (lower values of k_n). As expected, the gain with integral-dose driven scanning decreases with an increasing spot-to-spot distance, but Fig. 6.11 clearly shows that the integral-dose driven scanning suppresses the effect of random spot-to-spot errors even at longer values of $\delta x, \delta x$.
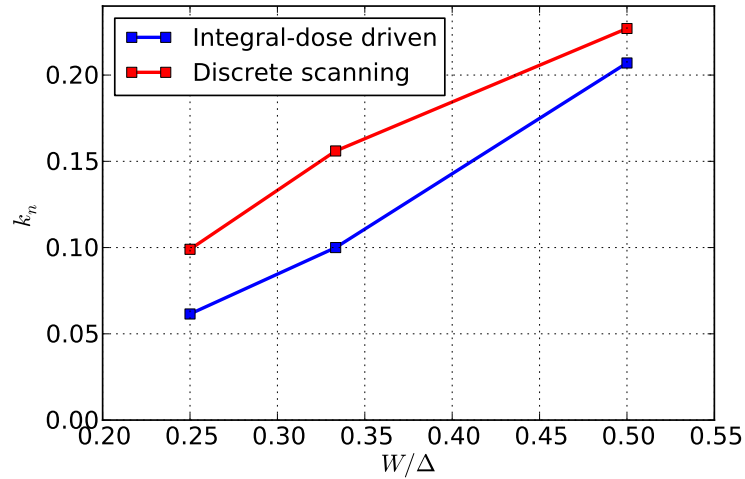
Figure 6.11: k_n vs. scanning step for the two scanning modes (carbon ions, 10 mm FWHM, Target 3.

Since integral-dose driven scanning is both the simplest and best scanning method to implement, a discrete scanning mode has been discarded in the following. All coefficients for Target 3 (single layer) are summarized in Tab. 6.3.

For a given value of E_n, a tight spot grid will produce a more homogeneous dose, which would imply that a tight spot grid ($\delta/W = 4$) is always preferred. However, reducing the spot-to-spot distance also reduces the number of particles per spot (quadratically). With low-weighted spots, the *relative* weight error tends to be larger.

The deep-seated target is less sensitive to random errors than the single-layer and superficial target. This is due to the larger spot sizes deeper in the target (see Fig. 5.20), which results in a higher degree of overlap between spots: random errors from spot to spot are then more smoothed out.

6.3.4.2 Beam position errors

Since the horizontal and vertical beam profiles are not identical, they have been evaluated independently to determine whether positional errors in one direction cause

larger dose inhomogeneities than in the other direction. An evaluation similar to the spot weight errors was made, revealing that the target inhomogeneity is linear to the maximum amplitude of the positional error:

$$H_{pos,x} = k_{pos,x} \times E_{pos,x}$$
$$H_{pos,y} = k_{pos,y} \times E_{pos,y} \tag{6.15}$$

The positional error coefficients are summarized in Tab. 6.3 (Cartesian and shifted grid give similar results).

Notable is that the horizontal beam position errors cause a slightly larger dose inhomogeneity than vertical beam position errors, in particular for carbon ion beams. This is due to the sharp horizontal beam profiles.

6.3.4.3 Random beam width errors

Random beam width errors in x and y have been evaluated like the positional errors, and again the inhomogeneity is linear to the maximum beam width error amplitude:

$$H_{W,x} = k_{W,x} \times E_{W,x}$$
$$H_{W,y} = k_{W,y} \times E_{W,y} \tag{6.16}$$

Tab. 6.3 summarizes the width error coefficients (Cartesian and shifted grid give similar results).

Notable is that proton beams, which are more smeared out due to scattering, are less sensitive to random beam width error; particularly for deep-seated targets, where the Bragg peak beam size is completely dominated by scattering, and thus insensitive to initial beam width errors.

6.3.4.4 Static beam width errors

Fig. 6.12 shows two example of the transverse dose distribution for Target 3 (carbon ions), being subject to a static horizontal beam width error of $\Delta W_x = -0.5$ mm, using a Cartesian (left) and shifted (right) grid. With the Cartesian grid, the width error causes stripes of systematic over- and under-dosage of up to about 1.3%, while these errors are significantly suppressed with the shifted grid.

The dose error distribution from a static beam width error is not normal distributed. The "hot" and "cold" stripes formed where the horizontal edges no longer overlap produce two spikes at the edges of the dose error distribution, with a more or less flat plateau in between. A conservative, but simple, model of this dose distribution is two dirac-delta funtions at the maximum dose error, $\pm\varepsilon_{max}$:

$$f(\varepsilon) \approx \frac{1}{2}[\delta(-\varepsilon_{max}) + \delta(\varepsilon_{max})] \tag{6.17}$$

An example of the relation between the ΔW_x and ε_{max} is shown in Fig. 6.13 for a Cartesian grid, 8 mm beam, 3.33 mm spot-to-spot distance. ε_{max} periodically drops to zero whenever the beam width $W_x + \Delta W_x$ is an integer multiple of the spot-to-spot

(a) Cartesian grid.

(b) Shifted grid.
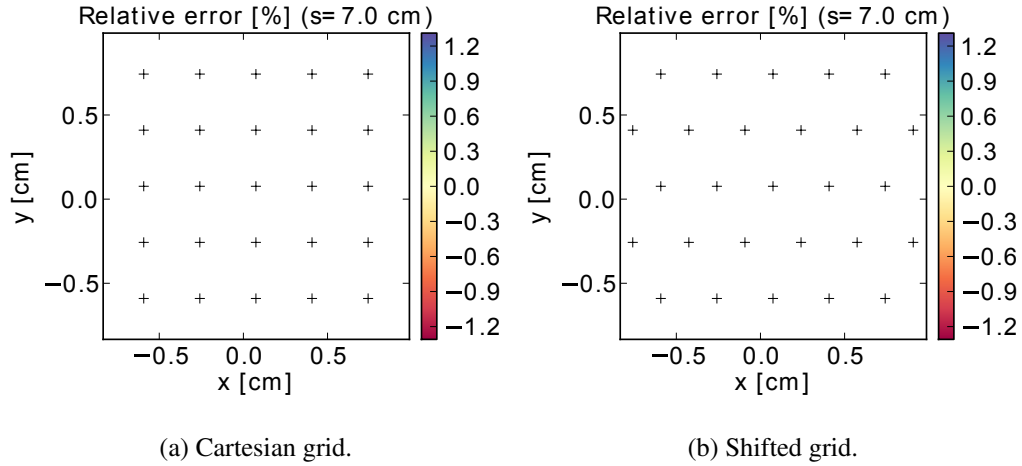
Figure 6.12: $\varepsilon(x,y)$ at Bragg peak, Target 3. $\Delta W_x =$ -0.5 mm, $\Delta = W/3$.

distance. For smaller-amplitude width errors, though, the relation is linear and ε_{max} can be conservatively estimated to:

$$\varepsilon_{max} = k_{\Delta W} \Delta W \tag{6.18}$$
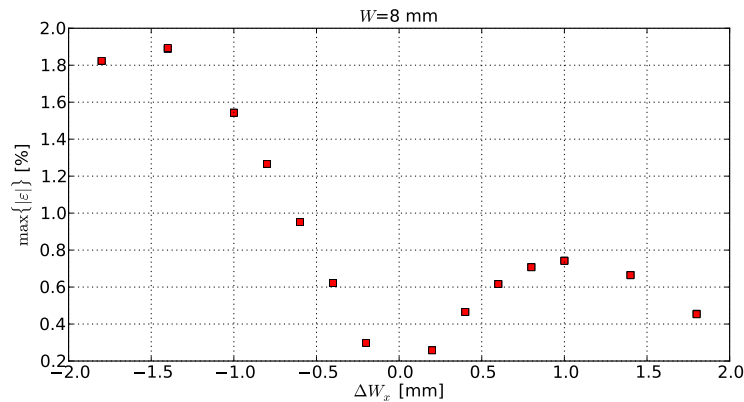
Figure 6.13: Max. relative dose error vs. ΔW_x ($W = 8$ mm).

Coeffients for Cartesian and shifted spot grids are presented in Tab. 6.3.

Moderate static beam width errors in the vertical direction will not cause any dose inhomogeneities, since periodic superposition of Gaussian profiles yield a homogeneous dose for spot-to-spot distances of up to [46]:

$$\Delta \leq 1.6 \times \frac{W'_y}{2.35} \approx 0.7 \times (W_y + \Delta W_y) \tag{6.19}$$

With a vertical spot-to-spot distance $\delta y = \Delta W_y/m$, this gives:

$$\frac{\Delta W_y}{W_y} \geq \frac{1}{0.7m} - 1 \tag{6.20}$$

133

For a moderate spot grid ($m = 3$):

$$\frac{\Delta W_y}{W_y} \geq -0.52 \tag{6.21}$$

The vertical beam profile can thus be up to 52% more narrow than planned, without affecting the dose homogeneity. *Positive* beam width errors (larger spots than planned) are primarily not limited by dose homogenity restrictions, but the width of the lateral penumbra.

6.3.5 Normal distributed errors

For completeness, the impact of normal-distributed errors has also been evaluated. The inhomogeneity remains linear to the error amplitude. Replacing the maximum error amplitude of the uniform error, E_i, with the standard deviation of the error type, σ_i gives a inhomogeneity of (from simulations):

$$H_i \approx 1.73 k_i \sigma_i \approx \sqrt{3} k_i \sigma_i = k_i E_i \tag{6.22}$$

i.e. a uniformly distributed spot weight error within $\pm 10\%$ causes the same dose inhomogeneity as a normal-distributed spot weight error with standard deviation 5.8%[3].

6.3.6 Combination of different errors

When different random errors are combined, the total effect can be calculated by adding their individual contributions quadratically (confirmed by dose calculations), i.e:

$$H = \sqrt{H_n^2 + H_{pos,x}^2 + H_{pos,y}^2 + H_{W,x}^2 + H_{W,y}^2} \tag{6.23}$$

With only random errors, the dose error distribution $f(\varepsilon)$ is Gaussian centered at $\varepsilon = 0$ with a standard deviation of H:

$$f(\varepsilon) = \frac{1}{\sqrt{2\pi}H} \exp\left(-\frac{\varepsilon^2}{2H^2}\right) \equiv G(0, H) \quad \text{[Random errors]} \tag{6.24}$$

However, static horizontal beam width errors ΔW_x approximately shifts half the field by $-\varepsilon_{max}$ and the other half by $+\varepsilon_{max}$:

$$f(\varepsilon) = \frac{1}{2}\left[G(-\varepsilon_{max}, H) + G(+\varepsilon_{max}, H)\right] \tag{6.25}$$

For a shifted grid, ε_{max} will be small, and $f(\varepsilon)$ is then well approximated with a single Gaussian.

---

[3]The factor $1/\sqrt{3}$ concides with the standard deviation of a uniform distribution within $\pm E_i$, which is $E_i/\sqrt{3}$.

6.3.7 Error coefficients

Error coefficients from individual errors (smallest and largest beam sizes) are presented in Tab. 6.3.

Table 6.3: Error coefficients (neg=negligible).

Target	Δ/W	W [mm]	Proton	Carbon
k_n [%/%] (integral-dose driven)				
1	1/4	4	0.010	0.018
1	1/4	10	0.022	0.023
1	1/3	4	0.018	0.029
1	1/3	10	0.033	0.035
1	1/2	4	0.031	0.063
1	1/2	10	0.072	0.080
3	1/4	4	0.015	0.043
3	1/4	10	0.039	0.061
3	1/3	4	0.027	0.077
3	1/3	10	0.075	0.100
3	1/2	4	0.061	0.162
3	1/2	10	0.151	0.207
k_{pos} [%/mm], x/y				
1	1/4	4	1.00/0.70	1.62/1.36
1	1/4	10	0.97/0.69	1.00/0.68
1	1/3	4	1.41/1.16	2.22/1.82
1	1/3	10	1.12/0.90	1.39/0.86
1	1/2	4	1.83/1.69	3.32/2.71
1	1/2	10	1.73/1.44	1.33/1.30
3	1/4	4	1.52/1.46	4.41/3.43
3	1/4	10	1.68/1.35	2.67/1.78
3	1/3	4	2.13/1.96	6.10/4.89
3	1/3	10	2.15/1.86	3.73/2.48
3	1/2	4	3.33/2.85	8.93/7.30
3	1/2	10	3.43/2.79	5.41/3.67
k_W [%/mm], x/y				
1	1/4	4	0.20/0.30	0.71/0.65
1	1/4	10	0.35/0.27	0.51/0.34
1	1/3	4	0.26/0.41	1.00/0.88
1	1/3	10	0.46/0.44	0.67/0.47
1	1/2	4	0.42/0.59	1.50/1.25
1	1/2	10	0.68/0.68	0.96/0.65
3	1/4	4	0.29/0.44	1.74/1.71
3	1/4	10	0.62/0.61	1.33/0.98
3	1/3	4	0.39/0.60	2.31/2.31
Continued on next page				

Table 6.3 – continued from previous page

Target	W [mm]	Δ/W	Proton	Carbon
3	1/3	10	0.86/0.85	1.75/1.30
3	1/2	4	0.55/0.85	3.57/3.37
3	1/2	10	1.29/1.26	2.56/1.96
$k_{\Delta W_x}$ [%/mm], Cartesian/Shifted				
3	1/4	4	neg/neg	neg/neg
3	1/4	10	neg/neg	0.48/neg
3	1/3	4	neg/neg	neg/neg
3	1/3	10	neg/neg	2.67/neg
3	1/2	4	neg/neg	1.00/neg
3	1/2	10	neg/neg	9.3/0.90

6.4 Beam delivery chain requirements

6.4.1 Longitudinal: Extraction energy accuracy

To achieve a longitudinal dose homogeneity of 1% rms or better in 86% of the cases, an energy accuracy of 0.18 MeV is required for protons, and 0.3 MeV/n for carbon ions, as shown in Sec. 6.3.2.2.

With a range accuracy of 0.25 mm, an energy accuracy of down to 0.1 MeV is required for (high energy) protons. The corresponding figure for carbon ions is 0.21 MeV/n.

6.4.2 Transverse: Permissible error amplitudes

As stipulated in Chapter 2, the combined effect of errors in beam size, positioning and spot weight should not exceed $\sigma_{tol} = 2\%$ (rms). With static beam width errors present, the relative dose error distribution is not Gaussian, so this condition is reformulated as a maximum dose error *variance* of σ_{tol}^2:

$$\int_{-\infty}^{\infty} \varepsilon^2 f(\varepsilon) \mathrm{d}\varepsilon \leq \sigma_{tol}^2 \tag{6.26}$$

Insertion of the expression for $f(\varepsilon)$ and integration gives the following simple constraint on all incoming errors:

$$H^2 + \varepsilon_{max}^2 \leq \sigma_{tol}^2 \Leftrightarrow \sum_i (k_i E_i)^2 \leq \sigma_{tol}^2 \tag{6.27}$$

where k_i are the error coefficients for error type i and E_i the maximum error amplitude (including static beam width error).

What we want is a set of permissible error amplitudes E_i that ensures a homogeneous dose for *all* foreseen beam sizes (4-10 mm). At a first unbiased stage, we

assume that all kinds of errors (6 in total) contribute equally to the dose inhomogeneity, i.e:

$$H_{pos,x}^2 = H_{pos,y}^2 = H_{W,x}^2 = H_{W,y}^2 = H_n^2 = \varepsilon_{max}^2 \qquad (6.28)$$

which yields:

$$k_i E_i = \frac{\sigma_{tol}}{\sqrt{6}} \approx 0.82\% \qquad (6.29)$$

From this, permissible error amplitudes E_i can be calculated.

The coefficients shown in Tab. 6.3 would give in total 24 different sets of permissible error amplitudes, depending on beam type, beam size, target and scanning step. In order to reduce this to a manageable amount, only two sets of coefficients is kept:

- Leave out the more relaxed coefficients for Target 1 and Target 2. The calculated permissible amplitudes E_i for Target 3 will be valid for all targets. Although Target 3 is somewhat "unrealistic", it is not unreasonable to design the beam delivery system such that it is capable of producing a homogeneous dose in a single layer. Furthermore, if the *distal* layer of a tumor is only painted once, the dose inhomogeneity in the distal part of the tumor would be similar to that of Target 3.

- With some hindsight, the $1/2$ scanning step is also omitted, since the coefficients are too high (in particular for the spot weight errors, k_n). It is then sufficient to deduce requirements for the $1/3$ scanning step, since these requirements will also satisfy the $1/4$ scanning step.

- Due to the high sensitivity to static horizontal beam width errors with a Cartesian grid (1 mm width error causes hot and cold "stripes" of up to 2.7% for a scanning step of $1/3$, carbon ions, 10 mm beam), it is strongly recommended that a shifted spot grid is used. Static beam width errors are then completely negligible up to the level where they either cause an unacceptably large lateral penumbra or make the beam smaller than the scanning step.

- The two error coefficient sets left are those for proton and carbon ions, shifted spot grid and a scanning step of $1/3$.

In order for the permissible error amplitudes to be valid for *all* foreseen beam sizes, the largest of the 4-mm FWHM and 10-mm FHWM coefficients have been used, i.e. $k_{pos,x} = 6.10$ %/mm for carbon ions (4 mm FWHM) etc.

The resulting requirements on spot weight beam position and beam width accuracy are summarized in Tab. 6.4.

To verify that the *delivered* dose was acceptable, one can only rely on the information given by the beam intensity, position and profile monitors: the reconstructed dose from the information given by these monitors should be within clinical specifications. The implication of this is that the uncertainty of the beam monitors should somehow be included into the permissible error amplitudes in Tab. 6.4. Two alternative methods have been used to this: one yielding strict conditions, assuming

Error type	Protons		Carbon ions	
	E_i	σ_i	E_i	σ_i
Spot weight	11.8%	6.8%	8.9%	5.1%
Pos, x	0.41 mm	0.24 mm	0.15 mm	0.087 mm
Pos, y	0.45 mm	0.26 mm	0.18 mm	0.10 mm
FWHM, x	1.03 mm	0.59 mm	0.39 mm	0.23 mm
FWHM, y	1.05 mm	0.61 mm	0.39 mm	0.23 mm

Table 6.4: Permissible error amplitudes, protons and carbon ions, Target 3 (σ_i indicates rms of Gaussian error).

uniformly distributed errors, and a more relaxed, assuming normal-distributed errors.

6.4.2.1 Requirements for uniformly distributed errors

The beam verification system should guarantee that the reconstructed dose of all approved spots yield an acceptable dose. The measurement error must thus be small enough that the fraction of "falsely approved" spots is sufficiently low. Falsely approved spots are all spots with an error larger than E_i (Tab. 6.4), but subject to measurement errors such that they are measured to be within tolerance. To account for measurements error, we introduce an *acceptance interval* for all error types of:

$$A_i = [-\kappa E_i, +\kappa E_i] \tag{6.30}$$

where $\kappa \leq 1$. Spots measured to be outside the acceptance interval are discarded. Depending on the measurement accuracy, some spots will still be falsely approved. Let q_e be the *measured* spot error, and q the "true" value. Since only q_e is available, $f(q|q_e)$ denotes the probability density function of the true value. The probability that a spot is falsely approved is thus (see "Consumer's risk" in [57]):

$$R(q_e) = \int_{-\infty}^{-E_i} f(q|q_e)\mathrm{d}q + \int_{E_i}^{\infty} f(q|q_e)\mathrm{d}q; \quad |q_e| \leq \kappa E_i \tag{6.31}$$

Furthermore, let $f(q_e)$ be the distribution of *measured* spot errors. The total fraction of falsely approved spots is then:

$$R_{tot} = \frac{\int_{A_i} R(q_e)f(q_e)\mathrm{d}q_e}{\int_{A_i} f(q_e)\mathrm{d}q_e} \tag{6.32}$$

With a normal-distributed measurement error, $f(q|q_e)$ is a Gaussian centered at q_e and a standard deviation of σ_{BVS}. As for the distribution of *measured* values, $f(q_e)$, we make the conservative estimation that these are also normal-distributed around zero, but with a very large spread, σ_{BDS}. Even in that case, one could require that the fraction of falsely approved spots is low. With $\sigma_{BDS} \to \infty$, the distribution

of measured values within the acceptance interval is approximately uniform, which gives:

$$R_{tot}(\sigma_{BDS} \to \infty) = \frac{1}{2\kappa E_i} \int_{A_i} R(q_e) \mathrm{d}q_e \qquad (6.33)$$

For an upper limit of 5% falsely approved spots, one possible combination of acceptance interval and measurement precision are:

$$\begin{cases} \sigma_{BVS} = 0.28E_i \\ \kappa = 0.8 \end{cases} \qquad (6.34)$$

That is, without making any assumptions on the performance of the beam delivery system, the uncertainty of the measurement should not be larger than $0.28E_i$ (Tab. 6.4), for ensuring that the accepted spots will result in an homogeneous dose distribution.

Of course, if the accuracy of the beam delivery system is low, a large fraction of the spots will be rejected. This probability is given by:

$$\gamma_{rejected} = \int_{-\infty}^{-\kappa E_i} f(q_e) \mathrm{d}q_e + \int_{\kappa E_i}^{\infty} f(q_e) \mathrm{d}q_e \qquad (6.35)$$

where $f(q_e)$ is the probability distribution of measured values, q_e. With uniformly distributed errors within $\pm \alpha E_i$ and a measurement precision of $\sigma_{BVS} = 0.28E_i$, the probability of a measurement outside the acceptance interval $\pm \kappa E_i$ is less than 0.1% if:

$$\gamma_{rejected} \leq 0.1\% \Rightarrow \alpha \leq 0.4E_i \qquad (6.36)$$

i.e. the accuracy of the beam delivery system should be 40% of the values in Tab. 6.4 to ensure that the probability of a spot to be measured outside the acceptance interval is less than 0.1%. A "typical" tumor size of the size of a fist ($\sim$0.3 l) requires about 10000 spots (W=1 cm, Δ=3.3 mm.). This would imply about 10 rejected spots per target.

Of course, if *all* of these requirements ($\sigma_{BVS} \leq 0.28E_i$, $\kappa = 0.8E_i$, $\alpha = 0.4E_i$) are fulfilled simultaneously, the dose homogeneity will be significantly better than the one stipulated in Chapter 2.

6.4.2.2 Requirements for normal-distributed errors

The other approach is to take the combined uncertainty of the beam delivery system and the measurements into account from the start. Since no exact data is available, we make the assumption that all measurement are normal-distributed around the *true* value[4] with a standard deviation $\sigma_{BVS,i}$, where i indicates the type of measurement (spot weight, beam position or beam width). Likewise, the true value is assumed to be normal-distributed around the specified value with a standard deviation of $\sigma_{BDS,i}$. The combined uncertainty is:

$$\sigma_{tot,i} = \sqrt{\sigma_{BDS,i}^2 + \sigma_{BVS,i}^2} \qquad (6.37)$$

---

[4]Systematic measurement errors are assumed to be eliminable via calibration.

and for a homogeneous distribution, $\sigma_{tot,i}$ should be less than the standard deviation in Tab. 6.4. Setting the measurement precision to:

$$\sigma_{BVS,i} \leq \frac{\sigma_{tot,i}}{2} \tag{6.38}$$

gives:

$$\sigma_{BDS,i} \leq 0.87 \times \sigma_{tot,i} \tag{6.39}$$

Tightening the measurement precision further will have small effect on $\sigma_{BDS,i}$. Requirements for proton and carbon ion are summarized in Tab. 6.5, and in the following, it is *these* requirements that will be used for performance specifications on the beam delivery system and the measurement precision (apart from the spot weight error, see Sec. 6.4.5).

Error type	Protons		Carbon ions	
	$\sigma_{BDS,i}$	$\sigma_{BVS,i}$	$\sigma_{BDS,i}$	$\sigma_{BVS,i}$
Spot weight	5.9%	3.4%	4.5%	2.6%
Pos, x	0.21 mm	0.12 mm	0.075 mm	0.043 mm
Pos, y	0.23 mm	0.13 mm	0.090 mm	0.052 mm
FWHM, x	0.52 mm	0.30 mm	0.20 mm	0.11 mm
FWHM, y	0.53 mm	0.30 mm	0.20 mm	0.11 mm

Table 6.5: Required precision (rms) of the beam delivery system and measurement.

6.4.3   Nozzle monitors

6.4.3.1   Precision

From Tab. 6.5, the rms error of the beam position measurement should not be larger than about 40 μm (carbon ions). From Fig. 5.23, this primarily sets a lower limit of the number of particles per beam position measurement:

$$N_{pos} \geq 10^4 \tag{6.40}$$

Beam movement during spot transition and finite monitor drift time will smear out the measured signal. For an accurate statement on the beam position, no spot should thus be shorter than:

$$T_{spot} \geq T_{move} + T_{drift} + \frac{N_{pos}}{\dot{N}} \tag{6.41}$$

Assuming $T_{move} \approx 200$ μs and $T_{drift} \approx 100$ μs gives minimum spot times, at highest and lowest foreseen extraction intensities, according to values in Tab. 6.6.

At $N_{spot} \geq 10^4$, the beam width measurement precision is sufficient to meet the 0.11 mm precision in Tab. 6.5.

As for the beam intensity monitor precision, the 2.6% required in Tab. 6.5 (carbon ions) is well above what is technically feasible. Since extraction stability fluctuations is expected to be the main contributor to dose inhomogeneities, this figure

	$\dot{N}$ [part/s]	$T_{spot,min}$ [ms]
Protons	1×10^8	0.4
	2×10^{10}	0.3
Carbon ions	4×10^6	2.8
	1×10^9	0.31

Table 6.6: Minimum spot times to make an accurate beam position measurement, assuming $T_{move} = 200\ \mu s$ and $T_{drift} = 100\ \mu s$.

can safely be tightened to $\sigma_{BVS,n} \leq 1\%$ (see e.g. [48]), which leaves almost the full spot weight error margin to beam intensity fluctuations.

Random gain errors of up to 2% (rms) per strip will not affect the calculated center-of-gravity and beam width noticeabley, as evaluated in Sec. 5.5.

6.4.3.2 Sampling rate

The sampling rate of the beam intensity monitors should be high enough that the sampling granularity does not cause significant weight uncertainty. With spot times down to 300 μs, a sampling rate of 500 kHz is sufficient ($\pm 0.33\%$ sampling granularity for 300 μs spots).

The beam position should be evaluated on a spot-to-spot basis, i.e. at about 3 kHz (300 μs minimum spot times). However, with a transition time of $T_{move} \approx 200\ \mu s$, a higher rate is needed, since the monitor signal will be "blurred" while the beam is moving. A spot position evaluation at 10 kHz could avoid this problem. If there are reasons to believe that the beam size varies more than acceptable during extraction (see Sec. 6.4.4), the beam width should be evaluated at the same rate as the beam position, i.e. at 10 kHz.

6.4.4 Scanning magnet precision

The beam position precision is slightly stricter in the horizontal plane (sharper beam profile). By specifying an overall current precision based on the required horizontal beam position precision, the beam width precision requirement is met automatically (since it is more relaxed). Insertion of $\sigma_{pos,x}$ from Tab. 6.4 as Δx into Eq. 3.116 to determine acceptable current ripple in the scanning magnet dipoles for protons ($B\rho/(B\rho)_{max} \geq 0.18$) and carbon ions ($B\rho/(B\rho)_{max} \geq 0.51$) gives a scanning magnet current precision of ($x_{max} = 10$ cm):

$$\frac{\Delta I}{I_{max}} \leq \begin{cases} 380\ \text{ppm (carbon ions)} \\ 380\ \text{ppm (protons)} \end{cases} \quad [\text{rms, 1}^{\text{st}}\ \text{iteration}] \qquad (6.42)$$

By chance, the required ppm accuracy of the scanning magnets is the same for both carbon ions and protons, due to the broader and less sharp proton spots.

As mentioned, there are other potential sources of beam position and size errors (e.g. magnetic field errors along the extraction line), which have not been evaluated.

To leave room for these errors, the scanning magnet power supply requirement can be further tightened to some 200 ppm:

$$\frac{\Delta I}{I_{max}} \leq 200 \text{ ppm (carbon ions)} \quad [\text{rms, 2}^{\text{nd}} \text{ iteration}] \tag{6.43}$$

The scanning magnet agility (delay and ramp rate) are considered in the next section, along with extraction stability.

6.4.5 Extraction stability and scanning magnet agility

With the intensity monitor uncertainty negligible, one can just as well use the original E_n from Tab. 6.4 to specify the extraction stability, rather than σ_n. In Eq. 3.104, E_n is given as a function of the scanning magnet agility (delay and beam velocity) and the synchrotron extraction stability. At writing moment, the foreseen communication rate with the scanning magnet power supplies is 50 kHz, so:

$$f_{PS} = 50 \text{ kHz} \Rightarrow \begin{cases} T_{PS} = 20 \ \mu\text{s} \\ \langle \tau \rangle = 30 \ \mu\text{s} \\ \tau_{max} = 40 \ \mu\text{s} \end{cases} \tag{6.44}$$

Reshuffling Eq. 3.104 gives the minimum irradiation time per spot for homogeneous irradiation:

$$T_{spot} \geq \frac{(1 + \xi_{max})\tau_{max} - \langle \tau \rangle + \xi_{max}T_{move}/2}{E_n} \tag{6.45}$$

Contour plots of T_{spot} as a function of ξ_{max} and v_{beam} is shown in Fig. 6.14. In order to evaluate it, we first need an estimate of how long a single spot is allowed to take:

Single spot, 2 liter target From Chapter 2, it should not take more than 3 minutes to give a dose of 2 Gy to a 2 liter target. For such a large target, the largest available spot size will be used (W=10 mm). Assuming the target is cubical with 126 mm sides and a layer thickness of 4 mm gives about 31 layers and 1800 spots per layer. With a dead time of about 1.5 s between layers, the average irradiation time per spot is:

$$\frac{180 \text{ s}/31 - 1.5}{1800} = 2.4 \text{ ms [Average, large target]} \tag{6.46}$$

Spot in distal layer, Target 2 The optimized layer weight of Target 2 (deep-seated) is approximately 5×10^8 protons/cm^2 and 4×10^6 carbon ions/cm^2. At highest foreseen extraction intensity (see Tab. A.2), and fractional doses of 2.0 Gy and 0.7 Gy for protons and carbon ions, respectively, the irradiation time per spot ($W = 10$ mm) would be:

$$T_{spot} = \begin{cases} 2.8 \text{ ms (protons)} \\ 0.44 \text{ ms (carbon ions)} \end{cases} \quad [\text{Distal layer, Target 2}] \tag{6.47}$$

Spot in proximal layer, Target 2 In the proximal layer, Target 2, the optimized layer weighs are 3×10^7 protons/cm^2 and 3.5×10^5 carbon ions/cm^2, resulting in spot times at highest extraction intensity of:

$$T_{spot} = \begin{cases} 160\,\mu s \text{ (protons)} \\ 40\,\mu s \text{ (carbon ions)} \end{cases} \quad \text{[Proximal layer, Target 2]} \quad (6.48)$$
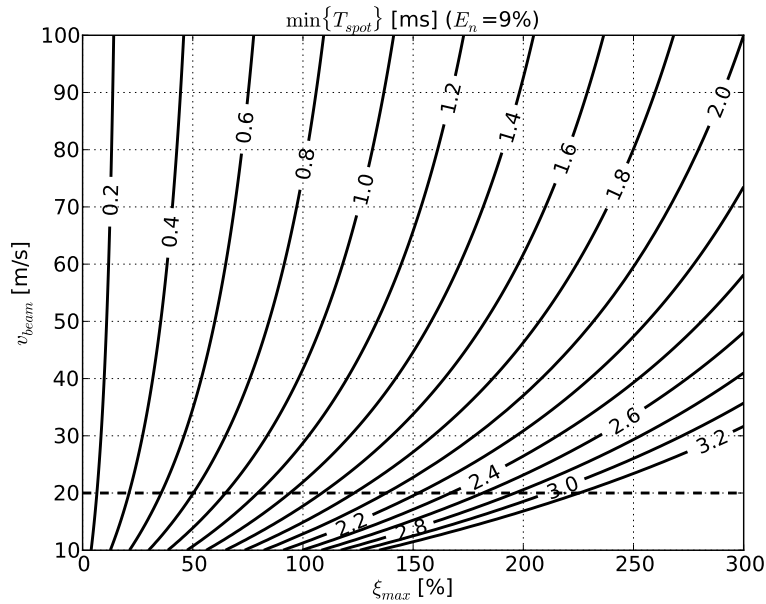

Figure 6.14: Minimum spot time to ensure a homogeneous dose distribution ($E_n = 9\%$).

With an average spot time of 2.4 ms (Fig. 6.14), one could accept up to 160% beam intensity ripple (v_{beam}=20 m/s). However, the lower fractional dose of carbon ions is somewhat problematic. Already in the most distal, high-weighted layer of Target 2 is the irradiation time per spot as low as 0.44 ms at highest extraction intensity. This is only marginally more than the minimum foreseen spot time of 0.3 ms. Taking into account that in a *real* treatment plan (non-rectangular target), high- and low-weighted spots are present in almost every layer [58], which means that the highest foreseen carbon ions extraction intensities can rarely be used at all: in most layers, a lower extraction intensity would have to be used in order not to systematically overdose the low-weighted spots. This is purely a limitation caused by the scanning magnets.

Nevertheless, to achieve the desired 9% spot weight accuracy for carbon ions, ξ_{max} should be less than 160% to safely irradiate spots at an average rate of 2.4 ms per spots (from the "2 Gy to 2 liters in 3-min" requirement).

The corresponding ripple tolerance for protons ($E_n = 12\%$) is $\xi_{max} \leq 220\%$. The required power supply accuracy for this is summarized in Tab. 6.7. The spill model

	$\Delta I/I_{max}$ ($\lesssim$1 kHz)		$\Delta I/I_{max}$ ($\gtrsim$1 kHz)	
	D	QD	D	QD
Protons ($\xi_{max} \leq 220\%$)	5 ppm	1 ppm	16 ppm	10 ppm
Carbon ions ($\xi_{max} \leq 160\%$)	5 ppm	1 ppm	8 ppm	5 ppm

Table 6.7: Max relative current ripple amplitude of the main ring magnet chains (D=Dipoles, QD=Defocusing quadrupoles, lowest k_{QB}).

has been used for medium frequencies (up to the kHz region), while results from particle tracking at 2 kHz are used for higher frequencies.

As mentioned in Sec. 4.3.2, ξ_{max} is below 10% at the quadrupole PWM frequency (12 kHz), so current ripple of several kHz are of little concern for the dose homogeneity.

6.4.6 Feasibility and possible improvements

The CNAO scanning magnet power supplies show an accuracy of 100 ppm (peak-to-peak) [59], so the here specified value of $\pm$200 ppm is technically feasible. As mentioned, this is also the case for the 1% beam intensity monitor precision, $\sigma_{BVS,n}$ [48].

Spot-to-spot beam width variations due to scanning magnet current ripple will be negligible with a $\pm$200 ppm precision, leaving some margin to other beam width error sources. The influence of static beam width errors (constant over the entire spill) will have limited effect on dose homogeneity if a shifted grid is used.

In [51], CNAO achieves an extraction stability equivalent to $\xi_{max} \approx 200\%$. Since the CNAO synchrotron (still under commissioning) is similar to the MedAustron synchrotron, this at least shows that the previously specified intensity stability of $\xi_{max} = 160\%$ is not unreasonable. However, the CNAO figure is achieved with RF-channelling and an air-cored quadrupole [12]. These techniques are also foreseen in the MedAustron ring, but primarily as a back-up solution: they are preferably avoided, in order to have an extraction mechanism that is as simple and fail-safe as possible. This uncertainty of technically achievable extraction stability makes the spot weight accuracy the most uncertain parameter and motivates tightening the requirements on the beam width and position. With the scanning magnets as only sources of beam position and width errors, the specified spot weight accuracies of 12% and 9% (protons and carbon ions) can be relaxed by approximately a factor 2, since the scanning magnet accuracy is tightened to 200 ppm instead of 380 ppm.

If the extraction stability requirements can not be met, the dynamic performance of the scanning magnets must be improved. Even if it *can* be met, improving the scanning magnet would be motivated for utilization of the highest foreseen extraction intensities. The relation between scanning magnet performance and minimum spot times and extraction ripple tolerance is shown in Fig. 6.15.

Increasing the scanning velocity to 50 m/s and f_{PS} to 100 kHz would be a vast improvement, allowing for 1 ms spots (or 400% beam intensity ripple). Such a scanning velocity is not unreasonable, but may come at the price of reduced beam
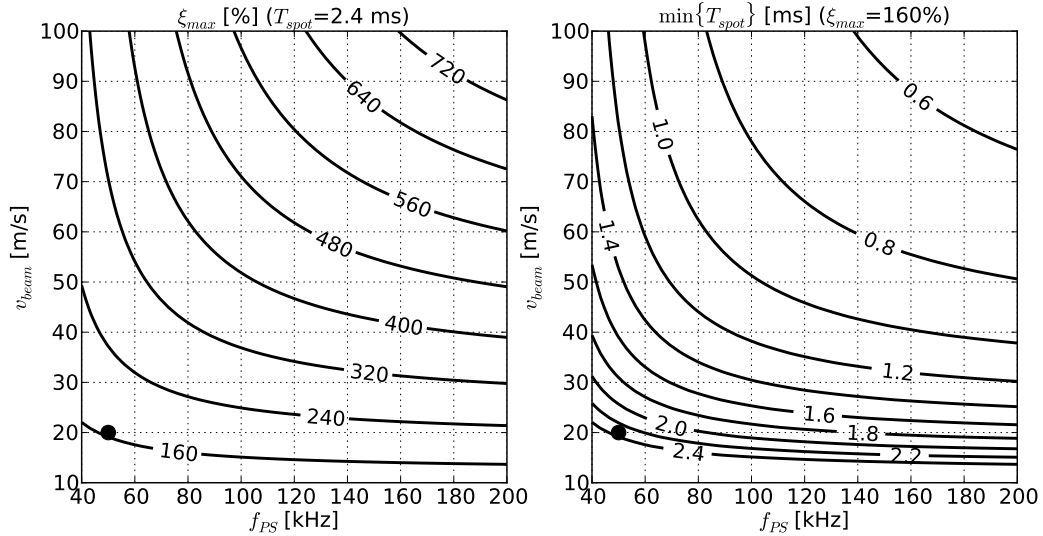
Figure 6.15: Permissible extraction ripple (left) and minimum spot times (right) as a function of scanning magnet agility.

positioning accuracy (the scanning magnets of e.g. HIMAC provide scanning velocities of up to 100 m/s [60]). Still, since the highest foreseen carbon ion extraction intensities can not be used safely without improving the scanning magnet performance, an improvement is motivated. Faster scanning magnets would in any case reduce the sensitivity to beam intensity fluctuations, and is therefore an alternative to supplementing the extraction mechanism with RF-channelling and an air-cored quadrupole.

6.4.7 Relaxation

By tightening the spot grid to $\Delta = W/4$, the requirements could be relaxed by roughly 30% (Tab. 6.3). However, this would reduce the spot weight by almost 50% (9/16), and the sensitivity to beam intensity ripple would be more or less the same. The lower spot weights would require reducing the beam intensity from maximum even further, causing longer irradiation time.

For Target 1, statistical "smoothening" in the proximal layers would also allow for a relaxation of the beam delivery requirements. For carbon ions, the requirements could be relaxed by at least a factor 2.3 ($k_{W,x}$, Tab. 6.3), while for protons, the requirements can be relaxed by between 1.5 (horizontal position) and 2.3 (spot weight).

6.4.8 Beam verification strategy and interlock conditions

If the error of some beam property at one spot is measured to be too large, an interlock to turn off the beam should be generated. However, with the assumption that errors are normal-distributed, occasional large-amplitude errors are to be expected,

but they will have little effect on the overall dose homogeneity.

Interlock conditions could be set e.g. at a 3-σ level of the expected error distribution (from Tab. 6.4). In order to verify that the error distribution is reasonably normal-distributed, one could set an upper limit of the number of spots per irradiation that is allowed to be between e.g. 2 and 3 σ of the tolerated standard deviation of the error (about 4%).

6.5   Dose model benchmarking at PSI, Gantry 2

Measurements have been made in Gantry 2 [13, 14] at PSI in order to benchmark the dose calculator presented in this work. Since the dose calculations are the basis for requirements along the beam delivery chain, the primary aim of the measurements was to confirm that the transverse dose distributions simulated at different target depths agreed with measurements. For this, an accurate modeling of the beam growth with depth is a prerequisite.

It should be stressed that when these measurements were made, Gantry 2 was still in the commissioning phase.

6.5.1   Preparations

Several single-layer treatment plans were prepared, consisting of spots in a Cartesian grid, 10×10 cm^2, with a beam energy of 173 MeV (about 20 cm range in water) and a dose of about 3 Gy at the Bragg peak. Four different spot-to-spot distances were used: 2, 3, 4 and 5 mm. The proton beam at PSI is approximately Gaussian in both planes, and the chosen spot-to-spot distances correspond roughly to 1/4 to 2/3 of the beam FWHM.

For each spot-to-spot distance, a reference plan with equal spot weight was generated, as well as plans with intentional random spot weight errors, in a discrete and integral dose driven manner. The maximum amplitude of the intentional random weight errors was set to 20% or 40% - high enough to overshadow any unintentional beam delivery imperfections.

6.5.2   Setup

A setup of the measurement is shown in Fig. 6.16: a CCD camera was positioned on the patient table, with the scintillating screen in a transverse plane at the isocenter. The resolution of the CCD camera is 17 pixels per cm.

To measure the transverse dose distribution at different depths Plexiglas plates were positioned on top of the CCD camera cover (5 mm plastics). In the first measurement series, the thickness of the Plexiglas plates were 80 mm, and in a second measurement 160 mm. The range of a 173 MeV proton beam in Plexiglas is about 17.9 cm, so the chosen thicknesses allow for a measurement of the transverse dose distribution in the plateau region (where the beam is largely unscattered) and closer to the Bragg peak (where scattering is more pronounced).
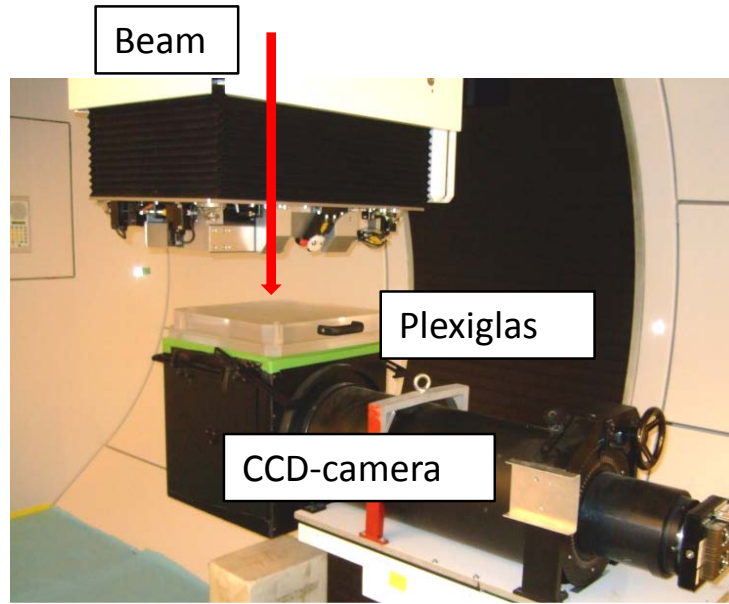
146

Figure 6.16: Setup in Gantry 2, PSI.

6.5.3 Data analysis

Fig. 6.17 shows the measured dose distribution from a reference plan, without any intentional spot weight errors ($\eta_i = 0$). Globally, the dose appears homogeneous, but a zoom on the center region and change of color scale reveals a granularity of a few percent. For this reason, the relative dose error when $\eta_i > 0$ is calculated according to Eq. 6.3.

A median filter (3×3) was applied to the CCD images to filter out malfunctioning pixels.

6.5.4 Measurements results

6.5.4.1 Beam growth due to scattering

Fig. 6.18 shows the measured profiles of 10 single spots, placed in a 5×5 cm grid[5], with 80 (left) and 160 (right) mm Plexiglas. The horizontal and vertical 1-σ widths of all spots (from fitting) are presented in Fig. 6.19a. The spread in spot sizes (1-σ) is in the order of 0.1 mm, and the horizontal beam profile is slightly wider than the vertical.

Fig. 6.19b shows a comparison between the simulated beam growth in Plexiglas of a 173 MeV proton beam, and the measured beam sizes at 8.5 and 16.5 cm depth[6]. The agreement between simulated beam sizes and measured is excellent.

---

[5]The 10th asymmetrically placed spot helps identifying which direction is horizontal and which is vertical in the CCD image files.

[6]The 5 mm thick plastic camera cover is modeled as Plexiglas.

(a) Raw image

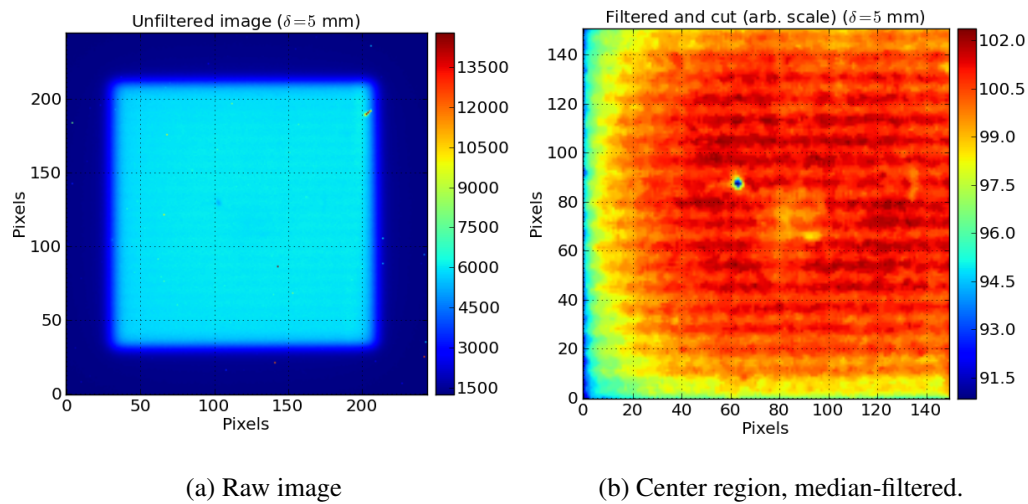(b) Center region, median-filtered.

Figure 6.17: CCD image of reference dose distribution ($\eta_i = 0$). The anomaly around coordinates (63,87) in the right figure is caused by a few malfunctioning pixels.
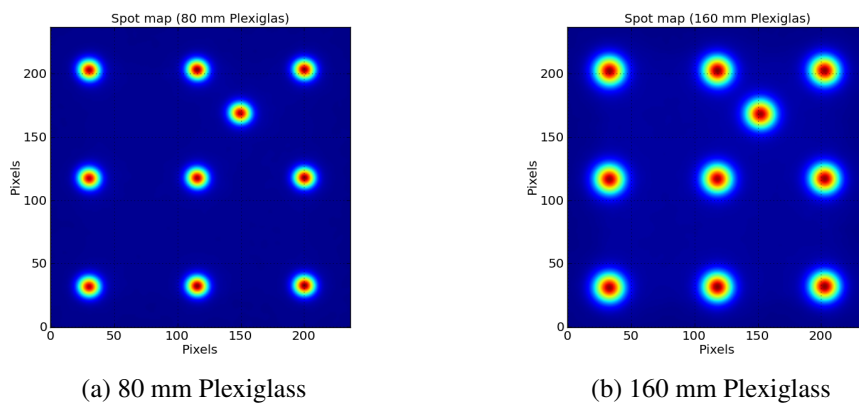

(a) 80 mm Plexiglass

(b) 160 mm Plexiglass

Figure 6.18: Spot profiles at different depths.

148

(a) Fitted beam sizes, horizontal and vertical.

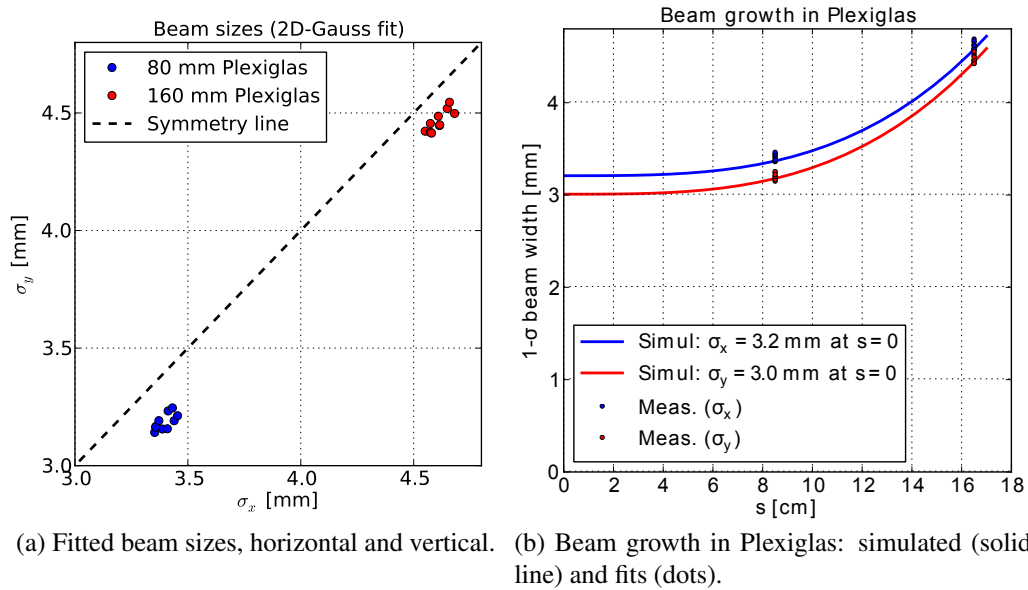(b) Beam growth in Plexiglas: simulated (solid line) and fits (dots).

Figure 6.19: Horizontal and vertical beam size (1-σ) with, 80 and 160 mm Plexiglas.

6.5.4.2 Field comparison

The agreement between measured and simulated transverse dose distributions was also good. One example is shown in Fig. 6.20 (8 cm Plexiglas, 3 mm spot-to-spot distance and up to 40% spot weight errors). Although there are minor differences in the dose distribution, the qualitative agreement is very good, as seen in Fig. 6.21.
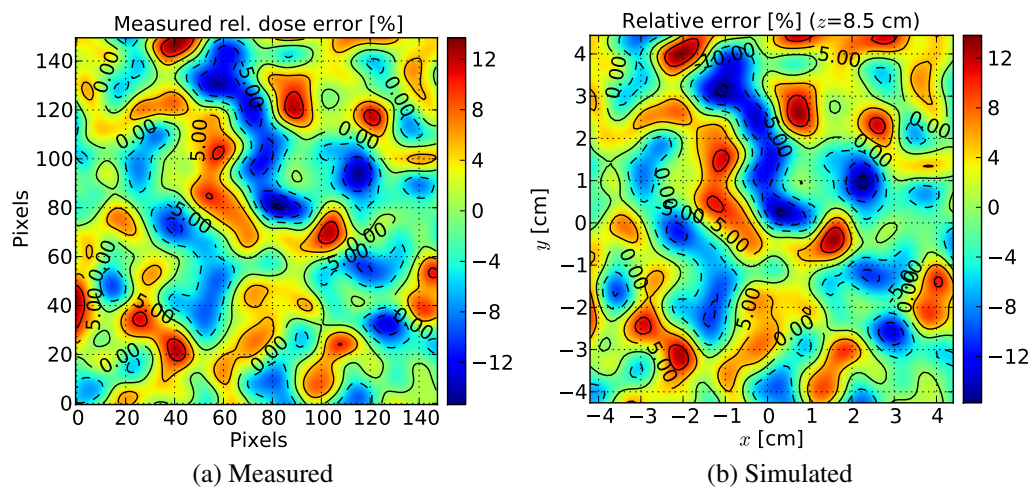

(a) Measured

(b) Simulated

Figure 6.20: Measured and simulated transverse dose distribution

An analysis of the measured dose homogeneity for different spot-to-spot distances is not included here, since the results were identical to those already presented in this chapter.
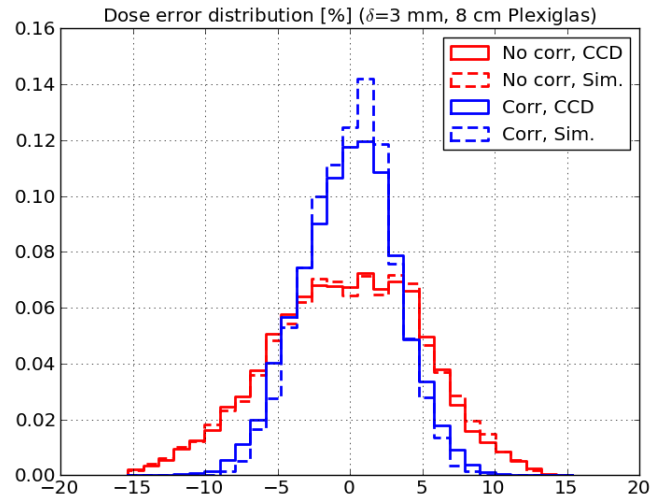
149

Figure 6.21: Simulated (Sim.) and measured (CCD) dose error histogram, discrete scanning mode (No corr) and integral-dose driven scanning (Corr).

6.6 Results and discussion

By combining the model of transverse beam growth due to scattering with parameterized Bragg curves, the impact of beam delivery system imperfections on target dose distribution has been evaluated. The dose calculation model was benchmarked in Gantry 2, PSI and agreed well with measured data.

With integral-dose driven scanning, slow beam intensity variations are automatically compensated for. Compared to discrete-scanning mode, random spot weight errors are suppressed more efficiently, which calls for implementation of a simple, integral-dose driven scanning algorithm.

Sufficient beam positioning accuracy for all foreseen beam sizes can be reached by a 200 ppm scanning magnet accuracy.

Beam intensity fluctuations should be kept below some 160% to ensure homogeneous irradiation. This results in a synchrotron power supply current accuracy of a few ppm at frequencies below the kHz range, while the required accuracy is more relaxed at higher frequencies.

For multi-layer targets, the constraints can be relaxed by at least a factor 1.5 (more than 2 for carbon ions). However, to ensure a homogeneous irradiation of the distal part of the tumor, Target 3 (single layer at 7 cm depth) requirements apply.

To avoid that small errors of the horizontal beam width cause significant dose errors, a shifted spot grid should be used.

By using a tighter spot grid ($\Delta = W/4$), an even better suppression of random beam delivery errors can be achieved. The disadvantage is that the spot weights are reduced, which sets an upper limit on the extraction intensity. However, if the target is small, the irradiation time per spot is not crucial: the total irradiation time will then be dominated by the time it takes to switch energy.

Allowing for 160% beam intensity fluctuations is based on the assumption that spots are irradiated in 2.4 ms, on average.

With 160% beam intensity fluctuations, spot irradiation times should on average be about 2.4 ms to ensure a homogeneous dose. This is more than five times longer than the nominal irradiation time of a single spot in the distal layer of a carbon ion target at highest foreseen carbon ion extraction intensity (0.44 ms at $\times 10^9$ ions/s). In other words: only a fifth of the maximum design intensity can safely be used. When taking into account that most layers of a "real" treatment plan contain spots of very low weight, this figure becomes even smaller. In an analysis of the treatment plans for over 600 fields for patient treatment at PSI, more than 90% of all spots had a weight which was less than 10% of the highest weighted spot [58]. To make use of the highest synchrotron extraction intensities, the dynamic performance of the scanning magnets must be improved: the gain with an improved scanning velocity and higher communication rate with the scanning magnet power supplies is shown in Fig. 6.15.

Increasing the scanning velocity to 50 m/s and f_{PS} to 100 kHz would be a vast improvement, allowing for 1 ms spots. Such a scanning velocity is not unreasonable (the scanning magnets of e.g. HIMAC provide scanning velocities of up to 100 m/s [60]), but may come at the price of reduced beam positioning accuracy[7]. Still, since the highest foreseen carbon ion extraction intensities can not be used safely without improving the scanning magnet performance, an improvement is motivated. Faster scanning magnets would in any case reduce the sensitivity to beam intensity fluctuations, and is therefore an alternative to complex corrections of the extraction method, such as RF-channelling and air-cored quadrupole.

---

[7]In [60], a current ripple and drift within ± 500 ppm is reported.

Chapter 7

Conclusions

In this work, a framework for linking target dose homogeneity to technical performance of the beam delivery chain has been developed. In Chapter 2, the clinical requirements were specified. In the following chapters (Chapter 3 to Chapter 5), the tools needed for understanding how different elements and processes affect the patient dose were developed. The current stability of synchrotron magnet power converters was translated into extraction stability; errors in spot position, size and weight were linked to the stability and agility of the scanning magnet power converters.

Different nozzle layouts and their impact on target beam size have been studied using a semi-empirical scattering model, which was extended to take beam optics into account, using standard Twiss functions. The optics extension made it possible to evaluate different options for the proton gantry with respect to beam size at the patient and discarding solutions with a significantly inferior penumbra. Vacuum windows at the gantry coupling point were consequently discarded due to the specifics of the MedAustron proton gantry optics. The implemented scattering model was also used to study different nozzle layouts with respect to minimum achievable beam sizes at the isocenter: in air and at the Bragg peak in a water phantom. Low energy beam sizes of 7 mm FWHM are achievable, but requires monitors that can be moved towards the patient, combined with helium bellows that reduce beam blow-up compared to air.

Using parameterized Bragg curves made it possible to study the impact of extraction energy errors on SOBP homogeneity, but also to evaluate and compare different kinds of ridge filters. For MedAustron, a solution with two orthogonal triangular ridge filters has been selected. This allows placing both ridge filters downstream of the monitors, avoiding double-calibration of the monitors depending on whether the ridge filters are used or not. An absolute energy accuracy of about 0.3 MeV/n is required for carbon ions to produce a sufficiently homogeneous SOBP: for protons, the corresponding accuracy is 0.3 MeV for deep-seated targets and 0.2 MeV for superficial targets. For a range accuracy of 0.25 mm, a relative energy accuracy ($\Delta E/E$) of down to 0.4×10^{-3} is needed for both protons and carbon ions.

In Chapter 6, the impact of various beam delivery errors (wrong beam position, wrong spot weight or wrong beam width) on target dose quality was investigated, us-

ing the implemented `dosecalc` module. The module joins the transverse scattering model with parameterized Bragg curves. This makes it possible to calculate the target volume dose distribution from a spot scanning session subject to various kinds of errors. Upper limits on acceptable beam delivery errors could thereby be set, based on clinical requirements. Using the previously developed framework, these limits could be directly translated into requirements on e.g. synchrotron and scanning magnet power converters, beam monitor performance or extraction energy accuracy. Requirements on synchrotron magnet power converter stability were estimated to be in the ppm region to limit beam intensity fluctuations to 160%. In practice, this is considered very strict. For the fast ramping scanning magnets, a power converter stability in the order of 200 ppm was deemed sufficient. The agility of the scanning magnet power converters - ramp rate and delay - determines the robustness of the scanning system with respect to beam intensity fluctuations. Faster scanning magnets would relax the constraints on the extraction stability, but also allow for using the full potential of the synchrotron more frequently: with the configuration parameters presented in this work, the highest available extraction intensities could rarely be used without risking unacceptable spot weight errors.

Several of the topics discussed in this work relates not only to the hardware of the beam delivery chain, but also to the software used in the preceding step: the treatment planning. It has been shown that the use of a "shifted grid" would dramatically reduce the sensitivity to static beam width errors in the horizontal plane. The treatment planning system should also aim at minimizing the number of spots of very low weight. The lowest weighted spot in a layer determines the maximum beam intensity that can be used, since the irradiation time of a single spot should never be shorter than the transition time. Very low weighted spots would thus put further restrictions on the extraction intensities that can be used which, in the end, makes the irradiation longer than necessary.

Appendix A

MedAustron synchrotron parameter list

The Twiss functions at the resonance sextupole (RS) and at the electrostatic septum (ES) are summarized in Tab. A.1. Main parameters of the MedAustron synchrotron and relativistic beam figures are presented in Tab. A.2 and A.3, respectively.

	RS	ES	
β_x	8.89864882	16.69113436	m
β_y	3.317467258	7.577337623	m
α_x	-0.1729102443	0.4283717337	-
α_y	-0.627889158	-0.2609741363	-
γ_x	0.11573	0.070906	m^{-1}
γ_y	0.42073	0.140961	m^{-1}
μ_x	5.40704126	9.38210643	rad
μ_y	6.18148638	9.64965671	rad
D_x	0.0	-3.954919	m
D_y	0.0	0.0	m
D'_x	0.0	0.6272660	rad
D'_y	0.0	0.0	rad

Table A.1: Twiss functions at resonance sextupole (RS) and electrostatic septum (ES).

155

Horizontal tune	Q_x	1.6666
Vertical tune	Q_y	1.78916
Horizontal chromaticity	Q'_x	-4.0
Vertical chromaticity	Q'_y	-1.09
Circumference	C	75.24 m
Resonance sextupole length	l_s	0.2 m
Normalized resonance sextupole gradient	k'	11.245 m^{-3}
Normalized sextupole strength	S	27.79 m$^{-1/2}$
ES position	x_{ES}	-35 mm
Momentum spread of beam	$\frac{\Delta p}{p}$	4×10^{-3}
Extraction duration	T_{extr}	1-10 s
Minimum proton extraction intensity		1×10^8 s^{-1}
Maximum proton extraction intensity		2×10^{10} s^{-1}
Minimum carbon ion extraction intensity		4×10^6 s^{-1}
Maximum carbon ion extraction intensity		1×10^9 s^{-1}
Beam size at isocenter (FWHM)	W_x, W_y	4, 6, 8, 10 mm

Table A.2: Main synchrotron parameters

	Protons		Carbon ions	
	60 MeV	250 MeV	120 MeV/n	400 MeV/n
ε_x, 1σ [π mm mrad]	1.4286	0.6679	1.4286	0.7325
ε_y, 1σ [π mm mrad]	1.4286	0.6679	1.4286	0.7325
β_{rel}	0.3415	0.6136	0.4640	0.7146
γ_{rel}	1.0639	1.2665	1.1289	1.4295
$B\rho$ [Tm]	1.14	2.43	3.25	6.34
T_{rev} [ns]	734	409	541	351

Table A.3: Geometric emittance, relativistic beam parameters β_{rel} and γ_{rel}, magnetic rigidity and revolution time for highest and lowest extraction energies.

Appendix B

Scattering parameters

Material	X_0 [g/cm^2]	P [g/cm^3]	X_0/P [cm]
Water	36.08	1.00	36.08
Air	36.66	1.205 g/l	304 m
Copper	12.86	8.96	1.43
Mylar	39.95	1.39	28.7
Kapton (polyimide)	40.56	1.42	28.6
Helium	94.32	0.1786 g/l	5280 m
Aluminum	24.01	2.7	8.89

Table B.1: Radiation length, density and radiation length divided by density. Numbers are taken from [40, Chapter 6].

Appendix C

Benchmarking of scattering model

C.1 Comparison with experimental results

Gottshalk [42] publishes a list of experimentally measured values of θ_M and θ_0 (fit) for different materials and thicknesses. A comparison between a few values of the fitted scattering angle, as measured in experiments, and that calculated with the scattering module `mpscatter` are presented in Tab. C.1. The agreement is within a few percent.

C.2 FLUKA

In order to validate the output from `mpscatter`, it has also been compared to simulation results from the MonteCarlo simulation package `FLUKA` [61, 62], using the graphical user interface Flair.

An example of the `FLUKA` input is shown in Fig. C.1, for a 250 MeV proton beam. The default parameters used are `NEW-DEFA`. In order to calculate the proton beam size at different longitudinal coordinate, the command `EVENTBIN` has been used. One such example is shown in Fig. C.2, where a 1-dimensional grid in steps of 0.2 cm perpendicular to the beam direction is defined 150 cm downstream of where the geometry starts. Using several `EVENTBINS` at different locations give the evolution of the beam size, which is initially taken to be zero.

Figure C.1: The first Flair input cards.

Thickness t [g/cm^2]	θ_0 [mrad], Measured	θ_0 [mrad], Calculated
Lexan ($\rho =$ 1.2 g/cm^2)		
0.094	1.762	1.58
1.455	7.436	7.45
Aluminum ($\rho =$ 2.7 g/cm^2)		
0.216	3.53	3.42
11.957	39.99	39.96
Copper ($\rho =$ 8.96 g/cm^2)		
0.045	2.204	2.02
10.13	49.068	47.8

Table C.1: Column 1: density normalized target thickness. Column 2: θ_0 as fitted to experimental results. Column 3: θ_0 as calculated using Kanematsu's integration of the scattering power.

Figure C.2: EVENTBIN card in Flair.

C.2.1 FLUKA analysis

The FLUKA simulations are done on a particle-by-particle basis. As soon as a proton passes through one of the segments in the defined EVENTBINSs, the track length in that particular segment is written to a file. If the proton trajectory is exactly perpendicular to the grid, the track length is equal to the longitudinal thickness Δz of the grid ($\Delta z = 0.01$ cm in Fig C.2). Protons with a non-zero divergence will yield a track length larger than Δz. An example of the output from one primary proton is shown below:

```
Binning n:    5, "           ",  Event #:  80, Primary(s) weight  1.0000E+00
Number of hit cells: 1
507  0.0100211296
```

This means that during the 80th simulated proton, a proton track length of 0.010211296 cm was recorded in segment number 507 in the 5th EVENTBIN.

The lateral proton beam profile is given by the number of protons (not the track length) recorded at each segment. A Gaussian beam profile can then be fitted to the central 98% of the protons. An example of the EVENTBIN histogram for a 60 MeV proton beam after passing through 2.4 cm water is shown in Fig. C.3. As seen, the Gaussian fit is a good approximation for the lateral beam profile, with a 1-σ beam width of 0.45 mm. In this case, mpscatter calculates the corresponding 1-σ beam width to be around 0.42 mm, an under-estimation of less than 10%.
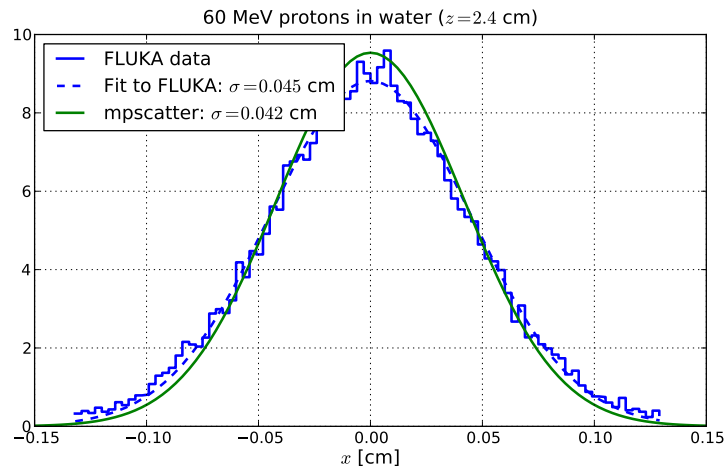
Figure C.3: Horizontal beam distribution of an initially zero emittance 60 MeV proton beam after traversing 2.4 cm water.

C.2.2 Beam widening in a homogeneous material

The widening of a 60 MeV proton beam in water is shown in Fig. C.4. As seen, FLUKA and mpscatter are in agreement to within a few percent. In Fig. C.5, the water has been replaced with helium and the initial beam energy raised to 250 MeV. The beam range is around 2.2 km, but the beam sizes from mpscatter and FLUKA are still in agreement within 10%.
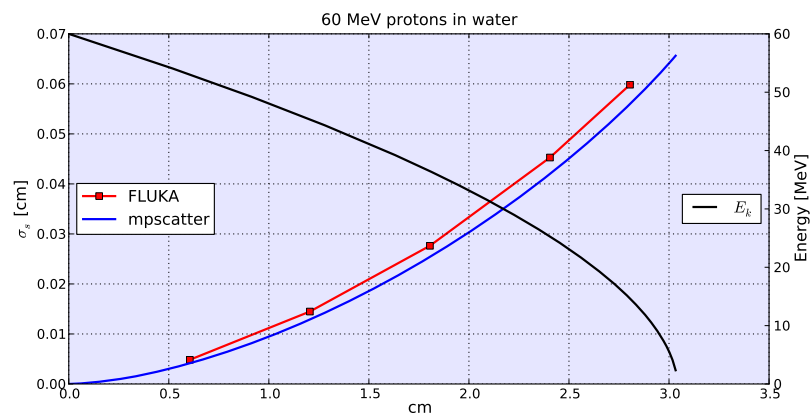

Figure C.4: Left axis: the 1-σ beam width versus depth for a zero-emittance 60 MeV proton beam entering water. Right axis: the beam energy as a function of depth.

C.2.3 Beam widening in a multi-slab geometry

mpscatter and FLUKA are still in agreement when the geometry is heterogeneous. Shown in Fig. C.6 is the beam widening of a 250 MeV proton beam. The geometry consists of a repeated pattern of 5 cm water, 1 cm copper followed by 96 cm helium.
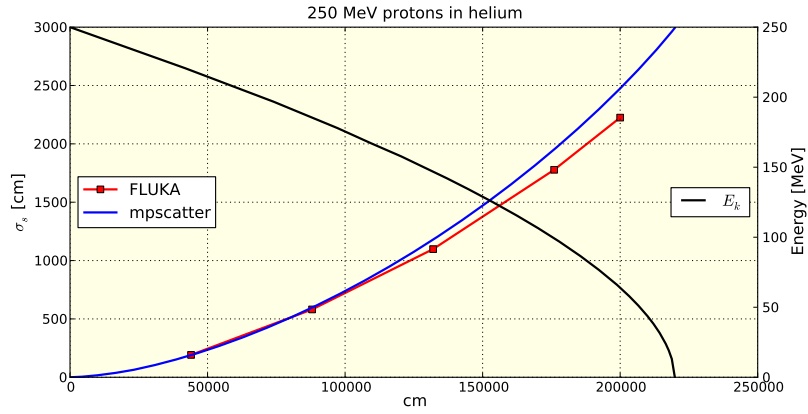
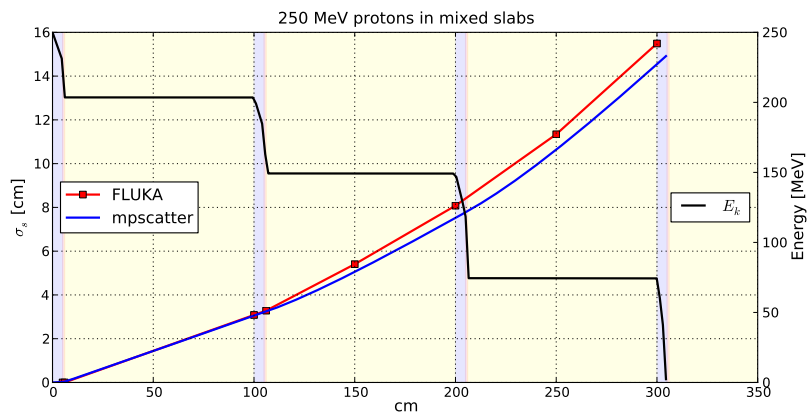Figure C.5: Beam widening of 250 MeV proton beam in helium

Figure C.6: Beam widening of a 250 MeV proton beam in a mixed-slab geometry (5 cm water, 1 cm copper and 94 cm helium repeated).

Bibliography

[1] V. Bozica and B. Bojana, "Radiotherapy: Past and present," in *Archive of oncology*, vol. 18, pp. 140–142, 2010.

[2] R. R. Wilson, "Radiological use of fast protons," *Radiology*, vol. 47, no. 5, pp. 487–491, 1946.

[3] HPaul, "Depth_Dose_Curves.jpg." http://commons.wikimedia.org/wiki/File:Depth_Dose_Curves.jpg, March 2009. Use permitted under the terms of the GNU Free Documentation Licence.

[4] ICRU, "ICRU Report 78 - Prescribing, Recording and Reporting Proton-Beam Therapy," *Journal of the ICRU*, vol. 7, no. 2, 2007.

[5] E. Pedroni, R. Bacher, H. Blattmann, T. Böhringer, A. Coray, A. Lomax, S. Lin, G. Munkel, S. Scheib, U. Schneider, and A. Tourovsky, "The 200-MeV proton therapy project at the Paul Scherrer Institute: Conceptual design and practical realization," *Medical Physics*, vol. 22, no. 1, pp. 37–53, 1995.

[6] IAEA, "Dose reporting in ion beam therapy," Tech. Rep. 1560, IAEA (International Atomic Energy Agency), 2007.

[7] T. Inaniwa, T. Furukawa, S. Mori, S. Sato, N. Kanematsu, T. Shirai, and K. Noda, "Dosimetric impact of multiple energy operation in carbon-ion radiotherapy," in *Proceedings of IPAC'11, San Sebastian, Spain*, September 2011.

[8] C. Grupen, "Tumour therapy with particle beams," Tech. Rep. physics/0004015. SI-2000-2, Siegen Univ. Fachber. Phys., Httental-Weinau, Apr 2000.

[9] M. F. Moyers and A. Ghebremedhin, "Spill-to-spill and daily proton energy consistency with a new accelerator control system," *Medical Physics*, vol. 35, no. 5, pp. 1901–1905, 2008.

[10] M. Benedikt and A. Wrulich, "MedAustron Project overview and status," *The European Physical Journal Plus*, vol. 126, pp. 1–11, 2011. 10.1140/epjp/i2011-11069-9.

[11] L. Badano, M. Benedikt, P. J. Bryant, M. Crescenti, P. Holy, A. T. Maier, M. Pullia, S. Rossi, and P. Knaus, "Proton-Ion Medical Machine Study (PIMMS), 1," p. 223, March 1999.

[12] P. J. Bryant, L. Badano, M. Benedikt, M. Crescenti, P. Holy, A. T. Maier, M. Pullia, S. Reimoser, S. Rossi, G. Borri, P. Knaus, F. Gramatica, M. Pavlovic, and L. Weisser, "Proton-Ion Medical Machine Study (PIMMS), 2," Aug 2000.

[13] E. Pedroni, R. Bearpark, T. Bhringer, A. Coray, J. Duppich, S. Forss, D. George, M. Grossmann, G. Goitein, C. Hilbes, M. Jermann, S. Lin, A. Lomax, M. Negrazus, M. Schippers, and K. G, "The PSI Gantry 2: a second generation proton scanning gantry," *Zeitschrift für Medizinische Physik*, vol. 14, pp. 25–34, 2004.

[14] E. Pedroni, D. Meer, C. Bula, S. Safai, and S. Zenklusen, "Pencil beam characteristics of the next-generation proton scanning gantry of PSI: design issues and initial commissioning results," *European Physical Journal Plus*, vol. 126, p. 66, July 2011.

[15] U. Amaldi AND M. Silari, ed., *The TERA Project and the Centre for Oncological Hadrontherapy*. INFN - LNF - Divisione Ricerca, 1994.

[16] W. Chu, J. Staples, B. Ludewigt, T.R.Renner, M. R.P. Singh, J. Collier, I. Daftari, H. Kubo, P. Petti, L. Verhey, J. Castro, and J.R.A1onso, "Performance specifications for proton medical facility," tech. rep., Lawrence Berkeley Laboratory, University of California, Accelerator & Fusion Research Division, 1993.

[17] A. Wambersie, "What accuracy is required and can be achieved in radiation therapy (review of radiobiological and clinical data)," *Radiochimica Acta*, vol. 89, pp. 255–264, 2001.

[18] ICRU, "Fundamental Quantities and Units for Ionizing Radiation (ICRU Report 85)," techreport 1, Oxford University Press, 2011.

[19] M. Urie, M. Goitein, W. R. Holley, and G. T. Y. Chen, "Degradation of the Bragg peak due to inhomogeneities," *Physics in Medicine and Biology*, vol. 31, no. 1, p. 1, 1986.

[20] T. Freeman, "Proton therapy: advances, challenges." http://medicalphysicsweb.org/cws/article/research/34341, May 2008. MedicalPhysicsWeb.

[21] ICRU, "ICRU Report 71 - Prescribing, Recording and Reporting Electron Beam Therapy," *Journal of the ICRU*, vol. 4, no. 1, 2004.

[22] E. B. Podgorsak, ed., *Radiation Oncology Physics: A Handbook for Teachers and Students*. IAEA, 2005.

[23] K. Bratengeier, M. Oechsner, M. Gainey, and M. Flentje, "Remarks on reporting and recording consistent with the ICRU Reference Dose," *Radiation Oncology*, vol. 4, no. 1, p. 44, 2009.

[24] D. Dubini, S. Gallo, F. Gerardi, R. Orecchia, S. Rossi, and V. Rotta, "The CNAO functional specifications (CNA-TNDC-050WXX-00250)," tech. rep., CNAO, 2003.

[25] "Loma Linda University Medical Center, Proton Therapy." http://www.protons.com.

[26] M. F. Moyers, G. B. Coutrakon, A. Ghebremedhin, K. Shahnazi, P. Koss, and E. Sanders, "Calibration of a proton beam energy monitor," *Medical Physics*, vol. 34, no. 6, pp. 1952–1966, 2007.

[27] E. Courant and H. Snyder, "Theory of the alternating-gradient synchrotron," *Annals of Physics*, vol. 3, no. 1, pp. 1 – 48, 1958.

[28] P. Bryant, "Introduction to particle accelerators." JUAS 2009, 2009.

[29] M. Benedikt, *Optical design of a synchrotron with optimisation of the slow extraction for hadron therapy*. PhD thesis, Vienna, Tech. U., Geneva, 1997. Presented on Oct 1997.

[30] J. B. Marion, *Classical dynamics of particles and systems*. 111 Fifth Avenue, New York, New York 10003: Academic Press, 2 ed., 1970.

[31] HPaul, "Electronic and nuclear stopping power for Al ions in Al." http://commons.wikimedia.org/wiki/File:Electronic_nuclear_stopping_Al_in_Al.png, March 2007. Use permitted under the terms of the GNU Free Documentation Licence.

[32] J. E. Turner, *Atoms, radiation, and radiation protection*. New York: WILEY-VCH Verlag GmbH & Co. KGaA, 3 ed., 2007.

[33] ICRU, "Stopping of ions heavier than helium," Oxford, 2005.

[34] W. T. Chu, B. A. Ludewigt, and T. R. Renner, "Instrumentation for treatment of cancer using proton and light-ion beams," *Review of Scientific Instruments*, vol. 64, no. 8, pp. 2055–2122, 1993.

[35] P. Kundrat, "A semi-analytical radiobiological model may assist treatment planning in light ion radiotherapy," *Physics in Medicine and Biology*, vol. 52, no. 23, p. 6813, 2007.

[36] J. F. and Ziegler, "SRIM-2003," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 219-220, no. 0, pp. 1027 – 1036, 2004. Proceedings of the Sixteenth International Conference on Ion Beam Analysis.

[37] R. Zhang and W. D. Newhauser, "Calculation of water equivalent thickness of materials of arbitrary density, elemental composition and thickness in proton beam irradiation," *Physics in Medicine and Biology*, vol. 54, no. 6, p. 1383, 2009.

[38] F. Moser, *Energy Verification in Ion Beam Therapy*. Doctoral thesis, Atominstitut, Technische Universität Wien, Vienna, Austria, 2012.

[39] T. Bortfeld, "An analytical approximation of the bragg curve for therapeutic proton beams," *Medical Physics*, vol. 24, no. 12, pp. 2024–2033, 1997.

[40] W.-M. Yao *et al.*, "Review of Particle Physics," *Journal of Physics G*, vol. 33, pp. 1+, 2006.

[41] G. Molière, "Theorie der Streuung schneller geladener Teilchen II. Mehrfach- und Vielfachstreuung," *Zeitschrift Naturforschung Teil A*, vol. 3, p. 78, 1948.

[42] B. Gottschalk, A. Koehler, R. Schneider, J. Sisterson, and M. Wagner, "Multiple coulomb scattering of 160 mev protons," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 74, no. 4, pp. 467 – 490, 1993.

[43] V. L. Highland, "Some practical remarks on multiple scattering," *Nuclear Instruments and Methods*, vol. 129, no. 2, pp. 497 – 499, 1975.

[44] N. Kanematsu, "Alternative scattering power for Gaussian beam model of heavy charged particles," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, vol. 266, no. 23, pp. 5056 – 5062, 2008.

[45] N. Kanematsu, "Semi-empirical formulation of multiple scattering for the Gaussian beam model of heavy charged particles stopping in tissue-like matter," *Physics in Medicine and Biology*, vol. 54, no. 5, p. N67, 2009.

[46] U. Weber and G. Kraft, "Design and construction of a ripple filter for a smoothed depth dose distribution in conformal particle therapy," *Physics in Medicine and Biology*, vol. 44, no. 11, pp. 2765–2775, 1999.

[47] S. Giordanengo, A. Ansarinejad, A. Attili, F. Bourhaleb, R. Cirio, M. Donetti, M. A. Garella, F. Marchetto, G. Mazza, V. Monaco, J. P. Montero, A. Pecka, C. Peroni, G. Russo, and R. Sacchi, "The CNAO system to monitor and control hadron beams for therapy," *Nuclear Science Symposium Conference Record, 2008. NSS '08. IEEE*, pp. 5609–5611, Oct. 2008.

[48] S. Lin, T. Boehringer, A. Coray, M. Grossmann, and E. Pedroni, "More than 10 years experience of beam monitoring with the Gantry 1 spot scanning proton therapy facility at PSI," *Medical Physics*, vol. 36, no. 11, pp. 5331–5340, 2009.

166

[49] M. T. Gillin, N. Sahoo, M. Bues, G. Ciangaru, G. Sawakuchi, F. Poenisch, B. Arjomandy, C. Martin, U. Titt, K. Suzuki, A. R. Smith, and X. R. Zhu, "Commissioning of the discrete spot scanning proton beam delivery system at the University of Texas M.D. Anderson Cancer Center, Proton Therapy Center, Houston," *Medical Physics*, vol. 37, no. 1, pp. 154–163, 2010.

[50] M. Crescenti, "RF empty bucket channelling combined with a betatron core to improve slow extraction in medical synchrotrons.," Tech. Rep. CERN-PS-97-068-DI, CERN, Geneva, Jan 1998.

[51] M. Pullia, "Status report of the cnao construction and commissioning," in *IPAC'11*, Sep 2011.

[52] A. Smith, M. Gillin, M. Bues, X. R. Zhu, K. Suzuki, R. Mohan, S. Woo, A. Lee, R. Komaki, J. Cox, K. Hiramoto, H. Akiyama, T. Ishida, T. Sasaki, and K. Matsuda, "The M. D. Anderson proton therapy system," *Medical Physics*, vol. 36, no. 9, pp. 4068–4083, 2009.

[53] F. Bourhaleb, A. Attili, R. Cirio, P. Cirrone, F. Marchetto, M. Donetti, M. A. Garella, S. Giordanengo, N. Givehchi, S. Iliescu, A. L. Rosa, J. Pardo, A. Pecka, and C. Peroni, "Monte Carlo simulations of ripple filters designed for proton and carbon ion beams in hadrontherapy with active scanning technique," *Journal of Physics: Conference Series*, vol. 102, no. 1, p. 012002, 2008.

[54] S. Fujitaka, T. Takayanagi, R. Fujimoto, Y. Fujii, H. Nishiuchi, F. Ebina, T. Okazaki, K. Hiramoto, T. Sakae, and T. Terunuma, "Reduction of the number of stacking layers in proton uniform scanning," *Physics in Medicine and Biology*, vol. 54, no. 10, p. 3101, 2009.

[55] R. Garcia-Molina, I. Abril, P. de Vera, I. Kyriakou, and D. Emfietzoglou, "Role of the interaction processes in the depth-dose distribution of proton beams in liquid water," *Journal of Physics: Conference Series*, vol. 373, no. 1, p. 012015, 2012.

[56] G. Arduini, R. Cambria, C. Canzi, F. Gerardi, B. Gottschalk, R. Leone, L. Sangaletti, and M. Silari, "Physical specifications of clinical proton beams from a synchrotron," *Medical Physics*, vol. 23, no. 6, pp. 939–951, 1996.

[57] I. Lira, *Evaluating the Measurement Uncertainty - Fundamentals and Practical Guidance*. Series in Measurement and Technology, Institute of Physics Publishing, 2002.

[58] A. J. Lomax, T. Bohringer, A. Bolsi, D. Coray, F. Emert, G. Goitein, M. Jermann, S. Lin, E. Pedroni, H. Rutz, O. Stadelmann, B. Timmermann, J. Verwey, and D. C. Weber, "Treatment planning and verification of proton therapy using spot scanning: Initial experiences," *Medical Physics*, vol. 31, no. 11, pp. 3150–3157, 2004.

[59] M. Incurvati, F. Burini, M. Farioli, G. Taddia, S. G. di Piano, M. Donetti, S. Toncelli, G. Venchi, I. D. Cesaris, C. Sanelli, F. Voelker, S. Giordanengo, and F. Marchetto, "Fast high-power power supply for scanning magnets of CNAO medical accelerator," in *Proceedings of EPAC08, Genoa, Italy*, pp. 3643–3645, 2008.

[60] T. Furukawa, T. Inaniwa, S. Sato, T. Shirai, Y. Takei, E. Takeshita, K. Mizushima, Y. Iwata, T. Himukai, S. Mori, S. Fukuda, S. Minohara, E. Takada, T. Murakami, and K. Noda, "Performance of the NIRS fast scanning system for heavy-ion radiotherapy," *Medical Physics*, vol. 37, no. 11, pp. 5672–5682, 2010.

[61] G. Battistoni, F. Cerutti, A. Fasso, A. Ferrari, S. Muraro, J. Ranft, S. Roesler, and P. R. Sala, "The FLUKA code: description and benchmarking," in *Hadronic Shower Simulition Workshop* (M. Albrow & R. Raja, ed.), vol. 896 of *American Institute of Physics Conference Series*, pp. 31–49, Mar. 2007.

[62] A. Ferrari, P. R. Sala, A. Fasso, and J. Ranft, *FLUKA: A multi-particle transport code (program version 2005)*. Geneva: CERN, 2005.

Marcus Palm

Curriculum Vitae

732 Rue de Paris
F-01170, Gex
France
✆ *+41 (76) 487 9580*
✉ *marcus.palm@cern.ch*
Date of Birth: Jan. 28, 1982

Education

2008–2013 **Doctor Technicae**, *Vienna University of Technology*, Austria.
Doctoral student at the Technical University of Vienna and at CERN.
- Design of and specifications for the MedAustron beam delivery system. A semi-analytical scattering model has been implemented, combining scattering theory with optical beam transport in order to calculate the transverse beam profile evolution along a transfer line in the presence of any scattering media, as well as target dose distribution. Important design choices of the proton gantry and the beam delivery nozzles are based on these evaluations. Various aspects affecting dose homogeneity in the patient have been investigated and quantified, such as permissible synchrotron magnet power supply ripple, required performance of the beam scanning magnets, beam optics precision at the iso-center and required beam interruption rate and time.

2010 **PSI Winterschool**, *Bad Zurzach*, Switzerland.
1 week course on proton therapy organized by Paul Scherrer Institute, Switzerland

2009 **Joint Universities Accelerator Scohol (JUAS)**, *Archamps*, France.
5 week course on accelerator physics organized by the European Scientific Institute

2002–2008 **Master of Science**, *Linköping University*, Sweden.
Applied Physics and Electrical Engineering (*Teknisk Fysik och Elektroteknik*)
- Specialization: Theory, Modeling and Visualization
- Master thesis at CERN, Geneva
  - Responsible for the detection of secondary charged particles in a proof-of-principle test of a novel kind of neutrino factory: a free mercury jet inside a 15 T solenoid.
  - Design, implementation and operation of diamond detectors and electron multipliers, peripheral equipment for these and a remote LabView interface for the system.
  - Simulations of the secondary particle flux, using FLUKA.
  - Analysis of the detector signals.
- Projects: Molecular dynamics-based simulator for investigation of properties and behaviour of gases (Computational Physics), Construction of an autonomous, IR-sensor guided robot (Electronics)
- Courses: Quantum mechanics, Quantum dynamics, Electrodynamics, Elementary particle physics, Physics of condensed matter, Computational fluid dynamics, Automatic control, Electronics and measurement technology, etc.

Experience

2012–Present	**System Engineer**, *MedAustron*, Wiener Neustadt, Austria.

Working on the realization of the beam delivery system for the Austrian hadron therapy facility MedAustron
- Analysis for the risk management process
- Investigating future advanced scanning modes
- Architecture design for the MedAustron Particle Therapy Accelerator

2006	**Tutor**, *Linköping University*, Linköping, Sweden.

Tutor for first year master students in a physics course

2001-2002	**Military Service**, *Göta Ingenjörregemente*, Eksjö, Sweden.

Military service at an engineering company. In command over a group of nine, with the assignment to establish and maintain inter-batallion communication.

Languages

Swedish	**Native**
English	**Fluent**
French	**Intermediate**

Computer skills

Software	**LabView, FLUKA, Matlab, Office, Latex, Root, CAD, WinAgile, MAD-X**
Programming language	**C++, Python, Assembler**
OS	**Windows, Linux**

Publications

- Physical Review - Accelerator and Beams
  - *Transfer Line Scattering Model of Therapeutic Hadron Beams and Applications to Nozzle and Gantry Optimization*, (Phys. Rev. ST Accel. Beams 16, 014702, 2013)
- IPAC 2011
  - *Design Choices of the MedAustron Nozzles and Proton Gantry based on Modeling of Particle Scattering*, 2nd International Particle Accelerator Conference, San Sebastian, Spain, 4-9 Sep 2011, pp. THPS081 (First author, selected for publication in Physical Review, Special Topics)
  - *Dose Homogeneity Driven Beam Delivery System Performance Requirements for MedAustron*, 2nd International Particle Accelerator Conference, San Sebastian, Spain, 4-9 Sep 2011, pp. THPS082 (First author)
- IPAC 2010
  - *Overview of the MedAustron Design and Technology Choices*, 1st International Particle Accelerator Conference, Kyoto, Japan, 23-28 May 2010, pp. MOPEA020
- EPAC 2008
  - *The MERIT High-Power Target Experiment at the CERN PS*, 11th European Particle Accelerator Conference, Genoa, Italy, 23-27 Jun 2008, pp. WEPP169